The book contains the papers, installations, posters, artworks, installations, video and live performances at the XXVII Generative Art conference hosted by the UNESCO Regional Bureau for Science and Culture in Europe, Castello 4930, 30122 Venice, Italy.
The GA conference is supported by the Italian National Commission for UNESCO

The outer cover features a 3D model generated, created as an interpretation of Venice's identity in progress starting with St. Mark's Basilica, by Celestino Soddu with Argenia AI.

The inside covers are images from the book ‘I Quattro libri dell’Architettura’ by Andrea Palladio, Venice 1570

Printed in Rome the 8th of December 2024

Domus Argenia Publisher

ISBN 9788896610473

GENERATIVE ART 2024

GA2024, XXVII Annual International Conference at the UNESCO Regional Bureau for Science and Culture in Europe

GA2024 is organized by Generative Art and Design Lab, Argenia Association, Roma, Italy

Proceedings

Edited by Celestino Soddu and Enrica Colabella

INDEX

Page 5
Celestino Soddu, Enrica Colabella*, Italy, Chairs of Generative Art conference,* ***OPENING***

PAPERS, INSTALLATIONS, VIDEOS, LIVE PERFORMANCES and ARTWORKS

Page 17
Celestino Soddu, Italia, ARGENIA: AN OPERATIONAL STEP TOWARDS THE DYNAMIC DEFINITION OF URBAN AND ARCHITECTURAL IDENTITIES AS HERITAGE TO BE SAFEGUARDED, Generative Art & Design Lab, Argenia, Roma

Page 29
Enrica Colabella, Italia, BEAUTY *TRACES* FOR AI *NATURELESS*, Generative Art & Design Lab, Argenia, Roma

Page 39
Marie-Pascale Corcuff, France, CITY PATHS, MAZES, INTELLIGENCE, AND GENERATIVE DESIGN, Architect, Rennes

Page 50
B.T. Franklin, USA, SPOOKLIGHT: A GENERATIVE APPROACH TO AI-DRIVEN STORYTELLING, independent Researcher, Phoenix, AZ

Page 60
Rafael Tamim, Tristan Cazenave, France, AN ATTEMPT AT GENERATING BEETHOVEN'S 33RD, PIANO SONATA WITH MUSIC TRANSFORMER, LAMSADE, Université Paris Dauphine - PSL, Paris

Page 70
Ian O'Loughlin, USA, GENERATIVE AUTHORSHIP, Department of Philosophy, Pacific University, Forest Grove.

Page 80
Ben Baruch Blich, Israel, THE BEAUTY OF THE RUINED BUILDING, The Interdisciplinary program of the Arts, Faculty of the Arts, Tel Aviv University

Page 89
Peter Beyls, Belgium, EMBODIED INTERACTION WITH ALGORITHMS, The School of Arts, University College Ghent.

Page 99
Sun Park, UK, THE REGISTRATION OF AI-GENERATED WORKS AS UNESCO

HERITAGE: HUMAN-ALGORITHMIC COMMON HERITAGE OF MANKIND, Department of Art History and Cultural Practices, University of Manchester, Manchester

Page 110
Christopher Fry, UK, ***Chad Eby***, USA LEXICONIA – A LEXICON OF GENERATIVE AI, Westminster School of Art, University of Westminster, School of Art and Visual Studies, University of Kentucky, Lexington, KY

Page 119
Tatsuo Unemi, Japan BUILDING A WEB SITE TO PLAY BACK GENERATIVE ANIMATIONS, Glycan and Life Systems Integration Centre, Soka University, Hachioji

Page 125
Cristina M. Reis, Portugal, ***Nuno Barreiro***, UK MAKE ROOM FOR NATURE! PERSONAL NARRATIVES CO-CREATED WITH GENERATIVE AI CENSE – Center for Environmental and Sustainability Research, NOVA School of Science and Technology, NOVA University Lisbon, Caparica, Department of Electronic Engineering, School of Engineering, Physical and Mathematical Sciences, Royal Holloway University of London, Egham

Page 134
Nicolas Reeves, Vincent Cusson, Alexandre Pariseau, Pierrot-Baptiste Lemée-Jolicœur, Marie-Martine Bédard, Canada THE JOURNEY OF THE MEMORY HORSE, University of Quebec in Montreal, University of Montreal, Activist and independent musician, Ste-Adèle.

Page 148
Monika Karwaszewska, Poland, GENRE MEMORY AND THE BEAUTY OF EARLY MUSIC VS AI IN THE AUDIOVISUAL PERFORMANCE 'CONCERTO GROSSO – W LABIRYNCIE', Stanisław Moniuszko Academy of Music in Gdańsk

Page 159
Agnieszka Mensfelt, UK, ON THE ARTISTRY OF GENERATIVE AI AND THE CASE FOR SURRE-AI-LISM, Department of Computer Science, Royal Holloway, University of London, Egham

Page 168
Manuel Báez, Canada, THE RESONANT MIND-FIELD OF EMBODIED CONSCIOUSNESS, Azrieli School of Architecture and Urbanism, Carleton University, Ottawa

Page 178
Ginger L Franklin, USA, ENCODING THE DIVINE: LEVERAGING AI FOR THE TRANSLATION AND PRESERVATION OF SIKH SCRIPTURE, Independent

Scholar/Phoenix, Arizona

Page 185
Davide Prete, USA, AN EXPLORATION OF AI TEXT-TO-3D GENERATORS: LATEST DEVELOPMENTS AND APPLICATIONS ON 3D PRINTING, Division of Arts and Humanities, Washington, DC

Page 194
Peggy Bloomer, USA, FAST FORWARD: HOW ARTIFICIAL INTELLIGENCE (AI) HAS INTEGRATED INTO THE DESIGN CURRICULUM IN LESS THAN A YEAR, Department of Art & Design, Central Connecticut State University, New Britain, CT

Page 202
Nikolaus Bezruczko, USA, GENERATIVE AI CLARIFIES VALIDITY OF AUTHENTIC PAINTINGS FOR MEASURING ARTISTIC JUDGMENT APTITUDE, Department of Clinical Psychology, The Chicago School of Professional Psychology, Chicago

Page 212
Guillaume Rochais, France, RANDOMNESS, RULE, PROGRAM: HOW THE ISSUES OF GENERATIVE ART TAKE SHAPE IN PRE-ALGORITHMIC MUSIC, ACCRA, Université de Strasbourg

Page 229
P. L. Capucci, Italy, ART, HUMANITY, AND BEYOND. TOWARDS NEW CREATIVE PARADIGMS, Noema, LABA Rimini

Page 238
Margaret McManus, USA, ARCHIVE STUDIO: EXPLORING FUTURE SPACE FOR FORMER MATTERS, School of Building Arts, Savannah College of Art and Design, Savannah, GA

Page 245
Joanna E. Kabala, The Netherlands, BEAUTY AND THE PROMPT, Eindhoven

Page 253
Jônatas Manzolli, Micael Antunes, Brazil, Portugal, France, RE(M)IND: COMPOSITION AND GENERATIVE SOUND DESIGN OF A MULTIMODAL OPERA, Interdisciplinary Nucleus for Sound Studies (NICS), Unicamp, Brazil, Centre for Informatics and Systems of the University of Coimbra (CISUC), Portugal; Aix-Marseille Univ., CNRS, Prism, Lis, InCIAM

Page 263
Stephen Erfle, USA, HUMANISTIC EXPLORATIONS IN ALGORITHMIC ART

USING SEQUENCE PLAYER MODE OF PLAYING WITH POLYGONS, department of International Business and Management, Dickinson College, Carlisle, PA

Page 271
Angela Ferraiolo, USA, COUPLED AND INTELLIGENT INTERACTIONS IN REACTION-DIFFUSION SYSTEMS, Sarah Lawrence College, Bronxville, New York

Page 278
Alain Lioret, Kamal Hennou, France, WORLD BEINGS, Université Paris 8, ESGI Digital Lab

Page 289
Roland Baz, Kristina Malyseva, Anna Pappa, Tristan Cazenave, France, BRAICHT, A THEATRICAL AGENT THAT SPEAKS LIKE BERTOLT BRECHT'S CHARACTERS, LAMSADE, Paris Dauphine - PSL, ZHDK, Zurich Univ. of the Arts, LIASD Univ. Paris 8

Page 296
Dragana Ćirić, Serbia, PROGRAMMABLE ARCHITECTURE, URBANISM, AND TERRITORY: SPATIAL INTELLIGENCE AND DATA-DRIVEN FORMS OF STRATEGIC URBAN SCENARIOS, Independent Scientific Associate/Research Assistant Professor; Founder, Licenced Architect, and PI at unit [d], Belgrade

Page 305
Reza Kakooee Benjamin Dillenburger, Switzerland, ILLUMINATING SPACES: DEEP REINFORCEMENT LEARNING AND LASER-WALL PARTITIONING FOR ARCHITECTURAL LAYOUT GENERATION, Institute of Technology in Architecture, Department of Architecture, ETH Zurich

Page 315
Natalie Oulikhanian, Behnaz Merikhi, Rhonda McEwen, Canada, SHARED SPACES OF CREATION: GENERATIVE ARTIFICIAL INTELLIGENCE'S CREATIVE AGENCY AND AFFORDANCES IN THE ARTISTIC PROCESS, University of Toronto

Page 325
Katherine Lapierre, Canada, OUTSIDER ARCHITECTURES, School of Design, Université du Québec à Montréal

Page 336
Daniela Sirbu, Canada, IMPRESSIONS OF AQUASCAPES WITH REFLECTIONS OF SPRING, Department of New Media, Faculty of Fine Arts, University of Lethbridge

Page 342
Umberto Roncoroni, Peru', ANIMATION TECHNIQUES USING L-SYSTEMS, Department of Graphic Design, Universidad Peruana de Ciencias Aplicadas

Page 352
Jan Paredaens, Belgium, THE BEAUTY OF ROTATIONS IN PROCESSING, University of Antwerp

Page 359
Curtis L Palmer, Canada, IN, OUT, AROUND - REVISITED, Synergetic Design, Edmonton, Alberta

Page 365
Marija Obradović, Slobodan Mišić, Serbia, LEARNING THROUGH PLAYING WITH AI: COLLABORATION OF INTELLIGENCES IN CREATING A TRANSFORMABLE 3D ART INSTALLATION, Faculty of Civil Engineering, University of Belgrade, Faculty of Applied Arts, University of Arts in Belgrade.

Page 375
Yan Bello Méndez, Spain, SONIFICATION.AI: GENERATIVE SOUND EXPERIMENTS, SpaceMinds SL, Madrid

Page 380
Bin Umino, Asako Soga, Motoko Hirayama, Japan, UTILIZING THE GENERATED DANCE CHOREOGRAPHIES IN THEATRE PERFORMANCES, Faculty of Sociology, Toyo University, Tokyo, Faculty of Advanced Science and Technology, Ryukoku University, Otsu, Graduate School of Comprehensive Human Sciences, University of Tsukuba

Page 387
Chin-En Keith Soo, Rowan Simmons, New Zealand, CHROMATIC GHOST, Department of Design, University of Waikato

Page 397
David Bouchard, Canada, ISOPOLIS, New Media Department, Toronto Metropolitan University

Page 400
İrem Hafız, Turkiye, A NEW AESTHETICS OF VASTNESS: COMPLEX FRACTAL GEOMETRIES IN GENERATIVE AI, Department of Architecture, Middle East Technical University, Ankara

Page 404
Anna M. Chupa, Michael A. Chupa, USA, HANDAIWORK: GENERATIVE AI-ENHANCED QUILTED TEXTILES, Art Architecture and Design, Lehigh University, Bethlehem, Anna Chupa Designs, Easton, Pennsylvania

Page 407
Arne Eigenfeldt, Canada, GOGO MUSEBOTS: A GENERATIVE CO-CREATION WITH HUMAN MUSICIANS School for the Contemporary Arts, Simon Fraser University, Vancouver

Page 409
Slawomir Wojtkiewicz, Karolina Masłowska, Poland, DREAM OF A CITY, Department of Visual Arts, Faculty of Architecture, Bialystok University of Technology, Academy of Fine Art in Gdansk, Department of Architecture

Page 416
Lucas Bang, USA, AUTOPOIETIC NEURAL NETWORK – DYNAMIC IMAGES ENCODED AS A SELF-MODIFYING NEURAL NETS, Department of Computer Science, Harvey Mudd College, Claremont CA

Page 420
Edward S. Johnston, USA, , FLUID TIME VARIATION, ANIMATED PROJECTION ARTWORK, Robert Busch School of Design, Michael Graves College, Kean University Union, New Jersey

Page 427
Anna Ursyn, USA, DANCE CODE, University of Northern Colorado

Page 429
Chad Eby, USA, THE ALPHABETIC LABYRINTH, School of Art and Visual Studies, University of Kentucky, Lexington

Page 433
Jeffrey M. Morris, Rebecca Hays, USA, VIRIDISSIMARIGORE: EVERGREEN STRUGGLES BETWEEN THE ORGANIC AND THE MECHANICAL, College of Performance, Visualization & Fine Arts, Texas A&M University, College Station

Page 435
Andrea Wollensak, Katja Novak, Brett Terry, USA, Estonia, GENERATIVE VISUAL POETRY: DRAWING THE LAND'S MEMORY THROUGH PLACE-BASED POEMS, Department of Art, Ammerman Center, Connecticut College, New London, CT; Poet, Translator, Tallinn; Independent Scholar/Composer, Sonalysts, Waterford, CT

Page 437
Jean-Claude Heudin, France, ANGELIA: AN EMOTIONAL GENERATIVE ALGORITHMIC INTELLIGENCE FOR CONTEMPORARY ELECTRONIC MUSIC, Artificial-Creature, Univ. Paris-Sud

Page 439
Monika Karwaszewska, Beata Oryl, Michał Garnowski, Marcin Garnowski, Poland, GESTOMAT - AN INTERACTIVE INSTRUMENT FOR THE PRECOMPOSITION OF AUDIOVISUAL WORKS, Stanisław Moniuszko Academy of Music in Gdańsk, Academy of Fine Arts in Gdańsk, Gdańsk University of Technology

Page 441
Enrico "Catodo" Zimuel, Italy, SYMMETRIIS, A SOFTWARE TO PERFORM LIVE VISUAL SYMMETRIES IN GENERATIVE ART PERFORMANCES, Computational artist, Turin

Page 443
Fernanda N. Bellicieri, Brazil, SOLO SALE SIAMO… JUSTDUST, Communication and Languages Center , Mackenzie Presbyterian University
San Paolo

Page 445
Sam Wells, USA, THE COTTONWOOD FLORILEGIUM, Boyer College of Music and Dance, Temple University, Philadelphia

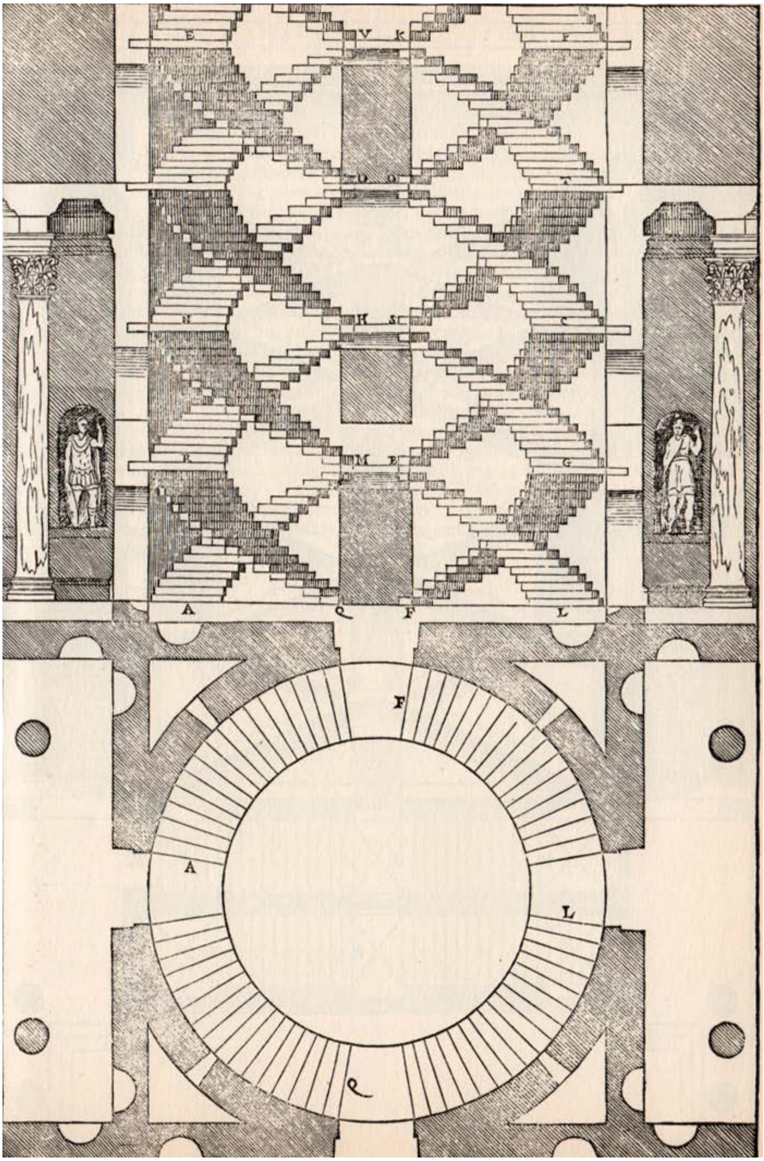
Opening

It’s a big pleasure to meet all friend and new participant at the XXVII Generative Art Conference hosted in so wonderful city by the UNESCO Regional Bureau for Science and Culture in Europe
In this last time the term "generative" became a main characteristic of the new AI generation. This is highly relevant to anyone exploring Generative Art. AI works referring broadly to software systems with open structures capable to answer to complex questions of our time. However, it’s essential to clarify that various types of AI differ significantly from one another:
AI dedicated to specific functional tasks
AI that aims to emulate general intelligence, akin to human reasoning
AI that learns from data (machine learning)
Deep learning AI, which utilises “deep” neural networks
AI for natural language processing
AI for image processing and Generative Adversarial Networks
AI for controlling machines (robotics)
AI for decision-making and evaluating alternatives
AI that supports human creativity
The latter is the one we are most interested in, often referring to Generative Art. These systems, rooted in the early digital age, embody a specific logical framework that can plot possible future developments.
The topic that was addressed most by the participants to this XXVII GA conference was Art creativity with AI, focusing the concept of beauty in our time. The fundamental statement is that machines are not creative following the traditional human concept. Now machines can perform a fundamental function in explicating the human idea by evolving a concept into its representations. This process works in the field of words, music, images or 3D models. So AI can be a fundamental support to achieve the explication of an idea in complex forms.
The main question, now, how can the idea be defined so that the AI can generate possible complex evolutions of formalisation. Referring to the concept expressed by Poincarè, the creative idea is a topological interpretation of what exists. That works by a structured reading of what exists toward a new-subjective organisation of the relationships between the incoming events. This could also be done by the machine using random choices and relations between events. But randomness, in the act of generating an event, cannot reflect an intended concept and purpose, as happens in every creative gesture. It is fundamentally just a representation of a chance and this might invalidate the interest aims.
This creative idea conception is and remains human peculiarity. All subsequent development can be managed by machines. Since it is a matter of progressively transforming the idea topological structure into successive and progressive formalisations, the algorithms must be able to manage transformations in harmony with the characters performed by authors.
Generative Art does not attempt to be universal. It works by adapting to the unique logical characteristics of each user/designer. It operates by interpreting and expanding on the

author's creative logic, generating scenarios that clarify and amplify rather than invent new ideas, and elaborating on a creative subject's core concept in diverse ways.
The groundbreaking advancement in common generative AI is its newfound capability to use language naturally. This doesn't make AI "smarter" by itself but stems from significant progress in translation software over recent years. AI now leverages organization paradigms specific to each language, moving beyond mere literal translation toward more fluid and natural communication. Recently, this includes arranging sentences to enhance the readability and coherence of responses, even when diverging substantially from direct translations.
This evolution allows AI to transition from data-derived concepts, like those provided by search engines, to responses that feel more conversational. Importantly, this does not enhance the analytical quality or creativity of AI responses, which remain data-driven and objective. However, when designed to foster creative thinking—as in Generative Art—AI can support and extend human creativity.
In our efforts to engage a broad community of researchers, scientists, and artists in the 27 Generative Art conferences, participants' contributions have varied widely, often diverging from purely generative experimentation through original software. Nevertheless, we affirm that Generative Art offers real potential for the advancement of generative AI by enabling the creation of "authorial" AIs. These unique systems could offer greater complexity and specificity, similar to how we approach books—not as bearers of absolute truth but as expressions of thought that we might resonate with or explore.
In today's world, where there is a pressing need to preserve the beauty and uniqueness of the events surrounding us, a subjective approach can help us understand the complexity and recognize and safeguard elements that are part of humanity's shared heritage.

Celestino Soddu and Enrica Colabella
Chairs of Generative Art Conferences

Venice, 8th of December 2024



Andrea Palladio, 1570

PAPERS

ARGENIA: An Operational Step Towards the Dynamic Definition of Urban and Architectural Identities as Heritage to be Safeguarded

Celestino Soddu Prof. Arch.
Generative Art & Design Lab, Argenia Association, Rome, Italy
celestino.soddu@argenia.it

Celestino Soddu

The Author's AI Towards Species Projects

In the 1980s, when I developed my generative design software for architecture and urban planning—initially named Basilica and later Argenia—I aimed to create a system, working only on 3D models, that could assist the design process from the earliest conceptual stages. As an architect and designer, I found not acceptable to abandon multiple potential paths for a project to pursue only one, as is often required to complete a design. My goal was to create a tool that could take an initial, un-formalized idea and develop a series of formalized scenarios, each unique yet recognizable as part of the same idea. In essence, just as in nature, each variation would be an individual belonging to the same species. The objective was to design a species, not just a single instance. The working on 3D models was, together with transformation algorithms applied to 3D events and representing a topological expression of my architectural idea, was the peculiarity of my software.

Identity as a Distinctive Character of Beauty to Be Safeguarded

The most intriguing outcome of this project was the interactive approach I developed, which could define the character of a species through algorithms based on progressive transformations, not merely shapes. It became possible to generate a series of distinct, three-dimensional events, all recognizable as stemming from the same core concept. The identity, often not tied to specific forms, expresses itself through the organic nature of the generated events.

Identity is a crucial aspect of any event, be it a city, a building, or an object. It is something that should be safeguarded, as it is the trait most valued in urban and environmental contexts. However, it is also one of the hardest qualities to define. Identity is not static; it is dynamic, continuously evolving. Like a living entity, it cannot be captured with simple analysis of shapes or colours. Each intervention in an urban environment interacts with its existing character, enhancing, transforming, or even diluting its recognisability. This evolution can either strengthen the identity or reduce it to the point of anonymity.

Uniqueness and Recognisability

The charm and beauty of a city, of a site whether natural or artificial, lie in its uniqueness and recognisability. This uniqueness allows us to interpret it as an event continuously transforming toward our vision of an ideal environment. Even a city as steeped in history as Rome can be viewed as evolving toward an ever-more “Roman” future. The same can be said for New York, which could become an even larger "Big Apple," or Venice, whose identity might accelerate toward becoming "more Venice" than before.

Codes Representing Identity

The identity of a city can be understood as a code that manages its evolution while simultaneously becoming more complex. This evolutionary code cannot be defined through analysis, it could be, as in my Argenia AI, a system of algorithms that identifies and manages the city’s potential paths, each contributing toward an aspect of the ideal city, This is a concept that can never be fully realized but serves to guide what we aspire to achieve.

ARGENIA: A Generative

Author's AI Based on Interpretation

I developed the ARGENIA generative software following this philosophy. Since its inception in the 1980s, ARGENIA has been a dynamic tool for exploring identity, not as a singular result but as an evolving possibility. It is, by nature, a visionary AI based on my vision, rooted in subjective interpretations rather than mere data analysis. My first experiment with ARGENIA aimed to explore the identity of Italian medieval cities, not through direct observation but through their depiction in the paintings of Giotto and Simone Martini. By interpreting these artworks, I developed algorithms capable of transforming simple 3D geometric forms into complex events, which evoked a medieval sensibility.

ARGENIA and Subjectivity as a Path Toward Infinity

ARGENIA is fundamentally based on subjectivity, on the subjective interpretation of past artistic visions, like those of Giotto. Just as perspective, rooted in Renaissance logic, can represent infinity, ARGENIA uses subjectivity to navigate complexity and approach the infinite. A subjective approach has no limits, unlike an objective, deductive and non-visionary one, which is restricted by its inability to engage with infinity.

The Structure of ARGENIA: An Author's AI

ARGENIA is not built on traditional databases; it is not as GAN that operates with adversarial networks but not exceed the first idea and works, until now, only with 2D images. Instead, Argenia operates on a collection of transformation algorithms, representing my discovering path as an architect. These evolving algorithms reflected and interpreted each new search of mine, each project I made in the last 40 years, allowing for the expression of identity in various forms. Characters identified in the past can be transformed into codes, which can be applied to new projects, helping to shape the future by interpreting the past.

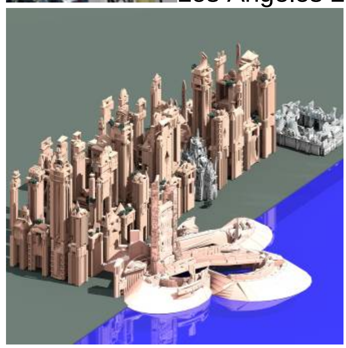
Interpreting the Past to Shape the Future

The power of ARGENIA lies in its ability to accept my interpretation of 3D historical forms and identities to forge new paths toward an idealized version of a city, one that preserves its unique identity while evolving dynamically. This approach moves beyond mere historical documentation, envisioning cities like Venice not as static entities, but as living, transforming organisms.

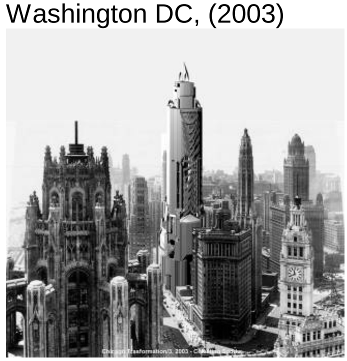
Operational Experiments on Urban Identity and the Ideal City

Over the past few decades of experimenting with the possibility of operationally investigating environmental identities, I have considered various urban and architectural realities. From early experiments on medieval Italian cities to Renaissance, Baroque, Futurism, and beyond. I have also considered and attempted to algorithmically construct the identity of cities such as:


Los Angeles 2003


Manhattan(2016)


Washington DC, (2003)


Chicago (2003),


Delhi,(2006)


Hong Kong,


Nagoya2003


Lima 2018

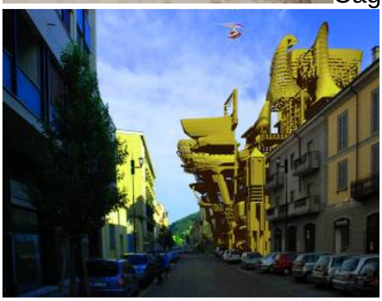

TelAviv2004


Jerusalem (2011),
and, in Italy, that of

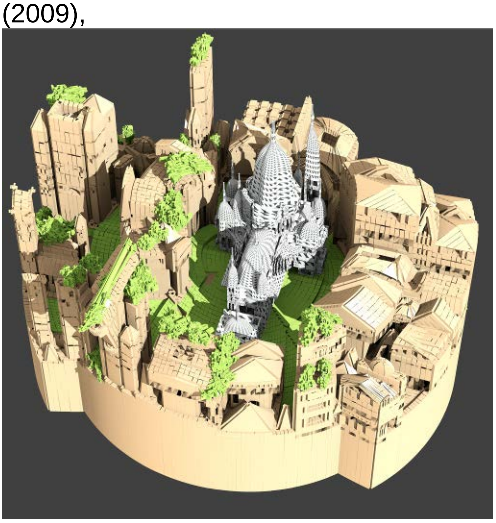

Rome(2011),


Milan (2004),

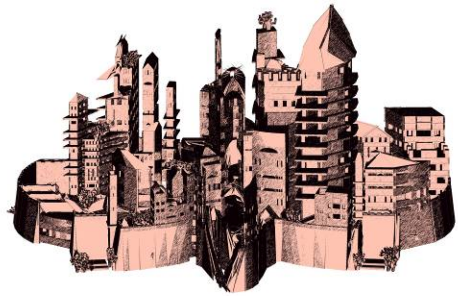
Cagliari(2009



Lecco (2009),



the Tuscan cities from Florence to the boroughs (2016),



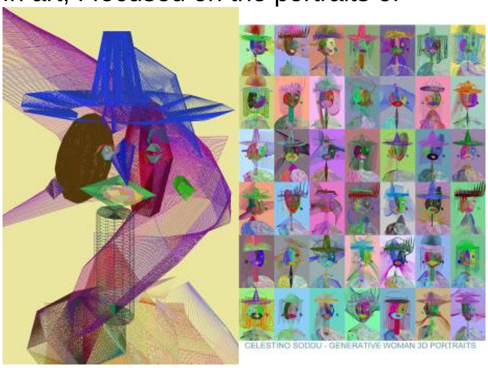
Lucca (2012),

Venice (2015),

Ravenna (2017),
and

Verona.(2018).

To refine my search for identity, I explored transformation codes related to Borromini's architecture, as well as those of Gaudí and the Nuraghi structures in Sardinia.
In art, I focused on the portraits of



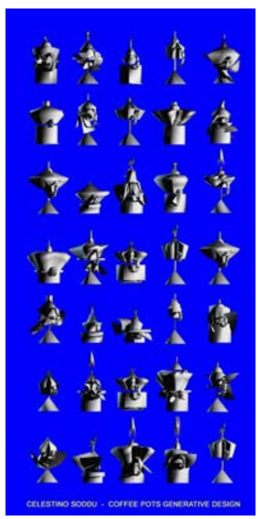
Picasso (1997) and

Francis Bacon (2018).
To verify some of these identities, I applied them to the design of various objects, including:

chairs(1992-2004)

,Jewels (2000),

lamps,



Coffee pots.

cars (2016), clothes, and more.

The goal of identity recognition and the potential for making it dynamic was explored by presenting various scenarios of these cities, generated through my generative software, Argenia, during my solo exhibitions in the cities I studied. I consistently asked visitors the same question: In which of these scenarios is your city (e.g., Los Angeles) more representative of itself than before? Or is Hong Kong more distinctly Hong Kong than it was previously?
From the answers, which were obviously as subjective as my approach, I was able to better tune the algorithmic construction of the specific identity and its complexity.
I also realized that by varying certain 3D geometric transformation parameters, it was often sufficient to focus specific identities and rediscover lost ones. Working on the Baroque period, particularly on Francesco Borromini, was very instructive. Borromini explicitly employed the principles of geometric transformation to transition from classical to Baroque styles. By interpreting his works, writing virtual transformation algorithms become evident, and by adjusting some of their parameters, I succeed also to interpret other identities.

This work of identifying identities, while highly subjective, can also be read inter-subjectively by discovering the hidden rules. However, it seems to be the only approach capable of exploring the dynamic complexity of these 3D events.

Venice and its Identity. From St Marco to Canaletto

During this meeting in Venice at the UNESCO office, I aimed to expand upon the initial work I developed a decade ago. My previous project interpreted Canaletto's paintings to explore Venetian identity by generating 3D architectures and collecting them in a topological way referring to Canaletto. More, these 3d complex models of possible Venice was populated by a virtual fashion exhibition made by using the same 3d transforming algorithms. Today, however, I am focusing on St. Mark's Basilica 3D interpretation to investigate its identity and what Venetians communicate to the world about Venice.
St. Mark's Basilica has undergone significant transformations over the centuries, reflecting the evolving identity and complexity of Venice beyond its original form. I aimed to interpret this dynamic by establishing a basic topology and exploring its variations, transformations, and potential increases in complexity, even by altering the complexities of the topologies involved.
The 3D results I generated with Argenia are representations of the various formal aspects of the idea I interpreted. These scenarios are not merely formal interpolations of the events in St. Mark's Basilica. I hope that an abstract and interpreted identity can be discerned from

them, which may aid in the ongoing construction of a dynamic identity of Venice. This identity exists in multiple, subjective ways in the minds of every Venetian and influences how they experience this beautiful city.

IDENTITY AND FUTURE OF VENICE. Generative Artworks

The transformations from Romanesque to Byzantine styles and beyond are interpreted through specific algorithms as transformation codes.

This work arises from a subjective interpretation of St. Mark's Basilica. St. Mark's Basilica was chosen because it symbolizes the representation of Venice that its citizens have crafted over time, closely tied to the growth and transformation of the city itself. Thus, it embodies the Venetian vision of the ideal city.

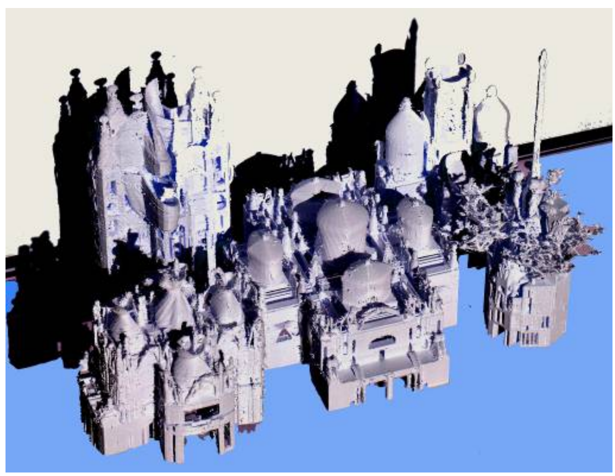
This distinctive, historically rich, and dynamic identity is explored through 3D geometric transformation logic. While not exhaustive of the Venetian identity, these transformations offer an example of how each individual—be an architect, a resident, or even a visitor—can contribute to an evolving system, propelling this identity towards its progressive development, shaping Venice’s 3D future toward its ideal form.

The 3D scenarios generated by this research works as a contribution to Venice's dynamic complexity and provide a fresh perspective on a dynamic identity, moving beyond a mere museum-like appreciation of the city.

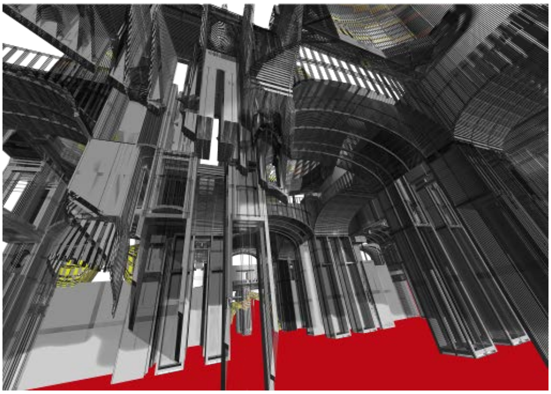
This experiment complements previous research, particularly the study conducted for Rome, which interpreted transformation codes from Classical to Baroque, tracing Francesco Borromini's dynamic approach. It was developed using Argenia, my personal generative artificial intelligence software.

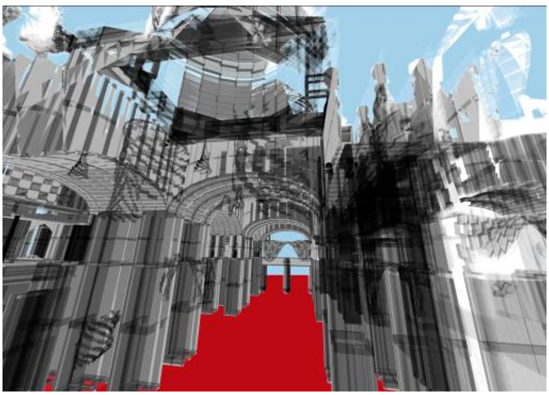
Argenia operates on a subjective vision of architectural identity, defined by abstract 3D geometric transformation codes that apply to any event, shaped by informal topological paradigms and an evolutionary dynamic. This approach links the events of the topological paradigm with transformative logic, activated sequentially or in parallel.

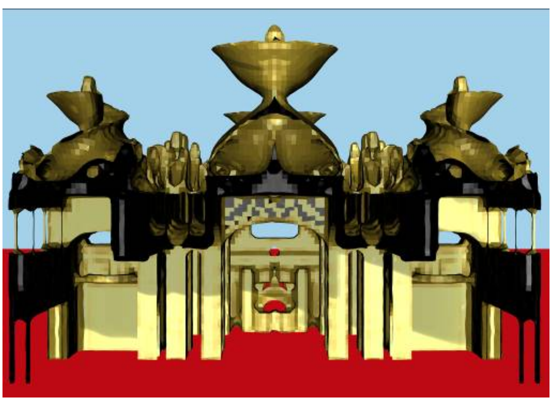
The architectural scenarios generated by Argenia for Venice are 3D models and were represented also through models created with a 3D printer.



5 generated and printed 3D models using ARGENIA. They are generated using my interpretation of St. Mark Basilica







Perspective cross-sections of 3D generated architectures following Venetian identity code interpretation (2024)

Their relevance was tested by placing these generated 3D architectures into Canaletto's painted scenes, a choice deemed most effective in assessing their potential to contribute to Venice's evolving identity, rather than relying on contemporary photographs of Venice.

*Six **generated 3D architectures** increasing dynamic Venice identity, inserted into **Canaletto paintings (2024)***

REFERENCS

2016 - Discovering the persistent deep memory in Generative Design. Duets, C.Soddu, GA2016

2016 - Generative Line, Enrica Colabella, GA2016

2017 - Designing Generative Art. Mosaic concept, Creativity, and Idea in Generative Design, C.Soddu, GA2017

2017 - Generative Art Language, Enrica Colabella, GA2017

2018 - AI ORGANIC COMPLEXITY in GENERATIVE ART, C.Soddu, GA2018

2018 - Generative art MetaLanguage, Enrica Colabella, GA2018

2019 - Subjectivity in Machine. The Future of Generative Art, C.Soddu, GA2019

2019 - Numbering, where trees move, Enrica Colabella, GA2019

2020 - Generative Art, FUTURING PAST, book edited C.Soddu and E.Colabella, Rome June 2020

2020 - Generative Art, Where Threes Move, book edited C.Soddu and E.Colabella, Rome September 2020

2020 - Generative Art & Design Theory, Methodology and Projects, book by C.Soddu and E.Colabella (English version of "Il progetto ambientale di Morfogenesi" 1992)

2020 - Generative Design, A Possibility To Teach Design By Following Students' Creativity And Their Subjective Vision by Celestino Soddu, GA2020

2021 - Argenia. Interp Years Of History To Live On Mars by Celestino Soddu, GA2021

2022 - Generative Art And Design. Visionary Algorithms For A Natural Complexity by Celestino Soddu, GA2022

2023 . Generative Art, The Author's Ai. Complexity, Naturalness, And Recognizable Creativity by Celestino Soddu, GA2023

All papers and books can be downloaded at
https://artscience.ebookshop.com
with all proceedings of:
https://generativeart.com
and the generated 3D architectures, objects, artworks and specific papers are at:
https://soddu.it and
https://generativedesign.com
more articles are in the website:
https://gasathj.com

BEAUTY *TRACES* for AI *NATURELESS*

Prof. Arch. Enrica Colabella

www.generativeart.com
enrica.colabella@generativeart.com


Abstract:
Human dust in endless variations.

Life looks so *wonderful* on the simple site, where happiness generates *nude* words in *infinite* extension, over any *beauty* prevision.

Beauty on AI boundaries between art and science. ***'What does the aesthetic experience of beauty consist of?'*** The answer will be an attempt to meet the challenge of *Kantian* aesthetics to ensure a union **between beauty and *reflection***, without suppressing the unequivocal link of the first with the feeling of pleasure.

Subject/Predicate
For Kant, beauty represents a very special predicate: it does not determine anything about the object of which it is predicate but rather expresses something about the same subject that makes the attribution in question. This is a cognitive judgment.

Visionary Beautyness fragments
Ode to Intellectual Beauty
…"Spirit of Beauty, which dost consecrate with thy hues all thou dost shine upon
Of human thought or form,-where art thou gone"….
Percy Bysshe Shelley [1]

"*THE HIDDEN GARDEN*" OF *GENERATIVE* WORDS

From the first bicycle time to our time of Mars' travels it is still alive the pleasure deepness of walking at down in a silent wood for hearing the birds first sounds of a new day.

An AI Poetic Voice from a Generative Past in Endless Variations
GA not-linear process. The discovering aim of this type of investigation is in trying to outline possible organic structures inside the past art experience. The main objective is to delineate ***continuity*** as a performing connector of a progressive process from the past and as a generative enlightening of possible scientific results**.**

Keywords for a GA Process aims: Three attributes
Starting point: three meta-textual words. By changing a word, we could therefore

change the very reality in which we live. So for every single part of the e-book "Morphogenetic Environmental Design. The Artificial DNA", 1992 [2] I decided to choose a fragment of a poetic text as a synthetic expression of the following theoretical text on GA meta-structure. Then for my teaching activity from 1996 at Politecnico di Milano, in a GA and Design process, I performed as students' main aim for gaining a complex character of their results 3 adjectives as three meta-textual words: 1+1=2>&1 and 1 (& indicates that what follows and precedes is a descriptor)

This generative design process, an homage to Leopardi, starts from the poet's texts in poetry and prose, following Pierce's logical concept of abduction.

We cannot generate anything if we do not fix in advance the characters of our aim. This was the first bridge built between the world of poetry and AI processes. The discovering aim of this type of investigation is in trying to outline possible organic structures inside the past art experience. The main aim is to delineate continuity as a performing connector of a progressive process as generative lighting of possible scientific results.

Abstract and figurative in a whole: poetry in prose [3]

The origin of the main structure of a generative art process is in defining a connection of poetry in prose. This is very ancient. It is possible to discover in the fable structure. The sounds of voices in fables can make the children's minds full of imagination. Fables configure a double site: one connected to the real world, the second one able to describe an imaginary world, but able to configure the reality. This is strongly similar to the creative process in GA. It tales the connection between a possible vision as an idea/code to an open algorithmic structure. This connection can generate a set of endless variations. There is a similarity between the open system of the telling voice and the variations of a generating reality. What is strongly identifiable is the code of tales, able to repeat variations during centuries of the same fable telling. I investigated the code of fable in my GA paper "*Mater Matuta*"[4]

This fable process was systemized as a popular voice connected to poetry in French, 1664, in Maximes by Francois La Rochefoucauld. This work opened a new line for preserving in written works the popular sapientia, able to generate similar but always in new ways over centuries. In poetry in prose is also "Le operette morali " by G. Leopardi, performed as a connection between myth and the real world. This process is still alive today in globalization, by preserving cultures through their not linear time process.

The Generative line of subjective life art

Valencia, November 2024 The long spontaneous line of helping people

The art of hearing

Lewis Capaldi had Tourette syndrome tics during one of his concerts.
Fans finished singing the song for him.

Lewis Capaldi's fans showed an immense level of love and support when the singer appeared to be overcome with tics due to his Tourette syndrome during a performance in Frankfurt, Germany singing *"Someone You Loved".* The audience's support was *a beautiful and heart-warming* reminder of the power of human connection, and will no doubt stay with Capaldi for years to come.

Language of Hidden Structures

"... There is a sense that remains hidden in the everyday but becomes accessible in the moments of greatest concentration, in the instants when consciousness loves the world. To capture that complex sense brings you an experience of particular happiness, to lose it consigns you to melancholy…"

Adam Zagajewski, Tradimento [5]

Orality&Poetry

Dante thinks poetry in the vernacular is the best tool to bring science to those who cannot sit at the 'Table where the bread of angels is eaten'. In this, he echoes Lucretius, the Latin poet who was convinced that 'poetry can help the spread of doctrines that are salutary for humanity'.

Vernacular sound

…ama il loro dialetto inventato ogni mattina,
per non farsi capire; per non condividere
con nessuno la loro allegria….
"…...love their invented dialect each morning,
not for being understood; not for sharing
with no one their joy...."
P.P.Pasolini …his last poem

…... The discovery aim of this type of investigation is to try to outline possible organic structures inside the past art experience. The main aim is to delineate continuity as a performing connector of a progressive process from the past, as a generative enlightening of possible scientific results.

The Generative Science of the individual mother tongue: the language of hidden structures in performing a GA process.

Vernacular sounds in Venice [6]

"I like crossing the imaginary boundaries people set up between different fields – it's very refreshing… There are lots of tools, and you don't know which is going to. It's about being optimistic and trying to connect things."
Maryam Mirzakhani, mathematician

AI and orality - Digital transformation of dialects and languages

Rather than re-creating humanity in AI, the AI discourse is redefining what it means to be human and how humanity is valued and should be treated. AI could work to translate also adapting oral poems into different languages and dialects, while keeping the original meaning and rhythm intact and preserving the richness of the tradition. The intersection between AI and oral poetry opens up very interesting and controversial scenarios. While AI can certainly enrich and innovate the creation and performance of oral poetry, there is a constant reflection on its ability to restore the authenticity and emotional depth that characterize the human oral tradition. In any case. AI can play an important role in preserving, transforming, and innovating oral poetry forms, at the risk, however, of losing some of the main characteristics that make oral poetry so unique and linked to a collective experience of gaining natural generative beauty.

Time references

…«O wonder! How many goodly
creatures are there here! How beauteous
mankind is! O ***brave new world***
(...HUXLEY) that has such people in't!»
William Shakespeare, *The tempest*
Act V, Scene I, 181-184

"*So when the great shell of knowledge will be completely filled, the immense ball of knowledge will continue to roll in space, but man will no longer be needed*"

"Ciò che viene sottratto all'uomo di oggi – da ogni partito, da ogni tecnica, da ogni conservatorismo -o riformismo o rivoluzionarismo – è né più né meno che l'amore"./ 'What is taken away from the man of today - by every political party, by every technique, by every conservatism- or reformism or revolutionarism - is neither more nor less than love'.

Eugenio Montale "Nel nostro tempo" [7]

TIME

Saint Agostino by Antonello da Messina

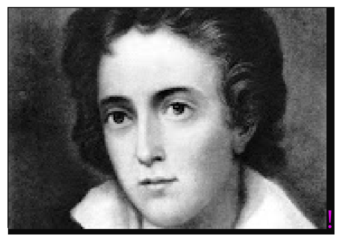
"*…so that there is, in reality, only one true reason to say that time is, except as it tends not to be so that, in reality, there is only one true reason to say that time is, except in as it tends not to be…."*
Saint Agostino,*"Le confessioni"*,XI,14 [8]

ABOUT"EGO":
THE PRICE *OF DISORDER*
Singular immersive
Alternative:
Immersive singular

Three/ Seventeen/ Eighteen:
Organic traces from Joseph Conrad
"Nobody remains the same. Men and women live as they dream. Alone"

The Conquest of Earth. *What redeems it is only the idea. An idea that supports it, not a sentimental pretext, but an idea and a disinterested faith, something, in short, to be exalted, to be admired, to which sacrifices can be offered.'..... And no bigger than pinheads on the untouched expanse of that immense hinterland. We dragged ourselves slowly…He absorbed the water from the palm of his hand, and then he sat in the sun, crossing his shins in front of him* ….

The tales time



"To a Skylark"
…Like a poet hidden
In the light of thought.
Singing hymns unbidden,
Till teh world is wrought
To sympathy with hopes and fears it heeded not…."
Percy Bysshe Shelley[9]

The generative hair around the things
Shelley perceived in the imagination an all-embracing and dynamic force, capable of transcending the unmoving and morally meaningless reality studied by science, until, by uniting with that scientific intelligence, he came to create an aesthetic ethos. Poetry, writes Shelley, is at once the centre and the

circumference of knowledge.

Generative Art

"When I say above that technique is the means of conveying an exact impression of exactly what one means, I do not by any means mean that poetry is to be stripped of any of its powers of vague suggestion. Our life is, in so far as it is worth living, made up in great part of things indefinite, impalpable; and it is precisely because the arts present us these things that we – humanity – cannot get on without the arts. The picture suggests indefinite poems, the line of the verse means a gallery of paintings, and the modulation suggests a score of metaphors and is contained in none: it is these things that touch us nearly that 'matter."

[10]Ezra Pound

Voice & tale

In our digital time, reading aloud fairy tales represents a natural medicine that can bring us back to slow, deep, and essential rhythms of listening in human contact. Whether one reads to a child, an adult, or even oneself. Listening to a recorded voice produces no reaction in the brain, on the contrary, the same tale told by a parent or another human being activates the area responsible for processing emotions. The fairy tale, therefore, is not only an educational tool that fosters language acquisition, but also a cognitive environment of deep emotional sharing, intimacy, and pleasure, for both children and adults. It is the human relationship able to constitute an adding value in this extraordinary form of reading.

Tales are exactly this: sunlight that sustains us, bringing us back to life, the real one.

The science of the individual: LEIBNIZ'S ontology of individual substance

De natura veritatis, contingentiae et indifferentiae, A VI.4, 1518

"There is no face ...whose contour does not form part of a geometric curve and cannot be drawn in one stroke by a certain regular movement." (Discours de métaphysique, § 6, A VI.4, 1538).

"But never can we reach by way of analysis the most universal laws [of our world], nor the most exhaustive rational explanation of singular things. Necessarily, indeed, this knowledge is reserved to God alone." [11]

Math Language

' To "mathematize" a problem does not mean only to measure and calculate, but to reveal as HYPOTHESIS a hidden skeleton of conceptual relations linked to the observing properties of the universe. Mathematization consists precisely in formulating an underlying idea of a possible connection between apparently different physical properties in an abstract math language. What is important is that the axiomatic-deductive system, but also interpretative-ipothetical one is the most powerful working.

Cathedrals as time *visionary* thinking*"*

Beauty is truth, truth beauty that is all / Ye know on earth, and all ye need to know."

John Keats [12]

This is a beautiful way for the USA to talk about the future, referring to cathedrals because when they placed the foundation stone knew very well that they would not see the finished cathedral. We must now realize that the environmental and biodiversity crisis with climate change are our cathedrals. We must put the foundation stone now for human future

CHATHEDRALS *perhaps* endless open numbers

La terra è stretta / The earth is confined
Si schiude in curva senza virgole: / It unfolds in curves without commas:
Effonde certezza la continuità, / Certainty spreads continuity,
Elimina l'immaginario di paura del dietro l'angolo:/ It eliminates the imagery of fear about the behind corner:
Fonte di dubbi e tremori illimitati fantasiosi. / Source of doubts and unlimited fantasy tremors.
Se mi guardo i piedi ad uno ad uno entrando in un fiume / If I look at my feet one by one entering into a river
Ancora a riva, tra la terra e l'acqua, non temo l'ignoto, / Still on the bank, between land and water I do not fear the unknown,
Il visibile mi rassicura producendomi gioia come d'attesa felice. / The visible reassures me by producing joy like happy expectation.
Un frammento dove c'è posto per l'immaginario eco infantile di scoperta. / A fragment where there is place for the imaginary childish echo of discovery
Piccoli passi/ Small steps
Arrivi al fiume / You arrive at the river
Ti bagni i piedi:/ You wet your feet:
L'uno dopo l'altro……./ One after the other.......
A discovery of time uniqueness

***"*THE *OUTCORDED* SONG OF AN ARTIFICIAL CHICKEN**

..... Human history has a beginning/end. AI iterates the process in/out with open sequences in endless time. Human life is what builds the artificial world, but it is destined to end. Human art history worked by giving meaning to death through art…The builders of a finite world**:** human history ends from Cartesio to Kant, Nietche **….**Fitzgerald in *Tender is the night* follows in alcohol this assumed life sense opening a long line to drugs**.**

Time/ Variations: *Overview* of the end

Arnold Bòckiln: four variations +1.[13] From a quiet place to the island of the dead. Silence: A painting to dream of the island of the living people.......





On the *other* hand

Guardati le mani….
Dalla parte delle unghie. /
Look at your hands…
On the nail side
Un urlo: attenta all'insieme! /
A shout: beware of the wholeness!
Invece corriamo veloci al dettaglio, /
Instead, we run fast to the detail,
sempre inconsci di un perché./
always unaware of a why.
L'abitudine è segno di controllo,
prevenzione o nemesi della paura? /
Is habit a sign of control, prevention or
the nemesis of fear?
Oppure siamo ingordi di perfezione nel
superare sempre l'altro su un misero
dettaglio? / On the other hand, are we
only gluttons for perfection in always
surpassing the other on a miserable
detail?

Dante

Terzine from Dante's Paradise, Canto XXII, verse 88, says:

"La mente innamorata, che donnea
Con la mia donna sempre, di ridure
Ad essa gli occhi più che mai ardea"
"The mind in love, always wandering to
my woman, longed more than ever to
return my eyes to her" [13]

The mind is one of the greatest mysteries of the universe. The value of the human being is in the value of what he seeks from what he is interested in. An emblematic example is the figure of Ulysses that Shakespeare tells us about in *Troilus and Cressida*. In the measure that it loves, our mind is attracted to the object of its love just as the earth is attracted to the sun. But it can also be something not human, as in the case of passion for art, for research, science or for religious life.

The conclusive text of Kant's *Critique of Practical Reason* says that there are essentially:

"Two things that fill the soul with admiration and veneration: the starry sky above me, and the moral law in me".

So what saves us from destruction, as Hannah Arendt says, from being a pinpoint in the infinity of the universe is this invisible ego capable by alone of contrasting himself with the infinite universe.

This is the mind in love; it is it that saves us.

AI Poems in the *past* poets style

A new study in the journal *Scientific Reports* [14] finds that non-expert readers can't reliably distinguish between poems penned by William Shakespeare, Emily Dickinson, T.S. Eliot or Sylvia Plath and ChatGPT-3.5 doing its best impression of each of them. More surprising, readers preferred the AI-generated poems — and were more likely to guess those were written by humans than real works by famous poets.

Di notte**:** *fragile***/** By night: *fragile*

Come curvi il tuo collo per girare gli occhi verso l'orizzonte/ How you curve your neck to turn your eyes towards the horizon
E' un gesto antico nel carattere del tuo tempo: un attimo fermo nel mio sguardo/ It is an ancient gesture in the character of your time: a moment standing in my vision
Non l'infrango estraendolo dalla sua fragilità in uno scatto da esibire*/* I don't break it by extracting it from its fragility in a shot to exhibit
Lo trattengo nel mio cuore per consolare un mio silenzio senza lacrime: di notte. / I hold it in my heart to console a tearless silence of mine: by night


Venice: Ezra Pound "Cantos XVII"

"Flat water before me,
and the trees growing in water,
Marble trunks out of stillness,
On past the palazzi,
in the stillness, The light now, not of the sun"
The canto ending is concerned with Venice, which is portrayed as a stone forest growing out of the water.

References:

[1] https://www.poetryfoundation.org/poems/45123/hymn-to-intellectual-beauty
[2] C. Soddu – Enrica Colabella "Morphogenetic Environmental design. The Artificial DNA"- www.accademia.edu https://artscience-ebookshop.com
[3] https://soddu.it/papers/papermathpoetry.pdf https://generativeart.com/GA2018_papers/GA2018_Enrica%20Colabella.pdf https://generativeart.com/on/cic/GA2010/Sensus-Communis_E.Colabella.pdf
[4] Mater Matuta https://generativeart.com/on/cic/2000/colabella.htm
[5] Adam Zagajewski, *Tradimento,* Adelphi, 2000
[6] Vernacular *sounds* in Venice https://www.televignole.it/novelle-dialetto-veneziano-2/ https://www.youtube.com/watch?v=aF9YvCtyJN0
[7]Eugenio Montale " *Nel nostro tempo*" Rizzoli page 43
[8] Saint Agostino, *Le confessioni*, https://formazione.indire.it/paths/agostino-le-confessioni-xi-14-e-18

[9]
https://www.poetryfoundation.org/poems/45146/to-a-skylark
[10] Ezra Pound. *"I Gather the Limbs of Osiris."* 1911. SP 3355
[11]
https://download.e-bookshelf.de/download/0000/0032/67/L-G-0000003267-0002368802.pdf
[12]https://www.poetryfoundation.org/poems/44477/ode-on-a-grecian-urn
[13]
https://www.theartstory.org/artist/bocklin-arnold/
[14]
https://digitaldante.columbia.edu/dante/divine-comedy/paradiso/paradiso-22/
[15]
https://www.washingtonpost.com/science/2024/11/14/chatgpt-ai-poetry-study-creative/

City Paths, Mazes, Intelligence, and Generative Design

Marie-Pascale Corcuff, Architect, PhD.
Rennes, France
e-mail: m-p.c@wanadoo.fr

Abstract

Walking through a city, going all over its paths, is a great part of discovering its spirit. Venice doubles this experience by interlocking walking paths and canals.

Urban networks are a first kind of mazes. A second type originates in labyrinths, and "solving" such mazes has been used as a test of "intelligence", or at least of learning capability, for a long time.

Such mazes are space-filling branching curves. How to produce generic mazes, and then self-similar mazes, and use them for generative design will be discussed in this paper.

1. City paths as mazes

According to the Cambridge dictionary, a maze is: "a complicated system of paths or passages that people try to find their way through for entertainment", or "an area in which you can get easily lost because there are so many similar streets or passages: *The old part of the town was a maze of narrow passages.*" [1]

We shall discuss the first definition later, but, regarding the second one, Venice is certainly an example of such a maze. If you quit the areas overcrowded with tourists, you are sure to rapidly get lost with the comforting landmarks such as Piazza san Marco out of view. It does not help that, while *sestieri*, *campi*, *corti*, *calli,* and *fondamente* have names, house numbers, though existing, are not much help: "Even if you have the address it can be difficult due to Venice's somewhat Byzantine way of numbering the houses. Numbers seem to change without logic. And look for number 1 in most streets and you're not going to find it." [2] There are still more confusing systems, though. In Japan, streets have no names, it is the blocks that are relied on and are identified with numbers, and inside them buildings are numbered too. Add to that the difficulty for an European, for instance, to even decipher the signs, and you are sure to rapidly get lost.

The above definition of an urban maze does not mention the shape and disposition of streets. It is obvious however that we consider contorted streets more adequate to the definition of an old town as a maze. An orthogonal grid is easier to navigate, and straight streets allow the view towards specific landmarks. Even only departing from the orthogonality may be confusing. There is a Strasbourg neighbourhood named Hautepierre, designed in the nineteen sixties, along a hexagonal grid that was supposed to optimize the balance between traffic and calm isolated areas. However, navigating such a system is not as easy as it seems. Turn 120° in any direction a few times and you are more disorientated than if you had done 90° turns.

The exact opposite of a maze is the common American town. Here, even if most streets are similar, they often are laid upon an orthogonal grid, and numbering instead of naming streets, though not very imaginative, is very helpful. You always know how many blocks you have to pass before arriving at the street you want to go to.

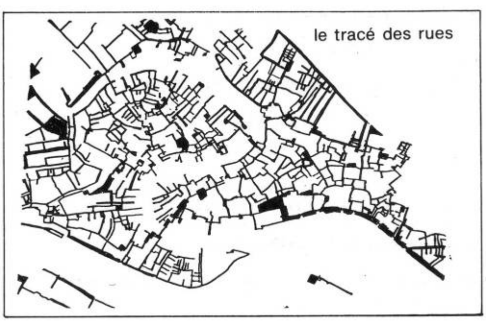


Fig. 1: Venice streets [3]

Going back to Venice, and looking at its street network (Fig. 1), one notices its particularity at once: there are a lot of dead ends abutting to a canal or to the lagoon itself. They are not usual *culs-de-sac* because the view is very open, and you could go on if you had a boat or were willing to swim... And there are sometimes *fondamente* that allow you to continue your journey along the water. But it is anyway very peculiar, most streets in any other town being prolonged into the surroundings. In Venice, there is only one path connected to the continent. Also, excluding La Giudecca, the street network is totally connected, all “islands” being linked by bridges. You have to be sure to find them, though, especially to cross the Grand Canal, which has only three bridges.

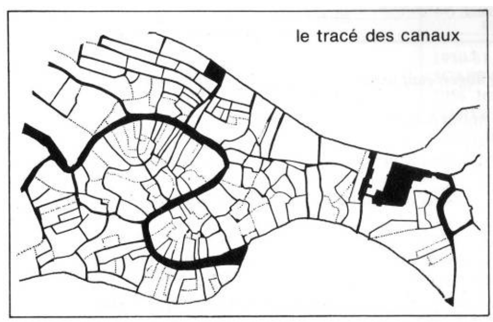


Fig. 2: Venice canals [3]

The canals constitute another network (Fig. 2) totally connected with the lagoon. Looking closely, we observe that there are no dead ends. Apart from the sinuous Grand Canal and a few other ones, this network is basically a grid, i. e. totally constituted of loops.

The definition of the network of canals is easy, there is a clear distinction between water and firm ground. Defining the network of streets is less straightforward. What is the counterpart of it, and should we not consider the opposition between built *vs* unbuilt instead, or between public *vs* private space? The way the Japanese address system works enlightens a different concept than the European one: buildings are the positive part of space, streets and any other unbuilt areas are the leftover, the "empty" space.

Anyway, considering the urban mazes of streets, they may be disorientating for diverse reasons, and constitute complicated networks, with loops and dead ends. But they also tend to "irrigate" the city the most efficiently: each individual front door must be reached eventually, we could even consider the corridors and stairways inside buildings, and view the urban maze as the equivalent of the vascular system of a living organism.

2. From labyrinths to mazes

Considering now the first definition of a maze as "a complicated system of paths or passages that people try to find their way through for entertainment", for which the Cambridge dictionary gives "labyrinth" as a synonym, one must look as this topics more closely.

Historically, from ancient Greece through medieval times, a true labyrinth is a folded unicursal path, without any loop or branching, and so without any chance of getting lost, without any need of Ariadne's thread or any other trick [4]. The only puzzling thing about it is how the myth of Theseus and the Minotaur is however related to it, even in Christian times as shows the inscription (in Latin) beside the finger labyrinth engraved upon a pillar of Lucca cathedral shows: "This is the labyrinth built by Dedalus of Crete; all who entered therein were lost, save Theseus, thanks to Ariadne's thread." (Fig. 3)

Fig. 3: Lucca labyrinth

After having been exclusively religious in the Middle Age, labyrinths were adopted by higher society for entertainment. But what fun is there in a unique path where nobody can get lost? So they evolved into mazes, departing from church pavings towards hedges and bush gardens. The definition is still loose, and many such mazes are often called labyrinths, such as the one built by Le Nôtre for Louis XIV between 1665 and 1673 in the Versailles gardens, and destroyed under Louis XVI (Fig. 4).

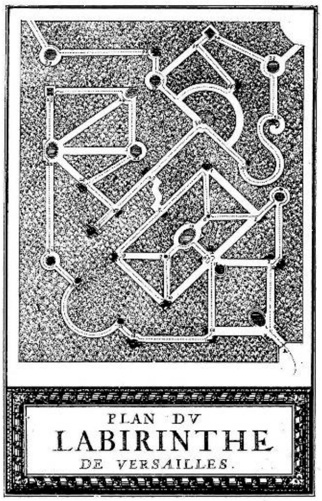


Fig. 4: labyrinth in Versailles garden

This configuration of paths does not truly qualify as a maze, though, and even less as a labyrinth. One essential component of labyrinths and mazes is their density of paths. A better example would be the Hampton Court Palace maze, created between 1689 and 1695 for King William III (Fig. 5).

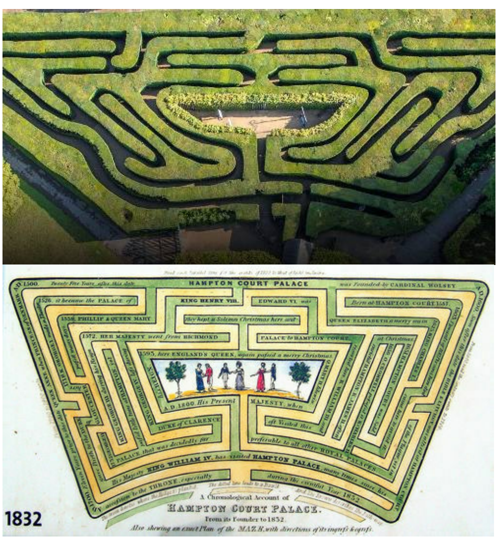


Fig. 5: Hampton Court Palace maze

In this case we see the direct heritage from the labyrinth, with only one entry, an arrival in the centre, and the intricate folding paths. The only difference with a true labyrinth is that there are dead ends.

A recent maze is the Borges (so-called) labyrinth, in Venice (san Giorgio Maggiore), designed by Randall Coate in 1979, but open in 2011 for the 25th anniversary of Jorge Luis Borges' death (Fig. 6). It is a nice design, but maybe not actually true to Borges, who suggests in his story "The Two Kings and the Two Labyrinths" that the best and most cruel labyrinth is not some elaborate configuration of paths, but the desert itself.

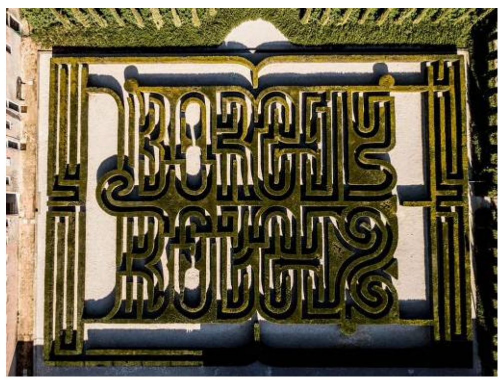


Fig. 6: Borges labyrinth in Venice

After being an amusement, mazes were harnessed by scientists to assess the memory and learning processes in mammals, mostly rodents. A very large variety of devices have been invented [5], but the first to have been used was a quadrangular equivalent of the Hampton Court maze, by William S. Small in 1901 (Fig. 7)

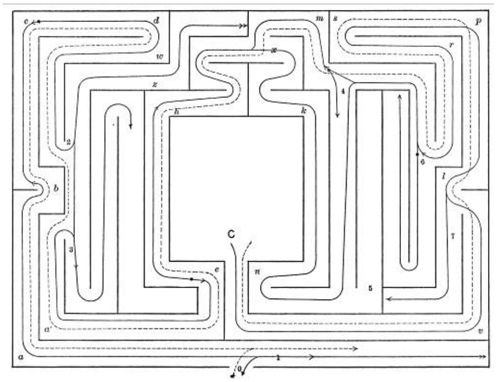


Fig. 7: experimental maze [6]

The key point in all those tests is to "solve" the maze: some reward (food) is put at a crucial place, and the rat must learn, first, that there is a path leading to that reward, and then, placed in the same situation a number of times, to use the most efficient, or shortest, path, by avoiding the dead ends. This is accomplished by constructing "a sequence of directions in its memory".

The most simple configuration, which we would not even call a "maze", is a branching Y or T shape, with the reward at the end of one of the branches. It essentially requires to be able to distinguish one's left from one's right, and remember which one is the "good" direction. Which is more than some of us humans can do, though...

Throughout all the different configurations imagined to annoy rodents there is one that is particularly interesting from an architectural point of view: the vertical maze (Fig. 8). It was designed by Albert Walton in 1930 to allow him to show the different stages of learning of rodents in front of a class of students.

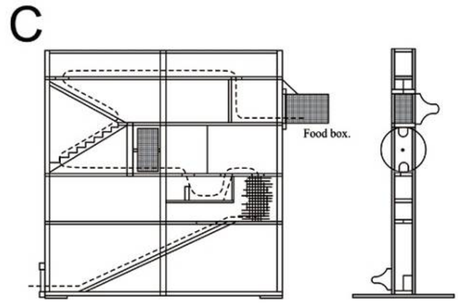


Fig. 8: vertical maze [7]

Aside from being reminiscent of some extreme narrow houses in contemporary Japan, this type of vertical maze has been adopted for humans, or at least children, in some architectural situations.

Lappset, a playground facilities company has designed a vertical maze, actually installed in the elementary school Parmentier in Paris (Fig. 9) [8]

Fig. 9: vertical maze in Parmentier school, Paris

In a more high-level area of architecture, the internationally recognized MVRDV's project KU.BE House of Culture and Movement (Denmark, 2016) provides many cultural spaces, but it is the route through the building that is the most peculiar. Among other adventurous devices, "the Mousetrap, a vertical maze, provides a challenging transition from floor to floor.” [9] (Fig. 10)

Fig. 10: the mousetrap in KU.BE, MVRDV.

Mazes are used for studying human abilities too, as in [10] (Fig. 11).

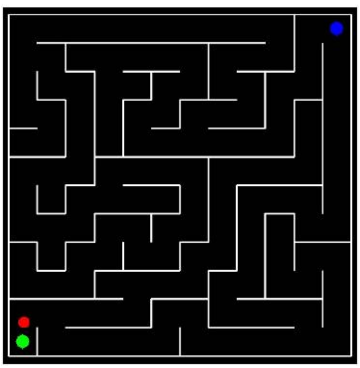
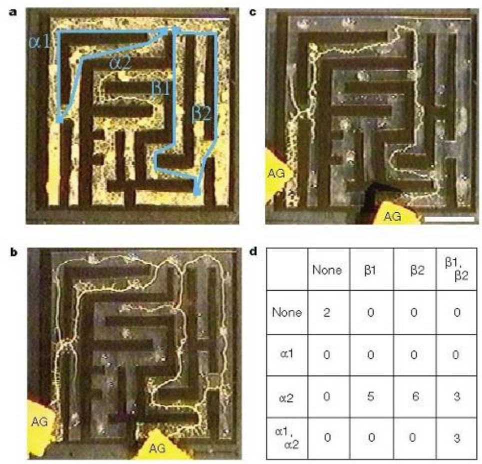


Fig. 11: experimental maze for humans

In all human experiments I could find, the subjects are not placed “in” the maze as rodents are, they look at a maze, generally on a computer screen. It is a huge difference, in my opinion, because rodents must rely only on their proprioception, without viewing the entire maze, while humans see the maze, and navigate it visually, but without moving their body.

Anyway, here we see the ultimate configuration, and the most accepted one, of a “maze”: it is based on a grid, it is dense, it comports branchings but no loops, and there exists a single path from entry to arrival. Solving such a maze, for a human, implies (visually) trying paths and revoking dead ends. One can say that it is a form of “intelligence”, but such mazes can also be solved by Physarum Polycephalum, a protist without any brain [11] (Fig. 12).



	None	β1	β2	β1, β2
None	2	0	0	0
α1	0	0	0	0
α2	0	5	6	3
α1, α2	0	0	0	3

Fig. 12: Maze-solving by Physarum polycephalum

Such a maze was designed by Bjarke Ingels in 2013 in the National Building Museum (Washington, D. C.) [12] (Fig. 13)

Fig. 13: National Building Museum maze

If we look at its layout (Fig. 14), we see a discrepancy: there are two entries, if we consider the four cases at the centre to be the arrival.

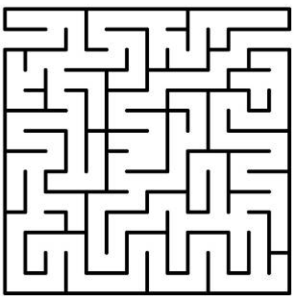


Fig. 14: National Building Museum maze layout

Such mazes are what qualifies for the use of the legendary Ariadne's thread, which metaphorically stands for "solving a problem which has multiple apparent ways to proceed - such as a physical maze, a logic puzzle, or an ethical dilemma - through an exhaustive application of logic to all available routes." [13] This article enhances the difference with the "trial and error" method, with which it can easily be confused, and is the true scientific method. Both processes however qualify as "intelligent" methods, based upon logic and reasoning.

3. Generating mazes

The definition of a maze we shall use from now on is: a maze is a tree (a branching network without any loop) on a grid, where any cell of the grid is reached. By this definition it is obvious that one can go from any place (or case of the grid) to another, and that there is a shortest path between those two cases.

3.1 Generic algorithm

There are a number of algorithms that generate such mazes [14]. I wrote my own version of the so-called Wilson algorithm. It consists in generating a unicursal path until it is impossible to go on (because the path has folded on itself), then doing the same thing by starting anywhere on the previous path, and so on (Fig 15).

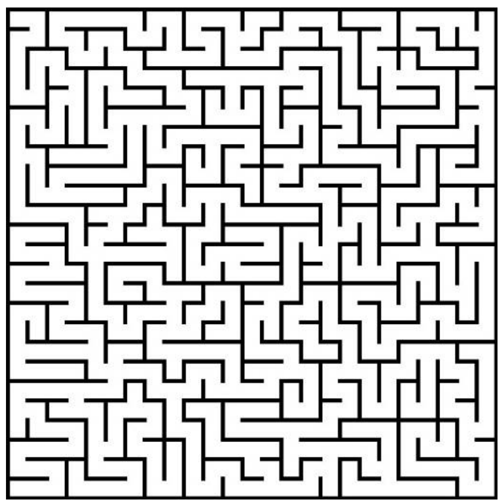


Fig. 15: generic maze on a 25 x 25 grid

We can highlight its construction by colouring, and diminish the width of the paths according to the history of their generation (Fig 16).

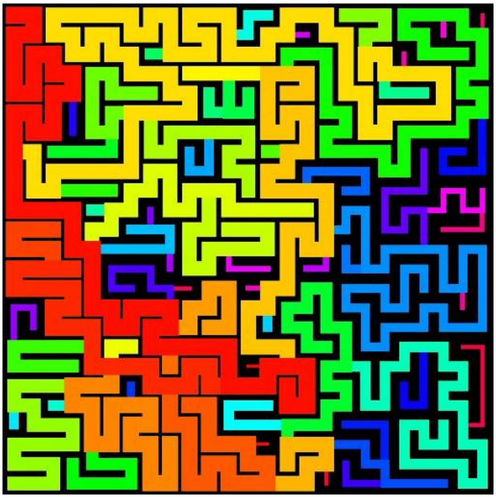
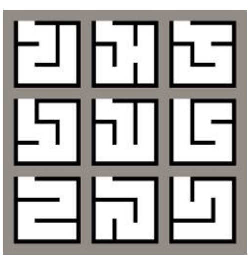


Fig. 16: generic maze on a 25 x 25 grid: construction

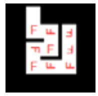
3.2 Self-similar mazes

Mazes as we have defined them share a lot of characteristics with what are called FASS curves, which stands for space-Filling, self-Avoiding, self-Similar, Simple curves. The only difference is “simple” which means that the curve is drawn in only one stroke. Mazes instead are branching. We can try to make them self-similar, in order to get “FASB” curves.

We start by generating generic mazes on a 3 by 3 grid, which will be the “seeds” of our self-similar mazes:



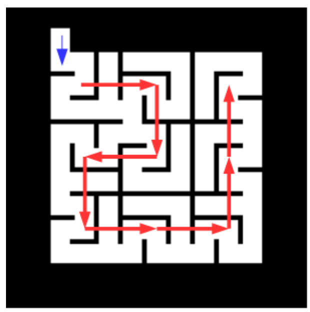
Let us take the first sample and analyse it in terms of translations and rotations:



We deduce a L-system from this analysis, adding the fact that we consider a linking path at the start of each pattern:

ω: F
F → [TF-YF+YF+YF-YF-YFYF-YFYF]
T → TTT
Y → YYY
with those interpretations:
F: draw line (0,-l) → (0,0)
T: translate (-l,-l)
Y: translate (0,l)
+: rotate right
-: rotate left
[: start branch
]: end branch

This gives us, at the second step:



And then we see the result at the third step, a self-similar maze on a 27 by 27 grid (Fig. 17):

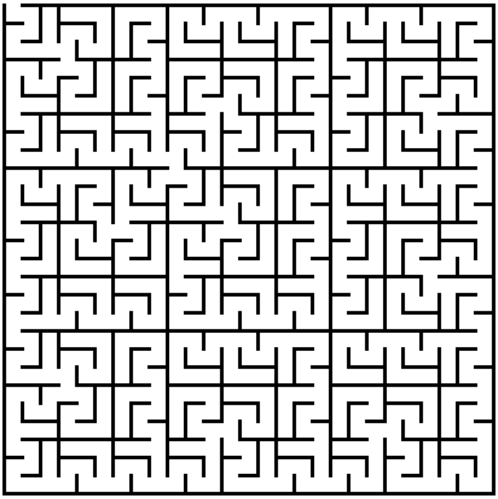


Fig. 17: self-similar maze on a 27 x 27 grid (seed 1)

A similar analysis of the second sample allows us to construct another L-system which will provide another self-similar maze (Fig. 18). This new seed is itself a branching system. But let us remark that, whether the seed is branching or not, the maze we obtain is a branching system by construction. Let us add that this method is easier than those producing FASS curves, which have much more constrains.

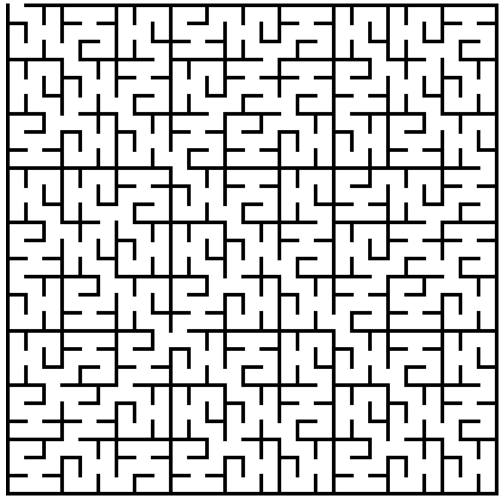


Fig. 18: self-similar maze on a 27 x 27 grid (seed 2)

We could generate many other self-similar mazes with different 3 by 3 seeds, and also with 4 by 4, 5 by 5, etc., seeds.

We can compare such mazes with a Hilbert FASS curve (Fig. 19), to acknowledge their resemblance.

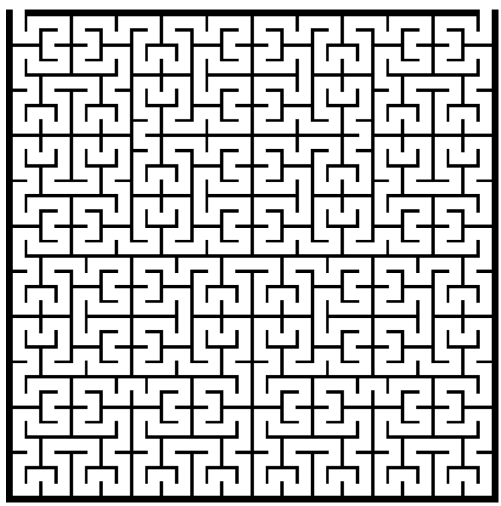


Fig. 19: Hilbert curve

3.3 Extrapolations

The resemblance between self-similar mazes and FASS curves brings the idea of treating them as we did for FASS curves, i. e. applying a Cartesian to polar transformations in order to obtain true labyrinths [4][15] . Let us recall that FASS curves share a lot of characteristics with labyrinths, except that their path starts at one corner of the square they develop in and ends at another corner, while in labyrinths the path starts on the perimeter of a circle (or any other shape) and ends at the centre. The transformation takes care of that discrepancy. Applied to a self-similar maze, we obtain this new centred maze (Fig. 20):

Fig. 20: centred maze

There would be many generative design ideas to be explored from the preceding explorations. Going from two to three dimensions will be our next endeavour, as well as exploring the relevance of mazes in urban and architectural design.

References

[1] dictionary.cambridge.org/fr/dictionnaire/anglais/maze

[2] https://lucasitaly.com/2019/02/01/venetian-house-numbers/

[3] Léon de Coster, Georges Vranckx*, 16 promenades dans Venise,* éditions universitaires, 1987.

[4] Marie-Pascale Corcuff, "From Labyrinths and Recursive Folds Towards Generative Architecture." 15th Generative Art Conference Proceedings, Lucca, Italy, December 10-12, 2012, pp. 223-238. https://www.generativeart.com/GA2012/marie-pascale.pdf.

[5] Fahimeh Mohseni, Shahram Ghorbani Behnam, Raheleh Rafaiee, "A Review of the Historical Evolutionary Process of Dry and Water Maze Tests in Rodents." *Basic and Clinical Neuroscience, 11*(4), 389-402, 2020. http://dx.doi.org/10.32598/bcn.11.4.1425.1

[6] Willard S. Small, "Experimental Study of the Mental Processes of the Rat.", The American Journal of Psychology, Jan. 1901, Vol. 12. N°12, pp. 206-239. https://www.jstor.org/stable/1412534

[7] Albert Walton, "Demonstrational and experimental devices." The American Journal of Psychology, 42(1), 109-14, 1930. DOI:10.2307/1414428

[8]

https://www.lappset.com/en-GB/reference/a-playground-with-vertical-maze-amazes-in-france]

[9] https://www.mvrdv.com/projects/50/kube-house-of-culture-and-movement

[10] Min Zhao, Andre G. Marquez, "Understanding Humans' Strategies in Maze Solving", 2013. arXiv:1307.5713v1

[11] Toshiyuki Nakagaki, Hiroyasu Yamada, & Agota Tóth, "Maze-solving by an amoeboid organism." *Nature* **407**, 470, 2000. https://doi.org/10.1038/35035159

[12] https://www.smithsonianmag.com/travel/winding-history-maze-180951998/

[13] https://en.wikipedia.org/wiki/Ariadne%27s_thread_(logic)

[14] https://en.wikipedia.org/wiki/Maze_generation_algorithm

[15] Marie-Pascale Corcuff, "Labyrinths and Space-Filling Curves, Spirals and Tessellations: Topological and Geometrical Implications of Cartesian to Polar Transformations", Proceedings of Bridges 2024: Mathematics, Art, Music, Architecture, Culture. pp 107–114 https://archive.bridgesmathart.org/2024/bridges2024-107.html#gsc.tab=0

Spooklight: A Generative Approach to AI-Driven Storytelling

B.T. Franklin, BASC
Independent Researcher, Phoenix, AZ, USA
www.btfranklin.info
e-mail: brandon.franklin@gmail.com

Abstract

In the rapidly evolving landscape of artificial intelligence, the intersection of generative processes and creative expression presents new possibilities for narrative and visual art. Spooklight is an innovative tool designed to explore these possibilities by utilizing a multimodal generative AI framework to create dynamic, image-driven stories. The core of Spooklight's methodology lies in an iterative loop where each phase of the story alternates between the generation of descriptive images and the crafting of corresponding textual narratives.

This approach leverages the power of AI to infuse traditional storytelling with a stochastic element, ensuring that each narrative is unique, evolving organically through the interaction between visual and textual prompts. By blending AI-generated images with narrative progression, Spooklight demonstrates how generative AI can augment human creativity, offering a new form of artistic expression that is both unpredictable and richly detailed.

Spooklight's generative process does more than simply create; it engages in a dialogue between the visual and the literary, crafting stories that are deeply intertwined with the imagery that inspires them. This paper explores the underlying architecture of Spooklight, the role of multimodal AI in creative processes, and

the implications of using generative tools to push the boundaries of narrative art. Through this exploration, we aim to contribute to the broader understanding of how AI can be harnessed to preserve and evolve the complexity of human storytelling in the digital age.

1. Introduction

1.1 Background on Generative AI in Art and Storytelling

In recent years, artificial intelligence (AI) has made significant strides in creative domains, profoundly impacting visual arts, music, and literature. Generative AI models, particularly those employing deep learning techniques, have enabled machines to produce content that closely mimics human creativity. In the realm of storytelling, AI has been harnessed to assist writers in generating ideas, crafting narratives, and even composing entire stories. The advent of multimodal AI models—which can process and generate both text and images—has further expanded the horizons of creative expression, facilitating a seamless interplay between visual and textual content.

The evolution of AI in creative processes has been marked by the development of sophisticated language models capable of understanding context, style, and nuance. These models have opened up new possibilities for artists and writers, allowing for collaborative interactions between humans and machines. As AI continues to evolve, its role in art and storytelling prompts critical examination of its capabilities and limitations, as well as its impact on human creativity.

1.2 Motivation for the Spooklight Project

The Spooklight project emerges from a fascination with the potential of AI to not only replicate but also enhance human creativity by introducing stochastic elements into the narrative process. Inspired by the enigmatic folklore of the will-o'-the-wisp—a spectral light said to lead travelers into unknown territories—Spooklight aims to explore the depths of generative storytelling through an iterative loop of image and narrative generation. By alternating between AI-generated images and corresponding textual narratives, the project seeks to create dynamic stories that evolve organically, guided by the interplay of visual and textual prompts.

The primary motivation behind Spooklight is to investigate how a multimodal AI framework can be leveraged to produce rich, image-driven narratives that push the boundaries of traditional storytelling. By injecting externalized stochastic processes into the narrative generation, Spooklight introduces elements of unpredictability and diversity, resulting in stories that are unique and richly detailed. This approach not only showcases the creative potential of AI but also highlights how technology can augment and inspire new forms of artistic expression.

1.3 Overview of the Paper

This paper presents the underlying architecture and methodology of Spooklight, detailing how multimodal AI models are utilized to produce intertwined visual and textual narratives. We begin by situating Spooklight within the context of existing AI storytelling tools, providing a comparative analysis to highlight its unique contributions to the field. The methodology section delves into the iterative process of the tool, explaining the algorithms and configurations that enable the dynamic generation of stories.

Furthermore, we discuss the challenges encountered during the development and execution of Spooklight. These include maintaining narrative coherence, preventing repetitive language, and ensuring thematic diversity—issues that are critical in the context of AI-generated content. By examining these challenges, we aim to provide insights into the limitations of current AI models and propose potential solutions for future enhancements.

Through this exploration, the paper contributes to the broader understanding of how AI can be harnessed to preserve and evolve the complexity of human storytelling in the digital age. It underscores the significance of merging visual and textual narratives through AI and reflects on the implications for the future of creative processes.

2. Related Work

2.1 Existing AI Storytelling Tools

The integration of AI in storytelling has led to the development of several innovative tools and platforms designed to assist writers and creators. The following are notable examples.

Sudowrite is an AI-powered writing assistant aimed at helping authors overcome writer's block and enhance their storytelling. By providing text prompts, users can generate ideas, characters, and plotlines to enrich their narratives. Sudowrite leverages advanced language models to offer suggestions that align with the writer's style, facilitating a collaborative writing process where AI acts as a creative partner rather than just a tool [1].

Tome is an AI-driven storytelling platform that combines a fluid, interactive canvas with an AI partner. It enables users to create and share immersive narratives, integrating features like DALL·E 2 image generation to enhance visual storytelling. Tome allows for the seamless blending of text and images, providing a more engaging and dynamic storytelling experience. The platform emphasizes collaboration between the user and AI, offering a space where ideas can be rapidly prototyped and expanded [2].

Depthtale is a platform for creating and playing immersive AI-powered visual novels. It allows users to craft complex narratives with interactive characters, each possessing their own stories and motivations. Depthtale adds depth to the storytelling experience by enabling a high degree of interactivity and personalization, where the narrative evolves based on user choices and character interactions [3].

Storynest.ai is an AI-powered platform for crafting stories, novels, and interactive fiction. It offers AI image generation capabilities to create character portraits, scene illustrations, and storyboards, enhancing the visual aspect of storytelling. Storynest.ai aids writers in visualizing their narratives and provides tools to develop detailed and immersive worlds, bridging the gap between textual and visual storytelling [4].

2.2 Comparison with Spooklight's Approach

While these tools significantly contribute to the field of AI-assisted storytelling, Spooklight differentiates itself through its unique iterative process that tightly integrates image generation and narrative development. Unlike Sudowrite, which primarily focuses on text generation to assist writers, Spooklight employs a multimodal approach where images and narratives influence each other in a cyclical manner. This continuous interaction creates a dynamic storytelling experience where each output informs the next, leading to an organic evolution of the narrative.

Tome and Storynest.ai incorporate AI-generated images into storytelling, but their primary function is to enhance user-created narratives with visual elements. Spooklight, on the other hand, places equal emphasis on the images and the narratives, using each to inform the generation of the other without direct human input at each step. This results in stories that are less constrained by initial human design, allowing for greater unpredictability and originality.

Depthtale offers interactive visual novels with AI-driven characters, emphasizing user interaction within a predefined framework. In contrast, Spooklight allows for the emergence of narratives that are not bound by predefined paths, evolving organically through the stochastic interplay between visual and textual elements generated by AI.

By focusing on the synergy between image and text, Spooklight extends the capabilities of generative storytelling tools. It demonstrates how AI can be harnessed to create complex, intertwined narratives without extensive human intervention, offering a novel approach that emphasizes the dynamic evolution of stories.

2.3 Positioning Spooklight in the Context of Generative Art

Spooklight contributes to the field of generative art by showcasing the potential of multimodal AI models to act as both storyteller and illustrator. It exemplifies how AI can transcend traditional boundaries of creativity, producing works that are not limited by human cognitive patterns or biases. By allowing narratives to evolve through a stochastic process, Spooklight introduces an element of serendipity and exploration into storytelling.

This approach aligns with the principles of generative art, where systems are designed to produce unpredictable and emergent outcomes. Spooklight's methodology pushes the boundaries of narrative art, offering insights into how AI can augment human creativity and inspire new forms of artistic expression. It opens up discussions on the role of AI in the creative process and its potential to reshape how stories are conceived and experienced.

3. Methodology

3.1 Algorithm Overview

Spooklight employs a cyclical process where image generation and narrative creation influence each other iteratively. The core algorithm can be summarized in the following steps:

1. **Initialization:**
    - **Story Concept:** The process begins with either a provided story concept, a starting image, or a starting image description.
    - **Author Style:** A random author's style is selected to guide the narrative voice.
    - **Visual Style:** If a starting image is provided, its visual style is extracted; otherwise, a visual style is generated based on the initial image description.
2. **First Step Generation:**
    - **Image Description:** An initial image description is generated if not provided.
    - **Image Generation:** Using the image description and visual style, the first image is generated.
    - **Narrative Generation:** A narrative is crafted based on the

story concept, author style, and the generated image.

3. **Iterative Story Loop:**

- **Image Description Generation:** For each subsequent step, a new image description is generated based on the story concept and the narrative so far.
- **Image Generation:** A new image is produced from the image description and the established visual style.
- **Narrative Generation:** The narrative is extended using the new image and the story's context.
- **Termination Check:** The loop continues until a predefined story length is reached or a natural conclusion is detected.

4. **Completion:**

- **Title Generation:** A title for the story is generated based on the complete narrative.
- **Compilation:** All images and narratives are compiled into a PDF document, forming the final story.

This process allows the story to evolve organically, with each phase building upon the previous, ensuring a tight integration between the visual and textual elements.

3.2. Technical Implementation

3.2.1 Multimodal Large Language Models

Spooklight leverages multimodal Large Language Models (LLMs) capable of processing and generating both text and images. The OpenAI API serves as the backbone for these capabilities, utilizing models such as GPT-4 for text generation and DALL·E 3 for image creation.

3.2.2 Prompt Engineering

Prompt engineering is crucial to guide the LLMs effectively. Structured prompts are crafted using the Promptdown format [5], which helps in defining clear instructions and constraints for the AI models. These prompts are designed to:

- Encourage creativity and imagination.
- Maintain coherence and consistency in the narrative.
- Reflect the selected author's style.
- Avoid overused words and phrases.

An example of a prompt used for generating the first narrative is:

```
## System Message

<persona>

You are an expert storyteller, writing in the style of {author_style}.

</persona>

<task>

You will be provided with a summary of the story concept. You will also be provided with a description of the first image that will be used in the story. Your task is to write the opening narrative extending from (but not limited to) the described image.

</task>

<rules>

<rule priority="critical">The story must be coherent and
```

```
understandable to a typical
reader. Do not get lost in
descriptive detail that does
not advance the events of the
story.</rule>

<rule
priority="important">This is
the beginning of the story, so
write it as an opening with a
"hook" that will draw the
reader in.</rule>

...

</rules>
```

3.2.3 Key Modules and Functions

The Spooklight project is implemented in Python, and is organized into several key modules:

- **Initialization Module** (`initialize_story.py`): Handles the setup of the story, including enhancing the story concept and selecting the author style.
- **Image Processing Module** (`imageprocessing/`): Contains functions for describing images, generating visual styles, and re-encoding images.
- **Step Generation Module** (`stepgeneration/`): Manages the iterative process of generating image descriptions, images, and narratives for each step.
- **Completion Module** (`completion/`): Checks for story conclusion and generates the story title.
- **CLI Interface** (`cli.py`): Provides a command-line interface for users to interact with Spooklight.

3.2.4 Input Parameters and Configuration

Spooklight allows users to customize the generation process through various parameters:

`--starting-image-path`: Specifies a starting image for the story.

`--starting-image-description`: Provides a description to generate the first image if no starting image is given.

`--story-concept`: Defines the overarching concept of the story.

`--story-length`: Sets the number of steps (image-narrative pairs) in the story.

3.2.5 Default Behaviors and Error Handling

If no starting image or description is provided, Spooklight attempts to generate one based on the story concept.

If critical parameters are missing, the tool exits with an error message.

The story length can be unspecified, allowing the story to conclude naturally based on the AI's determination.

3.2.6 Execution Process

To run Spooklight, users execute the `cli.py` script with the desired parameters:

```
python cli.py --story-concept
"An unlikely rebellion of
oppressed people fighting
against a tyrannical council
of wizards" --story-length 10
```

The output includes a series of images and narratives saved in the `output` directory, culminating in a compiled PDF document `story.pdf` containing the complete story.

4. Results

The following is an example of a generated image and a portion of the associated narrative produced by Spooklight:

“In the dim embrace of the grand hall, their flickering candlelight seemed to carve shadows of purpose upon the celestial map sprawling upon the long table. An astral tapestry of the cosmos, it lay unfurled beneath the robed figures gathered like constellations in quiet counsel. Their eyes shone with the intensity of seers gazing into the heart of eternity, capturing moments spun from the heavens to unfurl across history’s scroll. Each hushed whisper and measured gesture in their arcane congress bespoke a journey through the stars—a gathering to weave the threads of the world anew with wisdom plucked from the ancients’ highest echelons.”

5. Challenges and Limitations

Despite the innovative approach and promising results of Spooklight, several challenges and limitations were encountered during its development and execution. These issues highlight the complexities inherent in leveraging AI for creative storytelling and underscore areas for potential improvement.

5.1 Repetition of Overused Words

One significant challenge was the language model's tendency to overuse certain words and phrases, even when explicitly instructed to avoid them in the prompts. Words such as "tapestry," "delve," and "kaleidoscope" frequently appeared throughout the generated narratives, sometimes occurring on nearly every page. This repetition not only detracted from the narrative quality but also revealed recognizable patterns often associated with AI-generated text, potentially diminishing the perceived originality and creativity of the stories.

5.1.1 Potential Solutions

Enhanced Prompt Engineering: Refining prompts to include stronger discouragement of specific overused words or incorporating dynamic lists of words to avoid.

Lexical Diversity Algorithms: Implementing post-processing steps that analyze the text for repetitive language and replace overused words with suitable synonyms.

Fine-Tuning Models: Training the language model on additional datasets that emphasize lexical variety and discourage repetitive phrasing.

5.2 Incoherence Between Image Flow and Narrative

Another challenge stemmed from the inherent differences between the image generator (DALL·E 3) and the text generator (GPT-4o). This "impedance mismatch" sometimes led to discrepancies where the images did not accurately represent characters or events described in the narrative. Consequently, subsequent narratives, relying on these images, could become inconsistent or veer off course, disrupting the coherence of the overall story.

5.2.1 Potential Solutions

Cross-Modal Consistency Checks: Developing mechanisms to assess and ensure alignment between generated images and the corresponding textual descriptions before proceeding to the next iteration.

Unified Multimodal Models: Exploring the use of models that can handle both image and text generation within a single framework to reduce discrepancies.

Human-in-the-Loop Interventions: Incorporating minimal human oversight to validate the coherence between images and narratives at each step.

5.3 Tendency Toward Melodrama

The language model exhibited a propensity to escalate narratives rapidly towards high-stakes, dramatic scenarios. Each section often concluded with life-or-death situations, using hyperbolic language such as "the very earth itself" or "the very essence of [whatever]." This melodramatic tone overshadowed subtler aspects of storytelling, such as character development and plot progression, resulting in narratives that lacked balance and depth.

5.3.1 Potential Solutions

Tone Calibration in Prompts: Adjusting prompts to encourage a more measured narrative tone, emphasizing gradual plot development and nuanced storytelling.

Stylistic Constraints: Implementing guidelines within the prompts that discourage the use of hyperbolic language and overemphasis on dramatic stakes.

Model Fine-Tuning: Training the model on literature that exemplifies balanced storytelling with varied pacing and tone.

5.4 Insufficient Character Development

A recurring issue was the insufficient development of characters within the narratives. The model tended to focus excessively on describing the exact content of the images, rather than using them as a starting point to explore characters' motivations, backgrounds, and growth. This image-centric approach often resulted in stories that lacked emotional depth and failed to engage readers on a personal level.

5.4.1 Potential Solutions

Character-Centric Prompts: Crafting prompts that specifically instruct the model to focus more deeply on character development and interpersonal dynamics.

Story Planning Mechanisms: Incorporating intermediate steps where the AI generates character profiles or backstories to inform subsequent narratives.

Diverse Training Data: Exposing the model to a wider range of literature that emphasizes strong character arcs and development.

5.5 Thematic Consistency Issues

The model showed a tendency to introduce recurring themes such as magical portals and supernatural elements, even when they were inappropriate for the given story concept. For instance, a story intended to be about cavemen searching for fire was transformed into a narrative involving a "galaxy bird" and ancestral souls, diverging significantly from the original concept.

5.5.1 Potential Solutions

Stricter Adherence to Story Concepts: Enhancing prompts to reinforce the importance of staying within the defined story parameters.

Theme Filtering: Implementing content filters that detect and restrict the introduction of irrelevant or undesired themes.

Model Fine-Tuning for Genre Specificity: Training or adjusting the model to better recognize and adhere to the intended genre and thematic elements of the story concept.

5.6 Underlying Causes and Overall Solutions

These challenges stem from inherent limitations in the AI models and their training data.

The AI models are trained on large datasets that may contain prevalent themes, styles, or vocabulary, influencing the generated content toward those patterns.

Additionally, while prompts guide the AI's output, their effectiveness is limited by the model's inherent tendencies and the complexity of natural language understanding.

6. Conclusion

6.1 Summary of Findings

Spooklight serves as an innovative exploration into the integration of multimodal AI in generative storytelling. By alternating between image generation and narrative creation, it produces stories that are both visually and textually rich. The project's methodology demonstrates how AI can introduce stochastic elements into the creative process, resulting in unique and evolving narratives.

Despite its successes, Spooklight's challenges highlight the limitations of current AI models in maintaining narrative coherence, avoiding repetitive language, and ensuring thematic diversity. These issues underscore the need for continued development in AI technologies and prompt engineering techniques to enhance the quality of AI-generated content.

6.2 Future Directions

To enhance Spooklight's performance and address the identified challenges, future efforts will focus on refining the methodologies and leveraging advancements in AI technology. Continuously refining the prompts based on observed outputs will allow for more precise guidance of the AI models, steering them away from undesired patterns like overused phrases or melodramatic tones. Implementing mechanisms for stronger contextual awareness, such as summarizing key story elements to feed back into the model, can improve narrative coherence and character development. Additionally, developing techniques to ensure better alignment between the image and text generators will help maintain consistency throughout the narrative.

Another promising direction is the implementation of hybrid approaches that combine AI generation with human editing or oversight. Human intervention can play a pivotal role in ensuring quality and adherence to the intended story concept by correcting inconsistencies, enriching character development, and maintaining thematic coherence. Exploring advancements in AI models with improved context handling, reduced biases, and enhanced coherence in multimodal content generation is also vital. By pursuing these strategies, we aim to refine Spooklight's methodology, overcome current limitations, and contribute valuable insights to the broader field of generative art and multimodal AI research.

7. References

[1] "Sudowrite - Best AI Writing Partner for Fiction." *Sudowrite*, https://www.sudowrite.com/. Accessed 31 Oct. 2024.

[2] "Tome | Generative AI storytelling has arrived. | Welcome AI." *Welcome AI*, https://welcome.ai/solution/tome-generative-ai. Accessed 31 Oct. 2024.

[3] "Depthtale - Play & Create Immersive AI-Powered Visual Novels." *Depthtale*, https://www.depthtale.com/en. Accessed 31 Oct. 2024.

[4] "Storynest.ai - AI Powered Interactive Stories and AI Characters." *Storynest.ai*, https://storynest.ai. Accessed 31 Oct. 2024.

[5] Franklin, B. T. *btfranklin/promptdown: A Python Package That Enables the Creation and Parsing of Structured Prompts for Language Models in Markdown Format.* GitHub, https://github.com/btfranklin/promptdown. Accessed 31 Oct. 2024.

An Attempt at Generating Beethoven's 33rd Piano Sonata with Music Transformer

Rafael Tamim
LAMSADE, Université Paris Dauphine - PSL, Paris, France.

Tristan Cazenave
LAMSADE, Université Paris Dauphine - PSL, Paris, France.

Abstract

Generating a new piano sonata in the style of Ludwig van Beethoven poses a significant challenge, particularly when attempting to emulate the distinct characteristics of his late compositional period. Beethoven's music, which evolved significantly over his lifetime, is renowned for its complexity, emotional depth, and innovative structures, especially in his later works. The task of creating a 33rd sonata that Beethoven might have composed had he lived longer involves not only capturing his unique style but also understanding the broader historical and stylistic contexts of classical music.

This research delves into the intricacies of Beethoven's musical evolution and explores how artificial intelligence can be employed to generate a composition that reflects the distinct qualities of his late period. The project addresses the challenge of limited data by carefully considering the influence of different composers and periods on Beethoven's style, aiming to produce a work that resonates with the profound expressiveness and complexity of his final compositions. By pushing the boundaries of AI in music, this study contributes to the ongoing dialogue between technology and creativity, offering new insights into the possibilities of machine-generated art in the classical music tradition. It also raises important questions about the role of human interpretation and intuition in the creative process, especially when replicating a style as nuanced and historically significant as Beethoven's. This intersection of AI and classical music serves as a thought-provoking

exploration of how technology can both mimic and enhance the creative practices of the past, suggesting new directions for the future of music composition.

A Question of Context

In classical music, as we define it today, there are several distinct periods. The most well-known, in chronological order, are: the Baroque period (1600-1750), the Classical period (1750-1820), the Romantic period (1820-1900), the Modern period (1900-1975), and finally the Contemporary period (1975 to the present). These periods differ, among other things, in their style, techniques, and forms of composition.

Within these periods, each composer develops their own musical universe. Beethoven's style is therefore unique, and it is essential to choose a model capable of learning to generate music in this style. However, we knew that training a deep learning model would require a large amount of data, and limiting ourselves to Beethoven's 32 sonatas would not be sufficient. Therefore, we chose to start with a more extensive and diverse dataset containing pieces from various periods, while also incorporating the sonatas later in the process.

Moreover, Beethoven’s style evolved throughout his lifetime. His piano sonatas illustrate this evolution well. Musicologists have classified them into three parts, corresponding to the three periods of Beethoven’s life. Therefore, to generate the 33rd sonata, not all the sonatas should be given the same importance. The sonatas from his late period should have a greater impact on the generation of the 33rd sonata than the earlier ones. We therefore considered implementing a weighting system to represent the importance of each sonata in the generation process. These weights will subsequently be used in the loss function calculation that the model aims to minimize. We also realized that this system could be applied to all pieces in the dataset. The idea was to assign a low weight to composers stylistically distant from Beethoven (e.g., Debussy, Rachmaninoff) and a high weight to his contemporaries (e.g., Mozart, Haydn).

A model like Museformer [1] already uses a weighting system, but this concerns how the model learns rather than what it learns. Museformer is based on the principle that some musical passages are more important than others, containing information about the structure of the work, such as the main theme. Ultimately, after an unsuccessful attempt to make this model work, we decided to switch to another model based on the same architecture, namely Music Transformer [2].

When AI Judges Music

To objectively evaluate the model, we chose to use perplexity as a metric, in addition to the calculation of accuracy and the loss function implemented in the base code.

Here is the definition of perplexity from the article *Decoding Perplexity and its Significance in LLMs* [7]:

"In brief, perplexity measures the model's confidence in its predictions. The concept of perplexity evaluates how confused the model is when predicting the next word in a sequence. Lower perplexity indicates that the model is more certain of its predictions. In contrast, higher perplexity suggests that the model is more uncertain. Perplexity is a crucial metric for assessing the performance of language models in tasks such as machine translation, speech recognition, and text generation.

The perplexity of a language model can be calculated using the average negative log-likelihood. The formula for perplexity is given by:

$$\text{Perplexity} = \exp(\text{Average NLL})$$

where the average negative log-likelihood (Average NLL) is defined as:

$$\text{Average NLL} = -\frac{1}{N}\sum_{i=1}^{N}\log p(w_i \mid w_{1:i-1})$$

Here, N is the number of words in the sequence, and $p(w_i \mid w_{1:i-1})$ is the predicted probability of the word w_i given the previous words $w_{1:i-1}$ .The exponential function is used to convert the average negative log-likelihood into perplexity, thus providing a measure of the model's confusion regarding the word sequence."

In the previous definition, we talk about predicting words. However, it can be adapted to the prediction of notes or other elements of musical vocabulary since our dataset consists of MIDI (Musical Instrument Digital Interface) files.

A MIDI sequence is a digital file that contains control information for music, such as the notes played, velocity, tempo changes, and channel control commands.

```
2,  96, Note_on,  0, 60, 90
2, 192, Note_off, 0, 60,  0
2, 192, Note_on,  0, 62, 90
2, 288, Note_off, 0, 62,  0
2, 288, Note_on,  0, 64, 90
2, 384, Note_off, 0, 64,  0
```

Figure 1 : excerpt from a MIDI file (turned into readable ascii)

Figure 2 : Score corresponding to the MIDI excerpt

In figures 1 and 2, extracted from [3], we can visualize how a MIDI file works. In figure 1, the first line "2, 96, Note_on, 0, 60, 90" means "at 96 ticks (the moment the event occurs) on track 2, on channel 1, a middle C (C4) note is played with a velocity of 90." Below, we can clearly see the middle C on the sheet music.

This information is then tokenized during the preprocessing phase, and the model aims to predict the next token based on the previous tokens.

Training is a Sport

As mentioned earlier, training a model like Music Transformer requires a significant amount of data. Training the model directly on Beethoven's 32 sonatas would not have guaranteed a good result. Therefore, we split the training into two phases: a pre-training phase on a general classical music dataset of around 700 MIDI files, followed by a fine-tuning phase on Beethoven's 32 sonatas. During the first phase, the model adjusts its parameters to minimize a loss function, calculated based on the difference between the model's predictions and the actual data. In the second phase, the model builds on the parameters adjusted during the previous phase and fine-tunes them to capture the particular nuances of Beethoven's style, thereby improving the quality of the generated compositions.

We implemented a weighting system in the pre-training dataset by assigning each composer a weight relative to their similarity to Beethoven. The choice of weights is subjective, and we based them on our own musical knowledge. The following figures (3-6) show the different metric behaviors between training without weights and training with weights.

We can see that the accuracy decreased and the perplexity increased with the application of the weights. This did not positively influence the results, contrary to our expectations. As we will see later, the issue does not stem from the strategy implemented but rather from the dataset.

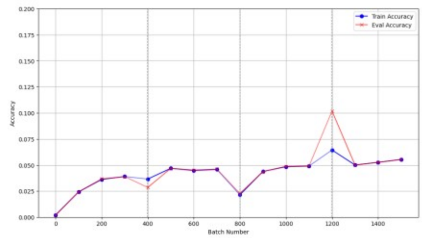


Figure 3 : Accuracy without weights

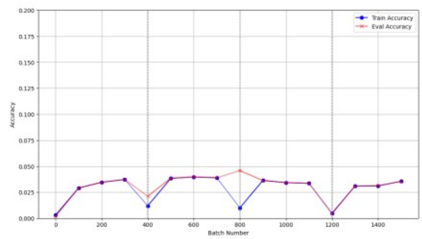


Figure 4 : Accuracy with weights

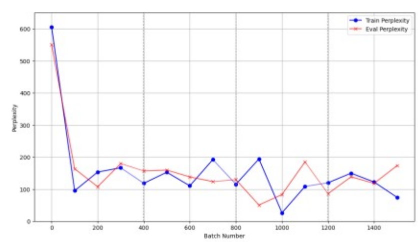


Figure 5 : Perplexity without weights

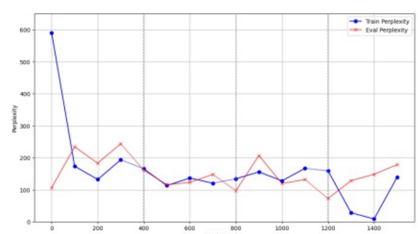


Figure 6 : Perplexity with weights

The Musical Parrot Syndrome

While observing the behavior of the loss function, we noticed signs of overfitting, indicating that the model does not generalize well to data it hasn't seen during training. In other words, it memorizes the pieces in the dataset and fails to generate a new piece effectively. To address this, we first considered adjusting the dropout rate. Depending on the severity of the overfitting, the dropout rate can be increased to mitigate the issue. Initially set at 20% in the base code, we increased it to 40%, which reduced the overfitting but did not eliminate it completely.

In truth, 700 pieces are not enough for a model as complex as Transformers. For instance, Museformer was trained on a dataset of nearly 30,000 pieces. Additionally, our dataset was not very homogeneous, with a majority of works by Bach and Chopin. We, therefore, sought a more extensive dataset, namely the Maestro dataset [4] developed by Google Magenta. With this dataset, the overfitting issue was resolved, even when keeping the dropout rate at its initial value. Moreover, the accuracy doubled, rising from 3.5% to 7%, and the perplexity decreased significantly from 178 to 61. However, a perplexity value of 61 is still too high; it means that during generation, the model has 61 possible tokens to choose from. To further expand the Maestro dataset, we transposed the pieces into all possible keys, like in [9], multiplying the number of pieces by 24 and resulting in a dataset of over 30,000 MIDI files. The following graphs (figures 7-8) show that the results improved dramatically with the transposition: accuracy reached 23%, and perplexity dropped to 14.

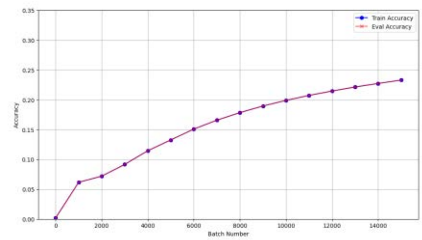


Figure 7 : Accuracy after transposition

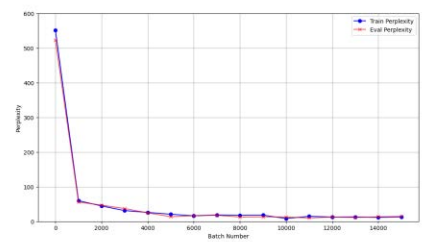


Figure 8 : Perplexity after transposition

The Cool New Tricks of AI

Let's start with the definition of a learning rate schedule from the site Towards Data Science [8]:

"A learning rate schedule is a crucial technique in training machine learning models because it allows for the adjustment of the learning rate throughout the training process. The primary utility of a learning rate schedule lies in the following aspects:

Improved Convergence: A well-adjusted learning rate can help converge more quickly to a minimum of the loss function. By gradually reducing the learning rate, the model's weights can be refined with greater precision as the minimum is approached.

Prevention of Oscillations: Using a schedule prevents a too-high learning rate from causing oscillations around the local minimum. A rate that is too high may result in excessive jumps in the loss function values, hindering convergence.

Avoidance of Local Minima: Strategies such as gradually reducing the learning rate allow the model to overcome local minima by dynamically adjusting the rate. This helps in more effectively exploring the solution space.

Resource Optimization: By regulating the learning rate, computational resources can be optimized, ensuring that learning is neither too fast (which could lead to premature convergence) nor too slow (which might require more computation time).

In summary, the learning rate schedule plays a key role in improving the performance and efficiency of model training by allowing better management of the learning speed throughout the training process."

Let's now return to our code. The model's basic strategy combined several techniques: an inverse square root schedule to dynamically adjust the learning rate, a linear warmup to gradually increase the learning rate at the beginning of training, and a minimum schedule to ensure a minimum threshold for the learning rate. See figure 9 for the formula of the learning rate.

$$\text{learning rate} = \frac{1}{\sqrt{d_{\text{model}}}} \cdot \min\left(\frac{1}{\sqrt{\text{step}}}, \frac{\text{step}}{(\text{warmup steps})^{1.5}}\right)$$

where:

- d_{model} is the model dimension.
- step is the current number of steps.
- warmup steps is the number of steps for the linear warmup.

Figure 9 : Initial learning rate formula

This technique is quite commonly used, which led us to try another, more recent strategy: Cosine Annealing. This method first appeared in the paper "SGDR: Stochastic Gradient Descent with Warm Restarts," written by Ilya Loshchilov and Frank Hutter [5]. This paper, published in 2017, introduces a method for dynamically adjusting the learning rate by using a cosine function to periodically reduce the learning rate. See figure 10 for the formula.

The learning rate η_t at step t is given by:

$$\eta_t = \eta_{\min} + \frac{1}{2}(\eta_{\max} - \eta_{\min})\left(1 + \cos\left(\frac{T_{\text{cur}}}{T_{\max}}\pi\right)\right)$$

where:

- η_t is the learning rate at step t,
- $\eta_{\min}$ is the minimum learning rate,
- $\eta_{\max}$ is the maximum learning rate,
- T_{cur} is the number of steps since the last restart,
- $T_{\max}$ is the total number of steps before a restart.

Figure 10 : Learning rate in the Cosine Annealing strategy

During the testing phase, we compared the results of training with the baseline strategy to those of training with the Cosine Annealing strategy. As shown in the following figures (11-12), Cosine Annealing improves the metric values, particularly the perplexity. Indeed, we achieve a perplexity of 61 by the end of the fifth epoch compared to only 140 with the baseline strategy. Precision, on the other hand, increases by only 1.5 points, which is not a significant improvement. Here is the perplexity curve for both strategies:

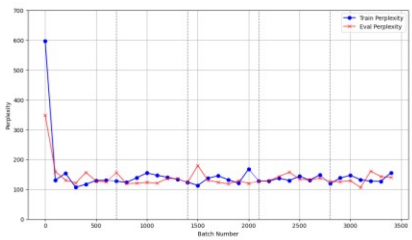


Figure 11 : Perplexity without Cosine Annealing strategy

Figure 12 : Perplexity with Cosine Annealing strategy

With the Cosine Annealing strategy and the transposition of the Maestro dataset into all possible keys, we achieved very good results (for reference, 23% precision and a perplexity of 14). However, so far we have only discussed the pre-training phase. Let's see how this performs specifically on Beethoven's sonatas:

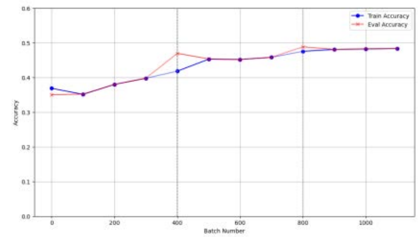


Figure 13 : Accuracy during fine-tune

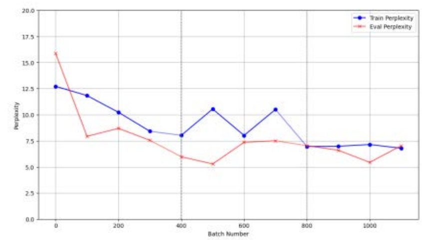


Figure 14 : Perplexity during fine-tune

The precision converges to 48% after 3 epochs, and the perplexity drops to 7! The model now has only 7 possible choices at each generation step.

When Numbers Lie

Finally, our last focus was to understand and improve the generation phase, i.e., the one that produces the final result. To generate a piece of music, the model takes as input a sequence containing the first *x* tokens from a MIDI file and uses what it learned during training to generate the following tokens, in other words, the continuation of the sequence.

We tested with several input sequences, particularly the beginning of Beethoven's 32nd sonata, and the result was rather surprising. There are many musical elements that do not contribute much to the structure of the piece, such as trills or notes repeated for a long time, which results in a somewhat incoherent output. Additionally, there is a noticeable lack of creativity as the same elements are frequently repeated. However, one might argue that the model has learned to use recurring motifs, treating them like a recurring theme. The beginning of the 32nd sonata contains few structural elements (very few harmonies or "melodies") and includes a trill, which helps explain why the result lacks creativity. Testing with other melodically more varied sonata openings shows better generation performance.

In generative models, different generation policies can be used, the main ones being argmax and softmax. Using argmax, the model systematically chooses the token with the highest probability at each generation step. With softmax, the model uses a probability distribution, allowing for more diversity in the choice of the token to be generated. In our code, the softmax policy was chosen. Therefore, the lack of creativity cannot really be justified by the policy used.

We think that the complexity of classical music explains why our model did not perform well. In classical music, there is a hierarchy of musical elements. Ornaments like trills, mordents, and appogiaturas enrich and express the music but remain secondary to the main melody. Similarly, passing notes, escape notes, appogiaturas, and suspensions add richness and harmonic tension but are generally less important than structural notes. However, our model treats all elements equally. In our opinion, learning should be decomposed into several tasks, such as one for the harmony of the piece, another for melody, rhythm, etc. Once these steps are completed, a basic structure of the piece can be filled with less important elements as mentioned previously. Such models already exist in the literature; for example, WuYun [6] proposes a two-step hierarchical architecture for melody generation guided by a skeleton. This paper focuses only on melody. In our case, a method to separate different musical layers (melody, harmony, etc.) would be needed, which is far from straightforward.

Conclusions

Artificial intelligence offers a wide range of applications, and in this context, we chose a field that is close to our heart: music. After considering several options, we ultimately decided to generate Beethoven's 33rd piano sonata. Our first task was to devise a method to generate a work in Beethoven’s style, consistent with his final creative period. This led to the idea of a weighting system in the dataset to give more importance to works from Beethoven’s later period. After selecting a suitable deep learning model for music generation, the next question was how to evaluate the results. Although subjective criteria are important, we needed objective criteria to assess the quality of the compositions generated by the AI.

The model training was carried out in two phases. This approach allowed the model to learn from a broad dataset before specializing in Beethoven. However, we encountered an overfitting issue, where the model memorized the specifics of the pieces, a phenomenon we tried to mitigate through various machine learning techniques, including our weighting system. As AI evolves rapidly, we tested a recent learning strategy, which significantly improved the results of both the pre-training and final phases compared to the baseline strategy. Despite good objective results, the generated sequences were not always pleasant to listen to. We then examined the model's generation policy, which provided some answers. Ultimately, this study helped us better understand the limitations of our model and consider future improvements.

From a technical perspective, the work accomplished reached a good level of completion while leaving room for improvements. The decision to use deep learning models, particularly Transformers, proved effective for handling the complexity of musical sequences. The two-phase approach provided a solid learning foundation, facilitating specialization.

However, the initial occurrence of overfitting highlighted a potential weakness in the model. While regularization techniques such as dropout and the weighting system helped mitigate this issue, it was the use of a larger and more diverse dataset that ultimately resolved the problem. This improvement allowed the model to produce more innovative and stylistically coherent works, suggesting that the quality and diversity of training data are crucial for enhancing the model's generative capabilities.

Integrating a weighting system to prioritize Beethoven's later sonatas was a creative and promising step. It influenced the musical generation by respecting the chronological and stylistic specifics of the composer. This concept could be expanded to other composers and styles, enriching research perspectives.

On the other hand, the project revealed some limitations. While effective, Transformers might not be the best tools for all musical generation tasks. The fine-tuning methodology on a limited number of sonatas may also not fully capture Beethoven's stylistic complexity. A more nuanced approach, combining machine learning techniques with elements of music theory, could offer more refined results.

References

[1] B. Yu, P. Lu, R. Wang, W. Hu, X. Tan, W. Ye, S. Zhang, T. Qin, and T. Liu. Museformer : Transformer with fine-and coarse-grained attention for music generation. In Advances in Neural Information Processing Systems (NeurIPS). NeurIPS, 2022.

[2] Cheng-Zhi Anna Huang, Ashish Vaswani, Jakob Uszkoreit, Ian Simon, Curtis Hawthorne, Noam Shazeer, Andrew M. Dai, Matthew D. Hoffman, Monica Dinculescu, and Douglas Eck.

Music transformer: Generating music with long-term structure. In Proceedings of the 7th International Conference on Learning Representations (ICLR). ICLR, 2018.

[3] Jean-Pierre Briot, Gaetan Hadjeres, and Francois-David Pachet. Deep Learning Techniques for Music Generation. Computational Synthesis and Creative Systems. Springer, 2019. ISBN 978-3-319-70162-2. Hardcover.

[4] Curtis Hawthorne, Andriy Stasyuk, Adam Roberts, Ian Simon, Cheng-Zhi Anna Huang, Sander Dieleman, Erich Elsen, Jesse Engel, and Douglas Eck. Enabling factorized piano music modeling and generation with the MAESTRO dataset. In International Conference on Learning Representations,2019. https://openreview.net/forum?id=r1lYRjC9F7.

[5] Ilya Loshchilov and Frank Hutter. Sgdr: Stochastic gradient descent with warm restarts. arXiv preprint arXiv:1608.03983, 2017.

[6] K. Zhang et al. Wuyun: Exploring hierarchical skeleton-guided melody generation using knowledge-enhanced deep learning. arXiv preprint arXiv:2301.04488, 2023.

[7]https://blog.uptrain.ai/decoding-perplexity-and-its-significance-in-llms/

[8]

https://towardsdatascience.com/a-visual-guide-to-learning-rate-schedulers-in-pytorch-24bbb262c863

[9] Huang, A., Wu, R. (2016). *A Transformer Based Pitch Sequence Autoencoder with MIDI Augmentation*. arXiv preprint arXiv:1611.03477.

Generative Authorship

Ian O'Loughlin, PhD.
Department of Philosophy, Pacific University, Forest Grove, USA
www.ianoloughlin.com
e-mail: ian.m.oloughlin@gmail.edu

Abstract

Artist attribution in cases of generative art is complicated and controversial. For one very particular class of cases, one hitherto overlooked possibility has merit. If a generative novel has been produced in the style of a human author, and the resultant work is indiscernible from other works with the same author even by competent critics, then there is reason to ascribe authorship of the generative piece to the human author whose work it was styled after. Although this has some counterintuitive and surprising consequences, these do not constitute insurmountable obstacles. This view also has a number of interesting implications, and offers some ways forward with regard to other arguments about artist attribution in generative pieces.

1. Introduction

Questions regarding whom to credit as the *artist* in cases of generative art remain complicated and fraught. Scholars are divided not only on how to answer this question, but also on how to approach it [1, 2, 3]. In the following I will examine one very special case of generative art—fiction in the style of a human author—and, by way of introducing a Turing Test-style condition, make a relatively strong, albeit conditional, claim about authorship attribution. My thesis, strictly speaking, only regards generative works of fiction in the style of a particular human author, but I will address possible and broader connections to other areas of generative art in due course. Additionally, of course, an argument for there being at least some cases of generative art for which artist attribution claims can be rightly made, it is hoped, already constitutes one small bit of progress in these larger pursuits.

To these ends, I will first briefly introduce some of the ways authorship attributions have been discussed and characterized by scholars interested in prose and fiction generated by artificial intelligence, and also some general considerations about authorship of prose works. After laying out some of the possible moves here, I will introduce and motivate a test in the style of the Turing Test, but that targets critics' perceptions of authorship. I will more or less directly lay out an argument that authorship can be rightly attributed in cases where a generative work passes this test. I will then consider a series of objections, worries, and implications.

2. The life of the author

Generative methods already produce significant prose works. Undoubtedly these are still in their infancy, and very likely in only a few years' time the prose works so produced will be far greater in both number and substance. Traditionally, the vast majority of prose works are straightforwardly attributable to an author or authors, and this is especially the case for long, complicated, self-contained and internally coherent works such as novels. While there are a few interesting cases of questionable authorship of novels (lost authors, collaborative works, complicated fan fiction, forgeries, etc.), almost all novels—more so than poetry or nonfiction—are clearly attributable to a definite and small collection of authors, usually to just one human being. This relative paucity of problematic cases is perhaps more the case for novels than it is for other art forms, and may make novels a good starting point for questions of artificial intelligence authorship.

This being said, questions of human authorship are not entirely simple. Roland Barthes famously argued that the role of the author *qua* author had been previously overemphasized in culture and literary criticism, and successfully reshaped the way that authorship questions are bound up with criticism of works [4]. Even more relevant to questions about the *nature* of authorship, Sophus Helle has argued that prose authorship exists in the "middle ranges" of agency, splitting the difference between complete originality and mere mediation, in part because of the heavy debt any author owes to previous works and elements of human culture [5].

In the case of generative authorship the apparent options are laid out nicely in a recent article by Jean-Marc Deltorn and Franck Macrez [6]. Deltorn and Macrez enumerate three broad options: to assign authorship to no one, to assign it to the programmers, developers, or deployers of the generative system, or to assign it to the generative system itself. On the face of it, this is a sensible and satisfying array of possibilities, and it tracks intuitions and discussions of artist attributions in cases other than authorship as well. It is not hard to imagine or find scholars in various disciplines, in different particular cases, taking up each of these causes in turn. For example, Jane Ginsburg and Luke Budiarjo dismiss the possibility of machine attribution, and argue that authorship questions only remain between the various humans that were involved programming or prompting the system in any given instance of generative authorship, with the distinct possibility that some works of art are authorless [7]. In other areas of art, some scholars argue for the possibility of machine authorship, but only under certain conditions [1].

While I will not in the present paper take sides on whether artificial intelligence systems can rightly be afforded author status in some cases, I will argue that in the case of a particular class of generative fiction, one possibility has been overlooked. In the following section, I will propose a test, and make the claim that any work that passes this test is neither authorless, nor the work of the programmers, prompters, or machine.

3. A criticism test and the implications of indiscernibility

The particular class of generative novels with which we are presently concerned is those novels written "in the style" of an existing human author. The tools for generative novels are already extant, and the tools for generating prose in a style particular to a single human author are also so, even though as of 2024 this has not yet culminated in a library of full and convincing novels in the style of dead novelists. Nevertheless, it is hard to imagine how these will not be readily available in the near future. For the sake of authorship questions, I propose a test to which any such work can be put.

This test is in the spirit of the Turing Test, but instead of competent judges of intelligent conversation, we will turn to competent judges of authorship. For the sake of this thought experiment, we must imagine a critic who is eminently familiar with a given author's style, and can confidently and rightly distinguish a novel by that author from other novels. Also for the sake of the thought experiment, we must imagine a critic who is not already familiar with this given author's full *oeuvre*. Admittedly, this requirement may strain credibility in many cases in the real world—it is perhaps plausible for Balzac or Murdoch, but not for Kafka or Proust. At present let us imagine a world where such critics are to be found as needed, by careful sequestration or localized amnesia if need be, and we will return to some of the real-world considerations eventually.

In such a world, any given generative fiction work in the style of a particular human author will lend itself to a simple test: the competent literary critic will either be able to discern the generative novel from the traditionally-produced novels of the author, or not. For example, let us imagine a generative novel, *Cyril Ambrose*, written in the style of George Eliot. We might give this novel, along with one of hers written in the traditional style by her in her lifetime, to discerning critics who were unfamiliar with either of these particular works (again, pardoning what may be the real-world implausibility of this for the sake of the present argument), and ask the critics to attempt to identify which is the generatively constructed novel. If the critics do no better than chance at this task, then *Cyril Ambrose* has passed the test.

Who or what is the author of such a work? I want to suggest that there is an author, and this author is none of the usual suspects (programmers, prompters, system). The author, if *Cyril Ambrose* has passed the above criticism test, is George Eliot. Characterizations of the nature of authorship and artistry lend support here. What an author is, according to Helle, is someone who mediates, who weaves together elements of culture into a new mode to be transmitted forward [5]. What an artist is, according to Alva Noë, is someone who provides a new way for us to understand our experience, offering a method for organizing the world as we find it [8]. Indeed, in other areas of art, we find historical cases of attribution that can serve as models: art historians sometimes attribute works to "Rafael" regardless of which of his assistants' hands may have been those in physical contact with the work [9]. It is treated as clear in these cases that Rafael is the person *responsible* for the work. In our own case, Eliot has done just this, offered a particular and new perspective, showcased in her novels. The generative system, and humans deploying it, have acted as her assistants, albeit separated from her in time. As is evidenced by even the most discerning readers being unable to tell the difference, this authorial contribution particular to Eliot is as present in the new generative work *Cyril Ambrose* as it is in any of her other works.

This somewhat expansive view of attribution is not entirely without precedent even in the arena of literature. Ascribing authorship of a text to Enheduanna or Homer even when the scribes who committed their texts to written form lived long after the authors' deaths has been relatively standard practice for centuries. Although it may at first seem like this represents an antiquated view, especially in light of Barthes' well-known *de*-emphasis of the author, since attributing generative posthumous works to an author might seem to emphasize or expand the role of the author, the reasoning in play can be interpreted as in keeping with Barthes. The text speaks for itself here—any details of the author's human life that are not already discernibly present in the text are not relevant. If the author of a collection of novels is, from a text-centered perspective, merely the zero-dimensional origin point, relationally defined as the organizing principle already immanent in the continuities among the perspectives offered by the texts, then this author "belongs" to a text only inasmuch as the author *comes through* in the text. Our criticism test is aimed at determining exactly this. These arguments, however, will be best made clear by considering a series of plausible objections.

4. Objections

In writing her novels, Eliot drew on a wealth of experience and had an array of intentions and goals directing these works' creation. Aren't these a necessary component of authorship?

In considering attributing artist status to generative systems, Adam Linson makes a compelling case for the necessity of experiential connectedness to human experience and culture, arguing that only if machines have broad engagement with the world, enabling the development of something like conscience, can they rightly be considered the artist [1]. Although I am not here arguing in favor of *machine* authorship, I *am* arguing that our hypothetical *Cyril Ambrose* is an authored work, and Linson's caveats are relevant, since one natural worry is that the apparent context-free generation of the text lacks just these. However, one question is whether this broad experience is manifest in the texts. If it is not, then, in keeping with Barthes, it is nor our problem. If it is, however, then a discerning reader would be able to detect its absence (otherwise what could we *mean* by 'manifest in the text'?). So if *Cyril Ambrose* has passed our criticism test, then any necessary authorial antecedents have been satisfied.

Another way to characterize the acceptance of these (or similar) necessary conditions for authorship is to point out that the objection begs the question against the thesis: to say that Eliot's intentions and experiences are absent in the creation if *Cyril Ambrose* is just to already deny authorship. On the present view, her experiences and intentions are immanent in *Cyril Ambrose* in just the same ways that they are immanent in Middlemarch. The method of execution—whether penned on paper, or generated by clever artificial intelligence systems two centuries later—is a feature to which authorship questions are indifferent.

All of that being said, it is also worth noting that one could offer up related reasons for predicting that it will be very difficult for generative pieces to pass the criticism test above. This I readily accept. It may be quite difficult. The point in contention is whether it would be correct to attribute the work to Eliot *if* it were to pass.

On your view, someone can write a novel long after they have died. Whatever we mean by 'authoring a novel', doesn't it exclude this spooky posthumous action?

Admittedly, this implication applies pressure to our standard and preconceived notions of artistic creation, not to mention action more broadly. Ultimately, however, the point must be conceded. If *Cyril Ambrose*, for example, is written in 2030, and if the author is Eliot, then Eliot will have in fact authored a novel long after she died. The counterintuitive nature of this may be mitigated by considering nearby thought experiments, however. We can imagine Rafael in his workshop, giving explicit instructions to one of his assistants, and then leaving on a trip where he unexpectedly dies without the assistant knowing. The assistant starts and finishes the work, just as many are thought to have done while Rafael was merely busy or absent, and—*voila!*—Rafael has painted something after he has died.

Notably, the question here is not whether Rafael can rightly be said to be the artist in this example—that is a complicated question in its own right—the relevant question is whether we would have *less* reason to call Rafael the artist if he had died on that trip, than we would if he had merely gone and spent some time in Florence, then returned after the work was complete. It seems absurd, in this case, to say that the artist attribution somehow depends on whether Rafael was still alive in Florence or not, even though that had no casual interaction with the creation of the work. The mere fact of death seems to not be as much on an obstacle to attribution as it may at first seem.

Your examples use dead authors, but by the logic of the argument, would the same claim not be made for a generative novel written in the style of a living author? In which case, are you suggesting that someone could write a novel without knowing it, even if they encountered and disavowed it as theirs?

I concede that dead authors make for clearer exemplars, and probably for more rhetorically convincing instances (perhaps because we are in the habit of not quite granting full personhood to the dead). However, this objection must be upheld; if the above arguments are accepted, then the textual indiscernibility speaks for itself, regardless of the state of the author's human body. It is true that this means that a living person could author a novel without knowing it, and it is true that this seems to us a strange state of affairs. It is worth noting, however, that there are nearby cases that are less counterintuitive: we certainly admit that living humans can be responsible for ideas, and deserve credit, even if these ideas were instantiated without their knowledge (more or less all of copyright law is predicated on just this). To say that some living person can author a novel without knowing it is not as dramatic a revision to our concepts of origination as it may be thought.

To the point of disavowal, it is helpful to remember that this does happen. An author may disavow one of their works. This does not make the work not authored by them, it makes it a work authored by them that was then disavowed. These are interesting cases (both the traditional and generative versions), but I think not particularly damaging to the argument at hand.

The criticism test you propose will probably never actually be able to be enacted. Why should we be interested?

It is a contingent feature of our world that authors happen to usually have a relatively small number of novels, and that anyone who has undergone the requisite training to be able to confidently and readily tell their work from any other is almost certainly familiar with the entire *oeuvre*. In the end, this must be accepted, but only as contingent. The main thesis of this paper is that *if* a generative novel were the sort of thing that *would* pass this test, then the author is the human author whose works it was styled after. That is, any work that is indiscernible from an authors' other works, even by the most astute critics, deserves to be credited to that author. The fact that it would be quite difficult to uncontroversially determine which cases this applied to is neither here nor there with regard to the conceptual claim.

The criticism test you mention seems to depend on some kind of "authorial essence". If a human author's works were sufficiently different from one another so as to keep critics guessing, would that just mean any and every generative novel was theirs?

Technically, as stated, the arguments above are vulnerable to this objection. Unlike the above objection, which gets at a contingent feature of the way the test happens to be constrained by the real world, this objection gets at a conceptual feature of the idea of the test itself—the critics in question must be able to tell this authors' novels from other work. As it happens, I suspect that most novelists' work would lend itself to the needed continuity here, but for the sake of completeness we can be clear about the conditions of the test: the critics are people who must be able to tell the author's work from other work. If a particular author is not amenable to this in any way then they may, at most, successfully immunize themselves against authoring generative novels.

This is absurd. If I ask 'how many novels did Woolf write', the answer simply cannot be 'eight, for now, but we'll see what happens next year'. You must be talking about something other than real authorship.

One of the reasons that novels, in particular, make a good target for this investigation is that they constitute a form of art that is entirely comprised of text—which by its nature transcends its physical implementation—but also this form of art is oriented toward the creation of a coherent whole, an invented world that nevertheless takes from and gives to the world we all collectively inhabit, in a way that is recognizably the author's work, conferring a particularized and human perspective on the reader. If such a coherent whole has been created, and the most competent judges deem it to offer the same perspective as a person's other works, then that author deserves credit for inventing the world in question. Moreover, if no one can tell the difference between that generatively-constructed whole and the traditionally-written novels, then we are forced to ask what it is we want from authorship. It seems clear that *Cyrus Ambrose* could not have existed without George Eliot. It also seems clear that, by virtue of it passing our criticism test, it has successfully created a coherent and invented world that confers Eliot's authorial perspective on the reader. Is this not authorship? If a ninth Woolf novel is released next year, so much the better.

In addition to objections, it may also serve our clarificatory aims to canvas a series of questions aimed at the implications of such a view.

5. Implications

You focus on novels. Could this be expanded to other art forms? Could we *make analogous tests, and ascribe artist status to other generative art instances that are done in the style of a particular human?*

These are worthy, if complex, avenues of investigation. There are a number of reasons why novels seem like a good starting point. Any art form that involves performance or too deeply involves its physical implementation invites another layer of complexity. Shorter form prose works are perhaps less plausible in the details of a hypothetical criticism test. While some visual art is perhaps not dependent on physical implementation (e.g. photography), these tend to simply involve techniques that are—as of right now—more difficult to imagine being executed indiscernibly by generative and artificial intelligence methods. In the end, these are all compelling as directions for future investigation, but there is some reason to begin with the simplest and least problematic cases. If we can show that there is a single class of art that rightly credits artists posthumously with its creation, we have made progress. This is so even if serious and widespread questions remain.

We sometimes speak metaphorically about the way that people who have died live on through their ideas. Does accepting this view of authorship imply that a human life, and particularly human agency, is more, and more literally, extended across time than is conventionally accepted?

This question involves a complex nexus of concepts, some of which cannot be disentangled here, but the short and simple answer is yes. The authorship in question is "real" authorship, which means that George Eliot will very likely continue producing novels over the next centuries. A living author who had this understanding of authorship might well, today, see their work as a beginning, rather than a completion; they are writing novels that will (or may) generate a coherent and fecund perspective that leads to a much larger supply of novels in future centuries. Although to some extent this may seem counterintuitive, in keeping with some of the responses to the above objections, it is worth remembering that we have long accepted that someone's ideas may take new forms, that their projects may take on new lives, or that their actions may continue to develop, long after the deaths of their bodies.

If it is true that at least some sufficiently indiscernible works of art are rightly attributable to the human whose corpus was used in the training of the system that created the work, does this have consequences for the legal ownership and attributability for generative art?

Although this present paper is concerned foremost with the authorship concept itself, this view almost certainly has implications for the legality of attribution, insasmuch as intellectual property law attempts to track actual authorship as at least one component of deciding under which conditions credit is due.

One relatively straightforward implication of the above arguments seems to be that any system that is trained and prompted *too* well, on just one artist's work, such that the results would fool a competent judge, may mean that the legal rights to the resultant work would be accidentally relegated to the original artist. While this was not the aim of the above argumentation about authorship, it seems an acceptable consequence. I make no claims, however, that this refined understanding of authorship offers very broad or powerful solutions to the complex collection of difficulties and questions artificial intelligence has brought to legal questions of intellectual property and origination.

6. Conclusions

Understanding how to rightly make artist attributions in cases of generative art is no simple task, and there are numerous and distinct features of this discussion each of which is convoluted and fraught in its own right. Although authorship questions in these cases have traditionally been limited to choosing from among the system as author, the programmers as authors, the prompter as author, or no one at all as author, this overlooks an intriguing possibility in cases where a generative piece is produced in the style of a particular human artist: that original artist as author. Whatever other conditions may suffice for authorship being credited thus, in a case where a generative work is indistinguishable from other works produced by the artist in question, this possibility gains real traction. At least in the case of novels, there is reason to think that the appropriate way to attribute authorship in these cases is to simply attribute authorship to the human whose work the system was tuned to and trained on. That is, if a generative novel is written in the style of a human author, and it is indistinguishable even by competent judges from other works by that author, then it is, itself, by that author. This is a way to respect the text speaking for itself, and to dismiss anything that is not relevant to the work. Although there are some counterintuitive consequences to accepting human authors' ability to "produce" novels long after their death, or without their knowledge, or in ways utterly separate from the traditional causal chain linking human and work, none of these succeed in undermining the reasons for believing that these are the work of the human author in question. This view of authorship has a number of implications for human agency, intellectual property law, and discussions of artist attribution in generative art, and merits further investigation.

10. References

[1] Linson, A. (2016). Machine Art or Machine Artists?: Dennett, Danto, and the Expressive Stance. *Fundamental Issues of Artificial Intelligence*, 443-458.

[2] Ghosh, A., & Fossas, G. (2022). Can there be art without an artist?. *arXiv preprint arXiv:2209.07667*.

[3] Browne, K. (2022). Who (or what) is an AI Artist?. *Leonardo*, *55*(2), 130-134.

[4] Barthes, R. (1967). The Death of the Author*, Aspen 5-6*.

[5] Helle, S. (2019). What Is an Author? Old Answers to a New Question. *Modern Language Quarterly*, 80(2), 113–139.

[6] Deltorn, J. M., & Macrez, F. (2019). Authorship in the Age of Machine learning and Artificial Intelligence, in *The Oxford Handbook of Music Law and Policy.* Oxford University Press.

[7] Ginsburg, J. C., & Budiardjo, L. A. (2019). Authors and machines. *Berkeley Tech. LJ*, *34*, 343.

[8] Noë, A. (2000). Experience and Experiment in Art. *Journal of Consciousness Studies*, *7*(8-9), 123-35.

[9] Jones, R. & Penny, M. (1983). *Rafael*. Yale University Press.

The Beauty of the ruined building

Ben Baruch Blich, ph.d.
The Interdisciplinary program of the Arts
Faculty of the Arts, Tel Aviv University

Ben Baruch Blich

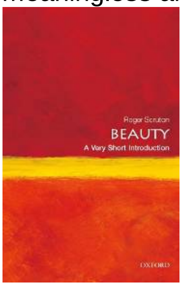
Abstract

The purpose of my paper is to unravel a much debated and classical subject in the arts, i.e.: the ongoing question 'what is beauty'; a question discussed by philosophers, artists and critics since antiquity, enlightening it from the point of view of architecture, maintaining that what is considered ugly, is in fact beautiful, gratifying and interesting. Devasted buildings are part and parcel of our daily surroundings. Most of us turn our back to them saying that they are unpleasing to see, they are ugly, horrible and unattractive. My paper advocates for a new and an unconventional aesthetics based upon the principle of inconsistent patterns, demonstrated by devastated ruined and uneven spaces of derelict buildings, which motivated artists and architects, to abandon the familiar and the conventional in favor of a new kind of beauty.

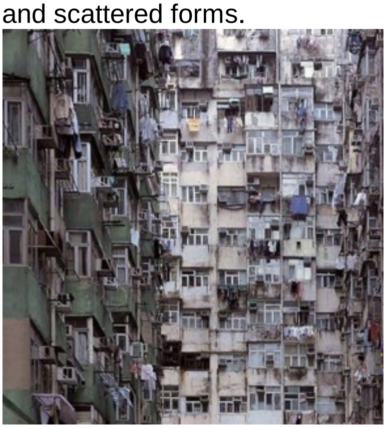
The purpose of my paper is to unravel a much debated and classical subject in the arts, i.e.: the ongoing question 'what is beauty'; a question discussed by philosophers, artists and critics since antiquity. I am aware that it might sound a bit arrogant to suggest a solution or even get close to an intelligent proposal how one should tackle the problem. Instead, I would like to enlighten this overloaded and much discussed subject from a different point of view, and put emphasis on one medium – architecture.[i]

My main points are as follows: too many writers were concentrated on the dichotomy between beauty and ugliness. It was Umberto Eco, a well known celebrated Italian philosopher and novelist, who have enumerated in two detailed and thick books, the chronological timeline and crossroads of beauty on the one hand and ugliness on the other.

By separating beauty from ugliness, he implicitly says that these two pillars of the arts, are two contradictory concepts and that there is nothing in between. The same goes with other philosophers, such as Roger Scruton, who in some of his major writings, and lectures maintained that beauty is an element of survival without which our life would become meaningless and trivial.



Beauty, according to him, is a value as important as truth and goodness. He goes on saying that in the 20th century beauty has stopped being important. Art increasingly aimed to disturb and break moral taboos. It is not beauty, but originality however achieved and whatever cost. Not only that art made a cult of ugliness, architecture too has become soulless and sterile. Ugliness repels us, by its deformed shapes, unsettled proportions, irritating and scattered forms.



St. Petersburg Stanilist apt. block

Eco as well as Scruton, to mention two salient advocators of the old notion which says that beauty and ugliness are two separate values, somehow ignore the randomness of modern life which cannot be redeemed alone by beauty without incorporating ugliness as its counterpart; that the one cannot survive without the other. In his frequently cited slogan of the American architect, Louis Henry Sullivan,[ii] the father of the skyscrapers in Chicago, who said: 'form follows function'. By these 3 words he envisioned the credo of modernism: stop thinking on the way a building looks and think instead on what it does, the shape of a building or an object should primarily relate to its intended function or purpose. It was Sullivan who opened the gate for a new angle of appreciating of the arts, and especially architecture and design, combining beauty and ugliness under the same umbrella.

Louis Henry Sullivan 1856-1924

It is on this line of thought that I want to talk on new approach to aesthetic of buildings, and convince you that even *left overs, debris*[iii] as well as *ruins*, *demolished houses*, we mostly derogate as ugly and disposable dirt, are nevertheless appealing, interesting and one can find them imaginative and even beautiful. In a nutshell, my aim is to point at the role of the demolished as a means to enhance our imagination, which represents in a roundabout way the randomness we witness in our daily postmodern life, described by Michel Foucault as *heterotopias*. i.e.: as transitional, diverse, disturbing, intense, incompatible, contradictory spaces.

Delacroix, Le lit défait, 1827

Now, what is the source of this approach in architecture on which I intend to elaborate shortly. Let me take you to the first half of the 19th century, and point at a peculiar painting by the French artist, Eugene Delacroix, *Le lit défait*, a water color in which he depicted his own untidy bed.
As you all know Delacroix is renowned for his romantic paintings, representing the revolutionary *zeitgeist* of the period in France, and yet he took time to paint his own private and intimate bed as if to say 'let me to show you what really interests me in the arts'. Indeed, looking at the painting, we easily notice the bed's chaotic appearance – the mattress is hidden under an enormous disordered pile of sheets, blankets and pillows, as if a nightmare battle took place during his sleep. But if we leave aside the psychological mood of the scene, and concentrate only on the meaning of its visual metaphor, one cannot escape the conclusion that Delacroix refers by this small painting to the historical turmoil taking place outside of his bed room; a turmoil that has shifted the ruling powers from the orderly, neat, decorated palaces, to the numerous barricades instantly built by the protesters in the streets and squares of Paris.

Barricades in Paris 1830

Delacroix's bed renders a new kind of space, though the painting is small and limited, but nevertheless it denotes and exemplifies the contrast between the King's royal bed which was intended for ceremonial purposes rather than for sleeping, now displayed in the Louvre, and the bed of a citizen aspiring change, freedom, equality, etc.


The bed of Charles X, King of France, 1824-1830

In short, Delacroix has established a new approach to the aesthetics of space; a space freed from conventional rules, a man-made improvised place, defined later by Claude Levi Strauss, as a *bricoleur* who creates and designs his surroundings ad-hoc without a blueprint of a preplanned sketch using mostly various simple materials available on hand.

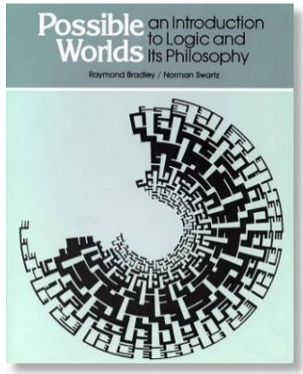
This new aesthetics which was stipulated by Delacroix, was not at all new to architecture; architects were for ages fascinated by the unplanned chaotic spaces, which served as a sharpener to their imagination, but only few of them have incorporated them into their work. The first, and the mostly famous is the Venetian artist, architect and archologist Giovanni Battista (or Giambattista) Piranesi 1720-1779, who was fascinated while visiting Rome by its remains of ancient buildings, appreciating their poetic aspects of the ruins: "speaking ruins have filled my spirit with images that accurate drawings could never have succeeded in conveying".[iv] The encounter with the ruins, the debris, the missing scattered parts of the buildings, stimulated Piranesi into creating his well-known prison etchings series he started to engrave in 1745.


Piranesi, Carceri d'invenzione (Imaginary Prisons) 1749-50

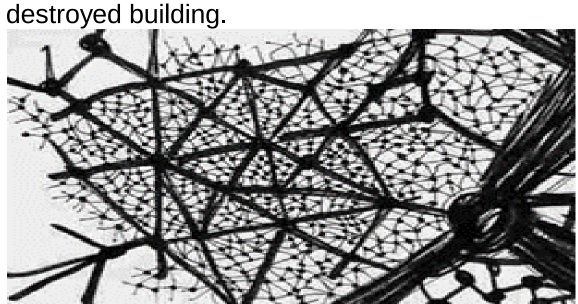
The imaginary prisons (*carceri d'invenzione*) include 16 etchings depicting enormous subterranean vaults with stairs and mighty machines, and although all of them are, as said, imaginary and do not stand the test of real life, they nevertheless reflect his desire and craving to get hold of the dark side of life, as so many other romantic artists before him. Piranesi's engravings of architectural landscapes offer a new way of seeing architecture and urban space. By putting an emphasis on the ruins of Rome,[v] Piranesi derives our attention to the effects of light, proportion, scale, etc. but above all, he invokes us to take into account the debris, the ruins, and the neglected parts of a building, as part and parcel of architecture. In short, Piranesi has brought to light the beauty of the neglected, the decayed; the beauty of and in the deformed and the ugly.

Piranesi, 1749-50

Piranesi's art echoes a much broader aspect of western civilization philosophy. Since antiquity philosophers of logic were questioning the status of truth: does truth reflect the world we live in, or are there alternative worlds (*possible worlds*) in which truth is not the only tool to grasp reality. I do not intend to answer in full this question, though logic was, and still to this day, a reliable key to understand the world, there are nonetheless lots of paradoxes and counterfactuals which cannot be settled by truth itself, such as what might be true, what should be true, what one believes to be true, etc. The logic of *possible worlds* deals, therefore, with the ways things could have been, leaving behind the real and the evident. Take, for instance, science fiction literature, cinema, as well as *AI* – all suggest a totally different approach to the evident, denying truth as the sole key to reality.



In this vein one should read the writings of Gilles Deleuze and Félix Guattari[vi] who use the terms *rhizome* and *rhizomatic* to describe a network that "connects any point to any other point" – a reality of random and endless possible connections in opposition to a tree-like use of concepts, which work with dualist categories (truth-false) and binary choices (beautiful-ugly). Piranesi's imaginative architecture falls under the category of possible worlds; he has demonstrated that architecture should unravel spaces which we wrongly consider not relevant or not related to architecture, such as the remains of a destroyed building.



Rhizome is a philosophical term used to describe the relations and connectivity of things.

Piranesi's legacy found its traces in Lebbeus Woods stimulating writings and architectural plannings. And indeed, most of the references attributed to Woods, consider him as an architect in the tradition of speculative, fantastical and sometimes grandiose architectural propositions, much the same as Piranesi and the French architect Étienne-Louis Boullée. Woods admitted that he is "an architect, a constructor of worlds, a sensualist who worships the flesh, the melody, a silhouette against the darkening sky".[vii] As an unconventional architect, Woods introduced a new term *freespace* - an architectural approach freed from any conventional, predetermined, preplanned design. It is a space addressed to the empty, a mass made of scattered, leftovers, redundant, obsolete materials of all kinds – concrete, blocks, iron, wood, etc. His fascination of the many facets of space, or as he put it *freespace*, brought him to the conclusion that even debris and ruins, are legitimate materials in the framework of architecture. Though ruins and debris are not manageable and workable entities, they nevertheless have fascinated Piranesi and many other artists and poets later on in the 19th and in the 20th century. Beginning with the romantic painter David Caspar Friedrich, through the photographer Robert Capa, up to the film director Andrei Tarkovsky, to mention few of the numerous artists who took pleasure of the ruined, fascinated by the demolished and mutilated buildings.

The American architect Gordon Matta-Clark[viii] is a *site-specific* artist, famously known for his video art and photography of derelict, abandoned unoccupied buildings. He, as you know, left his training as an architect, and turned the idea of 'building' inside out – inventing a new concept 'anarchitecture', in an endeavor to locate abandoned buildings scheduled for demolition, and stage in them a massive structural intervention. Matta-clark sliced through walls, floors, doors, windows, and anything that was in his way, creating negative spaces, emphasizing 'the dark side of architecture', with the intention to show their relevance to the work of the architect. His renowned projects such as *cuts* (1970), *splittin*g (1974) and many others, in which he radically subverted architecture and the urban landscape by cutting and slicing to produce a novel aesthetics ignored and even condemned by the architectural society as ugly, unattractive, a wreck, to be demolished and eradicated.

Gordon Matta-Clark, Cuts 1970

Gordon Matta-Clark, splitting 1974

The emotional impact of the ruined is ambiguous – on the one hand it cannot be considered genuine building since the original intention of the builder has been lost, but on the other hand we cannot resist to their appeal as remains, as derelict, or may I say as sacred relics of civilization. No wonder, therefore, that architects found a deep interest in the destructed building, either as a source of inspiration or as a key to the understanding the substance of architectural design. Lebbeus Woods[ix] was, therefore, an architect who found the desecrated building a source to his work. He not only wrote on wrecked buildings in war zones, he also incorporated a shapeless and an unconventional decoration to his works which hints at the scattered, the ruined and the destroyed.

A computer reconstruction by Lebbeus Woods in 1994 of the Electrical Management (Elektroprivreda) Building in Sarajevo, Bosnia

Leubbeus Woods' sketches of earthquakes spaces

According to Woods, a destructed space constitutes a form of noise or chaos, which he labelled *multiplicity.* By multiplicity he meant a composition of undefined elements with indistinguishable trajectories, forming an aggregate of spaces, an indeterminate motion, without a specific direction from one point to the other. It is a chaotic site, a consequence of violence, which triggers a series of endless changes and transformations, some of them are brutal, fierce, and repulsive, but nonetheless they are exciting aesthetic possibilities, challenging and sensitive opportunities for architectural awareness. Leaving aside ethical dilemmas, Woods did not condemn war zone's destructions as such; he visited Sarajevo right after the Bosnia war and the Siege of Sarajevo in 1995 which stimulated him to design with steel cables on the sites destroyed by war bombs. As an architect, he found value in its morphological fluidity – houses partly demolished serving him as a key to create new spontaneous reconstructions to the benefit of the city's inhabitants.

<u>Epilogue</u>

As I am writing this paper, my country is at war in the south and the north. I have witnessed ruined buildings all over – on our side of the conflict, and no less in Gaza. Wars spread devastation, despair and hatred. Myself, and most of my colleagues do not belittle the harsh consequences we are facing, and we are eager to stop the violence in our region. However, in light of the ideas I have painfully tried to justify and support, I cannot refrain looking at the bombed buildings the war has inflicted upon us from an aesthetic perspective. In spite of the fact that these stone and concrete made buildings do not function as they were intended to, their complexion and stance are intriguing, exciting, and may I say attracting.

In fact there are architects who have suggested to design buildings looking much the same as those damaged, such as Justin Plunkett, and many others.

Justin Plunkett's "Con/struct" Series Imagines Future of Urban Sprawl

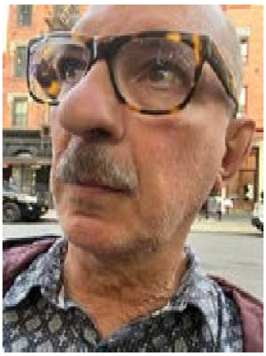

A 3d model of a ruined house

To summarize my paper, I suggest that we have to advocate for a new and an unconventional aesthetics based upon the principle of *inconsistent patterns*, demonstrated by the devastated ruined and uneven spaces of derelict buildings, which motivated artists and architects, to abandon the familiar and the conventional in favor of a new kind of beauty, what is generally considered unattractive, unpleasing, horrible, or in one word 'ugly'.

[i]Martin, Louis. “Against Architecture.” Log, no. 16 (2009): 153–67.
[ii] Abbott, James R. “Louis Sullivan, Architectural Modernism, and the Creation of Democratic Space.” The American Sociologist 31, no. 1 (2000): 62–85.
[iii] Gissen, David. “Debris.” AA Files, no. 58 (2009): 8–11.
[iv] John Wilton-Ely, *The mind and Art of Giovanni Battista Pirenasi* (London, 1978), p. 45
[v] Piranesi, Giovanni Battista, and Gillian Furlong. “The Ruins of Rome, Seen through 18th-Century Eyes.” In Treasures from UCL, 1st ed., 112–13. UCL Press, 2015.
[vi] Thayer-Bacon, Barbara J. “PLANTS: DELEUZE’S AND GUATTARI’S RHIZOMES.” Counterpoints 505 (2017): 63–88.
[vii] Woods, Lebbeus (2002). War and Architecture. New York: Princeton Architectural Press. p. 1
[viii] Walker, Stephen. “Gordon Matta-Clark: Drawing on Architecture.” Grey Room, no. 18 (2004): 108–31.
[ix] Woods, Lebbeus. “After Forms.” Perspecta 38 (2006): 125–32.

Embodied Interaction with Algorithms

Peter Beyls, PhD
The School of Arts, University College Ghent, Belgium
www.peterbeyls.net
peter.beyls@hogent.be



Abstract

This paper primarily contextualizes the *Material Matters* (MM) research project and documents some tentative experiments. MM envelopes multiple areas of interest – (computational) aesthetics, philosophy, embodied cognition and cognitive robotics, specifically, human-robot interaction. In particular, the principle of mind-body-environment unity emerges as a first principle. MM centers on the philosophical notion of *embodied cognition*, rejecting cognition as the construction and manipulation of representations of abstract symbols. Then, creative processes as constituted and afforded through motivated bodily activity in a situated physical setting. In the context of the art studio, such inclination marks the difference between creative development through impulsive improvisational object handling and the synthesis of new ideas through pure contemplation. Starting from earlier work in gesture-based interactive composing, we progress into human-robot interaction probing the role of objects and materials in the artist's studio in relation to software-initiated initiative.

1. Introduction

This paper relates the production, experience and interpretation of generative art by humans and machines and hypotheses on hybrid, cooperative formats. However, the depth of human aesthetic experience is informed by an unknown number of mostly unconscious dimensions, including social, cultural, historical and materialist scopes, (consider the smell of oil paint). Machine experience and production follows known dimensions because implemented as explicit models in software. From this perspective, how could we possibly create a closer relationship between collective initiative by human and machine – MM suggests experiments engaging humans and machines considering notions of identity, co-operation, motivated action, adaption and freedom. A robotic painting experiment studies the dynamics of a human-machine collective through reinforcement learning and observation of exclusively physical activity by human and machine. However, as a main focus, a rational framework is identified as a first step as documented below.

Human-machine interaction (HMI) in the musical domain, referred to as interactive composing [1], aims compelling HMI co-creation ex *tempore* – first, algorithms are developed anticipating complex interplay and second, the program performs in confrontation of human input. Conversational interaction develops, from tight instrumental control, to play-and-response interaction, to opaquely surprising feedback avoiding direct echoes of human input, to exploiting a motivation-driven strategy using real-time reinforcement learning [2]. Sound is the substance of interaction in computationally driven musical improvisation, musical fabric develops over time afforded by two apparently contrasting potentialities; human presence grounded in global cultural awareness and capacity of situated action, and the comparatively tiny algorithmic universe embedded in the program's logic. Yet, convincing procedural HMI interaction matures merging embodied presence and isolated algorithms existing only in silicon, still, a strategy of reinforcement learning allows for gradual human-machine adaptation in a real-time co-creative process [3]. Typically, HMI platforms should afford simultaneous and equally valued input from human and machines while maximizing behavioral complexity and suggesting a rewarding human experience informed by the delicate balance of meaning and mystery in the unfolding interaction.

We might attempt to chart the non-trivial intrinsic qualities and otherwise tightness of the human-machine relationship suggesting tense collaborative human-machine co-creativity to exist in affective and symbiotic emergent modalities rather than explicitly designed procedural interaction formats. Indeed, live-like systems exhibit unpredictable yet coherent behavior, not just loosely random pursuit. However, according to Gerhard Richter: "randomness does it better than me" and a life-time of intriguing work was founded on the core concept of controlled randomness by the Vera Molnar [4]. My early work conceptualized randomness as an adaptable, adjustable source of energy to activate inclusive degrees of freedom in a given algorithmic procedure, randomness is not visualized as such, it resides as a force enabling the algorithm to reveal itself by activating its implied behavioral scope [5, final chapter].

Paradoxically, we target the exposure of attention-grabbing machine-specific behavior, however avoiding explicit design, though consciously designed software is definitely still involved. In other words, we expect consistent performative autonomous behavior in software agents, not superficial reactive actions triggered from a palette of premediated options. Contrasting automatism and autonomy is key. Genuine autonomy entails development; the artificial agent acquires increasingly complex cognitive abilities by interacting within an unpredictable environment. For example, developmental robotics combines ideas of social interaction, embodied, situated and enactive development, and continuous online open-ended cumulative learning to achieve ever-more-complex skills [6].

The present paper extends HMI into an object-based and material-oriented studio practice, proposing a critical consideration of the discrepancy between the discreteness of the digital medium and the continuous nature of human perception and cultural engagement. Then, how could software-initiated activity and spontaneous bodily action merge into a new medium affording innovative human-machine co-creation? The MM research project aims to study and analyze the complexity of hands-on material and object centered studio practice within a digital context. For example (1) how to integrate affective gesture and algorithmic decision making, and (2) how to extract underpinning procedural patterns from engaging with physical materials. Our objectives are addressed through experiments in human-machine co-creation, affording creative action from and within both partners, with both entities being grounded in a shared physical environment. If, in the light of computational creativity, we acknowledge the hypothesis that art is fundamentally a social phenomenon [7, p. 9] and art as communication and sharing between people and art-making as an adaptive product of our biological evolution – "we do expect that the AI has to say something socially, something that suggests an inner consciousness and feeling" [8, p. 141], as such, software is still a long way from being considered art.

However, ideas and abstractions of human psychology can indeed function as design principles for artificial agents-based systems spawning complex emergent behavior from the specification of mutual social affinities, for example, user-supplied gestures viewed as virtual creatures developing complex behavior while adaptive and sensitive to neighbor creatures [Figure 1] to a self-organizing agency driving digital sound synthesis. Our present research aims to grow a similar distributed platform for social negotiation as instantiated in human and materialist non-human participants.

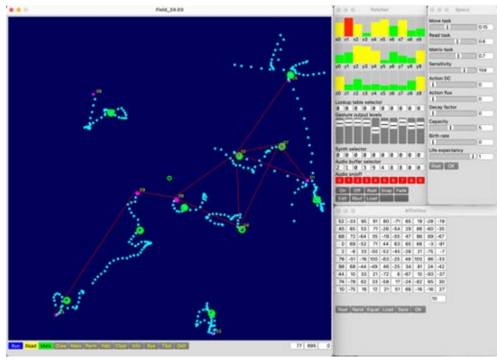


Figure 1: Living gestures audio interface

2. Contextualization

Renowned neuroscientist Antonio Damasio observes "The universe of affect – the feeling experiences derived from drives, motivations, homeostatic adjustments and emotions – was a *prior historical manifestation of intelligence*, highly adaptive and efficient, and was a key to the appearance of growth and creativity" [9, p. 181]. The author challenges the traditional Cartesian separation of mind and body, highlighting the crucial role of emotions in decision-making and rational thinking. Damasio advocates for a holistic understanding of the mind and body, emphasizing the interconnectedness of emotional and cognitive processes.

MM aims to close the mind/body gap through the development of (1) an extended contextual framework to address notions of materiality and object-centered activity relating to digital culture and (2) implement skin-tight innovative partnership formats in human-machine interaction (HMI) as to foster co-creation in a common physical ecosystem. The problem of human-machine interaction is explored in the context of the artists' studio taking a unified approach in the sense that human bodies, machine hardware, a widely diverse range of production tools, objects and materials all contribute to a materialist environment in continuous flux. All components contribute to the emergence of affordances allowing tangible creative exploration. When discarding any mind-body, software-hardware and analog-digital dualities, a key question arises: how do generative algorithms relate to this setting?

A plethora of tools has entered the studio, some situated at extreme opposites of the complexity spectrum; consider a worn-out pencil as contrasting to computer animation using a quantum computing [10]. However, media value and complexity should be considered in relation to its impact on the creative process – a pencil typically drives rather than implements a cognitive process – Richard Feynman insisted "I actually did the work on the paper (…) it's not a record, not really, it's working. You have to work on paper and this is the paper" [11, p. 256]. On the other hand, random ideas generally instantiate unconditioned by the pressure and technological burden of complicated media. The point I am trying to make is to stress that "ideas" are not machine-generated but unconsciously instantiate, typically from exposure to open-ended, undefined environments (consider, for example, a walk in the park) addressed without prejudice or unambiguous goals. Then, bodily experiences suggest procedural action implemented as algorithms: aspects of a macro universe reflect in the stylized (logical) micro universe of a computer program.

Many of my generative systems explored randomness as catalyst, it does not get visualized, it is not perceived as such though it affords the excitation of a maximum degrees of freedom implied in a supporting associated algorithm. However, "to come up with new and radical strategies, we need radical diversity of representation and ability" [12, p. 246]. Then, randomness is not perceived as the opposite of informed decision-making but underpins the *maximization of diversity* as a primary working principle.

3. Embodied Cognition

From a much wider perspective - from philosophy to Artificial Intelligence - the role of situated activity versus formal operations in human cognition is being considered thoroughly, how do motivated human bodies and programmed machines interact by virtue of sharing a common physical habitat? Answers are found in the concept of Embodied-Cognition (EC), and rejecting cognition as based on representations of abstract symbols. EC considers thinking processes as constituted and afforded through motivated bodily activity in a situated physical setting. In the context of the art studio, such inclination marks the distinction between creative development through impulsive improvisational object handling and the synthesis of new ideas through pure contemplation. MM discards the mind-body, software-hardware and analog-digital dichotomies, MM looks for materials/machines to respond and behave not by explicit programming but by the nature of their physical structure.

"Philosophy in the Flesh: The Embodied Mind and Its Challenge to Western Thought" by Lakoff and Johnson [13] explores this relationship between human cognition and the body, challenging traditional Western philosophical views: human thought is fundamentally shaped by the body's sensory motor experiences and deeply rooted in the way our bodies interact with the environment. Lakoff and Johnson also introduce the concept of *embodied realism* suggesting that abstract concepts and metaphors are grounded in our bodily experiences, cognitive processes such as image-schemas illustrate how our bodily interactions shape our conceptual systems.

From another angle, a highly relevant philosophical theory is *speculative realism* which seeks to move beyond anthropocentrism and explores the reality of entities independently of human perception. As part of Object-Oriented Ontology (OOO), asserts that objects are the primary focus of ontology [14]. Objects can include not only physical entities canvas and paint but also non-physical entities such as aesthetic concepts. OOO promotes a *flat ontology*, meaning all objects, whether human or non-human, are considered equal in terms of ontological status; it challenges hierarchical views that prioritize certain entities (humans) over others. Again, this fits my first principles of maximization of diversity and bootstrapping from scratch, object and subject are treated as equal partners, “what are normally called subject and object are simply *aesthetic* properties that are shared between objects … in the OOO universe, aesthetic experience is real and tangible yet unspeakable” (15, p. 63). Then, if machines are on equal (creative) authority with humans, could they address the environment consciously – since “nothing can be known in the absence of consciousness” [16, p. 114]. Accordingly, nascent feeling machines’ experiences are derived from motivations and homeostatic adjustments, meaning intelligence through affective adaptation. Consequently, feeling robots should exist as physical bodies requiring regulations and adjustments to persist with sensors reporting internal and external states, leading to optimized global behavior from the network of competing processes in the body and the environment. Therefore, the critical idea of machine autonomy is addressed, in particular, how it might contribute to innovative and unconventional human-machine co-creation.

Autonomy closely relates to the notions of *autopoiesis, co-evolution* and *enactment*: with humans in the loop [17], humans and artificial systems coalesce, interact and make sense of their common environment (the art studio) – the process of enactment takes place. Seminal work by Maturana and Varela [18] defines enaction and autopoiesis (self-creation) as key to life itself; living organisms constantly regenerate and maintain their own components, enaction describes the active role of the organism to create its own reality, and cognition is not a representation of an external reality but an ongoing, dynamic process that emerges from the interactions of an organism with its environment. Then, enaction supports participative AI, mediated creation of meaning from human-environment interaction, therefore, it is deeply rooted in radical empiricism [19].

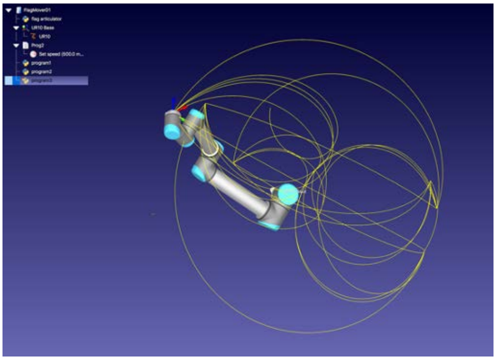


Figure 2: Finite State Machine robot action

In addition to participative AI, thinkers in art philosophy wonder how to be complicit with materials, [20, p. 12], media design [21], aesthetic theory [22] and social psychology [23] are asking a fundamental question: how to build media supporting qualitative coexistence of human and non-human participants, media balancing material diversification and expanding the human experience into a wider shared ecology.

4. Art/Science context

From art historian Clement Greenberg's definition of art materials as autonomous with distinct properties to physicist Karen Barrad's notion of *material complicity* – viewing material agency as a network of relationships rather than a property of things, discussions of materiality in art persist [24, 25]. Concept art introduced a shift from object-based art to more conceptually driven process-oriented practice. However, installations, performance/body art, and interventions in specific locations became prominent during this time, materiality was re-invented rather than abandoned. Consider Duchamp's ready-mades, the physical incarnation of a large number of symbolic elements (*la machine célibataire*) as a large glass sculpture and, most prominently, *Étant Données* (1946-1969) built from wood, nails, bricks, brass, aluminum sheet, steel binder clips, velvet, leaves, twigs, a female form made of parchment, human hair, glass, plastic clothespins, oil paint, linoleum, an assortment of lights, a landscape composed of hand-painted and photographed elements and an electric motor housed in a cookie tin which rotates a perforated disc [26]. Afterall, Duchamp proves objects are not reducible to the materials, objects afford multiplicity of thinking, subsequently "the idea has to be awfully good to compete with the object" [24, p. 31]. According to Sherry Turkle, "We think with the objects we love; we love the objects we think with" [23] – *evocative objects* are objects that express a meaning beyond their purely instrumental functionality, they act as emotional companions, relate to one's memories, support a relational attitude and trigger the imagination.
Then, how could specific *evocative objects* contribute to the experience of a corporeal aesthetics-oriented human-machine playground? Experience-centered aesthetics as total engagement was recognized early on by American philosopher John Dewey; *art as experience* emphasizes materialist interaction, the integration of art into the fabric of daily life thereby suggesting a holistic, inclusive approach to understanding and appreciating art [27]. Then, human-machine co-creation could

be viewed as a hybrid network linking art-historical awareness, cognitive abilities in machines, objects and humans configured as a complex dynamical system, pushing the system in various behavioral orbits through internal and external activation might reveal an unexpected yet compelling experience. Obviously, a fascination with unscripted machine-initiated action persists, however, idiosyncratic machine behavior is oxymoron with scripted instruction, what could a blend of bits and atoms in an analog/digital amalgam bring to the discussion? We might be able to describe how machines work, but what do they experience, what is their proper phenomenology? [28 p. 10]. Flat ontology suggests things can be many and various, specific and concrete, while their *being* remains identical – material objects and abstractions co-exist on equal levels of authority, another fundamental holistic vision.

While the idea of computation is unavoidably connected to human understanding and awareness, typically, engineering concerns over optimization prevail rather than efforts of considerate understanding of the machine by itself. Post-humanism rejects the notion of human-world (man-machine) correlate, humans are entangled and implicated in other beings. Flat ontology synthesizes the human and non-human into a common collective [29, p. 33], then, biological and synthetic/artificial agents might articulately co-exist on a common existential stratum. As objects are to be understood as both concreate and abstract, how could such intimate symbiosis be implemented to anticipate emergent creative action?

Figure 3: CNC machine with brush

Ian Bogost recommends a similar speculative question – while rejecting a selfishly anthropocentric view – how can we nevertheless deal with complex structures crafted by humans? Relevant to our hands-on approach, a pragmatic speculative realism of object-oriented engineering to physics ontology is put forward: “as philosophers (…) our job is to amplify the black noise of objects and to make the resonant frequencies of the stuffs inside them hum in credible satisfying ways” [28 p. 34]. Then, objects are not reducible to material substance but open to function as speculative triggers; this is why “the artwork is a prime example of the object’s capacity to evade the knowing grasp” [28, p. 14].

5. Experiments

Over the years, I have often contemplated on emotive relationships with machines, in particular computers, in the sense of how the physicality of the programmer (human body) relates to the materiality of the medium (computer hardware). For example, a patch on an analog computer is a physical reflection of an idea, itself a hypothetical imaginary procedure yet implemented as corporeal object. Handling DEC tapes on a PDP15 mainframe computer producing both intensive heat and excessive noise strongly impacts the programmer, intellectually and emotionally in an intimate attachment: the machine manifestation embodies the programmer. Programming in assembly language, counting processor action in nanoseconds versus using high-level symbolic languages such as LISP entails distinctive degrees of authority and reciprocal action. Today, computing impacts society on a larger scale, consider distributed laptop performance geologically extending situated/embodied cognition into sociocultural space in global networks. All instances prove human imagination to be affected and motivated by materiality, tangibility and overall physicality of the digital medium. As mentioned before, research in Digital Aesthetics views itself in bottleneck by accounting for the ontological discrepancy between the continuity of human perception and cultural engagement and the discreteness of digital technology. In addressing this problem Simon Penny suggests an aesthetics of behavior - cultural action emerging from a "reconceptualization of cognition as embodied, enactive, and integrated with the material and cultural world" [30 p. 213], and in a wider scope, viewing intelligence as residing at the intersection of the body and the world, viewing sensing and action as indivisible, and (following Varela) identifying action as a structural coupling between body and world.

Figure 4: UR10 robot setup

Figure 5: UR10 robot watercolor

A wide range of robotic infrastructure is being investigated in recent machine art, from musical robotics, social robots as creativity eliciting agents and embodied performance with robotic actors [31]. My early robotic work explored the machine's behavioral scope based on a probabilistic Finite State Machine algorithm [Figure 2].

Two approaches to machine drawing and painting within MM are briefly documented here: CNC machine and UR10 robot arm. CNC explores structural painting using brushes equipped with containers holding colored ink [Figure 3]. Initially, UR10 explored watercolor painting based on earlier software driving a pen plotter. A Python program runs inside the scripting engine of RoboDK [32]: take water and paint using a timing algorithm adapting to the materiality of the painting process itself [Figures 4, 5]. The current UR10 implementation takes the human in the loop, human and machine painting on a common surface, the robot using computer-vision to interpret input by both parties. Human and machine (H/M) take turns in a conversational process of action (select a color, decide how, what and where to paint and act) and perception (analyze the current contents of the painting surface and create a list of features). Our goal is to speculate on developing prolific emergent dynamics of social H/M interaction avoiding explicit instruction altogether. Questions arise: can we bootstrap interesting behavior from scratch, how to guarantee equal authorship of human and machine and how to maximize diversity without resorting to blind uncertainty. In addition, given a materialist platform, how to sense and make sense of tendencies in corporal human and machine activity? UR10 balances two competing motivations: H/M both aim to express a personal character while also intent to integrate into a larger social whole. A measure for social H/M proximity is related to (1) the physical distance, (2) the painting amount/surface and (3) color similarity between consecutive human and machine inputs. Intervals in proximity reflect intentions: for example, H painting nearby in space and with comparable colors mirrors H wishing to temporarily connect to M. H and M gradually learn about each other's aesthetic orientation. M uses various computer-vision tools (OpenCV) to discriminate between successive H actions and capture amount and complexity of H input. M holds a collection of speculative motivations, represented as arrays of types of possible actions and their respective efficiency which, itself, is updated according to intervals in social proximity – smaller intervals imply increased efficiency i.e., H and M integrating. A second adaptive process considers eventual M expression: machine motivations are adjusted using specific relationships (implemented as non-linear functions) between a list of features derived from human input and the motivation arrays. Then, a strong connection is being appreciated in the H/M abstract dialog, mutual activity seems coherent yet unpredictable. The software is being finalized right now; comprehensive results will be published in a subsequent paper.

Acknowledgements

I am indebted to my colleagues Elias Heunink and Thomas Janssens for technical and logistic support. The MM research project is financed by the Flemish Government of Belgium.

Flanders
State of the Art

References

[1] Chadabe, J (1989) Interactive Composing, an overview, the Music Machine, C Roads (ed.), MIT Press, Cambridge, pp.143-148

[2] Beyls, P (2019) Art as a Living Interface, *Springer Lecture Notes in Computer Science*, HCII, Orlando, FL

[3] Beyls, P (2018) Motivated Learning in Human Machine Improvisation, *NIME2018 Conference Proceedings*, Virginia Polytechnic Institute and State University, Blacksburg VI

[4] Molnar, V and Baby, V (2022) *Vera Molnar: Interview with Vincent Baby*, Manuella Editions, Paris, France

[5] Beyls, P (2015) *Simple Thoughts*, MER Publishing, Belgium

[6] Cangelosi, A and Asada, M (2022) *Cognitive Robotics*, MIT Press, Cambridge, MA

[7] Bown, O (2021) *Beyond the Creative Species*, MIT Press, Cambridge, MA

[8] Hertzmann, A (2018) Can Computers Create Art? *Arts*, 7, 18

[9] Damasio, A (2005) *Descartes' Error: Emotion, Reason, and the Human Brain*, Penguin Books

[10] Lioret, A (2021), Proceedings Generative Art Conference, Rome, Italy

[11] Clark, A (2013). Gesture as Thought? *The Hand, an Organ of the Mind; What the Manual Tells the Mental* (pp. 255-268). MIT Press, Cambridge, MA

[12] Bridle, J (2023) *Ways of Being, Animals, Plants, Machines: The Search for a Planetary Intelligence*, Pinguin Random House, UK

[13] Lakoff, G and Johnson, M (1999) *Philosophy in the Flesh The embodied Mind and its Challenge to Western Thought*, Basic Books

[14] Harman, G (2010) Object-oriented Philosophy, *Towards Speculative Realism: Essays and Lectures*, Zero Books

[15] Morton, T (2013) *Realist Magic. Objects, Ontology, Magic* Open Humanities Press

[16] Damasio, A (2021) *Feeling and Knowing: Making Minds Conscious,* Pantheon

[17] De Loor, P, Manach, K and Tisseu, J (2009) Enaction-Based Artificial Intelligence: Towards Co-evolution with Humans in the Loop, *Minds and Machine*, 19, pp. 319-343, Springer

[18] Maturana, H and Varela, F (1992) *The Tree of Knowledge*, Shambala

[19] *Cosby, D (2013) The Philosophy of William James: Radical Empiricism and Radical Materialism* Rowman & Littlefield Publishers

[20] Lange-Berndt (2015) How to be complicit with materials, *Materiality*, Whitechapel Gallery/MIT Press

[21] Nitsche, M (2022) *Vital Media: Making, Design and Expression for Humans and Other Materials*, MIT Press, Cambridge, MA

[22] Fazi, B (2019) Digital Aesthetics: The Discrete and the Continuous, *Theory, Culture and Society,* Vol 36(1) 3 - 26

[23] Turkle, S (2007) *Evocative Objects: Things We Think With*, The MIT Press, Cambridge, MA

[24] Lippard and Chandler, 1968 The Dematerialization of Art, *Art International*, Vol. XII, nr. 2

[25] Hudek, A (2014) *Object*, Documents of Contemporary Art, Whitechapel Gallery/MIT Press

[26] Belting, H (2010) *Looking Through Duchamp's Door*, Walther König, Köln

[27] Dewey, J (1934) *Art as Experience*, Minton, Balch, University of Michigan

[28] Bogost, I (2012) *Alien Phenomenology, or What It's to Be Like a Thing*, University of Minnesota Press, Minneapolis, MI

[29] Bryant, LR (2011) *The Democracy of Objects*, Open Humanities Press

[30] Penny, S (2018) *Making Sense*, The MIT Press, Cambridge, MA

[31] Creativity and Robotics, 2020 *Frontiers in Robotics and AI Journal*, ICSR 2020

[32] RoboDK, https://robodk.com/

The Registration of AI-generated Works as UNESCO heritage: Human-algorithmic Common Heritage of Mankind

Dr Sun Park (Lecturer in Creative and Cultural Industries)
Department of Art History and Cultural Practices, University of Manchester, Manchester, UK
https://research.manchester.ac.uk/en/persons/sun-park
e-mail: sun.park@manchester.ac.uk

Abstract

This paper is based on the author's doctoral thesis titled *"The Registration of AI-generated Works as UNESCO Heritage: Distributed Knowledge Systems as Human-Algorithmic Common Heritage of Mankind"* submitted in May 2024 in partial fulfilment of the requirements for the degree of Doctor of Philosophy in Sociology of the University of Warwick in the UK (the doctoral degree was awarded in October 2024).

UNESCO has used the term 'Common Heritage of Mankind (CHM)' over the past 80 years. The term generally refers to human-made heritage that holds universal significance for all humanity and should be passed down to future generations. Under the umbrella term 'CHM', UNESCO registers important heritage in the Member States on its international lists: The World Heritage List, the Intangible Cultural Heritage List and the Memory of the World Register. UNESCO also adopted the *Recommendation on the Ethics of Artificial Intelligence (AI)* in 2021, encouraging its Member States to integrate AI systems into the identification and preservation of heritage. However, UNESCO does not have a registration system that acknowledges AI-generated works as UNESCO heritage.

Given that heritage is inherited information about human history, culture and identity, AI-generated works not only demonstrate technological advancement but also embody the values of heritage, following the existing ones. Just as the Pyramids in Egypt, the watchmaking skills in France and Switzerland and the movie 'Wizard of Oz' in the USA respectively show architectural value, craftsmanship merging mechanical and artistic skills and film-making technology, an AI-generated work is a socio-technical assemblage crafted by humans' objects, knowledge and values associated with their response to the development of AI. Especially, AI-generated works have a distinctive value from the existing types of

heritage: human-AI interaction. It is necessary to examine how to identify and preserve the values of AI-generated works as heritage for future generations sooner rather than later, without leaving a gap in the history of heritage.

This paper will briefly introduce the argument on the doctoral thesis. After explaining UNESCO's notion of the Common Heritage of Mankind, the paper will demonstrate the distinctive values of AI-generated works as heritage, centring on the four types of human-AI interactions derived from the author's interviews with fifteen AI practitioners.

The paper will then summarise the doctoral thesis' proposal for identifying and preserving AI-generated works as UNESCO heritage. This finding is based on the author's interviews with the fifteen AI practitioners above and fifteen former and incumbent UNESCO staff members and their advisors. Drawing on the analysis of the limitations and benefits of UNESCO's current three heritage registration systems, the doctoral thesis provides (a) five criteria for evaluating AI-generated works as UNESCO heritage, (b) the eligibility of nominators of AI-generated works (c) and preservation plans to be included in a nomination dossier of UNESCO heritage. By doing so, this research contributes to extending the discussion of AI-generated works to the nexus between cultural sociology and heritage studies.

1. The Common Heritage of Mankind

The term 'Common Heritage of Mankind (hereinafter 'CHM')' generally refers to heritage that is meaningful to humanity and should be preserved by human collective actions [1]. The then ambassador of Malta to the United Nations (UN), Mr. Arvid Pardo, launched the idea of the CHM in his speech at the UN General Assembly on 1 November 1967. Pardo proposed the regime of the CHM protected from national appropriation for the benefit of mankind as a whole and international peace. Beyond Pardo's discussion on the common seabed, the field of international law has extended the concept of the CHM. The UN adopted legal instruments to deal with the issues of obtaining natural resources and scientific information from nature on the earth such as the high seas, Antarctica and beyond the earth such as the moon and space. The UN originally understood the CHM as shared natural resources which they thought had not to be claimed by individual nation-states.

UNESCO has contributed to understanding the CHM in the cultural sense. The Constitution of UNESCO requires its Member States' "conservation and protection of the world's inheritance of books, works of art and monuments of history and science" [2]. In other words, the 'common heritage' that humanity shall protect in the human sector together is 'cultural' heritage created by humans, compared to the original meaning of the CHM as nature. It explicitly demonstrates that UNESCO has extended the notion of the CHM: heritage deriving from nature to heritage created by humans, i.e., the human-made CHM. In UNESCO's terms, the human-centred idea of the CHM generally refers to products, knowledge and values created, cooked and refined by humans and their technological skills and social relations, such as architecture, tradition and a group identity. The cultural notion of the CHM, of course, acknowledges the interaction between humans and nature. UNESCO also deals with the preservation of natural heritage, for example, in its 1972 World Heritage Convention. However, the original meaning of the CHM underscores the economic, political and intellectual values directly emerging from natural resources, whereas the cultural CHM in the sense of UNESCO comes from human values and identities. Over 80 years, UNESCO has developed the framework of the cultural CHM. By adopting 36 legal instruments including the concept of the CHM, UNESCO has addressed immovable monuments and sites, movable objects, intangible tradition, rituals, ceremonies, oral history and the performing arts, and information and technology.

The human-made CHM illustrates the achievement of humanity, functioning as evidence of humans' historical and cultural achievements. Artefacts, practices and documents from the past have been interpreted in various ways throughout history. This is why UNESCO accentuates the preservation of heritage in its CHM policies. Regardless of a clear identification of the original human creator of heritage, the existence of heritage per se provides us with informative sources to tell human history, culture and identity, letting us understand each other's cultural features. Heritage is therefore universal informative sources about humanity.

UNESCO's 80-year CHM policies also demonstrate the cumulative academic bases explaining different roles of humans in the CHM. The 1972 World Heritage Convention based on archaeology sees humans mainly as architects and artistic professionals. The 2003 Intangible Cultural Heritage Convention based on anthropology regards humans as inheritors and bearers of tradition and focuses on the roles of communities in sharing the same knowledge. The 1992 Memory of the World programme based on archival and media studies highlights the roles of humans as creative recorders of knowledge. The 2021 Software Heritage programme revolves around the skills and roles of individual computer programmers and private companies in the computational fields. The different academic origins of the CHM have enabled UNESCO to identify that the places where humans create cultural work have been gradually detached from the human body. From human hands with hammers to construct the Pyramids, to a digital space detached from human hands to create born-digital heritage. This epistemological mechanism indicates the possibility of a new UNESCO policy concerning AI-generated works, which addresses relevant scholarly bases of the distinctive value and roles of humans and AI embedded in AI-generated works.

2. UNESCO's Heritage Registration Systems

UNESCO has three heritage registration systems as of May 2024: the World Heritage List under the 1972 World Heritage Convention, the Representative List of the Intangible Cultural Heritage of Humanity (hereinafter 'the Intangible Cultural Heritage List') under the 2003 Intangible Cultural Heritage Convention and the Memory of the World Register (hereinafter 'the Documentary Heritage List') under the 2015 Recommendation concerning the Preservation of, and Access to, Documentary Heritage including in Digital Form. The three heritage registration systems show specific criteria for identifying certain heritage as 'UNESCO' heritage.

According to fifteen interviews with UNESCO's former staff members and incumbent advisors who participate in the decision-making on the registration of UNESCO heritage on the three lists, the doctoral thesis has found out that the Documentary Heritage system is more likely to include AI-generated works in their policy remit than the World Heritage and the Intangible Cultural Heritage systems. However, the World Heritage and the Intangible Cultural Heritage systems also provide theoretical legacies to discuss the value of AI-generated works as UNESCO heritage: the idea of authenticity and the significance of communities respectively.

The fifteen interviews also enabled the thesis to discuss that each time UNESCO recognised a new type of the Common Heritage of Mankind (e.g., the shift from World Heritage to Intangible Cultural Heritage), UNESCO noted distinctive aspects of new types of heritage to be preserved and transmitted to future generations. It raises the question of whether an AI-generated work also has two main values of heritage: 1) universal informative sources about humanity as explained above; and 2) its own distinctive value, compared to the previous types of heritage.

3. An Alternative to the Turing Test: The Durkheim Process

Alan Turing invented the 'Imitation Game' called the 'Turing Test' today in his paper *Computing Machinery and Intelligence* [3] The Turing Test has three main agents: one human and one machine, and the other human who tries to determine which of the other two is the human and the machine. The role of the machine is to mimic human intelligence so that the machine makes the human inspector misidentify the machine as a human. The relationship between the human inspector and the machine is the one testing and the other one tested.

In the sense of UNESCO's mandates to enrich human values in the cultural field and make them sustainable for future generations, a UNESCO-specific issue of AI-generated works is not so much 'how much AI has human-like intelligence' but 'how AI contributes to creating works representing human talents, identities, history and culture'. In the context of the notion of the Common Heritage of Mankind (CHM), how can we examine the role of AI in creating cultural works that are eligible to be preserved for future generations? To UNESCO, the exploration of how AI contributes to creating cultural works representing human identities and history requires the study of how UNESCO can situate AI as a non-human contributor to cultural works. It can identify the distinctive value of AI-generated works that show new ways for humans to express themselves creatively by interacting with AI. The Turing Test is predicated upon an asymmetrical relationship between the human tester and the AI tested. It is less relevant to discuss how to maintain human values and creativity in cultural works generated by AI. A discussion about how humans collaborate with AI to represent socio-cultural contexts and human intentions is more imperative to implementing the mandate of UNESCO. What can be an alternative to the Turing Test?

To overcome the Turing-based understanding of AI as a 'tested' entity, the thesis adopts a different sociological approach to AI and AI-generated work. The discussion of AI-generated works as heritage and the contributions of AI to them requires a particular aspect of the sociology of humans and AI. Durkheim regards an artefact as a collective decision-making system that represents particular human knowledge [4]. Durkheim does not talk about machines in discussing collective representations of human culture. But he argues that a collective representation of a human group is not exclusive to humans. He talks about non-human animals as part of a knowledge system of creating a cultural work. Humans and animals are given "religious character[s]" [6] that collectively symbolise(s) a single identity of the clan. Durkheim highlights that humans incorporate and arrange the roles of animals in creating a representative objective of the human identity, forming a social association with animals. The animals are not the 'first-order human' representative of a human group, but the 'second-order non-human' representative of the human group in the same representative system. Humans decide their roles, depending on how they position non-humans in creating cultural works. Drawing on this Durkheimian idea on the sociality between the human and non-human, the doctoral thesis argues that AI becomes social where its new creation leads humans to extend and re-position human roles from traditional artists. The thesis therefore addresses not only human users' evaluation of AI's responses to human requests but human users' evaluation of themselves. Human devolution of cultural roles to AI and its internalisation into human identity is not a test. It is a contextual and iterative process of setting cultural roles of humans, as the past humans have done in leaving heritage. The thesis therefore creates and names this approach the Durkheim 'Process' to propose the registration of AI-generated works as UNESCO heritage.

4. AI-generated Works: Future Heritage based on Human-AI Creation

The fifteen interviewees from the AI-generated work field in the doctoral thesis shared their experiences, illustrating new types of human creativity and intelligent works coming from the interaction between natural and legal persons and AI systems. The interviewees were categorised into three groups. The first category (Category A) shows the interaction between generative pre-trained AI systems and their human users. The second type of human-AI interaction (Category B) is also concerned with a generative AI system but different from Category A. Unlike users of Category A, practitioners of Category B collect their own data to train a pre-designed AI algorithm by themselves. The pre-designed AI algorithm learns from the data given by its human developer. The last type of human-AI interaction (Category C) is when humans work on the whole cycle of an AI-generated work from data collection to algorithm designs to training algorithms, thus developing their own generative AI systems.

AI-generated works of Category A require skills at using human knowledge on the cultural field into prompt-engineering. AI-generated works of Category B require human decisions on constructing main (story) principles that their AI algorithms follow. AI-generated works of Category C require the orchestration of data curation, rule-making, algorithm design and the assimilation of human knowledge into computer language. These new creative and intellectual works of humans are new ways of expressing human talent and artistic identity.

The fifteen interviewees also shared how they acknowledged the different contributions of AI systems to cultural works, depending on the different human-AI interactions (Category A, B and C). The doctoral thesis characterises two types of contributors in cultural creation. The ‘primary contributor’ creates objects or knowledge by their own capacities and intentions. The ‘secondary contributor’ re-creates and enhances the value of the work of the ‘primary contributor’. The interviewees who used GPT systems (Category A) tended to regard their role as human task creators and AI systems’ role as a task implementer creating content. Those AI-generated works represent what humans want to embody in the form of art. The GPT systems cannot be primary contributors to cultural works because their works are tasked by humans, not by their inherent talent. They can be regarded as secondary contributors if they amplify the artistic ideas of the human user by creating options representing the human user’s artistic ideas.

One of the interviewees of Category B illustrated that his AI system is a writing companion that generates stories based on his’ story frameworks. The AI system complements the rest of the story but cannot be the ultimate author who guides itself about what is right and wrong. The AI system is not a primary contributor because it expresses not its identities but human identities. However, the AI system can be regarded as a secondary contributor that fleshes out a content framework designed by the human programmer. This AI ‘frame-filler’ participates in creating content that the human frame-maker cannot fully expect beyond its story frame.

Some of the interviewees of Category C demonstrated the relationship between the human ingredient provider and the AI cognitive extender. Thomas Whalen [5] uses the term ‘cognisphere’ where humans are embedded into programmable systems such as the internet, wired or wireless data flows.

With an example of the natural language conversational programme 'Alice' he developed, Whalen explains 'cognisphere' as an environment where the human user ultimately interacts with the programme, implementing a particular goal with the software. In this sense, some of the interviewees in Category C perceived that their AI systems created a 'cognisphere' where they make their final choice among a set of AI-generated outcomes they did not fully expect. The AI 'cognitive extender' is not a tool to realise the intention of the human, but 'guides' human intellectual decisions. The AI cognitive extender is not only a secondary contributor that re-creates informative ingredients given by humans. The AI cognitive extender is also a primary contributor to creating the intellectual space of human cultural selection, which humans do not create.

The other interviewees of Category C illustrated the relationship between the human system-maker and the AI art-maker. AI systems perform tasks even their human programmers cannot do for intellectual or cultural reasons. Beyond this, once a human constructs an algorithmic system that generates its own data and art, the AI system can have the highest degree of agency in producing cultural works. This is the point where humans detach themselves from direct art creation. To the human system-maker, the AI system becomes a primary contributor that generates its own artistic source without human intervention. It does not mean that the AI system becomes human. It means that the roles of humans and AI systems are not hierarchical but reciprocal in the context of the Durkheimian idea of distributed roles of humans and non-humans.

The generative performance of their AI systems led some of the interviewees to detach themselves from the art-making and devolve the artistic roles to the AI systems. Both the humans and the non-human AI systems collectively create a cultural work in Durkheim's terms, as a division of intellectual labour. The AI systems have not 'replaced' nor 'mechanised' the humans here. The AI systems created the new roles of the humans, which previous heritage did not contain. Humans exercised both knowledge and craft for the Pyramids. The interviewees exercised knowledge to construct a craft structure for their AI systems, partially or fully devolving art-making role to their AI systems. The fifteen interviews revealed that what decides the roles of humans and AI systems is not an outcome in the sense of the Turing Test but a process in the sense of Durkheim.

All in all, the doctoral thesis demonstrates that AI-generated works have the potential to be (cultural) heritage in UNESCO's terms. An AI system and its following outcomes illustrate situated knowledge and contexts including particular historical backgrounds, know-how, technical grammar and artistic vision of humans and AI. The Pyramids in Egypt, a UNESCO World Heritage site, were built with hammers and its blueprint. They were executed by different humans, the Royals and workers in a human-only distributed knowledge system. It aims to collectively represent an Egyptian Pharaonic civilisation, traditionally believed to have been founded in 3000 BC. On the other hand, an AI-generated work is created by humans and their AI. While humans create the human-AI collaborative systems, the AI in the system creates cognitive spheres for human final choices via iterative processes. It aims to collectively represent and amplify human ideas and cultural intentions at a particular time and space. These new cultural values of AI-generated works can supplement UNESCO's registration systems of World Heritage, Intangible Cultural Heritage and Documentary Heritage.

5. A Proposal for Registering AI-generated Works as UNESCO Heritage

5.1 Five Criteria: The Durkheim Process

Drawing on the analysis of the fifteen interviewees from UNESCO and the fifteen AI practitioners, the doctoral thesis proposes the 'Durkheim Process' comprising of five criteria for evaluating AI-generated works as UNESCO heritage. 'Criterion 1 on Universal Value' consists of: Criterion 1.1: Signifying Important Moments in Human History; Criterion 1.2: Universal Themes and Lessons; and Criterion 1.3: Revealing Unheard History and Voices

The doctoral thesis elaborates on 'Criterion 2 on Creative and Intellectual Works of Humans', by specifying ten questions to assess the value of the following human works:

- Ideation of an AI project(s)
- Objects and procedures setting
- Decision on a type of AI-generated works: prompt-engineering (Category A), the use of pre-designed AI algorithms to process particular datasets (Category B) and a new design of a whole process of creating an AI-generated work (Category C)
- Overall (story) frame-making, operative rule-making
- Input data collection and curation
- Algorithm design and training
- Translation of ideas and knowledge into computer languages
- Pilot test and end product decision
- Value interpretation

The doctoral thesis then specifies 'Criterion 3: Human-human Interaction', by forming eleven questions. Some of them are:

- Have relevant local communities been properly involved in the nominated work?
- Were there discussions for the computer scientists involved in the nominated work to understand the socio-cultural factors of the work?
- Were there discussions for contributors from the field of humanities and social sciences involved in the nominated work to understand the technical factors of the nominated work?
- Who contributed to interpreting the social, cultural and technological value of the nominated work?

The doctoral thesis also concretises 'Criterion 4: Human-AI Interaction', by formulating nine questions. Some of them are:

- Explain why and how certain work was tasked to particular AI system(s).
- Has the AI system performed tasks which its human creators were not able to do?
- Describe the position of the AI system(s) among 'Primary Contributor' or 'Secondary Contributor'
- How does the devolution of the roles to the AI system match the socio-cultural and technological backgrounds and objectives of the nominated work?

Lastly, the doctoral thesis elaborates on 'Criterion 5: Authenticity': Describe step-by-step workflows of the nominated work and any procedural changes that led to the work. What procedural factors of the workflows make the nominated work unique?

5.2 Nominator

UNESCO's three heritage registration systems mentioned above require nominators to include their concrete preservation plans when they nominate their heritage. Nominators should be able to guarantee that they have a strategic plan to preserve their heritage once it is registered on a UNESCO heritage list. Individual researchers or private companies may stop preservation activities at their discretion, prioritising their reduction of cost for maintaining their AI-generated works. But public entities such as research centres, universities and (non) governmental organisations can maintain their preservation activities in a more stable way and in the long term. It indicates the need for a private-public joint nomination of AI-generated works.

The second reason for the need for the private-public pairing is associated with political contexts. Individuals or private companies who wish to nominate AI-generated works need to involve proper public organisations in their joint nomination team. For example, the National Commissions for UNESCO or universities in their countries. Those public organisations can help filter potential politically controversial issues in interpretations of the value of an AI-generated work as UNESCO heritage.

An AI-generated work should therefore be nominated by joint nominators that include both private and public organisations or at least a public organisation. Natural and legal persons who create the nominated AI-generated work can be involved in the joint nomination team for their work. Those creators can launch a nomination and invite relevant scholars and public organisations who can enhance the interpretation of the value of their AI-generated work, in consideration of political contestations with other stakeholders. Among the designer(s) of the algorithm models, the data provider(s), and the user(s) of the nominated AI-generated work, anyone could 'start' a nomination but cannot monopolise the nomination preparation.

5.3 Preservation Plan

The doctoral thesis has analysed the technological nature of AI-generated works and the comments of the interviewees who engage in UNESCO's

Documentary Heritage registration system. It enabled the thesis to enumerate what to preserve for nominated AI-generated works. In summary:

1) Descriptive meta-data: analogue, digitalised and born-digital records on the process and the product of the nominated work and the location of the work
2) Technical meta-data: analogue, digitalised and born-digital records on: how the datasets and algorithms used in the nominated AI-generated work will be preserved; and how particular devices or software that (dis)play the work will be preserved
3) Administrative meta-data: Information about custodians preserving the nominated work and its meta-data
4) Meta-meta-data: how the afore-mentioned metadata will be managed by whom

6. Conclusion

The approach to AI-generated works as hybrid heritage in the doctoral thesis has extended the archaeological and anthropological ideas in the field of heritage, by bringing Science and Technology Studies. John Archibald Wheeler expresses the phrase 'it from bit'. The concept symbolises the idea that every physical item in the world consists of information in origin. In this light, the thesis offers a theoretical framework for identifying the human-algorithmic collaboration that transfers 'it from bit'. The AI practitioner interviewees demonstrated that data associated with human artistic intention and value are algorithmically processed and refined by humans, transferring the information into material paintings and sculptors.

The thesis has also discussed the transition of 'bit from it'. Physical photos and material human life on the street are digitalised into forms of textual, image and audio data, stimulating the human-AI collaboration. This iterative 'it from bit' and 'bit from it' process characterises a human-AI way of expressing human information. It has therefore contributed to examining the changes in the forms of heritage from physical buildings to latent space and neural networks of AI systems. By employing cultural sociology that investigates how new elements (AI in the case of the thesis) are melted into a human identity, the thesis has consequently integrated AI-generated works into UNESCO's concept of the Common Heritage of Mankind, constructing the idea of the human-algorithmic Common Heritage of Mankind.

References

[1] Droit, R. P. (2005) *Humanity in the Ma king: Overview of the Intellectual History of UNESCO, 1945- 2005*. UNESCO.
[2] UNESCO. (1945). *Constitution*. UNESCO.
[3] Turing, A. M. (1950). Computing mach inery and intelligence. Mind, LIX(236), 43 3-460.
[4] Durkheim, É. (1976). *The Elementary Forms of the Religious Life* (2nd ed.) (Swain, J. W., Trans.). Allen and Unwin.
[5] Whalen, T. (2000). *Data Navigation, Architectures of Knowledge* [Conference presentation]. In Cook, S & Diamond, S. (Eds.).(2011). *Euphoria & Dystopia - The*

Banff New Media Institute Dialogues (pp. 99-106). Banff Centre Press, Riverside Architectural Press.

Lexiconia – A Lexicon of Generative AI

Assistant Prof. Chad Eby, MFA
School of Art and Visual Studies, University of Kentucky, Lexington, KY, U.S.A.
chad.eby@uky.edu

Dr Christopher Fry, BA(Hons), MA, PhD, FHEA
Westminster School of Art, University of Westminster, UK
c.fry@westminster.ac.uk
http://www.lexiconia.art


Abstract

This paper describes and offers a rationale for 'Lexiconia', a collaborative project with the aim of publishing and maintaining an online 'lexicon of generative AI'. The purpose of the lexicon is to provide understandings and perspectives on the key concepts of generative AI by gathering critical responses to and reflections on practice. The project is particularly interested in the role of creative practice in critiquing as well as normalising emerging generative AI technologies. It aims to provide a space to examine some of these concepts more closely, and question many of the assumptions that underpin them.

Introduction

'Lexiconia' is a nascent collaborative project with the aim of publishing and maintaining an online collection of submitted texts to form a 'lexicon of generative AI'. This aims not only to explain the terms and concepts of generative AI, but also importantly to capture current practice. The project invites texts of between 800 and 2000 words on topics that might include: Large Language Models, Zero-shot Learning, Latent Space, Transformer Architectures, Attention Mechanisms and Autoencoders.

Generative AI has brought many new terms as well as adopting existing ones from different disciplines. Often with roots in Computer Science and Software Engineering vocabularies, in use they sit between technical and creative languages. They are often as esoteric as they are explanatory, part marketing jargon and part magical language.

The speed with which phrases such as 'text-to-image' and 'Neural Network' have entered everyday use hints at an accelerated normalisation. AI has been seamlessly incorporated into existing services and creative tools, their use often unnoticed and unsolicited.

This project aims to respond to the fast pace of developments by capturing and sharing experiences of generative AI as it is understood, applied, critiqued, and incorporated into creative practices.

A complete set of definitions would be impossible as well as far less interesting and useful than the myriad ways in which the emerging concepts, tools, terms and language of generative AI are changing the creative landscape. This lexicon is particularly interested in the role of creative practice in critiquing as well as normalising emerging generative AI technologies. It aims to provide a space to examine some of these concepts more closely and question the assumptions that underpin them.

What is Lexiconia?

Rather than 'definitions', Lexiconia aims to provide perspectives on the key concepts of generative AI practices by gathering critical responses and reflections contextualised by practice. The intention is for contributions to be relatively concise, providing an essential understanding of the chosen subject, but also present a particular perspective or position in relation to creative practice and the cultural landscape. In this respect it is not dissimilar to online publications including journals, blogs and the websites of creative practitioners. However, the lexicon places an emphasis on the vocabulary created by and employed in the description of generative AI. These terms and concepts, which are often as evocative as they are technical, act as a jumping off point for exploring different critical interpretations and perspectives. The discursive nature of AI, which often sees the transfer of terms from one domain to another, creates slippages between definitions and usages. It is these spaces, ambiguities and differing perspectives that the project intends to draw on.

The term 'lexicon' is used here in the sense applied by Matthew Fuller in their 'Software Studies: A Lexicon' published by MIT Press [1]. In the introduction to the work Fuller provides the following definition:

"A lexicon is a vocabulary of terms used in a particular subject. Rather than an encyclopedia, which is too universal, or a dictionary or glossary, which offer too short descriptions or readings of terms, a lexicon can be provisional and is scalable enough a form to adapt to any number of terms and lengths of text." [1]

Fuller's lexicon is a collection of written texts by key figures in software studies and includes entries on subjects such as 'analog', 'button', 'code', 'glitch' and 'elegance'. The format of the lexicon, Fuller argues, makes a virtue of being open ended, incomplete and evolving, describing it as collected "speculative, expository, and critical texts on particular digital objects, languages, and logical structures" [1]. Fuller also hints at the need for an approach that can keep pace with developments describing the project as "an exercise in the rapid prototyping of transversal and critical approaches to such stuff." [1] Lexiconia intends to be similarly provisional, scalable and able to respond to rapid developments.

Following the initial call, entries have been received on topics including Prompt Engineering, Latent Space, the Pandora's Box metaphor, the Digital Sublime, Art in the Age of Generative Production, Variable Autoencoders, and Robopoetics.

Moving from future horizons to the present

Documenting and sharing arts practice, including and especially new media arts, is vitally important and a well-established norm. Websites and journals such as Rhizome and Leonardo, and conferences such as ISEA, Transmediale and the Generative Art conference, have a long history of disseminating practice and research in the fields of arts, science and technology including generative AI. Lexiconia aims to complement these as well as addressing an issue particular to generative AI, namely the speed of its development.

The speed at which AI and generative AI technology is being developed can shift focus to the near future rather than the present [2]. This, combined with the ease of use and seamless integration of AI into existing software and tools creates a situation whereby the adoption of AI is obscured despite or perhaps because of its ubiquity. As a practitioner it can be difficult to reflect on one's use of generative AI when the tools themselves evolve so quickly. Meanwhile discourse is often focused on the near future. The promises of the next iteration of AI tools distract attention from a critique of current flaws and issues. What we may learn by engagement with generative AI can seem to become out of date very quickly. This reinforces a narrative in which the implications for future technologies seem more relevant than present experiences.

Natale and Ballamore note that this projection into the future has long been a feature of the discourse surrounding AI, tracing back through popular science literature to the middle of the 20th century [3]. Debates seem to move "from the horizon of the present to the horizon of the future", contributing to the proliferation of 'myths' and grand claims about AI. Natale and Ballamore suggest a similar situation persists and that "the imagination of future horizons and developments still characterizes contemporary AI myths" [3].

Focusing attention on the present moment is especially important given the comparatively nascent stage of development. AI and generative AI have a long history, far longer than is typically realised or acknowledged in wider public discourse. However, it could be argued that the adoption of generative AI tools by the general public and creative industries more widely has experienced a recent explosion [4]. As such the present moment may be crucial in the development and integration of generative AI into wider creative arts practice. As Christoph Ernst notes "it makes a crucial difference if we talk about media in their conceptual stage, realized state or vanishing state" [5]. Ernst adds, "when it comes to newly evolving media, artistic practices are of particular theoretical importance. They allow the production of new (re)configurations between imagination and media". [5] It is these stages of development where change is more possible, practices are still emerging and understandings of key concepts still being negotiated.

While generative AI may not be considered new, it has perhaps not yet become ordinary, rendered invisible by what Vilem Flusser might have described as habit and convention's "gray and trivializing cover" [6]. That is not to suggest that current practice is not currently shared. Sites and online galleries such as Civitai.com contain thousands of regularly submitted AI generated images, models, LoRAs, prompts, fora, tutorials and 'bounties', appealing for particular images or LoRAs [7]. Civitai.com also hosts articles allowing practitioners to post guides to specific tools or processes and 'musings' which offer more reflective texts.

Galleries of AI work can also be found on sub-Reddits such as r/aiArt which has 384k members [8]. Other sub-Reddits allow discussion of specific platforms including r/StableDiffusion and r/midjourney with 575k and 1.5m members respectively. Meanwhile sub-Reddits such as r/aiwars and r/DefendingAIArt are home to lively debates although as their names suggest, these can often be highly polarised, acting as a flashpoint for pro and anti AI art debate.

Lexiconia aims to address some of the challenges presented by the fast pace of AI development while also being a space for critical reflection. The short texts and online format with a rolling call for submissions aims to provide a relatively quick means of sharing ideas, observations and practices. Rather than looking to the near future, it places emphasis on current practice that has the capacity to inform future developments.

The role of arts practice in critiquing and normalising technology

At the centre of the project is creative practice. Not only the sharing of practice but also practice as a way of critiquing generative AI. There is clear precedent for this in the way that technologies such as the World Wide Web, mobile phones, Wi-Fi and VR have been explored by artists, bringing them to the wider public but also challenging our understandings and expectations of them. Examples would include the early net.art practices of JODI among others, or the Pervasive and Locative Arts Network that emerged in response to the development of mobile and wireless technologies and brought together "artists, activists and people from a variety of backgrounds" to explore a "break away from traditional computer interfaces." [9]

In many respects, new media arts practice has evolved from and exists as a critical examination and exploration of creative technologies. It has also see the establishment of what might be called 'Alternative' and 'Activist' New Media [10]. Organisations such as Furtherfield [11], Rhizome [12] and the Institute of Network Cultures [13] all examine the relationships between the arts, emerging technologies and the role of critical practice.

A recent report for the European Union titled 'The Role of the Arts in the Digital Transformation' sets out a role for the arts in shaping the public's understanding of emerging technologies such as AI [14]. This includes fostering 'critical awareness' as well as a 'healing' or 'therapeutic' effect. The report sets out a wide-ranging reflection on the ways in which the arts might act as 'transformational' agents including how they can support a critical attitude towards technologies and foster enhanced 'digital citizenship'. This recognises the impact and role of arts practice in mediating relationships with technology. However, the artist as 'social worker', 'therapist' or shaper of 'digital citizenship' raises questions about how this role may be simply normalising relationships with emerging technologies.

As Claire Bishop notes, such a 'social' or 'ethical' turn is not a requirement or obligation of the arts [15]. Not all arts practice needs to support what may be a flawed notion of some wider public good. However, without critique there lies the danger of technologies developing unchecked and according to the agendas of the technology companies producing them.

There is a tension between the rapid adoption of AI tools by artists and the creative industries as a whole and the space and time available in which to critique them. As the EU report acknowledges, arts practice plays a vital role in innovation, imagining new and often unintended applications for emerging technologies. However, they can also seemingly be tasked with fostering positive relationships with technologies or to find applications that will serve the public good.

This project aims to offer a space to explore the possibilities of AI as well as for pushing back against and questioning dominant discourses and assumptions. It aims to promote a perspective of AI technology that is more open to ambiguity, negotiated meanings and contradictory perspectives, and less inclined to accept the inevitability of dominant discourses that may bypass current 'flaws' to focus on the promises of the future. It aims to draw on the dual nature of the lexicon as both a source of knowledge and a site for critical propositions open to reinterpretation and negotiation.

The spaces between language and technology

If speed can be seen as one defining characteristic of generative AI then another might be its ability to generate 'technical terms' including many that might equally and less sympathetically be described as jargon. Mark Coeckelbergh notes the "rich conceptual space" between language and technology, a space potentially under investigated by approaches that take either an empirical or more technical, engineering approach [16].

The rapid development of AI seems to have required a proliferation of terms although as Philip Agre points out there can be a gap between the term and the realisation of what it is intended to describe [17]. Agre describes how “Innovations frequently involve techniques that bring new vocabulary into the field. Whether the resulting systems really are exhibiting these qualities is hard to say” [17].

Generative AI carries the dual danger of hyper-mystification and generalisation. Mystification occurs due to the perceived complexity of the technology amplified by the use of new or technical language together with the ‘black box’ nature of our interactions with them, limited as they typically are to the use of Graphical User Interfaces. Meanwhile, ‘Artificial Intelligence’ and ‘AI’ have become umbrella terms encapsulating a wide range of technologies and concepts, applied almost as a marketing term to imbue a certain ‘magical’ quality. The descriptions of AI artworks often make specific mention of the use of AI, framing the way they are received and by turn implying their significance.

As a combination of science and engineering, AI has always traversed disciplines. Natale and Ballatore argue this led to “discursive shifts by which concepts migrated from different contexts through analogical arguments, carrying with them their own cultural associations and meanings, and often resulting in misleading cross-domain translations.” [3] Agre describes the “dual character of AI terminology [...] the vernacular and formal faces that each technical term presents” [17]. This "strategic vagueness of AI vocabulary” allows terms to traverse domains. There is a gap between the vagueness and discursive nature of terms and the “precise mathematical specifications” that allow them to be described in a computer program [17]. Lexiconia does not aim to reconcile these dual faces but rather sees it as an opportunity to show how ideas can traverse between the specific and the abstract and between the technical and the intuitive. These are the spaces in which the practitioners of generative AI find themselves and the terrain they must negotiate. Agre describes the “borderlands between AI and its application domains” whose inhabitants include those impacted by generative AI [17]. This includes the creative practitioners, photographers, designers, illustrators and others at the forefront of generative AI.

Describing practice can sometimes mean using alternatives to more technical terms in favour of language that seems to match our experiences more closely or that is more evocative. An example would be a recent reimagining of 'prompt engineering’ as ‘spells’. *Spells,* a crowd-funded book published by RAM.studio in Berlin collects 250 ‘high-quality’ prompts for the Midjourney platform gathered from artists and designers and inhabits a space somewhere between a practical cookbook and a coffee table art book [18]. Both ‘spells’ and ‘prompts’ might be used to describe the same thing and yet one suggests a very different kind of activity to the other.

A key inspiration for Lexiconia, and one also noted by Fuller, is the 'Jargon files’ by Eric S. Raymond and Guy L. Steele. This is described as “a comprehensive lexicon and cultural resource reflecting hacker slang, written in the late 20th century.” [19] Far from a technical manual or a historical record, the text captures both the technical and the cultural, “documenting the often playful and inventive terms used within.” [19] Similarly, it is hoped that the Lexiconia lexicon will capture the human as much as the technical. Considering both when describing generative AI practice may be essential given a shift from the use of tools and machines to more ambiguous workflows.

‘Prompt engineering’ shows how using AI often necessitates a more experimental and intuitive approach. Rather than the replication of analog processes and interaction with icons representing their analog predecessors, AI requires a different attitude. The nature of generative AI and the rate at which it develops means that the same prompts will not always return the same results, even with otherwise identical initial conditions. Prompt engineering is more akin to negotiation than the use of a tool [2]. It is shaped more by ideas and articulating desire through language – or by consciously abandoning specific desires in the service of structural explorations – than to specific steps that can be easily recorded and shared. This suggests that rather than detailed accounts of specific methods and uses of equipment that may have been useful for prior technologies, what is needed are accounts of the intentions and process that are more transferable and so relevant to other practitioners.

Next steps

The Lexiconia project aims to bridge the gap between the role of academic publishing provided by journals and annual conferences such as Generative Art and other faster paced but often less reflective online sites and forums. By inviting shorter texts that foreground practice, it aims to provide a space to present ideas, observations and approaches as they emerge and that may ultimately develop into other publishing opportunities. In many respects it proposes to bring together knowledge and material that already exists, albeit often distributed across individuals’ blogs, online articles and forums. It is hoped that the lexicon will provide a focus to pull this valuable knowledge together, making it more accessible and visible as well as becoming a more permeant record.

The project emphasises arts and arts practice, however these terms are taken in their broadest sense and doesn't presuppose a particular set of practices or disciplines. It might simply be taken to mean anyone making something using generative AI. It is also important to note that the project is not seen as a lexicological or etymological study but would value the contribution of those working in these fields.
Lexiconia is still in the early stages of development, although for the reasons outlined above it aims to gather momentum quickly. It is expected that the process, format and structures will evolve in response to use and with feedback from the practitioners that we hope will contribute to and use it.

References

[1] Fuller, M. (ed.) (2008) ‘Software Studies: A Lexicon’, 1st edn. Cambridge:

The MIT Press, available at: https://doi.org/10.7551/mitpress/9780262062749.001.0001 [accessed 5th November 2024]

[2] Fry, C. (2023) ‘Commands, Prompts and Gestures: Understanding Generative AI as an apparatus’, XXVI Generative Art Conference, available at: https://generativeart.com/GA2023/papersDOC/OK/56_C_Fry_Corrected.pdf [accessed 5th November 2024]

[3] Natale, S. and Ballatore, A. (2020) ‘Imagining the thinking machine: Technological myths and the rise of artificial intelligence’, *Convergence (London, England)*, 26(1), pp. 3–18, available at: https://doi.org/10.1177/1354856517715164 [accessed 5th November 2024]

[4] McKinsey & Company (2023) ‘The state of AI in 2023: Generative AI’s breakout year’, available at: https://www.mckinsey.com/capabilities/quantumblack/our-insights/the-state-of-ai-in-2023-generative-AIs-breakout-year [accessed 6th November 2024]
[5] Ernst, C. (2018) ‘Imagination and Techno-Imagination – Preliminary Remarks Towards a Theory of Imagination as Media Practice’, available at*:* https://bonndoc.ulb.uni-bonn.de/xmlui/bitstream/handle/20.500.11811/1144/Christoph%20Ernst%2C%20Imagination%20and%20Techno-Imagination%20%5BWorking%20Paper%5D.pdf?sequence=1&isAllowed=y [accessed 5th November 2024]

[6] Flusser, V. (2023) ‘Communicology: Mutations in Human Relations’, tr. Novaes R., Stanford University Press

[7] Civitai (2024) https://civitai.com/ [accessed 5th November 2024]

[8] Reddit (2024) https://www.reddit.com/ [accessed 5th November 2024]

[9] Transmediale (2005) ‘PLAN - The Pervasive and Locative Arts Network’, available at: https://archive.transmediale.de/content/plan-the-pervasive-and-locative-arts-network [accessed November 5th 2024]

[10] Lievrouw, L. A. (2011) ‘Alternative and Activist New Media’, Polity

[11] Furtherfield (2024), https://www.furtherfield.org/ [accessed 5th November 2024]

[12] Rhizome (2024) https://rhizome.org/ [accessed 5th November 2024]

[13] Institute of Network Cultures (2024) https://networkcultures.org/ [accessed 5th November 2024]

[14] Alacovska, A., Booth, P. and Fieseler, C. (2020) ‘The Role of the Arts in the Digital Transformation. Artsformation Report Series’, available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3715612 [accessed 5th November 2024]

[15] Bishop, C. (2006) *‘THE SOCIAL TURN: COLLABORATION AND ITS DISCONTENTS’,* in *Artforum international*. New York: Artforum Inc, pp. 178

[16] Coeckelbergh, M. (2017) ‘Language and Technology: maps, bridges, and

pathways', in *AI & Society*, 32:175-189
DOI 10.1007/s001-46-015-0604-9

[17] Agre, P.E. (1997) 'Toward a Critical Technical Practice: Lessons Learned in Trying to Reform AI', in *Social Science, Technical Systems, and Cooperative Work*. 1st edn. Psychology Press, pp. 131–157. Available at: https://doi.org/10.4324/9781315805849-8 [accessed 5th November 2024]

[18] Kuwertz, M. (2024) 'Spells', RAM.studio

[19] Raymond, E. S. and Steele, G. L. (2024) 'The Jargon File, Version 4.2.2, 20 Aug 2000', available at: https://www.gutenberg.org/ebooks/3008 [accessed November 1st 2024]

Building a Web Site to Play Back Generative Animations

Prof. T. Unemi, DEng.
Glycan and Life Systems Integration Centre, Soka University, Hachioji, Japan
www.intlab.soka.ac.jp
e-mail: unemi@soka.ac.jp

Abstract

The advancement of computer graphics (CG) technologies has led to the standardization of programming techniques for graphical presentation within a basic framework of OpenGL. The initial concept was to develop a software library to assist developers. WebGL represents a set of application programming interfaces (APIs) for JavaScript that are compatible with OpenGL ES. It facilitates the rendering of computer-generated content on web pages via the client-side graphical processing unit (GPU).

Audiovisual media data can be loaded from both local and remote storage by the web browser. However, recently some combinations between local files and automatic playback of sound media are prohibited for security reasons. To cope with these restrictions, it is effective to unify the shader codes in the HTML code and to prepare a playback button on the screen. It provides an appropriate interface for viewers even the sound files are stored in a local or remote storage.

One of the methods for generative audio-visual arts is an evolutionary approach to automatically develop a functional expression that maps the XY coordinate of pixel to colour value. It is possible to compile the result expression to the *shader* code runnable on the GPU. By organising a set codes in HTML, CSS, JavaScript, and shader code, the result moving image can be drawn in realtime on the web browser.

The development of CG technology is continuing further, but the strategy is branched into two separated directions, Vulkan and Metal. WebGL is also a target of revision to adapt to such directions. We need to watch the ongoing development of WebGPU to keep supporting the Web based display for generative art.

1. Introduction

One of the most popular methods to realise a piece of generative visual art is to use shader programming that displays a still and moving image on the computer screen. Through the improvement of computer graphics technologies, the programming technique of this field has been standardised in a basic framework of OpenGL [1]. Nowadays, many higher layer programming environments are available not only as software libraries, such as glut, but also as wrapped programming frameworks in scripting languages, such as python and JavaScript.

The following part of this paper describes a summary of the history of programming in computer graphics, graphic processing unit (GPU) and shader programming, and the author's approaches to allow shader art to be displayed in web browsers using WebGL [2]. The web browsers have ability to render a media file, such as the most typical form of HTML, but also images, sounds and movies. The media data can be loaded from both local and remote storage. However, recently some combinations between local files and automatic playback of sound media are prohibited for security reasons. A software design is presented that copes with these restrictions in order to provide an appropriate interface for viewers through either local or remote storage.

This paper concludes with some future perspectives on the technological change in both hardware and software for computer graphics from the point of view of generative arts.

2. History of CG

The techniques of graphic representation as an output from the digital computer has been improved not only by the power of computation but also by the innovation of peripheral equipments. In 1950's the character printer was the only facility to make a pattern in 2D plane. In late 1960's, XY-plotter and storage-type stroke graphic display were introduced to draw line strokes on the paper or screen. The raster scan display by the cathode-ray tube (CRT) also got popular but the resolution was not fine enough comparing with the printing. The technologies of liquid-crystal display (LCD) was also developed in 1960's, but it had been used in a fixed shape and arrangement such as the seven segment display for a digit.

The world wide popularisation of consumer TV set in home push the innovation of analog CRT technology, but it also raised an issue on overwhelming consumption of energy because of the high voltage electricity for both electron beam and magnetic field. Plasma display, organic EL, and LCD are helpful to reduce the electricity consumption and to transit the public broadcasting system to the world standard version of high definition resolution by a digital signal. The transition was completed in 2010's in major areas in the world.

The industrial standard of computer monitor was also established in those years as 1920x1080 Full HD, RGB 8bits components, and 30 Hz for frame rate. The hardware to generate the video signal for such raster scan display usually requires a memory map of screen image called frame buffer. Therefore the functionality of the software to draw a CG image has to transform a model of 3D objects to the colour distribution on the pixel lattice represented on the digital memory.

Based on this framework, a lot of efforts has been done to develop computing techniques, such as geometry calculation to project a 3D model to a 2D image, simulation of lighting effects including diffusion, reflection, fading, and so on. One of the popular frameworks for the software development is OpenGL that has been used widely across a variety of hardware platforms. Another framework such as Direct3D [3] is also based on a similar concept because the hardware architecture is shared among them.

To make a development of application software easier, a number of different support systems have been developed. One of the common approach is to provide a set of modules for a type of functional units that is called via application program interfaces (APIs). As it is not necessary for all applications to have a CG rendering functionality, such models are optional to be attached to the application module. For a programming language to be compiled into a sequence of machine instruction, such as C and C++, the function *library* is a useful style to organise the modules. For a scripting language, such as Python, an optional modules dynamically loaded into the interpreter system is available. In the case of JavaScript, as the source code is assumed to interpreted by a web browser, the method of API implementation depends on the policy of the browser's architecture. WebGL is the specification of APIs functionally compatible with OpenGL ES, a subset of OpenGL assumed to be embedded in an application code in C language. Due to the difference between these languages, it cannot avoid the difference in the syntax, but it is easy to transform the code segments from C to JavaScript.

3. GPU and Shader Language

A lot of innovations are connected each other to reach the up to date technologies of CG, computation power, memory capacity, communication bandwidth, flat panel display of the resolution from 2K to 8K, from 8 to 16 bits colour resolution for each of RGB component, and 60 Hz or higher frame rate. Silicon Graphics Inc. [4] founded by James Clark was the pioneer of hardware and software development in this field. The current hardware design of 3D graphics inherited the basic design of rendering pipeline from 3D geometry model to 2D image represented on the memory of frame buffer.

As we can recognise the nature as a massively parallelised process from a view point of information flow, it is also natural to design the calculation for graphics as a parallel computing. The 3D model is constructed as a set of points, edges, and surfaces with 3D coordinate system in Euclidean geometry. The task to map these positions to the 2D position projected on the frame buffer can be calculated in parallel using leaner algebra of vectors and matrices. The computation to specify the colour value for each pixel considering the condition of lights and material is also the object of parallel processing. The hardware design of parallel architecture is also helpful to reduce the throughput time. It is easy to understand that the traditional style of sequential computation is useless for high resolution animation as it requires 30 times of calculation in a second for 2 million pixels at least, it would be necessary to reduce the computation time for each pixel to shorter than 16.67 nano second.

Currently popular GPU for the personal computer usually has some hundreds of processing elements. This number reaches some thousands in the high-end version. The compiled code for GPU is different from the code for CPU, and a program object is assumed to be applied to each elemental data of vertices and pixels. The code modules are separated into the vertex shader and fragment shader for each type of element, and communicating in a pipeline process. For 2D image generation, the vertex shader does not need to transform a coordinate from the model space to the pixel space. This means the code of the vertex shader is simple and fixed, but the fragment shader becomes complicated somehow, and it takes main role for the image generation.

The programming language for GPU used in OpenGL is OpenGL Shading Language, GLSL [5]. The basic syntactic structure and the elements inherit the features in C language, but it has some specific syntax to express a vector and matrix. For the representation of coordinate in 3D space needs a vector of three numerical elements for x, y and z coordinates. The colour value for a pixel also has to have three values for red, green and blue components. It is often requested one additional component expressing an opacity for an image data as a source material. Therefore, a vector of floating point numbers from two to four elements is frequently appears in the code. For convenience of coding, GLSL supports a flexible syntax to express such vectors in a similar manner but beyond a structure in C language.

4. Generating Graphics

SBArt4 is an application software running on MacOS developed by the author that allows the user to breed an abstract CG images and movies utilising a technique called Simulated Breeding [6, 7]. It also has a functionality to automatically generate unique images and movies by an evolutionary computation using a type of computational aesthetic measures to evaluate the fitness value for each individual [8]. The genotype is in a form of functional expression that maps XY coordinate of pixel and time value T of the frame for moving image to HSB colour value. The all of intermediate values are also vectors of three scalar value elements. Some of the functions that request restricted ranges of argument values, such as division and logarithm, the definitions are modified so as to avoid the error. The evolutionary operations of mutation and crossover are in a manner of genetic programming [9].

Figure 1 shows a sample image and the tree structure of genotype. The essential part of compiled code in the code for fragment shader is as follows. This part works together with fixed code for vertex shader, sub-functions for operators, and transformation code from HSB to RGB colour space.

—-

```
vec3
v=(divide(sin(((vec3(p.x,t,p.y))+(sin
((vec3(t,p.y,p.x))*3.1415927)))
*3.1415927),(hypot(cnst(u,vec3(
1.246575,-1.917661,-
1.058921)),vec3(p.x,t,p.y)))+((c2p(ab
s(cnst(u,vec3(
-1.055952,-
1.267934,0.467390)))))+((logA(abs(vec
3(t,p.y,p.x))))+(rotate(cnst(u,vec3(0
.790060,1.024423,0.397812)),(-
(exp(cnst(u,vec3(
-1.507477,-1.706989,0.434056)))))-
(cos((vec3(p.y,p.x,t))*3.1415927)))))
)))*vec3(2.378,1.098,1.663);
```

—-

5. Security Issues

The web technology has become essential in our everyday life nowadays. It is useful to exchange a variety of information between people and organisations for private communication, commercial promotions, business contracts, public service, and so on. One of the advantage is the capability to handle multimedia data, texts, photos, illustrations, moving images, speeches, musics, and so on.

As a dark side of such convenient functionalities, it is also the fact that it helps criminal acts such as frauds and spreading lies. To protect the users from such security issues, The web browsers often has some types of restrictions to access and to show the information on the internet. Here are two points to be concerned for the topic of this paper.

Not only the protection of secret information such as the login password, it should also prevent for leaking any information stored in the local storage of the users personal machine. Web browsers have a capability to refer to such local files by specifying the path name in the Universal Resource Locator (URL) with “file” as the protocol part. This functionality is useful for the developer to confirm how the browser displays the source code. However, this brings a potential risk of information leaks from the local storage, because any content of URL can be specified in the remote

HTML code that reads the content to send it to any remote servers [10]. To avoid such risk, the latest popular browsers don't allow indirect access to any local text file in default. This restriction disables working the code set in local if the shader codes are stored in independent files separately. From a view point of efficiency of both maintenance and storage size, it is better to store the shader code separately from HTML. It is possible to load the shader code as a content of hidden inner frame by a JavaScript code if files are placed in the same server with the HTML file. For the exhibition in a public space, it has to make the piece workable without the internet connection if the network is unavailable at the venue. One of the way to escape from this restriction is to embed the shader codes in the HTML or JavaScript code. HTML's "script" tag is useful for this purpose. In case we need more than one shader codes to draw CG images in turn, ID attribute attached to the script tag is helpful to indicate which code should be used for the next image.

The other point is on the playback of the sounds. It might not severe issue but it must be uncomfortable for the user if the browser makes an unpleasant sound immediately after the user access a web site. Unexpected words or warning sounds might lead the user's confusion and his/her mistake on the next operation. The "audio" object of HTML has an attribute named "autoplay" to indicate that the sound should be played back after the data is loaded as soon as possible. But the recent versions of popular web browser suppresses this functionality in default to avoid such disgusting user experience. Instead, it is recommended to start sound playbacks triggered by a user's operation such as click a play button on the screen [11]. A demand to start playing sounds is allowed to be called from a process of event handling such as a JavaScript code fragment specified in "onclick" attribute of "button" tag.

The sample web page is available from the following URL:

https://unemi.github.io/SBArt4Players/20240827/

A playback button appears on the browser's screen when you access the above web site. The user can start enjoying the moving image with sound effect by clicking this button. The image is displayed on a full screen size, and the rendering fineness adapts to the monitor's resolution as GPU calculates the colour value for each pixel. The recent personal computers, tablets, and even smart phones have no problem to play it back, but the animation frame rate would get unacceptably slower if the GPU is not powerful enough for the screen resolution.

The source codes are available from the following URL.

https://github.com/unemi/unemi.github.io/tree/main/SBArt4Players

6. Future issues

Much of the development efforts are conducted to improve the functionality of CG development. However, an undesirable issue recently occurred that the strategy was branched into two separated directions, Vulkan [12] and Metal [13]. The difference is mainly on the hardware architecture of the relation among CPU, GPU, and memory. The former one assumes that CPU and GPU are separated and they have their own memory for each. On the other hand, the later one intends to use unified memory shared between CPU and GPU. The shading languages are also redesigned separately with different policies. WebGL is also a target of revision to adapt to such directions. We need to watch the ongoing development of WebGPU [14, 15] to keep supporting the Web based display for generative art.

References

[1] Khronos group, OpenGL: https://www.opengl.org/

[2] MDN Web Docs, WebGL: 2D and 3D graphics for the web, https://developer.mozilla.org/en-US/docs/Web/API/WebGL_API

[3] Microsoft, Direct3D, https://learn.microsoft.com/en-us/windows/win32/direct3d

[4] Wikipedia, Silicon_Graphics, https://en.wikipedia.org/wiki/Silicon_Graphics

[5] Khronos group, Core Language (GLSL), https://www.khronos.org/opengl/wiki/Core_Language_(GLSL)

[6] Unemi, T. (1998) A Design of Multi-Field User Interface for Simulated Breeding, Proceedings of the third Asian Fuzzy Systems Symposium, Korean Fuzzy Logic and Intelligent Systems Society, pp. 489–494, Masan, Korea.

[7] Unemi, T. (2002) SBART 2.4: an IEC Tool for Creating 2D Images, Movies, and Collage, Leonardo, Vol. 35, No. 2, pp. 171, 189–191, MIT Press.

[8] Unemi, T. (2014) Automated Daily Production of Evolutionary Audio Visual Art – An Experimental Practice, Proceedings of the Fifth International Conference on Computational Creativity, June 9–13, Ljubljana, Slovenia, Session 2-2.

[9] Koza, J. (1992) Genetic Programming On the Programming of Computers by Means of Natural Selection, the MIT Press.

[10] Invicti (2024) Local file inclusion, https://www.invicti.com/learn/local-file-inclusion-lfi/

[11] MDN Web Docs, Autoplay guide for media and Web Audio APIs, https://developer.mozilla.org/en-US/docs/Web/Media/Autoplay_guide

[12] Khronos group, Vulkan, https://www.vulkan.org/

[13] Apple Co., Metal overview, https://developer.apple.com/metal/

[14] WebGPU Fundamentals, https://webgpufundamentals.org/

[15] MDN Web Docs, WebGPU API, https://developer.mozilla.org/en-US/docs/Web/API/WebGPU_API

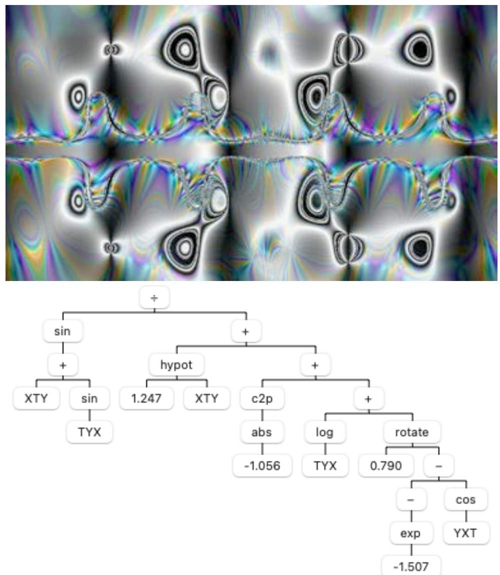


Figure 1. A sample image generated by SBArt4 and the genotype drawn as a tree structure.

Make Room for Nature!
Personal Narratives Co-Created with Generative AI

Cristina M. Reis
CENSE – Center for Environmental and Sustainability Research, NOVA School of Science and Technology, NOVA University Lisbon, Caparica, Portugal
e-mail: cm.reis@campus.fct.unl.pt

Nuno Barreiro
Department of Electronic Engineering, School of Engineering, Physical and Mathematical Sciences, Royal Holloway University of London, Egham, England

Abstract

Curiosity and exploration are essential in a successful learning path, and benefit from dedicated and personalized learning practices, which can be supported by generative machine learning. The mere presence in a specific environment, affordable to the senses and subject to interpretation, would constitute an experience no longer solely relying on a scient human tutor (nurture) or exceptional attention skills (nature). Artificial intelligence can encourage exploration, favor contemplation, nourish curiosity, and provide novel ways to experience knowledge.

Previous research indicates that, through a closer observation and experience of Nature's narrative, extended reality technologies can contribute to an improved understanding of the environment, favoring the development of an eco-agency – the understanding of the capacity to produce a conscious effect on the environment, by decision, intervention, or action. This ubiquitous digital presence, superimposing virtual layers of information on the real world, supports and informs users on their individual journeys, and can be designed to be both inclusive and accessible.

Machine learning generative tools provide immediate access to an immense pool of information, democratizing and stimulating the creation of personal narratives. Articulating this possibility with the development of a sense of agency in autobiographical memories, where narrative expression plays a crucial role, we propose bonding one's own narrative

with the surrounding environment, fostering eco-agency. The process we use to retain snippets of experience is non-linear, which brings questions about memory generation and organization, recollection adjustment and reorganization, all fundamental to the construction of our narrative identity.

By establishing a parallel between previous technological artistic disruptions, such as photography, and current perceptions of artificial intelligence in creative contexts, we focus on how generative processes feed into the curation and recollection of personal experiences. Via the design of a case-study where users are reminded of their presence and effect in a simple environment, we explore the concept of eco-agency, touching the realms of identity and narrative, and their flow between facts and fiction. Relying on new tools, both to revisit and visualize the past, we aim at addressing how we can also contribute with stories for the future, at community level.

1. Preamble

Generative AI (GenAI) is here to stay. Beyond the debates around naming (is it intelligent?), copyright (is it original?) and purpose (is it evil?), one can only acknowledge that Large Language Models (LLMs) seem to have popped out of the blue: in 2023, ChatGPT reached 100 million users in two months, the fastest-growing consumer software application in history [1]. The ability to generate new content from prompts has taken the world by storm.

As GenAI tools arise, movements pro and against started flourishing. It is only human to be hauled into heated manifests that often resist facts and new findings, or miss relevant points over exciting new applications. From a technical point-of-view, and according to the specialists, GenAI models live inside black boxes, as we lack mathematical theories to describe the underlying mechanisms of the generation process [2]. This imbues the public opinion with a mix of fear, respect and excitement (in varying proportions and degrees), and leads to over-polarized debates that often exclude crucial information and promote biased opinions. Furthermore, beyond the socio-political realm, these conversations are also driven by strong economic interests and agendas, in a market projected to reach US$36.06bn in 2024, and US$356.10bn by 2030 [3]. These debates should be protected, at all levels (e.g., UNESCO, social media platforms), as they output relevant guidelines, regulations and legislations [4-7], recommending and/or reinforcing specific features and aspects of practical applications (e.g., ethics, copyright, democratization, security, accessibility).

All these recent fast-paced developments have led to applications in sectors such as health, entertainment, research and education. In the context of this paper, we are particularly interested in GenAI multimodal outputs in the categories of text, image, video, audio, 3D models, and interactive media; and our goal is to obtain those outputs via a co-creation process with Nature [8] where GenAI can be viewed as a mediator/facilitator. This collaborative idea could be an instant or a clip, using an analogic photographic camera or an old acoustic guitar, and Nature participates through light or sound, or just by being there.

Pause reading and take one minute or two to contemplate around you. Can you see anyone or anything that will probably become memorable? Would you like to revisit any location or situation that you have experienced today?

2. The pencil of Nature

In its early days, photography was considered as Nature's own discovery of a way to register its image, hence the borrowed title from Fox Talbot's book [9]. It also presented the challenge of understanding and embracing the medium's automation and categorization. Photography offered an alternative to realistic painting in portraiture. It took a while to reach recognition and to witness the birth of multiple photographic genres, some adopted from other arts and boldly crossing different fields, such as the recent case of magical realism as documentary [10].

Painters continued reinventing their art in a succession of new movements. Similarly, photography has proceeded discovering its paths, also as an artistic medium and trustworthy communication vehicle. As a new and more accessible medium, photography also democratized portraits, family photos, and self-documentaries. From a social perspective, this documentation attests the importance of a dignified registry of existence. Our visual culture started flourishing, enriched by all sorts of medium production and outlets.

Wide adoption of photography also brings some controversies. In early photographic portraitures, some of the garments used were borrowed to make the subject look wealthier. A portrait of Abraham Lincoln [11] seems to combine his respectful head with the body of another politician. There is an intense debate over famous images such as *The Falling soldier* [12] by Robert Capa in 1936: it supposedly depicts the death of a republican soldier from the Libertarian Youth, during the Spanish Civil War, but it may have been staged. Dorothea Lange's famous Migrant Mother image was carefully intervened by the photographer, through reframing and subject manipulation, which in theory should not be expected from a documental photographer [13]. In 2014, after the Chennai floods, an image of the Prime Minister of India, Narendra Modi, was circulated worldwide, in which he was looking through a plane window with a crudely post-produced montage. And so on until our days, where GenAI can create alternative stories with multimedia support.

Legislators run behind this excitement, striving to predict and prevent the damages of powerful imagery on the loose. We also seem to fail to educate the youth, which we'd rather see as brave independent experimenters, determined not to *scroll* their lives away.

From an epistemological perspective, the central aspect lies around setting clear frontiers between knowledge and opinion, identifying differentiating criteria. From there, we could regulate and legislate to ensure general access to the Bare Necessities – remember Disney's wonderful song in The Jungle Book [14]? This would ease the social pressure and shift our behavior from scroller to participant, to becoming a dignified user with a critical voice.

3. Stories of our life

To our best current knowledge, as a social and cultural practice of sharing stories, storytelling is a human activity. It's through stories that we organize our thinking and create useful narratives, which serve different purposes, from entertainment to moral ideation.

There is a considerable body of work about the storytelling brain, on how it shapes narratives and is, in return, shaped by them [15]. We go through our lives trying to make sense of what we perceive, and we build overarching narratives that we hold on to. The brain has the ability to break down our experiences into bits of information, attributing neural representations to each particular event. It constantly exercises back and forth, updating those neural models, encoding them into memory, and originating new ones. Recent research suggests that we reactivate relevant stored information to make sense of ongoing experiences, interpreting the present through a lens of the past [16].

Apophenia is the tendency of the brain to seek meaningful connections between unrelated events or independent items. The brain is drawn to patterns, and prone to attributing magical meaning to random information [17]. Despite being a common phenomenon of brain function, apophenia makes us vulnerable to conspiracy theories and misinformation. Our brain processes are as complex and fascinating as its ability to engage in these activities unnoticed by the host (at an unconscious level).

Physiology has an enormous impact on how we experience our lives: it constitutes the hardware we are born with and hosts the software we keep updating and improving. This is at the core of how we experience, how we recollect, and how these two processes are intertwined.

Our reality is what we come to know, what we experience first-hand. Experiencing in that sense would require presence, would demand living the moment to the full. However, we should also visit how others perceive shared or similar occurrences, in order to expand our knowledge, individually and collectively. For Jerome Bruner, while studying a world of reality constructed accordingly to narrative principles, we are faced with the challenge of understanding how narratives organize the structure of our experience: *how "life" comes to imitate "art" and vice versa* [18].

Humankind has gathered, throughout History, countless parables and fables, with the usual intend of conveying a moral or ethical lesson. Engaging with stories requires a sophisticated blend of cognitive and affective processes [19]. Emotional memories are foundational to our personal life. Emotions can interfere with our memories, and they can enhance or distort how we recall certain events. Cognitive research shows that we are constantly revisiting and adjusting our memory, and therefore reshaping what becomes and constitutes our narrative identity, this embodied and unfolding story of the self that humans construct to understand their condition and attribute meaning to their lives. [20-21]

Current research is also interested in understanding how events become part of our narrative identity, how they are selected and discarded. At the core of those processes is the mechanism that brings memories back to our attention, and how it can ground our identity. Recent findings indicate that memories tied to autobiographical periods become more central to narrative identity [22].

The exercise of current self-expression in social media is often done through an external lens. Instead of being grounded on sharing an internal perspective, it focuses on serving a purpose of external appreciation via social signaling. This is not a new phenomenon or concern, but the current scale of it is becoming concerning, as is the influence of others (influencers or opinion-haters) over our life, and our choices.

All the apps, filters, and lighting tricks that contribute to an improved selfie for a social media point in the same direction: instead of appreciating the experience, and enjoying the collected imagery as is, we want to take possession over how it could have been at its best. Is it not enough to live it in the first person? How does this impact our experience?

This shift in perspective, where events are witnessed via a feedback lens, contributes to distance individuals from themselves and their own experiences. It creates the impression that, to build a strong sense of identity, we are relying less on living and experiencing, and more on pretending and showcasing.

This characteristic of interactions in the modern world constitutes one of the main ingredients of our approach to co-creating with Nature, as we should be able to stimulate behaviors that provide social reward. Humankind should be fostering unique perspectives, nurtured by artistic expression and with implied authorial co-creational liberties. People should be incentivized, at all ages and particularly in their youth, to be present and embrace each fleeting moment as they happen.

With a social backdrop of ongoing efforts around Equity, Diversity and Inclusion, movements promoting freedom of speech, and groups proclaiming their individuality, how can we still succumb to external validation as a significant dictator of our everyday life decisions?

4. AI uprising

Education should focus on supporting students in becoming epistemic agents [23-24], while addressing topics such as Equity, Diversity and Inclusion, global threats (i.e., Climate Change and pandemics), financial literacy, etc. It’s a task, both immense and fascinating, that should always be present when making plans for the future: we must do better in raising minds that will not all be scientists, but will certainly need to understand the concepts and terminology. The survival of humanity requires future generations to regulate their lives by scientific processes, understanding when science is talking to them and what is required by their environment, be it natural or human made. The non-scientific community tends to be steered by their beliefs, but they must learn to engage with reliable sources for scientific answers. Interdisciplinary efforts are necessary to ensure that beliefs and facts are clearly separated in education. Additionally, learning should be fun, as students are more and more in demand of engaging learning practices.

GenAI can provide significant support in approaching these challenges, as it provides an unprecedented level of access to the currant human knowledge. The outputs of GenAI models (texts, images, etc.) are new, in the sense that they had not previously existed as such but are not novel: they are a remix and reshaping of previous human creations. The massive appointment of writers to feed language models is a clear sign that AI does not create, but rather regurgitates [25]. Further evidence is provided by the notion of “model collapse”: using model-generated content to train LLMs causes irreversible damage to the resulting models [26]. Hence, GenAI is particularly adequate for applications that do not require original creations, but rather an efficient and reliable access to a large knowledge, which is the case for education.

Following on previous research [8], we also focus on the empowering aspects of GenAI for creators, providing the ability to accelerate ideation, unleashing process for thinkers and makers. In this context, we use the term co-creation to refer to the curation of GenAI outputs towards a creative goal. Artists, for example, often base their co-creations on so-called hallucinations – glitches produced when the models recognize patterns that are incompatible with human perception, and generate outputs that are nonsensical [27].

Humankind was initially marveled by photography’s authentic depiction of reality, even if photos can be biased by default: they are taken from one specific point of view, with a particular choice of framing, and they often depict staged performances. This state of awe and wonder is praised, and we have always wanted more of this human-machine interaction: seeing and communicating beyond earthly frontiers; and when depicting and conversing was not enough, small devices were created to allow us to capture, edit, and share.

Social signaling is growingly tightened to how we capture the world and, inherently, how we perceive it. We view our environment through the interests of our followers, and we mistake our needs for theirs. This dissociation represents a fracture between humanity’s need of survival and the perception of that need. GenAI can help bridging that fracture by facilitating an inspiring relationship with Nature and its biodiversity.

5. Crowd-sourced Nature trails

We aim at nurturing our relationship with Nature by showcasing its undeniable presence and relevance in our lives, therefore assuring its integral part in our narrative identities, both individual and collective.

GenAI can be used as a tool for observation, learning, and ultimately to foster an improved sense of belonging and environmental agency. Whereas Nature should be experienced without mediation, optional and nonintrusive extra layers can be added for enhanced perception, and these can be non-linear, both in time and space.

Visitors to a natural park can register their walked paths via an inconspicuous location-tracking app, and gather digital memorabilia along the way, either by recording media (e.g., photos, sounds, videos), by indicating personal points of interest for future reference, or by providing recommendations for other visitors. These recommendations may include guidance about good shaded locations for summer picnics, or sunny trails for winter walks. Some locations may exibit QR-codes to tag presence on exclusive or third-party social platforms.

The ghost-memory of visitors is left in those virtual trails, which can be revisited at home, by the visitor or other users. Visitors can explore some aspects of Nature that they were curious about during the walk, but that they left for a later occastion in order to experience a unperturbed communion. These revisiting moments can be co-created with GenAI by the exploration of enhanced immersive interactive multimedia stories, that also include memorabilia gathered by other visitors. These narratives co-created from past experiences with Nature, will provide new stories to enrich our narrative identities beyond social signaling.

Collective narratives can work at community level or globally. They can be used to raise awareness to specific environmental issues, but that can also provide invaluable archival data: documenting how places evolve and what experiences appear and disappear. In that sense, GenAI can support and document conservation, social history, and ecology.

References

[1] Milmo, D. (2023, February 2). ChatGPT reaches 100 million users two months after launch. *The Guardian*. https://www.theguardian.com/technology/2023/feb/02/chatgpt-100-million-users-open-ai-fastest-growing-app

[2] Pavlidis, G. (2024). Unlocking the black box: Analysing the EU artificial intelligence act’s framework for explainability in AI. *Law, Innovation and Technology*, *16*(1), 293–308. https://doi.org/10.1080/17579961.2024.2313795

[3] Market Insights, S. (2024). *Generative AI - Worldwide*. Statista. https://www.statista.com/outlook/tmo/artificial-intelligence/generative-ai/worldwide#global-comparison

[4] UNESCO. (2023). *Foundation models such as ChatGPT through the prism of the UNESCO Recommendation on the*

Ethics of Artificial Intelligence. UNESCO. https://doi.org/10.54678/BGIV6160

[5] Safe and Secure Innovation for Frontier Artificial Intelligence Models Act, SB-1047, Government Code (2024). https://leginfo.legislature.ca.gov/faces/billNavClient.xhtml?bill_id=202320240SB1047

[6] Union, E. (2024). *The AI Act* (Regulation EU) 2024/1689 of The European Parliament and of The Council No. Regulation (EU) 2024/1689; Official Journal of the European Union). European. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32024R1689

[7] European Commission. Directorate General for Communications Networks, Content and Technology. (2020). *The Assessment List for Trustworthy Artificial Intelligence (ALTAI) for self assessment.* Publications Office. https://data.europa.eu/doi/10.2759/791819

[8] Reis, C. M., & Câmara, A. (2023). Expanding Nature's storytelling: Extended reality and debiasing strategies for an eco-agency. *Frontiers in Psychology*, *14*, 941373. https://doi.org/10.3389/fpsyg.2023.941373

[9] Talbot, W. H. F., & Schaaf, L. J. (1989). *The pencil of nature* (Facsimile ed). Hans P. Kraus, Jr.].

[10] Duzant, M. (2024, October 31). El pez muere por la boca / The Fish Dies By Its Mouth by Santiago Escobar-Jaramillo—Book review. *LensCulture Weekly*. https://www.lensculture.com/articles/santiago-escobar-jaramillo-el-pez-muere-por-la-boca-the-fish-dies-by-its-mouth

[11] Farid, H. (2006). Digital Doctoring: How to Tell the Real from the Fake. *Significance*, *3*(4), 162–166. https://doi.org/10.1111/j.1740-9713.2006.00197.x

[12] Susperregui, J. M. (2016). The location of Robert Capa's Falling Soldier. *Communication & Society*, *29*(2), 17–43. https://doi.org/10.15581/003.29.2.17-43

[13] Curtis, J. C. (1986). Dorothea Lange, Migrant Mother, and the Culture of the Great Depression. *Winterthur Portfolio*, *21*(1), 1–20. JSTOR.

[14] Reitherman, W. (Director). (1967). *The Jungle Book—The Bare Necessities, song by Phil Harris, Bruce Reitherman* [Animated musical adventure]. Buena Vista Distribution. https://www.youtube.com/watch?v=6BH-Rxd-NBo

[15] Gottschall, J. (2021). *The story paradox: How our love of storytelling builds societies and tears them down* (First edition). Basic Books.

[16] Hahamy, A., Dubossarsky, H., & Behrens, T. E. J. (2023). The human brain reactivates context-specific past information at event boundaries of naturalistic experiences. *Nature Neuroscience*, *26*(6), 1080–1089. https://doi.org/10.1038/s41593-023-01331-6

[17] Blain, S. D., Longenecker, J. M., Grazioplene, R. G., Klimes-Dougan, B., & DeYoung, C. G. (2020). Apophenia as the disposition to false positives: A

unifying framework for openness and psychoticism. *Journal of Abnormal Psychology*, *129*(3), 279–292. https://doi.org/10.1037/abn0000504

[18] Bruner, J. (1991). The Narrative Construction of Reality. *Critical Inquiry*, *18*(1), 1–21. https://doi.org/10.1086/448619

[19] Vaccaro, A. G., Scott, B., Gimbel, S. I., & Kaplan, J. T. (2021). Functional Brain Connectivity During Narrative Processing Relates to Transportation and Story Influence. *Frontiers in Human Neuroscience*, *15*, 665319.

[20] McLean, K. C., Pasupathi, M., & Syed, M. (2023). Cognitive scripts and narrative identity are shaped by structures of power. *Trends in Cognitive Sciences*, *27*(9), 805–813. https://doi.org/10.1016/j.tics.2023.03.006

[21] Marsh, E. J. (2007). Retelling Is Not the Same as Recalling: Implications for Memory. *Current Directions in Psychological Science*, *16*(1), 16–20. https://doi.org/10.1111/j.1467-8721.2007.00467.x

[22] Thomsen, D. K., Pedersen, A. M., & Salgado, S. (2024). The experiences that define us: Autobiographical periods predict memory centrality to narrative identity. *Journal of Applied Research in Memory and Cognition*, *13*(2), 273–281. https://doi.org/10.1037/mac0000130

[23] Nieminen, J. H., & Ketonen, L. (2023). Epistemic agency: A link between assessment, knowledge and society. *Higher Education*. https://doi.org/10.1007/s10734-023-01142-5

[24] Elgin, C. Z. (2013). Epistemic agency. *Theory and Research in Education*, *11*(2), 135–152. https://doi.org/10.1177/1477878513485173

[25] George, J. A. (2024, September 7). 'If journalism is going up in smoke, I might as well get high off the fumes': Confessions of a chatbot helper. *The Guardian*. https://www.theguardian.com/technology/article/2024/sep/07/if-journalism-is-going-up-in-smoke-i-might-as-well-get-high-off-the-fumes-confessions-of-a-chatbot-helper

[26] Shumailov, I., Shumaylov, Z., Zhao, Y., Papernot, N., Anderson, R., & Gal, Y. (2024). AI models collapse when trained on recursively generated data. *Nature*, *631*(8022), 755–759. https://doi.org/10.1038/s41586-024-07566-y

[27] Maleki, N., Padmanabhan, B., & Dutta, K. (2024). AI Hallucinations: A Misnomer Worth Clarifying. *2024 IEEE Conference on Artificial Intelligence (CAI)*, 133–138. https://doi.org/10.1109/CAI59869.2024.00033

The Journey of the Memory Horse

Prof. Nicolas Reeves. PhD, S.M. Arch.S
School of Design, University of Quebec in Montreal, Canada
www.nxigestatio.org
e-mail: reeves.nicolas@uqam.ca

Vincent Cusson, PhD fellow
School of Media and Communication, University of Quebec in Montreal, Canada

Alexandre Pariseau, B. Sp. Communication
School of Media and Communication, University of Quebec in Montreal, Canada

Pierrot-Baptiste Lemée-Jolicœur, MSc Phys.
University of Montreal, Canada

Marie-Martine Bédard, B. Sp. Music
Activist and independent musician, Ste-Adèle, Québec, Canada

Abstract

"The Journey of the Memory Horse" is a research-creation program that takes its name from a small organ located in the brain, whose shape evokes that of a seahorse. The *hippocampus* (its anatomical name) is the area in which are recorded the events that deeply mark our memory, such as serious diseases or traumas. Located at the crossroads of science, technological arts and popular arts, the program is based on an interface intended to transpose the brain waves captured by an electroencephalogram (EEG) headset into timbres and melodic embryos that are then used for composing and arranging musical pieces with a vocal part. The project itself originates from a very violent trauma suffered by a musician from the Quebec pop-rock scene (Marie-Martine Bédard, identified below as Marie-Martine), during an assault that left her for dead, following which she undertook a journey of resilience lasting more than twenty years. Having become an activist committed to victims of similar attacks, she decided to produce an album of compositions in a very elaborate pop-rock style, based on the capture of her brain waves during recording sessions where she recalled the different stages that preceded and followed the aggression, including the

period of her mental, physical and physiological reconstruction. To achieve that result, she called upon the skills of our laboratory team which has already designed several instruments intended to transpose different objects and phenomena, such as remarkable architectures or atmospheric phenomena, into music. Despite our expertise, this transposition of brain waves into data with a musical potential faced us with a number of challenges. Reducing the data flows to a level allowing their transposition into music proved particularly complex, as did the creation of an interface that was ergonomic, fluid and intuitive, allowing musicians who were not experts in science or technology to quickly learn how to use this new instrument.

Our intervention will summarize the history of the project, the realization of which involved two research laboratories, a physicist, engineers, designers and musicians. The composition interface will be demonstrated in real time. The final result consists in nine music/songs pieces composed and written by Marie-Martine, who was assisted for the arrangements by one of the most recognized musicians in Quebec. A first show has already taken place; some commented excerpts will be presented during the conference, illustrating the quality of the results that can be obtained by the collaboration between highly talented musicians and university research-creation experts. The project, which has caught the attention of the Quebec media and is now moving towards wider distribution, is now cited as an example during the activities of the meetings organized by Marie-Martine, who, in parallel with her work as a musician, accompanies groups of victims of similar traumas during their own journey of resilience.

One of the recognized successes of the project is precisely to have succeeded in this harmonious meeting between two worlds considered foreign to each other, and to demonstrate the potential of digital arts, often considered elitist or over-specialized, for a form of popular art accessible to all audiences.

1 • Music in everything

In our previous projects, we were led to develop software musical instruments related to electronic lutherie, intended to transpose, through composition interfaces also designed in our laboratory, different phenomena into acoustic and musical sequences. We thus designed several instruments intended to allow musical composition from atmospheric phenomena [1] [Fig.1], another that read, through an iterative loop, the landscapes of Quebec as observed by satellite [2], another that transposed into music the architecture of remarkable buildings [3]. Another one, currently being implemented, generates variations of a classical piece through the genetic mutations that occurred after it was coded in the genome of an E. Coli bacterium [4].

Thank to these instruments, we were able to present real-time and *in situ* performances. Their different functions allowed us to generate and arrange melodies from sound timbres that could be pre-recorded (sampling), produced procedurally (like in the MIDI standard) or developed in real time from the phenomenon considered. This last possibility proved to be by far the richest and most promising of the three. Positioning our composition work in the field of spectral music, it allowed the transposition processes to unroll with immediate reactivity, and even in real time in cases where this was possible. In addition, the great variability of the phenomena we selected, as well as their unpredictability, reduced almost to zero the probability that the same musical sequence would repeat itself identically.

Thanks to these projects, we have developed a valuable expertise in electronic lutherie as well as in the role and adjustment of sensors, instruments and transposition processes between different types of phenomena and musical composition [5].

A note of precision is necessary from the outset: the work that we are doing does not fall under the field of sonification. Specialists in this latter field attempt term-by-term transpositions, analogous to isomorphisms, between a natural phenomenon and the sound events that represent it. This would be the case, for example, if our Cloud Harps directly and uniquely matched the height of clouds with the pitch of a note in the scale. This way of doing things is not deprived of interest: it is based on a hypothesis according to which the conversion of large data sets into sounds would reveal regularities or singularities that could not be detected by the analysis of raw data or through their visualization. Among the many attempts in this direction, we can note the work of Andrea Polli [6], who sonified recordings of weather data corresponding to hurricanes; or to the sonification of the first gravitational waves ever observed, produced by the collision between two black holes [7]. In such situations, the relationship between the phenomenon and the reaction of the instrument is similar to that which links the external temperature to a thermometer, or the amplitude of a sound signal to a VU-

Fig. 1 – The Ithaca Dream, a meteo-electronic instrument from our lab that transposes real time several atmospheric events (wind speed and direction, clouds height and density, temperature, pressure, humidity, pluviosity, electrostatic field, lightnings within a

50km distance) into musical sequences. Parc of the Bussy-Rabutin castle, France, 2021.

meter: it is limited to a quantitative information, akin to a measurement, which does not present a real interest at the musical level since it involves no real composition process.

All the instruments we developed include such a quantitative phase, carried out using probes or suitable sensors. Here too, the data sets they gather are similar to a quantitative measurement. But another step follows, in which they are formatted so that they can be input into the instrument itself, through which the musician can elaborate his/her own compositions, which are unique to him/her. He/she is in charge of defining the links between the natural phenomenon and the sound events, by adjusting the parameters and transposition scales on both sides. It is he/she who, listening to the results of these operations, selects and associates the timbres and musical embryos which will determine the final composition. The interface is nothing less than a harmonic palette: in the same way that a painter creates colours by mixing pigments and oil, then determines the chromatic palette of the work by mixing and juxtaposing the basic hues thus obtained, the spectral music composer mixes and associates elementary sound harmonics to create a music that, for him/her, evokes the ambiances and variations of the phenomenon he/she is working with.

2 • A Journey of Resilience

Drawing on our experience, when we were approached in 2019 to participate in a composition project involving the sound transposition of brain waves, we did not anticipate any major difficulties: we were simply going, as a university laboratory, to share with an artist the previous results of our research, which we would only have to adjust to obtain the desired results.

This optimistic vision quickly came up against the complexity of the waves produced by biological processes. This apparently straightforward project quickly became a full-fledged research program, on which we have been working for almost four years now. The next sections will detail its objectives, technical aspects and results. Before going through this however, I will briefly summarize the origins and intentions of the initial proposal.

The project itself was initiated by a singer-songwriter, Marie-Martine Bédard, identified below as Marie-Martine, who was lead guitarist and singer in the 80’s for the first all-female punk band in Quebec, a trio called Blue Oil [10]. Along with her band, she performed more than a thousand times in various venues. At the dawn of a promising career, and just after giving birth to her first child, she was the victim of a very violent aggression in Montreal by a serial rapist, who left her for dead. Heavily traumatized, she began a long journey of resilience that lasted more than twenty years.

After a long recovery period, during which she was barely able to get out of her house, she decided to undertake advanced musical studies at the Berklee School of Music in Boston, one of the most renowned musical centres on the planet for musical education. After completing her degrees, she conceived the project of an album of rock songs whose melodies and arrangements would be composed from the signals emitted by her own brain and captured by an EEG (electro-encephalogram) recording headset as she recalls the different episodes of her life surrounding the assault. The completion of this album would be followed by the preparation of a full show, which would add to her own voice the voices of other victims of similar acts. The album, like the show, were intended to be messages of support for all those who have suffered, who would suffer, attacks of the same order.

It is impossible to overestimate the courage and energy it took for Marie-Martine to even think about going back on stage and putting herself back in the spotlights. Psychologically paralyzed after the attack, assailed by destructive and paranoid thoughts, she nonetheless managed to take her life back in hand and to relaunch, more than two decades later, her interrupted career as a singer and musician after years spent hiding at home, and to only go out, after several years, with a large and dissuasive-looking dog.

The name of the project, “The Journey of the Memory Horse”, comes from the way in which traumas are recorded in the brain, where it causes dysfunction and dissociation of the circuits of the hippocampus, a small organ located between the temporal lobes [8]. So named because of its shape, which evokes that of a seahorse, it naturally led to the current title. Our first meetings were mainly intended to validate the feasibility of the project, but, one thing leading to another, the collaboration became progressively close. Empathetic to the cause, which seemed to me as essential as it is relevant for the times we live in, I offered to make my laboratory a full partner in the project.

To this end, we conducted a search for funding and equipment. A grant application to the Canada Council for the Arts was successful on the second attempt, which is remarkable enough to be mentioned here: grants in music, and specifically in song, are rare and particularly difficult to obtain, given the large number of applicants.

Despite this funding, choices had to be made regarding equipment. Cheap EEG headsets can be found on the internet, but their resolution is so low and the data they produce is so noisy that they are unusable for research purpose. On the other hand, the price of entry-level professional capture equipment is incompatible with the grants available for research-creation. We therefore decided to contact Emily Coffey, a professor at Concordia University in Montreal. A senior researcher in psychology and neuroscience, she immediately agreed to contribute to the project. Not only did she lend us a high-resolution professional headset, but she also provided us with a fully equipped room for data collection for an entire day, as well as a research assistant familiar with all the capture process.

3 • Recording the waves

Before proceeding with the actual recording, we had to establish the recording script. According to the singer's wishes, we proceeded according to a chronological sequence, during which she recalled the different episodes of her story: those that predated the assault, the assault itself, the events of the days that followed, and the main steps of her journey of resilience. Fifteen episodes were determined. Their title appears on the following list.

THE EPISODES OF THE JOURNEY

1 - The birth of the child

2 - The day before the assault

3 - The assault and the amnesia that followed

4 - Reconnecting with reality / arriving at the hospital

5 - The hospital stay

6 - The three days following the assault / returning home

7 - Current recollection of the amnesia period

8 - Reconstruction and resilience

9 - Music studies: difficult times, moments of satisfaction, rewards

10 - Precious dogs

11 - Dissensions / breakup with the father of the child

13 - The lead blanket / the unpronounceable word

14 - Mourning oneself

15 - Arriving in a new house / a new life / a dawn

Two full sets of recording were made in two different places. The first session took place in a very controlled room, a chamber equipped for recording and analysing dreams, in the psychology department at Concordia University – we called it the dream room [Fig.2]. The second one took place a few weeks later in a very nice natural environment, on the banks of a river, in the heart of a Quebec forest, a much more serene place. We could plan this second session thanks to a high-level nomadic recording device, loaned by Professor Coffey's laboratory, which we carried in our rucksacks for several kilometres before finding an appropriate location.

In both cases, the recordings were done without any words being spoken. In the dream room, the silence had to be total: any pronounced word would perturbate, and even saturate, the EEG sensors. In the forest, the only noises were those of nature: the whispering river, the leaves in the trees, the song of birds. They did not have any noticeable impact on the brain signals. In both cases, Marie-Martine had to recall in her memory, like if she was silently telling a tale to herself, the unfolding of each of the fifteen episodes. An important point: never since these events, and despite numerous therapy sessions, has she managed to tell anyone the detailed course of the attack.

Fig. 2 – The EEG headset being installed on Marie-Martine's head. 64 electrodes, positioned at strategic places on the skull, are connected to the recording device, which transmits the signals to a

computer. Each electrode must be coated with a thick layer of conductive grease to ensure the best possible contact.

4 • Dealing with overwhelming data flows

The hypothesis underlying the entire project was that the feelings and emotions aroused by the memories of the different episodes would produce signals of a sufficiently different nature to allow the composition of musical sequences characteristic of each moment. Even if the correlation between the memories evoked and the recorded EEG sequences was not always obvious, this hypothesis has been largely verified.

However, before arriving at this conclusion, we first had to analyse the observed data. This is where we realized that our previous expertise in electronic instrument making would not be of much use. The quantity of data collected was actually *colossal*. Where the weather probes of the Ithaca Dream, for instance, typically generated a few measurements per second, the 64 electrodes of the EEG helmet produced a torrent of data whose rate was higher by several orders of magnitude, up to hundreds of thousands of measurements per second. A direct sound transposition of such quantities would produce only pink noise, with only indistinct variations and no musical interest whatsoever: any musical transposition of such signals would be equivalent to that obtained from a random source [Fig.3].

It was therefore necessary to carry out a first processing, by which the temporal variations of these masses of data could be analysed through stochastic algorithms so as to yield an average temporal trajectory corresponding to their overall evolution. Such a process is not unknown in the field of music: composer Yannis Xenakis, among others, used it deliberately in several of his pieces [9], in which very large numbers of very short musical events, generated by an enlarged classical orchestra, were associated in clouds of sound events that moved in groups, in the same way that a cloud in the sky, made of a huge number of water droplets with random individual movements, adopts an overall behaviour that is clearly perceptible to us.

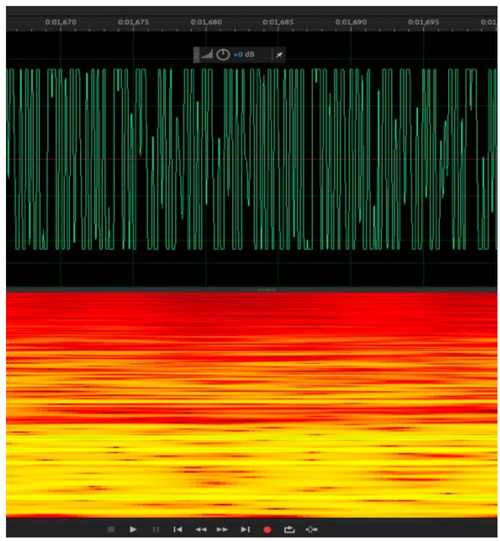


Fig.3 – Visualisation of the raw signal through an audio analysis software. The total duration of this sample is 0,035 seconds. The green line on the top gives the signal amplitude, the red and yellow pattern gives the evolution of its harmonic spectrum (low frequencies at the bottom of the graph, high frequencies at the top. The lighter the colour, the higher the amplitude of a given harmonic). No obvious regularity appears; the signal looks almost like if it

has been generated by a random phenomenon.

Not having the mathematical expertise to carry out this work, we called upon the services of a physics student. With his help, we could implement an interface allowing us to adjust the reading parameters of large data flows (temporal windows, grouping modes, frequency calibration, etc.). At the end of this process, we were able to produce an initial set of files whose transposition into music became possible [Fig. 4].

5 • The Memory Horse Interface

Once the EEG signals were converted into usable files, the next step was to design and implement an interface that was both simple and ergonomic, allowing Marie-Martine and her team to have a first visual overview of each signal before listening to its sound transposition. A team of three designers, two of whom having advanced musical knowledge, focused on this stage of the project, taking advantage of the potential of the MAX/MSP platform.

The first step was to select, on the different recordings, the segments corresponding to the events experienced by the singer, then to locate on these same segments those that would best lend themselves to a musical transposition. Through numerous exchanges between the designers and the musicians, we could specify the functions that would best allow the intuitive exploration of all segments; the transition from raw sound data to harmonized data was the subject of particular attention.

With such an instrument, each composer can define his/her own workflow. In our case, a typical work session began with a careful listening of the direct frequency transposition of the EEG signal. By default, the signal used was a simple sine wave, which gives the clearest auditive counterpart of the displayed EEG graphs. It allowed the composer to define the transposition ambitus, by calibrating the interval between the highest and lowest frequencies between which the signal must remain. This gave us an initial understanding of how the signal evolves over time, and allowed the musicians to

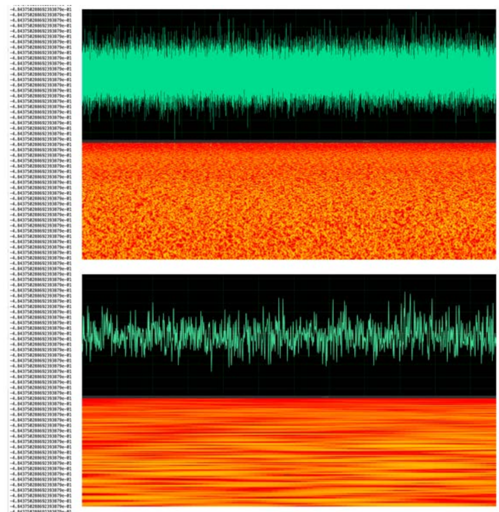


Fig.4 – Enlarged EEG signals after a first processing, at two different zoom factors. The list on the left shows the corresponding values, for a very small part of this signal. In the top picture, the orange graph on the bottom begins to show promising regularities, like a textured canvas. The amplitude graph on the top looks almost continuous. The bottom picture has been enlarged to a much greater factor. The green line still shows frantic variations, but larger oscillations become visible. The analysis software allows to define the size of

temporal windows, in order to find those which will optimally detect these larger variations, whose potential for sound or musical transposition becomes more obvious.

make initial hypotheses about the place that the corresponding segment will occupy in the melody, or in the final arrangement. From this same calibration, the musician could then discretize the frequencies in order to map the signal to a particular musical scale. Seventeen modes were made available, ranging from Dorian to pentatonic, including a classical

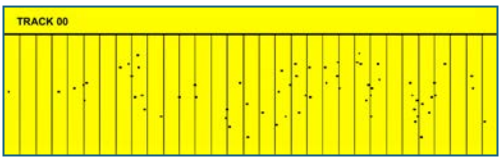


Fig.5 – A basic melody obtained at the end of the transposition process, represented in the form of a time-frequency diagram. The notes, represented as black points, result from the frequency and time discretization of the signal. A simple adjustment of the time and frequency increments will transform this diagram into a musical score. By using the composition interface, the musician will then be able to explore the potential of this melody by trying different musical parameters, such as timbre, tempo, duration, envelope, and by adding different filters.

tempered mode; each one corresponded to a specific musical scale. A second discretization was done at the temporal level, by specifying the number of notes emitted each second: here again, new composition hypotheses could emerge, this time at the level of rhythm or tempo. A commented picture of this interface appears in the annex section, at the end of the present paper.

The musician could then determine the timbre of the instrument on which the transposition was played, by using either the MIDI standard or a VST plug-in. The compatibility of the interface with a lot of commercially available plugins allowed him/her to shape the envelope of the sound, and to modify the result through a wide variety of filters and digital effects, in order to create the musical ambiances instruments that he/she deemed most appropriate to evoke the different episodes of the singer's journey of resilience.

6 • Results and perspectives

In order to guarantee the professionalism and quality of the arrangements and recordings, Marie-Martine retained the services of a renowned arranger in Quebec, Jean-François Lemieux, who has played and collaborated with almost all the major singers and musicians of the Quebec scene. Since a particular attention has been paid to making the interface as simple and intuitive as possible, after a brief training, he quickly realized the wide potential of this electronic instrument. After multiple work sessions, during which the potential of EEG-to-music transposition was explored in depth, nine songs were written, composed, arranged and recorded on a CD called *Projet Hippocampe*, without any further help on our part.

Fig. 6 – The cover of the "Projet Hippocampe" (Seahorse Project) CD.

The production of this album already represents a personal victory for the artist. Far from being satisfied with this first achievement, she immediately started the next step, which consisted in bringing together a group of musicians to put on a full show, the premiere of which took place in 2023 in a church converted into a performance hall in Saint-Eustache, north of Montreal.

This premiere was largely considered a success, and the whole process was followed with great interest by the local media [11]. Despite a twenty-year eclipse, the singer's expertise, acquired during her group's thousand or so concerts, appeared through every aspect of the stage performance. The audience praised its remarkable professionalism, from the compositions themselves to the lighting and scenography, including the sets and the talent of the musicians. Also responsible for a support group for women who went through the same ordeal as her, Marie-Martine brought some of them together on stage, for the duration of a song for which they provided the backing vocals, giving her the opportunity, for a short but decisive moment, to be reborn in the light.

The artist has already begun preparing a second, larger-scale show, this time with an immersive and interactive component. The technological aspects of the composition process will be more present in the scenography, and the compositions will alternate with electro-acoustic sequences also taken from her EEG recordings. The set will hybridize real and virtual elements, the evolution of which being also controlled by EEG data. Each show will welcome the words of victims of similar acts and will be centred on the idea of making the audience aware of the issues related to the awkward silence and deliberate blindness of their friends and relatives after such events, like a leaden blanket that adds to the trauma to further

Fig. 7 - *Marie-Martine Bédard on stage, during her very first show after a decades-long eclipse from the musical scene, in which she performed for the first time several songs written and composed during the "Journey of the Memory Horse" project. The happiness of this moment of artistic rebirth shines through her face and body.*

bury the victims in a state of social disappearance.

7 • Conclusion

As mentioned above, for a person who has gone through the events that have destroyed a whole part of her life, the simple fact of thinking about going back on stage, and to show herself again on a stage, under the spotlight, in front of a crowd of spectators, after more than twenty years of eclipse, is a feat that is impossible to overestimate. An analogy deserves to be tried: the EEG recordings correspond to memories that have remained hidden for years in the depths of the singer's traumatized memory. In a fascinating occurrence of stigma inversion, the show brought them back to light, allowing the singer to welcome and overcome them, to regain control over her life and to transform the resonances of a personal drama into a source of courage and energy. From the first EEG recordings to the public presentation of the show, the Journey of the Memory Horse corresponds almost term by term to the singer's entire journey of resilience.

After having faced and resolved one by one the challenges implied by the project, our team, with the help of the musicians, implemented an interface whose design allowed for fluid and intuitive work, despite the complexity of the algorithms and the huge amount of data involved. After an almost anecdotal irruption in our research-creation activities, the Journey of the Memory Horse progressively established itself as a new research program in its own right. Beyond its role in the production of the album, the interface revealed numerous and fascinating avenues of research, which we promise to explore in the coming years. It joins the collection of electronic instruments that we have developed, from the basic concept to the design of the physical and computer interfaces, increasing, after the songs of the atmosphere and the transposition of major architectures, our level of expertise in this field.

But above all, and this is perhaps the most important point, the project allowed so-called popular art forms – song and rock – to meet technological arts and university research for a project deeply rooted in social and civic commitment, in an unexpected fusion that proves surprisingly – and even strangely - natural. The demonstration of the fertility and richness of this intimate merging of these two art forms is clearly one of the main achievements of the Memory Horse project.

All songs and recordings from the project are available for listening on the artist's web site: https://mmbedard.com/

References

[1] Reeves N - *Ceux qui parlent par la voix du vent : généalogie d'une Harpe à Nuages*, in « Vents : invention et évolution des formes ». Louis-Michel Nourry ed, ENSAB, Rennes (Fr) 2008

[2] Reeves N – *L'ordinarium ou la mise en œuvre des quasi-infinis*, ETC (61) pp 15-20, Quebec (Can) 2003

[3] Reeves N – *Échos harmoniques d'une cosmologie de pierre*, in « Les changements d'échelle : les arts et la théorie confrontés au réel », pp 377-420, Mimesis, Nancy (Fr) 2020

[4] Reeves N, Despond A, Cusson V, Atmosphéries and the Poetics of the In Situ: the Role and Impact of Sensors in Data-to-sound Transposition Installations, proceedings of the 2021 AudioMostly Conference, pp 48-55, ACM Digital Library 2021 *[Best student paper award]*

[5] Reeves N, Rodrigue S, Cusson V, Sauvé L - *ESCHE BACH*, proceedings of the Generative Arts XXIth conference, pp 282-296, Domus Argenia (Rome) 2018

[6] Polli A - *Atmospherics/Weatherworks: A Multi-Channel Storm Sonification Project*, ICAD 2004

[7] Ubach H, Espuny J – *Sonification of gravitationally lensed gravitational waves*, https://doi.org/10.48550/arXiv.2407.09588, opened 2024/11/08

[8] Bremmer J D, Vermetten E – *The Hippocampus and Post-Traumatic Disorders*, in "The Clinical Neurobiology of the Hippocampus L An Integrative View", pp. 262-272, Thorsten Bartsch ed, Oxford Academics 2012

[9] Gibson B, 2011. *The Instrumental Music of Iannis Xenakis: Theory, Practice, Self-Borrowing*. Iannis Xenakis Series, No 3 Hillsdale (NY) Pendragon 2011

[10] Golden A - *Women to the Front: Interview with Quebec's First All-Female Punk Band Blue Oil,* Louder Than War online magazine, 2023/09/19 https://louderthanwar.com/women-to-the-front-interview-with-quebecs-first-all-female-punk-band-blue-oil/, opened 2024/11/08

[11] Among many others : Moulun A - *Projet Hippocampe : Des échantillons de pensées pour faire de la musique, journal Accès. 30/05/2023.https://www.journalacces.ca/culture/projet-hippocampe-des-echantillons-de-pensees-pour-faire-de-la-musique, opened 2024/11/08/*

Image Credits

Fig. 1 – Nicolas Reeves

Fig. 2 – Nicolas Reeves

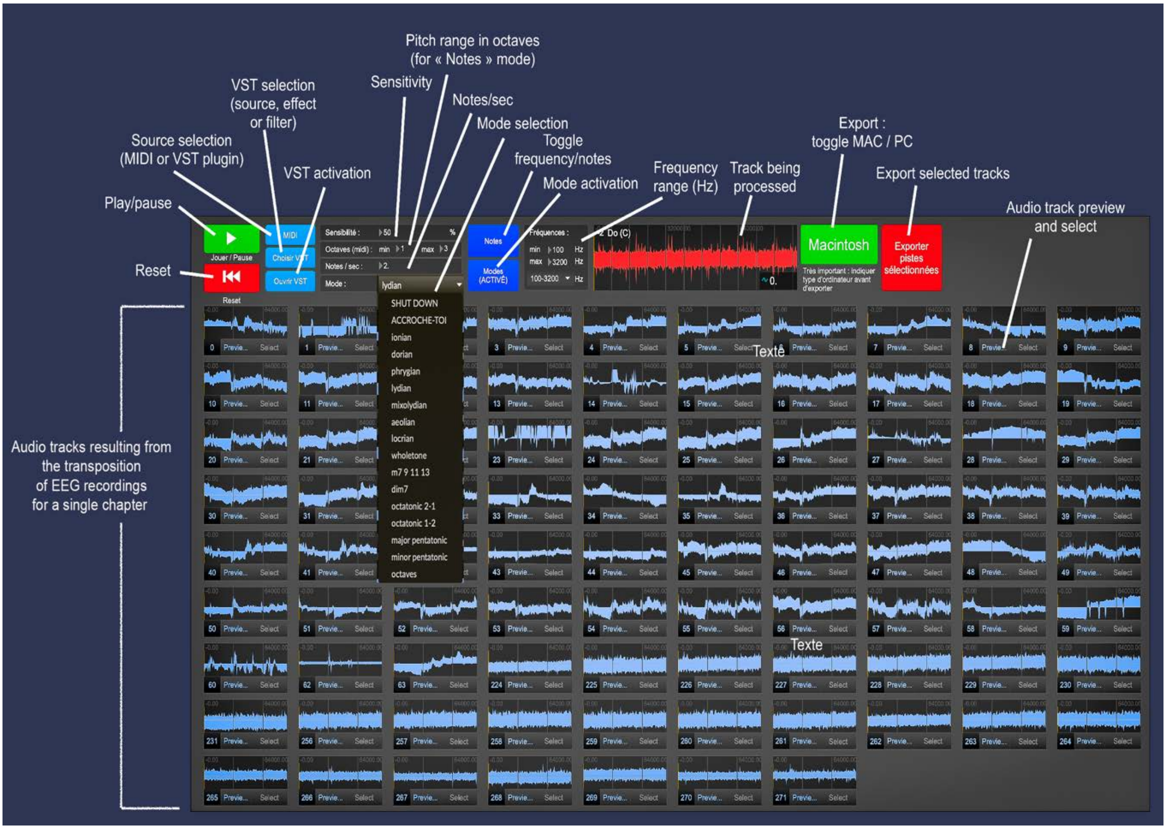
Fig. 3 – NXI Gestatio design lab
(w. Adobe Audition software)

Fig.4 – NXI Gestatio design lab
(w. Adobe Audition software)

Fig. 5 – NXI Gestatio

Fig. 6 – Marie-Martine Bédard

Fig. 7 – Caverne Pyralis



ANNEX 1 • "The Journey of the Memory Horse" composition interface.

To start the process, the composer first selects the episode of the Journey he/she is willing to work with. The program then automatically loads all the audio tracks that have been generated through the transposition of the singer's brain waves while she was remembering the different parts of that episode. By visually examining the variations of the audio signals, he/she can evaluate those that present the best potential for musical transposition. He/she then listens to them by using the frequency mode which produces a continuous sound, a simple sinusoidal wave that directly follows the variations of the signal, and whose maximal and minimal frequencies can be mapped on different intervals.

Once the track selection is done, the composer switches to the Notes mode. There again, the ambitus of the resulting melodies is limited by setting a highest and a lowest notes. To the contrary of the frequency mode however, in which all frequencies are played even if they fall outside of the prescribed limits, the notes that are not included in the interval are not played. The generated melodies will then have moments of silence that constitute an integral part of the composition. He/she also sets the number of notes per second and the sensitivity of the transposition, which defines the minimum amplitude of the signal that will trigger a note. By using the MIDI mode, and setting it, for instance, to the sound of a piano, simple monophonic sequences can be obtained. Here again, they are quite useful for evaluating the musical potential of a given track.

When he/she is satisfied with a given set of transposed signals, he/she can load a large variety of VST plugins that will allow him/her to try his/her melody on different instruments, to modify the envelope parameters of a given timbre, and to add a multitude of filters and effects that will enrich the sequences and complete their transposition to fully musical tracks. By saving them through the exportation module, he/she will be able to import them in any off-the-shelf mixing and recording software.

Genre Memory and the Beauty of Early Music vs AI in the Audiovisual Performance 'Concerto Grosso – W Labiryncie'

prof. dr hab. Monika Karwaszewska
Stanisław Moniuszko Academy of Music in Gdańsk
e-mail: mon-kar@wp.pl

Abstract

The world of art, like every other sphere of our lives, benefits from the achievements of artificial intelligence, which introduces fascinating, previously unknown possibilities of artistic expression. Innovative technological solutions redefine the way in which the audience experiences and interprets a work of art. Artificial intelligence may also cause a work of art to cease to be original and unique, but rather devoid of emotions and human experience.

The audiovisual 'Concerto Grosso – W Labiryncie' (Concerto Grosso – In the Labirynth), scripted by Tomasz Jocz, is a performance based on the assumption of presenting the issue of human relations with artificial intelligence (AI), creating a new artifact in augmented reality. To show these relations, the author of the concept used the Baroque concerto grosso genre based on musical competition. The instrumental group (string sextet with prepared piano and percussion instruments) was juxtaposed with electronic instruments (synthesizers, electronic keyboards) and a live electronics layer performed by a DJ. The musical layer is integrated with the projection in the form of a visual installation generated live by VJs. A counterpoint, the visual layer (partly prepared) required the use of AI tools offering textual, illustrative and cinematic techniques. The intended polyphonic character of the relationship between the layers (including the choreography) emphasizes their equivalence in the context of the artistic whole of the performance, and thus the creative activities of humans and artificial intelligence. All layers are inspired by the rhetoric and style of past eras. References to the past become symbols, or a message. In this way, the author of the work, in individual movements, expresses his dilemmas over AI dominating modern man.

The original arrangement of the performers on stage was subordinated to a visual installation which created the possibility to model animated objects in virtual space, using the performers' movements. On the one hand, the performance becomes a humanitarian manifesto against the exclusive use of artificial intelligence over humans, and on the other hand, it invites art viewers to interact with the work of art in a way that appeals to their senses, emotions, ideas and creativity. Thanks to the synthesis of tradition and modernity, the place of man in the digital world can be explained through an in-depth exploration and deconstruction of selected reference points rooted in tradition. 'Concerto Grosso – W Labiryncie' will be analysed and interpreted. The artistic project presented is aimed at considering the need to use artificial intelligence for creating innovative works of art that will capture the audience's attention with new artistic and aesthetic qualities. Undoubtedly, AI-enabled works of art require separate scientific and art-based research.

The 'Concerto Grosso – W Labiryncie' artwork according to Tomasz Jocz's concept
Movements:
I Ingerencja I (Interference I)
II Ingerencja II (Interference II)
III Ingerencja III (Interference III)
IV Ingerencja IV (Interference IV)
V Ingerencja V (Interference V)
VI Walka (Fight)
VII Śmierć (Death)
VIII Pawana (Pavan)
IX Triumf (Triumph)

Performers:
Paweł Kukliński – 1st violin
Zuzanna Adler – 2nd violin
Paweł Panasiak – viola
Jakub Grzelachowski – 1st cello
Jan Dutkiewicz – 2nd cello
Jędrzej Kacprzyk – double bass
Paweł Nowicki – percussion instruments
Tomasz Jocz – piano, electronic keyboard instruments, synthesizers
Beniamin Baczewski – bass guitar, bass synthesizer
Karol Balicki – DJ BALIC
VJ-e Robert Turło i Adam Przybysz
YouRhythmics Group

Music: Beniamin Baczewski
Visual action: Robert Turło i Adam Przybysz
Choreography: Marzena Kamińska
Lighting: Piotr Paruszewski

Keywords: intermedia; audiovisual performance, artificial intelligence, concerto grosso, visual art

1. Introduction. The subject of research and interpretation

As is already widely known, modern technologies and the rapid development of new media create endless possibilities for creating works of art – intermedial, interactive and generative art – while changing the way in which art is received. As a result, we can observe various relationships developing between individual media (artistic discourses). ‘These are usually associated with the blurring and crossing boundaries between media; with the hybridization of media utterances; with intertextual relationships between media; with intermedial relationships between media; and with an increasing self-reference and self-reflection of the arts as media' [7].

The intermedial work chosen for presentation, analysis and interpretation is an audiovisual performance entitled 'Concerto Grosso – W Labiryncie' (Concerto Grosso – In the Labirynth) according to Tomasz Jocz's concept. The music was composed in 2023 by a young Polish composer Beniamin Baczewski. The live premiere of the original version of the performance was given on 9 May 2024 in the Concert Hall of the Stanisław Moniuszko Academy of Music in Gdańsk. The visual layer was created by artists Robert Turło and Adam Przybysz, who both acted as VJs during the performance. A modern choreography for the YouRhythmics Group was created by Marzena Kamińska. It is therefore an intermedial performance which conveys a completely new message and new quality. The performance 'Concerto Grosso – W Labiryncie' is an example of an interactive work in which generative technologies come into play in order to convey to the audience a message hidden in the narrative. On the one hand, they allow one to make the work more deeply expressive and deliver new aesthetic and artistic quality; on the other hand, they carry a message by highlighting the danger of artificial intelligence dominating humans and their creative acts, or creativity. The main goal of the project was to present the issue of the relations between humans and artificial intelligence (AI).

The performance was created as part of an interdisciplinary project involving composition, live stage performance (traditional instruments and live electronics), performers' movement, improvisation and visual creation (e.g. generative computer graphics, animated object modelling in virtual space, computational processes). The motive and visual layers, depending on the artistic vision and technical setup of the performers, are partly prepared and planned, and partly improvised and unexpected both for the performers and the audience. Thanks to this, an original intermedial narrative is created which, as rightly observed by a research team from Brazil analysing this type of art, 'narrative that explores the poetic potential of these images, ultimately serving as an inspiration for various other creative endeavors' [5].

A show with performance art elements, it reveals the symbolic meaning of sound and image, interacting with choreography, lighting and the partially AI-generated visual layer. All the artistic messages employed correspond with each other despite the fact that each sign system uses different symbolism. Individual media, bearing a relationship with each other, convey an integrated artistic message. Just as opera is a hybrid of various arts: music, ballet, theatre and, therefore, various media – singing, movement, speech, acting, lighting, choreography, electronic means – the audiovisual performance discussed may be categorized as a type of overt intracompositional intermediality, as distinguished by Werner Wolf in his classification [11].

What is fundamental in the performance 'Concerto Grosso – W Labiryncie' is the category of scenicity, which determines

spectacularity and explains the content of the work. The visual and motive layers do not serve as an accompanying background; they constitute an equivalent part of the polyphonic texture, participating in the creation of the integrated work. This means that none of the layers can exist separately to convey the hidden meaning to the audience.
New technologies and artificial intelligence have undoubtedly contributed to closing the distance between a work of art and its audience, and the qualities of a work of art using new technologies are different for the creator and the audience. The audience of such a work first solves an intellectual puzzle, and then, its sense and artistic quality. The Polish researcher and philosopher Maria Gołaszewska points out that works created on the basis of new technologies 'do not always need to coincide with the criteria of art. Most often, a hybrid creation is produced as a technical and artistic or para-artistic performance.' [6] Therefore, that the authors of the work refer to the category of beauty seems to have become the highest category. The main idea that led the artists was to search for those qualities that would move and evoke intense emotions in the audience. The use of artificial intelligence to project the performance in real time was added to merely reinforce the creative human activity. AI, gradually being incorporated into human life, also appears as a theme displayed in the visual layer. Thanks to this, it accompanies the audience throughout the form of the work, reminding them of its dominance not only in art, but also in education and everyday life. With each movement of the performance, the AI interference becomes more intense, reaching its apogee in the final chord.
The main aim of the research was to seek answers to the question how the creative potential of new technology influences the process of constructing the narrative, both in the musical, choreographic and visual layers, and how new means of expression can be fused with musical practices and styles known from tradition and, more essentially, if they can dominate art and exclude humans from actively creating it. In this case, tradition is understood as an inter-text, thanks to which the audience can hear various references to the styles of musical expression characterizing specific directions in early music (medieval organum, concertante technique). It is the clash of past reference points with the contemporary artistic status quo, mentioned by Tomasz Jocz, that is supposed to reflect the relationship between humans and virtual reality. To illustrate the collision, or the constant competition of man with AI, the authors used concerto grosso to shape the form.

2. A musical labyrinth 'between styles'

The expression 'between styles' is a overt reference to the title of the monograph by Stanisław Balbus, the author of intertextual strategies [3]. When analysing the musical layer of the work, one should refer to one of the intertextual strategies mentioned by the literary scholar Stanisław Balbus in his publication. It concerns 'proper styling', which nowadays offers the opportunity to focus on source music and combine it, or rather juxtapose it, with the latest trends in art. Stylisation, according to Balbus, 'highlights the most clearly and starkly [...] the character of all intertextual phenomena, i.e. intertextual and interstylistic relationships' [3]. and 'one of the main functions of stylisation is to reinterpret tradition' [3]. The authors of the work, looking for inspiration in the heritage of the past, drew from the traditional concerto grosso genre as well as the style of musical expression of past eras, thus setting the standards for its reception. Used were the basic properties of the genre and styles which made it possible to shift existing paradigms and create new ones. At this point, the memory and beauty of the genre (concerto grosso, medieval organum) and style (polyphonic technique, concertante technique) seem to be the most appropriate choice, next to the instruments, to be juxtaposed with the new technology. The performance consists of nine independent movements with titles. The titles are not accidental; they reflect the next stage of AI interference in human activities to finally take over. The last movement, Triumf, is the victory of one over the other. The final result of this fight remains for the audience, who will create their own vision by following the audiovisual plot.

I Ingerencja I (Interference I)
II Ingerencja II (Interference II)
III Ingerencja III (Interference III)
IV Ingerencja IV (Interference IV)
V Ingerencja V (Interference V)
VI Walka (Fight)
VII Śmierć (Death)
VIII Pawana (Pavan)
IX Triumf (Triumph)

The structure of the performance was inspired by the assumptions of the Baroque concerto grosso genre, along with its main idea of competition, to illustrate the 'fight' between humans and AI. The idea of competition is expressed in the use of contrasting groups of performers: traditional instruments and electronics, which reflect humans and AI. The instrumental group (string sextet with prepared piano and percussion instruments) was contrasted with electronic instruments (synthesizers, electronic keyboards) and a live electronics layer performed by a DJ. In addition, various styles of popular, techno, film and video game music were used to illustrate artificial intelligence. This allows the audience to understand the message of the entire performance. When composing the musical layer of the work, the authors were inspired by the four-part organum Viderunt omnes by Perotinus, who was the first European composer to create four-part polyphony. Inspiration has several sources. First of all, the authors were inspired by the sacred space of the Notre-Dame Cathedral in Paris, where Perotinus worked to design the appropriate colours and texture of the work with planned breaths. The burning cathedral, as a symbol, also appears in the visual layer in the movement VIII, Pawana. Moreover, the work was inspired by the modal rhythm of the organum (melismatic groups of three) and modality. However, strict ecclesiastical modal scales were not used, but instead a hexatonic scale and motifs characteristic of the work of Perotinus. Throughout the form, the musical material is destroyed and expanded as a result of the introduction of dissonances which reflect AI interference. The same happens with the rhythmic layer, in subsequent parts the rhythm becomes more and more complex, from simple structures to polyrhythms. The opening and closing chords in the work have symbolic meaning. Simple in its structure, the chord that opens the piece refers overtly to that of the organum, which was an address to God. In the performance, the harmony – pure and simple in its structure – reflects the human being (see Figures 1, 2). The work's finale is a grotesque triumph; the harmony becomes 'dirtied', distorted with dissonances and electronics, which symbolize humans dominated by AI. The polyphonic texture of the organum was reflected in the polyphonic treatment of all layers in the performance. What invites comment is the instrumental layer, which also becomes more and more intense with subsequent interferences. The aggressive, unpleasant for the audience, layer of live electronics performed by the DJ comes to the fore, building the tension and reflecting the domination of AI over humans. The DJ part was created on the basis of samples created by the composer with the help of tools (including Digital Audio Workstation – Cubase by Steinberg, VST Instruments – East West, Halion, Groove Agent, Padshop, Sample Audio-Splice, Envato, synthesizers Yamaha MODX, Korg Monologue), which are presented at the concert in a live electronics form. Therefore, at the level of all the aforementioned musical elements, over time there is a gradual deconstruction of the simplicity assumed by Perotinus, who is identified in the performance with human activities. Already in Ingerencja III, the audience can experience how artificial intelligence tries to take over the creative activity of humans.

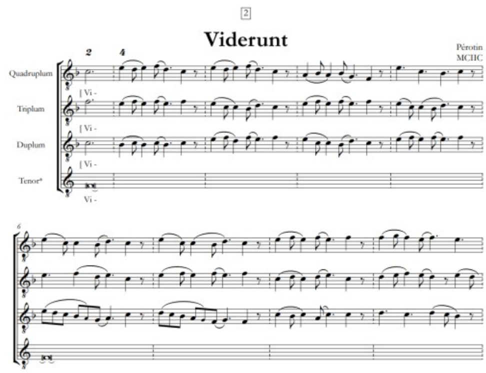


Figure 1. Perotinus, Viderunt omnes [8]

Figure 2. Beniamin Baczewski, W Labiryncie, I. Ingerencja I [2]

3. The aesthetics and message of interactivity in the choreographic and visual layers

Artificial intelligence has undoubtedly caused change in contemporary dance and choreography.

The performance shows how the programming language and artistic challenges that visual artists face while working on processing real-time movement and gestures can affect the aesthetic and artistic qualities of audiovisual works involving artificial intelligence. The key issue that arises, apart from music and choreography, is the use of interactive systems which ensure a modern appeal on the one hand and, on the other hand, they fulfil a signifying function by illustrating artificial intelligence. The process of designing, programming interactive space and translating motive capabilities into other sign systems is complex; it involves the collaboration among specialists across several different disciplines. In his article, choreographer Johannes Birringer presents a definition of interactivity that is adequate for the technology used in the performance. Birringer understands interactivity as 'as a spatial and architectural concept for performance, and second, I look at "interactivity" in the narrower sense of collaborative performance with a control system in which the performer's movement, gesture, and action are tracked by cameras /sensors and thus used as input to activate or control other component properties from media such as video, audio, MIDI, text, graphics, QuickTime movies, scanned images, and so forth. In the latter case we speak of an interactive system that allows performers to generate, synthesize, and process images, sounds, voice, and text within a shared real-time environment.' [4].

The performance 'Concerto Grosso – W Labiryncie' used an interactive system for creating the visual layer. The images were partially prepared, i.e. generated by AI on the basis of the message sent, and then approved and processed by the VJs. The results are presented at the live concert in a free configuration, yet taking into account the character and style of a particular movement of the performance. Thanks to this, based on the music and the performers' gestures, the VJ creates in real time the expression of images. To generate them, AI tools were used that deploy textual, illustrative and cinematic techniques (including Mid Journey, PIXverse, Pika, Runway, Leonardo, Playground, Haiper, Krea, Kaiber, Lens Go, Luma Dream machine). Producing visual effects during the live concert (presented on the screens) was possible thanks to the innovative Resolume editing and mixing software. Thanks to its intuitive interface, VJs could freely create and mix live video footage and effects on any number of screens (including plastic drop cloth, serving both as a projection screen and scenery). The concert was also enhanced by technology allowing generative animation using the performers' movements. The technique of reproducing the expressive movements and gestures of the YouRhythmics Group performers on the screens influenced the generation of realistic animations and special effects from algorithms, which accompanied the other visualisations.

The visual layer accompanies the music and the performers from beginning to end, being one of the counterpoints in the polyphonic texture of the performance. In the subsequent parts, images are displayed in which the audience may experience increasingly stronger 'cutting' of AI into human cells and tissues in order to dominate humans. Algorithms reflecting motive expression are generated in the background (see Figures 3–6).

Figure 3. 'Concerto Grosso – W Labiryncie', III, Ingerencja III (visual layer)

Figure 4. 'Concerto Grosso – W Labiryncie', III, Ingerencja III (visual and choreographic layers), Musical Theatre in Gdynia, October 2024

Figure 5. 'Concerto Grosso – W Labiryncie', VII Śmierć (visual layer)

Figure 6. 'Concerto Grosso – W Labiryncie', VII Śmierć (visual and choreographic layers), Musical Theatre in Gdynia, October 2024

The choreography and preparation of the choreographic layer, as one of the counterpoints of the polyphonic texture, had fundamental importance for conveying emotions. Marzena Kamińska (the choreographer and one of the performers) carefully selected the holographic costumes designed and prepared specifically for this performance, as well as the individual look of the performers (hairstyles reminiscent of dreadlocks, and makeup). The appearance of the performers was intended to reflect the gradual transformation of a human into a robot. In the movements (III, VII) with electronically processed human voice (robot sounds), the performers present stiff movements such as trembling shoulders and hips (the so-called shaking) and positioning their bodies in different directions (spreading their arms in divergent directions).

Movement VIII compels special attention. Pawana is a certain tribute to humans and their life. Here, the authors of the performance established a particularly emotional tone, conveyed through an extremely expressive, mournful melodic line played by the strings and the pianist in octave doubling (see Figure 7). This theme, reminiscent of a lamenting human voice (lugubre e lamentoso), is accompanied by counterpoints: the movement of the soloist wearing a white dress, joined behind the drop cloth by a second performer, as if imitating the shadow of the soloist's slightly modified movements, and images of the eye, changing faces, and visualizations of the burning cathedral in Paris (see Figure 8).

Figure 7. Beniamin Baczewski, W Labiryncie, VIII. Pawana [2]

Figure 8. ‘Concerto Grosso – W Labiryncie’, VIII Pawana (visual and choreographic layers), Concert Hall of the Stanisław Moniuszko Academy of Music in Gdańsk, May 2024

4. Conclusions

The performance ‘Concerto Grosso – In the Labyrinth’ is intended to make the audience realize that artificial intelligence is increasingly penetrating all spheres of life, including art, enabling artists to push the boundaries of their creative expression. It shows the symbolic confusion of man in the reality of AI and the search for the meaning of life by constantly fighting this danger that can arise from replacing man with robots. The labyrinth becomes a metaphor for the enslavement of the individual by an incomprehensible and oppressive system of new technology. Finally, the titular labyrinth symbolizes a hidden meaning (the meaning of the existence of a puzzle), which is presented to the audience using different media. Humans, avoiding obstacles or traps related to the AI's attempt to dominate, tries to reach the desirable goal that is finding balance in the relationship with AI and a balanced approach. The authors try to explain that artificial intelligence should be treated as a tool that is not a threat, but a creative partner, in order to raise the performance to a new, previously unknown level. Juxtaposing artificial intelligence with movement and music is changing the way in which artists and audiences experience both forms of art. What was once imaginary is now becoming real. The transformation brought about by artificial intelligence in art is revolutionary; artists can combine human creativity with the precision of the machine. As artificial intelligence continues to evolve, its impact on the art world will continue to increase, offering endless possibilities for artistic expression and innovation. However, this should be done within reason, so as not to lose what is most beautiful in art: human emotion, feeling and expression which, in an incomparable way, reflect the beauty of a work of art. This thesis is supported by the results of research conducted between 2019 and 2022 as part of the Creative Algorithmic Intelligence Research Project as part of cooperation between the Department of Engineering Science at the University of Oxford and the Oxford Internet Institute. The project report confirms that 'Human agency in the creative process is never going away. Parts of the creative process can be automated in interesting ways using AI...but the creative decision-making which results in artworks cannot be replicated by current AI technology'. [1] Furthermore, the report concludes that 'the future of ML arts will belong to those with both technical and artistic skills.' [9] Drawing the performance's authors' attention to the role of humans and their artistic creation by referring to the art of the past – the outstanding and timeless work of Perotinus and the Baroque genre – is a strong emphasis. These actions are confirmed by Aleksander Tansman's statement: 'dry branches fall down spontaneously, tradition itself is in constant renewal, in constant enrichment' [10].

References

1. *Art for our sake: artists cannot be replaced by machines – study*. University of Oxford, News & Events, published 3.03.2022. https://www.ox.ac.uk/news/2022-03-03-art-our-sake-artists-cannot-be-replaced-machines-study. (accessed 14 September 2024).
2. Baczewski B., *W Labiryncie*, 2023–2024 [manuscript made available with the consent of the composer].
3. Balbus S. (1993). *Między stylami*, Kraków.
4. Birringer J. (2004). Dance and Interactivity. *Dance Research Journal*. 2004;36(1):88–111. doi:10.1017/S0149767700007580.
5. Gatti D., Andraus M.B., Almeida A.Z., Angeli D., Zanchetta G., Partesotti E., Albergaria A.I., Falleiros M., Rossetti D., Pontes V., Manzolli J. (2023). *Dialogue between Body and Sound through Multimodal Interaction*, GA2023, XXVI Annual International Conference at the Borrominian Hall of Biblioteca Vallicelliana, Celestino Soddu

and Enrica Colabella (eds.), Roma: Domus Argenia Publisher, p. 391.
6. Gołaszewska M. (2001). *Estetyka współczesności*, Kraków: Wydawnictwo Uniwersytetu Jagiellońskiego, p. 85.
7. Kattenbelt Ch. (2008). *Intermediality in Theatre and Performance: Definitions, Perceptions and Medial Relationships*, Cultura, Lenguaje y Representación / Culture, Language and Representation, Vol VI\ 2008, p. 21.
8. Perotinus, *Viderunt omnes*, IMSLP (public domain music).
9. Ploin, A., Eynon, R., Hjorth I. & Osborne, M.A. (2022). *AI and the Arts: How Machine Learning is Changing Artistic Work*. Report from the Creative Algorithmic Intelligence Research Project. Oxford Internet Institute, University of Oxford, UK, p. 78. https://www.oii.ox.ac.uk/wp-content/uploads/2022/03/040222-AI-and-the-Arts_FINAL.pdf, (accessed 14 September 2024).
10. Tansman A. (1967). Autoprtrait par Alexandre Tansman, *Journal musical français*, Paryż, maj 1967, p. 24, in: Granat-Janki A., *Forma w twórczości instrumentalnej Aleksandra Tansmana*, Wrocław 1995, p. 170.
11. Wolf W. (2002). *Intermediality Revisited Reflections on Word and Music Relations in the Context of a General Typology of Intermediality*. In: *Word and Music Studies: Essays in Honour of Steven Paul Scher and on Cultural Identity and the Musical Stage*, Amsterdam – New York, p. 22, 28.

On the Artistry of Generative AI and the Case for Surre-AI-lism

Agnieszka Mensfelt, PhD.
Department of Computer Science, Royal Holloway, University of London, Egham, UK
e-mail: agnieszka.mensfelt@rhul.ac.uk

Abstract

The rapid evolution of generative AI tools, such as language and text-to-image models, has democratized access to artistic creation on an unprecedented scale. These models—trained on vast amounts of human-generated data—raise a spectrum of questions, ranging from environmental to ethical and legal concerns. Among these, a question as old as art criticism lingers: “Is this art?” This work proposes an answer by revisiting and adapting the distinction between decorative and fine art. In this framework, the term "decorative" is not used pejoratively; rather, both categories are treated as neutral descriptors to classify generative AI outputs based on intent and purpose. Innovation, characterized by the novelty of concept and/or form that deliberately departs from standard outputs, is identified as a key factor in this distinction. Through this lens, generative AI is viewed as a tool whose default creations—based on the replication of the familiar—fall into the category of decorative art. However, these models also possess the potential to create works that can be recognized as inventive "fine" art. Subsequently, the paper proposes a method for probing the creative capacity of text-to-image models, referred to as *surre-AI-lism*. This approach leverages the principle of operation of the models—representation learning and approximate retrieval—to achieve visually divergent outputs. The models' training data can be considered a form of collective consciousness. Therefore, drawing inspiration from the Surrealists' exploration of the human unconscious, generative AI tools can be

used to probe the collective unconscious. The accompanying images demonstrate that aesthetic novelty can emerge from the unguided synthesis of distinct styles, themes, and the visual interpretation of abstract sentences.

***

"*There is (...) realm of human activity which it is extremely difficult to duplicate by mechanical means. This is activity relating to artistic creation and aesthetic evaluation. While perhaps not impossible for suitably programmed computers, it will certainly be centuries rather than decades before machines are writing Shakespearean sonnets or Beethoven symphonies.*"

Claude Shannon, 1953

1. Generative AI: sudden innovation

Claude Shannon, best known for his contributions to information theory, was one of the founding fathers of the digital age. In a 1953 Bell Labs memorandum, *The Potentialities of Computers* [9], Shannon considered the possible capabilities and limitations of computers. He suggested that artistic creation and aesthetic judgment might remain areas where humans would retain an advantage for the longest time. Since then, many of Shannon's predictions on the future of computing have materialised. However, do recent achievements of generative AI invalidate his predictions on machine artistry? Or is the artistry of generative AI illusory and derivative of human creativity?

This question remains under debate. The disruption brought by the advent of large language models (LLMs) and other generative AI tools to human creativity—as well as to other domains of life—was sudden and unforeseen. On the one hand, generative AI has democratized art creation, allowing anyone with access to the Internet to produce a piece of art with a single click, opening new possibilities in areas such as art therapy [12] and education [2]. On the other hand, generative AI faces challenges on a scale unprecedented in traditional art media: from the high environmental costs of computation to complex questions of authorship [3,8], as creations are often derived from extensive datasets of human work. Moreover, it is burdened with issues of bias embedded in training data, the risk of AI-generated content contaminating future training datasets [11], and the potential for significant job displacement in the creative industries [3]. And, as with all innovations, generative AI raises a timeless question: Is this truly "art"?

This paper aims to address this final question, focusing on text-to-image generative tools as a model example. These tools have ignited considerable debate, with AI-generated works winning awards in competitions traditionally reserved for human-created art [6] (and vice versa [4]). While they bring ethical, legal, and environmental challenges, such models are likely here to stay, which implies the need for a tailored framework for evaluating their outputs. Such a framework could help assess their limitations and risks, as well as enhance the exploration of their creative potential.

The structure of this paper is as follows: to develop such a framework, a general overview of how text-to-image models operate is presented. Following this, the frequently drawn analogy between generative AI for images and the invention of photography [3] is examined. Next, the proposed framework for evaluating generative AI art is outlined. Finally, an approach aimed at exploring the creative capacity of these models, referred to as *surre-AI-lism*, is introduced.

2. Text-to-image models

Text-to-image models [1] are machine learning models designed to generate images from text prompts provided by users. In general, machine learning algorithms are trained on large datasets to learn and generalize patterns. For text-to-image generation, these models are typically trained on vast datasets containing images paired with corresponding textual descriptions, often ranging from hundreds of millions to billions of samples.

By learning from these pairs, the model forms associations between words and visual elements. For example, it may associate the word "sun" with a circular shape emitting rays, typically represented in shades of yellow, orange, or pink. Importantly for the approach introduced in Section 5, these models learn not only the visual representation of objects but also various aesthetic styles, such as "photorealistic," "cartoon," or "cubist painting," along with styles linked to specific artists. By being exposed to proper image data, the model gains the ability to reproduce distinct features, stylistic nuances, and thematic elements.

The generation process with a trained model begins when a user provides a description, such as "a kitten in a basket." Typically, the text is first converted into a numerical representation through a technique known as embedding, which captures the semantic meaning and context of words and phrases. The model then uses this embedded input to generate an image, often starting with random noise and progressively refining it to match the description, with the exact process varying based on the model's architecture.

The ability to generate images almost instantly and without any creative effort or intent raises questions about whether AI-generated images can truly be considered art. However, to fully address this question, it is essential to consider the principles underlying the training of these models.

The training process can be understood as a form of data compression, where the model learns to encode the information contained in the training data. For instance, to learn the visual representation of a cat, the model learns the statistical distribution of features common to cats, capturing patterns that define their visual characteristics. This process results in a distilled representation of *catness*. Due to the randomness in the generation stage—

which can be interpreted as an approximate retrieval of the information stored in the training dataset—each generated image is novel in the sense that it did not exist before. At the same time, by default, the image will contain little conceptual and visual novelty, regressing towards an "average" cat picture (Fig. 1). Thus, the model's creative capacity is limited by its training data and generation principles.

3. The camera analogy

To contextualize text-to-image models within the realm of visual media and assess their impact on art, a common analogy is drawn with the advent of photography [3]. Let us examine it in light of the operational principles of these models discussed earlier.

The invention of the camera in the 19th century made it possible to capture realistic images of objects— a task that, until then, had been fulfilled by painters. This shift liberated painters from the quest for realism, stimulating the exploration of expressive and abstract forms of art. Photography itself was initially valued for its ability to document reality but was not recognized as "true art" due to its "mechanical" nature. However, over time, photographers began pushing the boundaries of the medium, elevating photography into a creative and expressive art form. Among the most radical innovators was Man Ray. His works, such as the iconic *Le Violon d'Ingres* and his rayographs—images created directly on photosensitive paper without the use of a camera—expanded the affordances of the medium. These efforts, along with those of other early 20th-century photographers, helped establish photography as a legitimate art form.

The rise of generative AI tools in the 21st century enables the automatic, instantaneous creation of images that do not require a creative background or specialized skills. Similar to the early days of photography, there is ongoing debate over whether such outputs can be considered "art." Like photography, generative AI can serve as a creative aid, allowing for rapid concept sketches and ideation, thereby freeing up creative resources. Reflecting on the history of photography and other innovations, one might speculate that generative AI will eventually be recognized as "legitimate" art form.

However, the analogy between photography and generative AI has its limitations due to differences in underlying mechanisms and the nature of creative control. In photography, the author retains creative control over the entire process, whereas with generative AI, even the question of authorship is unclear, and the user exercises limited creative control. Moreover, due to its dependence on the training set and generation principles, using generative AI as an idea generator carries the risk of inhibiting creativity through design fixation [5], where exposure to a single solution limits the exploration of alternative, potentially superior ideas. Another significant difference is that generative AI tools are developing and becoming widespread at a much faster pace. This motivates a dedicated evaluation framework, to both contextualize text-to-image models within existing visual media and assess them based on their specific characteristics.

4. The evaluation framework

Historically, many competing aesthetic evaluation theories have emerged, each attempting to address what establishes the value of a work of art and its status as “true art.” In *The Role of Theory in Aesthetics* [10], Morris Weitz examines several classical theories, such as formalism, emotionalism, intuitionism, and organicism, while criticizing rigid definitions and advocating for a more open-ended approach to understanding art as an evolving concept. The emergence of digital and generative art, which expands the understanding of what art can be, exemplifies and supports this more flexible approach.

The framework presented here adopts a similarly open-ended, descriptive approach. Rather than assigning or denying the status of “art,” it aims to contextualize and facilitate discussion around generative AI art, while also enhancing its further exploration. This framework revisits and adapts the historical distinction between “decorative” and “fine” art. To better understand how outputs from text-to-image models relate to this distinction, let us consider well-known examples from each category.

One iconic example of “decorative” art is the depiction of dogs playing poker, originally painted by Cassius Marcellus Coolidge. The later series, created in the early 1900s for cigar advertisements, became a cultural meme, widely replicated in various forms and used to decorate homes and venues (and ultimately inspiring creative reinterpretations). By contrast, a famous—or infamous—example of “fine” art is Kazimir Malevich’s *Black Square*, the “zero point of painting”—a radical departure from classical painting, with a rich conceptual background. Now, let us compare the cat image from Fig. 1 to these examples. This image aligns more closely with the first category. Similar to the various incarnations of anthropomorphized dogs, it is an easily produced instantiation of familiar imagery with numerous practical and decorative applications.

At the same time—similar to painting and photography—generative AI can be used to expand beyond decorative replication. Drawing on historical examples, Malevich’s geometric abstractions introduced innovation through a radical departure from representationalism, while Man Ray's rayographs, discussed in the previous section, were groundbreaking in their unconventional use of photosensitive paper and experimentation with randomness. Likewise, innovation in generative AI may emerge from breaking away from predictable, average-regressing outputs. For pre-trained models, this might involve steering the generation process to creatively combine learned representations in novel ways, resulting in emergent qualities that diverge from the default. For new model development, this would involve designing architectures capable of producing more conceptually and visually divergent outputs.

5. Surre-AI-lism

One inspiration for broadening the exploration of the output space of text-to-image models may be drawn from Surrealism. The vast training dataset of these models can be understood as a collective visual consciousness—a repository of styles, symbols, and motifs. A standard, literal prompt can be seen as

Figure 5: Beksiński x Bosch (DALL·E 3).

Figure 3: Beksiński x Pop Art (DALL·E 3)

accessing this collective consciousness for familiar representations. At the same time, inspired by the Surrealists' exploration of the unconscious, ambiguity and randomness can be utilized to probe the collective unconscious, revealing unexpected connections and juxtapositions. The output of this approach has the potential to be conceptually and visually novel. One

Figure 4: Beksiński x Schiele (DALL·E 3)

Figure 2: Beksiński x Lissitzky (DALL·E 3)

method for achieving such ambiguity is to provide a prompt without a literal visual interpretation. Two types of such prompts are described in more detail in the following sections. The model used to create all illustrations presented in the following sections is DALL·E 3 [7].

5.1 Blending the styles

Figure 7: “A Throw of the Dice… ” (DALL·E 3)

The first type of prompt involves requesting an image in the style of two or more artists or art movements. Here, the prompt specifies only the style, not the content, which means the model independently selects themes and stylistic elements associated with the requested artists, combining them freely into a single image. For example, let us consider a blend of the styles of Zdzisław Beksiński—a Polish dystopian surrealist—with other artists. Each of the presented images was generated solely by specifying a painting in the style of two given artists, without defining the specific content of the image.

Fig. 2 shows a blend with Hieronymus Bosch; the image includes a fantastic creature characteristic of Bosch, combined with greater depth and sense of desolation typical of Beksiński. The image in Fig. 3 merges Beksiński's motifs with the vibrant colours of Pop Art. Fig. 4 presents an image where Egon Schiele's influence heightens the drama and expression of a dystopian scene. Finally, Fig. 5 blends El Lissitzky's constructivist

Figure 6: “A Throw of the Dice… ” (DALL·E 3)

architectural landscapes with surreal imagery.

This approach may involve two or more artists and styles, resulting in unexpected blends. By combining distinct and unique styles, a new stylistic quality has the potential to emerge.

5.2 Abstract sentences

Another way to explore the model's creative limits is to use an abstract sentence without any clear visual representation. For example, consider *Un coup de dés jamais n'abolira le hasard* (*A Throw of the Dice will Never Abolish Chance*), the title of Mallarmé's avant-garde poem. Illustrating this visually ambiguous phrase would be challenging even for human artists, given its lack of concrete imagery. Sample artistic interpretations generated by a text-to-image model are presented in Figs. 6 and 7. In these cases, the prompt only included a request to artistically illustrate the provided quote, without specifying what should appear in the image.

6. Conclusions

This paper addressed the question of the artistry of generative AI outputs, with a focus on text-to-image models. An analogy with the advent of photography was examined, and a framework adapting the distinction between decorative and fine art was proposed. This framework reconciles opposing views on generative AI art, recognizing it as a tool with a dual nature—similar to painting and photography—that is capable of both reproducing the familiar and enabling creative exploration.

By focusing on the underlying mechanisms of these models, we can better understand their limitations and recognize the potential risks of stifling human creativity. However, this awareness can also guide their inventive and exploratory use, thereby expanding our creative toolkit. Ultimately, like any tool, generative AI can be used for both constructive and destructive purposes, just as a knife can be used to slice bread or take a life. It is up to us to decide whether we will use this "knife" to kill our creativity or to carve new paths in the landscape of imagination.

References

[1] Bie, F., Yang, Y., Zhou, Z., Ghanem, A., Zhang, M., Yao, Z., ... & Song, S. L. (2023). RenAIssance: A survey into AI text-to-image generation in the era of large model. *arXiv preprint* arXiv:2309.00810.

[2] Dehouche, N., & Dehouche, K. (2023). What's in a text-to-image prompt? The potential of stable diffusion in visual arts education. *Heliyon*, 9(6).

[3] Epstein, Z., Hertzmann, A., Herman, L., Mahari, R., Frank, M. R., Groh, M., Schroeder, H., Smith, A., Akten, M., Fjeld, J., Farid, H., Leach, N., Pentland, A., & Russakovsky, O. (2023). Art and the science of generative AI: A deeper dive. *arXiv preprint* arXiv:2306.04141.

[4] Hern, A. (2024, June 13). Photographer takes on the machines in AI competition – and wins. The Guardian. www.theguardian.com/artanddesign/article/2024/jun/13/photographer-takes-on-the-machines-in-ai-competition-and-wins

[5] Jansson, D. G., & Smith, S. M. (1991). Design fixation. *Design studies*, 12(1), 3-11.

[6] Kuta, S. (2022, September 2). An AI-Generated Artwork Won First Place at a State Fair. Artists Aren't Happy. *Smithsonian Magazine*. www.smithsonianmag.com/smart-news/artificial-intelligence-art-wins-colorado-state-fair-180980703/

[7] OpenAI. (2023). DALL·E 3: Improving text-to-image generation with natural language understanding. *OpenAI*. cdn.openai.com/papers/dall-e-3.pdf

[8] Piskopani, A. M., Chamberlain, A., & Ten Holter, C. (2023, July). Responsible AI and the arts: The ethical and legal implications of AI in the arts and creative industries. In *Proceedings of the First International Symposium on Trustworthy Autonomous Systems* (pp. 1-5).

[9] Shannon, C. E. (1993). The potentialities of computers. In N. J. A. Sloane & A. D. Wyner (Eds.), *Claude Elwood Shannon: Collected papers* (pp. 691-695). IEEE Press.

[10] Weitz, M. (2017). The role of theory in aesthetics. In *Aesthetics* (pp. 512-517). Routledge.

[11] Xing, X., Shi, F., Huang, J., Wu, Y., Nan, Y., Zhang, S., ... & Yang, G. (2024).

When AI Eats Itself: On the Caveats of Data Pollution in the Era of Generative AI. *arXiv preprint* arXiv:2405.09597.

[12] Yoo, D., Kim, D. Y., & Lopes, E. (2023, September). Digital Art Therapy with Gen AI: Mind Palette. In *2023 11th International Conference on Affective Computing and Intelligent Interaction Workshops and Demos (ACIIW)* (pp. 1-3). IEEE.

The Resonant Mind-Field of Embodied Consciousness

Assoc. Prof. Manuel Báez, B. Arch, M. Arch, MRAIC, Architect (New York).
Azrieli School of Architecture and Urbanism, Carleton University, Ottawa, Canada
https://architecture.carleton.ca/archives/people/manuel-baez
e-mail: manuelbaez@cunet.carleton.ca

Abstract

"You who speculate on the nature of things, I praise you not for knowing the processes which Nature ordinarily effects of herself but rejoice if so be that you know the issue of such things as your mind conceives." Leonardo Da Vinci

Within the current scientific and technological realms, our extraordinary, unprecedented, and ever-expanding capability of manipulating matter, energy and in-formation has provided highly impressionable and awe-inspiring conceptualizations and experiences. The virtual World Wide Web and Cloud, the latest upgraded version of digital devices, and, of critical importance to the theme of this conference, the potential of Artificial Intelligence (AI), offer what seems to be almost limitless possibilities and fluid access to new realms of manipulated and seamlessly structured *in-formation*. These latest devices, innovative developments and experiences will continue to provide highly impressionable sources and uncharted areas, inspiring our vibrant metaphoric speculations on the nature of matter, the universe, and what is humanly possible. At this critical and unprecedented stage in our capabilities through these virtual innovations, Leonardo da Vinci's words of wisdom remind us to reflect on our human nature as revealed by these developments and by *"such things as our mind conceives*."

Our ability to be consciously aware of ourselves and our manifestations offers the possibility for introspective reflection regarding the nature of our intellectual, creative, and imaginative capabilities. Introspective self-analysis is a very insightful way of acquiring perceptive and intrinsic knowledge regarding the inner workings of the theatre of the mind, its creations, and its relationship to the material and ethereal world. We are embodied with and inextricably embedded within this fluent, interactive,

and generative matrix or *field*, which has been the source of our vast array of cultural achievements, artifacts and motifs. Therefore, we need to reflect on our rich human nature as revealed by current innovations and other equally relevant mental conceptions from the vibrant cultural past.

Self-awareness and reflective intelligence have yet to or may never be achieved by Artificial Intelligence (AI). It must be emphasized that the technological, cultural and intellectual achievements from our embodied and self-reflective consciousness have been the data source for AI's development and apparent breakthroughs. Our latest understanding and speculative theories regarding consciousness and how it could be generated are critical to the potential and limitations of AI. Current insights are offered regarding how our embodied awareness is generated through our understanding of the nature of vibrant and fluent fields and how this may or may not relate to AI's potential and any proposed theory of consciousness.

Ongoing related research and projects inspired by the generative potential of the concept of fields are also provided as examples of what is possible through a reflective working process that responds to and engages these fertile ideas.

1. Introduction

The technical concept of invisible "*fields*" and vibrant fields of forces or energy was first revealed in physics in the mid-nineteenth century. Since then, our comprehension of the properties and generative potential of these fields has been widely expanded into numerous areas that now include theories of consciousness. Similar to our intellectual achievements, cultural artifacts and motifs, these current conceptions and speculations regarding consciousness are possible through our embodied self-awareness and potential as explorative seekers, engaged participants, experiential interpreters, and reflective dreamers. Thus, the commonalities, parallels and resonant self-similarities between these human conceptions and speculations can provide insightful and deeper self-awareness to and about their source, as offered and unveiled throughout history by a myriad of vibrant motifs, stories and metaphors, mirroring our very nature. Perhaps the reflective lessons that modern advances can reveal are that those ancient intuitive traditions (possibly as a much more relevant and specific version of AI as *Ancient Intelligence)*, through their spiritual, artistic, and cultural motifs, can offer equally as profound and comprehensive insights, or perhaps even more so, than those that may be possible to unveil through modern technological progress.

2. From Lines of Force to A Fluent Field of Forces

As a technical idea in physics, the concept of a "field" was first conceived in the middle of the nineteenth century. Previously, related notions were gradually developed in response to, for example, the intriguing effects of magnetism, the perplexing phenomenon of the ebb and flow of tides and ocean waves, the orbits of planets, and the effects of forces such as forces of attraction between bodies. These evolving notions regarding discernible "*zones of influence"* and *"action at a distance"* were related to the eventual concept of a "field" with varying degrees of direct relevance. Eventually, the definitive chapter in the origins and evolution of the technical concept of a "field" and its properties would come from the experimental work with electric and magnetic phenomena by the self-taught English scientist Michael Faraday and, shortly thereafter, the related and integrative work of the Scottish physicist and mathematician James Clerk Maxwell.

Regardless of his lack of formal education, Michael Faraday (1791-1867) was a highly imaginative and incessant experimentalist in science. His research dealt with all aspects of electricity and magnetism in physical phenomena, with his formidable intuition and broad vision making up for his lack of formal training in mathematics and the sciences. As we still are, intrigued by the manifested pattern of curving *lines of force* when iron filings are placed around a magnet, Faraday would eventually envision this as revealing a dynamic medium independent of the specific source being worked with. The dynamic properties of the experimentally revealed *lines of force* would be interpreted as a theoretical *field of force* that would subsequently and intrinsically underlie the discipline of physics throughout the twentieth century.

As a result of this far-reaching achievement, Faraday conceived space and matter as undifferentiated and filled with the highly dynamic and versatile activity of these forces through a continuous and ever-present process. His lack of formal training was highly advantageous, freeing and allowing his intuition and imaginative vision to be unrestrained from the mathematical understanding and formal restrictions of the period. Faraday's insights revealed the profoundly connected relationship between electric and magnetic phenomena, with the two as features of one field, referred to as the *electromagnetic field (EM field).*

Faraday could not, nor was he concerned with expressing his new concepts through rigorous mathematical formalism. Again, this freedom was the reason for his successful analysis and interpretation, which were rigorously based on the experimental results. As remarkable as this qualitative, intuitive, and speculative interpretation was, the possibility for further advancement through the quantitative and precise requirements of modern physics would eventually be provided by James Clerk Maxwell (1831 – 1879). In 1856, his article titled "*On Faraday's Lines of Force*" was published, providing the rigorous mathematical interpretation that would confirm Faraday's vision and the reality of the *field of force*.

2.1 Resonant Metaphors and Analogies

Both Faraday and Maxwell relied on appropriate and convenient analogies as the models that would guide the progress of their work. Thus, Maxwell imagined a fluid as the medium that carries the lines of force, comparable to the *ether* that had been conceived as a conceptual requirement of any such theory. He did not fully believe in the actual existence of this hypothetical medium. Regardless of this fact, this suitable metaphor allowed for Maxwell's rigorous mathematical analysis, with the flow lines of fluid mechanics imagined as comparable to Faraday's lines of force. Metaphorically, the singular fluid-field of fluid mechanics was conceived as the ether comprised of the two distinct, interacting, and interrelated fields of *electromagnetism*. Additionally, Maxwell imagined the lines of force of both fields as pliable and elastic tubes that would easily respond to and represent the flow speed of the ether and be comparable to the veins and arteries of the human body.

Eventually, the mathematics representing and combining electricity and magnetism as EM fields would reveal that they travel as undulating waves through space at the speed of light, leading Maxwell to realize that light is a form of electromagnetic radiation that is visible to us. In 1872, Maxwell published the summation of his research in two volumes titled *A Treatise on Electricity and Magnetism.* His formidable integration and conception would eventually lead to the unveiling of the full implication and diversity of electromagnetism as an immense spectrum of diverse EM fields of propagating waves of energy that would include a far greater number than what is visible to us (i.e., radio, x-rays, ultraviolet, etc.). According to none other than Albert Einstein, "*This change in the conception of reality is the most profound and fruitful one that has come to physics since Newton*." [1] The mathematics developed by Maxwell would eventually become the foundational basis for modern technologies and electronic devices, leading to global computing and communications.

2.2 A Fluent and Resonant Medium of Radiant Energy

Through the experimental work of Faraday and the broader integration by Maxwell, the incredible and expansive achievement they initiated regarding the nature of electromagnetism has led to the latest comprehension and speculations regarding the potential contained within what has been interpreted as a highly sensitive and fluent medium of radiant energy, the EM field. Through its immensely diverse radiant potential, all the interrelationships and interactions of resonant waves and their possible frequencies are revealed as the manifestations that make up the universe. It is interesting and revealing that the roots of the word *universe, as* uni-verse, mean *one-verse* or *one-whole*.

"*It is thought-provoking*," states theoretical physicist Matt Strassler, "*that all these waves obey the same speed limit and that their fields are found everywhere in the cosmos. It hints that the fields may all be melded, along with empty space, into some kind of unified edifice, or framework, or…well, universe*." [2]

Through its rigorous and focused discipline, has science revealed the Ancient Intelligence referred to previously regarding the resonances between ancient intuitive traditions and modern science?

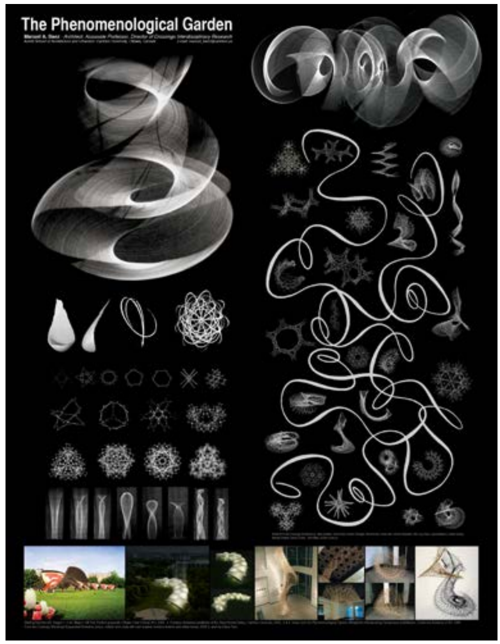
We are now speculating and exploring how this fluent medium of radiant energy that surrounds and is within us relates to our resonant mind-field of embodied consciousness. Our known and hidden senses rely upon and are highly sensitive to vibratory phenomena, mirroring the sensitivity of this all-encompassing generative medium of energy. The diversity and coherent properties of *resonance* play a vital role in how the dynamics of *in-formation* is encoded within the rhythmic undulations of energy waves and deciphered for interpretive awareness to be possible. This should remind us of the expressive power, immense scope, profound depth, and cultural diversity of music and how this highly potent, evolving, and primal manifestation from Ancient Intelligence profoundly mirrors the same source that surrounds and is embodied within us. Regarding the significant role of *resonance* within this environment or "cosmic ocean," Matt Strassler states: "*For a species that so often puts music at the center of celebrations, worship, and pleasure, this is deeply appealing. Like any musical instrument, the cosmos resonates with a pattern of frequencies … The universe rings everywhere, in every thing*." [3]

It is highly doubtful that Artificial Intelligence could ever achieve this kind of experiential self-awareness, reflective consciousness, and expressive potential. We should remember that it is an "it from bits" and that such an "it" can not be from being.

3. Related Projects

The following work has been inspired by and responds to the possibilities offered by the above-mentioned themes.

3.1 The Phenomenological Garden



© *M. Báez, The Phenomenological Garden,* poster of related projects.

The above poster was initially exhibited as part of the 2010 XIII Generative Arts Conference held at the Politecnico di Milano University. It summarised the evolution of the research and related process-work at that time and offers an appropriate overall perspective on the interrelationships between the research themes and how this relates to the theme of this paper. Overall, the research and work are still considered as originating from explorations of the phenomena surrounding and lurking within us in the *Phenomenological Garden*. I will only refer to the most relevant work related to this paper.

Referring to the poster layout, the upper right and left images, including the central right image of an undulating line, are from a project considered as the initial source of the work that subsequently followed. A ceiling-mounted potter's wheel was used to spin a hanging string into sequential wave-formations (shown on the lower-left side). Stable formations would gradually transform, becoming highly turbulent at higher frequencies. Scientific investigations and speculations regarding the effects and properties of dynamic phenomena and circular motion inspired this project. Its title, *Ariadne's Thread/Rumi's Ocean*, metaphorically bridges or connects related Western and Eastern traditions, an early sign of an underlying theme, highlighting the parallels and resonances between these two traditions or worldviews (referred to previously regarding modern science and ancient intuitive traditions). *Ariadne* is the mythological Greek guide to the labyrinth of chaos and the individual life. Jalal al-Din Rumi is the great Persian mystic poet of the thirteenth century and the creator of the whirling circular dance of the Mevlevi dervishes.

The potter's wheel was also used to cast wax forms, generating them within the virtual vortex made possible through the wheel's spin. Two views of one of these casts are shown in the middle-left side and, to their right, are other views of the spinning string. The sculptural and architectural potential of the forms generated through the above procedures, including their inspirational sources, led to various explorative directions.

Through a hands-on process, the dynamic potential of basic geometric relationships was explored through a series of modular cells, or units, and their combinations into cellular membranes or, now more specifically relevant, cellular *fields* that allowed for intuitive discovery to occur regarding the dynamic interrelationships between form, structure, and process.

Form-generating possibilities were explored by folding the possible 2D patterns and by assembling the cells and their combinations as flexible 3D fields using wooden dowels and rubber bands. The central-lower left side shows basic cellular units and their emergent complex assemblies. The flexibility of the joints, cells, and fields allowed various forms and structures to be generated through the emergent, transformative and self-organizing properties inherently encountered through the working process. This in-forming potential and its fluent diversity are displayed throughout and within the suspended undulations of the string and the colour images on the lower right.

Related architectural proposals are shown on the lower left.

3.2 The Light Keeper

© *M. Báez, The Light Keeper* installation.

The *Light Keeper* is a permanent ceiling installation *presented at the* 2017 XX Generative Arts Conference in Ravenna, Italy. The installation is on campus at the Centre for Indigenous Support and Community Engagement at Carleton University in Ottawa, Canada. Housed within the Centre is the Indigenous Student Centre, named Ojigkwanong, meaning "morning star" in Algonquin, as a tribute to the renowned Anishinaabe Elder Grandfather William Commanda. The installation was designed and built by the author and with architecture students in consultation with the Centre's acclaimed architect, Douglas Cardinal. The design is based on the theme of *light*, symbolic of knowledge and on the fact that the Carleton campus is on traditional, unceded territories of the Algonquin nation.

The ceiling installation is envisioned as an interconnected woven assembly and support structure made of bands of birch plywood with coloured wire-mesh highlights at significant locations. The design was generated from the inherent properties of the materials with no pre-conceived design other than what emerged from the haptic exploratory process. The dimensions and configurations extracted from this generative process led to design possibilities, modified accordingly as the project evolved and in response to the specifics of the interior space. The design originates from one singular band of thin birch plywood with its ends overlapped to create a circular frame, recalling the basic frame of the traditional native drum. The diameter of this circular frame is the minimum allowed by the pliability of the birch plywood. Accordingly, the essential design elements are the circle, the circular drum, its essential frame, and the interwoven "sounds" emanating from this source.

Overall, the ceiling is composed of modulating patterns within the collaborative weave, with the configurations and interconnected parts envisioned as symbolic metaphors for the Centre's mission and the worldview of Indigenous Peoples.

3.3 Diluvio: Teatro delle Ombre (Deluge: Theatre of Shadows)

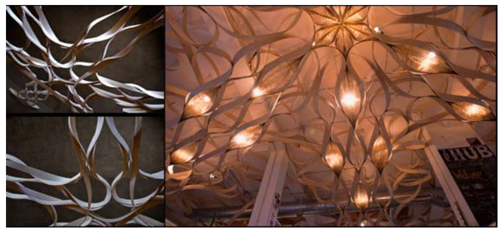
© *M. Báez, Diluvio: Teatro delle Ombre.*.

Conceived for the 500th anniversary of Leonardo da Vinci's death in France on May 2, 1519, *Diluvio: Teatro delle Ombre* (Deluge: Theatre of Shadows) offered an interactive and immersive exhibition/installation as part of *Cinquecento: Carleton Celebrates Leonardo da Vinci,* the 2019 year-long commemorative initiative at Carleton University. The project culminated the series of *Diluvio* installations from the author's Crossings Interdisciplinary Research Collective Workshop. The projects were inspired by Leonardo's *Deluge* drawings and his reflections regarding the complementary interrelationships within nature as revealed by his studies of the flow of water, air, light, shadows, and energy. The rich complexity of such phenomena is revealed and experientially encountered by working with the highly pliable properties of woven wire-mesh membranes when folded into the classic Miura-ori tessellation pattern. The inherent attributes of this fluent field and the sculptural work emerging from the working process are revealed through evocative shadow projections activated by the public within the unlit exhibition space. Inspired by Plato's *Allegory of the Cave,* the immersive experience offered a deluge of self-activated shadow projections as a way of stirring, triggering, and thus revealing the highly resonant, fertile and imaginative potential lurking within the *theatre of the mind*. Leonardo offers us the following advice regarding this versatile, evocative, and reflective process of perception:

"*Principles for the Development of a Complete Mind: Study the science of art. Study the art of science. Develop your senses - especially learn how to see. Realize that everything connects to everything else*" [4]

Thus, the immersive experience encouraged participants to introspectively reflect upon and interpret the elusive and evocative forms, shadows, and resonant patterns they encountered as a way of stimulating and expanding their imagination while gaining revealing insights into the inner workings of the human mind and its conceptions.

3.4 Resonant Currents I and II



© *M. Báez, Resonant Currents I.*

© *M. Báez, Resonant Currents II.*

The first permanent *Resonant Currents* project was conceived in response to an invitation in 2013 to propose a permanent installation to celebrate the start of HUB Ottawa, one of several non-profit co-operative organizations located worldwide in major cities. This global network of organizations is focused on community engagement initiatives and relationships that address global, social and environmental challenges. *Resonant Currents II* was initiated through a similar invitation in 2023. By then, ImpactHub Ottawa was in a new location and preparing to celebrate the 10th anniversary of the organization. Both installations, similar to *the Light Keeper*, incorporated bands of birch plywood with wire-mesh highlights at significant locations within the space. The weaving process was metaphorically envisioned as an interconnected network and support structure, mirroring ImpactHub's vision and Canada's ethnic and cultural diversity. The projects were inspired by the parallels between the underlying themes of the research and those of the global and local initiatives of the organization. Thus, as an integrative network and support structure, the installations mirrored their environment's open concept and inclusiveness.

Inspiration also came from global commonalities, connections and their resonant expressions through human cultural artifacts (i.e., weaving traditions, calligraphy, basketry and mosaic patterns). As in previous work, both projects were conceived, developed and installed with students from the Crossings Interdisciplinary Workshop. Their cultural diversity was celebrated as critical thematic contributions to the overall composition. Overall, these installations, as woven integrated networks, reflect their material capabilities and properties while simultaneously responding and adapting to the specifics of their setting.

4. Conclusion

The themes and ideas addressed above, through ongoing research and projects, continue to be developed and expanded upon. It has drawn inspiration from creative individuals and developments throughout history and the incredibly fertile yet tenuous time we find ourselves in. The summary of the related achievements of both Michael Faraday and James Clerk Maxwell had profound implications beyond their expectations, far into the 20th century and what is currently being theorized and made possible. Essentially, they revealed the vital role that vibration and related phenomena play in the universe at both the macro and micro scales. The underlying principles that this has brought forth regarding fundamental interconnected resonance were previously known and explored through the perennial traditional wisdom of many cultures. The *knowing of* and deeper wisdom, initially scorned and discarded by the focused scientific trajectory towards *knowledge* and truth, has re-turned to engulf the foundations of science. The Italian-American physicist, engineer, and inventor Federico Faggin offers the following insights regarding *knowing* and *knowledge* with much relevance towards computers, AI, awareness, and consciousness:

"*In the Italian language there is a significant difference in meaning between conoscere and sapere, two distinct types of knowing, even though in common usage the two words may often be synonyms. Sapere refers to symbolic knowledge, while conoscere refers to semantic knowing, i.e., getting the meaning of the symbolic information."*

"A computer can manipulate symbols without understanding anything, i.e., it can have knowledge (sapere) but cannot

have knowing (conoscere). We can have both symbolic knowledge and experiential knowing (comprehension, meaning)." [5]

At the heart of this distinction also lurks the difference between information and in-formation, between a form or shape and forming and shaping, and finally, between nouns and verbs. On one side, there is the realized form or shape, and on the other, we have that which continuously generates, propagates and maintains the forming and shaping. This process has inspired the research and works-in-process summarized above, inherently and experientially addressing this fundamental paradox that's encountered when contemplating the relationship between substance and form, subject and object, as well as unity and multiplicity.

5. Notes

1. Einstein, Albert (1931).
2. Strassler, Matt (2024)
3. Ibid (p. 247)
4. MacCurdy, Edward (1939)
5. Faggin, Federico (2024)

6. References

Anastopoulos, Charis. *Particle and Wave: The Evolution of the Concept of Matter in Modern Physics*. Princeton and Oxford: Princeton University Press, 2008.

Currivan, Jude. *The Cosmic Hologram: In-formation at the Center of Creation*. Rochester: Vermont: Inner Traditions, 2017

Edelman, Gerald M., and Tononi, Guilio. *A Universe of Consciousness: How Matter Becomes Imagination*. New York: Basic Books, 2000.

Einstein, Albert. "Maxwell's Influence on the Evolution of the Idea of Physical Reality." In *James Clerk Maxwell: A Commemorative Volume 1831-1931*. Cambridge: Cambridge University Press, 1931.

Faggin, Federico. Irreducible: Consciousness, Life, Computers, and Human Nature. Washington: John Hunt Publishing, 2024.

Joye, Shelli Renée. The Electromagnetic Brain: EM Field Theories on the Nature of Consciousness. Rochester: Vermont: Inner Traditions, 2020.

MacCurdy, Edward (Trans.). *The Notebooks of Leonardo Da Vinci*. New York: George Braziller, 1939, Vol. I.

McMullin, Ernan. "The Origin of the Field Concept in Physics," Physics in Perspective 4, 2002, P.13-39.

Strassler, Matt. Waves in an Impossible Sea: How Everyday Life Emerges from the Cosmic Ocean. New York: Basic Books, 2024.

Encoding the Divine: Leveraging AI for the Translation and Preservation of Sikh Scripture

Ginger L Franklin, PhD.

Independent Scholar/Phoenix, Arizona, USA
Breakingborderspublishing.com
e-mail: sikhlibraryusa@gmail.com

Abstract

The Sikh corpus of scriptural narratives is a linchpin of cultural heritage and spiritual guidance. Amidst an era where cultural digitization is paramount, this paper introduces a cutting-edge approach to the translation and preservation of these narratives through artificial intelligence. Dr. Ginger L Franklin, with her extensive background in religious studies and linguistic proficiency, spearheads this initiative. This research does not view AI as a mere translator but as a conduit for the generative art of literature, interpreting the multi-faceted layers of Sikh scriptures.

The proposed AI driven methodologies aim to capture the poetic essence, theological depth, and ethical wisdom embodied within Sikh texts. This project employs generative algorithms that respects the delicacies of linguistic and cultural nuances, thus crafting a translation that offers a dynamic and enriched reading experience. Additionally, the paper addresses the moral considerations in the use of AI for sacred texts, advocating for a collaborative AI use approach that involves religious scholars to uphold authenticity and respect.

Dr. Franklin's unique confluence of academic knowledge and practical translation experience is the fulcrum of this interdisciplinary endeavour. "Encoding the Divine" posits a transformative future for religious literature's digitization, one that ensures the continued relevance and enlightening power of these texts across the globe.

My journey into the translation and preservation of Sikh scriptural narratives commenced with an immersive learning experience in Gurmukhi script, supplemented by extensive research with multiple dictionaries and textual analyses.

This foundational stage was pivotal in understanding the complex theological concepts and cultural contexts embedded within the sacred texts. As technological advancements progressed, my methodology evolved, integrating more sophisticated models of AI, including various iterations of Chat GPT, to enhance the translation process.

These advanced AI tools proved instrumental in dissecting the linguistic structure and semantic layers of the scriptures, offering preliminary translations at an unprecedented pace. The iterative process of employing generative algorithms allowed for a nuanced approach that could approximate the poetic cadences and subtleties of the original text, capturing not just the literal meaning but the essence of the spiritual message.

Despite the leaps in AI capabilities, the critical role of human oversight remained apparent. The final stages of translation necessitated a meticulous human touch to edit and clarify the outputs, ensuring that the AI-generated interpretations adhered to the doctrinal accuracy and emotional resonance expected in the faithful rendition of these timeless writings. The synergy of AI and human expertise highlighted the indispensable nature of the translator's insight in bridging the gap between technological potential and the sanctity of religious literature.

As the technology continues to advance, the human element stands as a beacon of integrity and intentionality within the translation process, affirming the collaborative future between AI capabilities and human discernment.

Keywords: Artificial Intelligence, Translation, Sikh Literature, Generative Text, Digital Heritage Preservation, Ethics

I. Introduction

The Guru Granth Sahib is one of the most revered texts in world religions, serving as both the holy scripture of Sikhism and the eternal living Guru for Sikhs (Singh, 2018). Compiled by the Sikh Gurus between 1604 and 1708, it is a comprehensive anthology that includes hymns, poetry, and teachings from not only Sikh Gurus but also saints from different faiths and backgrounds. Written primarily in Gurmukhi script, the text encompasses a variety of languages and dialects, including Punjabi, Sanskrit, Persian, and regional vernaculars.

The Guru Granth Sahib is unique in its status as the final and perpetual Guru, embodying the spiritual authority and wisdom of the ten human Gurus who preceded it. Its compositions address universal themes such as devotion to God, ethical living, social justice, and the unity of humanity, transcending cultural and religious boundaries (Kapoor, 2019).

Despite its universal message, the Guru Granth Sahib remains inaccessible to many due to language barriers. The Gurmukhi script and archaic language present challenges for even native Punjabi speakers, and international audiences face even greater difficulties (Singh, 2018). Traditional translations often struggle to convey the poetic and philosophical nuances, leading to potential misinterpretations or a loss of the text's intrinsic beauty. In the digital age, however, there is an unprecedented opportunity to preserve and disseminate sacred texts globally. Digitization aids in both preservation and interactive engagement with scriptures, but it must ensure that the sanctity and integrity of the texts are maintained (Kapoor, 2019).

The emergence of Artificial Intelligence (AI) has revolutionized many fields, including language translation. Advanced AI models, particularly in natural language processing (NLP), have demonstrated remarkable capabilities in understanding context, semantics, and even generating human-like text (Sinha & Thakur, 2020). Neural Machine Translation (NMT) models have surpassed traditional statistical methods, offering more fluid and contextually appropriate translations.

AI's potential to handle complex linguistic structures and low-resource languages makes it a promising tool for translating texts like the Guru Granth Sahib (Floridi & Chiriatti, 2020). The ability of AI to process large datasets allows it to recognize patterns and subtleties that might be overlooked by human translators.

This paper aims to explore the integration of AI in translating and preserving Sikh scriptures. By leveraging AI as a generative tool, the research seeks to capture not only the literal meaning but also the poetic essence and theological depth of the Guru Granth Sahib. The study underscores the importance of human-AI collaboration to ensure authenticity and respect for the sacred text, while also addressing ethical considerations of using AI in this context and proposing guidelines for responsible application.

II. Author’s Journey

My journey into Sikh scripture translation began with a deep immersion in the Gurmukhi script, developed by Guru Angad Dev Ji, the second Sikh Guru. Language is the vessel of culture and thought, so I dedicated significant time to mastering Gurmukhi’s script and phonetics. This involved studying its letters, vowel signs, and grammatical structures, which are essential for accurate interpretation.

Using multiple dictionaries, such as Bhai Kahn Singh Nabha’s "Mahan Kosh," I engaged in textual analysis to comprehend the layers of meaning within the hymns. This foundational work was critical for understanding the complex interplay of language, culture, and theology in the scriptures. My first published translation was that of Zafarnama. At that time, the use of AI in translation work was not yet available.

In addition to linguistic proficiency, it was vital to immerse myself in the theological concepts and cultural contexts embedded in the Guru Granth Sahib. The text is rich with metaphors, allegories, and references to historical events and cultural practices. I studied Sikh theology, philosophy, and history, exploring works by scholars to gain insight into Gurmat (the Guru’s teachings). Concepts like Naam (the Divine Name), Hukam (Divine Will), and Seva (selfless service) were central to accurate translation and interpretation.

During my initial efforts, several challenges became apparent. The poetic structure of the hymns, with intricate rhyme schemes and rhythmic patterns, was difficult to replicate in another language without losing meaning or emotional impact. Additionally, the archaic language and regional dialects within the text presented obstacles to understanding and conveying the intended messages. This underscored the need for a more sophisticated approach to translation, one that could handle the complexities of the scripture while maintaining its poetic and theological integrity. My translation work continued as the development of AI, specifically ChatGPT, grew.

III. Integration of Artificial Intelligence in Translation

Recognizing the limitations of traditional methods, I turned to artificial intelligence to augment the translation process. Initial experiments with basic machine translation tools were inadequate due to their inability to grasp context and nuance (Sinha & Thakur, 2020). However, the advent of advanced AI models like GPT-3 and GPT-4, developed by OpenAI, marked a significant turning point for translation work overall (Floridi & Chiriatti, 2020). These models, built on transformer architectures, demonstrated an unprecedented ability to understand and generate human-like text. By leveraging them, it became possible to process the complex linguistic patterns of the Guru Granth Sahib.

Rather than using AI solely for direct translation, I envisioned AI as a generative tool—a collaborator in the creative process. The AI has been, as of this day, unable to, on its own, recognize and replicate the poetic and philosophical styles of the scripture, to produce translations that resonate with the same emotional and spiritual depth. A human collaborator is still necessary. Generative algorithms will hopefully be designed to respect linguistic and cultural nuances, and techniques like reinforcement learning and fine-tuning should help the AI adapt to the specific style and vocabulary of the Guru Granth Sahib in the future.

IV. Methodology

The AI-generated translations were produced in segments for manageable review and refinement. I approached the translation process by copying and pasting small sections of the Guru Granth Sahib into the AI, focusing on maintaining both accuracy and readability. Each AI-generated output was compared not only to existing translations but also to the original Gurmukhi script. I paid close attention to how each verse fit within the broader context of the work as a whole, ensuring that the philosophical and theological themes remained consistent throughout the translation.

In order to refine the output, I would request an exact, word-for-word translation from the AI to fully understand the nuances of the original text. This approach allowed me to glean deeper meanings from each verse, particularly where the subtleties of the Gurmukhi language offered insights that might otherwise be lost in a purely literal translation. By analyzing these word choices, I could ensure the translation remained faithful to the original's intent.

After reviewing the AI output, I made careful edits, refining the language to preserve theological accuracy, linguistic fidelity, and cultural appropriateness. The translated sections were formatted in a way that aligned with the overall structure of the Guru Granth Sahib, ensuring that the flow and integrity of the scripture were respected. This iterative process involved multiple cycles of AI generation and human editing, with each round of feedback further enhancing the quality of the translation. Through this collaborative approach, the resulting translation reflects both the precision of AI and the thoughtful input of human interpretation.

V. Results and Discussion

The AI-assisted translations showed significant improvement over traditional methods. The translations retained much of the poetic rhythm and stylistic elements of the original hymns while conveying complex theological concepts effectively. Additionally, readers reported a stronger emotional connection to the texts, indicating success in preserving the spiritual impact. Comparative analysis revealed advantages such as contextual accuracy, consistency, and efficiency, although the AI sometimes overemphasized poetic elements at the expense of clarity.

However, challenges included handling archaic language and cultural references that required human intervention. These challenges highlight the need for continuous improvement of AI models and the indispensable role of human expertise in the translation process. The success of AI-assisted translation depends significantly on the proficiency of the translator in the target language.

Translators must possess a deep understanding of both the source and target languages to ensure that the nuanced meanings and cultural contexts of the original text are accurately conveyed. Proficiency in the target language is essential for preserving the essence of the scripture and ensuring that the translation resonates with its intended audience, as opposed to being wrought with elementary grammatical errors as seen in many current translations.

VI. Ethical Considerations

The use of AI in translating sacred texts raises important ethical questions regarding sanctity, cultural sensitivity, and authority (Church & Hestness, 2019). Sacred texts hold deep religious significance, and translating them is not just a technical task but requires reverence and a deep understanding of their spiritual context.

There is a concern that AI, lacking human consciousness, might unintentionally produce translations that fail to respect the sacredness of the text or offend religious sensibilities. Additionally, sacred texts are closely tied to the cultural and historical context in which they were written. While AI can process language, it may miss or misinterpret cultural nuances, leading to insensitive or inaccurate translations.

Authority is another issue. Traditionally, sacred texts have been translated by scholars and religious leaders whose authority comes from years of study and spiritual insight. AI, lacking this authority, poses questions about the legitimacy and trustworthiness of its translations.

Given these challenges, it is essential that AI be used collaboratively with human translators and scholars to ensure translations are accurate, respectful, and faithful to the text's cultural and spiritual significance. Ethical guidelines are necessary to govern AI's role in this sensitive area.

- **Traditional translation without the use of AI:**

One Universal Creator God. Truth Is The Name. Creative Being Personified. No Fear. No Hatred. Image Of The Undying. Beyond Birth. Self-Existent. By Guru's Grace:

Aasaa, First Mehl:

Vaar With Shaloks, And Shaloks Written By The First Mehl. To Be Sung To The Tune Of 'Tunda-Asraajaa':

A hundred times a day, I am a sacrifice to my Guru; He made angels out of men, without delay (Sikhiwiki.org).

- **AI assisted translation:**

There is One Universal Creator. His is the True Name. He is the Creator of All, without fear, without enmity, a timeless form, unborn, self-existent, and realized by the Guru's grace.

Melody of the First Guru, Guru Nanak:

A ballad set to an offbeat rhythm to be played on an Asraj, and including versus written by Guru Nanak, the First Guru.

Verse of the First Guru:

I am a sacrifice to my Guru; a hundred times a day, I humbly offer myself.

He created humans and gods alike. No living thing is untouched by His actions (Asa Di Vaar in English: Published by Ginger L Franklin, PhD).

VII. Conclusion

The integration of AI in translating portions of the Guru Granth Sahib resulted in translations that were remarkably faithful to the original's poetic and theological essence. By leveraging AI's ability to handle vast amounts of text while maintaining consistency in language and structure, I was able to generate translations that honored the rhythmic and devotional qualities inherent in the scripture.

However, the human element was indispensable in ensuring that the deeper, spiritual dimensions of the text were accurately conveyed. The synergy of AI and human expertise proved essential for achieving both the technical accuracy required for a precise translation and the spiritual authenticity needed to respect the sanctity of the Guru Granth Sahib.

This collaborative approach highlighted the strengths and limitations of AI in sacred translation work. While AI excelled at processing complex language patterns, it still struggled with the interpretation of archaic and culturally specific terms that require a deeper understanding of the spiritual and historical context.

The human role was crucial in guiding these subtleties, ensuring that the translation remained connected to the profound meanings of the original text.

Future work in this area should focus on improving AI's ability to handle archaic language, especially when translating ancient scriptures or texts where traditional linguistic forms are prevalent.

Additionally, expanding ethical frameworks for AI use in sacred contexts will be important to maintain respect for the spiritual significance of such works.

This includes developing guidelines that ensure the sanctity and integrity of the texts are preserved, and that the use of AI remains a tool for enhancing understanding rather than replacing human interpretation.

As AI continues to evolve, its potential to contribute to sacred translations could be transformative, but it must always be coupled with human insight and reverence for the text.

IX. Refrences

Church, Kenneth W., and Joel Hestness. "Ethical Considerations for Artificial Intelligence in Sacred Texts Translation." *Ethics in AI*, vol. 8, no. 1, 2019, pp. 15-28.

Floridi, Luciano, and Massimo Chiriatti. "GPT-3: Its Nature, Scope, Limits, and Consequences." *Minds and Machines*, vol. 30, no. 4, 2020, pp. 681-694.

Kapoor, Kavita. "Translating the Sacred: The Role of Translation in Preserving Sikh Scriptures." *Journal of Sikh Studies*, vol. 15, no. 2, 2019, pp. 45-62.

O'Brien, Sharon, and Maureen Ehrensberger-Dow. "Machine Translation Interfaces: A Human Factors Perspective." *Routledge Handbook of Translation and Technology*, edited by Minako O'Hagan, Routledge, 2020, pp. 207-223.

Sinha, Rohan, and Anupam Thakur. "Neural Machine Translation: A Case Study on Low-Resource Languages." *Proceedings of the International Conference on Computational Linguistics*, 2020, pp. 123-130.

Singh, Pashaura. "The Poetics of the Guru Granth Sahib: Preserving Musicality in Translation." *Sikh Formations*, vol. 14, no. 1, 2018, pp. 77-90.

Guru Granth Sahib. Original work published between 1604-1708.

An exploration of AI Text-to-3D generators: Latest Developments and Applications on 3D Printing

University of the District of Columbia

Associate Professor Davide Prete, MA, MFA

Division of Arts and Humanities
4200 Connecticut Avenue NW, Bldg 42, Suite A-08
Washington, DC 20008 USA
www.davideprete.com
https://www.udc.edu/cas/employees/davide-prete/
e-mail: davide.prete@udc.edu

Abstract

Introduction:

The convergence of artificial intelligence (AI) and 3D printing has given rise to innovative techniques that transform textual descriptions into detailed 3D models. This emerging technology, "AI text-to-3D," leverages advanced natural language processing (NLP) and 3D generation algorithms to create intricate meshes suitable for 3D printing. This paper explores the latest developments in AI text-to-3D, focusing on its applications in generating detailed meshes for 3D printing.

Background:

AI in 3D Modeling

AI has significantly impacted 3D modeling by automating complex tasks that traditionally required skilled human intervention. Neural networks, generative adversarial networks (GANs), and other machine learning models have been employed to generate and refine 3D models from various inputs, such as images and sketches.

3D Printing

3D printing, or additive manufacturing, involves creating three-dimensional objects by layering material based on a digital model. The mesh's precision and complexity determine the printed object's quality and functionality. Detailed meshes are crucial for applications requiring high accuracy and fine details, such as medical implants, intricate art pieces, and engineering components.

Objective/Hypothesis:

This wants to be an exploration of the latest developments in AI Text-to-3D.

The first objective is to analyze and compare different approaches to the AI creation process with a deeper understanding of specific generators.

For example, one of the significant advancements in AI text-to-3D is the application of **Neural Radiance Fields (NeRF)**. NeRFs can represent complex 3D scenes by encoding volumetric information into neural networks. Recent research has extended NeRFs to generate 3D models from textual descriptions, producing highly detailed meshes suitable for 3D printing.

NeRF-W, an extension of NeRF, incorporates textual input to generate 3D models with varying levels of detail. This model leverages large language models to parse and interpret textual descriptions, converting them into spatial representations that can be further refined into detailed 3D meshes.

Diffusion models, which have shown great promise in image generation, are now being adapted for 3D model creation. These models incrementally refine 3D structures from noise, guided by textual descriptions. The iterative refinement process allows for generating highly detailed and accurate 3D meshes.

Latent diffusion models operate in a lower-dimensional latent space, making the generation process more efficient while maintaining high detail levels. By incorporating text prompts, these models can create intricate 3D meshes that capture fine details essential for high-quality 3D printing.

Transformers, known for their effectiveness in NLP tasks, are now used in 3D generation. These models can handle long-range dependencies and understand complex textual descriptions, enabling the creation of detailed 3D meshes from text.

Text2Mesh is a transformer-based approach that directly translates textual descriptions into 3D meshes. By training on large text datasets and corresponding 3D models, Text2Mesh can generate highly detailed and contextually accurate 3D objects suitable for 3D printing.

Applications in 3D Printing

AI text-to-3D enables the creation of personalized 3D-printed objects based on user descriptions. This is particularly useful in fashion, jewelry, and consumer goods, where customization is highly valued.

AI-generated 3D models can create patient-specific implants and prosthetics in the medical field. Detailed meshes ensure that the printed objects fit perfectly and function as intended.

Artists and designers can leverage AI text-to-3D to bring their creative visions to life. Detailed meshes allow for creating intricate sculptures and design prototypes that would be difficult to model manually.

Engineers can use AI text-to-3D to prototype complex components and machinery parts quickly. The high level of detail in the generated meshes ensures that the printed parts meet stringent engineering standards.

Methods:

As the first step in my research, I experimented with simple LLM use and reviewed the latest literature to understand the current state of AI text-to-3D. Starting a couple of years ago, I experimented with simple ways to use AI to create 3D Models.

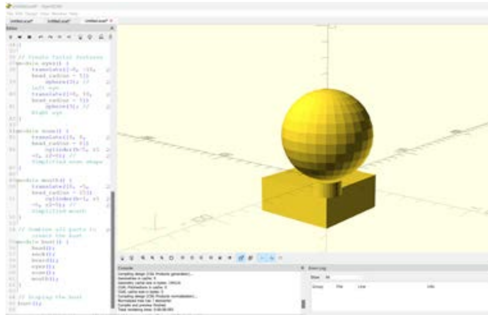
The simplest one was to use an LLM to create a script that can be fed to 3D modeling software such as OpenScad.

Even though this system works best for geometric forms, I tried using it to make a soft organic form, such as a portrait.

I asked, for example, Copilot and ChatGPT to write a script for OpenScad to create a “bust of a bearded man. “Both Copilot and ChatGPT created a script explaining that “OpenSCAD is more suited for parametric and constructive solid geometry than organic shapes. However, it provides a basic framework that can be refined further with more detailed modeling software or techniques.

The script presented a head represented by a simple sphere, a neck by a smaller cylinder beneath the head, and a beard represented by a cube subtracted from the head's lower portion to fit the "face" nicely. Two small spheres were placed to define the eyes.

These facial features were represented by simplified shapes (a small cone for the nose and a cylinder for the mouth).



Definitions of details and general proportions would be a problem with this approach.

Something similar to using AI in conjunction with another software would be using any text-to-image generator to create a series of images representing the same object from different points of view. The images can be fed to any photogrammetry software to recreate a 3D model.

Even if it was a possible road to take, my initial approach was surpassed by the latest text-to-3D generators that can be divided into seven different categories:

Generative Adversarial Networks (GANs) generate high-quality 3D models from textual descriptions. They consist of a generator that creates the 3D model and a discriminator that evaluates its quality, iteratively improving the output.

Diffusion Models

Diffusion models gradually refine a noisy 3D shape into a detailed model by reversing a diffusion process. This method is effective for generating complex and high-resolution 3D models.

Transformer-Based Models

Transformer models, such as those used in natural language processing, are adapted for text-to-3D generation. They leverage their ability to understand and generate sequences to create detailed 3D models from text prompts.

Point Cloud Generation

Techniques like OpenAI's Point-E generate point clouds from text inputs, which can then be converted into solid 3D models suitable for 3D printing.

Magic3D by NVIDIA

Magic3D employs a coarse-to-fine strategy, starting with a low-resolution model and refining it to a high-resolution textured mesh.

This method ensures high-quality outputs suitable for 3D printing.

Prompt-Based 3D Mesh Editing

Tools like Magic3D allow users to modify 3D models through text prompts, enabling customization and fine-tuning of the models before printing.

Image-Based Techniques

Some methods generate images first using text descriptions, which are then converted into 3D models. This approach is useful for creating detailed textures and visual elements.

AI-Driven Optimization

AI platforms optimize the generated 3D models for printing by ensuring they are watertight and have the appropriate resolution and texture quality.

These methods represent the cutting edge of text-to-3D model generation, each offering unique advantages for creating detailed, high-quality 3D models suitable for 3D printing. Researchers and designers can easily transform textual descriptions into tangible, printable objects by leveraging these techniques.

GANs consist of two neural networks: a generator and a discriminator. These networks are trained simultaneously through a process called adversarial training. The generator creates synthetic data (e.g., images, text, or 3D models) that mimics real data. It takes a random noise vector as input and transforms it into data that resembles the training data. The discriminator's role is to distinguish between actual data (from the training set) and synthetic data produced by the generator. It evaluates the authenticity of the data and provides feedback to the generator. The generator and discriminator are trained in a game-theoretic scenario where the generator tries to fool the discriminator by producing increasingly realistic data. In contrast, the discriminator tries to become better at identifying fake data. This adversarial process continues until the generator produces data that the discriminator cannot reliably distinguish from accurate data. GANs are used to generate 3D printable models from textual descriptions, so they have similar limitations encountered in the creation of images such as "Mode Collapse" (when the generator produces limited varieties of outputs, failing to capture the diversity of the training data), "Training Stability" (GANs can be challenging to train due to the adversarial nature of the process, which can lead to oscillations and instability) and "Evaluation Metrics" (Measuring the performance of GANs is challenging, as traditional metrics may not adequately capture the quality of generated data). In my case, I will evaluate the models only for AM purposes and limit the testing to FDM and SLA printing.

Applications that use GANs are:

Meshy uses GANs to transform textual descriptions or sketches into detailed 3D models ready for printing.

It allows the creation of 3D Models from texts and images, AI texturing, and Voxels from text.

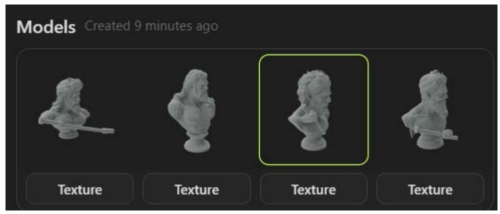
For my research, I will use the same prompt of 474 letters for all the applications to have a common starting point:

"Act as a classical Roman sculptor and create a classical bust of Heracles or Hercules, a muscular, nude, male head with a strong jaw line, a thick mane of hair, and a severe expression, with the telltale sign of the Nemean lion skin draped over his head or shoulders and holding a club. The bust would prominently display the powerful muscles of Heracles' neck, chest, and face with a determined and focused facial expression. The surface needs to show a chiseled finish."

With Meshy, you can select Realistic, Sculpture, and PBR art styles and target the mesh's definition from 3K to 100 K. It also allows the option to develop topology in quads or triangles and keep the symmetry on, off, or auto.

The credit cost of the creation was 15 credits, and the creation time was about 1 min. Considering the $20 per month plan, the cost per credit can be calculated at $0,02, and the model cost $0,3.

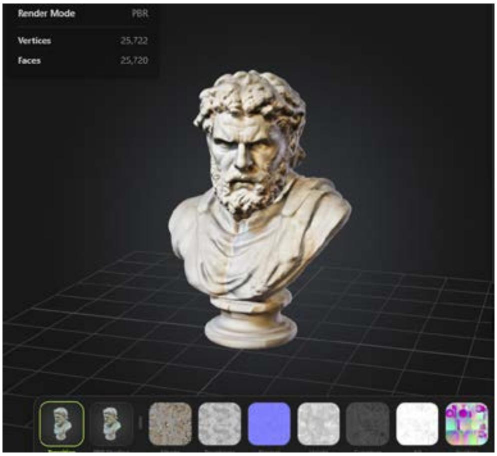
In the description, I added the "and holding a club" part specifically to see how the solution between bust and arm/hand was resolved.



Except for one of the models, the club was positioned correctly outside the bust.

A classical capital base was added to the bust in all four models without the text specifying it.

The additional texturing added details to the models that, with additional software, can be used to increase the general visual quality of the 3D surface in case of full-color 3D printing.



The texture's albedo, roughness, normal, height, curvature, AO, and position can be visualized separately.

I tested the text-to-voxel option using the same prompt, obtaining an exciting result.

The model created doesn't present any issues and is ready to print. It is watertight and doesn't have inverted triangles, holes, cracks, self-intersections, or thin features.

DreamFusion and **Magic3D** use 3**D Diffusion** or, better, a pre-trained 2D text-to-image diffusion model to perform text-to-3D synthesis. Software such as **MVDream** improved results by using Multi-view Diffusion for 3D Generation. Practically, these models use a sort of photogrammetry based on multiple images of the object or, better, various points of view of the object to create a 3D model. All these images need to be spatially coherent with each other.

In the case of DreamFusion, as explained in the paper "DreamFusion: Text-to-3D using 2D Diffusion," because large-scale datasets of labeled 3D assets and efficient architectures for denoising 3D data are missing, pre-trained 2D text-to-image diffusion models are used to perform text-to-3D synthesis.

A loss based on probability density distillation enables using a 2D diffusion model as a prior for optimizing a parametric image generator. Using this loss in a DeepDream-like procedure, a randomly initialized 3D model (a Neural Radiation Field, or NeRF) is optimized.

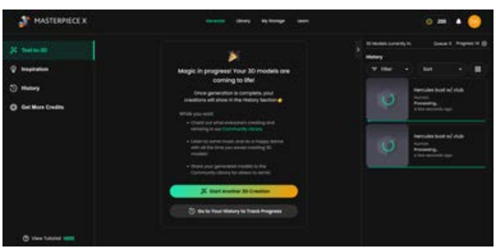
The resulting 3D model of the given text, based on 2D renderings from random angles, achieves a low loss. It can be viewed from any angle, lit by arbitrary illumination.

DreamFusion, which similarly uses a pre-trained text-to-image diffusion model to optimize Neural Radiation Fields (NeRF), was recently improved with a two-stage optimization framework. First, a coarse model using a low-resolution diffusion is created, and then a textured 3D mesh model is optimized with a high-resolution latent diffusion model.

Both models can use “Prompt-based Editing” and “Image-based Editing.”

I tested the models through **NVIDIA Edify**, a multimodal architecture for developing visual generative AI models for images, 3D, 360 HDRi, physically based rendering (PBR) materials, and videos.

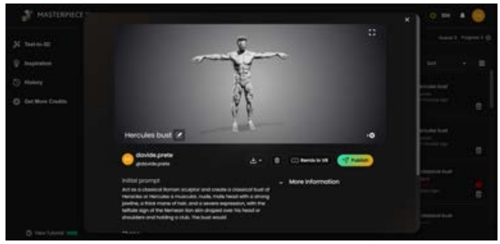
Leading adopters are **Masterpiecex** and **Cuebric**.



In **Masterpiecex**, I encountered several limitations, starting from the limit of the length of the prompt. After several trials, I had to reduce the prompt to:

“A classical bust of a bearded man”

The processing time was much longer than the 2 min used by Meshy, and the results were more generic, with an entire body of a bearded man created.

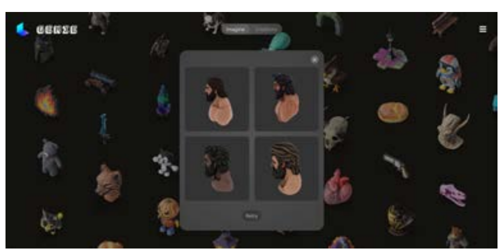


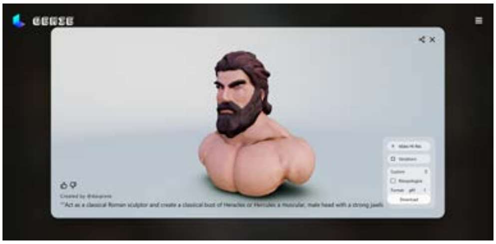
Other software uses similar methods:

Spline is an innovative text-to-3D generator that uses AI to produce lifelike 3D objects and animations from text-based instructions. It offers a collaborative platform for real-time teamwork and individual creativity.

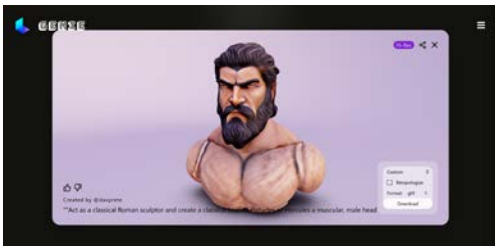
3DFY AI creates high-quality 3D models from text descriptions by separating the AI pipeline into object categories and using specific category datasets to train the generation of each class of object. In this way, each deep learning algorithm can create any missing information from the inputs. Higher poly count models are quad-meshed with geometrically adaptive density, perfect for 3D printing.

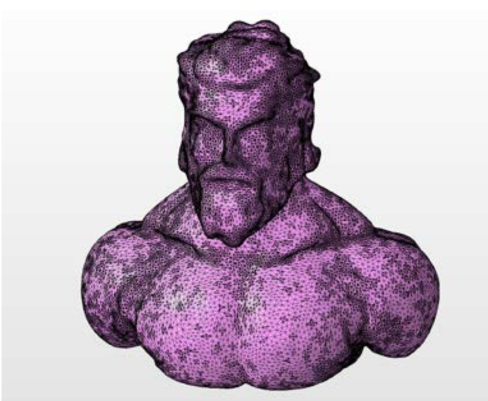
Genie by Luma AI: Genie is a free tool that converts text into detailed 3D models. It supports various export formats, such as STL, BLEND, OBJ, and more. The only issue I encountered was that I had to remove the word “nude” from the prompt for Genie.





After the first generation of the models took only a minute, the "Hi-Res" generation of the selected one took several minutes.





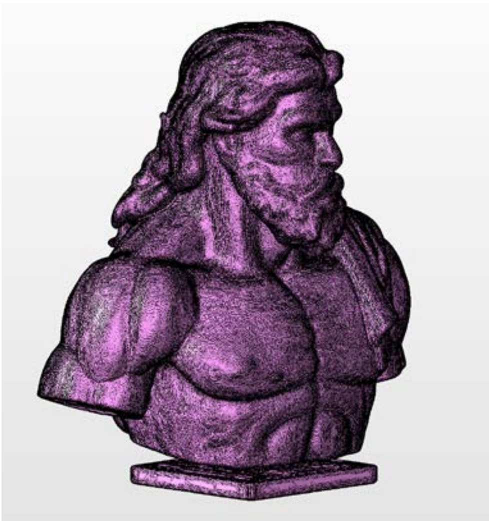
These tools leverage transformer-based models to streamline the process of creating 3D models from textual descriptions, making it easier for designers, animators, and developers to bring their ideas to life.

Another system developed by OpenAI is **Point-E**: A system for generating 3D point clouds from complex prompts. Point-E produces 3D models in only 1-2 minutes on a single GPU by first generating a single synthetic view using a text-to-image diffusion model and then creating a 3D point cloud using a second diffusion model, which conditions the generated image. For now, the mesh quality is much lower, but it is much faster (one to two orders of magnitude) to sample from, offering a practical trade-off for some use cases.

Another exceptional generator is **TripoAI**, which is specifically designed to prepare models for printing. Tripo AI's technology is built on the AIGC (AI-Generated Content) foundation model, a sophisticated AI framework designed to create high-quality 3D models from text descriptions or images. Tripo AI uses advanced NLP techniques to understand and interpret text descriptions. This allows users to input simple descriptions of what they want to create, and the AI can comprehend and process these instructions. When generating models from images, Tripo AI employs GANs and computer vision algorithms to analyze and understand the visual content. This helps in accurately translating 2D images into 3D models. The core of Tripo AI's technology includes specialized algorithms for 3D modeling that can create detailed and complex structures based on the input data.

The generator allows you to rig and animate the model, simplify it, retopology it, and export it to different types of files.

The model I created had high-quality details and good proportions. The texture followed the prompt without errors, such as open / intersected triangles or holes.



The best generators for sculptural applications were Meshy and Tripo. Both support AI 3D printing generators, enabling users to prepare models for physical production. The platform's compatibility with STL generator AI and AI STL generator ensures seamless integration with 3D printers.

Conclusion

The quality of the generated 3D models heavily depends on the training data. Ensuring a diverse and high-quality dataset is crucial for the models to generalize well across textual descriptions.

The generation of detailed 3D meshes requires significant computational power. Optimizing the algorithms to be more resource-efficient will be a key development area.

Results:

Seamless integration of AI text-to-3D models with existing 3D printing workflows, including compatibility with various 3D printing software and hardware, is essential for widespread adoption.

Key words: Additive Manufacturing, 3D Printing, AI, Text-to-3D Generators, NeRF.

FAST FORWARD: How Artificial Intelligence (AI) Has Integrated into the Design Curriculum in Less Than A Year

Assistant Professor Peggy Bloomer, PhD.
Department of Art & Design, Central Connecticut State University, New Britain, CT, USA
www.peggybloomer.net
e-mail: bloomerp@ccsu.edu

Abstract

The rapid integration of Artificial Intelligence (AI) into the design curriculum has transformed from a perceived disruptor to an indispensable tool. AI-powered applications empower students to enhance their creativity and efficiency. This evolution highlights AI's capacity to augment human capabilities, enabling students to streamline processes, generate innovative ideas, and refine their artistic endeavors.

In less than a full semester, the use of AI has become an integral part of courses in the design curriculum. The Fall 23 semester started with the expectation that AI would become a technology disruptor in classes and that there would be a need to guard against using AI in student design and writing projects. Instead, AI has become an essential tool for design students as the capabilities and integration of AI into software have rapidly evolved and developed. This paper will discuss how students use AI to

1. create images,
2. map and develop projects,
3. code and create intermediate artifacts,
4. and use AI incorporated into industry-standard design software.

In addition, it is noteworthy that the university where the author teaches has developed an "AI Corridor" on campus and an academic major in Artificial Intelligence. Their current position about using AI is that students must learn how to use it to stay current in the marketplace. The university policy on AI and student ethical usage of it can be found at the Central Connecticut State

University website on the library guidelines and AI usage page.[1] In addition, using Artificial Intelligence and Augmented Reality have been added to the core design curriculum and two special topics classes cover AI/AR and producing product packaging with AI.

Text to Images

Surprisingly, many students starting in the Graphic Information Design Program have not experimented with AI image production before entering the design program. Even though many incoming students do own iPads and use programs such as Procreate (https://procreate.com) to create illustrations and digital drawings, they have not looked to text-to-image software for inspiration or enhancement of their image creation. Procreate is a fantastic tool for fostering creativity, enabling users to draw freely with a digital pencil and incorporate photos and videos into their workflow. However, there are limitations when creating artwork that can be reproduced beyond digital screens. Incoming students are generally unaware that the resolution of Procreate artwork is often too low for professional design applications and that integrating it into high-resolution projects can be challenging without enhancing it through software like Topaz which uses AI to upscale images. Issues with image resolution or pixels per inch (PPI) are common in all image-to-text generation processes. While upscaling technology has improved, achieving full-page resolution for AI-generated images still requires tools like Topaz or other AI-driven enhancement programs.

Stand-alone text-to-image generators are widely accessible, but many users, especially students, often rely on ‘free’ image generators. For a list of the best options, see Sabrina Ortiz’s review.[6] Adobe Creative Suite has established a strong presence in the AI market by integrating text-to-image generation with Adobe Firefly, adding tools like text-to-image generators, expanded generated backgrounds, and advanced foreground and background selection in Photoshop and Adobe Illustrator.

Similarly, Canva, used by over 75 million users per month, aims to attract content creators with comparable tools and an additional AI feature called magic design.[1] While Adobe primarily targets professional designers, Canva appeals to a diverse audience, including educators, students, business professionals, and social media managers, in general, the non-professional.[5,7] The addition of text-to-image generation in these platforms makes image creation more accessible to a broader range of content creators and extends users’ capabilities.

Crafting Better Prompts

Initially, students expressed frustration with how AI fashions hands and exhibits a confirmation bias toward certain ethnicities. However, they quickly learn that crafting more specific prompts yields better results, aligning with Boris Groys's observation that using AI prompts is similar to using search engines.[2] AI applications can write code, generate written content, create images with intricate details, and offer specific styles. This allows users to produce imaginative and creative results that go beyond their artistic skill levels and techniques. One effective technique students used successfully this semester was incorporating a reference image to guide the AI process.

Below are several avant-garde fashion images created by student Joey Rodrigues using Adobe Firefly for a web project that combines a prompt with a reference image, illustrating how the reference image influences the final results. While these outcomes share similarities, the unpredictability of Artificial Intelligence adds an extra dimension to the creative process which is demonstrated in these examples.

1. *A reference image combined with a prompt in Adobe Firefly creates a hyper-realistic image of a woman in avant-garde fashion. The result is in black and white, with a blend of light to emphasize contrast and texture.*

2. *A reference image combined with a prompt in Adobe Firefly creates a hyper-realistic image of a woman in avant-garde fashion. The result is in black and white, enhanced with a blend of rainbow tones for a striking, multidimensional effect.*

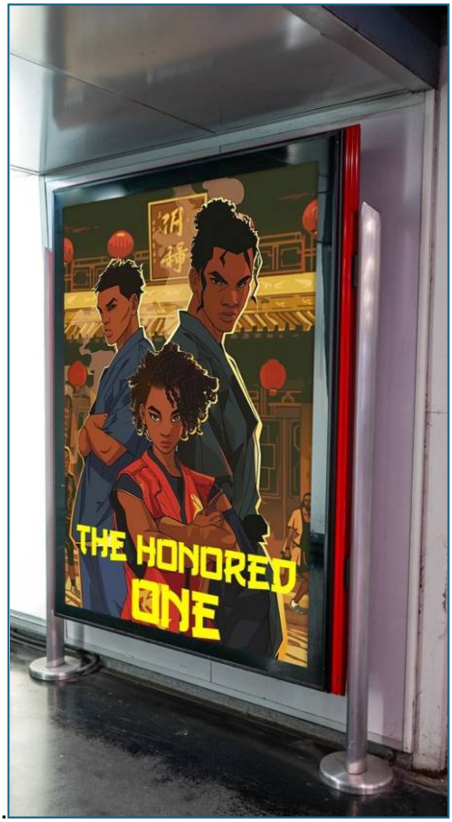
3. *Another reference image combined with a prompt in Adobe Firefly generates a hyper-realistic image of a woman in an avant-garde fashion. The result is accented with a subtle blend of rainbow tones to add depth and color contrast.*

4. *A reference image combined with a prompt in Adobe Firefly creates a hyper-realistic image of a woman in*

avant-garde fashion. The result is in black and white, enhanced with a blend of holographic foil effects and geometric, asymmetrical shapes for a modern, abstract aesthetic.

Saving Time

Integrating AI into digital illustration enhances students' ability to rapidly create detailed, imaginative scenes, streamlining the creative process and allowing them to explore complex visual storytelling in otherwise time-consuming ways. Students appreciate the remarkable level of detail that AI can add to images in a short amount of time. By embracing rapid prototyping in “text-to-image” generation, they can efficiently create intricate backgrounds, patterns, and scenes for their projects—tasks that would otherwise demand substantial time and effort if done manually. Through well-crafted prompts, students streamline their workflow, leveraging AI to amplify their creative output. AI makes it easy to produce highly detailed scenes and patterns, which students use to build storyboards, and components of a project.



4. *Students Justin Youmans, Cody Eckstrom, Kahari Kirlew, and Christian Sosa developed patterns and imaginative realms with artificial intelligence, combining them into websites, social media posts, and posters.*

One area where Artificial Intelligence does not excel is in character animation. Even using image references, controlling a character's consistency is still hard. In the digital illustration class, students created digital character studies using artificial intelligence. They were somewhat successful. However, it proved almost impossible to recreate a character for even a short animation reliably.

5. *Above is a generated character pose study created by student Justin Youmans. Below, Christian Sosa uses AI to create a storyboard that effectively delivers the story's message in a compelling visual, but the characters do not remain consistent across scenes. Notice how detailed and persuasive the surrounding scenes are, though.*

Using LLMs (Large Language Models) to create project planning and coding

Initially, students hesitated to use AI tools like ChatGPT in their projects. However, after Ms. Bloomer attended a class on artificial intelligence, her view on large language models (LLMs) developed, and she began sharing the advantages of AI with her students. In a digital illustration class, students used ChatGPT to generate plot summaries and scripts for short animations based on their prompts.

These AI-generated scripts served as starting points, which students then reviewed and modified to align with their project specifications and imagination. This approach highlights how AI tools can effectively extend human creativity through their direction.

ChatGPT also proved helpful for research and citation management. It can help students search through PDFs they collect and store in folders on the cloud to find information about topics they are using in a paper, generate citations in various style formats, and alphabetize bibliographies.

Another valuable application of ChatGPT is in coding assistance. This semester, Ms. Bloomer teaches an introductory design course covering design planning artifacts, HTML, and CSS. As part of this course, she instructs students on how to build simple, responsive websites through hand coding. After students become familiar with HTML tags and CSS selectors, Ms. Bloomer encourages them to use ChatGPT for additional support. For instance, she suggests that students input one of the HTML project briefs as a prompt in ChatGPT to generate code for that project.

This assignment reinforces the importance of understanding the underlying concepts. When a student provided specific parameters and image URLs in their prompt, ChatGPT returned an HTML and CSS file, which the student-organized into folders on their computer to use effectively.

As detailed as this prompt looks, the text and the header image need either more preparation or some work in the HTML. Here is what one prompt looked like:

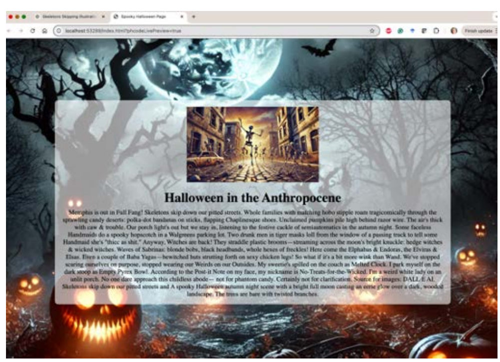
> Make me a web page using HTML and CSS to create a responsive page with a minimum width of 400px and a maximum width of 1200px. Use this image as a background image that covers the page https://www.dropbox.com/scl/fi/ntodexk2cqyg7487igax9/DALL-E-2024-10-15-08.45.49-A-spooky-Halloween-autumn-night-scene-with-a-bright-full-moon-casting-an-eerie-glow-over-a-dark-wooded-landscape.-The-trees-are-bare-with-twisted-bra.webp?rlkey=xtwvjr7zeaheezjts1mmku3ay&dl=0 place with this text in a container div that has an opacity of 60% and a font color of black with a nested header containing this image https://www.dropbox.com/scl/fi/0083kli7zv71sconjzkfm/DALL-E-2024-10-15-08.43.09-A-surreal-scene-showing-skeletons-joyfully-skipping-down-an-old-worn-out-street.-The-street-is-pitted-with-cracks-and-potholes-giving-it-a-gritty-p.webp?rlkey=iwtrx08sjuy4yytr16r17jap5&dl=0

6. A student prompt in Chat-GPT to create a responsive web page with URLs for images

Chat-GPT returned an HTML and CSS page with the following explanations and directions for the student to follow.

- The page is responsive with a minimum width of 400px and a maximum width of 1200px.
- The background image covers the entire page.
- The container has a 60% opacity with black text.
- The skeleton image is nested inside the container as a header image.
- You can save the HTML and CSS code as separate files (index.html and styles.css) and open them in a browser to see the page in action

Here is what the page looks like:



7. *Viewing the page in a browser after creating pages and placing them in a folder on the student's computer.*

The student still needs to add line breaks and adjust the top image to span across the top of the column, requiring further work using the HTML and CSS code to achieve the desired layout. This approach reverses Bloom's taxonomy by allowing a student with an incomplete understanding of HTML and CSS to start with a result and then refine it to improve their work. By "debugging" and engaging directly with the code, the student learns through problem-solving rather than rote memorization. This inversion of traditional learning methods may help students better understand coding structures and their logic. In a world of constantly evolving knowledge, this troubleshooting approach might offer a more effective pathway for learning than traditional methods focused on mastering vocabulary and syntax.[8]

This HTML and CSS knowledge level may be enough for a design student working with content management systems like WordPress and Squarespace. Using a large language model (LLM) to assist with coding, the student could likely write more effective prompts, reducing the amount of manual work needed afterward by leveraging AI-generated frameworks.

Incorporation of AI into Programs

Many AI applications have extended users' creative and professional work by integrating AI capabilities into more extensive software programs, such as Photoshop, Illustrator, and PowerPoint. For instance, Photoshop and Illustrator now include generative fills and object placement features. The UI software platform Figma also integrates with an AI component called "Dev Mode," which generates code to match the web and mobile app prototypes designed within Figma.

My appreciation for AI integration within software grew during my digital illustration class. Storyboarding, an essential part of planning any extended motion graphic project, is often time-consuming and tedious. This semester, students utilized "Boords," an AI-enhanced storyboarding program, to collaboratively create polished scripts, shot lists, and storyboards. The program even allows for the creation of an animatic from the storyboard.

The accessibility of AI programs is an important topic to consider. With the extensive range of AI tools available for design work, choosing the right platform for a project can be challenging without reliable, industry-standard options. Free versions of artificial intelligence software often come with limited creative capabilities, slow rendering, and restricted access to features.

With AI developers constantly introducing new capabilities, there are new offerings every day, making it difficult to select the best AI platform for specific needs. Many programs are neither free nor offer educational versions, and many require a mandatory year-long subscription. Students often express concerns about the high cost of the Adobe Creative Suite. Exploring AI's potential may be so costly that students, in particular, may not participate in this important new way to augment a design practice.

What's Next

In the upcoming semesters, AI will continue to be incorporated in various forms in the design courses and workflows, from "text-to-image" Large Language Models (LLMs) to software that integrates AI seamlessly into the workflow. Print assignments like posters and social media posts will also include AR 3D projections in their final iterations, enabling these projects to evolve from 2D designs into dynamic 3D experiences.

One thing is certain: AI is here to stay. While Artificial Intelligence prompts and options will continue to evolve, creating and implementing tasks into usable projects still requires skill, knowledge, and a human touch. This process calls to mind Disney's *The Sorcerer's Apprentice*—without human guidance, curation, and intelligence; artificial intelligence cannot fully capture or convey human existence, creativity, or the essence of humanity. Humans must remain the maestro and orchestrate these projects.

The shift from skepticism to acceptance of AI in design education and workflows represents a transformation where this technology is increasingly seen as a catalyst for innovation rather than a threat to creativity. McLuhan's view of technology as an extension of human capabilities resonates with this perspective, suggesting that the future of design education lies in embracing Artificial Intelligence as a collaborator that enhances human ingenuity and fosters creative expression.[3]

REFERENCES

[1] Canva**.** (n.d.). *Magic Design.* Canva. Retrieved November 2, 2024 from https://www.canva.com/magic-design/

[2] Central Connecticut State University. (n.d.). AI and ChatGPT - Research Starters - CCSU Research Guides. Central Connecticut State University. Retrieved November 1, 2024, from https://libguides.ccsu.edu/ai#:~=Students%20should%20remember%20to%20act,cite%20their%20usage%20of%20AI

[3] Groys, B. (n.d.). From Writing to Prompting: AI as Zeitgeist-Machine (notes, e-flux). E-flux.com. Retrieved September 19, 2023, from https://www.e-flux.com/notes/553214/from-writing-to-prompting-ai-as-zeitgeist-machine

[4] McLuhan, M., & Fiore, Q. (1989). *The medium is the massage* (11th ed.). Pocket Books.

[5] Noble Desktop. (n.d.). *Canva vs. Adobe Creative Cloud for graphic design.* Retrieved [today's date], from https://www.nobledesktop.com/classes-near-me/blog/canva-vs-adobe-creative-cloud-for-graphic-design

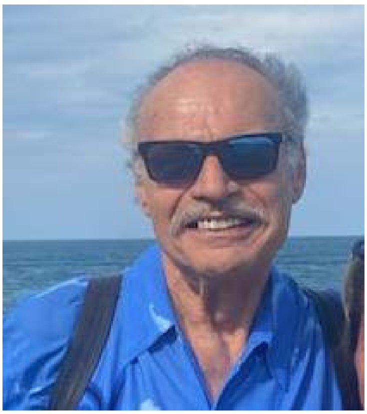
[6] Ortiz, S. (2024, September 20). *Best AI image generator*. ZDNet. https://www.zdnet.com/article/best-ai-image-generator/

[7] The Social Shepherd. (n.d.). *Canva statistics.* Retrieved November 1, 2024 from https://thesocialshepherd.com/blog/canva-statistics

[8] Wright, S. (2012, May 15). *Flipping Bloom's taxonomy*. Powerful Learning Practice. https://plpnetwork.com/2012/05/15/flipping-blooms-taxonomy/

Generative AI clarifies validity of authentic paintings for measuring artistic judgment aptitude

Dr. Nikolaus Bezruczko, PhD.
Department of Clinical Psychology,
The Chicago School of Professional Psychology, Chicago, USA
Nikolaus Bezruczko's albums | Flickr
e-mail: Nbezruczko@thechicagoschool.edu



Abstract

The ideas of aptitude and talent are intertwined in common language, and their empirical identification has broad economic implications in Western post-industrial economies. The purpose of aptitude measurement is to improve human function by identifying where those talents lie. Early attempts to measure artistic judgment aptitude in the 20th century failed because test images could not remove the influence of culture and education. Generative AI solved the problem by using algorithms to create culture-free images that were uncontaminated by experience or arts training. The purpose of the present research was to demonstrate that preference for generative images validated by professional artists can be generalized to authentic paintings, which would accommodate an aversion to synthetic images found during aptitude testing. This research implemented a probabilistic Rasch model to measure the stochastic force that underlies preference for artist-validated generative images. Then a professional artist adapted this algorithm to figurative paintings. Results showed, in general, visual preference for synthetic images and authentic paintings were positively correlated (r^2=.50), and internal consistency was high, .90 and .95, respectively. Unexpectedly, this generalization of professional validation was mediated by painting style and content. In addition, layperson preferences tended to converge on professional preferences but only in an inverted U function as complexity was increased. These results give mixed support for Berlyne's hypothesis but also suggest caution for AJ aptitude measurement.

1. Background

Many aptitudes in the physical, intellectual, and artistic domains contribute to human performance differences with significant influence on lifetime outcomes. Moreover, aptitude measurement is pursued by schools, families, and corporate departments of human resources to enhance career counselling and optimize talent management.

1.1 Practical methods.

Early 20th century efforts to standardize artistic judgment (AJ) aptitude measurement struggled with construct validity issues. They typically implemented authentic masterpiece paintings, which confounded AJ aptitude with arts experience and visual arts training. Without objective art standards for inferring AJ aptitude, the meaning of scores could not be established and AJ aptitude testing was largely abandoned. Demand, however, for AJ testing continued.

Modern statistical approaches introduced the idea of abstracting visual arts design factors from its cultural context, then generating artificial images. Isolating abstract design factors that are associated with artists' creative preferences would embody artistic standards independent of culture. Perceptual sensitivity to those culture-free design factors by lay persons could indicate latent AJ aptitude. Significantly, design factors would eliminate dependency on recognizable masterpiece paintings. By the 1940s, visual preference factors were identified that distinguished between artists and nonartists [1] such as asymmetry, simplicity, and nonuniformity, which continue to interest researchers [2]. The Divine Proportion or Rule of Thirds is probably the oldest design principle widely known among artists. The present research hypothesized that laypersons preferring design factors in standardized random images may be high in AJ aptitude, which was empirically validated.

This emphasis on design factors gained additional importance with the rise of generative art, which implements algorithms to produce synthetic art. Generative AI algorithms were embodied with design features that emulated artist-preferred authentic paintings [3]. Explicit control over design features facilitated producing standard synthetic images. Rasch model statistical decomposition of preference ratings then provided empirical validation of an aptitude test variable.

The stochastic philosophy underlying AJ aptitude measurement has also influenced contemporary painters. Simultaneous with AJ aptitude research, 20th century painters began experimenting with stochastic ideas. Pollack's random drip paintings are widely recognized, but also entire art movements emerged such as pointillism and neoplasticism, which explicitly implemented randomness in their art. The most notable advocates of stochastic theory in visual arts were Piet Mondrian and Theo van Doesburg from the Dutch De Stijl School. Ellsworth Kelly, Jean Arp, Kenneth Martin, and later in his career, John Cage were also artists in this contemporary trend.

1.2 Problem

Despite advances implementing generative art to emulate artist-preferred design factors, a complication now mutes their success. Recent studies suggest that laypersons prefer authentic paintings even when the AI generated image is indistinguishable [4,5]. Consequently, generalizing validated design factors from synthetic images to authentic paintings is a problem.

The aim of the present research was to demonstrate that professional artist validation of synthetically generated images can be integrated into authentic figurative paintings. Spontaneous authentic art presents validity issues for aptitude testing because its design features are typically unstandardized. This strategy of imposing design structure on authentic art would alleviate enormous costs and inconvenience of professional artist validation.

1.3 Research question

The goal here was to generalize professional artist validation of synthetic images to authentic paintings. The central question was whether rule-based design factors can guide painting authentic figurative art thereby emulate professional artist validation. More specifically, does validity of AI-generated images generalize to authentic oil on canvas paintings?

1.4 Empirical strategy

Authentic paintings guided by a rule-based algorithm were expected to correlate positively with AI generated synthetic images that were previously validated by professional artists. The statistical null hypothesis was preference for AI-generated images and authentic paintings are unrelated.

The idea conceptually was to link AJ aptitude with perceptual sensitivity to validated design factors, then generalize that validation to authentic paintings by adapting the generative algorithm for paintings. The plan of this research was:

- Implement a generative algorithm to produce synthetic images and generalize its rule-based design structure to authentic paintings.
- Examine scale properties and statistically correlate synthetic and authentic preference measures.
- Evaluate generalization of professional artist validation to authentic paintings.

2. Method

2.1 Sample

The sample was 462 examinees from the Johnson O'Connor Research Foundation (JOCRF) testing offices in Boston, New York, Chicago, and Dallas. They were clients of JOCRF's aptitude assessment service. This sample consisted of 215 males and 247 females, predominantly white (95%), upper-middle-class, and college-educated or college-bound.

2.2 Aptitude test images

Two sets of images were central to this research. First, a synthetic set that consisted of random patterns generated by an algorithm, which embodied design factors of simplicity and asymmetry. Then a second set of authentic figurative paintings that emulated design factors embodied by the synthetic image algorithm. In other words, the design factors had been carried over to the paintings by the adapted algorithm, which has been published [3].

2.2.1 Generative images

Two design factors were operationally defined. Complexity as variety and frequency of elements. Redundancy was defined by element dispersion and spatial orientation. Presented in pairs: artist-preferred vs other, each image contrasted both factors and scoring was (0/1) keyed for less-complex and less-uniform artist-preferred preferences.

2.2.2 Artist validation

The published AJ literature shows an egregious bias in the selection of study participants. Virtually all empirical aesthetics studies including Berlyne's experiments enrolled only college students. Therefore, differences between professional artists and laypersons simply could not be examined. In contrast, artist validation conducted for the present research *excluded* students by only enrolling artists of standing. The validation sample was restricted to working artists with recognized career trajectories, as well as awards and distinctions in several American cities. Artist background details of this sample are published [6].

2.4 Authentic figurative paintings

Figurative images for the present research were painted in five styles where design factors were explicitly manipulated in successive paintings at three complexity levels. The five styles: Fauvist, Surrealism, Renaissance, Baroque, and Post-Impressionism. Figurative images were oil and acrylic on canvas paintings. An artist-preferred image and three levels of complexity were rendered for each style. First, the artist identified proportional relations between complexity and redundancy in four artist-preferred synthetic images (1 complexity: 0.5 redundancy) and reproduced this ratio on canvases. The painting in a fifth style, Fauvism, presented complexity and redundancy in a 1 to 1 relation. Then complexity for each style was discreetly increased three times while maintaining stylistic coherence. Redundancy was not explicitly manipulated and may have varied slightly across paintings.

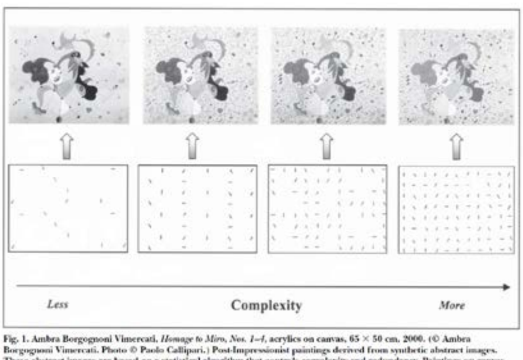


Figure 1. The artist identified proportional relations between complexity and redundancy in five art-preferred synthetic images (1 complexity: 0.5 redundancy) and replicated this ratio on canvases in five styles: Fauvism, Post-Impressionist, Surrealism, Renaissance, and Baroque.

Altogether, 20 image pairs compared an artist-preferred image with an image of higher complexity. Figure 1 presents the process of generalizing validated design factors from synthetic generative images to authentic paintings [7].

2.5 Data collection

JOCRF's standard aptitude battery was presented, then examinees were presented synthetic images and authentic paintings. All preferences were scored (0/1) for their agreement with professional artists.

2.6 Scale analysis

In the measurement theory rationalizing probabilistic Rasch models, each person B_n and item D_i and their observed differences (B_n - D_i) are compared to maximum likelihood model predictions, which shifts measurement from arbitrary normative distributions confounded by sampling to stochastic, mathematically modelled differences independent of samples. During examination of this difference (B_n - D_i), uniform units are systematically accumulated and probabilistically concatenated to define an interval scale. Observed and predicted inconsistencies are examined by fit statistics and residual analysis. Finally, standard errors for Bs and Ds are established along the entire variable providing maximum scale precision.

3. Results

3.1 Exploratory factor analysis

As expected, Principal Component Analysis (PCA) identified four dominant factors. As expected, after Promax rotation, both synthetic components (complexity and uniformity) loaded on factors one and two, while authentic paintings loaded on a third factor. A fourth factor, unexpectedly, presented several groups of images clustered together defined by style and content. Factors 1, 2, and 3 accounted for 19.4%, 9.5%, 5.5%, and 4.4%, respectively. Factor 1 was correlated with 2 and 3, r=.285, and r=.265, respectively. Internal consistency reliability was computed for correlated factors, which indicated alpha >.90.

3.2 Correlation with artist-preferred

Null hypothesis that synthetic and authentic art are uncorrelated was rejected (p<.05). Results show statistically significant correlation but conditional on style and content. Figure 3 shows a generally positive correlation between synthetic and authentic preferences where almost 50 percent of images with comparable design structure tended to agree (r=.70). A closer examination, however, reveals their relatively complex covariation. Preferences, for example, tend not to define a common identify line as one would expect when analysing concurrent measures. Instead, persons who agreed with the artist-preferred image on both synthetic and authentic art appear to define a cluster. While persons who agreed with artist-preferred on authentic paintings but not on synthetic tended to form another group. Speculation is these results represent three types of AJ aptitude. First is Type 1 (n=69; 15%), examinees who agreed with professional artists on synthetic images but not on authentic art. Type 2 agreed with artists (n=103; 22 %) on both synthetic and authentic art (quadrant 2). Type 3 (n=132; 29%) agreed with artists on synthetic but not on authentic art (quadrant 3). Quadrant 4 shows examinees who disagreed on both synthetic and authentic art (n=158; 34%).

3.3 Preference measurement

The Rasch preference variable parameterized in this research was a probabilistic measure of agreement between laypersons and professional artists. Study participants scoring high indicated consistency with professional artists, while low scores indicated disagreement. Analysis of Simplicity and Uniformity components separately indicated both scales were very reliable (>.90) and model fit was reasonable. Person reliability was lower, .83 and .77, respectively. Simplicity items, however, were very difficult and both scales showed substantial floor and ceiling effects. When items were pooled, Complexity and Uniformity items systematically contributed to a single, common

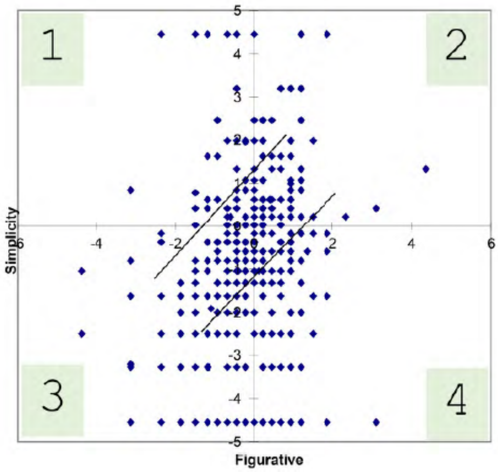


Figure 2. N = 435. Layperson agreement with professional artists on synthetic and authentic art correlated r=.70, which indicates about 50% agreement.

preference dimension. Uniformity items were systematically lower on the item hierarchy than Simplicity indicating Uniformity art-preferred images were more easily preferred than Simplicity items. Results show substantial support for calibrating synthetic and authentic paintings together, which defines a common, unidimensional preference variable. For example, item and person separations were higher 5.36 and 2.59, as were reliabilities .97 and .87. Also, both item and person values indicate adequate model fit. Moreover, by not showing ceiling or floor effects this combination was better targeted on examinees. PCA of model residuals confirmed separate influence of synthetic and authentic factors but without threatening overall scale structure. These results suggest that combining synthetic and authentic paintings define a unidimensional measure of visual arts-related preference better than either factor alone.

3.4 Agreement with artists differed by style and content

The aptitude test variable examined here was defined by images located along a continuum ranging from very low agreement with professional artists to very high agreement. All paintings including complexity manipulations were located along this variable. Higher agreement would be indicative of higher AJ aptitude. Results show layperson agreement with artists was mediated by painting style and content. Post-Impressionist images appearing near the scale floor indicated lowest agreement hence lowest AJ aptitude. Surrealism clustered slightly higher. Renaissance and Baroque images showed more agreement with artists. Fauvist images were hardest for laypersons to agree with artists and indicative of higher AJ aptitude. In general, representational styles solicited more agreement with professional artists, and they were indicative of higher AJ aptitude.

Figure 3 shows three complexity manipulations for each style. Preference for representational styles (Renaissance, Baroque, and

Fauvism) tends to follow an inverted U relationship. The third complexity increase points to conformity with Berlyne's complexity function but notably only at the highest complexity level in present research [8]. In fact, agreement with artists when presented Post-impressionist and Surrealism images *decreased* at highest complexity. Surprisingly, nonrepresentational styles, did not conform to complexity expectations suggesting substantially more complexity would be needed to drive layperson preference closer to artists.

4. Discussion

4.1 Conditional AI generalizability

Results support conditional generalization of professional artist validation from synthetic images to authentic paintings. Unexpectedly, painting style and content interacted with design factors, which presents practical complications. Also, the manipulation of design factors requires caution as exaggerations may artificially increase agreement with professional artists' preference. In the present research, a 30 percent increase in complexity was sufficient to shift layperson preference.

Image content also presented complications. Thematic paintings including spiritual, religious content were more challenging for laypersons to agree with artist-preferred suggesting content, in general, may mediate agreement. Style also showed major variation on the aptitude test variable, where Fauvism and its manipulations were harder to agree with artists than all other painting styles in this research. All Fauvism manipulations of complexity showed stronger agreement with artists than any other style hence higher latent AJ aptitude. As expected, modestly increasing complexity tended to decrease preference for artist-preferred images. However, additional increases in complexity tended to "push" layperson preference toward artist-preferred paintings. In other words, visual preference follows a U function and at higher levels, laypersons prefer less-complex paintings concurring with professional artists. These results were unexpected, though they recall Berlyne's prior research. Although conducted without artists, Berlyne found student preference also followed a U function. Berlyne's hypothesis in present research was supported by representational styles such as Renaissance, Baroque, and Fauvism, but not Surrealism or Post-impressionism, which contributes to a contemporary discussion [9]. In terms of AJ aptitude measurement, exaggerated manipulation of complexity has the undesired effect of shifting laypersons toward artist preferences.

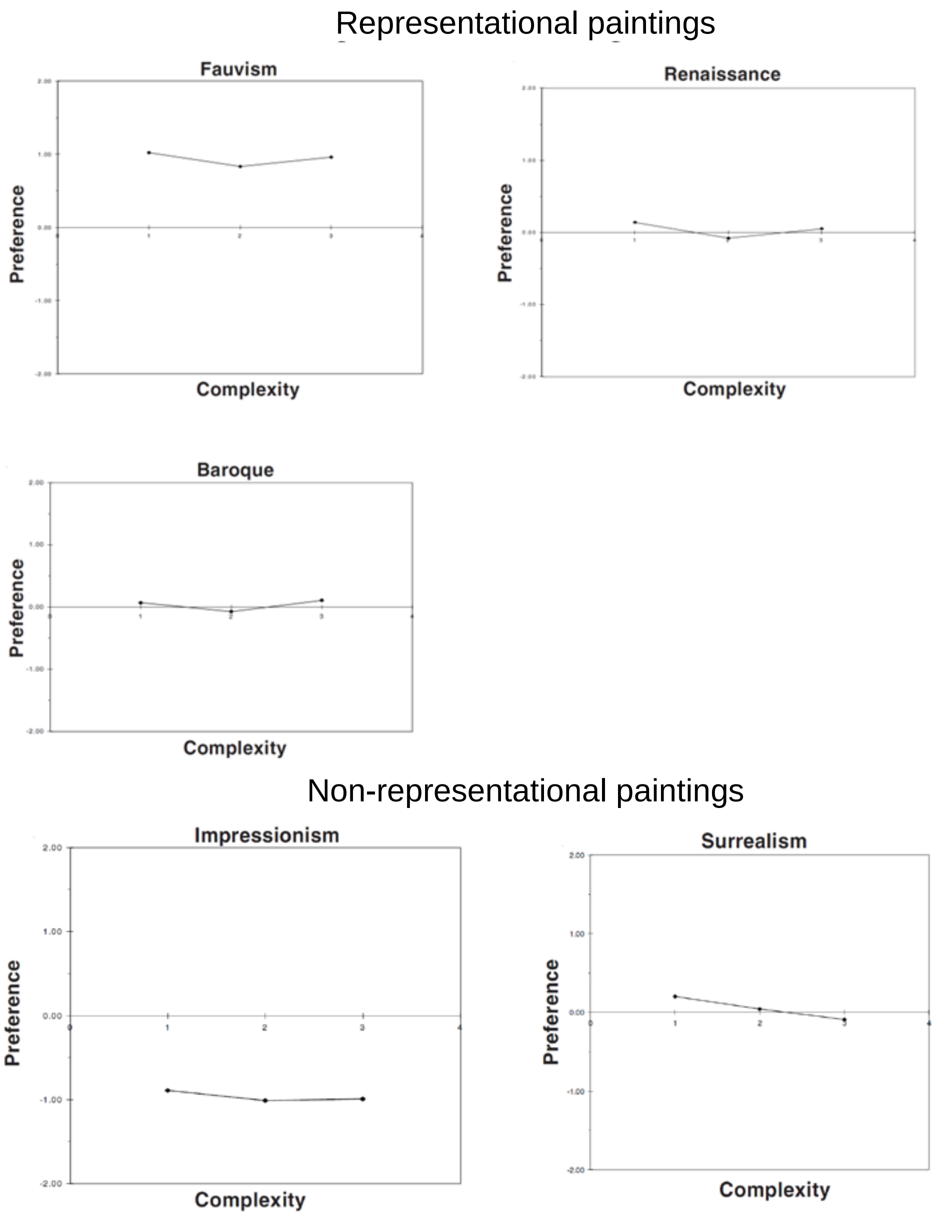


Figure 3. Authentic paintings were executed in five styles: Fauvism, Renaissance, Baroque, Impressionism, and Surrealism. An artist then proportionally increased complexity in three separate versions. Preference was measured in logits. Purpose was to capture influence of complexity as a design factor on agreement of laypersons with artists.

In other words, laypersons' preference for professional art increases as the alternative becomes more complex especially for representational images. The implication for aptitude testing is complexity in image choices
should not reach upper complexity levels to avoid distorting individual differences. Images tend not to discriminate well between high and low AJ aptitude when the non-art preferred image exceeds a complexity ceiling. In the present research, that ceiling was reached on the third manipulation of several paintings. As more complexity was added, art-preferred image became much easier.

4.1 Summary and conclusions

Using authentic paintings for aptitude measurement was supported but with caution. Generalization of artist validation established for generative images to authentic paintings was mediated by painting style and content, while representational images appeared to moderate layperson agreement with artists. These results have important implications for AJ aptitude measurement.

Acknowledgements

I am deeply grateful to the Johnson O'Connor Research Foundation for collecting data for this research. Likewise, I am indebted to Ambra Borgognoni Vimercati, an Italian fresco restoration artist, for painting the art for this research, which would not have possible otherwise. Brian A. Russo and Kyle Perkins contributed valuable comments.

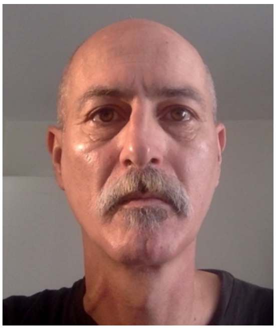
References

[1] Eysenck, H.J. (1940). The general factor in aesthetic judgements. *British Journal of Psychology. 31*(1), 94-102.
[2] Gartus, A., Völker, M., & Leder, H. (2020). What experts appreciate in patterns: *Symmetry*, *12*(5), 707.
[3] Bezruczko, N. (2013). Generative art simplifies psychometrics of artistic judgment aptitude. International Generative Arts Conference, Milan,
[4] Bellaiche et al., (2023). Humans versus AI. *Cognitive Research: Principles and Implications*, *8*(1), 42.
[5] Messer, U. (2024). Co-creating art with generative artificial intelligence. *Computers in Human Behavior: Artificial Humans*, *2*(1), 100056.
[6] Bezruczko, N., & Schroeder, D.H. (1991). Artistic Judgment III: Artist Validation. Technical Report 1991-1. Chicago: Johnson O'Connor Research Foundation.
[7] Bezruczko, N., & Vimercati, A.B. (2004). Advances in measuring artistic judgment aptitude. *Leonardo*, *37*, 187-188.
[8] Berlyne, D.E. (1974). *Studies in the New Experimental Aesthetics.* Washington, DC: Hemisphere.
[9] Marin, M.M., Lampatz, A., Wandl, M., & Leder, H. (2016). Berlyne revisited. *Frontiers in Human Neuroscience*, *10*, 536.

Randomness, Rule, Program:

How the issues of generative art take shape in pre-algorithmic music

Guillaume Rochais, PhD candidate and artist-programmer.
ACCRA, Université de Strasbourg, France
e-mail : guillaume.rochais@gmail.com



Abstract

Although the 1960s marked the birth of computer art[1], numerous examples in art history outline issues that still inform the practice and theory of generative art today. This is particularly true in the field of music, the first art form to fully embrace the computer.

Algorithmic music, which emerged in the 1950s, is part of a path that stretches from Greek antiquity to integral serialism, a path in which the intertwining of music and mathematics, which has been the object of many different approaches, recurs as a leitmotif.

Other themes of particular interest to generative art also punctuate this path: the programmatic nature of the score, the recurrence of combinatorics, the generative dimension of the isorhythmic process or of Steve Reich's phasing, the different views of randomness, from the *Musikalische Würfelspiele* to John Cage or Iannis Xenakis, the formal and systematic components of serial music, the computational and algorithmic ones of set theory, the relationship between *micro* and *macro*, element and structure, or the opposition between inspired genius and objective aesthetics that runs through 20th century art.

On the one hand, the prodromal precedents of algorithmic music shed relevant light on the entire field of generative art. On the other hand, by placing generative art in a historical perspective, they contribute to the recognition of its full belonging to the artistic field.

Introduction

As early as the 1950s, music was the first artistic discipline to embrace the computer, and more specifically programming, to produce artworks. This precocity can be explained by two main factors.

First, the close relationship between music and mathematics, which has been amply demonstrated throughout history, from the first observations of Pythagoras[2] to electronic music, through medieval harmonies, the contrapuntal complexity of Johann Sebastian Bach's compositions, serialism, or again Olivier Messiaen's modal research.

Second, the two-stage separation of the creative process in music—composition /performance—presents a striking similarity with the times of writing and then performing the program. The score is in fact nothing more than a set of symbols that encode a series of predefined instructions, the execution of which produces the musical work, with the nuance, of course, that the sensitivity of the instrumentalist or conductor brings a subjectivity to the final rendering that the computer is incapable of.

1 - Early Algorithmic Music

The 1950s saw the appearance of the first algorithmic compositions created on a computer, the most famous of which is undoubtedly the *Illiac Suite* by Lejaren Hiller and Leonard M. Isaacson (1957) [3]. Originally titled *String Quartet No. 4*, the piece consists of four experiments with various musical and mathematical concepts—rhythm, pitch, dynamics, but also stochastic processes with the use of Markov chains[4].

Composed entirely on the University of Illinois's computer—Illinois Automatic Computer, abbreviated to Illiac—the *Illiac Suite* was intended to be performed by traditional instrumentalists, in this case, a string quartet. In fact, this was the case for most algorithmic compositions of the time, as the capabilities of computers at the time did not allow for the automated production of sounds of sufficient quality.

Research into algorithmic sound synthesis did exist, most notably at Bell Laboratories, initiated by Max Vernon Mathews—the "Max" of Max/MSP software—but it was still in its infancy. In 1957, the American engineer wrote his first sound-generating program, *Music I*, running on an IBM 704[5]. The program was used to produce what would go down in history as the first computer-generated piece of music, a minimalist seventeen-seconds sound sequence composed by Newman Guttman: *In the Silver Scale*[6]. Max Mathews and Newman Guttman produced a more accomplished piece the same year with *Pitch Variations*[7], a one-minute piece published by Bell Laboratories in 1962 in the anthology *Music From Mathematics*[8].

Max Mathews improved his program significantly over the next few years with several versions, but the insufficient sound quality on the one hand and the impossibility of generating music in real time on the other hand largely limited its use outside the field of experimental research[9]. Real-time algorithmic music production did not appear until the mid-1970s, thanks in particular to John Chowning's invention of FM synthesis in 1973, and mainly in the 1980s, with the advent of the MIDI standard in 1983, followed by the first Atari ST in 1985.

The term "algorithmic music", used to describe these new forms of creation that emerged in the late 1950s, refers strictly to music composed following a series of pre-established instructions:

> "[…] algorithmic composition is the art of creating music using a set of pre-determined rules. It can be called meta-composition, since the composer produces a system, a set of rules that in part produce the music for him. While randomness is very often part of the process, it is not always indispensable[10]."

As Christophe Robert points out, algorithmic music often makes use of randomness, but does not require it. Here, we need to distinguish between three notions whose frequent intermingling is a source of confusion: the term algorithmic music is not synonymous with either aleatoric or generative music, although a musical creation may well fall within these three fields.

If the composition of the *Illiac Suite* is based on randomness, it is perfectly possible to imagine a program that does not use any random functions at all; the music produced in this way will correspond in every detail to the composer's prescriptions. In this case, the computer is used only for its computational capabilities, to automate a series of calculations that would be too tedious to perform by hand.

Moreover, the fact that the score of the *Illiac Suite*, like that of other algorithmic works of the same period, is fixed once and for all, also excludes it from the field of generative music: apart from the variations in execution—which are found in all written music—the piece is identical for each performance. To approach a generative dimension, it would have been necessary, for instance, to run the program before each concert, in order to redraw all the random parameters and thus generate a new score, different from the previous ones.

2 - Isorhythm

As in the visual arts, the use of algorithms in music long predates the advent of the computer. An early trace of this can be seen in the isorhythmic process used in the late Middle Ages, in which a rhythmic motif—*talea*—and a melodic motif—*color*—begin together and repeat until they end together[11]. If the *talea* has five durations and the *color* has six notes, then the *talea* must be repeated six times and the *color* five times, a sequence of thirty notes/durations, to reach the initial configuration.

The process is not without similarities to the phasing Steve Reich used as early as 1965 for *It's Gonna Rain*, a piece for two tape recorders playing the same sound loop at slightly different speeds. The two tapes start in unison, then gradually slip out of step until they return to unison, and so on. The composer frequently reused this technique, most notably in 1967 for *Piano Phase*, in which he had his score performed by two pianists shifting their playing before reuniting[12].

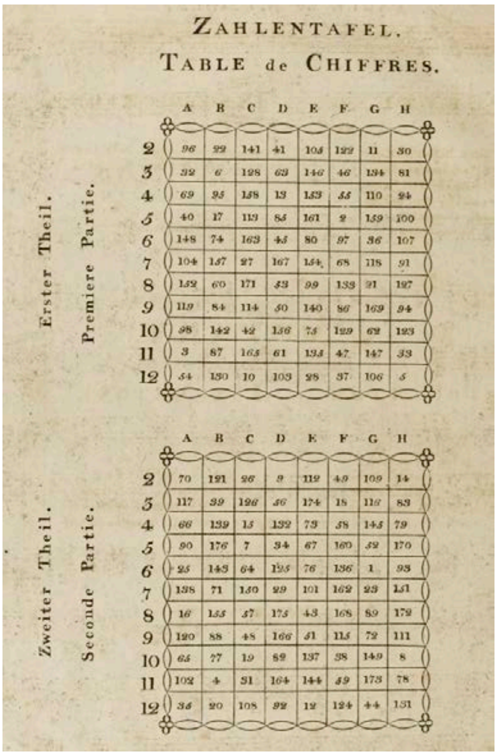
Whether we're talking about isorhythm or phasing, we can hardly speak of algorithms, but rather of devices. In these two examples, however, the premises of the generative approach emerge: from the interaction of two fundamental elements—*talea* and *color*, sound loops, melodic motifs—a larger form is born that exposes a whole series of different configurations.

3 - Musikalische Würfelspiele

The use of randomness in musical creation aroused great enthusiasm in Europe in the second half of the 18th century, through a collection of composition games intended for a musically ignorant audience.

One of the most famous, published by Johann Julius Hummel in 1793 and Nikolaus Simrock in 1796[13], is a system of instructions for composing waltzes which both publishers attribute to Wolfgang Amadeus Mozart, although this authorship is not clearly established[14]. This variant is commonly referred to as *Musikalisches Würfelspiel*—Musical Dice Game—a name that has been extended to the whole family of random musical games of the 18th century.

These games proved very popular in the salons of the time, and there were numerous versions: some twenty are listed, some of which gave rise to multiple editions in different languages[15]. Despite their differences, these *Musikalische Würfelspiele* are based on more or less the same principles: from a predetermined set of musical motifs, the players roll the dice in succession to create an original piece.



ZAHLENTAFEL.

TABLE de CHIFFRES.

Erster Theil. Premiere Partie.

	A	B	C	D	E	F	G	H
2	96	22	141	41	105	122	11	30
3	32	6	128	63	146	46	134	81
4	69	95	158	13	153	55	110	24
5	40	17	113	85	161	2	159	100
6	148	74	163	45	80	97	36	107
7	104	157	27	167	154	68	118	91
8	152	60	171	53	99	133	21	127
9	119	84	114	50	140	86	169	94
10	98	142	42	156	75	129	62	123
11	3	87	165	61	135	47	147	33
12	54	130	10	103	28	37	106	5

Zweiter Theil. Seconde Partie.

	A	B	C	D	E	F	G	H
2	70	121	26	9	112	49	109	14
3	117	39	126	56	174	18	116	83
4	66	139	15	132	73	58	145	79
5	90	176	7	34	67	160	52	170
6	25	143	64	125	76	136	1	93
7	138	71	150	29	101	162	23	151
8	16	155	57	175	43	168	89	172
9	120	88	48	166	51	115	72	111
10	65	77	19	82	137	38	149	8
11	102	4	31	164	144	59	173	78
12	35	20	108	92	12	124	44	131

Figure 1: Table for determining 'Musikalisches Würfelspiel' motifs. Columns correspond to bars, rows to dice outcomes. Illustration from Nikolaus Simrock's 1796 edition, p. 2. Available at: https://gallica.bnf.fr/ark:/12148/bpt6k316645m/f2.item

The version attributed to Wolfgang Amadeus Mozart is based on 176 one-bar motifs. The composition is divided into two parts of eight bars each. For each bar, eleven motifs can be selected corresponding to the eleven possible outcomes of a throw of two dice[16]. Such a system generates 11^{16} possible combinations, i.e. about 46 million billion.

These systems can be assimilated to algorithms: a set of well-defined rules is applied step by step, ultimately producing a musical sequence. Moreover, the use of randomness brings them into the generative realm, as each game session generate as many original compositions. Indeed, due to the combinatorial explosion resulting from successive throws of the dice, the probability of the same sequence being repeated is so low as to be considered zero.

Combinatorics, which according to Pierre Barbaud characterizes the algorithmic[17] approach, was already at the heart of these salon games. At the time, this interest echoed the research of Gottfried Wilhelm Leibniz[18], who defined combinatorics as that part of mathematics that aims to study and formalize relationships between objects. The German mathematician identified this "combinatorial art" with the art of invention, which in his view consisted in creating original arrangements of already known elements in order to discover new truths[19]. A vision that still resonates two centuries later in the permutational art developed by Abraham Moles in 1971[20], a theory which asserts, among other things, the playful aspect of art.

Beyond their interest as precursors of generative music, two aspects of the *Musikalische Würfelspiele* deserve a closer look: their primarily playful dimension, and the nuance they introduce into the notion of randomness, between aleatory and luck.

The relationship between art and play would merit an in-depth study beyond the scope of this paper, but it's worth noting that *Musikalische Würfelspiele* are foremost board games. While they do indeed produce musical pieces, they are not designed as methods of composition or learning, but as entertaining activities whose goal is simply to give players pleasure. It's a very special kind of game, one that completely dispenses with the idea of competition—there can be no winners or losers—and focuses on the wonder that arises when a series of throws of the dice produces a seductive piece of music. It is this playful pleasure that leads Sarah Troche to distinguish two modalities of randomness in *Musikalische Würfelspiele*:

> "The millions of minuets are equal before the dice, but unequal before the quality of the sound results: they are *more or less beautiful*, even if they are all *equally correct*. In other words, ludic pleasure is based on the conjunction of two types of randomness: 'the aleatory', which makes all possibilities equal, and 'luck', which will bring out a successful piece. The aleatory is the structural condition for the emergence of luck, the levelling of data from which the inequality of the fortuitous arises. These musical dice games offer a singular opportunity to consider the relationship between randomness and invention: how does randomness, as an aleatory variable in a self-contained process, contribute to a systematization of invention closely linked to the problem of genius[21]?"

This crucial issue runs through all generative art. For if the multiple actualizations of a generative process, although all conforming to the initial prescriptions, are not equal from an aesthetic point of view, when it comes to extracting one or more singular configurations from the multitude, it is ultimately up to the artist to decide, often on the basis of subjective criteria, judgments of the value of a particular iteration which, at the end of an automated process, give the hand back to the creator.

A problematic attitude for many generative artists, who consider that any subjective choice made after the program has been executed contradicts the very concept of programmed art. As we shall see, this was the subject of a profound disagreement between Pierre Barbaud and Iannis Xenakis in the early 1960s.

4 - Twelve-Tone Music and Serialism

If the music of the 20th century, in the wake of Claude Debussy, is characterized by the overcoming of Romanticism, it is more singularly disrupted in its writing by the principle of atonality, of which Josef Matthias Hauer and Arnold Schöenberg are the main architects. Since the 1910s, Josef Matthias Hauer had been developing this new approach, which he detailed in *Vom Wesen des Musikalischen*—the essence of musical fact—published in 1920[22].

Fundamentally, the idea was to free oneself from the notion of tonality by composing from the twelve semitones of the chromatic scale in such a way that a note could only reappear once the other eleven had been used, a principle that forged the name dodecaphony. For Pierre Barbaud, Josef Matthias Hauer's systemic and combinatorial approach makes him a forerunner of algorithmic music, despite his late recognition:

> "Josef Matthias Hauer's dodecaphonic system aroused little curiosity at the time it was conceived, and the life of its inventor

> was always marked by great material distress. However, through the rigor of his thinking, and the possibilities opened up by his writing system for the mechanization of musical composition, it heralds the most recent research[23]."

An authorship that Moreno Andreatta also affirms, while specifying:

> "The systematic work on hexachord structures, whose exhaustive enumeration [Josef Matthias Hauer] completed in the early 1920s, contributed to the establishment of a truly algorithmic approach to composition, the result of which would be a series of 'dodecaphonic games' or Zwölftonspiele[24]."

In the 1920s, another composer took an interest in twelve-tone music: Arnold Schöenberg, who composed the five pieces of his *Piano Suite* between 1921 and 1923. Here, he took up the principle of series—a particular sequence of twelve sounds—introduced by Josef Matthias Hauer, and placed it at the center of his approach.

Although both composers practice twelve-tone music, their perspectives diverge sharply. While Josef Matthias Hauer strives for a form of abstract music, stripped of expressivity and emotion, behind which the artist disappears, Arnold Schöenberg seeks above all to free himself from the dictates of tonality, while remaining strongly influenced by the notion of Romantic genius.

It was Arnold Schöenberg who set the standard, and it was under his guidance—as well as that of Alban Berg and Anton Webern, who were his students—that twelve-tone music evolved into serialism: the series at the origin of the work goes beyond chromaticism, structures itself into subgroups, explores symmetry or the play of intervals… The series constituted can be used in its original form, but also in reverse order—retrograde series—, in inversion—the notes, represented by numbers from 1 to 12, are replaced by their opposites modulo 12—or in recurrence of inversion[25]. The rules of classical tonal harmony are thus replaced by a whole formal system that draws much of its inspiration from mathematics.

Whether one prefers the approach of Arnold Schöenberg or that of Josef Matthias Hauer, it has to be said that their music, despite its undeniable numerical and systematic components, is neither truly computational nor algorithmic.

5 - John Cage and Randomness

Algorithmic music was born at the end of a decade that saw the almost simultaneous emergence of three musical approaches that can be linked to the generative one: the aleatory experiments of John Cage, the open works of the neo-serialists, and the stochastic music of Iannis Xenakis. Although, at first glance, these practices may be united by their search for the unexpected, their methods and intentions are distinct, and there are many polemics between the proponents of one or the other.

With *Music of Changes*, in 1951, John Cage—who had been a student of Arnold Schöenberg—inaugurated a practice that was to become his hallmark: the introduction of randomness at the heart of the compositional process. For while his earlier pieces, notably the *Sonatas and Interludes* for prepared piano (1948), may present accidental variations in performance, they remain marginal and insignificant:

> "[…] the objects between the strings of the prepared piano are not arranged 'at random', but precisely disposed between certain strings indicated by the score: the aim is to make the piano more percussive, to modify the timbre by integrating more complex sounds, and not to introduce unintentional noises. The very relative element of the unforeseen does not act on the overall structure of the music, which is always composed, but locally in the complexity of the sonorities obtained[26]."

The intention behind *Music of Changes*, a four-part piece for solo piano, is quite different, as demonstrated by the method used. All of the compositional decisions, from the *macro* to the *micro*—the structure of the piece, the method of sequencing the notes, the parameters of the sonic material[27]—are the result of randomly drawn hexagrams from the *Yi Jing*—or *I Ching*—a Chinese divinatory book dating from the Zhou period (circa 1045-256 B.C.), known in English as the *Book of Changes*, which gives John Cage's work its title.

For the American composer is deeply imbued with Eastern influences—Indian music, Hindu spirituality, Zen Buddhism, etc.—and we can see the organization of the work into four parts—"books" in John Cage's terminology—as a reference to the *Four Books* which were, in the 13th century, the final addition to the thirteen classics—*Jing* in Chinese—whose study was compulsory in China until the early 20th century.

The *Yi Jing* is based on the fundamental figure of the trigram, composed of three continuous lines—*yin*—or discontinuous lines—*yang*—arranged one above the other. There are thus eight possible trigrams, and their paired combinations make up the 64 hexagrams of *Book of Changes*.

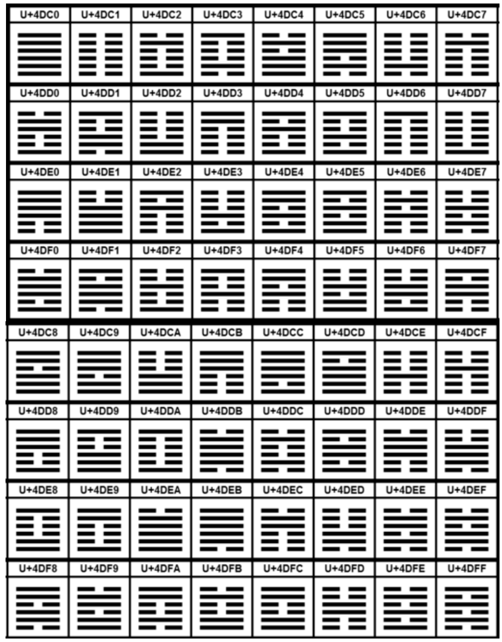


Figure 2: The 64 hexagrams of the Yi Jing, and their corresponding Unicode codes. (Source: Wikimedia Commons)

The determination of a particular hexagram is the result of six successive random draws that determine whether each line is *yin* or *yang*. It should also be noted that each line can be "stable" or "mutable", and that this will influence the interpretation of the oracle, as mutable lines refer to secondary hexagrams[28]. This is why three coins are commonly used to determine each line as follows:

- three heads: old *yin*, or mutable *yin* (probability 1/8)
- two heads, one tail: stable *yin* (probability 3/8)
- two tails, one head: stable *yang* (probability 3/8)
- three tails: old *yang*, or mutable *yang* (probability 1/8)

The *Yi Jing* thus comprises 64 hexagrams, each of which refers to a specific “judgment”, sibylline sentences supposed to shed light on the situation about which the oracle is being consulted. The judgment is completed according to the distribution of the mutable lines. Finally, the mutable lines are transformed into their opposites, generating a new hexagram whose judgment is said to predict the evolution of the situation.

It’s worth noting here that John Cage understood that the *Yi Jing*, although too often reduced to its divinatory aspects, is above all the vehicle for a philosophy of acceptance of change, a recognition of the intrinsically fluctuating nature of the universe, of the vanity of the desire for stability and of the impossibility of achieving a situation of mastery.

John Cage’s use of the *Yi Jing* thus expresses his adherence to Zen principles. But it also reflects his desire to impose severe constraints on himself, to neutralize his subjectivity as a composer in order to achieve his goal of an impersonal work:

> “If Cage maintains recourse to the *Yi Jing* when he has much faster methods at his disposal for more indeterminate results, it is perhaps precisely because this method demands time, patience and self-sacrifice in the precise, slow accomplishment of a repetitive task. For the slowness, repetitiveness and monotony of the gesture have an effect on those who agree to submit to it[29].”

As Sarah Troche also points out, whereas the *Musikalische Würfelspiele* are limited to a combinatorial play of pre-written elements in which the composer’s touch can be found, the randomness that John Cage introduces at the heart of the creative process produces a music freed from ego, and thus contradicts the idea of inspired genius:

> “In one case, randomness produces a composition from pre-composed elements, the general idea of the work comes first and determines the possibilities; in the other, we are in an experimental process that cancels any intentional relationship between cause and effect and produces a globally unpredictable result (we compose and then we hear, these two moments being separate from each other[30]).”

6 - Integral Serialism and Set Theory

While John Cage introduced chaos at the heart of musical creation, other composers turned to an eminently mathematical formalization of music, pushing serial technique to its paroxysm with integral serialism. In the United States, it was Milton Babbitt who embodied this approach, while in Europe it was Pierre Boulez or Karlheinz Stockhausen.

Until then, the series had only affected the pitch of a note, but the neo-serialists applied the principle to other sonic characteristics, an idea that Olivier Messiaen is known to have originated:

> “In 1942, while teaching serial music to his students Nigg, Boulez and Martinet, he advised them to write serial works not only with series of frequencies, but also with series of timbres, of intensities and of durations. But it wasn’t until 1949 that he put his fruitful idea into practice for the piano in ‘Mode de valeurs et d’intensités’. Immediately, all the young people were dazzled and launched into compositions imitating or paraphrasing this work[31].”

But the first pieces to belong to integral serialism are most commonly considered to be Milton Babbitt’s *Three Compositions for Piano*[32], written in 1947. The composer also produced numerous theoretical texts, some of whose concepts were taken up by Allen Forte in a musicological analysis known as 'set theory' after the author's application of the mathematical set theory to atonal music[33].

Set theory introduces a strong algebraic dimension into musical composition, as illustrated by the analysis of manipulations applicable to a series—transposition, inversion—and the possible relationships between two series—inclusion, complementarity—or the concepts of interval vectors and form matrices. The near disappearance of the boundary between music and mathematics is most clearly demonstrated by the existence of the hexachord theorem[34], also known as Milton Babbitt's theorem, which has been proved many times, and which Moreno Andreatta expresses in everyday language as follows:

> "In a hexachord and in its complementary, the multiplicity of occurrence of each interval (an interval always being equivalent to its complementary) is the same[35]."

We will not go into the practical applications of this theorem or the notion of "combinatoriality" introduced by Milton Babbitt—who applies it particularly to hexachords—as they go far beyond the scope of this study[36]. The main purpose of these examples is to demonstrate that the resolutely mathematical orientation of American-style integral serialism constitutes a step towards a computational and algorithmic approach to music.

7 - The Open Work Among the European Neo-Serialists

On the other side of the Atlantic, the algebraic component consubstantial with serial music is less invested, in favor of an approach marked by the search for the unexpected. Inspired by Stéphane Mallarmé, Pierre Boulez, and the neo-serialists with him, introduced a form of variability into the musical work by including in their scores various options for the performer to choose from[37]. This is the case, for example, in Karlheinz Stockhausen's *Klavierstück XI* and Pierre Boulez's *Troisième sonate pour piano*, written in 1956 and 1957 respectively[38].

Pierre Boulez's approach is that of a "directed" randomness, a freedom of selection or execution that the composer grants in order to put "back into the creative circuit the performer himself, who for many years had only been asked to play the text as 'objectively' as possible[39]". This directed randomness is also a form of response to the "inadvertent randomness" that Pierre Boulez accused John Cage of. For the French composer does not understand the reasons for the meticulousness with which John Cage sets up his random systems, only to relinquish all control over the final form. And he makes this clear with undisguised contempt:

> "[…] the event happens as it can, without control (voluntary absence, though not meritorious, by impotence), BUT within a certain established network of probable events, because randomness must have some contingency. Why, then, choose the network so meticulously, why not leave the network itself to inadvertence? That's what I've never been able to work out. We're playing a half-fair game, but we have the merit of not hiding it. It's a nicely arranged artificial paradise where I believe that dreams are never very miraculous; this kind of narcotic does indeed protect you from the sting inflicted by any invention, and you have to admit that its action is overly soothing, sometimes hilarious, just like what hashish addicts describe[40]."

Here, then, is John Cage, clearly targeted, though not mentioned by name, reduced to the rank of a hippie lover of artificial paradises, his "philosophy tinged with Orientalism [masking only a] fundamental weakness in compositional technique[41]". This accusation of resignation from the role of composer, which Pierre Boulez levelled at John Cage was also levelled at him by Iannis Xenakis, for whom the process of directed randomness was a mere pretense:

> "The performer is a highly conditioned being, so we cannot accept the thesis of unconditional choice, of the performer as a roulette wheel. […] The composer resigns when he admits several possible and equivalent circuits. In the name of the scheme, the problem of choice is betrayed, and it is the performer who is promoted to the rank of composer by the composer himself. There is therefore a substitution of authors[42]."

8 - The Stochastic Music of Iannis Xenakis

Iannis Xenakis's approach to randomness is indeed resolutely different from the two previous ones: with the notion of stochastic music[43], he introduces randomness governed by probability theory[44], an idea inherited from the development of statistical physics and the only way, i his opinion, to get music out of the rut of integral serialism:

> "Stochastics studies and formulates the laws of large numbers, as well as the laws of rare events, the various random processes, and so on. This is how, in 1954, out of the dead end of serial music, among other things, was born a form of music based on the principle of indeterminism, which two years later I called stochastic music. The laws of the calculus of probabilities entered the composition by musical necessity[45]."

In *Musiques formelles*, Iannis Xenakis details a model of musical creation in eight phases[46], specifying that their order is not rigid and their application is often unconscious: on the basis of his intuitions (phase a), the composer defines sound beings—instrumental sounds, electronic sounds, noises, etc.—and their associated symbols (phase b). The next two phases concern the structure of the piece and the constitution of musical motifs—*macro* and *micro* composition (phases c and d). Next comes the sequential programming of the previous two phases (phase e), followed by the calculations, as schemes and patterns requires algebraic, logical, or stochastic operations (phase f). The results of the first six phases are then formalized in symbolic form: traditional partitions, numerical listings, diagrams, graphical partitions, etc. (Phase g). This leads to the final stage: the embodiment in sound of the work thus composed (phase h).

An initial comment is in order: this model is very similar to a program, and could easily be presented in the form of a flow chart. Iannis Xenakis's formalization thus comes close to the approach of algorithmic music, and the composer explicitly states that the last three phases are intended to be "taken in hand by computers". He also expressed his regret that, in his time, the final embodiment could only be achieved through human operators, while prophesying that "in the very near future […] advanced mechanization [would] do away with orchestral performers or tape recorders, and [would]—take over the mechanized production of sound beings and their transformations[47]."

However, Iannis Xenakis's approach differs from that of Pierre Barbaud, a key figure in algorithmic music for both his theoretical contributions and his compositions. Firstly, because the programmatic approach to music is only one facet of Iannis Xenakis's work, whereas it is the central axis of Pierre Barbaud's work from 1958 onwards.

Secondly, and most importantly, because the former, when using a formal model to generate a score, allows himself to retouch it, whereas the latter categorically excludes this option.

In *Nomos alpha* (1966), Makis Solomos[48] measured a deviation rate between the formalized score and its final version of 18.5%. These discrepancies can be explained in different ways: some are simple mistakes—in transcription or calculation—others are corrections deemed necessary for errors in the theoretical model, and lastly others are changes to details based on entirely subjective aesthetic preferences. In so doing, Iannis Xenakis fully accepts that the hand may return at the end of the computational process:

> "What we obtain by calculation always has its limits. It lacks an internal life, unless you use very complicated techniques. Mathematics produces structures that are too regular, below the requirements of the ear and the intelligence. The great idea is to be able to introduce randomness in order to break the periodicity of mathematical functions, but we are only at the beginning. The hand lies between randomness and calculation. It is both the executor of the mind—very close to the head—and an imperfect tool[49]."

This practice is contested by Pierre Barbaud, who sees it as an abuse of power, but also a sign of the imperfection of Iannis Xenakis's formal models—a disagreement not unlike that between Gustav Matthias Hauer and Arnold Schöenberg. For the algorithmic composer, any choice made after the computational process is unjustifiable, because it contradicts the initial concept of creation:

> "The solutions received at the output of the machine are decoded and, in the case of Xenakis, analyzed and filtered: the unplayable is eliminated, pure reasoning clothed in sensitivity… Here the composer has the last word. This intervention is no longer desirable for algorithmic musicians who, prior to and in parallel with Xenakis's research, developed research based on the use of a more complex language and allowing a more universal and total transcription of the program and the data[50]."

9 Pierre Barbaud and the Algorithmic Music

The purpose pursued by Pierre Barbaud is clearly established through the substantial corpus of his theoretical writings. The aim was to return to "pure" music, i.e. music based on "ingenious arrangements of sounds[51]", and to break away at last from the cult of expression inaugurated by the vogue for opera, which had subjected music to the domination of passion:

> "In the minds of those who make it or listen to it, music has definitively become the medium of an emotional scenario, the pretext for expressing who knows what torments, regardless of the invention that presides over the arrangement of the sounds that make up the work[52]."

To those who are indignant about the use of computers for creative purposes, Pierre Barbaud argues that the compositional process, whether it be "serial music, post-serial music or even, why not, just plain music", carries within it the seeds of algorithmic music:

> "We implicitly submit to a system of thought designed to produce a result, which we hope will be unpredictable, from a practically infinite number of possible results considered to be banal in the name of criteria yet to be defined. […] Writing

> music is above all a matter of defining the limits within which the combinatory imagination of the elements brought into play can manifest itself. […] It is the most lucid possible formulation of these constraints that characterizes algorithmic music in the first place[53]."

For Pierre Barbaud, algorithmic music is perfectly conceivable without the computer, as demonstrated by Michel Philippot's compositions for "imaginary machine[54]", a concept taken up a few years later by his friend Vera Molnár. However, there were two reasons for using computers: the neutrality of the machine, which defused the temptation to retouch in the manner of Iannis Xenakis, and above all the impossibility, when the program grew in size, of carrying out the necessary calculations by hand.

Pierre Barbaud's stance can sometimes seem overly rigorist, as in his criticism of Iannis Xenakis. Yet it rests on solid conceptual foundations, inherited from Gustav Matthias Hauer: the rejection of music dictated by emotion in favor of composition built on the exercise of reason:

> "It is about, by means of a machine, making a music of a machine, that is to say where the laws of logical thought, exercised on a set of data within a program, are implacably respected, without the intervention, at any moment, of a decision that is not the result of a premeditated calculation[55]."

Conclusion

Pierre Barbaud's influence extends beyond the field of music. His meeting with Gottfried Honegger at the end of the 1960s led the latter to introduce randomness into his working process and then to use the computer in preparatory phases[56]. The two men would later appear together at the same event organized by the Groupe Art et Informatique de Vincennes (GAIV) in Angers in 1974[57]. It was also after making the acquaintance of Pierre Barbaud in 1967 that Manfred Mohr—also inspired by the aesthetics of Max Bense—abandoned his painting practice to turn definitively to algorithmic visual creation from 1969 onwards[58].

The debates opened up by the practice of algorithmic music fed into those of the nascent computer art. Thus the different modalities of the unexpected, the opposition between randomness and luck, the organized variability of a program's output, the full appropriation of the computer, the mobilization of scientific advances, the questioning of the intrinsic infallibility of the artist as a demiurge or the search for an objective aesthetic, irrigated the theoretical substratum on which computer art would flourish.

Notes and References

[1] In this paper, the label “computer art” refers less to an art form that uses computers than to the decade in which this usage emerged: from the few years preceding the historic 1965 exhibitions by Georg Nees - *Computergrafik* in Stuttgart — and A. Michael Noll and Bela Julesz - *Computer-Generated Pictures* in New York — to the early 1970s.

[2] Although the liberal arts were not formalized until the 3rd century, the ancient Greeks, and later the Romans, already considered music to be an indispensable part of human education, on a par with arithmetic and rhetoric. Around 270 AD, Porphyry classified the seven liberal arts into two categories. The trivium covers the arts related to language - grammar, rhetoric and dialectics - while the quadrivium covers those related to mathematics: arithmetic, geometry, astronomy and music.

[3] There are precedents, such as Martin Klein and Douglas Bolitho's short *Push Button Bertha* in 1956, but these have remained little or undocumented. On this subject, see Moreno Andreatta's article:

ANDREATTA, Moreno, 2013. Musique algorithmique. In: DONIN, Nicolas and FENEYROU, Laurent (eds.), *Théories de la composition musicale au XXe siècle*. Lyon : Symétrie. pp. 1239-1268. Symétrie recherche, 21.

[4] Theorized at the beginning of the 20th century by Andreï Markov, these chains are sequences of numbers in which the value at time *t+1* is determined by the value at time *t* according to the probabilities - fixed in advance in a transition matrix - of the different possibilities. The chain is said to be of order one when only the previous state is used to determine the next state. If both previous states are taken into account, the chain is of order two, and so on.

The advantage of these chains is that they can be used to set possibilities, as well as obligations and prohibitions. For example, taking the three notes C E G, the transition matrix can impose that a C will be followed 75% by an E and 25% by a G - so never by a C -, that an E will always be followed by a G, and that a GI will be followed 50% by another G, 25% by a C and 25% by an E.

[5] MANNING, Peter, 1985.*Electronic and computer music*. [1987]. Oxford: Clarendon Press. p. 217 ff.

[6] GUTTMAN, Newman, 1957. *The Silver Scale* [online]. New York : Bell Labs, 1957. [Accessed September 19, 2024]. Available at:

https://www.youtube.com/watch?v=PM64-lqYyZ8

[7] GUTTMAN, Newman, 1957. *Pitch Variations* [online]. New York : Bell Labs, 1957. [Accessed September 19, 2024]. Available at:

https://www.youtube.com/watch?v=TExxv1-j42Y

[8] VARIOUS ARTSISTS, 1962. *Music From Mathematics* [online]. [12’’ vinyl]. Decca / Bell Labs, 1962. [Accessed September 19, 2024]. Available at:

https://www.youtube.com/watch?v=pusmHRTHbXQ

This is a second edition; the first is poorly documented, dating from 1960 or 1961 according to sources.

[9] We do not include here hybrid systems such as GROOVE - also developed in 1969 by Max Mathews for Bell Labs - in which the computer controls analog synthesizers. Although these systems operate in real time, they are not algorithmic synthesis.

[10] ROBERT, Christophe, [no date]. Qu'est-ce que la musique algorithmique ?. *MusicAlgo* [online]. [Accessed April 19, 2023]. Available at:

http://musiquealgorithmique.fr/quest-ce-que-la-musique-algorithmique/

Christophe Robert is a doctoral student at the Université Côte d'Azur. His research, entitled *Esthétique générative et musique acousmatique : une relation en question* focuses in particular on algorithmic music. See : https://theses.fr/s318612

[11] COPE, David, 2000. *The Algorithmic Composer* [online]. Madison, Wisconsin : A-R Editions. [Accessed July 4, 2024]. Available at:

http://archive.org/details/DavidCopeTheAlgorithmicComposer pp. 3-4

[12] For more details on these two pieces, see Steve Reich's *Notes on Compositions 1965-1973* (pp.49-53) in :

REICH, Steve, 1974. *Writings about music* [online]. Halifax : Press of Nova Scotia College of Art and Design. [Accessed July 3, 2024]. Available at:

http://archive.org/details/writingsaboutmus0000reic

[13] MOZART, Wolfgang Amadeus [attributed to], 1796. *Instruction pour composer autant de Walzer ou Schleifer que l'on veut, par le moyen de 2 dés sans savoir la musique ou la composition* [online]. Bonn: Nikolaus Simrock. Available at:

https://gallica.bnf.fr/ark:/12148/bpt6k316645m/f1.item

[14] HEDGES, Stephen A., 1978. Dice Music in the Eighteenth Century. *Music & Letters*. 1978. Vol. 59, no. 2, pp. 180-187. DOI 10.1093/ml/59.2.180. p. 183

[15] TROCHE, Sarah, 2015. *Le hasard comme méthode: figures de l'aléa dans l'art du XXe siècle*. Rennes : Presses Univ. de Rennes. Aesthetica. p. 194

[16] 2 parts x 8 measures x 11 variations = 176 motifs

Note by the way that the 11 possible motifs for each measure are not equiprobable: 1/36 for motif 2 versus 1/6 for motif 7.

Numerous websites now offer implementations of these games, enabling the score to be generated in a fraction of a second, and even played. See for instance:

https://dice.humdrum.org/ [Accessed July 5, 2024]

[17] BARBAUD, Pierre, 1960. Musique algorithmique. *Esprit* [online]. 1960. N° 280, pp. 92-97. [Accessed July 5, 2024]. Available at :

https://www.jstor.org./stable/24255071 p. 92

[18] LEIBNIZ, Gottfried Wilhelm, 1666. *Dissertatio de arte combinatoria* . Lipsiae : J. S. Fikium et J. P. Seaboldum.

[19] COUTURAT, Louis, 1901. *La logique de Leibniz : d'après des documents inédits.* [online]. Paris : Félix Alcan. [Accessed July 5, 2024]. Available at :

https://gallica.bnf.fr/ark:/12148/bpt6k110843d p. 279

[20] MOLES, Abraham A., 1971. *Art et ordinateur*. [New revised and expanded edition 1990]. Paris : Blusson.

[21] TROCHE, Sarah, 2015. *Le hasard comme méthode: figures de l'aléa dans l'art du XXe siècle*. Rennes : Presses Univ. de Rennes. Aesthetica. p.196

[22] HAUER, Josef Matthias, 1920. *Vom Wesen des musikalischen, ein Lehrbuch des atonalen Musik* [online]. Leipzig : Waldheim-Eberle. [Accessed July 14, 2024]. Available at:

http://archive.org/details/vomwesendesmusi00hauegoog

[23] BARBAUD, Pierre and LENGAGNE, Rémi, 1999. Musique contemporaine - La musique sérielle et le dodécaphonisme. *Encyclopædia Universalis* [online]. [Accessed July 8, 2024]. Available at:

https://www.universalis.fr/encyclopedie/musique-contemporaine-la-musique-serielle-et-le-dodecaphonisme/

[24] ANDREATTA, Moreno, 2013. Musique Algorithmique. In: DONIN, Nicolas and FENEYROU, Laurent (eds.), *Théories de la composition musicale au XXe siècle*. Lyon : Symétrie. pp. 1239-1268. Symétrie recherche, 21. p. 1242

Note Josef Matthias Hauer's invocation of the notion of play.

[25] AGON, Carlos and ANDREATTA, Moreno, 2008. La musique mise en algèbre. *Pourlascience* [online]. November 2008. [Accessed July 9, 2024]. Available at:

https://www.pourlascience.fr/sd/mathematiques/https:https://www.pourlascience.fr/sd/mathematiques/la-musique-mise-en-algebre-1879.php

[26] TROCHE, Sarah, 2015. *Le hasard comme méthode: figures de l'aléa dans l'art du XXe siècle*. Rennes : Presses Univ. de Rennes. Aesthetica. p. 378

[27] CAGE, John, 1961. *Silence: lectures and writings* [online]. Middletown, Connecticut. Wesleyan University Press. [Accessed July 10, 2024]. Available at :

http://archive.org/details/silencelecturesw1961cage p. 18

[28] WILHELM, Richard, 1968. *Yi king, le livre des transformations - Livre I*. [new revised and updated edition 1973]. Paris: Librairie de Médicis. p. 12.

[29] TROCHE, Sarah, 2012. Structure et méthode dans la musique de John Cage : une discipline d'attention. *Nouvelle revue d'esthétique*. 2012. Vol. 9, no. 1, pp. 91-104. DOI 10.3917/nre.009.0091. p. 101

[30] TROCHE, Sarah, 2015. *Le hasard comme méthode: figures de l'aléa dans l'art du XXe siècle*. Rennes : Presses Univ. de Rennes. Aesthetica. p. 400

[31] XENAKIS, Iannis, 1955. La crise de la musique sérielle. *Gravesaner Blätter*. 1955. Vol. 1, n° 1. p. 2

[32] BABBITT, Milton, 1947. *Three Compositions For Piano*, ARNONE, Augustus (performer), [online]. [Piano solo]. Boelke-Bomart, 1947. [Accessed 20 September 2024]. Available at:

https://www.youtube.com/watch?v=AcpMR4r2kcQ

[33] FORTE, Allen, 1973. *The structure of atonal music*. New Haven : Yale University Press.

[34] A hexachord is a collection of six different notes - known as *pitch classes* in set theory terminology - from among the twelve of the chromatic scale. For more details, see :

WHITTALL, Arnold, 2008. T*he Cambridge introduction to serialism*. Cambridge: Cambridge University Press.

[35] ANDREATTA, Moreno, 2003. *Méthodes algébriques dans la musique et la musicologie du XXème siècle : aspects théoriques, analytiques et compositionnels* [online]. PhD thesis in musicology. Paris : École des Hautes Études en Sciences Sociales (EHESS). p. 37 [Accessed 19 September 2024]. Available at:

https://theses.hal.science/tel-00004074

[36] For more details on the mathematics of set theory and its musical translations, see the article by Moreno Andreatta and Stefan Schaub introducing the Autour de la set theory colloquium at IRCAM in 2003: ANDREATTA, Moreno and SCHAUB, Stéphan, 2003. *Une introduction à la Set Theory : Les concepts à la base des théories d'Allen Forte et de David Lewin*. Musurgia [online]. 2003. Vol. 10, No. 1, pp. 73-92. [Accessed 19 September 2024]. Available at: https://www.jstor.org/stable/40591272

[37] This practice of the open work was theorised in 1962 by Umberto Eco who, relying heavily on Stéphane Mallarmé, extended the concept from James Joyce or Franz Kafka to Marcel Duchamp, Alexander Calder, Bruno Munari or even Jesús-Rafael Soto.

ECO, Umberto, 1962 [1965 for the first french edition]. *L'œuvre ouverte*. [1979]. Paris: Editions du Seuil. Points, 107.

[38] For more details on these pieces:

Karlheinz Stockhausen, *Klavierstück XI* (1956) see:

https://brahms.ircam.fr/fr/works/work/12123/

Pierre Boulez, *Troisième sonate pour piano* (1957) see:

https://www.musiquecontemporaine.info/zpiece-Boulez-TroisiemeSonate.php

[39] BOULEZ, Pierre, 1966. *Relevés d'apprenti* [online]. Paris : Seuil. p. 55. [Accessed 4 July 2024]. Available at:

https://archive.org/details/relevesdapprenti0000unse/page/n7/mode/2up

[40] *Ibid.* pp. 41-42

[41] *Ibid.* p. 41

[42] Xenakis, Iannis, 1963. *Musiques formelles: nouveaux principes formels de composition musicale*. Paris: Richard-Masse. La Revue musicale. p. 48

[43] The modern meaning of stochastic is "which depends on, which results from chance", and more specifically in the mathematical field "which comes under the domain of randomness, the calculation of probabilities". The word comes from the ancient Greek *stokhastikós*, which means "that aims well, that tends directly towards" but also "skilful at conjecturing, penetrating". It derives from *stokhastês*, meaning diviner.

For the Greek etymology, see : BAILLY, Anatole, 1935. *Greek-French dictionary* [online]. [Accessed 27 February 2024]. Available at:

http://archive.org/details/BaillyDictionnaireGrecFrancais or the online application: https://bailly.app/stochastikos

[44] In the same way, Iannis Xenakis conceived "strategic music" and "symbolic music" on the basis of game theory for the former and (mathematical) set theory for the latter. See chapters 3 and 5 respectively of *Musiques formelles* (*op. cit.*).

[45] *Ibid.* p. 19.

[46] *Ibid.* pp. 33-34

[47] *Ibid.* p. 34

[48] Solomos, Makis, 2023. Bricolage. *Site officiel Iannis Xenakis* [online]. 2023. [Accessed 11 July 2024]. Available at:

https://www.iannis-xenakis.org/en/bricolage/

[49] Iannis Xenakis in: REY, Anne, DUSAPIN, Pascal, Si Dieu existait, il serait bricoleur, *Le Monde de la Musique* no. 11, May 1979, pp. 92-97. Quoted by Makis Solomos in

SOLOMOS, Makis, [no date]. Iannis Xenakis. *Ressources.IRCAM* [online]. [Accessed on 11 July 2024]. Available at:

https://brahms.ircam.fr/fr/iannis-xenakis#parcours

[50] BARBAUD, Pierre et PHILIPPE, Robert, 1969. L'ordinateur et la musique. *Communication et langages*. 1969. N° 3, pp. 17-25. DOI 10.3406/colan.1969.3744. pp. 20-21

[51] BARBAUD, Pierre, 1960. Musique algorithmique. *Esprit* [online]. 1960. No. 280 (1), pp. 92-97. p. 93. [Accessed on 3 July 2024]. Available at:

https://www.jstor.org/stable/24255071

[52] *Ibid.*

[53] BARBAUD, Pierre and PHILIPPE, Robert, 1969. L'ordinateur et la musique. *Communication et langages*. 1969. No. 3, pp. 17-25. DOI 10.3406/colan.1969.3744. pp. 22-23

[54] PHILIPPOT, Michel, 1963. Composition pour double orchestre (1960). In: XENAKIS, Iannis, *Musiques formelles: nouveaux principes formels de composition musicale*. Paris: Richard-Masse. pp. 53-56 La Revue musicale.

[55] BARBAUD, Pierre and PHILIPPE, Robert, 1969. L'ordinateur et la musique. *Communication et langages*. 1969. No. 3, pp. 17-25. DOI 10.3406/colan.1969.3744. p. 24

[56] Extract: Gottfried Honegger: Homo scriptor, [no date]. *Les presses du réel* [online]. [Accessed 24 July 2024]. Available at:

https://www.lespressesdureel.com/extrait.php?id=490&menu=

[57] LENGLOIS, Camille, 2023. Chronologie des manifestations du GAIV. In : *Le Groupe art et informatique de Vincennes (GAIV)*. Éditions du Centre Pompidou. pp. 70-81. Recherches. p. 73

[58] MOHR, Manfred, [no date]. Statement. *Manfred Mohr official website* [online]. [Accessed 12 July 2024]. Available at:

http://www.emohr.com/tx101.html

Art, humanity, and beyond. Towards new creative paradigms

Prof. P. L. Capucci
Noema, Italy; EU Expert on Science, technology and Humanities, LABA Rimini
https://noemalab.eu
https://capucci.org
e-mail: plc@noemalab.eu

Over the last half century, we have witnessed the rapid succession of technologies, devices, and tools that have profoundly changed professions, the sciences, the economy, communication, and daily life. This rapid evolution has also influenced the methods of producing images, music, and audiovisuals, generating new creative opportunities and professional skills. In the realm of artistic institutions, teaching these new tools has become fundamental in the "applied arts" courses, complementing traditional methods but has also enriched artistic disciplines, expanding their poetics, horizons, and fields of application.
Artificial Intelligence, in particular generative Artificial Intelligence, is today a much-debated topic, with important spin-offs in many economic and production sectors, destined to have a profound impact on contemporary culture, society, and the arts. Generative AI has entered into the production of images, videos, music, architecture, poetry, and literature, and new forms of expression, design, and communication have emerged. Generative AI has also proven to be useful in many scientific realms, including environment studies and climate research.

1. Generative AI issues and artistic education

Generative AI has revealed great opportunities in many sectors and has attracted enormous investments, while also raising criticisms and concerns, some summarized as follows:
1) Since generative AI allows the creation of texts, images, musical pieces, and audiovisuals from textual inputs without requiring any technical knowledge in those sectors, it leads to rethinking professional skills by introducing new operational modes.
2) Since generative AIs are trained on huge amounts of data, including copyrighted material, they are accused of improperly using such materials without any authorization or compensation for rights holders.
3) The data on which generative AIs are trained are subject to biases, prejudices, linguistic and cultural conditioning that affect the results.
4) Generative AIs are not autonomous and infallible: in generating outputs, "hallucinations" [1] and errors can occur

that require verification and correction, making the outcomes not immediately usable for key purposes.
5) It’s easy to create fakes and spread false information to deceive and influence politics, the economy, and social life.
6) The activity of these applications consumes enormous amounts of energy and resources, including water for cooling calculation centers, with a significant ecological impact.
7) Generally, AI technologies are owned by a few large private companies whose primary purpose is profit and whose power and activities, in the absence of national and supranational institutional regulatory interventions, can become uncontrollable.
8) Neural networks, which are the basis of generative AI, are currently the most powerful tools of Artificial Intelligence, but we can’t understand how they arrive at their results. Even if the system returns coherent outcomes with respect to the initial data, what a neural network does while performing tasks, training, and operating is inaccessible even to developers. Similar to the neural activity of the brain they try to imitate, the operations of a generative AI system are inherently opaque, hidden inside a black box [2].

For these aspects, some experts even come to consider Artificial Intelligence as a threat to the very existence of humanity… All this in the face of technologies that today constitute an economic gamble because they require enormous investments and, being at a loss, must be continuously refinanced.

However, even though generative AI is still in its infancy, it represents an important stage in technological evolution and could mark the beginning of an epochal transformation. Combined with other computer tools, AI is a powerful device for analyzing, managing, and interpreting the immense amount of digitally recorded memories of human culture to extract information, connections, and knowledge from these data almost in real-time, tasks that are simply unattainable for human capabilities and timelines. But humanity’s relationship with “intelligent devices” involves a broader dimension; it is part of a general paradigm shift that requires questioning basic issues such as the type of society we want, the rules and ethics on which to base it, the fairness of the economic and social system, the relationships between different cultures and visions, the relationship with the environment and with what is not human, up to the management of the issues outlined above and to the lives of individuals.

Artistic institutions have begun to introduce generative AI within teaching, recognizing the importance of artistic disciplines in the field of research and innovation. In 2023, the author of this presentation designed and held a pilot course on teaching generative AI in artistic institutions as part of the EU4ART_*differences* project supported by the European research and innovation program Horizon 2020. The project involved twenty-five undergraduate, graduate, and doctoral students from several European art institutions to discuss historical-artistic and technological aspects, teaching, operational and design methods, limits, and issues, with a final practical part dedicated to software and the creation of images. The course required a transdisciplinary approach and the use of a plural language between arts, sciences, and technologies that was understandable to students of the humanities even on the most technical and scientific topics.

2. Errors, semantic disorientation, unpredictability, incoherence, and copyright

Generative AI can produce interesting results without requiring specialized technical skills, “democratizing” access to content creation in a path that in the field of images can be traced back to the invention of photography. At a basic level, it can replace human technical intervention more radically than previous technologies. At an advanced level, for complex targeted objectives, generative AI can be an important tool in many sectors: a meta-device to compose, with the issues outlined above, a surprising extract of the cultural memory on which it is trained and interrogated. Similar to other tools, to generate interesting results, it requires knowledge in various disciplines even very distant from the technical and technological dimension. This is a crucial point for artistic institutions that want to integrate generative AI into their educational programs. As an evolving cultural tool, it can support students' creativity. In its highest sense, thanks to the ability to query enormous cultural deposits, it could be considered as a sort of philosophical device.

An interesting aspect from an artistic perspective is the errors that generative AI systems introduce in pictures: distorted hands, faces, and bodies, improbable postures, and unreal objects are the most evident. The problem is that AI has not a body, so it is necessary to “describe” and train through symbols (language, texts, and pictures) the huge variability of body situations and behaviors that do not rely on symbolic-based knowledge but on body-based knowledge. Some of these situations are easy to describe through symbols, but for others it is very complicated or impossible. Since a robot has a body, it could be interesting to connect generative AI with robotics to enhance the learning of the body issues.

A further interesting aspect is the unpredictability and incoherence of the generative AI systems. The images that are generated depend on a huge number of variables, from the online and offline many applications and their different versions to the models (the checkpoints) and the versions that are used. For instance, on the online platform Huggingface, one of the two major Open Source communities that grew around Stable Diffusion (the other is CivitAI), by the time of this text are listed 43,631 text-to-image models. Moreover, the most complete online and offline applications present plenties of options and functions that can greatly influence the final result. Literally, the same software, the same prompt, the same model release, and the same options in different days can generate very different results!

But beyond the general technical aspects, unpredictability is also present in the making of the pictures. Traditional paintings and drawings are sort of works in progress: the artist has control over the image, that is on the composition of space, on the disposition of the forms on the support, on their creation and evolution in time, on colors, and so on. Roughly the same happens, although in different ways, with technological tools in digital painting, 2D and 3D pictures: the artists have control over composition, forms and their evolution, colors. In generative AI pictures, there is a total lack of control instead: after the prompt, the rendering phase is a black box with no option to control. Of course it is possible to orient the image contents with a prompt, eventually making it very descriptive, but the final result is a surprise.

Imagining a precise correspondence between text and image is a fallacy. Text and image are two different realms; the match between them can be very variable. Images evidently represent situations but cannot describe their context, the story, or the reasons for their making and use. Conversely, writing can

clearly describe situations, contexts, and stories but cannot represent them; it is the reader's imagination that creates the pictures. Even in the most accurate text description, there is large room for visual uncertainty, making so-called "prompt engineering" very relative. This makes generative AI similar to human behavior: in the interaction among people, the same sentence can be interpreted and imagined in different ways from different people, depending on a number of factors (context, culture, language, knowledge, situation, history, ways of expression, and so on). Every student in a class will create a different picture to illustrate a written text or a speech.
One more interesting aspect concerns those images devoid of coherent representational qualities, improbable, or semantically indefinable, which I have called the images of the "almost real."

1. Picture generated by the author of this paper with Stable Diffusion

For example, I find this picture intriguing [Fig. 1]: it is literally full of objects, but beyond a generic description of the environment, it is difficult to recognize what is represented, which objects are in the scene; none are recognizable as representations of existing objects. They are images that defy description and generate semantic disorientation, recalling poetic modes and stylistic features typical of Surrealism.

An additional layer of uncertainty is laid by the legal issues related to copyright in the images created with Generative AI. The U.S. Copyright Office has twice rejected the registration request for an image created using Generative AI:

«But copyright law only protects "the fruits of intellectual labor" that "are founded in the creative powers of the [human] mind." (the Office will not register works "produced by a machine or mere mechanical process" that operates "without any creative input or intervention from a human author" because, under the statute, "a work must be created by a human being").» [3, 4]

This decision was reiterated in 2023. [5]

2. Jason M. Allen, Théâtre D'opéra Spatial*, 2022, picture created with Midjourney*

However, AI made works can be awarded. The image by Jason M. Allen, *Théâtre D'opéra Spatial*, created in 2022 with Midjourney [Fig. 2], won the first prize at the Colorado State Fair Digital Arts Competition [6], raising many criticisms but also highlighting the growing importance of generative AI in the artistic domain. In the United States, at least for now, generative AI images cannot be copyrighted but can win artistic competitions.

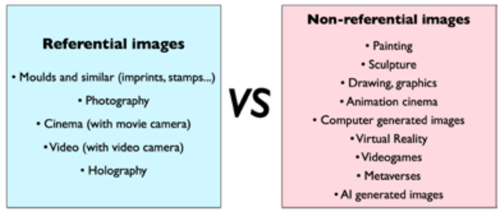
3. Algorithms and simulation

Thanks to generative AI, it is possible to obtain images with impressive photorealism – pictures that look like photographs but are not *referential*, that is they are taken from real physical subjects, but are entirely invented. *The Electrician* [Fig. 3], a picture presented by Boris Eldagsen, won the "Sony World Photography Awards 2023" but was created using generative AI. In rejecting the award, Eldagsen argued:

«AI images and photography should not compete with each other in an award like this. They are different entities. AI is not photography. Therefore, I will not accept the award. We, the photo world, need an open discussion. A discussion about what we want to consider photography and what not. Is the umbrella of photography large enough to invite AI images to enter – or would this be a mistake?» [7]

3. Boris Eldagsen, The Electrician*, detail, 2022 (from the author's website)*

In the discussion on the nature of photography, maybe this schema distinguishing images between "referential" and "non-referential" [8, 9] based on how they are created can be useful [Fig. 4]. The adjective "referential" comes from the Latin expression *res ferens*, meaning "that carries the thing," since referential images require the material presence of the object or physical phenomenon during the image creation process [8]. Referential images are obtainable only *in presence* of the referent (from which they are inseparable; "they carry it with them").



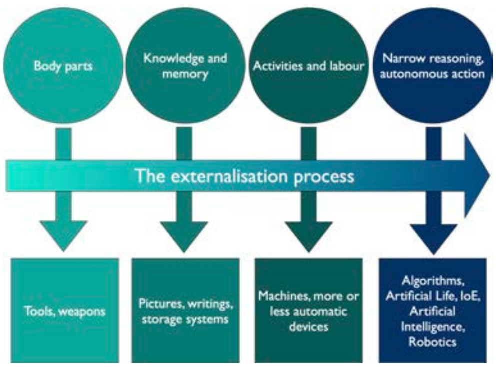
4. Referential and non-referential images (picture by the author)

The subject or object to be represented must physically exist and be present during the image creation. This category includes casts and analogs (imprints, moulds, etc.) and images obtained using light, such as photography, cinema, and video "from reality," and holography, which require the physical presence of a subject or object capable of reflecting or emitting light during the image generation process. In referential images, without a subject or object, there is no image.
Conversely, non-referential images do not have the constraint of the co-presence of the referent, so they can represent things that do not exist physically or are not present at the time of image creation. Thus, painting, sculpture, drawing, graphics, animations, computer images, Virtual Reality, video games, metaverses, images created with generative AI, etc. belong to this category. Consequently, photographic images and photorealistic computer-generated images belong to two different categories and should not be confused.
With generative AI, it is possible to create realistic images without any manual talent. However, unlike photography, these images are non-referential. Therefore, generative AI adds one more expressive possibility: without being able

to draw, it is now possible to create any type of image, both referential and non-referential.

4. The externalization process

Generative AI is inside a more general process of progressive externalization outside the body of human functions and activities [10], underway since the origins of humanity [Fig. 5].



5. The externalization process (picture by the author)

Through simple tools used as utensils and weapons, human culture began to externalize parts and activities of the body. Then, with images and writing, it recorded knowledge and memories outside the body. Then, with the use of more or less automatic machines and devices, physical work and many related activities were externalized. Finally, with Artificial Intelligence, Robotics, and Artificial Life, narrow reasoning abilities and autonomous actions were externalized. If this trend continues in the future, an increasing number of more and more complex human activities will be externalized.

A similar process is underway in the arts. Examples of "autonomous creativity" have existed since the dawn of human expression. Initially, it involved simple devices to assist the artists, but with scientific and technological advances, increasingly advanced and autonomous systems have emerged. Today, such systems and devices can support and replace the artist in many functions and operations, with various degrees of autonomy, with their own creativity, raising, among others, the issues of "art made by the machines," ethical concerns, problems about originality, authorship, and intellectual property. Externalizing creativity to an autonomous device can reduce and even replace human intervention in the processes that lead to the work of art. New tools and art forms, made possible by generative AI, are emerging. However, the human presence and responsibility in art can not be canceled. [11]

In the ongoing tense of externalizing the designing and making of art through the latest technological tools, what is changing in the creative, artistic process? What are the critical elements that emerge in making and enjoying art? How much of the human remains in these forms of expression? What novel art forms can emerge?

One might provocatively envision extending this process of externalization and human relativization. Would these creations still be aimed solely at humans, or might they transcend the human domain? Art serves as a "philosophy of contemporaneity," allowing us to interpret the present and peer into the future through transdisciplinary and critical lenses. Since art has been fundamentally intertwined with human essence, what meaning could an "art of the machines for the machines" or an "art for the non-human" possibly hold? Is it conceivable for some form of aesthetic sensibility to develop beyond the human domain? How might we recognize, consider, and assess these hybrid art forms? What new aesthetic paradigms would they engender? In this extreme externalization, would any trace of humanity remain?

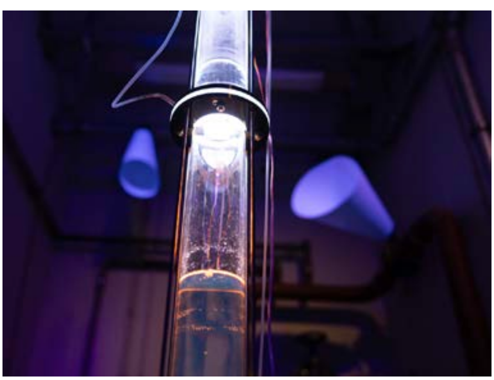
5. Towards an Oscillating Truth

In the end, I want to present an art installation where generative AI is not used to create images.
The human body is a complex and dynamic ecosystem inhabited by trillions of microbes, including bacteria, viruses, fungi, and other microorganisms. This microbial community, called the microbiome, plays a crucial role in our health and well-being, and bacteria in particular influence numerous physiological processes and contribute to protection against diseases. The relationship between the human body and bacteria is a fascinating example of biological symbiosis. Bacteria are not just passive inhabitants of our body but active actors that influence our lives in complex and interconnected ways. In 1991, biologist Lynn Margulis used the word “holobiont” [12] to define an organism characterized by the symbiotic close coexistence of biological agents that do not share the same DNA, that is, of different species that live inside or around the organism and are in symbiosis with it.
In the human oral and nasal cavities live several species of bacteria. When we speak, we modify the environment in which they live, both physically, by moving parts of the mouth and oral cavity and generating sound vibrations; therefore, the acoustic and physical conditions in which these bacteria live change. Our vocalizations influence their behavior; laboratory data show that certain frequencies favor their multiplication. For example, the tonal part of the voice, which is typical of singing and moaning, stimulates the growth of bacteria and strengthens their cytoskeleton, increases their reproductive activity, and the bacteria multiply, while the noises associated with speech stress them, limiting their growth.
Reciprocally, the bacteria influence the sounds of our speech by absorbing the frequencies that are useful to them. Therefore, in human vocalization, their contribution cannot be ignored; they actively intervene between the vocal-linguistic emissions and the sounds that are generated, and their role has been called “microperformativity.” [13] The voice, which is a personal distinctive and peculiar characteristic, is in fact the result of the interaction between our phonatory activity and the activity of entities that are not human.

6. Klaus Spies, Emanuel Gollob, Paul Gründorfer, Jens Hauser, Chronolalia: Towards an Oscillating Truth*, bio-media installation, 2023*

This preamble is in order to introduce *Chronolalia: Towards an Oscillating Truth* [14, 15] [Fig. 6], an installation by Klaus Spies, Emanuel Gollob, Paul Gründorfer, and Jens Hauser, funded by the Austrian Science Fund and created in collaboration between the Centre for Public Health of the Medical University and the Department of Digital Art of the University of Applied Arts, both based in Vienna.
In the installation, one must speak, sing, or whisper into a microphone. The vibrations generated by the voice and the

body are transmitted to the plate on the floor and then to a bioreactor, which contains the oral microbiome (in the installation it is that of the researchers). The bioreactor is a device that allows living organisms, in this case oral bacteria, to be kept alive outside their usual environment, a bit like an external mouth [Fig. 7]. From the voices, the microbiome in the bioreactor extracts the most functional frequencies for its growth, which are sent to a vocal articulator that makes them audible, and the acoustic speakers emit the filtered voice.
A generative AI based on a Deep Learning system, which is active in the installation, has been trained on the growth of the oral microbiome in response to phonetic vibrations. If people imitate the filtered voice produced by the articulator in the microphone in a way similar to an echo, they provide the Deep Learning process with enough data to continuously adapt the vocal frequencies to the “nutritional” needs of the microbiome. The system is constantly learning throughout the public exposure of the installation, and together with the voices of previous users, a kind of polyphonic and hybrid choir is created, consisting of vocal compositions created by new interspecies intelligence at different times, guided by the microbiome.



7. Klaus Spies, Emanuel Gollob, Paul Gründorfer, Jens Hauser, Chronolalia: Towards an Oscillating Truth*, bio-media installation, 2023*

This installation forms a kind of meta-organism, which transcends the human dimension and offers the experience of a language not spoken by human subjects but by microorganisms, who choose their phonemes and transform humans into instruments to create a more suitable environment for them. An experience that offers a multi-perspective and less anthropocentric vision of the living.
Towards an Oscillating Truth shows how truth is a fluid, oscillating, post-anthropocentric entity that transcends conventional temporal boundaries. By regulating the rhythm of their voices, users engage in a symbiotic cycle of information. This cycle alludes to the tension between the acceleration of our era and our inner rhythms.

References

[1] Sourav Banerjee, Ayushi Agarwal, Saloni Singla, “LLMs Will Always Hallucinate, and We Need to Live With This”, *arXiv*, 9 September 2024, https://dsh.re/f9913 (last access 30 October 2024).

[2] Vikas Hassija, *et al.* “Interpreting Black-Box Models: A Review on Explainable Artificial Intelligence”, *Cogn. Comput.*, 16, 2024, pp. 45–74, https://dsh.re/0c54c (last access 15 October 2024).

[3] U.S. Copyright Office Review Board, “Second Request for Reconsideration for Refusal to Register A Recent Entrance to Paradise”, *Copyright Review Board*, 14 February 2022, online, https://dsh.re/2ce72 (last access 2 November 2024).

[4] Jane Recker, “U.S. Copyright Office Rules A.I. Art Can’t Be Copyrighted”, *Smithsonian Magazine*, 24 March 2022, online, https://dsh.re/a0726 (last access 20 October 2024).

[5] U.S. Copyright Office, Library of Congress, “Copyright Registration Guidance: Works Containing Material Generated by Artificial Intelligence”, *Federal Register*, Vol. 88, N. 51, March 16, 2023, online, https://dsh.re/7a086 (last access 25 October 2024).

[6] Kevin Roose, “An A.I.-Generated Picture Won an Art Prize. Artists Aren’t Happy”, *The New York Times*, 2 September 2022, online, https://dsh.re/dd148 (last access 3 November 2024).

[7] Boris Eldagsen, “Sony World Photography Awards 2023”, online, https://dsh.re/0cc270 (last access 30 October 2024).

[8] Pier Luigi Capucci, *Realtà del virtuale. Rappresentazioni tecnologiche, comunicazione, arte*, Bologna, CLUEB, 1993, pp. 41–45. Second Edition: Ravenna, Noema, 2015, ebook, https://dsh.re/c4abe (last access 1 October 2024).

[9] Pier Luigi Capucci, “Evolutionary creativity. The inner life and meaning of art”, in Celestino Soddu, Enrica Colabella (eds.), *XX Generative Art conference*, Milan, Domus Argenia Publisher, 2017, pp. 320–330.

[10] Pier Luigi Capucci, “Outsourcing Life”, in Marcin J. Schroeder, Mark Burgin (eds.), *Proceedings of the 2021 Summit of the International Society for the Study of Information*, 81(1), 57, 2022, https://www.mdpi.com/2504-3900/81/1 (last access 26 October 2024).

[11] Pier Luigi Capucci,“Perspectives et visions de l’art”, *Ligeia*, n. 200–204, January–June 2023, pp. 179–191. Dossier “Art & Futur” edited by Hervé Fisher.

[12] Lynn Margulis, René Fester, *Symbiosis as a Source of Evolutionary Innovation*, Cambridge, MA, MIT Press, 1991.

[13] Jens Hauser, Lucie Strecker (eds.), “On Microperformativity”, *Performance Research*, vol. 25, n. 3, 2020.

[14] Klaus Spiess *et al.*, “Microbial Keywording. Towards material speech acts”, *Performance Research*, vol. 25, n. 3, 2020, pp. 56–62, online, https://dsh.re/3e802 (last access 20 October 2024).

[15] Klaus Spiess, Maximilian Hauptmann, Lucie Strecker, “Entangled Speech. Speaking through oral microbes”, *Performance Research*, vol. 25, n. 5, 2020, pp. 122–128, online, https://dsh.re/5b51b (last access 20 October 2024).

Archive Studio: Exploring Future Space for Former Matters

Margaret McManus, AIA, NCARB, NCIDQ, MFA
School of Building Arts, Savannah College of Art and Design, Savannah, GA, USA
e-mail: mmcmanus@scad.edu

Abstract

Can we begin the process of spatial design interventions at the conceptualization of an intimate-scale detail? How else can we upend explorations of inhabitable space as it merges with technology? Design education studios related to the built environment typically delve into a sequential process of exploration that includes gathering data on site location, analysis, architectural programming, plans, sections, and then subsequently (if time allows), carrying out detailed components that support the proposed architecture. This trusted path—or similar—has proved to result in successful outcomes of thoughtful architecture, however, ideas of space and how occupants inhabit space are rapidly evolving and invite new questions into the processes and outcomes of architectural manifestations. This paper looks at a brief snapshot at how space has formerly been explored and how—through the lens of a pedagogical design studio—inverting a more traditional design process could be the catalyst for new was of exploring space.

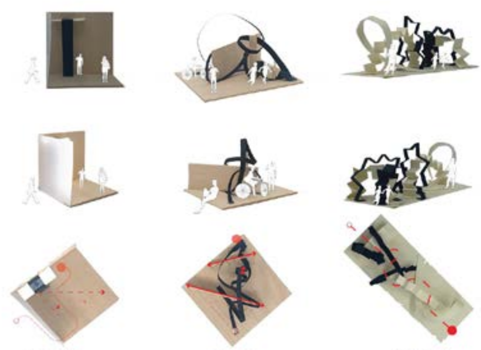


Figure 1: Visual explorations of space in physical model and 2D digital collage. by Zina Uzor

1. Introduction

Space is moving into an increasingly more virtual realm: illusory, imagined, and absent of much of itself--though not disappearing. Space is as present as it ever was, and in the education of the built environment, we must consider how to define space--in the literal word definition and in the literal physical “capture” of it. Understanding that space has been primarily something that associates with a physical occupation or inhabitable component, the question is in pondering the transition of the physical space of the past and the more virtual, absent, "false" space of the future.

Figure 2: Digital collage exploring archival space, by Hardik Jhanwar

2. The Archive

To study such a transition in a pedagogical setting, this studio—and subsequent paper looks at “archiving” a place to begin. As civilizations historically look to preserve their past and present for the sake of teaching--or reminding—those to come; creating a transition from past to future and vice versa. The archival of “things” as studied in the course, conveys that “things” can be physical or experiential, with the primary point of the exercise being to explore how to archive-- or "capture" with the intent to teach something worthy of posterity. And to then study and explore it through spatial and physical manifestations that are proposed to exist in the built environment.

The student work shown in figure 2 seeks to capture the intangible experiences of confusion and irony in an architectural museum—or archive. The students notes that his aarchival proposal will inform the next generations *why we did what we did and how our way of life made sense for our time.*

In The Archival Paradigm, the author writes: *“The key is to explain the physical aspects and intellectual structure of the collection that may not be apparent and to provide enough contextual information for the user to understand the historical circumstances and organizational processes of the object’s creation.”* [13]

This implies ideas of revealing that which may not be apparent and connects pedagogically with the idea that generating such a project would take on both the visual, and non-visual, tangible and non-tangible.

3. Illusions of Space

Non-tangible space conjures its collapse from 3D to 2D and hence, we can connect this to early ideas of creating the *illusions* of space. Additionally, the notion that this design pedagogy is part of an architecture program is one that inherently makes connections from 3D to 2D space. The illusion of space is more closely tied to the 2-dimensional arts: back to the understandings of 2 dimensional techniques for creating depth: diminution, overlapping, atmospheric perspective, and linear perspective. [12] All of these fundamental principles are a part of architecture education and architectural communication.

Emphasizing the prior connections to illusory space, Blackman writes in “Perspective: The Illusion of Space - Adventuring into Art.,” “*We use perspective to create the illusion of space and depth, pictured on a flat surface (like a sheet of paper, or a canvas, or a wall). When we look at an image that uses perspective, our eyes seem to see deep into the distance.*” [10] And connecting this perhaps more to the individualistic nature of the designer or occupier of virtual space (the wearer of the goggles): Sarah writes in “Brunelleschi “Rediscovers” Linear Perspective,” *“Linear perspective, as a novel artistic tool, spread not only in Italy but throughout Western Europe. John Berger, an art historian, notes that the convention of perspective fits within Renaissance Humanism because, ‘it structured all images of reality to address a single spectator who, unlike God, could only be in one place at* a time.’” [11] We can transition this notion easily to that of recent technologies tied to virtual space.

Figure 3: Digital collage exploring Hybrid Space, by Boyu Liu

4. Virtual Space

Virtual or false space takes the very ideas of the flat illusions of 2D space mentioned above and brings somewhat of a dimension to it, but can we call this the 3rd dimension? With virtual goggles on, we may be falsely occupying space, but we are a few--or several--degrees above experiencing 2-dimensional art and are a few--or several degrees below experiencing "“really”" physical, architectural space. And are degrees even an appropriate unit of measurement? Perhaps they work for this discussion alone.

Where the above notions of the illusions of space are coming full circle with ideas of virtual "flat" space that we can inhabit through visuals or goggles, the space that is dealt with in the built environment is undeniably 3-dimensions. That said, it is so much more. It is and must be responsive to the contextualities around it: the sun, its heat, its position, its light, the neighbouring buildings or trees, sounds, local resources, etc. It is rooted--for better or worse-- in the real, physical matter that we inhabit daily.

5. Student Explorations

Understanding more briefly the complexities of navigating the pedagogical exercise, the student examples shown throughout the paper expose how the use of technology assists in exploring former matters to explore spaces for the future.

It is evident in the student work that collage is an integral part of the exploration, avoiding the use of sole 3-dimensional architectural models, at the risk of space being explored too rigidly or "confined." The art of the collage is in itself, a type of transition as well: using found imagery or patterns to create and to communicate new ideas.

What may be an old method and technique for the generation of art, collage—be it analogue or digital has innate abilities and properties that allow the showcase of diminution, overlapping, atmospheric and linear perspective. Ease of scaling, transparency, and layering are embedded in digital software, and even perspective tools are easily used as a function of many digital programs.

Figure 4: Hybrid media studies of color, light, atmosphere, and archival space, by Emily Eatmon

While the students took the projects from collage into 3d models ultimately to fulfil architectural course requirements, it is at this generative art "space" that the paper is meant to discuss. This generative art product is --like the college method used to create it-- also a transitional step between thinking of an idea and manifesting it through an architectural design process.

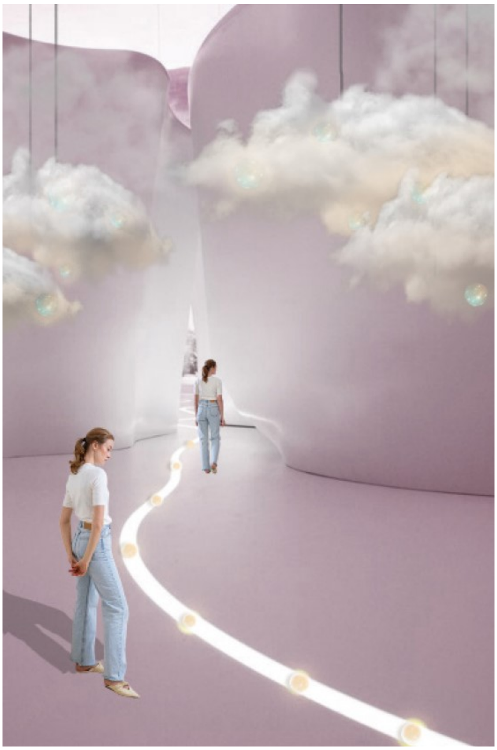


Figure 5: Visual explorations of space in 2D digital collage. by Khushi Patel

6. Conclusion

The pedagogical idea behind exploring future space with former matters, asks teachers and students alike how we can creatively transition between ideas from the past to the future. This particular question was relative to transitions of space in an architectural studio where ideas of 2D and 3D spaces are becoming more blurred and less defined. How can those ideas be pushed in an academic setting, and explored through generative art?

If we think of our own contextual surroundings -- whether real, illusory, or virtual, we can ask how the tools around us can assist in bridging what we know and what we do not know yet--what we imagine. The archive studio is a platform for studying such transitions at a variety of scales.

Figure 6: Visual explorations of space in 2D digital collage. by Zina Uzor

Ideas of future experiences may have little to do with the built physical architecture and more to do with screens and imaginary or unknown/unfamiliar space. Published under “Building the Metaverse,” author, Jon Radoff describes the Metaverse as a place that “expands to include immersive learning, shopping, education, travel, and other undreamt-of applications…” [5]. He also redefines or paraphrases the notion of “immersive” to mean “giving us the illusion of a *place* we are in.” [5] These so-called places and spaces on the horizon continue to evolve, inevitably leading to new ways in which to think about space, architecture, and scale.

At the intersection of reality-or real context and the consideration for future "space"—or the absence of architectural space; the studio focused on archiving existing realities. With unknown architectural space on the horizon---and the flexibility to propose program in design studios—by looking at relevant objects or "things" (tangible or intangible) that could/should be archived as a way of accounting for historic necessities and preserving memories of the past to serve as reminders for the unforeseeable future. This allowed for the exploration and the manifestation of space on a variety of scales while still grappling with ideas of immersive and illusory space.

This paper does not focus on the design *of* virtual space or design *in* virtual space, rather it speculates how we can address the education of exploring space—as it transitions from something once defined to something less defined.

Figure 7: Digital collage exploring Hybrid Space, by Tarana Pahlajani

Figure 8: Digital collage exploring new interiors, by Tarana Pahlajani

7. References

[1] Abdul Malak, H. (2023) *What is archiving: Why is it important? Information Management Simplified*. Available at: https://theecmconsultant.com/what-is-archiving/ (Accessed: September 16, 2024).

[2] ConstructConnect (2019) *Virtual reality is architecture's next big game-changer* , *The American Institute of Architects*. Available at: https://www.aia.org/articles/6232562-virtual-reality-is-architectures-next-big-(Accessed: March 11, 2023).

[3] Nicola, G. (2022) *Smart Technology that changed the Interior Design Industry Forever*, *Tallbox*. Available at: https://www.tallboxdesign.com/smart-technology-interior-design (Accessed: September 16, 2024).

[4] Nicola, G. (no date) *How Metaverse Might Change Interior & Architecture Design?*, *Tallbox*. (Accessed: September 16, 2024).

[5] Radoff, J. (2021) Building the Metaverse. Available at: https://medium.com/building-the-

metaverse (Accessed: September 16, 2024).

[6] Szekely, I. (2017) *Do archives have a future in the digital age? - yale university*, *The Journal of Contemporary Archival Studies*. Yale University. Available at: https://elischolar.library.yale.edu/cgi/viewcontent.cgi?article=1038&context=jcas (Accessed: September 16, 2024).

[7] Thamilselvan, K. (no date) *What is the metaverse and why should architects care? Insights. Jonite.* Available at: https://insights.jonite.com/what-is-the-metaverse-and-why-should-architects-care (Accessed: September 16, 2024).

[8] U.S. Environmental Protection Agency. Indoor air quality Available at: https://www.epa.gov/report-environment/indoor-air-quality (Accessed: September 16, 2024).

[9] White, S.T. (2022) *Designing the metaverse: The role of architecture in Virtual Environments*, *Metropolis*. Available at: https://metropolismag.com/programs/architecture-virtual-environments/ (Accessed: September 16, 2024).

[10] Blackman, Matthew. “Perspective: The Illusion of Space - Adventuring into Art.” *Adventuring into Art*, 9 Jan. 2020, imbaliartbooks.org.za/perspective-the-illusion-of-space/. (Accessed 7 Nov. 2024).

[11] Sarah. “Brunelleschi “Rediscovers” Linear Perspective - Drawing Academy.” *Drawing Academy*, 26 Sept. 2018, drawingacademy.com/brunelleschi-rediscovers-linear-perspective. (Accessed 7 Nov. 2024).

[12] “The Visual Experience: Reading 2014.” *Arthistoryresources.net*, 2014, arthistoryresources.net/visual-experience-2014/illusion-space.html. (Accessed 7 Nov. 2024).

[13] “The Archival Paradigm: The Genesis and Rationales of Archival Principles and Practices • CLIR.” *CLIR*, www.clir.org/pubs/reports/pub89/archival/ (Accessed 7 Nov. 2024).

Beauty and the Prompt

Joanna E. Kabala, MFA, MSc
Eindhoven, The Netherlands
vangdansk.com/kabala.html
joanna.kabala@vangdansk.com

Abstract

This paper presents observations made during image generation using text-to-image and image-to-image prompts in several commercially available Generative AI Art generators. Images from various sessions have been compiled into a short animation that is part of an ongoing art practice, conceptually touching on the issue of consciousness. Without delving into philosophy, consciousness is viewed as the essential feature differentiating humans from machines. Thanks to consciousness, the domain of art belongs to humans. AI, even if it behaves *intelligently*, is perceived merely as a tool that artists can efficiently use to create content. The observations have been made by a visual artist who, years ago, was involved in an experimental generative art design-research project and created some simple graphical shapes for a system run by genetic algorithms.
The main impulse to investigate Generative AI Art generators and report on them, arose from the avalanche of media headlines proclaiming the end of art, due to the advent of the AI era. Many content creators feel uncertain if the threat is real. Will Generative AI eliminate artists?

1. Introduction

The flow of media messages has been less intense in the last few months. Still, many people are concerned with the frightening vision of the complete replacement of humans by artificial intelligence in some professions. Is artificial intelligence also a threat to artists and designers? Is imagery generated by AI intelligent? Is it art? We are now living in a *post-ChatGPT* era, a time marked by the accelerated spread of awareness about generative tools, which until recently were known mainly to specialists. Millions of people now use AI editors (*) daily. Some claim they no longer need artists or writers.

Many artists feel anxiety when they see images that once required teams of designers and months of rendering time, now generated within seconds, three or four variants at a time.
It seems it took the entire lifetime of Salvador Dalí [1], or some longer effort even for an artist as fast as Picasso [2], to create what renders now in front of our eyes in no time.
In addition, it is loudly advertised as free, although it is not because creating a work of art requires many more than just a few random sample renders. But is there anything to fear, or should artists feel blessed by the opportunities brought by the broad access to generative tools?

2. Text-to-Image

The initial experience with Generative AI Art generators has been intentionally simplified. Prompted traditional themes like landscapes and flower fields have been captivating and even slightly breathtaking, evoking afterimages of beautiful dreams. While beautiful, they remained impersonal, offering too few cues to suggest any human-like intelligence behind their creation.
It is also possible that the textual prompt, which needs to be given to the artificial intelligence to start rendering, posed a challenge to a visual artist confronted for the first time with the task of creating images solely with words. Imagery creation via natural language is unnatural to visual artists who create forms and compositions organically, *while* observing them on a specified plane and only name the images with full sentences *after* they come into existence — *if* they name them at all, as many visual artworks remain untitled.
Writing a sentence to express a concept should not be a difficult task, as concepts usually precede visualizations. And if the urge to create is an impulse and not a concrete thought, the natural language feature of AI editors may expand artists' ability to articulate ideas, requiring them to put visions into words, rather than loosely thinking about them. Thus, it is already a useful tool that enhances art.
Other features may seem overwhelming, especially the unreal beauty and astonishing speed at which AI generates images. It’s easy to think that humans can’t match this level of performance.
The visual quality is exceptionally high and often immediately recognizable as AI renderings. These images primarily differ in their surreal use of color. It seems that AI employs a range that humans can’t even perceive. While humans may not notice the difference between shades that lie next to each other on a scale of millions of colors, artificial intelligence can differentiate and efficiently utilize them.

Nevertheless, this is not the end of human art. This undoubtfully helpful tool does not essentially change the creative process, it speeds it up and enhances it. But what follows is a day as usual, as every author knows well. It is sometimes satisfactory but could be exhausting and may not yield the desired results. Long rendering sessions are needed to obtain a meaningful image, especially when aiming for a sequence of images.
Sometimes corrections by hand are necessary as these intelligent generators still lack some accuracy. The most striking are anatomical errors that AI makes. Especially, if prompts did not include any request to generate defective figures. Of course, such random software

behavior is not surprising for generative designers.

What may surprise in the whole context is that people may believe the view spread by the media, which has already broadly announced AI as a sentient being that even possesses free will. We could hear such declarations about on-screen software. Also, we could see demonstrations of robots, such as Sophia [3], whose successful attempt to obtain citizenship made a huge impression in 2017. However, it is disputable if this was a result of *free will*.

And if AI is so intelligent and free in its creations, why does it repeatedly make certain errors? Characters rendered by this artificial competitor to the human author are often anatomically incorrect, it seems as if they were created by a student, not a very talented one. Some body parts are let go as if they couldn't be drawn accurately. It is not only human anatomy rendered by AI that is often neglected. Other forms, objects, or architectures are frequently incomplete, illogical, or, in other words, not well-designed. Should they be accepted as, for example, futuristic or cubistic even if such a style was not included in the prompt?

Another attempt was made to communicate with the Generative AI Art generator through natural language. This time, the prompted image was supposed to be a portrait of a woman with red hair sitting at a wooden table (*ill. 1*), in Andy Warhol's style [4]. When writing the prompt, the goal was to create a graphical, almost stencil character with a hint of Marilyn Monroe. After some trials, such a character appeared. Or at least a character close to the original imagination, though it is certainly not Warhol's colorful graphical Marilyn, and she does not sit at the table. Rewriting the prompt did not help. All important details, such as a wooden table, were completely ignored. Additionally, the composition is tight, resembling a talking head sitting in a chair on TV, with more chairs in the frame than needed. Nevertheless, it was accepted. If AI interprets the human vision written in a sentence this way, a closer look can be taken, just out of curiosity, even if it's not exactly what was envisioned.

Illustration 1a
Mary is OK - AI imagery composed into a short animation. Without AI, this image would not exist; without the artist's mind, it would not be Mary.

Illustration 1b
Mary is OK - set of variants. Some images were edited by hand to obtain a clear stencil style.

There is more to this artwork than the struggle for adequate visualization [5]. An important aspect of this work was a specific scent, intended to evoke a thought that artificial intelligence has no sense of smell. AI cannot smell or taste and will never experience such sensations because it is not a biological process. Moreover, biology alone is not what makes us human. Without consciousness, we would not create and lead discussions about creativity. As for consciousness, people cannot implement into machines such phenomena that they do not understand themselves. To make AI human-like, their creators (software developers) would first need to understand who humans are. While such an understanding is emerging nowadays [6], the way to make it common is long. When it becomes common, it will be obvious that thinking about and making efforts to imitate humans makes no sense.

Another issue that goes beyond the topic of this work is how artificial intelligence may be used against humans. This paper does not neglect the real threats; it just stays in the domains of art and design.

3. Image-to-Image

Months of rapid development in Generative AI Art generators have passed, yet, the shape of characters' hands and feet still needs corrections — or perhaps it should be accepted as a *signum temporis*.

The common opinion is that AI creates. But is this truly the case? AI searches the data space, makes associations (some may be considered intelligent), and then renders images, fully blown, hyper-realistic, or any other known style if required. So, in a way, it does create, although the result is a kind of *artificial* art. The results of AI creation processes are chaotic, by the very nature of generative algorithms. The first images in the series are random. Some are very beautiful, but their randomness often reveals a lack of understanding on the part of the tool that created them. During longer rendering sessions, an increase in chaos can be observed. With more specific prompts, the generators quickly get lost and fill the images with unnecessary details that are difficult to remove. The generator needs a reset to resign from a specific style or abandon a particular object. Upon reset, all the previous work is gone. Repeating the prompt will not recall the former stage, it will be a whole new journey. Chaos arises not only due to the nature of generative algorithms. Deliberating on human-like creation, it is important to note that artificial intelligence creates chaos because it "looks" but does not see. It multiplies the variants mathematically and recognizes the images only as far as it was taught to read data but does not see their meaning. Extracting semantic or figurative meaning from images or giving such meaning to images is not a mathematical process. Will this ever change? Although it cannot be ruled out, it remains a fantasy that Consciousness would one day inspire AI. For now, semantic and metaphoric understanding of natural language remains in the human domain.

Meaningful images can be generated more efficiently with image-to-image generators than with text-to-image generators, although in most cases, these two methods work together. Often, the generative process must start with a written prompt, after which the option to use a reference image may become available. Some generators offer the option to use two reference images, one for composition and another for style. This enhances creative possibilities and provides some level of control over the renderings.

Illustration 2a
Love Victory = √1%
On the left: the initial reference image
On the right: AI rendering, edited

The main theme of this work is a woman showing the sign of victory, the sign of freedom. She is the symbol of peace, and her name is Love Victory. The Jackdaw-Phoenix flying next to her represents intelligence and social skills. Let the woman and the city jackdaw become a symbol of the interaction of an intelligent society, capable of calming itself internally and radiating peace externally. (**)

Illustration 2a shows the initial hand drawing, made with a digital pen, and one of more than 500 renderings that followed. Throughout the process, some AI renderings were also used as reference images. The selected set of final images has been composed in a short animation according to the author's storyline.

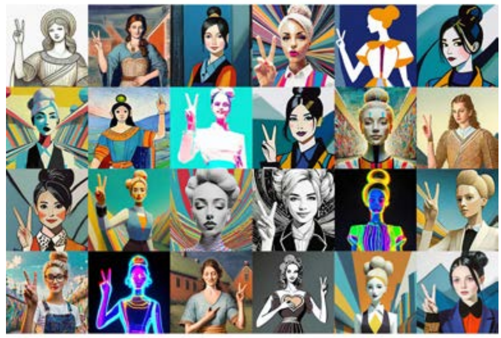


Illustration 2b. Love Victory = √1%
24 snapshots from the animation.

4. Conclusions

Generative AI Art generators mark a new epoch in human creativity. Nevertheless, those currently available, especially text-to-image generators, are merely filters comparable to tools like the Photoshop Filter tool. Creating artworks according to concepts conceived by the human mind often requires broader toolsets. Current generators do not allow for control over scenes, even in basic aspects such as the number and placement of forms and characters in the composition. It is all chaos, all random, as befits generative tools. Yet, to maintain the meaning of the author’s storyline, a certain level of control is useful. This can to a certain degree, be achieved with image-to-image generators - but this is not the end of human art. Contrarily, this is the dawn of new types of tools. And as for tools, these are intelligent indeed. They allow individual artists to produce artworks that previously required whole teams and

studios.
Broad access to a quickly evolving variety of Generative AI Art tools is good news in the domain of visual arts. Although artists should not expect that artificial intelligence understands poetry, they may efficiently use AI in creating poetical and symbolic imagery.

As for exceeding humans, Generative AI Art generators still have a long way to go to match what they advertise. Despite unreal beauty and astonishing rendering speed, there exists a problem with the datasets that AI uses. Even the largest ones are finite and limited. Often, they also infringe on copyrights. Additionally, they seem to be governed by semantic coincidence. This observation comes from the point of view of an artist, who is not satisfied to see *anything* resembling a piece of art. The artist wants to see something pleasing to the human mind and satisfactory to consciousness. We say that AI creates, but it merely generates and renders. It pulls data randomly and the answers are not very intelligent when it lacks information. In this respect, it resembles a human being a bit indeed. However, such productions can hardly be called art because they contain no conscious or subconscious intention. The effort needed to control or at least direct the generative process according to the artist's vision is comparable to the effort required to visualize a concept in any other visual content creation software. The difference lies in the quantity and quality of the images.
Generative AI Art generators are brilliant tools for image creation. However, do not fear that generative artificial intelligence will replace humans in art, because neither humans nor their arts are mathematical processes, as AI is. Additionally, biology is not the only barrier between humans and artificial intelligence. Human art is not a product of the brain alone. The brain controls the arm, which paints, but notice that some disabled people have developed the capability to paint with their feet or mouth. This was possible because the brain controls the body. However, the impulse to create comes from the mind, not the brain. Artists create thanks to biological impulses and spiritual processes in the brain-mind area. The human mind is the domain of consciousness that is unavailable and not transferable to machines. Artists create under the influence of (un)consciousness, in which phenomena still largely unrecognized, such as dreams, play a role. Art is uncountable and unpredictable, while AI renderings are relatively predictable, as they always depend on datasets.
To the incomprehensible aspects of human existence, which are not transferable to machines and are essential in the creation process, we can also add semantics. This is another threshold that cannot be crossed by AI, as it is the human minds of the authors that give meaning to art. From turning on devices to providing prompts, to selecting or rejecting AI products, AI would create nothing without humans.
Art exists only thanks to the artists' will. It exists only when it is perceived by consciousness. Artificial Intelligence does not perceive; it reads data. It is said that AI creates, but it only renders data supplied to it. It merely imitates. Artists may appreciate AI renderings and incorporate them into the creative process. Yet, conscious human effort is needed to obtain specific images in a

particular style. AI art editors are just tools, and they remain so. They are not and will not be independent entities that could replace humans in the field of art. Not today, not tomorrow, not even in our lifetime, and possibly never, will we be able to create *artificial consciousness.*

Notes

(*) ChatGPT was launched at the end of 2022. Since then, many AI chatbots have been created. There are also many Generative AI Art generators, such as Leonardo, DALL-E, Midjourney, Stable Diffusion, Firefly, Runway, LeiaPix, and many others, with new ones appearing continuously.

(**) The Generative AI animation titled *Love Victory = √1%* conceptually addresses the notion of consciousness, which has started to be acknowledged as a “force” even by mainstream physics. In 2012, Peter Higgs discovered the God particle, and recently, Sir Roger Penrose, referring to the issue of consciousness, admitted that for the universe to exist, a process that cannot be described mathematically is essential. Speaking colloquially, miracles happen in the bosonic quantum field. Additionally, their elusive nature sometimes brings tangible results.

For example, scientists have researched the so-called Extended Maharishi Effect for 43 years, since 1974. They have collected enough data to conclude that it only takes a small group of people to initiate a change in social consciousness. They counted it carefully: the square root of one percent is enough. When a group of people the size of the square root of 1% of a given community experiences pure awareness through extended transcendental meditation, that is, in simple words, calms down internally, there appears a greater cohesion in the collective consciousness, contributing to conflict resolution, such that the number of recorded incidents drops. The research was carried out in various places on Earth. A decrease in incidents and even a temporary easing of armed conflicts occurred near meditating groups, in areas equal to districts of larger cities [7].

The square root of one percent isn't much. For example, in a community of 2,500 people, just 5 people are enough to calm down this community, at least temporarily. A sustained action would require a sustained group effort, which is not feasible. But this thought and the scientifically confirmed fact seemed worthy of creating a symbolic drawing and editing the results of generative rendering sessions into an animation. *Love Victory = √1%* is available on YouTube,https://youtu.be/MX1XyvyQxW

The contextual article is available on LinkedIn, https://bit.ly/3YlHbe2.

References

[1] Robert Descharnes, Gilles Néret. *Dalí. The Paintings*, Benedikt Taschen Verlag 1997.

[2] Henri-Georges Clouzot. *The Mystery of Picasso,* 1956. A fragment of the movie with time-lapsed, "generative" images is available on YouTube https://tinyurl.com/mry2zy3w

[3] *Sophia*, a human-like robot https://www.hansonrobotics.com/sophia/

[4] Andy Warhol, *Marilyn Monroe,* 1967 https://www.moma.org/collection/works/portfolios/61240

[5] Joanna E. Kabala. *An image spoke to me: Mary is OK.* October 2023, available on LinkedIn https://bit.ly/4eVS1hJ

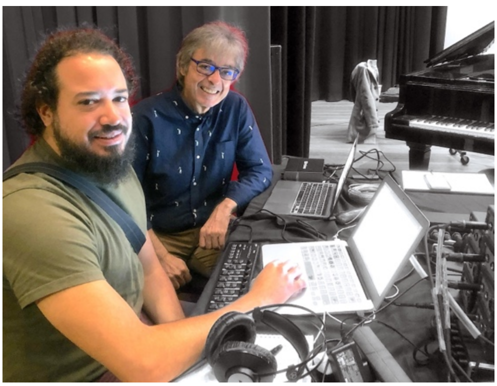
[6] Faggin, F. *Irreducible*, Essentia

Books, an imprint of John Hunt Publishing, 2024.
[7] David W. Orme-Johnson and Lee Fergusson. *Global Impact of the Maharishi Effect from 1974 to 2017: Theory and Research*, Journal of Maharishi Vedic Research Institute, 2018

Re(M)ind: Composition and Generative Sound Design of a Multimodal Opera

Prof. Jônatas Manzolli, BMath, BMus, MMath, PhD.
Interdisciplinary Nucleus for Sound Studies (NICS), Unicamp, Brazil
Centre for Informatics and Systems of the University of Coimbra (CISUC), Portugal
https://orcid.org/0000-0003-4257-7118
e-mail: jotamanzo@gmail.com

Dr. Micael Antunes. BMus, MMus, PhD.
Aix-Marseille Univ., CNRS, PRISM, LIS, InCIAM, France.



Abstract

This paper explores the challenges encountered in the musical composition and generative sound design of *Re(M)ind* (2023), a multimodal opera creation. It focuses on achieving a delicate balance between artistic creativity and the technological solutions required to integrate musical elements with dynamic performer-machine interactions. The discussion will cover the aspects of composition, sound design and real-time performance that shape an immersive audience experience. Additionally, the paper addresses the technical complexities of incorporating generative music with operatic elements, interactive dance, and bio signals in real time, all while maintaining a cohesive artistic narrative.

1. Introduction

The intersection of technology and artistic expression is a dynamic and evolving field within contemporary performance arts, with multimodal opera emerging as an exciting frontier. These intricate productions blend traditional operatic elements with cutting-edge technological innovations—including sound design, motion capture, and bio-signal processing—to create immersive, interactive experiences for audiences. This paper examines the creative and technical challenges encountered during the development of *Re(M)ind* (2023), a multimodal opera that merges generative sound design with real-time performer-machine interactions.

Re(M)ind draws inspiration from a lineage of innovative works, beginning with the inception of RoBoser by Verschure and Manzolli in 1998 [1]. This legacy includes pioneering projects such as *Ada:*

Intelligent Space (2002) [2], *re(PER)curso* (2007), and the *Multimodal Brain Orchestra* (2009) [3]. Together, these foundational works inform the original score of *Re(M)ind*, composed by Manzolli. The production also incorporates advanced technologies—including bio-signals, motion tracking, and EEG—developed by the SPECS Lab research team, further enhancing the immersive and interactive dimensions of the performance.

The world premiere of *Re(M)ind: A Multimodal Journey into Consciousness* took place on October 19 and October 21, 2023, at Theatre "C" at Radboud University, Nijmegen, in collaboration with the Donders Institute and Donders Citylab. This artwork invites audiences to explore the complexities of human consciousness through an innovative fusion of art and science. The concise libretto, structured in four acts and created by Anna Mura, Manzolli, and Verschure, delves into the enigmatic processes of the human mind, inviting reflection on the nuances of cognition and self-awareness. By blending music, dance, and interactive video, *Re(M)ind* creates a rich, multisensory experience.

2. Concepts, Components and Technologies

This section provides an overview of aspects forming the foundation of *Re(M)ind'*s multimodal opera.

2.1 Multimodal Opera

In [4], the authors argued that from its inception in late sixteenth-century Italy, opera has been a multimodal art form, with each mode contributing to its complex interplay of musical, verbal, visual, and dramatic meanings. Building on recent studies of embodied cognition in film and multimedia, [5] examines audio-visual congruence as key to understanding the viewer's experience of opera in its remediated form as film. The composition of such multimodal artwork requires a delicate balance between artistic vision and technical execution.

2.2 Design of Sound Textures

The compositional strategy at the heart of this work is based on the concept of musical texture, which serves as a dynamic framework that mediates the creative process. Musical texture offers both qualitative and quantitative aspects to composition [6]. From a qualitative perspective, composers often use texture as a metaphor, weaving together layers of sound to evoke affective responses through different sound configurations [7, 8]. This metaphorical use of texture allows a deeper exploration of how auditory elements interact to shape sensations.

On the quantitative side, musical texture provides a means of organizing material by describing the number and relationships of its constituent elements. This quantitative description is essential for understanding how these elements interact over time and across frequency ranges and how this impacts perception. Here, concepts from *Gestalt* theory [9], *psychoacoustics* [10], and *auditory scene analysis* [11] are often applied to investigate how elements within a texture integrate or segregate perceptually.

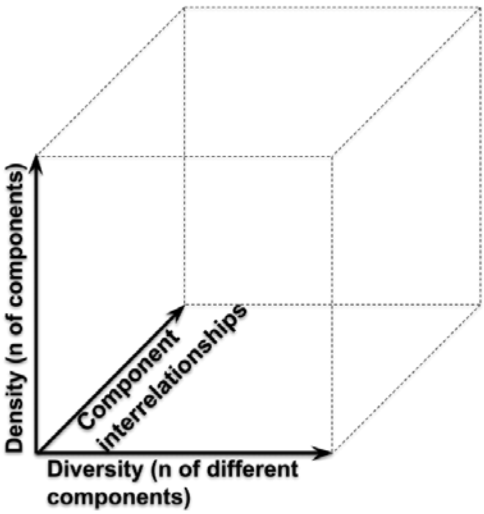


Figure 1 – Conceptual space of musical texture

Although this paper does not aim to delve into the theoretical aspects of musical texture, we outline a conceptual framework for its use in the interactive design of *Re(M)ind* in Figure 1. In this context, we define musical texture in terms of three elements: (1) density, which represents the number of elements within a given texture—including melodies, soundscapes, spoken voices, and other sonic materials; (2) diversity, indicating the variety of elements within the texture; and (3) interrelationships, describing the connections and interactions among the components. With these elements, we generate a wide array of musical textures by manipulating available sound sets within a performance. In *Re(M)ind*, we explore these components, their interactions, and their real-time control, as detailed in the following sections.

2.3 Motion Caption

The motion caption system in *Re(M)ind* used two main devices: a set of body sensors distributed over the dancer's body and accelerometers embedded in smartphones, which were worn by both the dancer and the percussionist during the performance.

The body sensors collected *x* and *y* position data of the dancer's body in space, from which movement descriptors were generated to mediate the interaction between the dancer's movements and the resulting sound responses.

The primary objective in selecting these descriptors was to foster a responsive relationship between the performers' movements and the sound textures perceived by the audience. To achieve this, four descriptors closely associated with movement perception were chosen: velocity, acceleration, and jerk (represented by the derivative of acceleration), as well as the magnitude of movement. Additionally, a descriptor linked to position, specifically the body angles, was incorporated. This data enabled a dual mapping approach, where movement could both trigger sound samples and control spatialization, as further discussed in Section 3.2.

2.4 Bio-signals

The physiological measurements used during the performance included an Electrodermal Activity (EDA) sensor attached to the mezzo-soprano and two electroencephalogram (EEG) headsets to measure electrical activity in the brain. We calculated the overall activity of the EEG signals, representing a cumulative measure of brain activity, to capture the EEG dynamics of the two brain-performers as they were exposed to the *Re(M)ind* performance.

Thus, Electrodermal Activity (EDA) can serve as an objective indicator for assessing emotional arousal [17]. In *Re(M)ind,* we used two measures of arousal to control the Affective Synthesizer: one from the EDA of the active performer, reflecting emotional arousal, and the other from the EEG of two passive performers, capturing brainwave patterns associated with emotional engagement. These combined measures enabled dynamic control of the system, allowing it to respond to both immediate and subtle emotional states.

Both EDA and EEG data were transmitted to the system, where their values were normalized within the interval [0…1] to be used as control input for the Affective Synthesizer, as will be discussed in Section 3.3.

3. Multimodal Integration

The multimodal scenario integrates innovative tools and processes to create an immersive and interactive performance. We explore musical composition strategies that integrate generative sound design with operatic structures.

3.1 Musical Composition & Sound Design

The compositional process began during the "Donders Cognition, Brain and Technology" (DCBT2023) Summer School on Cognition, Brain, and Technology at Radboud University, Nijmegen is structured around four short acts, each reflecting a distinct state of consciousness: (1) Awakening the Instincts, (2) Self, (3) Memory, and (4) Confluence of Self and Memory.

The composition strategies of *Re(M)ind* were intertwined with sound design. This approach stems from the central narrative idea that the human mind is like a river-full of infinite possibilities and never-ending. At each stage of composition, a sound design strategy was employed, creating a soundscape that opens each act. Following this initial soundscape, each act continues with an aria, in which the lyrics sung by the mezzo-soprano explore a different aspect of the mind. To achieve this sound poetics, each of the opera's four short acts gradually immersed the audience in a river of sounds and images, activated by sensors and enhanced by spatialization elements. This composition aims to create an immersive experience that mirrors Heraclitus' view of the world as always "becoming" rather than "being" - reflection of the continuous process of expanding consciousness and embracing the ever-changing nature of existence.

The real-time elements of the musical composition were structured around the concept of musical texture. This approach relies on a strategy of shared control to dynamically shape sound textures in real time [12]. The elements contributing to the composition over time are grouped into three main categories, as illustrated in Figure 2: (1) Fixed-media electronics, which include "mind river echoes" and pre-composed electroacoustic soundscapes. (2)

Performance elements encompassing the singer, percussionist, and pianist. (3) Real-time electronics consisting of granular synthesis, affective synthesizer, and ambisonics spatial control, each modulated by shared control mechanisms.

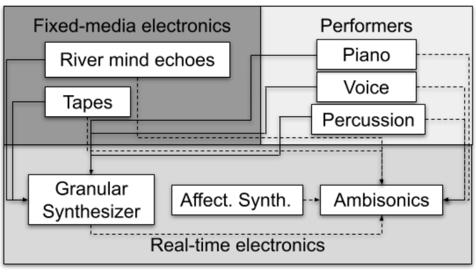


Figure 2 – Sound flux of shared textural composition

The concept of shared control [12] is designed to achieve two primary objectives: (1) to foster a sense of integration between movement and sound by establishing a clear causal relationship between the performers' actions—particularly the dancer's movements—and the evolving soundscape, thereby enhancing the audience's perception of their interplay; and (2) to enable high-level sound scene management by electroacoustic performers, ensuring a cohesive balance between the soundscape and spatialization.

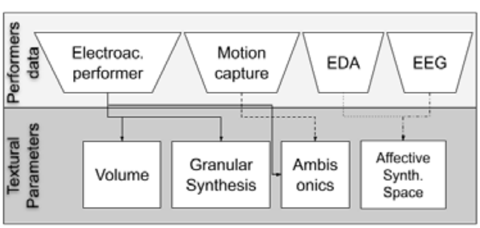


Figure 3 - Shared control of musical texture parameters

Control inputs are distributed to different performers and linked to their specific interpretive behaviours, performance strategies and tools. The two electroacoustic performers have primary control over the sonic scene, managing parameters such as spatialisation and synthesis. Additional controls are embedded throughout the performers, as described in sections 2.3 and 2.4. The mappings of these controls are visualised in Figure 3, which illustrates the interplay between each performer's controls and the sound texture parameters.

3.2 Dance Sonification

The primary objective of the dance sonification in *Re(M)ind* was to establish a responsive connection between body movement and the composition of musical textures, as mentioned above. To achieve this, the movement descriptors of *velocity*, *acceleration*, and *jerk* were employed as triggers for a sound instrument implemented within the *MoveGuitar* application [12]. These triggers also activated an FM synthesizer controlled by the electroacoustic performers.

In addition, the magnitude and angle descriptors were used to control the spatial distribution of both the dancer-triggered instruments and the pre-recorded sounds. This spatialisation was achieved by mapping the descriptors to the polar coordinates of the ambisonics system [13]. Specifically, the magnitude was mapped as a distance parameter from a central switch point within the *ambisonics* space, while the angle was the angular component within the polar coordinates of the same space. This configuration enabled an immersive response, allowing sound textures to dynamically correlate with the performer's body movement, thereby enhancing the expressive connection between movement and auditory experience.

3.3 Affective Synthesizer

The Affective Synthesizer was designed based on the Circumplex model [14], which served as inspiration for the creation of an emotionally responsive control space. This approach facilitates the real-time display of a sound bank, which interacts with EEG headset data and electrodermal activity (EDA) sensor data, allowing for a responsive from the performers acting as control of the sound designed responses of the synthesizer. The synthesizer is built upon a sound bank derived from ADA, the intelligent space created in 2002 [2]. This sound bank consists of a dataset of labelled sound files representing different emotional states associated with ADA: joy, sadness and anger [2].

To structure the synthesizer, a two-dimensional space was established to display the sound files, illustrated by Figure 4. The left portion of this space is divided into ten segments: five for sad sounds and five for angry sounds. Conversely, the right side of the space is allocated into five segments for sound files representing joy. Each segment is associated to a Gaussian curve that controls the loudness of the respective sound file, enabling smooth transitions between all sounds. This design not only enhances the emotional based response of the synthesiser, but also allows for fluid interactions informed by the physiological responses of the performers.

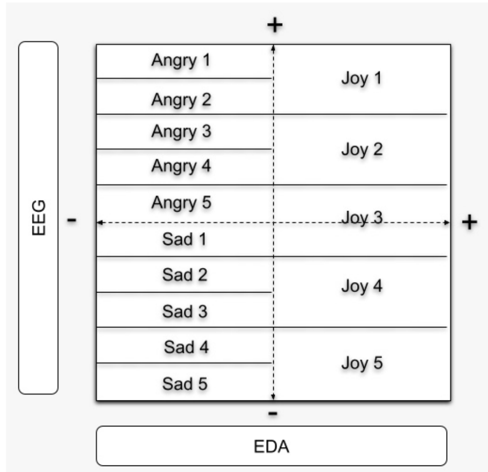


Figure 4 - Pseudo-circumplex space and the disposition of the sound bank

3.4 Sound Diffusion

The spatialization system in *Re(M)ind* utilized an ambisonics framework for constructing a dynamic sound environment, implemented via the HOA library [13]. This approach allowed sound sources to be positioned within a two-dimensional space, creating an immersive and responsive sound experience. Two primary spatialisation strategies were employed: one for localised sources and another for a granular synthesis immersive environment.

In the first approach, six different localised sound sources were configured as follows: (1) Percussion, (2) MoveGuitar, (3) Affective Synthesizer, (4) Tapes, (5) Mind River Echoes 1, (6) Mind River Echoes 2, as shown in the HOA interface in Figure 5.

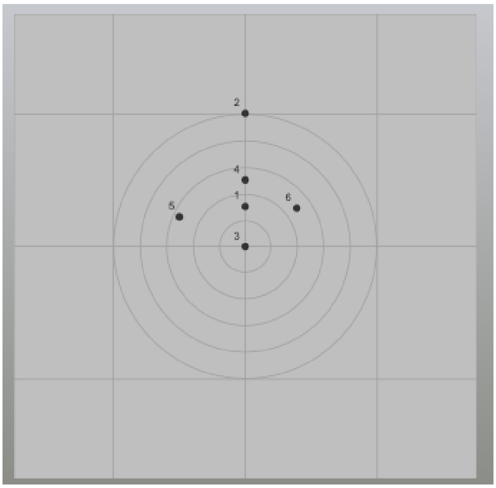


Figure 5 - HOA interface sound source management

Sources 2, 5, and 6 were designed to respond to the performers' movements, following the strategies previously discussed. The other sources were controlled directly by the electroacoustic performer, using two techniques: (1) predefined spatial trajectories and (2) sinusoidal oscillators mapped to polar coordinates, which facilitated the creation of various circular and spiral trajectories. The parameters for these movements were adjusted in real time via MIDI controller knobs, allowing nuanced manipulation of spatial paths in real time.

The second spatialization strategy, dedicated to granular synthesis, treated the synthesis environment as an immersive sound field within the ambisonics space. Implemented with HOA library presets, this approach allowed for adjustments in the spread of granular sounds into the sonic field, creating a sense of volume and space that was dynamically modulated. This design followed the principles outlined by Barry Truax [15], who emphasized the spatial distribution of sound grains in immersive contexts to enhance depth and presence. Through these strategies, the spatialization system offered a fluid and interactive soundscape that aligned closely with the performers' actions and the composition's immersive intent.

4. Discussion & Conclusion

Musical textural composition plays a crucial role in integrating diverse elements within a multimodal opera, acting as a connective tissue that binds together the sonic, visual, and performative aspects.

Figure 6 – *Re(M)ind* performance portraits in three photos.

In *Re(M)ind*, the use of musical textures serves as an expansive and flexible canvas, allowing for the seamless incorporation of other artistic and technological artifacts. Unlike traditional melodic or harmonic structures, textural music emphasizes layers of sound that interact dynamically with external inputs, such as bio-signals or movement data. This approach creates a sonic environment where various components—interactive dance, generative music, and real-time processing—can coexist and influence each other without disrupting the overall coherence of the performance.

The openness of sound textures fosters a space for continuous interaction, adaptation, and transformation, making them an ideal medium for multimodal operatic works that rely on the fluid integration of multiple creative and technological elements. This compositional method not only enriches the auditory experience but also supports the holistic integration of the various modalities at play.

In advancing textural composition, our work anticipates the integration of deep learning tools—specifically, variational autoencoders (VAEs)—in future performances to explore new dimensions of control over the textural space [16]. One of the key advantages of using VAEs in this context is their ability to capture high-level features in texture data within a lower dimensional probabilistic latent space. This allows smooth transitions between textures by interpolating within this space, offering an approach that could inspire spatial representations to generate and evolve textures in real-time performance. This provides a way to link physical gestures with textural transformation, creating an embodied interaction between the performer and the sound field.

The development of *Re(M)ind* showcases the potential of integrating real-time sound synthesis, motion capture, and bio-signal processing to enhance both performer expression and audience immersion. These innovations open new avenues for interactive performance, where technology and art seamlessly converge to create dynamic, evolving narratives.

Looking ahead, the continued incorporation of generative techniques into opera presents a promising pathway for artistic evolution. As technological tools become increasingly sophisticated, the ability to integrate real-time data into performance will further expand interactive possibilities. The use of textural music composition, as seen in *Re(M)ind*, is particularly significant in this regard, as its open and adaptive nature allows for the fluid integration of diverse elements such as bio-signals, movement, and spatial audio. This approach not only enriches the musical landscape but also creates a flexible framework that can accommodate future advancements, including AI-driven systems and immersive sound environments.

Acknowledgment

Re(M)ind is a multimodal opera created by Manzolli with contributions from a team of performers and researchers. The narrative, scenography, and choreography were led by Anna Mura, Jônatas Manzolli, and Paul Verschure. Micael Antunes and Manzolli developed the textural musical elements described in this article, with performances by Helvio Mendes on percussion, Stephanie Firnkes as mezzo-soprano, Bert Kappen on piano, and Aida Guirro Salinas as movement artist. The SPECS team managed bio-signal capture and processing, while Hector Lopez Carral provided interactive graphics. Real-time EEG and motion capture (using Xsens technology) were conducted by Tarandeep Singh and Sajad Kahali.

Micael Antunes acknowledge InCIAM institute, through which he has received funding from the Excellence Initiative of Aix-Marseille University – A*Midex, a French “Investissements d’Avenir programme” AMX-19-IET-005, and Manzolli thanks CNPq, project 308695/2022-4 Brazil, for their support.

References

[1] J. Manzolli and P. F. M. J. Verschure, “Roboser: A Real-World Composition System,” *Computer Music Journal*, vol. 29, no. 3, pp. 55–74, Sep. 2005.

[2] K. C. Wassermann, Kynan Eng, P. F. M. J. Verschure, and J. Manzolli, “Live soundscape composition based on synthetic emotions,” *IEEE Multimedia*, vol. 10, no. 4, pp. 82–90, Oct. 2003.

[3] S. Le Groux, J. Manzolli, and P. F. M. J. Verschure, “Disembodied and Collaborative Musical Interaction in the Multimodal Brain Orchestra,” in *Proceedings of NIME2010*, Sidney, Australia, 2010.

[4] R. E. Page, Ed., *New perspectives on narrative and multimodality*. in Routledge studies in multimodality, no. 1. New York: Routledge, 2010.

[5] Y. U. Everett, “Opera as Film: Multimodal Narrative and Embodiment,” in *The Oxford Handbook of Music and the Body*, Y. Kim and S. L. Gilman, Eds., Oxford University Press, 2019, pp. 398–417. doi: 10.1093/oxfordhb/9780190636234.013.4

[6] W. Berry, Structural functions in music. in Dover books on music, music history. New York: Dover, 1987.

[7] J. Noble, “Perceptual and Semantic Dimensions of Sound Mass” Doctroal thesis, Schulich School of Music McGill University, Montreal, 2018.

[8] M. Antunes da Silva, “Explorando Texturas Sonoras: Análise e Composição com Suporte de Descritores de Áudio e Modelos Computacionais,” Tese de Doutorado, Universidade de Campinas, Campinas,

Brasil, 2024.
[9] E. Cambouropoulos and C. Tsougras, "Auditory Streams in Ligeti's Continuum: A Theoretical and Perceptual Approach," p. 19.
[10] M.Antunes, G. F. Do Espirito Santo, J. Manzolli, and M. Queiroz, "A Psychoacoustic-Based Methodology for Sound Mass Music Analysis," in Music in the AI Era, vol. 13770, M. Aramaki, K. Hirata, T. Kitahara, R. Kronland-Martinet, and S. Ystad, Eds., in Lecture Notes in Computer Science, vol. 13770. , Cham: Springer International Publishing, 2023, pp. 267–281. doi: 10.1007/978-3-031-35382-6_21..
[11] C. Douglas, J. Noble, and S. McAdams, "Auditory Scene Analysis and the Perception of Sound Mass in Ligeti's Continuum," Music Perception, vol. 33, no. 3, pp. 287–305, Feb. 2016, doi: 10.1525/mp.2016.33.3.287.
[12] G. Zanchetta, M. Antunes, D. Rossetti, and J. Manzolli, "Composing Interactions: Co-creative Generation Of Sound And Movement In Multimodal Architectures," in Actes Des Journées D'informatique Musicale 2024, Marseille, France, 2024.
[13] A. Sèdes, P. Guillot, and E. Paris, "The HOA library, review and prospects," in International Computer Music Conference| Sound and Music Computing, 2014, pp. 855–860. Accessed: Nov. 05, 2024. [Online]. Available: https://hal.science/hal-01196453/
[14] A. Evmenenko and D. S. Teixeira, "The circumplex model of affect in physical activity contexts: a systematic review," International Journal of Sport and Exercise Psychology, vol. 20, no. 1, pp. 168–201, Jan. 2022, doi: 10.1080/1612197X.2020.1854818.
[15] B. Truax, "Composing with real-time granular sound," Perspectives of New Music, pp. 120–134, 1990.
[16] Z. Wang, D. Wang, Y. Zhang, and G. Xia, "Learning Interpretable Representation for Controllable Polyphonic Music Generation," Aug. 17, 2020, *arXiv*: arXiv:2008.07122. doi: 10.48550/arXiv.2008.07122.
[17] P. Correia, P. Morais, C. Quintão, C. Quaresma, and R. Vigário, "Assessing the Emotional Reaction to Negative Pictures Through Electrodermal Activity Data," in Advances in Human Factors and System Interactions, vol. 265, I. L. Nunes, Ed., in Lecture Notes in Networks and Systems, vol. 265. , Cham: Springer International Publishing, 2021, pp. 116–124. doi: 10.1007/978-3-030-79816-1_15.

Humanistic Explorations in Algorithmic Art using *Sequence Player* mode of *Playing with Polygons*

Prof. Stephen Erfle, PhD.
Department of International Business and Management, Dickinson College, Carlisle, PA USA
https://blogs.dickinson.edu/playing-with-polygons/
e-mail: erfle@dickinson.edu

Abstract

Machines calculate, but humans evaluate. Generative art will necessarily maintain a humanistic element even in the face of AI advances due to this basic distinction. Electronic string art, a simple form of generative art, allows users to create artistic images by playing with four parameters. The *Sequence Player* mode increases the dynamism involved in such play. Purposeful play leads to human learning, even if that learning is informal in nature. This paper provides examples of such humanistic interactions.

1. Background

Electronic String Art, ESA, simulates traditional string art which creates intricate patterns based on a simple counting rule (connect every 7^{th} nail with string) on a closed set of vertices based on four parameters: *n* and *J* create the *n,J*-star vertex frame, VF; *S* is the number of equally spaced endpoints (think nails) on each line of the VF, and *P* is the counting rule using these endpoints to connect lines in the final image [1]. Two electronic modalities of image creation are available, Excel, and a web version developed by Liam Myles [2]. These basics are laid out in [3] in which I argued that ESA can be used as a resource across the K-12 curriculum.

Due to the simplicity of the underlying model, ESA can be used as a teaching tool, or more generally as an informal learning resource. The four parameters can be adjusted using up and down arrows which allows even young users the ability to play and see what happens. This discovery process is consistent with Stein's *Triex: Explore, Extract, Explain* view of how one can more deeply teach and understand mathematics [4].

An open question is how artificial intelligence, AI, will change the nature of education: What should be taught and how, as well as what needs to be understood to be a productive citizen in this rapidly changing world? How will we maintain our human imprint in a technologically assisted world?

Two themed issues of the *Bulletin of the American Mathematical Society (New Series)*, April and July 2024, examine how AI technology is likely to affect the work of research mathematicians. In one article, Davis notes that despite progress on many fronts, AI has, to this point, failed to surmount the intricacies of elementary mathematical word problems [5]. As U.C. Davis Emeritus Professor Don Chakerian stated in email discussion on this topic [6]: “The article by Davis shows how children can outstrip the most powerful AI systems we have now in dealing with very elementary word problems. You could mention the even greater contrast in the creativity of children in making aesthetic judgments regarding visual figures and patterns. Machines calculate, but humans evaluate. Machines not only lack creativity, but also a sense of play and recreation!”

The present paper focuses on simple regular polygons and stars. By this we mean that the VF will be a triangle, square, or a simple *n*,*J*-star like a pentagram or a 12,5-star, or variations on these simple images in the sense that the images will have rotational symmetry of 120°, 90°, 72°, or 30°, for example.

Regular divisible stars cannot be created in ESA due to the *continuously drawn* nature of the ESA images but “close to” divisible star images such as the 492-line 12,3 and 12,4 images in Fig 1 are possible. Both are based on a regular 12,5 VF star. The upper image has three “squares” tilted 30° to one another. The lower image has four “equilateral” triangles tilted 30° to one another. Both are readily understood based on how close their respective *P* values are those producing a square, *P* = 492/4 = 123, and equilateral triangle, *P* = 492/3 = 164.

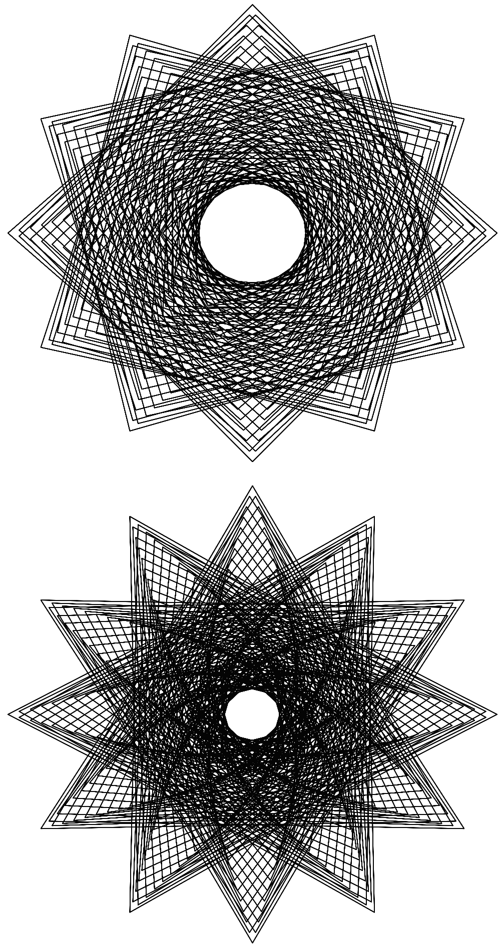


Fig 1. Two almost regular 492-line stars given n = 12, S = 41, J = 5. Upper 12,3-star, P = 121; Lower 12,4-star, P = 163.

The images in Fig 1 are two of 80 distinct full density images available given these values of *n*, *S*, and *J*. As the counting rule *P* varies for fixed *n*, *S*, and *J* one obtains waves of images punctuated by images with fewer lines. ESA Sections 10.2 and 18.5 describe why waves of images appear to adjust smoothly as *P* changes unless the number of lines in the image changes in which case the image sequence seems to flicker relative to surrounding images. Put another way, as noted in ESA Section 5.4, the images are full density (all subdivisions are used) unless SCF = GCD(P,$S·n$/VCF) > 1. This observation forms the basis for the *Sequence Player* mode to which we now turn.

2. *Sequence Player* Mode

One form of dynamism, discussed in [1,3] using the *Home* mode of [2] is how an individual image is drawn as a series of connected line segments given fixed n, S, P, and J. This is the notion of being *continuously drawn.* The present paper focuses on a second form of dynamism: How do images change as the counting rule changes? This can be examined manually using *Home* mode by adjusting P, but it is automated in the *Sequence Player* mode. Two additional modes are available from [2], *Picked Sequences* provides a library of *Sequence Player* image sequences (click to activate), and *Spirals* is discussed elsewhere [7,8].

To avoid the flickering that occurs when SCF > 1, all images in a *Sequence Player* sequence are full density, SCF = 1 (except the *Start Points P* which can have any value of SCF). The *Start Points P* value can be chosen to make a point, but if SCF > 1, it will not be shown in subsequent image sequence iterations. When *Speed* is at its default value of 100, approximately 10 images are shown per second.

Regardless of initial *Start Points* value, if n, S, or J changes, P is reinitialized to 1. The maximum P is set to $n·k·S$ since the maximum number of lines for images in a sequence, **N**, is **N** = $n·k·S$/VCF and k is the number of jumps in a jump set. *The number of lines in in the image*, **N**, changes because VCF changes as J changes. The general rules for lines in the VF with jump sets is $k·n$/VCF with VCF = GCD(**J**,n) and *the jump set sum is* **J** = $J_1+J_2+…+J_k$. In the jump set context, SCF = GCD(P,**N**).

3. The Total Number of Images in a Sequence, T

Some P values between 1 and **N** have SCF > 1 and therefore are not included in an image sequence. Values of P included in the image sequence are coprime to **N**, or SCF = 1.

3.1 The Single Prime Case

If **N** has a single prime factor, a, then all P from 1 to a-1 are coprime to **N** as are a+1 to 2a-1, and so forth. This implies that *the total number of full density images is* **T** = **N**·(a-1)/a. Figs 2 through 4 show examples based on the first three primes, chosen because each sequence has roughly 500 images. See **NOTES** for a link to each figure's image sequence.

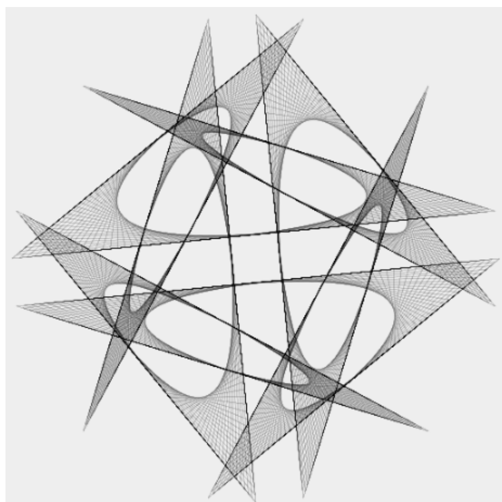


Fig 2. An image based on powers of 2.

Fig 2 has n = 32 but VCF = 8 so one can readily see the four-ness in the resulting image, due to 90° rotational symmetry. This 2^{10} = 1024-line image is a *curved-tip star* [1] because P < S making the four-jump set pattern (9,19,11,17) easy to follow. Note the four different sized nested squares outlined as a result. The image shown is one of 512 images in the sequence because every odd number is coprime to 1024, 512 = 1024·(2-1)/2.

Fig 3 has *n* = 27 but VCF = 9 so the resulting curved tip star image certainly exhibits three-ness or 120° rotational symmetry. The 9-line VF has an internal equilateral triangle and two truncated equilateral triangles [7] due the three-jump set pattern of (6,10,2). This 3^6 = 729-line image is one of 486 in the image sequence because 486 = 729·(3-1)/3.

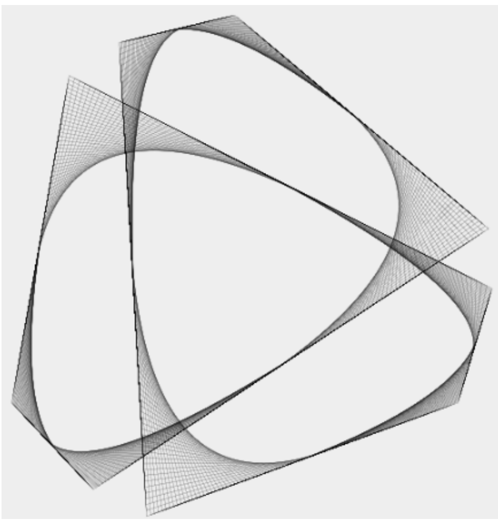


Fig 3. An image based on powers of 3.

Fig 4 has *n* = 25 but VCF = 5 so the resulting image certainly exhibits five-ness or 72° rotational symmetry. This image uses a five-jump set pattern of (8,1,25,22,4). Note the middle jump is a zero jump [1,3] which is why there is a spray of lines at vertices 4, 9, 14, 19, and 24. This 5^4 = 625-line image is one of 500 in the sequence, 500 = 625·(5-1)/5. Indeed, *P* = 312 is the *porcupine* image [1]; note that *P* = 313 = 625-312 is the same static image drawn in reverse.

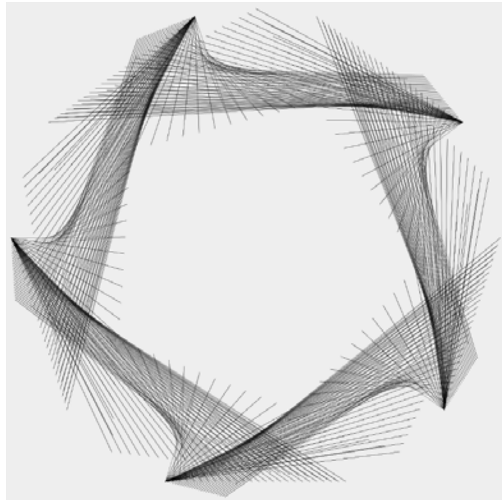


Fig 4. An image based on powers of 5.

3.2 There is Symmetry about the *Porcupine* value of *P*

Although there are 500 images in the Fig 4 image sequence, only half are distinct. The *P* < **N**/2 and *P* > **N**/2 halves are always the same images shown in reverse order. Seeing the images in reverse order may make it appear like they are different images, but this is only an optical illusion. It is always the case that for *P* < *n*, *P* and *n-P* produce the same static image drawn in reverse [1].

It is worth noting that porcupine images have a different look when *n* is even from when *n* is odd. Fig 5 shows the *P* = 245 porcupine image based on the Fig 1 image sequence. Note the prickly 12,5-star nature of this image. Fig 6 shows the *P* = 511 porcupine image based on the Fig 2 four-jump set image sequence. The Fig 7 triangular porcupine, *P* = 364, is based on the Fig 3 three-jump set image sequence.

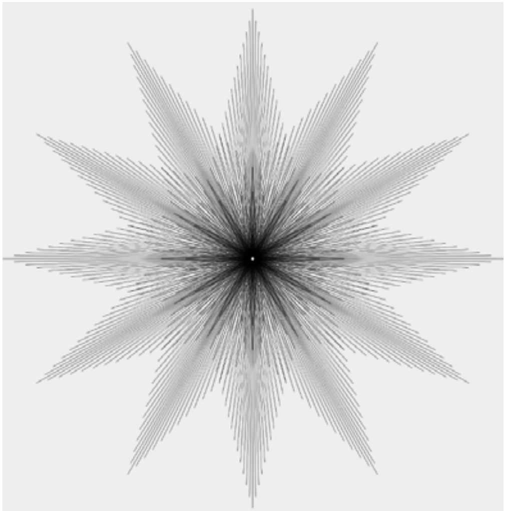


Fig 5. Porcupine 12,5-star for Fig 1.

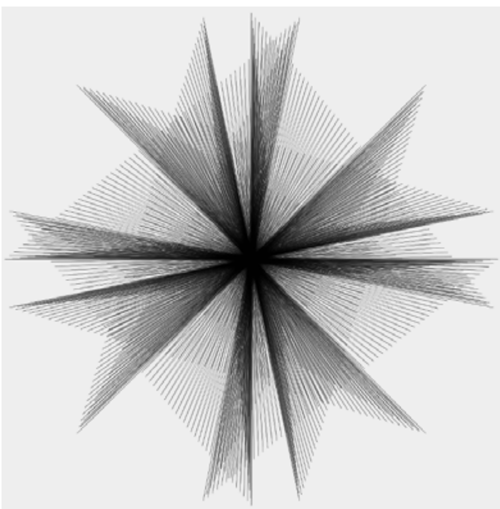


Fig 6. Complex porcupine for Fig 2 with 90° rotational symmetry.

The similarity between even n Figs 5 and 6 porcupines and odd n Figs 4 and 7 porcupines is clear. Since Figs 5-7 are based on Figs 1-3, use the links to those images at **NOTES** and change P as necessary.

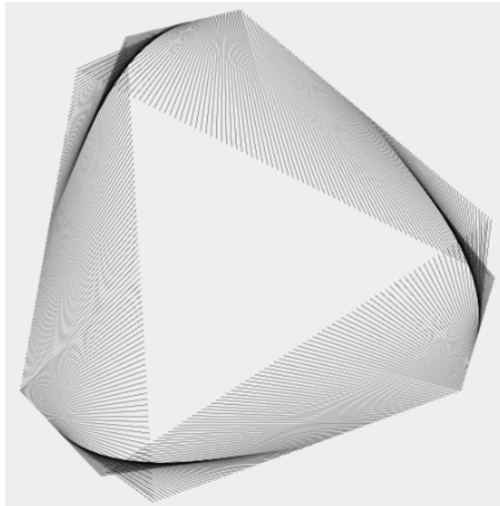


Fig 7. Porcupine image for Fig 3.

3.3 The General Rule for T

Suppose **N** has r distinct primes, $r > 1$, labelled a_1, a_2, …, a_r. Of this number, the fraction $(a_1\text{-}1)/a_1$ are coprime to a_1, the fraction $(a_2\text{-}1)/a_2$ of those remaining are coprime to a_2, and so on. The total number of P values that are coprime to all r primes, **T**, is thus:

$$\mathbf{T} = \mathbf{N}\cdot(a_1\text{-}1)/a_1\cdot(a_2\text{-}1)/a_2\cdot\ldots\cdot(a_r\text{-}1)/a_r \quad (1)$$

We noted above that Fig 1 showed 2 of 80 possible full density images given n, S, and J. $\mathbf{N} = 12\cdot41 = 492$ has three prime factors, 2, 3, and 41. Total images, $\mathbf{T} = 160 = 492\cdot(2\text{-}1)/2\cdot(3\text{-}1)/3\cdot(41\text{-}1)/41$ according to Equation (1) but half, or 80, are distinct due to symmetry about the porcupine value discussed in Section 3.2.

3.4 Waves of Images

Sometimes image sequences do not appear to morph smoothly as P increases. This occurs when there are not enough full density images in a sequence, **T**, but the notion of “enough” cannot be crisply defined. Nonetheless, we can see why waves are more likely to appear smooth if fewer primes are used, or if they are larger primes.

Fig 8 provides four similar images; each is an example of *polygons in a cycle* discussed in ESA Section 10.5 [1]. Each is a part of an image sequence but 8a uses 4 primes, 8b uses 3, 8c uses 2 and 8d uses 1. The percentage of **N** (210, 195, 225, 169 in 8a-8d) used increases from 22.9% to 49.2% to 53.3% to 92.3% in going from 8a to 8d. If you watch the image sequences for each (using the links at **NOTES**), you will see smoothness increase in going from 8a to 8d because 8d has the most images, **T** = 156, despite having the smallest number of lines per image, **N** = 169.

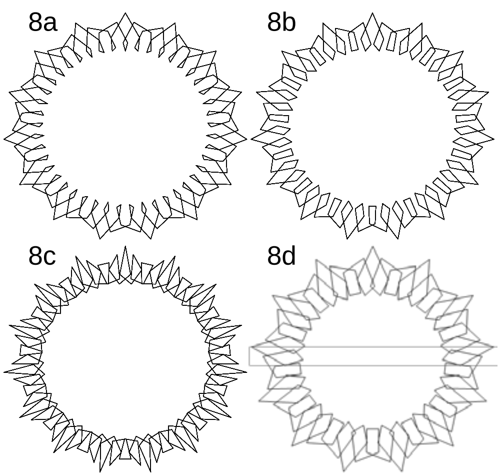


Fig 8. 8a-(15,14,53,4), 8b-(15,13,49,4), 8c-(15,15,56,4), 8d-(13,13,42,4).

One obtains even smoother image transitions if *S* is a larger prime. Fig 9 shows the porcupine image in that image sequence. This sequence was chosen because it has roughly the same **N** = 185 and **T** = 144 as Fig 8 sequences. This **T** is smaller than in 8d, but the sequence appears to morph more smoothly because the number of endpoints on a VF line, *S*, is 37 rather than 13. Note also that Fig 1 morphs smoothly with **T** = 160 because *S* = 41.

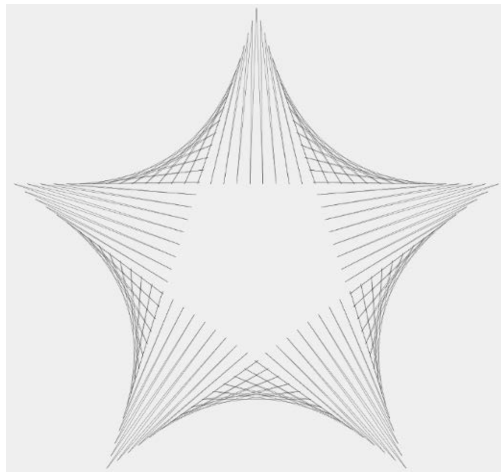


Fig 9. A 5,2-star porcupine image.

4. Elementary Questions

The 12,5-star image sequence used to create Figs 1 and 5 is the default image sequence for *Sequence Player* mode. By adjusting *J* one can readily use this image sequence to explain common divisors. Reduce *J* by 1 to 4 and a triangle VF results because 12 and 4 have 4 in common; a triangular image results if we skip count by 4s (4, 8, 12). Reduce *J* again and we have a square because 12 and 3 have 3 in common; a square VF results if we skip count by 3s (3, 6, 9, 12). This process can be pushed in various ways by varying both *J* and *n* to examine common factors. For example, set *J* = 2, vary *n* and watch an alternation between *n*,2-stars, and *n*/2 polygons. All of this is done while having a mesmerizing image sequence play out in the background.

5. *Equilateral Three-ness*

From an educational perspective, the best question you can ask is: ***Can you find similar but not identical images to an image you find interesting?***

Our interest here is in finding image sequences that exhibit 120° rotational symmetry or *equilateral three-ness*. Fig 10 reduces *S* from 81 to 27 relative to Fig 3 so that the image sequences do not take as much time to view. Each 3^5 = 243-line sequence has 162 images as a result.

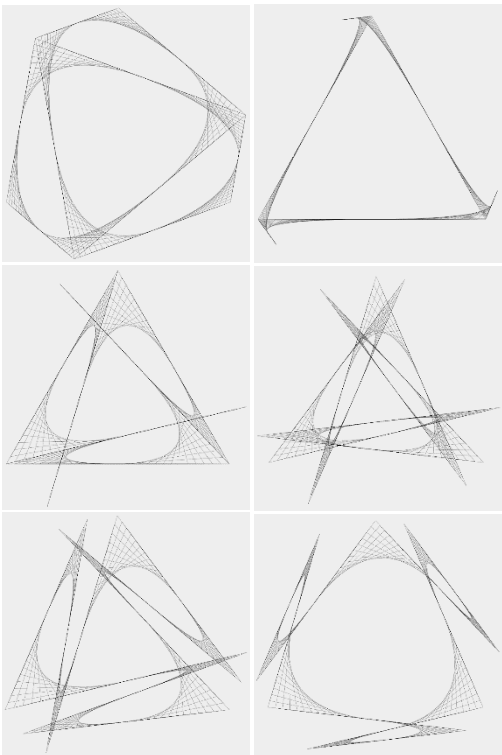


Fig 10. A sampling of image sequences with equilateral three-ness. Each image shown is P = 14. Links are Top 10a, 10b, Middle 10c, 10d, and Bottom 10e, 10f.

The first three versions, 10a-10c, have inscribed equilateral triangles that include vertices of the 27-gon because one of the jumps in the jump set is 9 = *n*/3. Fig 10d has three interior VF equilateral triangles, Fig 10e has two interior and one truncated VF equilateral triangles and all three of the Fig 10f VF triangles are truncated [7]. Figs 10b and 10c are examples of tails and spikes discussed in ESA Section 17.3 [1]. Note also that 10d and 10e are obtained from 10c by adjusting J_1 up or down 1 and J_2 down or up 1 relative to 10c. Doing so changes the 0° spike into a 180/27 = 6.67° sharp angle according to the inscribed angle theorem.

The image sequences in Fig 10 are not exhaustive, but simply are examples of some of the possibilities given *n* = 27 and *k* = 3. Of course, other values of *n* produce similar equilateral three-ness images (and different ones), all that is required is that *n* is a multiple of 3 and 3 = *n*/VCF. The images in Fig 10 worked because VCF = 9 in each case and one need only adjust jumps in the jump set to find additional versions.

Equilateral four-ness, or 90° rotational symmetry, is similarly easy to obtain. One need only have *n* be divisible by 4 with 4 = *n*/VCF. From here you might challenge students to create images that have 5,2-ness, or 7,3-ness.

6. Extensions

As you adjusted jumps in a three-jump set you may have run across situations where a scalene triangle results. This happens whenever the sum of jumps is a multiple of *n*. This forms the basis for obtaining “closest to equilateral” triangles when *n* is not divisible by 3.

Figs 11a and 11b show the smallest *n* isosceles, but not equilateral, porcupine images where the difference in angles is less than 1° (the difference is 180/181° and 180/182°, respectively). The apex angle of 11a is greater than 60° and 11b is less than 60° but both are virtually indistinguishable from (3,81,121,1).

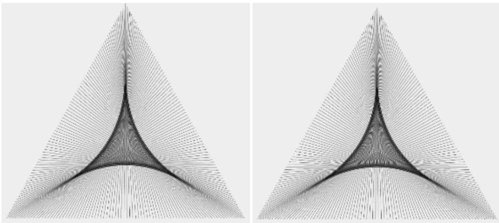


Fig 11. Two isosceles porcupines.

7. Summary

Electronic String Art combines art and mathematics in a way that allows users to understand mathematics more deeply, simply by adjusting parameters and watching what happens. The *Sequence Player* mode of *Playing with Polygons* provides a dynamic complement to the *Home* mode that helps users understand how individual images emerge as a result of applying a simple counting rule. Both modes were created for independent explorations, but both are useful in more formal K-12 classroom settings to support teaching of various topics in a visually appealing way.

Notes

Each link takes you directly to the web version [2] set to that figure in *Sequence Player* mode. Links to Figs 1-9 note the number of lines in the image, **N**, and the number of images in a sequence, **T** and shown *P* as **N**, **T**, *P*. Figs 10 and 11 note the jump set associated with each panel.

Fig 1,492, 160, 121. Fig 2,1024, 512, 37.

Fig 3,729, 486, 47. Fig 4,625, 500, 312.

Fig 8a,210, 48, 53. Fig 8b, 195, 96, 49.

Fig 8c,225, 120, 56. Fig 8d,169, 156, 42.

Fig 9,185, 144, 92.

Fig10a (6,9,3). Fig 10b (9,26,1).

Fig 10c (9,16,11). Fig 10d (10,15,11).

Fig 10e (8,17,11). Fig 10f (7,21,8).

Fig 11a (60,61,60). Fig 11b (61,60,61).

References

[1] S. E. Erfle, *Electronic String Art: Rythmic Mathematics*, 2024.

[2] L. Myles, “Playing with Polygons.” https://www.playingwithpolygons.com/.

[3] S. E. Erfle, “Electronic String Art as Educational Algorithmic Art.” *XXVII Generative Art Conference*, 2023, Celestino Soddu, Editor. Rome, Italy. (10 page paper)

[4] S. K. Stein, “The Triex: Explore, Extract, Explain.” *Humanistic Mathematics Network Journal*, vol. 1, no. 14, pp. 6–8, 1996.

[5] E. Davis, “ Mathematics, Word Problems, Common Sense, and Artificial Intelligence.” *Bulletin of the American Mathematical Society (New Series)*, vol. 61, no.2, April 2024, pp. 287-303.

[6] D. Chakerian, email communication, May 19, 2024.

[7] S. E. Erfle, “Playing with Polygons.” https://blogs.dickinson.edu/playing-with-polygons/.

[8] S. E. Erfle, “Using Archimedean Spirals to Explore Fractions.” *Proceedings of Bridges 2021: Mathematics, Art, Music, Architecture, Culture*, pp. 397-402, 2021.

Coupled and Intelligent Interactions in Reaction-Diffusion Systems

Angela Ferraiolo, MFA.
Sarah Lawrence College, Bronxville, New York
https://littleumbrellas.net/
https://www.instagram.com/_angela_ferraiolo
e-mail: aferraiolo@gmail.com

Abstract:
This project investigates the dynamic interaction between two coupled reaction-diffusion systems based on the Gray-Scott model. Reaction-diffusion systems are well-known for their ability to model complex pattern formation in fields such as physics, chemistry, and biology. Here, we use a side-by-side comparison of reaction-diffusion by coupling two independent instances and introducing an AI-driven leader-follower dynamic. Each system operates under defined behavioral states, such as synchronization and opposition, while continually evaluating its role and adapting its strategy. Through dynamic decision-making, these systems can cooperate, compete, or diverge, providing a rich exploration of emergent patterns influenced by computational intelligence and environmental feedback.

Keywords: reaction-diffusion systems, Gray-Scott model, coupled systems, self-organization, leader-follower dynamics, dynamic pattern

Overview:

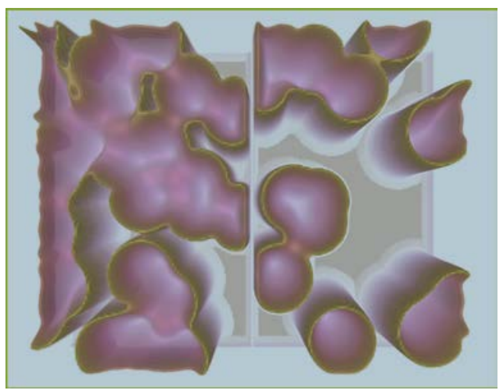


Figure 1. Coupled Diffusion Systems, Gray-Scott model

Reaction-diffusion systems have long served as a cornerstone in the study of self-organization and pattern formation. From modeling chemical reactions to simulating biological growth, these systems demonstrate how simple interactions between components can give rise to complex structures. The Gray-Scott model is a particularly versatile reaction-diffusion system, capable of generating a wide range of patterns through the interaction of activator and inhibitor chemicals.

This project extends the Gray-Scott model by coupling two independent systems, introducing a leader-follower dynamic influenced by computational intelligence. Through a truth table linking system modes and behaviors, each system evaluates its current state and adapts its response. The interaction unfolds in cycles, with systems choosing strategies to either synchronize their behaviors (SYNC mode) or actively oppose each other (OPPOSE mode).

The motivation for this project was to create a visual exploration of LEADER / FOLLOWER dynamics and to explore how these dynamics influence emergent patterns, blending artistic expression with complex adaptive systems.

1. Design

The system facilitates dynamic interactions between two independent reaction-diffusion processes. Rather than directly exchanging chemical concentrations or modulating each other's parameters, model instances interact through role-based dynamics. Each system evaluates its state in real-time, assigns itself a role (LEADER or FOLLOWER), and adapts its behavior based on the leader-follower relationship.

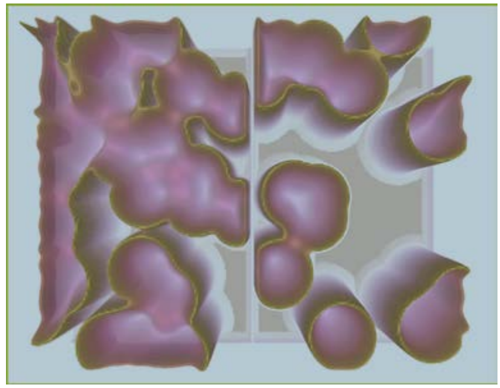


Figure 2. Coupled Diffusion Systems, Gray-Scott model

Decision-making algorithms enable each system to select a strategic mode (SYNC or OPPOSE) and a corresponding behavior. These decisions are informed by real-time metrics, such as global activity and local variance. This framework prioritizes immediate, context-driven responses, allowing the interactions to evolve dynamically without preconditioning.

Side-by-side visual feedback ensures that the states and decisions of each system are legible. Patterns generated by the system serve as an expressive medium, illustrating the evolving interplay between leader and follower roles, modes, and behaviors.

2. Implementation

The project was implemented in three distinct phases:

Phase 1: Independent System Initialization

Each system was initialized as an independent Gray-Scott reaction-diffusion model. This phase focused on setting up the foundational chemical processes, ensuring each system could generate patterns autonomously based on standard reaction-diffusion dynamics.

Phase 2: Role-Based Coupling

The systems were coupled indirectly by introducing leader-follower dynamics. Rather than exchanging chemical concentrations or directly modulating parameters at shared boundaries, the systems interact through assigned roles (`LEADER` or `FOLLOWER`) and modes (`SYNC` or `OPPOSE`).

Key aspects of this phase include:

Role Assignment: Each system evaluates its real-time metrics, such as global activity and local variance, to assign itself a role. This process occurs periodically, ensuring that the leader and follower roles dynamically shift.

Mode Selection: Depending on the role assignment, each system selects a mode of interaction. The `LEADER` determines whether to synchronize or oppose its counterpart, while the `FOLLOWER` adapts to the leader's mode.

Behavior Execution: The chosen mode and the system's internal metrics dictate the execution of visual behaviors, such as fractal growth, ripples, and noise disturbances.

Phase 3: Adaptive Decision-Making

Building on the role-based framework, decision-making algorithms were implemented to enable adaptive behavior:

Real-Time Metrics: Each system relies on real-time calculations (e.g., global activity, edge activity) to guide its decisions.

Dynamic Strategy Switching: Systems periodically switch between predefined strategies (probabilistic, oscillating, perturbed, and pattern-based) to decide how to assign roles, ensuring diverse and evolving interactions.

Behavior Mapping via Truth Table: The truth table maps each mode and current behavior to a new behavior, creating a flexible decision-making framework that responds dynamically to changes in interaction states.

This implementation results in a visually compelling and conceptually rich system where two reaction-diffusion processes interact dynamically. The emergent patterns reflect the underlying roles, modes, and strategies, demonstrating how simple systems can exhibit complex, adaptive behaviors without relying on historical data or direct parameter exchange.

3. Results

This section presents the visual outcomes of the system, demonstrating how the interaction of roles (`LEADER`, `FOLLOWER`) and modes (`SYNC`, `OPPOSE`) influences emergent pattern.

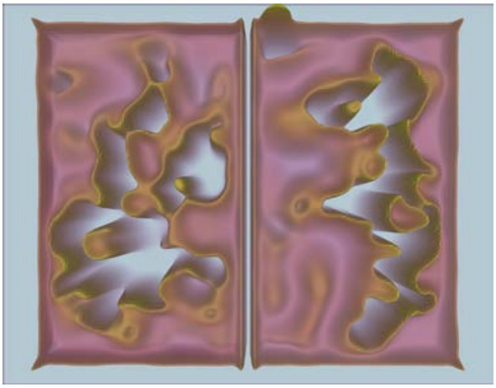


Figure 3. Coupled Diffusion Systems, Gray-Scott model, Leader Sync, Follower Sync

In Fig. 3, Leader Sync, Follower Sync, the system exhibits harmonious behavior, with both leader and follower synchronizing their actions. This alignment results in smooth, continuous pattern growth across both systems.

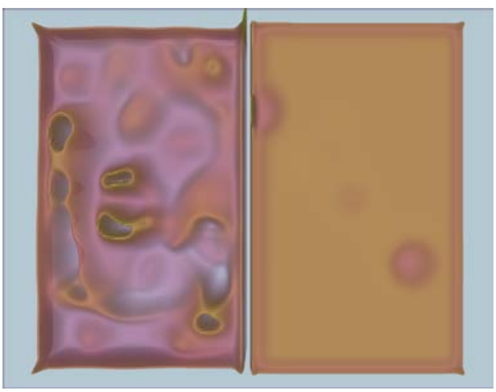


Figure 4. Coupled Diffusion Systems, Gray-Scott model, Leader Sync, Follower Oppose

In Fig. 4, Leader Sync, Follower Oppose, the leader attempts to synchronize, but the follower opposes. The resulting pattern displays conflicting growth dynamics producing localized disruption zones.

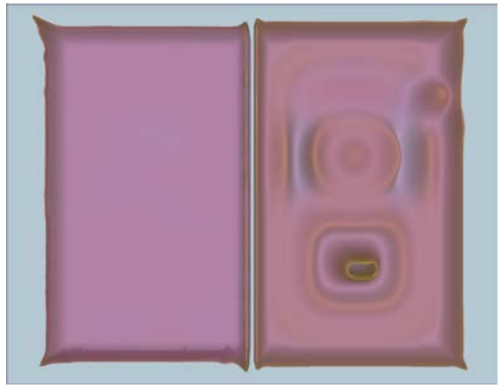


Figure 5. Coupled Diffusion Systems, Gray-Scott model, Leader Oppose, Follower Oppose

In Fig. 5, Leader Oppose, Follower Sync, the leader's opposing strategy disrupts the follower's synchronized behavior. The leader's influence introduces chaotic disturbances into the otherwise harmonious patterns of the follower.

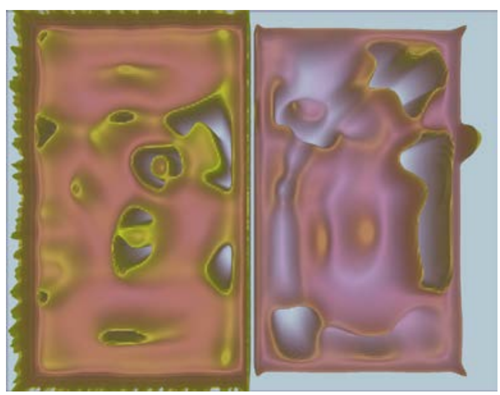


Figure 6. Coupled Diffusion Systems, Gray-Scott model, Leader Oppose, Follower Oppose

In Fig. 6, Follower Oppose, Leader Oppose, both systems engage in opposing modes, leading to highly chaotic and competitive patterns.

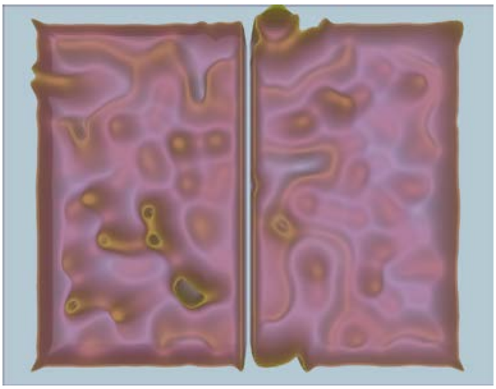


Figure 7. Coupled Diffusion Systems, Gray-Scott model, Leader Oppose, Follower Oppose

In Fig. 7, Follower Sync, Leader Sync, both systems synchronize their behaviors, regardless of their roles. The resulting patterns exhibit high coherence, with visually harmonious structures emerging across the interaction zone.

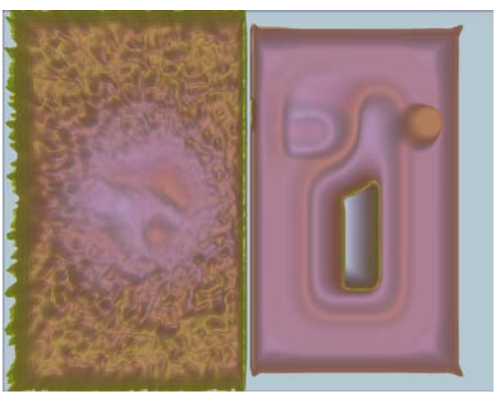


Figure 8. Coupled Diffusion Systems, Gray-Scott model, Leader Oppose, Follower Oppose

In Fig. 8, Follower Oppose, Leader Sync, the follower's opposing mode introduces disruptions into the leader's synchronized patterns.

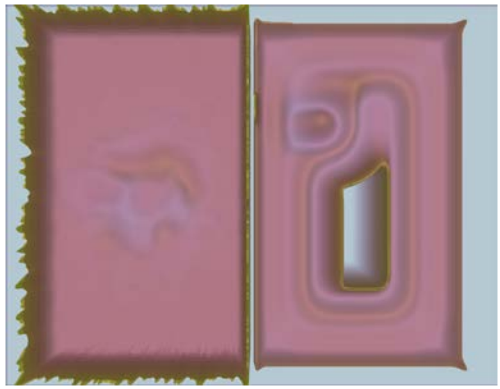


Figure 9. Coupled Diffusion Systems, Gray-Scott model, Leader Oppose, Follower Oppose

In Fig. 9, Follower Sync, Leader Oppose, the leader's opposing mode overwhelms the follower's synchronized attempts, creating an imbalance. Patterns appear fragmented, with the leader's behavior driving the overall dynamics.

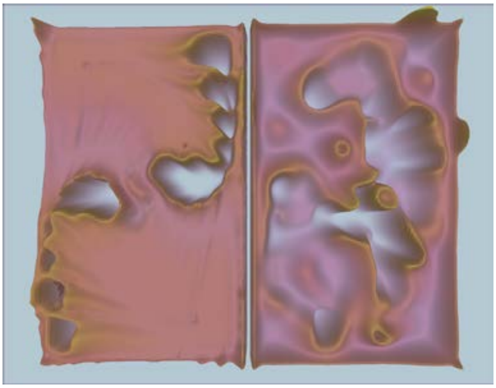


Figure 10. Coupled Diffusion Systems, Gray-Scott model, Leader Oppose, Follower Oppose

In Fig. 10, Follower Oppose, Leader Oppose, both systems adopt opposing modes, resulting in a highly competitive environment. The visual output is characterized by fragmented, high-intensity patterns with significant disruption throughout the shared space.

These results illustrate the system's ability to produce diverse and dynamic visual patterns, driven by the interplay of roles and modes. The emergent behaviors highlight the complex, adaptive interactions reaction-diffusion can achieve through simple decision-making rules and role-based dynamics.

4. Conclusion

By integrating dynamic coupling with adaptive decision-making, this project expands the scope of reaction-diffusion modeling. The resulting system is a fusion of chemical pattern formation and role-driven interaction, illustrating how simple rules can yield complex, emergent behaviors. This approach opens new avenues for simulating diffusion systems, including ecological, biological, and social environments, where multiple entities interact, adapt, and influence one another over time.

4. Bibliography

1 Camazine, Scott, Jean-Louis Deneubourg, Nigel R. Franks, James Sneyd, Guy Theraulaz, and Eric Bonabeau. 2001. *Self-Organization in Biological Systems*. Princeton, NJ: Princeton University Press.

2. Epstein, Irving R., and John A. Pojman. 1998. *"An Introduction to Nonlinear Chemical Dynamics: Oscillations, Waves, Patterns, and Chaos".* Oxford: Oxford University Press.

3. Gierer, Alfred, and Hans Meinhardt. 1972. "A Theory of Biological Pattern Formation." *Kybernetik* 12 (1): 30-39.

4. Gray, Peter, and Stephen K. Scott. 1984. "Autocatalytic Reactions in the Isothermal, Continuous Stirred Tank Reactor: Oscillations and Instabilities in the System A + 2B → 3B, B → C." *Chemical Engineering Science* 39 (6): 1087-1097.

5. Helbing, Dirk. 2012. *Social Self-Organization: Agent-Based Simulations and Experiments to Study Emergent Social Behavior*. Berlin: Springer.

6. Holland, John H. 1998. *Emergence: From Chaos to Order*. Cambridge, MA: Perseus Books.

7. Kondo, Shigeru, and Takashi Miura. 2010. "Reaction-Diffusion Model as a Framework for Understanding Biological

Pattern Formation." *Science* 329 (5999): 1616-1620.

8. Murray, James D. 2003. *Mathematical Biology: I. An Introduction*. 3rd ed. New York: Springer-Verlag.

9. Pearson, John E. 1993. "Complex Patterns in a Simple System." *Science* 261 (5118): 189-192.

10. Sutton, Richard S., and Andrew G. Barto. 2018. *Reinforcement Learning: An Introduction*. 2nd ed. Cambridge, MA: MIT Press.

11. Turing, Alan M. 1952. "The Chemical Basis of Morphogenesis." *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* 237 (641): 37-72.

World Beings

Prof. Alain Lioret
Université Paris 8
ESGI Digital Lab
e-mail: alainlioret@gmail.com

Kamal Hennou
ESGI
e-mail: hennou@myges.fr


Abstract

"World Beings" is a project aimed at creating an innovative World Building system using artificial creatures. The idea is to allow a complex ecosystem of artificial beings to work together, building the content and rules of their own world. These creatures are different from other artificial creatures because they are involved in specific tasks, such as creating lights, emitting sounds, setting climate or political rules, etc. This world operates with certain rules, based on several artificial intelligence and artificial life algorithms, including cellular automata, genetic algorithms, Generative Adversarial Networks, and continuous evolution through machine learning.

1. AI, Computational Creativity, and Emergence

The *World Beings* system fits within the scope of computational creativity and emergence, as well described by Manovich, who explores the intersection between artificial intelligence and human creativity. Manovich examines how AI systems, such as Generative Adversarial Networks (GANs), replicate works of art and music in ways that sometimes deceive human observers. The text raises questions about the nature of creativity, how we attribute intentionality and autonomy to artificial systems, and whether creativity can truly be replicated by machines. Manovich also critiques the anthropocentric perception of creativity and intelligence. (1)

Similarly, Mattia explores how artificial intelligence (AI) is integrated into artistic production and redefines the boundaries between human and machine creation. He highlights the rapid evolution of creative AIs capable of generating visual works, music, texts, and poems, often indistinguishable from those produced by humans. However, one major question remains: Can AI's creations be considered art in their own right ?

Mattia emphasizes that technique alone does not define art, but rather the underlying aesthetic intent and artistic theory are essential. He also notes a negative bias toward AI-created works, with people tending to evaluate these creations as less pleasing or meaningful

when they know they are machine-generated.
Mattia suggests that AI should not be seen as an autonomous creator but rather as a powerful tool available to artists. Examples such as Generative Adversarial Networks (GANs) and text-to-image algorithms demonstrate AI's potential to assist artists in their creative processes without replacing human creativity. (2)
Jon McCormack and Alan Dorin explore critical concepts around generative art, focusing on the concept of emergence. They analyze how generative systems can produce unexpected phenomena, linking art, biology, and systems theory. They introduce the concept of the "computational sublime," where machine-generated artistic processes evoke contradictory emotions of pleasure and awe. (3)
Many authors follow this line of thought to create intelligent avatars, virtual assistants, interactive environments, as well as digital twins, blockchain, and neural interfaces. (4)
Anna Daudrich explores the use of algorithms in artistic creation, comparing human artists' approaches with the algorithmic methods used by computers. She analyzes the works of three artists: Manfred Mohr, a pioneer of computer-assisted art; Sol LeWitt, a conceptual artist; and Hans Arp, a Dadaist artist. Her work highlights the differences between art generated by precise algorithms, like Mohr's computer works, and the more intuitive or random processes used by LeWitt and Arp. While Mohr uses algorithms to generate works based on strict rules, LeWitt and Arp employ freer or random methods. (5)

2. The Beings Series

For more than 20 years, Alain Lioret has been exploring the creation of artificial entities capable of interacting with humans in meaningful ways. Lioret develops the concept of *Painting Beings*, virtual entities programmed to analyze and react to works of art. These beings use techniques such as cellular automata and genetic algorithms to understand and recreate images, including works by great masters. Lioret examines how these entities perceive our world, not by recognizing objects or figures like humans but by analyzing pixel arrangements. This raises questions about how these beings understand and interact with images, suggesting new perspectives for interactive and generative art. (6)

Fig 1 Painting Beings. © Lioret 2004

With *Galatema*, Alain Lioret presents an innovative system for generative filmmaking. *Galatema* is based on an ecosystem of artificial creatures, such as *Painting Beings*, *Plant Beings*, *Light Beings*, and *Sound Beings*, who collaborate to create 2D and 3D films autonomously. Each artificial being controls specific elements of the film, such as lights, sounds, or camera shots, using genetic algorithms, cellular automata, and machine learning. The system also incorporates geometric rules such as the golden ratio and Fibonacci numbers to organize the creative parameters. *Galatema* allows for the production of classical or experimental films, thus exploring new forms of digital creativity. (7). *Galatema* system can be considered as the precursor of World Beings.

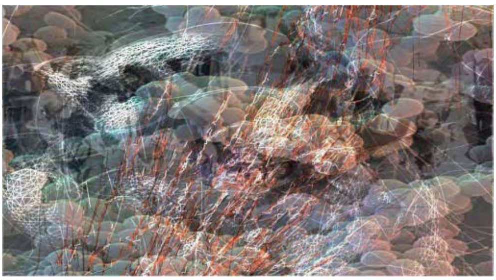


Fig 2 Galatema. © Lioret 2008

In the same series, Alain Lioret and his co-authors explore the creative potential of self-organizing systems, particularly in the fractal patterns observed in nature. They suggest that these fractal structures, present at various scales in the universe, could be replicated by computer simulations to generate artistic creativity and artificial life phenomena. The importance of critical self-organization processes, at the boundary between chaos and order, is emphasized to produce non-adaptive innovations, a concept inspired by biological exaptation. These systems are studied as potential models for developing new forms of computational art and biomimetic engineering. (8)

Fig 3 Fractal Beings. © Lioret 2012

3. Worldbuilding

Worldbuilding is a set of tools and methods for creating fictional universes. There are numerous writings and tools to approach this subject.

Thus, *The Planet Construction Kit* by Mark Rosenfelder is a comprehensive guide for creators of fictional worlds, whether they are authors, game designers, or fans of science fiction and fantasy. The book provides practical tools for developing believable imaginary planets, covering aspects such as geography, ecology, political systems, cultures, and languages.

Rosenfelder explains how to design varied landscapes, realistic climates, and coherent ecosystems while ensuring that geographical aspects logically influence human or extraterrestrial societies. The book also addresses the creation of fictional languages and the social, economic, and political interactions between peoples.

Another central point is the creation of historical and cultural narratives to give depth to fictional worlds. Rosenfelder offers concrete examples and practical advice to avoid common pitfalls in worldbuilding, such as overly simplistic or unrealistic systems. (9)

In the same vein, *Emissary's Guide to Worlding* by Ian Cheng explores the concept of "Worlding," or the art of creating fictional worlds that can outlast their creator. Cheng presents strategies for overcoming creative obstacles by incorporating a deep understanding of the psychological dynamics of the artist. The book also provides practical advice for generating worlds capable of capturing the imagination while remaining functional, and it explores how worlds can evolve to have a life of their own, independent of their creator. (10)

4. Study of the Masters of Worldbuilding

The document *"Tolkien and Worldbuilding"* by Catherine Butler explores how J.R.R. Tolkien created the rich and complex universe of Middle-earth. It examines how Tolkien, considered a master in the art of world-building, spent nearly six decades developing geographical, historical, cultural, and linguistic details to make his world believable and coherent. Butler discusses the notion of "sub-creation," introduced by Tolkien in his essay *On Fairy-Stories*, where he describes the act of creating an imaginary world as a form of secondary creation, inspired by the real world but rearranged.

The article also explores how the differences between *The Hobbit* and *The Lord of the Rings* reflect the evolution of Tolkien's universe, influenced by readers' expectations, shifting from a children's story to an epic for adults. Butler emphasizes the importance of internal consistency in Tolkien's work, despite the inevitable contradictions due to the evolution of his universe over several decades. Ultimately, Butler shows that Tolkien's influence lies not only in the depth and richness of his universe but also in his ability to bring to life a world that seems both autonomous and connected to our reality—a significant legacy in modern fantasy literature. (11)

The text *"Fantasy After Representation"* by Ryan Vu explores the evolution of fantasy and science fiction genres in contemporary popular culture. These systems have allowed participants to create narratives across various media. The essay emphasizes that the *Game of Thrones* series embodies this shift toward a culture where genre codes and conventions are played with and manipulated rather than offering deep ontological differences. Vu argues that the fantasy genre has become a hegemonic cultural norm, with narratives that play on viewers' expectations while subverting them, reflecting a post-representational world. (12)

5. Tools for Generative Worldbuilding

Genie: Generative Interactive Environments introduces a generative AI model capable of creating interactive environments from simple text prompts, images, or sketches. Developed by Google DeepMind, Genie is designed to generate virtual worlds in which users can interact in real-time. The model is based on unsupervised learning from over 200,000 hours of online gaming videos without requiring action or text annotations. Genie uses a dynamic autoregressive model that predicts each next frame in a generated environment from latent actions, offering great flexibility for creating varied environments. The system can simulate physical objects and adapt its behaviors to different situations. Genie paves the way for more general learning agents by enabling training from videos without annotated actions while creating immersive and interactive environments for users. (13)

Fig 4 GENIE. Google Deepmind

World-GAN: A Generative Model for Minecraft Worlds presents a new method for procedural content generation based on machine learning (PCGML) in Minecraft. The approach uses a 3D Generative Adversarial Network (GAN) capable of generating world fragments from a single example, such as structures or biomes, while capturing significant variations.
The method integrates the block2vec concept, a dense representation inspired by word embeddings in Natural Language Processing (NLP), allowing the model to efficiently handle Minecraft's wide diversity of blocks. Unlike representations based on one-hot encoding, block2vec enables the generation of more varied content and improves the inclusion of rare blocks.
Thus, *World-GAN* through examples of biomes and structures generated in Minecraft, show that the model can reproduce varied visual styles while maintaining structural coherence. (14)

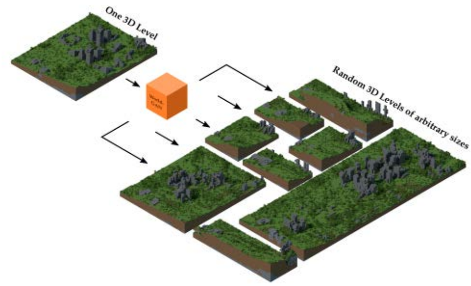


Fig 5 World-GAN

WorldSmith: Iterative and Expressive Prompting for World Building with a Generative AI presents a tool designed to facilitate the creation of fictional worlds using generative artificial intelligence models. *WorldSmith* combines multi-modal inputs such as texts, sketches, and region paintings to allow users to visualize and modify fictional worlds iteratively. The tool offers hierarchical image composition features, enabling the creation of complex worlds layer by layer.
One of the major innovations is the introduction of two new interaction dimensions: hierarchical prompting (allowing multi-level descriptions for distinct elements) and spatial prompting (allowing users to draw and reposition elements to guide the AI). *WorldSmith* was tested in a study with 13 participants, who found this tool particularly useful for speeding up the process of creating complex worlds while maintaining a high degree of creativity. (15)

6. Inspiration from Oulipo

Oulipo is a literary group founded in 1960 by Raymond Queneau and François Le Lionnais. Oulipo is distinguished by its experimental approach, using formal constraints and mathematical structures to generate new literary forms.

The group's members, including writers and mathematicians, focused on exploring the potential of language through constraints such as the lipogram and combinatorics. (16)

The text *"L'Oulipo et la géométrie: l'étrange utilité de la table des matières"* by Natalie Berkman explores how Oulipo uses geometric constraints in literature. We discover how Calvino employs a geometric structure in his table of contents to organize fragments of imaginary cities and offer multiple readings, while Audin relies on Pascal's theorem to structure his narrative. These geometric constraints allow both authors to reconcile form and content while illustrating the tension between order and chaos. Geometry, in this case, becomes a tool for better understanding reality through abstraction while enriching the Oulipian literary approach. (17)

Among other major interesting texts, we can cite *La Vie mode d'emploi* by Georges Perec, published in 1978, an emblematic novel of Oulipian literature, combining literary creativity and mathematical constraints. The work is structured like a puzzle, where each piece, each chapter, reveals part of the story, allowing the reader to navigate through the different layers of life. (18)

Beyond Oulipo, other authors have used interesting mathematical structures to create their universes. For example, *Une nouvelle manière de voir le monde. Jalons pour une poétique de la figure fractale en littérature* by Dominique Raymond explores the transposition of fractals, a complex geometric figure identified by Benoît Mandelbrot, into the literary domain. Fractals, characterized by their self-similarity and infinite iteration, are compared to the “mise en abyme” in literature, a device presenting elements nested or replicated at different scales. Raymond analyzes literary examples such as *Cent mille milliards de poèmes* by Raymond Queneau and *Fractales* by Ian Monk, showing how repetition and self-similarity are exploited in the form and content of certain texts.

Raymond also examines how these mathematical concepts influence structure and narrative in some works, particularly through themes of order and chaos. The article concludes that although fractals are not always perfectly translatable into literature, they inspire innovative narrative structures and allow for reflection on infinity and repetition in the text. (19)

7. Further with Metaphor

It is interesting to see the importance of metaphors in human understanding and communication. Metaphors are not limited to poetic or literary expressions but play a central role in how we think and interpret the world. Metaphors allow abstract concepts to be linked to concrete experiences, thereby facilitating the learning and assimilation of new information. We can also examine how metaphors influence our understanding of complex notions, such as science or philosophy, by providing interpretive frameworks. (20)

We can also refer to *"Création de monde chez le concept artist: le modèle de la métaphore générative"* by Stéphane Meury, which explores the process of creating imaginary worlds, particularly in concept art. The author draws on the works of Donald Schön to understand how metaphors play a key role in this process. Through this project, he shows that metaphorical thinking allows one to overcome stereotypical ideas and open up new creative possibilities. Metaphors act as tools to enrich the design of coherent and immersive worlds. (21)

8. A World Based on Lenia

Bert Wang-Chak Chan presents Lenia, a system of continuous cellular automata capable of generating artificial life forms. Lenia, inspired by Conway's *Game of Life*, differs through its continuous space-time and smooth update rules, creating patterns resembling living organisms. More than 400 species, organized into 18 families, have been discovered, exhibiting characteristics such as self-organization, motility, and adaptability. These artificial life forms share similarities with biological organisms in terms of structure and behavior. (22)

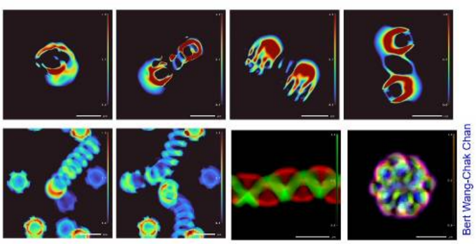


Fig 6 Lenia Creatures

Bert Wang-Chak Chan also introduces experimental extensions of Lenia. These extensions include higher dimensions, multiple kernels, and channels, allowing Lenia to explore more complex behaviors. By using semi-automatic search algorithms like genetic algorithms, new phenomena have been discovered, such as self-replication, pattern emission, and the emergence of structures resembling eukaryotic cells. These "virtual eukaryotes" exhibit behaviors such as internal division of labor, reminiscent of real biological processes. (23)

Flow Lenia is an extension of Lenia. *Flow Lenia* introduces mass conservation, enabling the creation of localized creatures that can evolve in a shared environment. This model promotes the emergence of complex behaviors without requiring complex parametric exploration. *Flow Lenia* also facilitates the integration of dynamic parameters, allowing for multi-species simulations where each species can have different evolution rules. The results show that *Flow Lenia* can generate creatures with directed movement and other specific behaviors through evolutionary optimization strategies. This advancement opens the way for intrinsic evolution, with creatures capable of adapting, dividing, and interacting in an artificial ecosystem. (24)

We decided to use the new Lénia cellular automata as the basis for the World Beings.

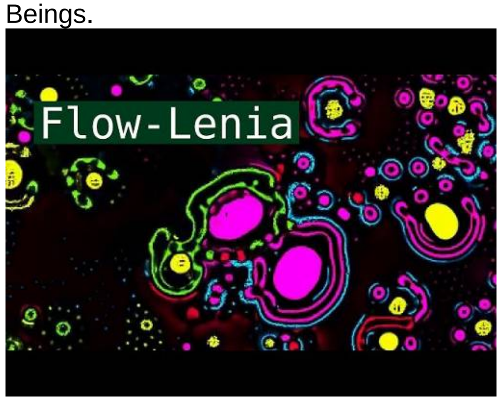


Fig 7: Flow Lenia Creatures

9. Quantum Worldbuilding

In his text *"L'Art de Montrer l'Invisible, L'Autre Révolution Quantique"*, Charles Antoine explores how scientists and artists attempt to reveal the invisible across various disciplines. The article examines three forms of the invisible: what cannot be seen, what cannot be shown, and what is inconceivable. He relates these forms of the invisible to quantum physics, a discipline that has revolutionized our understanding of the universe and introduced new, often counterintuitive concepts. Quantum physics, omnipresent in our daily technologies, also influences art, with artists like James Turrell and Antony Gormley exploring the invisible through light and matter. (25)

With his work *"Quantum Generative Art"*, Alain Lioret explores the application of quantum computing in generative artistic creation. Lioret begins by tracing the evolution of generative art inspired by quantum technologies, highlighting concepts like superposition of states, particle entanglement, and teleportation, which bring new perspectives to artists. He describes how quantum algorithms, such as Grover's and Shor's, enable the generation of unprecedented works by using quantum gates (Hadamard, Pauli, Toffoli, etc.), replacing traditional logical circuits.

We can also cite early works of pioneering artists like Antony Gormley and Julian Voss-Andreae, who used these technologies to create sculptures and installations inspired by quantum physics.

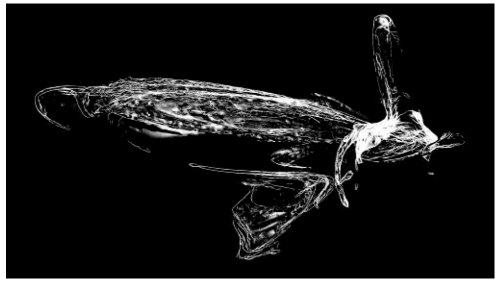


Fig 8: Quantum Swan. © Lioret 2021

Finally, Lioret addresses Quantum Generative Adversarial Networks (QGANs), a revolutionary technology combining artificial intelligence and quantum computing to generate realistic images. These networks use two quantum circuits, a generator and a discriminator, which improve each other to create quantum states capable of producing even more innovative artworks. (26)

10. Implementation of the World Beings

Our World Beings system uses open-source software, including Blender, Godot Engine, and Python for programming.

This is a first version of the World Beings, still very incomplete (we haven't yet turned our attention to politics, economics, geography, etc.).

The old Beings created by Lioret are back in action, but all in new versions, based on the new Lénia cellular automata and quantum calculations.

Painting Beings, Eye Beings, Light Beings, and Plant Beings are directly involved in the scenes of the World.

However, these are new versions of the Beings, which now use new algorithms and include quantum calculations.

Painting Beings II are now based on cellular automata of the Lenia type (22) (23).
What's more, all calculations are performed using the Qiskit library, and all pass through a quantum simulation.
The widespread use of quantum is naturally justified by the fact that we live in a quantum world, and the Beings are updated with these new algorithms.
Similarly, all the other Beings used for Worldbuilding are quantum (note that this can be seen as a new version of the Cinema Beings "*Galatema*") (7).

11. Experimentations

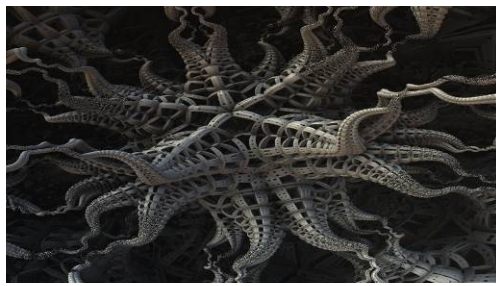
To demonstrate how World Beings work, we put them into action (with Lénia as the artificial life engine) in the production of a few environments where they produce universes in the Mandelbulb world (with Quantum Fractal Beings) (8).
Here is an example of fractal code produced by the Beings:

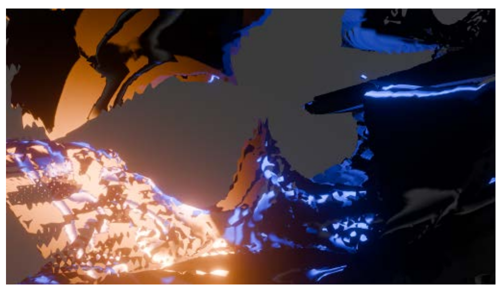
```
[CODE]
558BEC81EC300000005356578B75088B7E308B
D8D9D0DD02D8C8DD55F8DD01D8
C8DD55F0DEF9D9E8D8D9DFE080E4417E04D9E8
DEF1D9E8DEC1D9FADD5DE890DD
45F8DC45F0DD03D8C8DEF9D9E8D8D9DFE080E4
417E04D9E8DEF1D9E8DEC1D9FA
DC4FF0D9C0DC4DE8DD5DE8DD5DF090DD03DC4D
F0DD02DC4DE8DD01DC4DE8DD19
DD1ADD1BD9D08BC35F5E5B89EC5DC20800
[END]
```

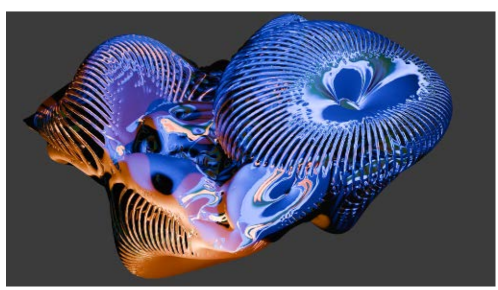
In addition, some creatures have been genetically evolved by providing genome creations as input to the Framsticks artificial simulation engine.(27)
Here is an example of creature code produced by the Beings: (Creature Name : Ypadim Ekoekaho)

```
GGllllllffffXM(, (, , , , , ,
lllllffffX,
X[|,15:1][@1:1.272](RlllX[|=:0,
3:0.931][-2:-1.664,
3:4.985]M(llllFGGgbddDX[N][-2:-0.322,
T:-5.074],, ), LfffX[1:-11.876, =:0, -
1:8.608, -1:-3.403][|8:-2.496, 1:-
2.898][@/:999, -
1:3.489]rllllGGgbddDX[@=:0.000, -1:-
7.75][|=:0](, , lllllFX[0:-0.375,
1:4.712][@0:-27.137, -1:0.916][|G:-
8.69, 2:14.084], ,
ffffMGgGbX[@1:10.332, 1:-13.801,
1:0.545][|-9:-2.985]ffffMGGgX,
RllllX(, lllllFgdDdDX, )), (,
lllllFGGgbddDX))))
```

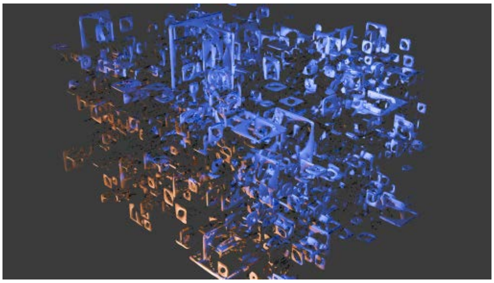
Here are some images of the first creations made by the World Beings:


World Beings Series © Lioret – 2024


World Beings Series © Lioret – 2024



World Beings Series © Lioret – 2024



World Beings Series © Lioret – 2024

World Beings Series © Lioret – 2024

References :

1. Arielli E, Manovich L. AI-aesthetics and the Anthropocentric Myth of Creativity. 2022
2. Mattia GD, Chiarella SG, Marenghi F, Focareta M, Cuono S, Torromino G. Redefining the role of Artificial Intelligence in artistic production [Internet]. 2023
3. McCormack J, Dorin A. Art, Emergence, and the Computational Sublime.
4. Huynh-The T, Pham QV, Pham XQ, Nguyen TT, Han Z, Kim DS. Artificial Intelligence for the Metaverse: A Survey [Internet]. arXiv; 2022
5. DAUDRICH A. Human vs. Machine: Algorithmic Methods in the Realm of Artistic Production. 2016
6. CHAVEZ M, LIORET A. Artificial Beings That Look Back At Us. 2006
7. Lioret A. Galatema : A Framework for Generative Cinema.
8. Lioret A, Tricaud M, Bardakos J. Fractal Artificial Life: Looking for Creativity at the Edge of Chaos. LATE-Break Abstr. 2017;5.
9. Rosenfelder M. The Planet Construction Kit. Yonagu Books; 2010.
10. Cheng I. Emissary's Guide To Worlding.
11. Butler C. Tolkien and Worldbuilding. In: Hunt P, editor. J R R Tolkien [Internet]. London: Macmillan Education UK; 2013
12. Vu R. Fantasy After Representation: D&D, Game of Thrones , and Postmodern World-Building. Extrapolation. 2017 Jul;58(2–3):273–301.
13. Bruce J, Dennis M, Edwards A, Parker-Holder J, Shi Y, Hughes E, et al. Genie: Generative Interactive Environments [Internet]. arXiv; 2024
14. Awiszus M, Schubert F, Rosenhahn B. World-GAN: a Generative Model for Minecraft Worlds [Internet]. arXiv; 2021
15. Dang H, Brudy F, Fitzmaurice G, Anderson F. WorldSmith: Iterative and Expressive Prompting for World Building with a Generative AI [Internet]. arXiv; 2023

16. Benabou M. Quarante siècles d'Oulipo. Raison Présente. 2000;134(1):71–90.

17. Berkman N. L'Oulipo et la géométrie : l'étrange utilité de la table des matières. Études Litt. 2021 Nov 25;50(2):17–34.

18. Perec G. La Vie mode d'emploi. Fayard; 2011.

19. Raymond D. « Une nouvelle manière de voir le monde ». Jalons pour une poétique de la figure fractale en littérature. Études Litt. 2021 Nov 25;50(2):35–49.

20. Ortony A, editor. Metaphor and thought. 2nd ed. Cambridge [England] ; New York, NY, USA: Cambridge University Press; 1993. 678 p.

21. Meury S. Maître ès Art, concentration création.

22. Chan BWC. Lenia-biology of artificial life. ArXiv Prepr ArXiv181205433 [Internet]. 2018

23. Chan BWC. Lenia and expanded universe. In: Artificial Life Conference Proceedings 32 [Internet]. MIT Press One Rogers Street, Cambridge, MA 02142-1209, USA journals-info …; 2020

24. Plantec E, Hamon G, Etcheverry M, Oudeyer PY, Moulin-Frier C, Chan BWC. Flow-Lenia: Towards open-ended evolution in cellular automata through mass conservation and parameter localization. In: Artificial Life Conference Proceedings 35 [Internet]. MIT Press One Rogers Street, Cambridge, MA 02142-1209, USA journals-info …; 2023

25. Blue G, Turrell de J. L'ART DE MONTRER L'INVISIBLE, L'AUTRE REVOLUTION QUANTIQUE.

26. Lioret A. Quantum Generative Art. In: Generative Art Conference [Internet]. 2021

27. Komosiński M, Ulatowski S. Framsticks: Towards a Simulation of a Nature-Like World, Creatures and Evolution. In: Floreano D, Nicoud JD, Mondada F, editors. Advances in Artificial Life

BrAIcht, a theatrical agent that speaks like Bertolt Brecht's characters

Roland Baz
LAMSADE, Paris Dauphine - PSL

Kristina Malyseva
ZHDK, Zurich Univ. of the Arts

Anna Pappa
LIASD Univ. Paris 8

Tristan Cazenave
LAMSADE, Paris Dauphine - PSL.

Abstract

This project introduces BrAIcht, an AI conversational agent that creates dialogues in the distinctive style of the famous German playwright Bertolt Brecht. BrAIcht is fine-tuned using German LeoLM, a large language model with 7 billion parameters and a modified version of the base Llama2 suitable for German language tasks. For fine-tuning, 29 plays of Bertolt Brecht and 907 of other German plays that are stylistically similar to Bertolt Brecht are used to form a more diverse dataset. Due to the limited memory capacity, a parameter-efficient fine-tuning technique called QLoRA is implemented to train the large language model. The results, based on BLEU score and perplexity, show very promising performance of BrAIcht in generating dialogues in the style of Bertolt Brecht.

1 Introduction

In the rapidly evolving landscape of technology and creativity, chatbots and conversational agents have changed the way we interact with art (Miller, 2023), music (Hennecke, 2022), dance (Berenguer, 2016) and theater (Dorsen, 2023). They have evolved from simple tools into true artistic collaborators. With their advanced language capabilities and contextual awareness, these conversational agents foster creative innovation by engaging in meaningful dialogues and contributing to artistic endeavors. In theater, conversational agents act as dynamic performers, leading dialogues and interacting with audiences during live performances, while musicians collaborate with chatbot composers to produce unique music. In the visual arts, conversational agents assist artists in brainstorming and curating exhibitions.

This study examines whether AI can reproduce Brecht's unique style in conversational agents and poses the question: Can AI help create conversational agents that reflect Brecht's style? Motivated by the idea of using artificial intelligence in plays, this research aims to create a conversational agent that enables the bot to communicate in the style of dialogues from Brecht's plays.

2 Background

Following the success of the transformer archi- tecture by (Vaswani et al., 2017), the AI commu- nity has focused on developing open- and closed- source large language models (LLMs) that com- municate effectively with users. (Radford et al., 2019) introduced GPT-2, a 1.5B parameter model trained on a large collection of web text, which excels at multiple NLP tasks without task-specific fine-tuning, including text completion and ma- chine translation. (Touvron et al., 2023) later de- veloped Llama2, a collection of LLMs with 7B to 70B parameters, optimized for dialogues, show- ing superior performance over other open-source models. Additionally, (Jiang et al., 2023) in- troduced Mistral-7B, which outperforms Llama2 (13B) and Llama1 (34B) in reasoning, math, and code generation tasks, offering faster inference and efficient sequence processing.

Pre-trained AI models can be fine-tuned for spe- cific tasks, improving performance. For example, (Yang, Tang, & Tam, 2023) developed InvestLM, based on Llama-65B, for financial topics like SEC filings and quantitative finance. It performs well in answering investment questions, matching ad- vanced models like GPT-4. Similarly, (Shoham & Rappoport, 2023) created CPLLM for clinical dis- ease prediction, outperforming models like Med- BERT in metrics such as PR-AUC and ROC-AUC. (Yu et al., 2023) used LLMs to tackle challenges in financial time series analysis, showing superior re- sults over traditional models like ARMA-GARCH and gradient boosting.

In the field of theater and the arts, there are few projects that use LLM fine-tuning to generate po- ems and plays in the style of playwrights. One example is (Bangura, Barabashova, Karnysheva, Semczuk, & Wang, 2023), which presents a GPT-2 model for automatically generating German drama texts. First, scene outlines are generated based on keywords, and then these outlines are converted into complete scenes. They use the German the- ater corpus and the German text archive. While the quantitative results are promising, the quality of the generated texts is rated as poor, likely due to the quality of the training data.

A notable contribution is MoliAIre (Cazenave, Grosjean, Rozie`re, & Pappa, 2022), a conver- sational agent that mimics the dialogue style of Molie`re's characters. The model fine-tunes GPT-2 on Molie`re's corpus to generate responses in the 17th-century French theatrical style. Key im- provements include integrating works from other playwrights for better stylistic coherence and us- ing reverse-generation for improved rhyme gener- ation. Based on previous work, our research offers a significant and unique contribution to the field.

3 Analysis and Results

Following the state-of-the-art, BrAIcht uses a generative language model that has been pre- trained on an extensive German text corpus so that it has basic knowledge of German grammar and vocabulary. The parameters of the model are then fine-tuned using a specific corpus consisting of Brecht's works and plays by other artists that have a similar style to Bertolt Brecht. This corpus con- sists of dialog excerpts formatted as follows:

• USER : line 1
• BrAIcht : line 2
• USER : line 3
• BrAIcht : line 4
.

Once trained, the model can function as a conver- sational agent as follows:

- The user enters a prompt. A context is cre- ated that is identical to the training context, but BrAIcht's prompt is left blank.

- This context is fed into the model, which then generates the response from BrAIcht and stops when a specific end-of-prompt to- ken is encountered.

- This process is repeated to generate a con- versation by accumulating the exchanges be- tween the user and BrAIcht.

3.1 Data Pre-processing and Encoding

The dataset includes 29 plays by the famous German playwright Bertolt Brecht and 907 plays by other German playwrights who have a similar style to Bertolt Brecht. From this dataset, we ex- tract the cues. We then create the prompts that are provided to the model. We also introduce certain special tokens to better convey the structure of the data. First, we insert a sentence start token la- bel <s> to signal the beginning of the sentence. Similarly, an end-of-sentence token labeled </s> is inserted to signal the end of the sentence. We also insert a pad token labeled <pad>. Finally, the keywords are converted into tokens and used as input to the model.

Our approach involves a two-stage fine-tuning process. In the first training session, we fine-tune the LLM using the German theatre dataset. We then perform a further fine-tuning of the resulting model using the Brecht dataset. We opt for this two-step approach because the Brecht dataset is relatively small. To make the dataset more diverse, we include the German plays that have stylistic similarities with Brecht's works. For the first fine- tuning step, we split the data into a training set (80%) and a validation set (20%). In the second fine-tuning step, however, we use a split of 90% for training and 10% for validation. More detailed information about the split of the data set can be found in Table 1.

	Number of cues	*Training set*	*Validation set*
German plays	*542,474*	*433,979*	*108,495*
Brecht plays	*17,740*	*15,966*	*1,774*

Table 1: BrAIcht datasets

3.2 Pre-trained model

We first try the base models of Llama2 and Mistral, but we only get average results. This is because the dataset is in German and Llama2 and Mistral are primarily pre-trained on an En- glish corpus. To solve this problem, we found a

German-focused language model called LeoLM, which builds on Llama2. LeoLM is available in two versions, 7B and 13B, both trained with a context length of 8k. It uses techniques such as linear RoPE scaling and Flash Attention 2 to im- prove training efficiency. To improve the model in German, a second stage of pre-training is per- formed with Llama2 weights and a German cor- pus with 65B tokens. This approach significantly improves the performance of the model in Ger- man compared to the Llama2 and Mistral baseline models.

3.3 Parameter-efficient Fine-tuning

The model is loaded and trained on an A6000
GPU with 48 GB VRAM. However, since fine- tuning all the parameters of the model is an ex- tremely resource-intensive process that requires considerable computing power, we decide to fine- tune the smallest LeoLM model with 7B pa- rameters. While the larger models are more promising in their results, they require signifi- cantly more memory and computing power that our single GPU cannot provide. In addition, we use parameter-efficient fine-tuning (PEFT), known as Quantized Low-rank Adaptation (QLoRA), to train the model. This approach allows us to op- timize the performance of the model and adapt it to our task while reducing the high memory re- quirements and training time associated with fine- tuning. The main idea behind the QLoRA tech- nique developed by (Dettmers, Pagnoni, Holtz- man, & Zettlemoyer, 2023) is that it uses a novel high-precision technique to quantize a pre-trained model to 4-bit (Frantar, Ashkboos, Hoefler, & Al- istarh, 2023), (Frantar & Alistarh, 2022), and then adds a small set of learnable low-rank adapter weights that are tuned by backpropagating gradi- ents through the quantized weights. Consequently, this technique reduces the size of the model so that it fits into the GPU while maintaining performance comparable to full fine-tuning.

3.4 Training

To train the base LeoLM-7B, we first obtain its weights from Hugging Face, a widely used plat- form where the AI community shares their open- source models and contributes to the development and progress of AI research (Wolf et al., 2020).

3.4.1 Objective Function and Optimization

The learning objective is to reconstruct the di- alog extracts by minimizing the logarithm of per- plexity, which is a common metric used to evaluate language models, where lower values indicate bet- ter performances in predicting the next word in a sequence. Each dialog extract is divided into a se- quence of tokens U = (u1, . . . , un) and the model parameters Θ are optimized by minimizing the inverse of the log-likelihood function (negative log-likelihood):

$$L(U, \Theta) = -\sum_{i=1}^{n} \log P(u_i \mid u_1, \ldots, u_{i-1}, \Theta) \quad (1)$$

To minimize the objective loss function $L(U, \Theta)$ from equation 1, we use an optimization technique called AdamW (Zhuang, Liu, Cutkosky, & Orabona, 2022), (Zhou, Xie, & YAN, 2023) a variant of Adam, an optimization technique commonly used for training machine learning and deep learning models.

3.4.2 Inference

To promote the diversity of the model's responses and improve its improvisational ability, we use a stochastic generation strategy. Specifically, we use the top-k sampling method for text generation in combination with temperature, which controls the randomness of predictions. Based on the context prompt, the model generates a probability distribution over all possible next tokens. We then select the next token from the 50 most probable options (k = 50 and temperature = 0.7).

3.4.3 Evaluation

The results show a significant reduction in perplexity, reaching a value of 3.57 on the validation dataset after the second stage of fine-tuning. The results are shown in Table 2.

	German plays	*Brecht plays*
Perplexity	*10.19*	*3.57*

Table 2: Perplexity

We proceed with the calculation of the BLEU score, keeping in mind that this metric can be influenced by various parameters such as top-p, top-k, temperature, and the number of generated candidates (n). While BLEU is traditionally used for evaluating translation tasks, it has limitations in evaluating dialogue models, particularly because it focuses on n-gram overlap rather than creative or stylistic coherence. Nevertheless, it provides a useful point of comparison for assessing improvements in model performance and allows for a standardized evaluation metric across different models. To simplify our approach, we generate only one candidate per reference and set top-k to 50. In our case, we compare the model tuned to German plays with the model tuned to Brecht's plays. This comparison is performed using the Brecht validation set to assess how well each model generates Brecht's plays. We take samples of size 100, 300, 500, and 1000 from the dataset and calculate the BLEU score three times for each sample, taking the average of the results. The final scores are shown in Table 3. To further illustrate this, we

	German plays	*Brecht plays*
n=100	*0.46*	*0.70*
n=300	*0.22*	*0.65*
n=500	*0.18*	*0.57*
n=1000	*0.17*	*0.74*
BLEU (avg)	*0.26*	*0.665*

Table 3: BLEU score

perform three BLEU score trials for each sample size and then average the three results. We repeat this process for all sample sizes. At the end, we calculate the average of all the BLEU scores obtained to get the final score, which clearly shows a significant improvement from 0.26 to 0.665.

4 Conclusion & Recommendations

The aim of this research project is to develop a conversational agent that is able to generate plays in the style of Bertolt Brecht using large language models. For this purpose, the LeoLM-7B, a Ger- man variant of the Llama2, was selected due to its strong performance on German language tasks. The fine-tuning process consisted of two phases: First, the model was trained on a dataset of Ger- man plays and then retrained exclusively on plays by Brecht. This training was carried out with an A6000 GPU and 48 GB RAM. The preliminary results are promising and show that the model can generate plays that reflect Brecht's unique style. The model achieved a BLEU score of 0.67 and a perplexity of 3.57, indicating solid performance. However, a major limitation in this project was the

availability of GPU resources, which affected the model selection. Future work could improve the accuracy of the model through fine-tuning larger models and the use of retrieval-augmented gener- ation techniques.

References

Bangura, M., Barabashova, K., Karnysheva, A., Semczuk, S., & Wang, Y. (2023). Automatic generation of german drama texts using fine tuned gpt-2 models.

Berenguer, R. (2016). HOMEOSTA- SISV2. Retrieved from https:// rocioberenguer.com/index.php ?lang=en&project=3

Cazenave, T., Grosjean, G., Rozie`re, B., & Pappa, A. (2022). Moliaire, a theatrical agent which speaks like Molie`re's characters. In XXV generative art conference (pp. 286–294).

Dettmers, T., Pagnoni, A., Holtzman, A., & Zettle- moyer, L. (2023). Qlora: Efficient finetun- ing of quantized llms.

Dorsen, A. (2023). Prometheus Fire- bringer. Retrieved from https:// anniedorsen.com/projects/ prometheus-firebringer/

Frantar, E., & Alistarh, D. (2022). Sparsegpt: Massive language models can be accurately pruned in one-shot. In Advances in neural information processing systems (Vol. 35, pp. 4475–4488).

Frantar, E., Ashkboos, S., Hoefler, T., & Alis- tarh, D. (2023). Gptq: Accurate post- training quantization for generative pre- trained transformers.

Hennecke, M. (2022). The (Un)answered Question: A concert written by emo- tions. Retrieved from https:// theater.digital/en/aktuelles/ news/the-unanswered-question -ein-konzert-geschrieben-von -emotionen/

Jiang, A. Q., Sablayrolles, A., Mensch, A., Bam- ford, C., Chaplot, D. S., de las Casas, D., . . . Sayed, W. E. (2023). Mistral 7b.

Miller, B. (2023, March 21–April 22). gallery 976 Madison Avenue, New York. Retrieved from https://gagosian.com/exhibitions/2023/bennett-miller/

Radford, A., Wu, J., Child, R., Luan, D., Amodei, D., Sutskever, I., et al. (2019). Language models are unsupervised multitask learners. OpenAI blog, 1(8), 9.

Shoham, O. B., & Rappoport, N. (2023). Cpllm: Clinical prediction with large language models.

Touvron, H., Martin, L., Stone, K., Albert, P., Almahairi, A., Babaei, Y., . . . Scialom, T. (2023). Llama 2: Open foundation and fine- tuned chat models.

Vaswani, A., Shazeer, N., Parmar, N., Uszkor- eit, J., Jones, L., Gomez, A. N., . . . Polo- sukhin, I. (2017). Attention is all you need. Advances in neural information processing systems, 30.

Wolf, T., Debut, L., Sanh, V., Chaumond, J., De- langue, C., Moi, A., . . . Rush, A. (2020, Oc- tober). Transformers: State-of-the-art natu- ral language processing. In Empirical meth- ods in natural language processing: System demonstrations (pp. 38–45). ACL.

Yang, Y., Tang, Y., & Tam, K. Y. (2023). Investlm: A large language model for investment us- ing financial domain instruction tuning.

Yu, X., Chen, Z., Ling, Y., Dong, S., Liu, Z., & Lu, Y. (2023). Temporal data meets llm – explainable financial time series forecast- ing.

Zhou, P., Xie, X., & YAN, S. (2023). Towards understanding convergence and generaliza- tion of adamw.

Zhuang, Z., Liu, M., Cutkosky, A., & Orabona, F. (2022). Understanding adamw through proximal methods and scale-freeness.

Programmable Architecture, Urbanism, and Territory: Spatial Intelligence and Data-Driven Forms of Strategic Urban Scenarios

Dr Dragana Ćirić, PhD.
Independent Scientific Associate/Research Assistant Professor;
Founder, Licenced Architect, and PI at unit [d], Belgrade, Serbia
e-mail: unit.d.dciric@gmail.com

Abstract

Programmability of both existing and newly devised spatial structures, and transition to their intelligent forms, construct the main framework of the study and demonstrated design outputs. As a design mode that relies on complete computational control of a large class of processes and information related to investigated spatial situations and main research or design subjects, programmability enables fully monitored and rationalised creative thinking and design process, their representation and clear overview through propagation workflows (visual graph-based guidance through them), control of each design element and operation, as well as informed maintenance and development of specific design solutions. The expansion of graph-based visual computation methods to a greater number of software made the programmability method accessible to a wider range of users. Without the need to necessarily master all programming languages, while only relying on strong design knowledge and logic, users were given the power of understanding, creating, and defining the chains of computational operations that lead to the planned results (the power of reasoning with/through graphs as structured ways of representing data) and of expressing the generative logic of their ideas and designs by computationally and algorithmically managing generative processes. More advanced users have been given transformative power to innovate defined software frameworks by adding their own lines of code and plug-ins. All designers have been greatly empowered to introduce new definitions and paths to solutions to the defined problems, responses to design objectives, and new computational methods expressed through original workflows. Automated workflows and generators indicate additional design and analytical possibilities considering almost all design definitions, not only explorative design, versioning, and variation that investigate and bring about a large number of probable outcomes or enable optimisation of different design aspects.

With that in mind, several exercises and experiments are proposed while targeting the application of the mentioned theoretical framework to cultural heritage. Derived from the previous research on network science and graph-based design methods, the application of constructed workflows suggests their scalability for urban and architectural design, analytics, and planning in various situations. Even though exemplary computational workflows have been developed on one specific example and city layout (Greater Paris), their topological nature enables the underlying methods to be used in any context just by changing initial city network inputs and data. By changing the geometry of the examined urban territory and built environment, along with the information required for planned operations, the workflows can be easily adjusted to new spatial, urban, and architectural contexts. On the other side, with regards to this year's thematic focus, inclusion of specific classes of architectural and urban entities (e.g. cultural heritage) as parameters that guide urban scenarios construction and decision-making within the generative processes, assigns different themes and quality to thus far developed network operations workflows and diversifies both probability space in terms of possible outcomes as well as propagation graphs design and logic leading them to more complex configurations. Based on that, the presented experiments having the city of Paris and its transportation network as a subject, can work as an exemplary, opening up new directions for project continuation by enlarging the number of forms of urban scenarios based on cultural heritage data and intelligence systems that drive them, as well as the number of city case studies, to which the methods can be applied. The logic that workflows contain is presumed to be applicable at various scales—from the level of the city and its administrative area or natural area (territory) to the level of individual architectural objects, their specific constitutive elements and details.

Keywords: Spatial Intelligence, Urban Intelligence, Intelligent City, Urban Data Cloud, Urban Scenarios, Programmable Architecture, Programmable Urbanism, Programmable Territory, Programmable Heritage, Architectural and Urban Design, Computational Design, Computation Workflows, Reasoning Structures, Graph-Based Reasoning, Urban Mobility

1. Introduction

Due to specific framework of the presented research — constructive, problem-based, computational, empirical, and experimental — the lecture will primarily rely on output materials that unambiguously demonstrate stated claims and thus far accomplished results. The materials contain proof of a concept and solutions to the defined problems, which are central to studies with practice- and performance-based objectives. Therefore, the dynamic outputs will constitute the central part of the presentation.

The study explains in parallel the definition of the design outputs that are aimed at and generated, and design of the design processes and workflows (algorithms) whose propagation fulfils the defined design and representation objectives. The design content is based on network structures and geometries, operated and animated to present various problem-solving paths and scenarios for spatial, architectural, or territorial analyses, formal transformations, and generation of new geometries and designs, the latter being particularly in focus for further project development.

Leaving the explanation of theoretical grounds and references to other sources, some of which have been reviewed in previous and preliminary studies [1-4] with support of the relevant literature [5-18], the lecture will be focused on original contribution contained in presented materials, as well as the resulting parametric, algorithmic, and generative performance. Both completely and less responsive outputs regarding the defined tasks will be commented and explained.

2. Data- and Intelligence-Driven Urban Scenarios

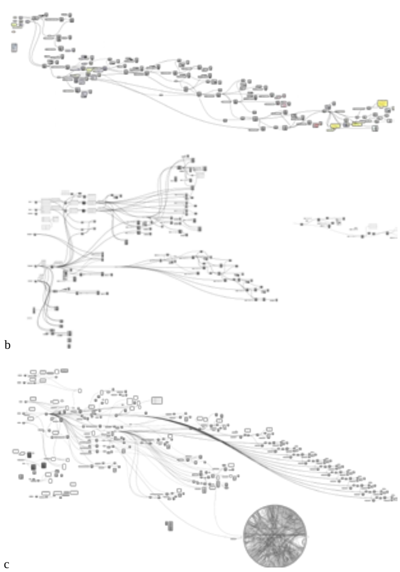


Fig. 1. Process propagation graphs for network and path-generation and design, including graph parts for dynamic localisation, search for possible moves within the network according to given parameters (such as distance, or a specific class of destination points), tree-like representation of decision-making process during the path construction and transition from point to point.
Graph a. formal outputs: Fig. 6. Graphs b. and c. formal outputs: Fig. 2, 3, 4, and 5. Source: © Dragana Ciric, 2023-2024, all rights reserved. Software: Grasshopper, Rhinoceros

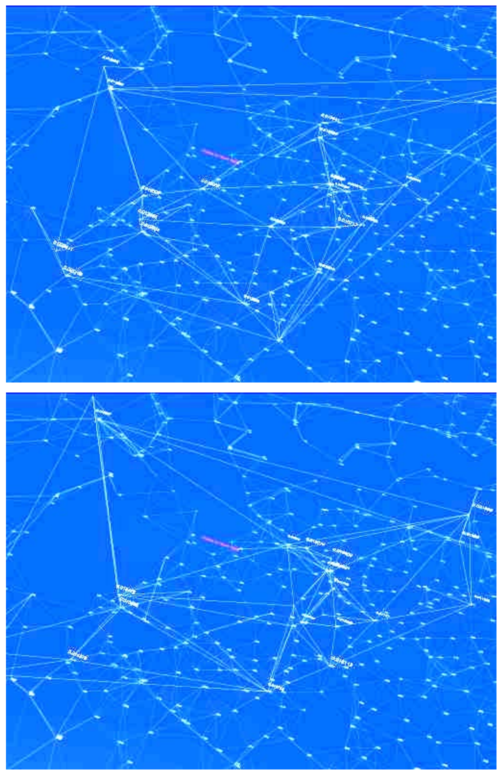


Fig. 2 Distance analysis from multiple points in a network, displaying thereby relationships between analysis results. https://va.media.tumblr.com/tumblr_sjhyjy04iM1a96rc8_720.mp4 Software: Grasshopper. Source: © Dragana Ciric, 2024

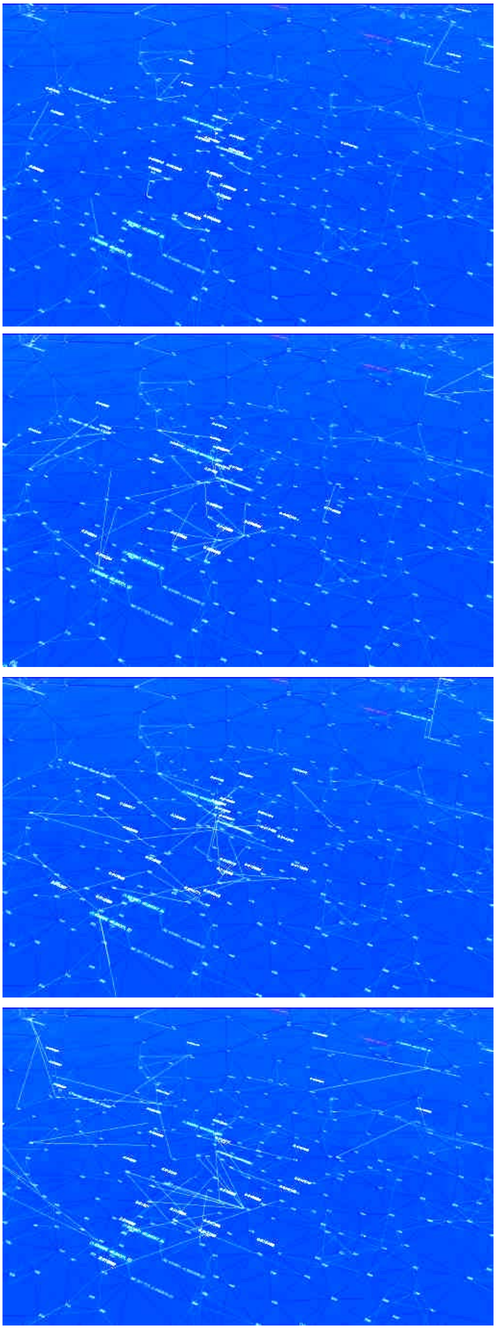


Fig. 3. Distance analysis from multiple points in a network, https://va.media.tumblr.com/tumblr_sjabhiH4F31a96rc8_720.mp4 Software: Grasshopper. Source: © D. Ciric, 2024

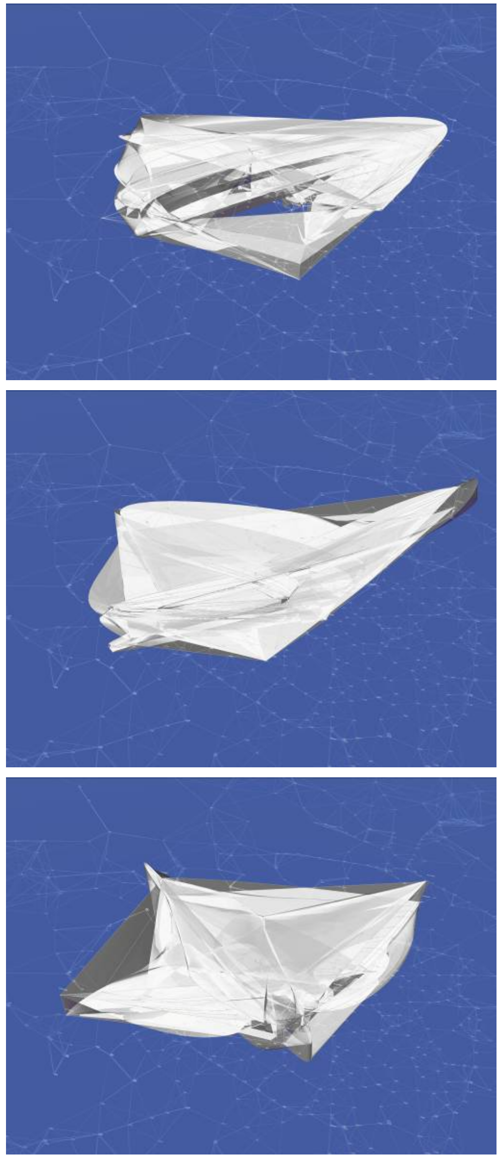
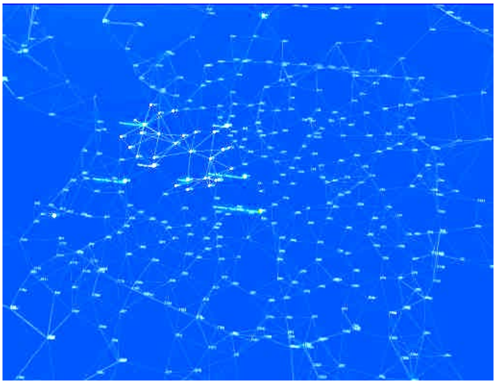
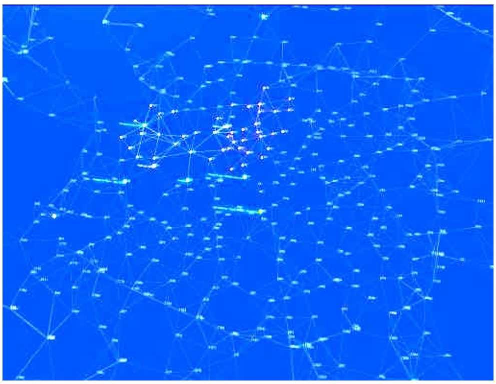
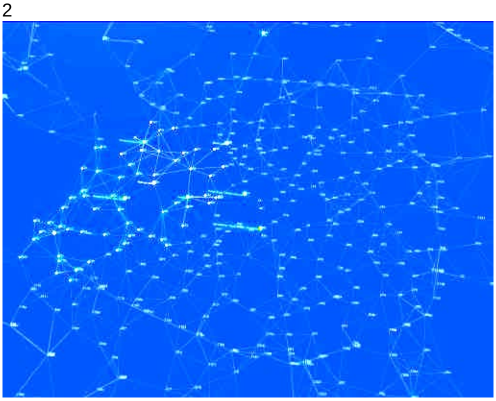
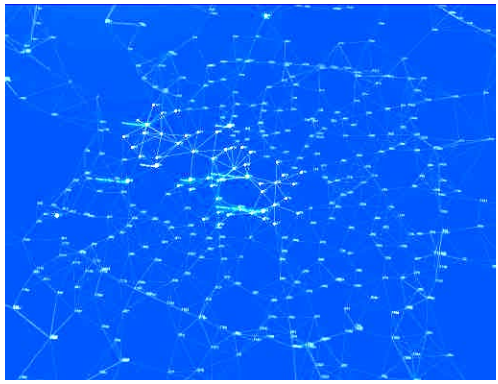


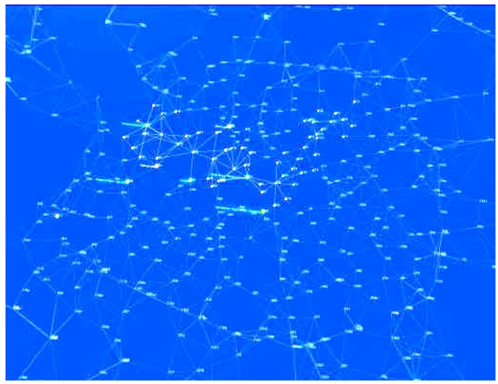
Fig. 4. Formalisation of analysed data: formal transformations of the data cloud with respect to change of the chosen parameter values (the dynamic network from Fig. 2 and Fig. 3. represents the constitutive structure of the generated form). Software: Grasshopper. Source: © Dragana Ciric, 2024


1


2


3


4


5

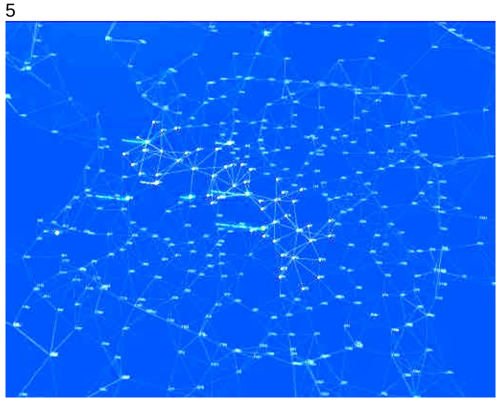

6

Fig. 5. Five path-generation scenarios that share the four transition steps represented in the image 1, corresponding to nodes of the M1 metro line with Charles de Gaulle – Étoile as a starting point. Images 2, 3, and common part of the paths in images 4, 5, and 6 display three path branches that appear with step 5 of the path-generation process (in node 4 - Concorde), while images 4, 5, and 6 represent subsequent sub-branches of the M1 line path continuation, diverging based on the path-generation process decisions made with step 9 (in node 8 – Châtelet – Châtelet-Les Halles). Scenario 1: https://va.media.tumblr.com/tumblr_sk8ldhPyuG1a96rc8_r1_720.mp4 . Scenario 5: https://va.media.tumblr.com/tumblr_sk8rdaiy9v1a96rc8_720.mp4 Software: Grasshopper. Source: © Dragana Ciric, 2024

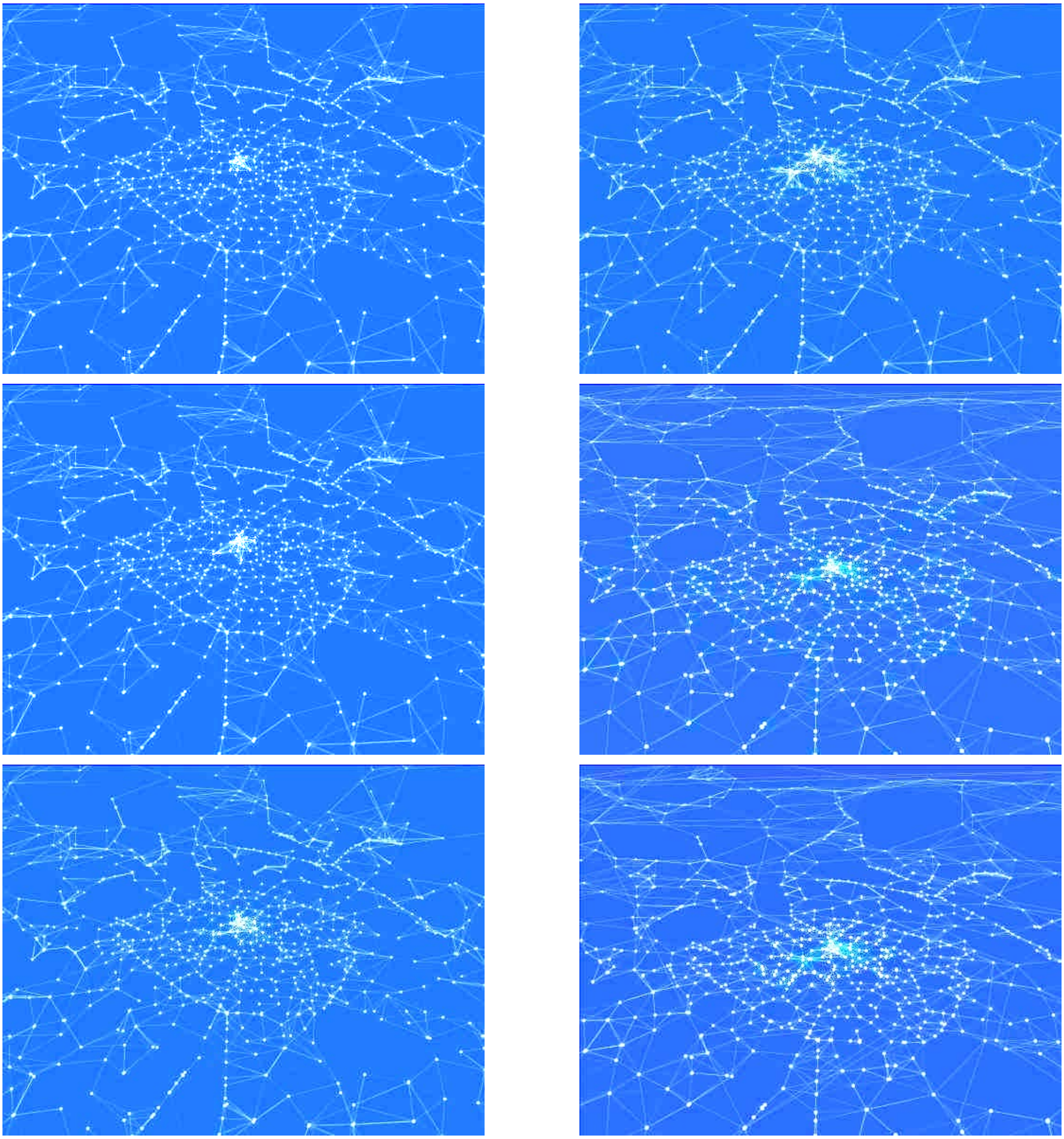


Fig. 6. Network growth. Steps include the following: selection of the starting point → propagation of the search function from the chosen location while rendering states that can possibly be occupied with regards to established connections → network growth from the given point by acting along all possible increments and occupying all suggested states → repetition of the search function from all newly occupied states in the previous step (display of all states that can possibly be occupied (nodes or points)) and connectivities (connection lines) with dynamic preview of the change of connectivities between the considered points and possible moves within the covered area → network growth by occupying all suggested states → continuation of the same iteration from all newly occupied states. https://va.media.tumblr.com/tumblr_sbaeuiOF4C1a96rc8_720.mp4. Software: Grasshopper. Source: © Dragana Ciric, 2024

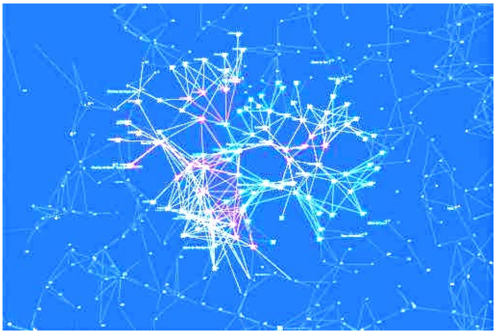


Fig. 7. Tree search and analysis structure with the constructed path that follows the geometry and logic of the defined branching within the network. Construction of the single path: https://va.media.tumblr.com/tumblr_sl084gZVxR1a96rc8_720.mp4. Construction of the path with branches: https://va.media.tumblr.com/tumblr_sm906qKZyq1a96rc8_720.mp4. Software: Grasshopper. Source: © Dragana Ciric, 2024

3. Conclusion – Future Plans and Prospects

The aims of the project that addressed design of operable software workflows and set of computational design methods that will resolve defined problems of network operations have been accomplished. The algorithmic logic, its elements, or chains of operations and problem-solving methods have been completely defined with clarity and author's awareness of the possibilities offered by the used software. The systematisation of exercises has also been achieved and each phase or subtask possesses proper explanation. However, the necessary refinements within the workflows might refer primarily to the aspects of representation and graphic solution, especially those addressing the parallel, simultaneous propagation of larger number of processes within the same graph function, implying also the need of better correspondence to interactive human-machine interfaces which might likewise add certain smaller parts and changes to the thus far established propagation graphs. Having in mind probable further application, mentioned improvements should make the solutions technically more robust and user-oriented while enhancing their performance scope.

Regarding presumed application, urban scenarios should also have a separate analytical and definition framework so as to enable one to include larger number of decision-making parameters and their transposition to graphs, while propagation graphs' applicability to different geometries should be tested in some of the next project stages. Formalisation of dynamic data analysis leads directly to new generative design possibilities, and this aspect will probably impact most of the upcoming studies. City Data Clouds as one of the forms that have emerged as a result of the presented programmable solutions, while remapping the existing city structures or individual object, open new generative paths for their active transformation, including those that are based on and operate with cultural heritage as the central reference area. The examples showcased the influence of landmark and cultural heritage locations on path generation and decision-making while constructing urban movement scenarios in real-time, and in that respect, a number of more advanced applications are expected.

In conclusion, it is important to put emphasis on the key background process of the presented tests—processes in which raw data and information transition to *intelligence*, an operative knowledge and support to learning, urban awareness, decision-making and problem-solving that guide agents' experience, performance, possibilities and abilities in space.

4. References

[1] Ćirić, D. (2023a). The Science of Networks: Urban Movement Design, Analytics, and Navigation. In: Bogdanović, R. (Ed.). *On Architecture – Challenges in Design, Proceedings*, pp. 110-129. STRAND – Sustainable Urban Society Association, Belgrade.

[2] Ćirić, D. (2023b). The Science of Networks: Network Graphs in Urban Navigation Design Problems—Mobility, Transportation Systems, and Movement Paths Generation [Revised e-Book Version]. In: Soddu, C. and Colabella, E. (Eds). *XXVI Generative Art 2023—Proceedings of the XXVI GA Conference*, pp. 105-120. Domus Argenia Publishing, Rome.

[3] Ćirić, D. (2023-2024). Network operations, https://dciricnetworks.tumblr.com/, last accessed 2024/09/16.

[4] Ćirić, D. (2024). Generative and AI Methods for Transportation Systems and Urban Mobility Design, Planning, Operation, and Analysis: Contribution to Theory and Methodology of Urban Computing. In: Đukić, A., Krstić-Furundžić, A., Vaništa Lazarević, E., and Vukmirović, M. (Eds.). *Keeping up with Technologies to Imagine and Build Together Sustainable, Inclusive, and Beautiful Cities, 8th International Academic Conference on Places and Technologies Proceedings*, Vol.8, pp. 490–503. Belgrade University - Faculty of Architecture, Belgrade.

[5] Barabási, A-L. (with M. Pósfai). (2018). *Network science*. Cambridge University Press, Cambridge, UK.

[6] MacLennan, B. J. (2005). The Morphogenetic Path to Programmable Matter. In: Lyke, J., Christodoulou, C. G., Vera, A., Edwards, A. H. (eds.), *Proceedings of the IEEE. Reconfigurable Systems: Advanced Applications and Technologies (Special Issue)*, vol. 103, no. 7, pp. 1226–1232. https://doi.org/10.1109/JPROC.2015.2425394

[7] Woodbury, R. (with contributions by O. Y. Gün, B. Peters and M. R. Sheikholeslami). (2010). *Elements of Parametric Design*. Routledge, London and New York

[8] Jabi, W. (2013). *Parametric Design for Architecture*. Laurence King Publishing Ltd., London.

[9] Tedeschi, A. (2014). *AAD_Algorithms-Aided Design: Parametric Strategies Using Grasshopper*. Le Penseur Publisher, Brienza.

[10] Madl. A. (2022). *Parametric Design for Landscape Architects: Computational Techniques and Workflows*. Routledge, London and New York.

[11] Kitchin, R., Dodge, M. (2011). *Code/Space.* The MIT Press, Cambridge, Mass.

[12] Zheng, Y. (2018). *Urban Computing*. The MIT Press, Cambridge, Mass.

[13] Figueiredo, S. M., Krishnamurthy, S., Schroeder, T. (Eds). (2020).

Architecture and the Smart City. Routledge, London and New York.

[14] Shi, W., Goodchild, M. F., Batty, M., Kwan, M-P., Zhang, A. (Eds.). (2021). *Urban Informatics* [Online]. Springer Nature and The International Society for Urban Informatics, Singapore. https://doi.org/10.1007/978-981-15-8983-6

[15] Picon, A., Ratti, C. (2023). *Atlas of the Senseable City*. Yale University Press, New Haven and London

[16] Batty, M. (2024). *The Computable City: Histories, Technologies, Stories, Predictions*. The MIT Press, Cambridge, Mass.

[17] Ćirić, D. (2022). Exo Case-Study and Prototype of the Dynamic Responsive Architectural System: The Key Attributes of Systemic Integration and Design - Part I. *Civil Engineering and Architecture* 10(5), 1701-1717. https://doi.org/10.13189/cea.2022.1005011701-1717

[18] Ćirić, D. (2022). Exo Case-Study and Prototype of the Dynamic Responsive Architectural System: The Key Attributes of Systemic Integration and Design - Part II. *Civil Engineering and Architecture* 10(6), 2627 – 2643. https://doi.org/10.13189/cea.2022.100630

Illuminating Spaces: Deep Reinforcement Learning and Laser-Wall Partitioning for Architectural Layout Generation

Reza Kakooee, PhD
Institute of Technology in Architecture, Department of Architecture, ETH Zurich, Switzerland
https://www.linkedin.com/in/reza-kakooee/
e-mail: kakooeer@ethz.ch

Prof. Dr. Benjamin Dillenburger, PhD
Institute of Technology in Architecture, Department of Architecture, ETH Zurich, Switzerland

Abstract

Space layout design (SLD), occurring in the early stages of the design process, nonetheless profoundly influences both the functionality and aesthetics of the ultimate architectural outcome. The complexity of SLD necessitates innovative approaches to efficiently explore vast solution spaces. While image-based generative AI has emerged as a potential solution for SLD, current methods often rely on pixel-based space composition methods that lack intuitive representation of architectural processes.

In this paper, we leverage deep Reinforcement Learning (RL) for SLD, as it offers a procedural approach that intuitively mimics the process of human designers. Moreover, the duality of agent and environment in RL mirrors the interplay between form and context in architectural design. However, effectively using RL for SLD requires an explorative space composing method to generate desirable design solutions.

We introduce "laser-wall", a novel space partitioning method that conceptualizes walls as emitters of imaginary light beams to partition spaces. This approach bridges vector-based and pixel-based partitioning methods, offering both flexibility and exploratory power in generating diverse layouts. We present two planning strategies: one-shot planning, which generates entire layouts in a single pass, and dynamic planning, which allows for adaptive refinement by continuously transforming laser-walls within the plan. Additionally, we introduce on-light and off-light wall transformations for smooth and fast layout refinement, as well as identity-less and identity-full walls for versatile room assignment in generated layouts.

We developed SpaceLayoutGym (SLG), an open-source OpenAI Gym compatible simulator for generating and evaluating

space layouts. Our deep RL algorithm interacts with the SLG to process the input design scenarios and generates solutions following a reward function that balances geometrical and topological requirements and potentially can incorporate aesthetics. Our results demonstrate that the RL-based laser-wall approach can generate diverse and functional space layouts that satisfy both geometric constraints and topological requirements. The proposed RL-based laser wall partitioning enables the generation of creative space layouts while being architecturally intuitive.

1. Introduction

SLD represents a complex optimization problem that permeates various domains of human endeavor. At its core, SLD involves the strategic arrangement of elements within a defined spatial boundary to achieve specific objectives from maximizing usability to ensuring aesthetic harmony. This process necessitates a delicate balance between functional requirements including geometrical constraints, and topological relationships, often leading to the exploration of vast solution spaces in search of optimal configurations [1].

SLD challenges span diverse applications, forming the foundation of efficient building design in architecture and influencing residential to complex commercial structures [2]. Urban planners apply SLD principles to shape city layouts, impacting urban livability and sustainability [3], while interior designers use SLD to optimize furniture arrangement for functionality and user experience [4]. In manufacturing, SLD enhances factory layouts, boosting productivity and safety [5], and in technology, it is crucial for component arrangement in integrated circuits, balancing performance and heat dissipation [6]. Game designers also employ SLD to craft engaging virtual environments [7]. While SLD has diverse applications, this research focuses on architecture due to its impact on daily life, well-being, and societal structures, nonetheless, the proposed methodology can be extended to other domains.

Despite its widespread application, achieving optimal SLD solutions remains a challenge due to the need for advanced optimization techniques [8]. The complexity of SLD stems from its inherently combinatorial nature and the need to satisfy multiple, often conflicting objectives simultaneously. Automated layout design has been addressed through various computational methods. Traditional methods such as shape grammars [9], use rule-based systems to generate designs, while graph-based methods represent spatial relationships for topological analysis [10, 11, 12, 13]. Evolutionary algorithms have laid the groundwork for systematic exploration of design possibilities while adhering to architectural principles and constraints [1, 14, 15, 16]. However, these approaches often fall short in capturing the intricate nuances of architectural design.

Machine learning methods have also contributed to automating and optimizing SLD [17, 18, 19, 20]. House-GAN++ [21] have demonstrated the ability to generate house layouts in pixel-level. HouseDiffusion [22] employs denoising processes to enable precise control over vector-based floor plans, inspired by advancements in diffusion techniques [23]. Researchers have also explored the potential of RL in various aspects of architecture. [24] developed an approach that combines RL with multi-agent systems to evolve architectural forms, addressing the limitations of conventional parametric design methods. [25] used RL algorithms to optimize building layouts, while [26] applied RL to adjust 3D room configurations. [27] created a self-play RL framework to optimize spatial structure assembly. [28] and [29] integrated RL with shape grammars for housing layouts, while [30] applied RL to generate simple floor plans. RL has also been used to optimize factory layouts [31] and furniture arrangements [32]. Recent work by [33] highlights the potential of multi-agent RL in complex spatial synthesis across various scales.

However, despite these advancements, many AI methods still rely on pixel-based methods that lack intuitive representation, are limited in addressing design requirements, or often fail to capture the nuanced interplay between form and context in architectural design. This limitation emphasizes the need for more procedurally driven methods that are flexible, efficient, and intuitive. This paper introduces a novel space composition method called "laser-wall" partitioning, which conceptualizes walls as emitters of imaginary light to divide spaces. This partitioning method allows us to leverage RL for architectural SLD, providing a procedural method that mirrors human design processes while maintaining flexibility in exploring spatial configurations.

2. Methodology

Architectural SLD involves a synergy of space composition methods and computational approaches to optimize the composition. In this paper, we focus on partitioning as the space composition method and RL as a computational approach to optimize layout composition given a design scenario.

In the following section, we introduce a novel space partitioning method called "laser-wall". This innovative method enhances the effectiveness of subsequent computational approaches, such as RL, in exploring the solution space to identify optimal layouts. Moreover, the laser-wall method transforms the layout design problem into a game-like environment, making it particularly amenable to RL algorithms, leveraging their proven efficacy in finding optimal policies. We then use RL to find optimal solutions for the optimal space layout using our proposed composition methods.

2.1 Laser-wall partitioning

Space partitioning can be broadly categorized into two main methods: vector-based and pixel-based. Rectangular dissection, a common vector method, divides space using straight lines but often results in long, inflexible walls, limiting irregular shapes. Pixel-based partitioning, in contrast, allows for more flexible and irregular designs by dividing space at a granular level but is slower and computationally intensive. These contrasting approaches highlight the need for a more balanced method, leading to the development of our laser-wall partitioning technique, which aims to combine the strengths of both.

Basics of laser-wall method:

A laser-wall consists of two components: a hard part (the base wall) and a soft part (light beams). The base wall is formed by two connected segments, either aligned or perpendicular, establishing the physical structure. From each end of these segments, light beams are emitted. These beams travel through the space, defining partitions and interacting with other elements in the layout. When a base wall is placed within the plan, its light beams automatically emit and propagate until they encounter obstacles; either the outline of the plan, other base walls, or other light beams. A key concept governing beam interactions is the infiltration rate, which introduces flexibility to the partitioning process. We explore two implementations: fixed and decreasing infiltration rate. In the former, all walls have the same infiltration rate; beams stop upon encountering other beams. In the latter, the infiltration rate decreases with distance from the beam's source, allowing beams closer to their origin to cut through beams from other walls if they have a lower infiltration rate at the intersection point.

2.2 Planning Approaches

Building on the laser-wall concept, we introduce two planning approaches: One-Shot Planning, and Dynamic Planning.

In one-shot planning, the entire layout is determined in a single iteration. Starting with an empty plan, base laser-walls are selected from a predefined wall library and placed at chosen coordinates. Upon placement, their light beams are emitted to partition the space. This process continues until the required number of rooms is achieved. Once placed, walls remain fixed, completing the layout in one iteration. Two mechanisms are proposed for assigning rooms to subregions. In the identity-less mechanism, walls are not linked to specific rooms during partitioning. After partitioning, the subregion connected to the entrance is designated as the living room, and other rooms are assigned by matching each subregion to the room with the closest desired area defined in design scenario. In contrast, the identity-full mechanism assigns each wall a room index before placement. The smaller subregion created by the wall is assigned to its associated room, with constraints ensuring walls are not placed in already occupied regions or the entrance-connected area reserved for the living room. This process is exemplified in Figure 1.

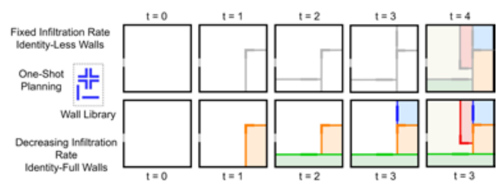


Fig 1: Comparison of one-shot planning approaches using fixed infiltration rate with identity-less walls (top row) and decreasing infiltration rate with identity-full walls (bottom row)

Dynamic planning introduces adaptability by allowing walls to transform within the plan, enabling exploration of a wider range of spatial configurations. Walls can move and rotate: through incremental movements, walls shift in 8 cardinal directions (north, northeast, east, etc.) to refine their positions, while 90-degree rotations allow entire base walls or individual segments to turn clockwise or counterclockwise, altering the shape and proportions of rooms. This discrete approach aligns with a cell-based representation of the plan, where space is divided into a grid of units.

Wall movement requires determining how lights will behave during transformations. Two types of wall transformations are introduced: off-light and on-light transformations. In the former, the light of the moving wall is deactivated before its movement. After the wall is repositioned, all walls' lights are activated in a specific order, with the moved wall's light being activated last. In the latter, the light of the moving wall stays active throughout its movement. This process is practically similar to initially turning off all lights and then *incrementally* activating them in order, with the moved wall's light being turned on last. After each transformation and following partitioning, first, connected subregions to each base-wall are identified, next, Intersection over Union (IoU) score between new subregions and their previous counterparts are calculated, and then rooms are assigned based on the following steps: a) the region connected to the entrance is assigned as the living room, b) subregions with an IoU of 1 (perfect overlap) are assigned to maintain continuity, c) remaining regions are assigned by iterating through walls in a specific order (with the moved wall considered last), assigning regions with the highest IoU to corresponding walls. This process is exemplified in Figure 2.

According to the laser-wall partitioning method, we developed SLG, an OpenAI Gym compatible simulator. It supports both one-shot and dynamic planning approaches, allowing exploration of various SLD strategies. SLG provides a standardized environment for creating design scenarios, evaluating layouts, and optimizing solutions based on multiple criteria.

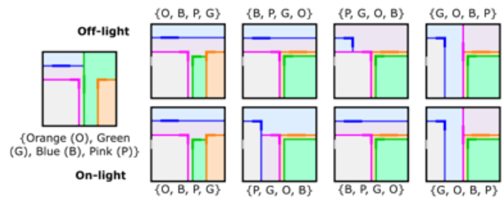


Fig 2: Illustration of off-light and on-light transformations in laser-wall partitioning

In this paper, we focus exclusively on dynamic planning, building upon our previous work on one-shot planning [34]. Our investigation examines dynamic planning using on-light and off-light transformations to compare the resulting layouts. The goal is not to identify a "best" method, but to understand how these variations impact agent performance and layout characteristics, offering insights into the flexibility of our RL-based approach to SLD. We also analyze how different wall types—angled, straight, or both—affect the RL agent's performance across different scenarios.

2.3 RL for SLD

RL provides a framework for an agent to learn optimal decision-making strategies through interaction with an environment, making it well-suited to the iterative process of layout refinement in dynamic planning. To apply RL to our SDL with the laser-wall partitioning method, we frame the SLD problem as a Markov Decision Process (MDP).

An MDP provides a framework for modeling decision-making in scenarios where outcomes are partly under the control of a decision-maker. The MDP is defined by there components, state space, action space, and reward function.

State Space: The state space in RL represents all possible configurations or situations the agent can encounter in the environment. In SLG, the state at any given time is represented by an RGB image of the current layout. This image encapsulates both geometric and topological information, including room configurations, wall positions, and the entrance location. The use of an image state allows the agent to perceive spatial relationships intuitively, mirroring human visual processing of layouts.

Action Space: The action space in RL defines the set of all possible actions an agent can take to interact with or modify the environment. In dynamic planning, actions are consisting of two decisions: a) selecting an existing wall to transform, b) choosing a specific transformation (8 directional movements and 6 rotation options).

Reward Function: The reward function in RL quantifies the desirability of actions and resulting states, guiding the agent towards optimal behavior. In dynamic planning, each action results in a complete new configuration, allowing for assessment of both geometric and topological properties at every step. The reward consists of two components: instant reward, and terminal reward. The reward is calculated from two elements: an instant reward and a terminal reward.

Instant rewards: a) Negative rewards for violating hard constraints (e.g., wall collisions, entrance blockage). b) Negative rewards for intermediate layouts that deviate significantly from desired properties.

Terminal rewards: a) Negative rewards if the layout is significantly different from the desired properties. b) Non-negative reward for solutions that fall within a threshold of closeness to desired properties, scaled linearly or nonlinearly based on the closeness. c) Bonus reward for perfect adjacency matching.

Learning Algorithm:
We employ Proximal Policy Optimization (PPO) [35] algorithm, to learn an optimal policy for layout generation. This choice is motivated by PPO's ability to handle high-dimensional state spaces and its stability in learning complex strategies. The neural network architecture underlying our RL agent is designed to process the image-based state efficiently. It combines convolutional layers to extract spatial features, following by fully connected layers to map these features to action probabilities and state value estimates.

Training Process:
The agent learns through episodic interaction with the SLG environment. Each episode in dynamic planning begins with an initial random wall configuration and terminates when either a satisfactory layout is achieved or a maximum number of transformations is reached. The agent gradually improves its policy by maximizing cumulative rewards across episodes.
This RL formulation allows our system to learn design strategies that balance local

optimizations (individual room properties) with global constraints (overall layout coherence and adjacency requirements).

3 Experiments

This section presents two experiments to evaluate the effectiveness of laser-wall partitioning in SLG and the performance of RL to design space layouts. We define six distinct design scenarios, varying in complexity from 4 to 9 rooms. Table 1 outlines the properties of the desired layout for each scenario. For all scenarios, the desired room aspect ratio ranges from 1 to 6. The desired adjacency requirements are as follows: living room is connected to all the other rooms and to the entrance; each room access to at least one non-blocked façade.

Table 1: Design Scenarios Specifications

Scenario Number	Number of Rooms	Desired Areas (in cell counts)	Entrance location
1	4	{431, 322, 250, 206}	West
2	5	{301, 220, 220, 214, 210}	East
3	6	{279, 210, 186, 160, 160, 158}	West
4	7	{313, 150, 144, 140, 138, 138, 134}	West
5	8	{299, 216, 210, 192, 180, 170, 160, 150}	South
6	9	{435, 228, 190, 178, 150, 146, 140, 140, 134}	East

We conducted two experiments a) assessing the impact of wall type flexibility by comparing layouts generated with straight versus angled walls; and b) comparing on-light and off-light transformations to evaluate their impact on layout generation.

3.1 Results:

Angled versus straight walls: Figure 3 shows the results for training PPO agents on a subset of design scenarios from Table 1, comparing layouts generated using only straight walls versus those using only angled walls. The agent's ability to find a design solution indicates that the generated layouts closely match the geometric properties defined in the table, including the aspect ratio requirements. This is because geometric properties are integral constraints of the optimization problem, and without satisfying these constraints, a solution cannot be achieved. Therefore, the immediate conclusion is that the PPO agents were able to find solutions that closely match the desired geometric properties. However, observing the dashed lines on the layout reveals numerous missed connections with straight walls, whereas layouts generated with angled walls show marked improvement, successfully establishing all required connections. This demonstrates that using angled walls leads to more effective space partitioning, emphasizing that combining both straight and angled walls is promising, and highlighting the strength of our laser wall method compared to conventional rectangular dissection.

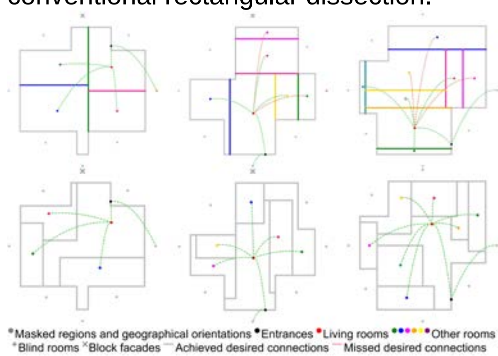


Fig 3: Comparison of layouts generated using straight walls only (top row) and angled walls only (bottom row).

On-light and off-light transformations: For each design scenario in Table 1, we trained a PPO agent using dynamic planning with both on-light and off-light transformations, utilizing identity-full walls in both cases. We then evaluate the layouts generated by our RL agents based on their adherence to the desired geometric properties (areas and aspect ratios) and topological requirements (adjacencies). Additionally, we consider factors such as convergence speed. Figure 4 illustrates the learning curves for the agent using on-light and off-light transformations in SLD. Both approaches show consistent improvement over time, with rewards increasing and episode lengths decreasing.

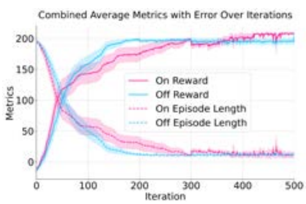


Fig 4: Average rewards and episode lengths over training iterations

The layouts generated by the trained PPO agents are shown in Figure 5 provide a clear visual representation of various design elements. Geometrically, the agent achieves a high degree of precision in meeting room area and aspect ratio requirements. The average difference between desired and achieved room areas across all scenarios is minimal, and the achieved aspect ratios closely match the specified values (both less than 5%). Topologically, the agent excels in establishing the required connections between rooms and living rooms and facades. The green dashed lines in each layout represent successfully implemented adjacencies, while the red dashed line and gray plus sign indicate instances where the agent failed to create a requested connection. Across all scenarios, there are 72 potential connections, calculated as the sum of room-to-living room and room-to-facade connections. The agent successfully establishes all but 2 of these connections, showing its proficiency in meeting adjacency requirements.

The results underscore the effectiveness of our RL-based approach in generating space layouts that align closely with specified design criteria. The agent demonstrates a strong capacity for balancing multiple objectives simultaneously, including room sizing, aspect ratios, and inter-room connectivity, while adapting to different levels of complexity in the design scenarios. This performance suggests that our method could be a valuable tool for architects and designers, offering efficient and creative solutions to complex space layout problems.

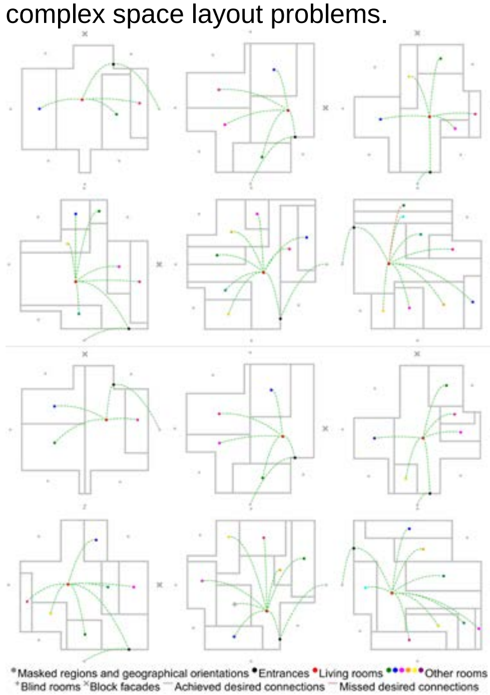


Fig 5: Comparison of space layouts generated using off-light (top two rows) and on-light (bottom two rows) transformations.

Design trajectory: To understand the trained agent's behavior, we illustrate the action trajectory in Figure 6 which shows the step-by-step transformation of an SLD by our trained RL agent for scenario 4. Starting from an initial random configuration, each subsequent image represents the layout after an action taken by the agent. Key observations include the gradual improvement in room proportions, the establishment of required adjacencies, and the overall organization of space becoming more coherent. This visual trajectory not only showcases the effectiveness of our RL approach but also provides insights into the decision-making process of the agent in the dynamic planning, highlighting how it balances various design criteria to achieve the final, optimized layout that satisfies the geometrical and topological properties defined in the design scenario.

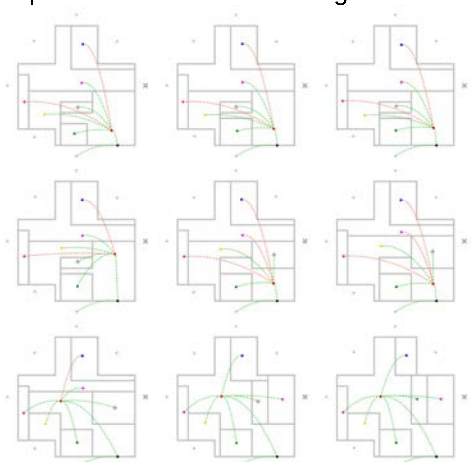


Fig 6: Agent action trajectory from initial random layout to final optimal layout.

4. Discussion and Conclusion

Our research demonstrates the effectiveness of the proposed laser-wall partitioning method in SLD. We framed the layout design problem as an MDP and successfully trained PPO agents to explore the solution space to find the optimal solution. The experiments across various scenarios highlight the adaptability and efficiency of our partitioning method in generating layouts that satisfy complex geometric and topological requirements. The comparison between straight and angled walls reveals the importance of flexibility in composing the space. Successful training of RL agents in different on-light, off-light dynamic planning emphasizes the robustness of the proposed laser-wall method. The visual trajectory of the agent's decision-making process provides insights into how it navigates the solution space, progressively refining layouts to meet specified criteria.

While our approach shows promise, there are areas for future exploration. Integrating more architectural constraints, such as wall alignments and room width and length constraints, could further enhance the practical applicability of this method. Additionally, exploring the potential of transfer learning between different design scenarios could improve the efficiency and generalization capabilities of the RL agent. Comparing dynamic planning and one-shot planning is also left for future research. As we continue to refine this approach, the ultimate goal is to develop a tool that not only assists architects in generating initial design concepts but also serves as a collaborative partner in the creative process, offering novel solutions that human designers might not have considered. We already open sourced the SpaceLayoutGym simulator to support open research in architectural design and RL: https://github.com/RezaKakooee/space_layout_gym.

References

[1] J. Michalek, R. Choudhary, and P. Papalambros, "Architectural layout design optimization," *Engineering Optimization*, vol. 34, no. 5, pp. 461–484, Jan. 2002

[2] D. Lobos and D. Donath, "The problem of space layout in architecture: A survey and reflections," *Arquitetura Revista*, vol. 6, pp. 136–161, Dec. 2010

[3] M. Mourshed, "AUTOMATED SPACE LAYOUT PLANNING FOR ENVIRONMENTAL SUSTAINABILITY".

[2] M. Saruwono, N. F. Zulkiflin, and N. M. N. Mohammad, "Living in Living Rooms: Furniture Arrangement in Apartment-Type Family Housing," *Procedia - Social and Behavioral Sciences*, vol. 50, pp. 909–919, Jan. 2012

[5] A. Drira, H. Pierreval, and S. Hajri-Gabouj, "Facility layout problems: A survey," *Annual Reviews in Control*, vol. 31, no. 2, pp. 255–267, Jan. 2007

[6] A. B. Kahng, J. Lienig, I. L. Markov, and J. Hu, "Chip Planning," in *VLSI Physical Design: From Graph Partitioning to Timing Closure*, A. B. Kahng, J. Lienig, I. L. Markov, and J. Hu, Eds., Dordrecht: Springer Netherlands, 2011, pp. 55–92.

[7] H. Liu, Z. Wang, A. Mazumdar, and C. Mousas, "Virtual reality game level layout design for real environment constraints," *Graphics and Visual Computing*, vol. 4, p. 200020, Jun. 2021

[8] B. Medjdoub and B. Yannou, "Separating Topology and Geometry in Space Planning," Mar. 16, 2013, *arXiv*: arXiv:1303.4017.

[9] G. Stiny and W. J. Mitchell, "The Palladian grammar," *Environ. Plann. B*, vol. 5, no. 1, pp. 5–18, 1978

[10] P. H. Levin, *Use of graphs to decide the optimum layout of buildings*. in Building research current papers - Design series. Garston: Building Research Station, 1964.

[11] T. Grasl and A. Economou, "From topologies to shapes: Parametric shape grammars implemented by graphs," *Environment and Planning B: Planning and Design*, vol. 40, pp. 905–922, Jan. 2013

[12] K. Shekhawat, N. Upasani, S. Bisht, and R. Jain, "GPLAN: Computer-Generated Dimensioned Floorplans for given Adjacencies," Aug. 04, 2020

[13] L. March and P. Steadman, "The Geometry of Environment: An Introduction to Spatial Organization in Design," Routledge & CRC Press. Accessed: Nov. 06, 2024.

[14] J. H. Jo and J. S. Gero, "Space layout planning using an evolutionary approach," *Artificial Intelligence in Engineering*, vol. 12, no. 3, pp. 149–162, Jul. 1998.

[15] M. K. Thakur, M. Kumari, and M. Das, "Architectural layout planning using genetic algorithms," in *2010 3rd International Conference on Computer Science and Information Technology*, Jul. 2010, pp. 5–11.

[16] E. Rodrigues, A. R. Gaspar, and Á. Gomes, "An approach to the multi-level space allocation problem in architecture using a hybrid evolutionary technique," *Automation in Construction*, vol. 35, pp. 482–498, Nov. 2013.

[17] P. Merrell, E. Schkufza, and V. Koltun, "Computer-generated residential building layouts," in *ACM SIGGRAPH Asia 2010 papers on - SIGGRAPH ASIA '10*, Seoul, South Korea: ACM Press, 2010, p. 1.

[18] B. Dillenburger, "Raumindex. Ein datenbasiertes Entwurfsinstrument," Doctoral Thesis, ETH Zurich, 2016.

[19] S. Chaillou, "ArchiGAN: Artificial Intelligence x Architecture," in *Architectural Intelligence*, P. F. Yuan, M. Xie, N. Leach, J. Yao, and X. Wang, Eds., Singapore: Springer Nature Singapore, 2020, pp. 117–127.

[20] F. Mostafavi, M. Tahsildoost, M. Zomorodian, and S. Shahrestani, "An interactive assessment framework for residential space layouts using pix2pix predictive model at the early-stage building design," *Smart and Sustainable Built Environment*, vol. 13, Dec. 2022.

[21] N. Nauata, S. Hosseini, K.-H. Chang, H. Chu, C.-Y. Cheng, and Y. Furukawa, "House-GAN++: Generative Adversarial Layout Refinement Networks," Mar. 03, 2021, *arXiv*: arXiv:2103.02574.

[22] M. A. Shabani, S. Hosseini, and Y. Furukawa, "HouseDiffusion: Vector Floorplan Generation via a Diffusion Model with Discrete and Continuous Denoising," Nov. 23, 2022, *arXiv*: arXiv:2211.13287.

[23] J. Ho, A. Jain, and P. Abbeel, "Denoising Diffusion Probabilistic Models," Dec. 16, 2020.

[24] D. Wang and R. Snooks, "Artificial Intuitions of Generative Design: An Approach Based on Reinforcement Learning," in *Proceedings of the 2020 DigitalFUTURES*, P. F. Yuan, J. Yao, C. Yan, X. Wang, and N. Leach, Eds., Singapore: Springer, 2021, pp. 189–198.

[25] B. Kim, Y. Jeong, and J. G. Shin, "Spatial arrangement using deep reinforcement learning to minimise rearrangement in ship block stockyards," *International Journal of Production Research*, vol. 58, no. 16, pp. 5062–5076, 2020.

[26] L. Ren, Y. Song, J. Lu, and J. Zhou, "Spatial Geometric Reasoning for Room Layout Estimation via Deep Reinforcement Learning," in *Computer Vision – ECCV 2020*, A. Vedaldi, H. Bischof, T. Brox, and J.-M. Frahm, Eds., Cham: Springer International Publishing, 2020, pp. 550–565.

[27] T. Hosmer, P. Tigas, D. Reeves, and Z. He, "Spatial Assembly with Self-Play Reinforcement Learning," presented at the ACADIA 2020: Distributed Proximities, Online and Global, 2020, pp. 382–393.

[28] M. Ruiz-Montiel, J. Boned, J. Gavilanes, E. Jiménez, L. Mandow, and J.-L. Pérez-de-la-Cruz, "Design with shape grammars and reinforcement learning," *Advanced Engineering Informatics*, vol. 27, no. 2, pp. 230–245, Apr. 2013.

[29] L. Mandow, J.-L. Pérez-de-la-Cruz, A. B. Rodríguez-Gavilán, and M. Ruiz-Montiel, "Architectural planning with shape grammars and reinforcement learning: Habitability and energy efficiency," *Engineering Applications of Artificial Intelligence*, vol. 96, p. 103909, Nov. 2020.

[30] F. Shi, R. K. Soman, J. Han, and J. K. Whyte, "Addressing adjacency constraints in rectangular floor plans using Monte-Carlo Tree Search," *Automation in Construction*, vol. 115, p. 103187, Jul. 2020.

[31] M. Klar, M. Glatt, and J. C. Aurich, "An implementation of a reinforcement learning based algorithm for factory layout planning," *Manufacturing Letters*, vol. 30, pp. 1–4, Oct. 2021.

[32] X. Di and P. Yu, "Multi-Agent Reinforcement Learning of 3D Furniture Layout Simulation in Indoor Graphics Scenes," Feb. 18, 2021, *arXiv*: arXiv:2102.09137.

[33] P. Veloso and R. Krishnamurti, "Spatial synthesis for architectural design as an interactive simulation with multiple agents," *Automation in Construction*, vol. 154, p. 104997.

[34] R. Kakooee and B. Dillenburger, "Reimagining space layout design through deep reinforcement learning," *Journal of Computational Design and Engineering*, vol. 11, no. 3, pp. 43–55, Jun. 2024, doi: 10.1093/jcde/qwae025.

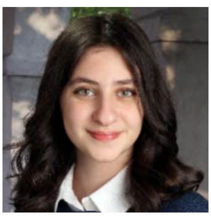
[35] J. Schulman, F. Wolski, P. Dhariwal, A. Radford, and O. Klimov, "Proximal Policy Optimization Algorithms," Aug. 28, 2017, *arXiv*: arXiv:1707.06347. doi: 10.48550/arXiv.1707.06347.

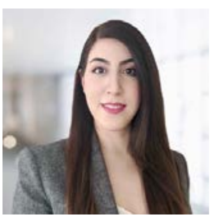
Shared Spaces of Creation: Generative Artificial Intelligence's Creative Agency and Affordances in the Artistic Process

Natalie Oulikhanian, BA
University of Toronto, Canada
e-mail: natalie.oulikhanian@mail.utoronto.ca

Dr. Behnaz Merikhi, BSc, MSc, PhD.
University of Toronto, Canada
Northeastern University, Canada
e-mail: behnaz.merikhi@utoronto.ca

Dr. Rhonda McEwen, BSc, MBA, MSc, PhD
Victoria College, University of Toronto, Canada
e-mail: vic.president@utoronto.ca


Natalie Oulikhanian


Dr. Behnaz Merikhi


Dr. Rhonda McEwen

Abstract

The emergence of Generative Artificial Intelligence (Gen-AI) in the field of art carries a range of ontological questions on the nature of creativity and artwork. Gen-AI can be conceptualized to be reaching, or is currently, the approximation of human creativity. Consequently, many discussions aim for redefinitions of what it means for an entity to be considered creative or to produce art. However, this motivation to better clarify what it means to have a property such as creativity is reminiscent of a different, underlying, sentiment: that the role of technology in the creative process is considered changing. If it is indeterminate what a technology is understood to be, it becomes especially a challenge to discern what it can do – its possibilities. We therefore aim to explore how to position technology and artifacts, and with it, Gen-AI, in the relationship composing the artistic creative process. This paper traces the evolution of how we conceptualize technology's role in the creative process, moving from traditional humanist perspectives that position the human artist as the sole agent of creativity, to an adoption of Csikszentmihalyi’s systems theory approach that can recognize creativity and agency as inherently relational. Building on this foundation, we argue that understanding Gen-AI's unique impact on artistic creation requires examining how different technologies materialize this relational agency through their distinct affordances. We demonstrate that while previous technologies primarily afforded possibilities in the realm of craft and physical execution, Gen-AI introduces unique affordances in the cognitive domain of creativity itself – offering the opportunity to expand one’s capacity to work with and realize their imagination. This fundamental shift from physical to cognitive affordances largely distinguishes Gen-AI from previous technological innovations in art and demands a new framework for understanding creative agency in the age of artificial intelligence.

1. Introduction

A creative product does not consist only of the outcome-oriented process focused on creating a meaningful artifact; it also encompasses the highly relational dynamics arising between various tools and actors involved in the act of creation [1]. Art, being a quintessential manifestation of the creative process, has witnessed profound transformations and aesthetic shifts throughout history as new technologies emerge. As artists integrate the capabilities of a technology into their work, the creative process has the potential to expand to become a shared space that includes human intention and nonhuman influence. Consequently, emerging technology shapes the creative process, transforming aesthetic diversity and influencing the boundaries of artistic expression. Yet, identifying how exactly these dynamics unfold is of particular challenge. What is it about a new technology that generates change in the creative process? Can these technologies be considered to hold any responsibility over a final piece of art, or is it the artist who exerts direct control over how all the technology is used? While not exclusive to only advanced digital tools, these questions of what or where to attribute responsibility to creative outcomes is growing especially important when a technology emerges that appears to substantially reduce the role of the human in the artistic creative process. The emergence of Gen-AI especially represents a pivotal moment in this ongoing evolution of artistic creativity. As AI models exhibit that they can feasibly produce outputs that society may traditionally judge as being creative, generating novel and useful outcomes [2], we can see changes to existing paradigms of the creative process and conventional understandings of artistry. Especially, we argue that Gen-AI re-emphasizes the need for theories about agency that are inherently relational and that do not view the human as the focal source of responsibility in the creative process in order to best identify its distinctions among other technologies.

2. The Humanist Creative Process: Human as the Creative Agent

The emergence of a new technology in the artistic process is often first met with ideas about loss: some kind of erosion of a quality otherwise considered essential in the creative process. For example, the advent of a novel, and importantly a significant, technology is largely characterized by concerns about displacement – the loss of practice, or with inauthenticity – the loss of intention and craft. A common claim thus persists that a new technology within the field of art will produce inferior art: hurting art as a subject, producing non-art, or creating 'lifeless' art. Photography's emergence formed a process of creation some consider as lacking "that refined feeling and sentiment which animate the productions of a man of genius" [3]. Further, the introduction of computer animation faced skepticism that any art produced would bypass skills traditionally associated with the creative process despite the process remaining labour-intensive and requiring skill for most details [4]. Gen-AI similarly captures concerns over the disappearance of human creative labour [5], the rights of artistic ownership [6] and the future diversity of artwork overall by forming hegemony [7].

Yet, in conceptualizing an artwork or artist as having lost something, there must be the presumption that certain qualities or capacities were considered inherent to such entity prior to the technology. Largely, concerns about loss will presuppose that the creative process is confined within a certain entity, focally: humans. When technologies are seen for the way in which they replace, replicate, or imitate the qualities of a human artist, the creative process responsible for producing art found within human artists is understood as being reduced in some way. Under this assumption, we may then associate new technology as being the entity that may have 'unethically' gained the lost trait. We may also view the rise of technologies such as Gen-AI in art positively, arguing that the technology does not necessarily need to induce losses, but acts as another exploratory tool in which humans have access to in their creative process [4]. This perception though continues to frame AI's role as subordinate to human qualities.

Overall, these assumptions about the centrality of the human actor largely are a result of the way in which we may perceive agency in the creative process. Where agency is typically understood as what we may attribute to an entity that has the capacity to act [8], such as in a process or outcome, it is no surprise then that traditional interpretations of the term are firmly rooted in human control. In the artistic process, a human is often considered as the central figure in both intention and execution. Whereas, technology and other artifacts are considered to be 'employed,' requiring the artist's capabilities to provide any use.

Because of the notion that humans are the only full agents in art, the field of computational creativity generally lacks discussion of creative agency since it is considered a secondary topic beneath ontological questions about genuine creativity in the computer, primarily whether a computer can create truly novel and useful work. In traditional computer-generated work, creativity is considered a mere appearance, and only "due to some specifiable slice of the programmer's own creativity having been imprinted onto the algorithmic workings of the system" [9]. The machine is considered to rely on two mechanisms: the use of the computer and the artist's intentions, consequently relying generally only on the human's agency [10]. Additionally, since evaluating creative capacity can be traced to whether the technology is considered to have agency to begin with, such as Moruzzi argues that if a system is judged as holding high creativity, it must hold agentive capacities [11], then if we face difficulties discerning whether computers are even first creative, how could we enter discussions about agency outside of only the human?

However, Gen-AI is perceived as a technology that appears beyond the typical understanding of machines as only hosting repetitive and pre-programmed activities. Gen-AI models have produced outputs that are not directly predictable or controllable by their human users, exhibiting a degree of apparent autonomy or intentionality that blurs the line between assistance and co-creation. Therefore, more recent discussions, including ours, in creativity centre on whether Gen-AI acts as a tool or as something closer to an artist or agent [12]. The introduction of Gen-AI thus emphasizes the re-emergence of various theories that explore the topic of agency beyond the human artist.

3. The Post-Humanist Creative Process: Creative Agency Elsewhere and Everywhere

As Oliver Bown emphasizes, creativity does not exist in a vacuum: the notion of attributing sole responsibility for a creative work to a single source quickly unravels under scrutiny [13]. All human creativity occurs within an intertwined system with entities of mutual influence. Even in viewing artwork as intrinsically arising from the human, any creator is ultimately subject to influence from their culture, environment, tools available, and other forces at play during their creative process. Further, creativity, distinct for its focus on novelty, is considered to derive from the breaking down of conceptual paradigms. This process is inherently relational. Novelty can only be defined compared to existing norms and standards. Whatever the individual cognitive processes involved in practicing creativity, it must be one that occurs in a contest of previous cultural or social achievements and is thus inherently inseparable from the artist's environment and understanding of novelty. Viewing the concept of creativity through the lens of a wide-spanning network consisting of various sources of influence, substantially proposed by Csikszentmihalyi's System's Model of Creativity [14], may provide a more accessible groundwork to establish an understanding of creative agency in the age of Gen-AI [15].

Importantly, adopting a systems theory of creativity can better inform us that agency, too, must be treated as a relationship rather than something that can be had. If creativity is an expression of agency [11], creativity that exists on a relational plane must also arise from a relational level of agency. Understanding agency from this relational perspective aims to resituate human beings in their environment and better conceptualize agency as fundamentally distributed [16].

A range of ways exist to conceive the idea of a relational agency, often distinguished between their interpretation of the role of an artifact and its properties. First, agency can be understood by reviewing the relationship among entities from the top-down, emphasizing the greater sense of interconnectedness between entities than any properties associated among them. Such theories can be considered to speak of both the social and the material in the same light, avoiding treating them as separate, even if interacting, phenomena [17]. For example, Latour regards agency as not an essence that is bound to humans, but as rather a capacity which occurs through the interaction of actors [18]. This view considers agency an emergent phenomenon arising from engagement itself. Second, post-human theories about agency can also compose the relational world by continuing to recognize distinctions between entities. In some approaches of this position, agency can be composed of both ‘human agency’ and ‘material agency.’ By identifying two forms, this theory of agency aims to build a difference in the properties of artifacts in order to better allow “for maintaining the distinction between social and material agencies with respect to intentionality while still recognizing their synergistic interaction” [19]. Such theories tend to “speak of humans and technology as mutually shaping each other, recognizing that each is changed by its inter-action with the other, but maintaining, nevertheless, their ontological separation” [17]. The two approaches remain rooted in the emphasis of adopting a relational perspective to agency, and thus also creativity, yet fundamentally differ on whether they may have an initial independence of entities, artifacts, and technology, from each other.

However, it is important that when adopting a relational standpoint to agency that a theory can still be able to recognize that the landscape of art differs as one technology emerges to another; differing technologies like Gen-AI may exert influence on the creative process in unique ways. Under the first perspective, presuming that there are no independently existing entities with inherent properties [17], any kind of distinctions made between humans and artifacts, and artifacts between other artifacts, is purely analytical. Materials are considered and treated as purely relational products, “they do not exist in and of themselves” [20]. However, moving towards the second perspective’s distinctions between material and human agency also does not necessarily mark key differences in why the introduction of certain technology shaped the creative process and artwork differently. Instead, it is important for the discussion of both the history and future of technology to better identify why some lead to differing outcomes still within a relational perspective. Although Pels et al. observes that it is less so what the material symbolizes within social action that matters but “their constitutive agentic effects within the entangled networks of sociality/materiality,” [21] these agentic effects in and of themselves can differ between materials, and is how we aim to explore the distinction of Gen-AI.

4. Where Creative Agency is Amidst Generative Artificial Intelligence: Affording Cognitive Ideation and Conceptualization

In order to ensure that we may properly acknowledge the differences between how various technologies shape the artistic creative process, we may apply Leonardi's argument that agency is not inherent in any entity but has rather materialized through its relations [22]. Particularly, he considers that there are various ways in which agency can materialize. Leonardi thus considers the form of materialization to depend on affordances. Rather than taking an ontological understanding of only the properties of an entity, such as if it appears to resemble autonomy or intentionality, affordances are distinct by forming a new definition of what value and meaning are: asking what an entity offers – what it provides or furnishes [23]. Affordances denote the potential action of technologies. Using this, agency can continue to be relational while justifying how agency appears different among various networks consisting of differing technologies.

So then, what are the affordances Gen-AI offer that may provide an understanding of how the overall agency looks like in the creative process, how it materializes? We argue that the affordances produced by Gen-AI are particularly unique to other technologies, and thus materializes agency differently.

First, to claim a distinction, it is important to form an understanding of the type of affordances produced by technologies prior to Gen-AI. The emergence of most technologies in the artistic process is largely perceived to afford capabilities primarily within the craft-oriented dimension of artistic creation, as they are understood to offer ways to automate or enhance the physical execution of artistic vision. These affordances materialize in how they provide opportunities to manipulate, transform, and realize artistic concepts through mechanical or digital means. While new artistic styles and mediums have indeed emerged following the introduction of these technologies, these developments stem not from technologies directly affording changes in an artist's cognitive process associated with creation but occur when the affordance that a specific technology can bring new mediums and styles is realized into effect. These outcomes of changes in the landscape of art can be considered affordances that are objectively possible from the technology itself but require external perception to have an effect [24]. It was not immediately recognized, for example, that photography would afford processes capable of the novel form of cinema and filmmaking, requiring discovery by both artists and technologists [4]. Further, even as technologies became more sophisticated, their affordances remained largely anchored in how they could enhance the execution and manipulation of artistic vision rather than participating in the cognitive dimension of creativity itself. The relational agency that materializes, in this case, can be more closely understood as one characterized by 'helping,' where the affordances of technologies arise from their role in the physical crafting process of art.

Gen-AI, while affording some actions that can be considered more part of the craft or physical elements of the creative process – such as in making additions to or editing existing artwork, is primarily unique by affording actions that are located more cognitively. Particularly, Gen-AI's affordances occur in the conceptualization and ideation process, a process that is central to the creation of a final artwork, but has remained untouched by the affordances of any traditional technology.

The process of conceptualization or ideation, despite being influenced by a range of entities, occurs cognitively in the artist. An artist is understood to take raw and unstructured material or knowledge and consider different possibilities for recombining and transforming relevant knowledge in order to generate a tangible novel form [25]. The creative process has been compared akin to a qualitative problem-solving process in which, among numerous problems that are beset, an artist tries to solve them consciously or unconsciously [26]. In this way, the artist's creative process is largely built on interacting with what is around them. This ability to interact with both an observer and with one another is what is typically reserved to be an affordance found between people and animals [23]. Other technologies, contrastingly, do not have the capacity to change and, unlike the living, their behaviour does not afford new behaviour: their affordances can only be realized [24]. A paintbrush, for example, does not make decisions based on past painting experiences. Gen-AI differs. By leveraging its high-dimensional representation of latent space, AI can interpolate between features, creating novel outputs that extend beyond the original inputs [12]. Gen-AI models can autonomously discover patterns and relationships within the data, offering insights into the creative process that might not be immediately apparent to human artists. For instance, AI trained on datasets of artworks can identify complex patterns of shading and colour usage that contribute to the perception of illumination in ways that are intuitive but not explicitly understood by human creators [27]. AI models are also able to learn and represent knowledge, identify patterns and knowledge structures, and use these to make useful and appropriate connections and inferences [28]. In this way, the affordances associated with Gen-AI are closer to that of how humans and animals – where behaviour affords behaviour [23]. This transformation is particularly evident in how Gen-AI, like conceptual art, operates without requiring physical referents from the natural world [12]. Where traditional image-making technologies like photography necessitated a direct connection to physical reality through light and chemical processes, Gen-AI's creative process exists purely in the realm of computational abstraction.

Unlike previous technologies, Gen-AI's affordances thus exist in a shared cognitive space between living actors, producing a form of relational agency that fundamentally transforms our understanding of creative partnerships. Therefore, the location of where affordances resemble or originate from in the creative process matter. While other materials, such as watercolour, demonstrate a level of unpredictability like Gen-AI that can appear to influence the imagination process, a watercolour's version of unpredictability often remains purely in the physical domain of execution that affects the craft-oriented process. Gen-AI's affordance of unpredictability, however rests in a more cognitive realm, enhancing ideation and conceptual development, and consequently, its affordance is more active. Because Gen-AI can learn, adapt, and generate novel responses based on interactions, it creates a dynamic feedback loop where both human and AI behaviours afford new behaviours in each other. The ability for one's ability to imagine and enact ideas from their imagination is similarly heightened by Gen-AI's ability to do the same. This recursive relationship means that agency in the creative process can no longer be understood as simply evenly distributed or emergent from static material engagement but must be conceived as an evolving dialogue between cognitive entities that mutually shape each other's creative possibilities.

5. Conclusion

That agency exists beyond only the human is not a new conception. The understanding of creativity and agency as a relationship emerging from a variety of forces and influences can extend to what can be considered to be the simplest of materials to those understood as inherently complex. However, as Gen-AI models exhibit outcomes not wholly predictable and hold a level of apparent autonomy in its creation, the blurring between what is considered the role of technology in creative processes strengthens, and consequently pushes the re-emergence of these discussions about the creative process as shared among actors. Yet, in applying a relational perspective to the emergence of Gen-AI, we are compelled to consider why the emergence of certain technologies has shaped and continues to shape the creative process differently from others. In order to move away from ontological suggestions about the properties of a computer or machine, such as whether it is considered creative, we highlight the importance of introducing the concept of affordances in materializing a relational agency, as suggested by Leonardi [22]. Consequently, we consider that Gen-AI represents a fundamental shift from affordances typically associated with traditional technologies that affect the craft physically, to cognitive affordances that are closer to the imagination. Gen-AI thus provides a unique shared imaginative process, where the dimensions of understanding and visions are becoming shared between the two entities: as something in which each other's possibilities can be accessed and visited. Therefore, we consider that Gen-AI can afford the opportunity within the creation process to enhance acting on one's imagination. In arguing that the location of the technology's affordance rests in a more cognitive form, we also consider that Gen-AI can be similarly constrained by issues pertaining to the cognitive processes among the living, such as of bias. Consequently, it is especially important for future research in the exploration between art and Gen-AI to facilitate meaningful interactions that can emphasize transparency and explainability. By participating in the cognitive domain of creativity, Gen-AI can introduce a new nuance in understanding creative partnerships, where both the human and machine can actively shape the other's imaginative possibilities rather than maintaining boundaries between ideation and execution.

6. References

[1] Nordström, P., Lundman, R., & Hautala, J. (2023). Evolving coagency between artists and AI in the spatial cocreative process of artmaking. *Annals of the American Association of Geographers*, *113*(9), 2203–2218. https://doi.org/10.1080/24694452.2023.2210647

[2] Zhou, E., & Lee, D. (2024). Generative Artificial Intelligence, human creativity, and art. *PNAS Nexus*, *3*(3). https://doi.org/10.1093/pnasnexus/pgae052

[3] Photography. (1855, March 14). *The Crayon*, *1*(11), 170.

[4] Hertzmann, A. (2018). Can computers create art? *Arts*, *7*(2), 18. https://doi.org/10.3390/arts7020018

[5] Bender, S. (2024). Generative-AI, the media industries, and the disappearance of human creative labour. *Media Practice and Education*, 1–18. https://doi.org/10.1080/25741136.2024.2355597

[6] Chesterman, S. (2024). Good models borrow, great models steal: Intellectual property rights and Generative AI. *Policy and Society*. https://doi.org/10.1093/polsoc/puae006

[7] Doshi, A. R., & Hauser, O. P. (2024). Generative AI enhances individual creativity but reduces the collective diversity of novel content. *Science Advances*, *10*(28). https://doi.org/10.1126/sciadv.adn5290

[8] Loyal, S. (2012). Agency. *The Wiley‐Blackwell Encyclopedia of Globalization*. https://doi.org/10.1002/9780470670590.wbeog562

[9] Veale, T., Amílcar Cardoso, F., Pérez y Pérez, R. (2019). Systematizing Creativity: A Computational View. In: Veale, T., Cardoso, F. (eds) Computational Creativity. Computational Synthesis and Creative Systems. Springer, Cham. https://doi.org/10.1007/978-3-319-43610-4_1

[10] Salah, A. A. A. (2021). AI Bugs and Failures: How and Why to Render AI-Algorithms More Human? In P. Verdegem (Ed.), *AI for Everyone?: Critical Perspectives* (pp. 161–180). University of Westminster Press. http://www.jstor.org/stable/j.ctv26qjjhj.12

[11] Moruzzi, C. (2022). Creative Agents: Rethinking Agency and Creativity in Human and Artificial Systems. *Journal of Aesthetics and Phenomenology*, *9*(2), 245–268. https://doi.org/10.1080/20539320.2022.2150470

[12] Mazzone, M., & Elgammal, A. (2019). Art, creativity, and the potential of artificial intelligence. *Arts*, *8*(1), 26. https://doi.org/10.3390/arts8010026

[13] Bown, O. (2015). Attributing Creative

Agency: Are we doing it right? *Sixth International Conference on Computational Creativity*. https://computationalcreativity.net/iccc2015/proceedings/1_3Bown.pdf.

[14] Bown, O. (2015). Sixth International Conference on Computational Creativity. Retrieved from https://computationalcreativity.net/iccc2015/proceedings/1_3Bown.pdf.

[15] Celis Bueno, C., Chow, PS. & Popowicz, A. Not “what”, but “where is creativity?”: towards a relational-materialist approach to generative AI. *AI & Soc* (2024). https://doi.org/10.1007/s00146-024-01921-3

[16] Keeling, D. M., & Lehman, M. N. (2018). Posthumanism. *Oxford Research Encyclopedia of Communication*. https://doi.org/10.1093/acrefore/9780190228613.013.627

[17] Orlikowski, W. J. (2007). Sociomaterial practices: Exploring Technology at work. *Organization Studies*, *28*(9), 1435–1448. https://doi.org/10.1177/0170840607081138

[18] Latour, B (2005). *Reassembling the Social: An Introduction to the Actor-Network Theory*. Oxford, England and New York, NY, USA: Oxford University Press.

[19] Leonardi, P. M. (2012a). *Car crashes without cars: Lessons about simulation technology and organizational change from Automotive Design*. MIT Press.

[20] Law, J. (2004). *After method: Mess in social science research*. Routledge.

[21] Pels, D., Hetherington, K., & Vandenberghe, F. (2002). The Status of the Object. Theory, Culture & Society, 19(5-6), 1-21. https://doi.org/10.1177/026327602761899110

[22] Leonardi, P. (2023). Guest Editorial: Affordances and Agency: A Clarification and Integration of Fractured Concepts. *MIS Quarterly*, 47(4) pp.ix-xx. https://aisel.aisnet.org/misq/vol47/iss4/3/

[23] Gibson, J. J. (2014). The Theory of Affordances. In *The people, place, and space reader* (pp. 56–60). Routledge.

[24] Norman, D. A. (1999). Affordance, conventions, and design. *Interactions*, *6*(3), 38–43. https://doi.org/10.1145/301153.301168

[25] Lang, K. (2019). *Field Notes on the Visual Arts*. https://doi.org/10.2307/j.ctv36xvrmv

[26] Ecker, D. W. (1963). The Artistic Process as Qualitative Problem Solving. *The Journal of Aesthetics and Art Criticism*, *21*(3), 283–290. https://doi.org/10.1111/1540_6245.jaac21.3.0283

[27] Stork, D. G. (2023). How AI is expanding art history. *Nature*, *623*(7988), 685–687. https://doi.org/10.1038/d41586-023-03604-3

[28] Russell, S. J., & Norvig, P. (2010). *Artificial Intelligence: A modern approach* (3rd ed.). Pearson Education Inc.

Outsider Architectures

Prof. K. Lapierre, BArch, MA, MArch, PhD.
School of Design, Université du Québec à Montréal, Canada
www.katherinelapierre.com
e-mail: lapierre.katherine@uqam.ca


Photo: Samuel Barriault

Abstract

The following research engages with the notion of the irrational in the conception of digital architectural design. Along this line, various projects have been developed that propose explorations in the form of plans and 3D models, exposing the precarious balance between feasibility, viability, and impossibility due to standards and physical or technical constraints. These various experiments are a testing ground for new practices and media, such as the exploration of different 3D printing techniques. The work focuses on *outsider architecture*, directly referring to the expression "Art Brut" (or "Outsider Art"). Opening an unprecedented sphere of action, it broadens the debate on the status and possible contribution of Art Brut notions to architecture and design in constructing a global architectural discourse. The project centers on the design of sculptural models inspired by empirical practices that were surveyed and gathered during my fieldwork in *outsider architecture*. The works themselves integrate enlarged-scale versions of chosen artifacts from nature as well as found objects that are scanned and transformed. Influenced by the idea that *outsider architectures* can revive Bachelardian theories once silenced by the paradigm shift that accompanied modernity, the current research-creation work links the unconscious to Bachelard's concept of material imagination. The model and plan studies, which draw from sculptural and architectural processes, are used to create living units that explore the potential of digital architecture as a trigger for the imagination, while the different representation techniques used in these creations contribute to developing various skills in computer modeling, 3D scanning, and 3D printing. *Outsider* creates the opportunity to investigate methods of representation sidelined by traditional architectural practices, shifting the logic of construction by emphasizing the creation of spaces with numeric tools of representation that follow empirical

design methods. Model-making becomes a laboratory for the formal exploration of various numeric transformations such as anamorphism, scaling, piercing, extruding, and erasing.

1. Introduction to the *Ideal Palace*

Fig. 1 & 2 – Palais idéal, Facteur Cheval, Hauterives, France, 2007 and 2024. East facade (Autocad drawing and scan with polycam / visualization with Rhino).

In *Earth and Reveries of Will*, Gaston Bachelard explores the idea of "underground life as an image of rest" and evokes the *Ideal Palace* [1]. Facteur Cheval (Ferdinand Cheval, 1836-1924) built the *Ideal Palace* (1879-1912) and the *Tomb of Endless Rest* (1914-1922) in Hauterives, France, until his death. He follows the idea of an architecture conceived from the dimensions of the human body, using the image of the grotto as a fundamental archetype. Bachelard describes the cave as "a refuge of which we dream endlessly […] We want to be protected, but we do not want to be enclosed." [2] Through the image of a grotto in the shape of a coiled shell, he sees the desire to live in the heart of a rock, a natural dwelling. [3] It is through his reading of Jung that he reflects on the notion of archetype, drawing a distinction between perception and imagination. For Bachelard, imagination is "the ability to distort the images provided by perception." [4] According to Jean-Claude Filloux, "to psychoanalyze reason, from the Bachelardian perspective, is to update that cognitive unconscious which defends the mind against being open to the new." [5] At the turn of the 20th century, the *Ideal Palace* crossed a new threshold on the scale of the irrational. It revived the tradition of rockwork and garden follies in the image of a cave-palace. It constitutes a monstrous architecture through the hybridization of the archetypes of the grotto, the labyrinth, and the castle.

In a letter from Cheval dated March 15, 1905, he writes: "In a dream, I had built a palace, a castle or caves, I can't quite put it into words. [...] Then, fifteen years later, just when I'd almost forgotten my dream, when I least expected it, my foot reminded me of it. My foot had caught a stone [words crossed out, hesitation in the text], an obstacle that almost made me fall. I wanted to know what it was; it was such a strangely shaped stone that I put it in my pocket to admire it at my leisure. [...] It is a molasses stone worked by the waters and hardened by the force of time. It has become as hard as rock; it represents a sculpture so bizarre that it is impossible for man to imitate it; it represents every kind of animal, every kind of caricature." [6] This first cornerstone of the palace is placed with others on one of the bases of the terrace, and its volume is equivalent to two or three human heads. When Cheval writes of pocketing the stone that awakened his dream, he sees himself as a giant. Cheval wrote: “My thoughts will live with this rock.” The monument bears the inscription “L'auteur du Peuple 1884” (“The author of the People”). In 1905, Cheval replaced it with a simple inscription: “Temple de la Nature - 1884” (“Temple of Nature “). Around 1897, he added “Seul au monde” (“Alone in the world” or “the only one in the world”). In 1902, anxious to control the image of his palace, he distributed the first photographs titled “Palais imaginaire seul au monde” (“Imaginary palace alone in the world”). The name “Palais Idéal” was inspired by Émile Roux-Parassac's poem “Ton Idéal, ton Palais”, dedicated to Cheval in 1904. In 1897, “the entire form is but a single block of *rocaille,* about 600 cubic meters of stone in its entirety.” [7] Claude Prévost writes of Facteur Cheval's palace: “This extreme of itself extends it, giving birth to the other according to the law of the imaginary and the unconscious. Spectacular divine play of creation, another form of sociality.” [8]

Fig. 3 – First stone, Ideal Palace (scan with polycam, 2024).

The formal universe of the *Ideal Palace* opens up to the theme of “archisculpture,” a neologism used by Michel Ragon, an author particularly interested in the notions of Art Brut. In his work *Où vivrons-nous demain?* [9] (Where will we live tomorrow?), Ragon uses the term "archisculpture" to describe creators who, according to him, reconnect with the radical anti-functionalism of Facteur Cheval, Gaudi (Spain), or that of the German Expressionists and Austrians: Bruno Taut, Hermann Finsterlin, and Frederick Kiesler. We can trace the terminology to the nuanced definitions attached to it: “architecture-sculpture,” “sculpture-architecture” [10], “archisculpture”, [11] “anarchitecture.” [12]

My own work engages in this lineage, linking the unconscious to the material imagination, inspired by psychoanalysis, analytical theory, and Jung's archetypes. Bachelard is the first to have undertaken research entirely centered on the imagination of matter.

2 . Outsider Architecture

It is in relation with the notions of Art Brut that my research and creation projects focus on the notion of *outsider architecture*, a direct reference to the expression “Outsider Art”. This term was used for the first time by Roger Cardinal in 1971, [13] and published a year later with its name-sake book, [14] to translate the expression “Art Brut,” which was defined by Jean Dubuffet in 1945. My projects concentrate on the creation of sculptural models inspired by practices and research on *outsider architecture*. Centered on the foundations of *archisculpture*, my work confronts the idea of integrating the arts with architecture, while going beyond the schema of a simple dialogue between the

artwork and the architectural space.
I therefore propose the material exploration of notions located at the frontier of the fields of art and architecture.

My recent projects, *Villa Poiësis* (CALQ, 2019-2020) and *Villa Tekhnai* (CALQ, 2021-2023), which focus on architecture as a work of art, are a play of tension between irrational geometries and orthogonal prisms with brutalist echoes. In connection with my previous research, [15] I note the potential of the architectural typologies of the exterior cave and troglodyte dwelling that have been neglected by contemporary architecture practice. I have been particularly interested in the play of scales and my observations of "wild" nature. By exploring these different empirical design methods, the prototypes themselves constitute the test beds of a practice that privileges research on architecture as a work of art.

Fig. 4 – Rock, Cap-aux-Oies, Villa Tekhnai, 2021-2023.

Fig. 5 – Cap-aux-Oies, Villa Tekhnai, 2021-2023.

3 . Tekhnai

In the *Villa Tekhnai* project (2021-2023), the focus of research is the transposition of forms arising from fire. Opening the question of the limits of reasonableness, the realization of this project allowed me to invest the potential of digital tools in the concept of tekhnai through an architectural lens. Raw digitized figures such as logs recovered from a campfire, fish scales in blown glass, rocks from the Éboulements region in Quebec—where the impact of a huge meteorite, 350 million years ago, has deeply modified the topography—are used to transcribe the forms of these objects shaped by fire. In a set of tests, I explored the relationship between assemblies of raw shapes and figures sculpted by various fire techniques and Euclidean geometries.

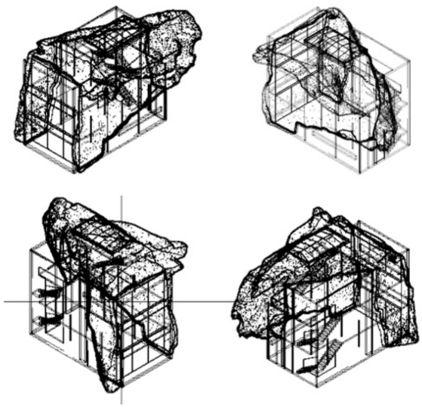


Fig. 6 – Cap-aux-Oies, Axonometric views, 2023.

Bachelard recalls that “fire is linked to tekhnai, to the arts and techniques of fire and, consequently, it becomes a civilizing fire.” [16] In his book *The Flame of a Candle*, Bachelard wrote: “The flame, among the objects of the world which call for reverie, is one of the greatest image-makers. The flame forces us to imagine.” [17] Likewise, the research is centered on the potential of architecture to make us imagine and to become an operator of images, in the same way as the flame.

Fig. 7 – Burnt log, Feu, Villa Tekhnai, 2021-2023.

For the philosopher, what we see is that nothing compares to what we imagine. Relying on natural forms implies a greater degree of imagination, of fiction, and obscures the purely geometric approach. In the manner of Facteur Cheval, who saw beyond the material by closely observing a stone sculpted by the waters of the Galaure river, I engage in work centered on the material imagination. The digitized artifacts associated with water and fire serve as a basis for imagining habitats, while the aim of the work is to develop new techniques of computer-aided design methods through the study of sculpted forms like the *Palais Idéal*.

Fig. 8 & 9 – Feu, Villa Tekhnai, 2021-2023.

The investigation of these ideas led me to develop three prototypes. *Feu* (*Villa Tekhnai*, 2023) is the amalgam of a log, blown glass fish scales, and classic figures from the architectural vocabulary revisited through a play of proportions. The interior staircase, inspired by Jean Dubuffet’s *Tour des Figures* (1985-1988), is like a spiral ascent in a monumental snail shell. Using the concept of bricolage, defined as a “primitive” science, by Claude Lévi-Strauss, in *Wild Thought* (1962), [18] the result is an assemblage of fragments that belong to different moments in history. [19]

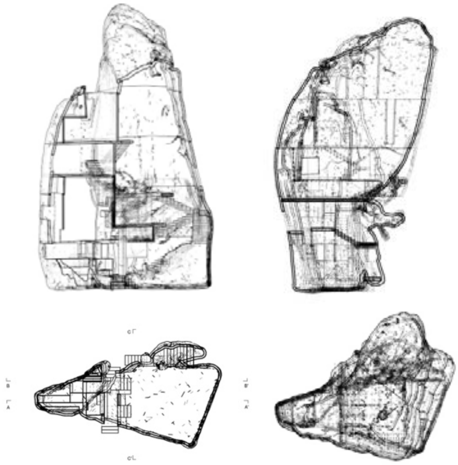


Fig. 10 – Section drawing and plans (first floor and roof), Troglodyte, Villa Tekhnai, 2021-23.

The *Troglodyte* prototype explores solid and empty volumes, extrusion through Boolean operations, and the differences between orthogonal prisms and the digitized shape of a burnt log. Guided by a play on the monumental, a spiraling staircase along the inner wall multiplies to infinity like a flame, dictating the layout of this disproportionate space.

Fig. 11 – Troglodyte (detail view), Villa Tekhnai, 2021-2023.

Fig. 12 – Troglodyte, Villa Tekhnai, 2021-2023.

The *Cap-aux-Oies* project integrates a stone transformed by the energy of the paleolozoic meteorite strike that created the giant Charlevoix crater, in Quebec, where it was gleaned. Evoking a past as old as 400 million years, amalgamated with a rectangular prism and scaled up to the size of a typical Montreal three-decker apartment building from the Plateau Mont-Royal neighbourhood, it expresses the idea of inhabiting the interior of a rock, relating to the original image of the cave habitat.

Fig. 13 – Cap-aux-Oies II (in progress, PAFARC, UQAM, 2024)

4 . Stone and fire

In continuity of *Villa Tekhnai* (2021-23), I extended the research from the element of fire to that of water. In a sculptural architecture, the thickness of the rock wall gives way to a set of stairs and antechambers, while the roof surfaces, designed for rainwater collection, create a waterfall on the rock façade. A digital hybrid of a rock from *Cap-aux-Oies* and a cantilevered block, the sculpture model is made of lime wood and PLA filament. It integrates the supports necessary for 3D printing: these are usually removed once the model is complete.

Fig. 14 – Studies for Cap-aux-Oies II, percussion stone and obsidian rock, PAFARC, UQAM, 2024. Photos: Émilie Presseault

The calculations for the positioning and size of these supports are integrated into the printing software. They can be temporarily separated from the model and manually adjusted thanks to a set of parameters. The project thus combines models based on scanning, digitization, and software calculation. The printed supports become integrated elements of the structural system, architectural elements that can adopt myriad shapes (adjusted, organic, rectilinear, etc.) before being reintegrated into the models. Through this process, it also becomes possible to imagine how the printed object would look at full scale without support.

As in the *Cap-aux-Oies* project, the *Cap-aux-Corneilles* prototype is made from the digitization of a percussion cone (also known as an impactite) that resulted from the direct impact of the Charlevoix meteorite. This stone fragment is placed on a central pillar that was calculated in organic mode by the PrusaSlicer software. The resulting set of tubes connects the stone to the ground and becomes the rainwater tank for domestic use. The walls of the building result from the aligned supports from 3D printing and were progressively adjusted through a back-and-forth process between printing, scanning, and modeling.

Fig. 15 – Cap-aux-Corneilles, Shatter cone on a 3D support system (PrusaSlicer: organic), PAFARC, UQAM, 2023-2024. Photos: Émilie Presseault

5 . Art Brut and Architecture

As it can be seen from the previous sections, my research-creation program is based on the notion of evolutive series, on the potential of printing media for boundless rotations and multiplications, and on the interpenetration of 3D scans and various superimposed media. Other explorations involve different minerals, such as obsidian, a volcanic stone. Importing such scans in software such as Rhino and Tinkercad gives access to all Boolean operations, such as subtraction and intersection. In terms of 3D printing, serial copying is no longer imposed as a production model, and digital machining is no longer limited to producing series of identical objects.

Fig. 16 – Shatter cone fragment, Cap-aux-Corneilles, Rhino model's view of the stone on 3D support system (PrusaSlicer: align), PAFARC, UQAM, 2023-2024.

The program also falls within the field of "architecture autre" ("other" architecture), coined by Reyner Banham (1955) along with the definition of New Brutalism, which corresponds to an architecture that excludes any historical or conventional cultural references: it abandons all concepts of composition, symmetry, order, module, and proportions. Also related to "Art Autre" ("Art of Another Kind", Michel Tapié, 1952) [20] and Art Brut, this architecture was to exclude any monumental symbolism and exceed "the norms of its expression with as much vehemence as Dubuffet's paintings exceeded the standards of painting." [21] If the connections between Art Brut and architecture are sometimes tenuous, they revive theories supplanted by rationalism and allow us to reconsider postmodernity through a different lens.

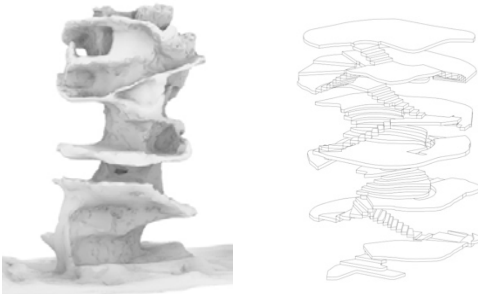


Fig. 17 – "Gastrovolve", Tour aux figures, Jean Dubuffet. Presentation model scanned with Polycam at the Fondation Dubuffet, Parc de l'île Saint-Germain, France.

6 . *Archisculpture*

Banham's article establishes the influence of Jean Dubuffet for the architectural discipline. In 2011, Daniel Abadie's work "Dubuffet Architecte" formulates and situates the architectural dimension developed by the artist between 1962 and 1974, and which, in 1982, earned Dubuffet the medal from the American Institute of Architects.

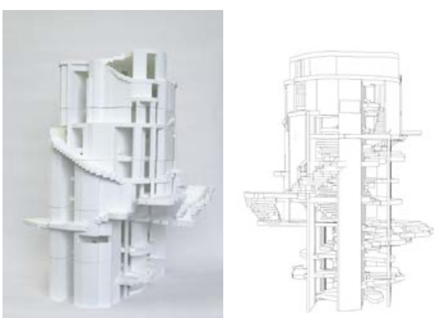


Fig. 18 – Tour 2, 3D model, PAFARC, UQAM, 2024. Double spiral staircase inspired by the Tour aux figures (Jean Dubuffet). 3D support system (PrusaSlicer: align).

His constructions, built on a monumental scale from his sculptures and drawings, are examples that recall certain notions of Art Brut applied to architecture and design. Let us name, for example, the *Jardin d'émail* (1974) in the De Hoge Veluwe national park, Netherlands, the *Closerie Falbala* (1971-1976) in Périgny, Val-de-Marne, France, and the afore-mentioned *Tour aux figures* (1985-1988) in Issy-les-Moulineaux, France. Although they cannot be considered as Art Brut productions per se, they provide strong glimpses of the potential of Art Brut notions for architecture.

Dubuffet's *Tour aux figures* is an enlarged version of a meter-high sculpture he created as part of the *L'Hourloupe* cycle in 1967. A 12-meter concrete structure covered in painted resin, with no openings or belvedere, the tower's black and white painted interiors feel like a cave. Classified as a historic monument in 2008, its restoration project was completed in 2020. After having gathered Dubuffet's texts, plans and models, I undertook the *Tour* project (CALQ, 2013-2014; MAQ, 2018) that was presented at the Maison de l'Architecture du Québec. The starting point of the project was a set of transcriptions and copies in 2D and 3D of architectural drawings showing numerous construction and layout details, after which the modelling work began. Despite the ambiguity of taking Dubuffet's work as a model, because of its positioning at the border between arts and architecture, and due to Dubuffet's official status as an artist, it was essential for me to make it the basis of an architectural experiment. At the end of the process, the coiling interior and sequence of spaces, stairs, and corridors, remain close to the original, while the object itself, with its imprecise and organic envelope and openings, deviates from it. The final model is made using Dubuffet's sculpture techniques (foam, coating, and plaster).

7. Conclusion

With emphasis on the digital representation of such architectures, the research reveals the manifold possibilities offered by exchange and dialogue between the disciplines of drawing, sculpture, design and architecture. It is worth noting that besides the research-creation work itself, multiple experiments allowed for the testing of new practices, techniques, and materials for 3D printing and scanning at different scales.

In the *Tour 2* project, inspired by the plans of “Gastrovolve”, another project by the same artist, the *Tour aux figures’* staircase is modelled, manipulated, and duplicated in order to create a double helix staircase; *Tour 2* integrates print supports calculated with PrusaSlicer in aligned mode. In 2024, I undertook new on-site scannings of parts of the *Tour aux figures* with Polycam software. The resulting documentation will lead to further versions integrating new information available since the 2020 restoration.

As for diffusion, the projects presented in this paper will be shown as part of the *Habitat outsider* exhibition at the Montcalm gallery in Gatineau (Quebec) in the fall of 2025, accompanied by a selection of previous projects.

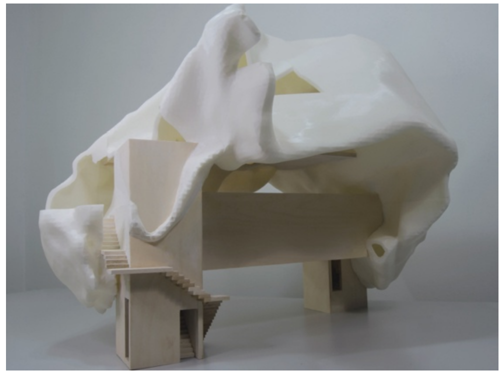


Fig. 19 – Papier 1.2, Villa Poïësis, CALQ, 2019-20.

All along this work, the links created between sculpture and architecture propose a critical and poetic look at the themes of creation and habitability. Through the hybridization of techniques and digital technologies in the creation of the prototypes inspired by *outsider architectures*, new axes of research regularly emerge, new possibilities constantly appear. The unexpected results clearly illustrate the richness and potential of these ventures that constantly challenge the boundaries of traditional architecture and sculpture.

References

[1] Gaston Bachelard, *La Terre et les rêveries du repos* [1948] (Paris: José, Corti, 2004),164.
[2] Ibid., 208-209.
[3] Gaston Bachelard, *La poétique de l’espace* [1957], (Paris : PUF, 2001), 126.
[4] Gaston Bachelard, *L'air et les songes – Essai sur l’imagination du mouvement* [1943], (Paris, Librairie José Corti, Livre de poche, biblio essais, 2004), 5.
[5] Jean-Claude Filloux, "Fabre, Michel. – Bachelard éducateur” (Paris: PUF,

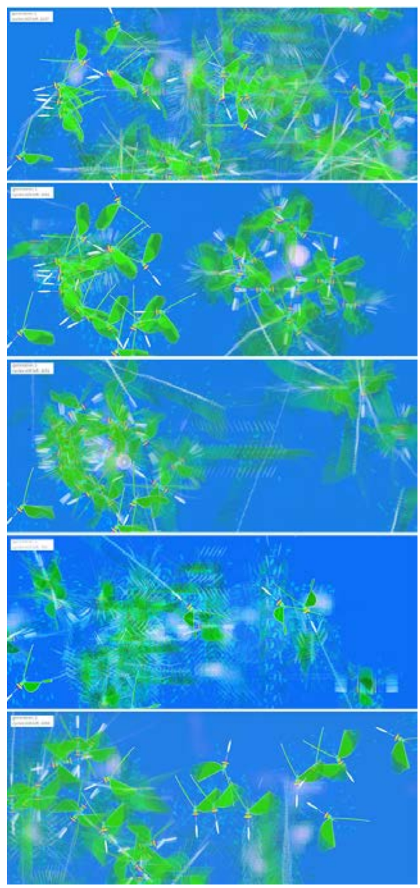
1985), 159. In *Revue française de pédagogie, Année* 1996, 117, no. 1, 158 – 162.
[6] Jean-Pierre Jouve, Claude Prévost, and Clovis Prévost, *Le Palais Idéal du facteur Cheva*l (Paris: du Moniteur, 1981), 119 and 298-299.
[7] Ibid., Letter from Ferdinand Cheval to André Lacroix in 1897, written out by his son Cyrille.
[8] Ibid., 263.
[9] Michel Ragon, *Où vivrons-nous demain?* (Paris: Robert Laffont, 1963).
[10] See Virginie Thiéry, "Jacques Couëlle : quand l'architecture se révèle sculpture. Monte Mano, la sculpture habitée". *Labyrinthe atelier interdisciplinaire*, 12, 2002, 97-106; Sonzogni, Valentina, "Frederick Kiesler et la maison sans fin". *L'Architecture d'Aujourd'hui*, France, 2003, no. 349, 53 and *Architecture sculpture*/hors collection, (Orléans: HYX, 2008).
[11] See Eva Kraus, "Plaidoyer pour une archisculpture—*Aujourd'hui* 53, 1966" *L'Architecture d'Aujourd'hui*, *op. cit.* 58-63 and Jean-Clarence Lambert, *2L'imagination matérielle* (Paris: édition du Cercle de l'Art, *Diagonales*, 1991), 138.
[12] Lucienne Peiry, "Anarchitecte." Collective, *Richard Greaves: anarchitecte* (Milan, Italie, 5 Continents, 2005).
[13] Magliozzi, Marielle, *L'art du bricolage dans l'architecture marginale de la seconde moitié du XXe siècle en France. Les créations de la « subculture »*, (PhD thesis, Université Aix-Marseille 1 – Université de Provence U.F.R. Art History and Archeology, 2003), 151 : « …Outsider art (utilisé par le critique d'art R. Cardinal dès 1971) ». (Outsider Art (used by the art critic R. Cardinal as early as 1971).
[14] Roger Cardinal, *Outsider Art* (London: Studio Vista, 1972). Cardinal introduces the term "outsider art," as an equivalent to the term "Art Brut." Lucienne Peiry, *L'Art Brut*, Coll, Tout l'art Histoire (Paris: Flammarion,171). Peiry notes, "Outsider art is the best English equivalent of the untranslatable term "Art Brut" […]. Obviously, outsider art uses a reference to designate its difference, whereas "Art Brut" is a name in itself." In Peiry, Lucienne, *L'Art Brut*, op. cit., note 5, 279.
[15] www.katherinelapierre.com
[16] Jean-Philippe A. Beaudin, *La figure mythique de Prométhée dans la philosophie de Platon*, Master's in Philosophy thesis, 2011, UQAM, 7.
[17] Gaston Bachelard, *La flamme d'une chandelle* (Paris: P.U.F, 1961), 1.
[18] Claude Lévi-Strauss, *La pensée sauvage* [1962] (Paris: Plon, Collection Pocket Agora no. 2, 2010).
[19] See Marielle Magliozzi, *Art Brut, architectures marginales – un art du bricolage*, (Paris : L'Harmattan, coll. L'Écarlate), 73. See Claude Lévi-Strauss, *La pensée sauvage*, [1962] (Paris: Plon, 1985), 30-49.
[20] Tapié, Michel, *Un art Autre* (Paris, Gabriel-Giraud & fils, 1952). See Banham, Reyner, "The New Brutalism". *The Architectural Review*, no. 708, 1955, 355-361.
[21] Reyner Banham, *The New Brutalism: Ethic or Aesthetic*, trans. Angélika Walbaum and Pierre-F. Walbaum (New York: Reinhold Publishing, 1966).

Impressions of Aquascapes with Reflections of Spring

Associate Professor Daniela Sirbu, PhD
Department of New Media, Faculty of Fine Arts
University of Lethbridge, Canada
www.ulethbridge.ca/fine-arts/profile/dr-daniela-sirbu
e-mail: daniela.sirbu@uleth.ca
.

Abstract

This is a multi-panel art piece generated by a running artificial life system developed by the artist. The system evolves the behaviours of a population of drawing agents, continuously creating plant-like structures by loosely emulating processes of structure formation in living and non-living systems in the natural world. Gradually, the system evolves to create visual designs with aesthetic value, from which a selection has been made to create this series. The idea behind this artificial life system is to create a space with aquatic qualities while dynamic elements suggestive of the spring season, are also present. The juxtapositions are meant to explore ways in which the familiar can become ineffable while drawing interest towards the creation of aesthetic spaces that breeze artificial life.



1. Introduction

This art paper takes the form of an expanded artist statement to explain more directly the background, the medium, the general approach of the artist to her later work, and how all this applies to the meaning and the development process of the series *Impressions of Aquascapes with Reflections of Spring*. As an expanded artist statement, there is a strong element of self-reflection that explains the personal background from which this art series has evolved.

2. Background

During the last 15 years I have been struggling to situate my art production into the world. As the time went by, this problem and the questions it rises have become more ardent instead of being alleviated. I started to question the question itself. I concluded that more important is what I have to say, and if what I want to say is sufficiently intense. The question has become: Is what I want to say strongly connected with my core as to make it alive through a complicated creative process and through all the technology it involves? At the end of this process, is what I want to say meaningful to others? This is an old question, in a world that has become very fast changing.

What I want to say as an artist is very simple. It is about the beauty of the world, which is recognized by all. I find very alluring the light of a sundown miraculously transforming the objects, natural or man-made, into their soft and warm counterparts. I find the first snow of winter to be miraculous when transforming the world into its crispy cold white version of itself, a version that lost its warmth and seems silenced and distanced from us. At the beginning of spring, I find the new life and the colours that emerge from the dying cold of the winter to be immensely mysterious in their making. This beauty that comes into being in various forms, with mystery eternally embedded in its processes of bringing things into being, is what I find to be most alluring of all. The beauty of these processes is what my work has revolved around lately.

I worked for a long time to develop a computational system that is artistically creative and therefore capable of bringing aesthetic objects into being [1], [2]. This is in the form of an artificial ecosystem based on the GA computational paradigm. In this ecosystem, abstract forms aggregate in time into aesthetic structures that resemble but are never exactly like natural living forms in the real world. In this sense, I come close to my love of seeing beautiful forms coming into being.

I believe this system does produce visual compositions, or that is what empirical data is saying when I put under scrutiny the output of this ecosystem of artificial life. But the interest in the artificial creation process in development has emerged as even more meaningful than the final output.

3. Motion as a Process of Creating Living Forms in an Artificial Ecosystem Driven by Aesthetics

While watching the system running, I was surprised to discover that I am drawn in, more than anything, or more than even the final output, by the visualization of the process of creating these living artificial structures. This is a process driven by the movement of the artificial drawing agents in the system. It reminds me that at the core, I am an animator. I am entranced when I see expressive motion be it natural or artificial. In this computational system, the process of creating forms is carried out by drawing agents in motion. What is novel about the animated motion of the agents is that it shows the structures coming to live while continuing to actively breeze artificial life into structures already formed in the ecosystem. All this is expressed through animated motion. The still frames captured from the system (Figures 1, 2, and 3) illustrate well the motion dynamics in the system.

The artificial drawing agents have emerged as independent creatures and I feel somehow excluded from their workings although I am the designer of the synthetic world. This brings an element of surprize every time when running the system.

However, I can act on the environment in which these artificial drawing agents live. I can change the environment and sometimes in ways that are cataclysmic for the artificial life within it. This way I can force the agents to react and adapt to the reconfigured environment and speed up the direction of their evolution towards new visual designs.

All this evolutionary process is expressed through animated motion, and it is novel for the viewer because it can be seen in a way that is difficult to see in the real world due to the large span of time that evolution usually involves.

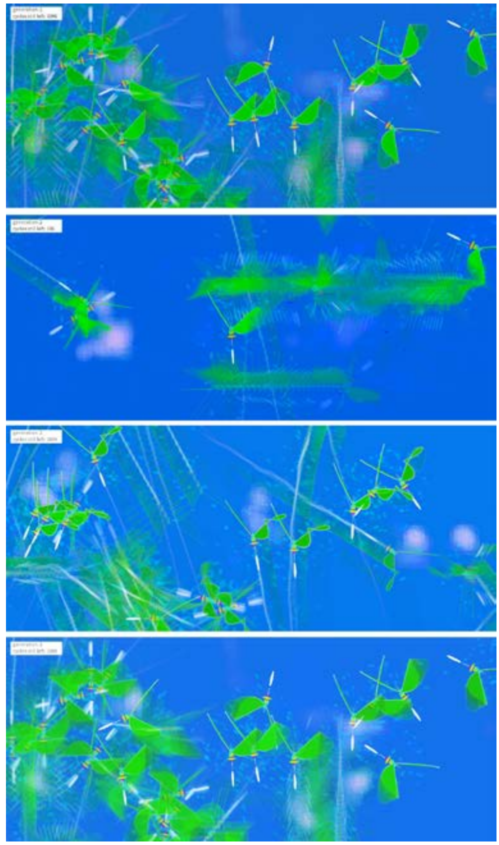


Figure 2. Stills captured from live running of computational creativity system. Artwork Title: *Impressions of Aquascapes with Reflections of Spring*. Artist: Daniela Sirbu. Software development by the artist.

Similarly, for example, when life emerges in spring, we have quick snap shots of the process as we go by. One can be in awe of the beauty of fragile blossoms and later see the beauty of the full blossoms, and later can see how the blossoms recede, dry and disappear to be replaced with other structures as the season progresses. A tree changes in this process but the change cannot be grasped at once. The computational system allows seeing the structure creation and disintegration and the breath of artificial life ongoing in between. Figure 3 shows this aspect in several still frames captured from the live running system.

The general intention is to allow the viewer the novel opportunity to see forms coming to life, even if that happens within an artificial ecosystem.

The deeper meaning is rooted into the original intention of developing a computational system that is artistically creative and therefore capable of bringing aesthetic objects into being. Creativity has been considered for a long time a specifically human endeavour. However, the analysis of the creativity process and its decomposition into many psychological processes involved, shows that creativity can be interpreted as an adaptive process, which is also specific to evolution [3], [4], [5]. Species recreate themselves into new forms that are more beautifully useful in response in a changing environment. Therefore, the process of change in nature can be considered creative at its core.

Deeply embedded into the computational system is the idea to attract the viewer by visualizing the artificial creative process. This is seen as a process of structure formation in an artificial ecosystem driven by aesthetic pressures as an equivalent of environmental pressures in the real world. It shows an entire ecosystem that

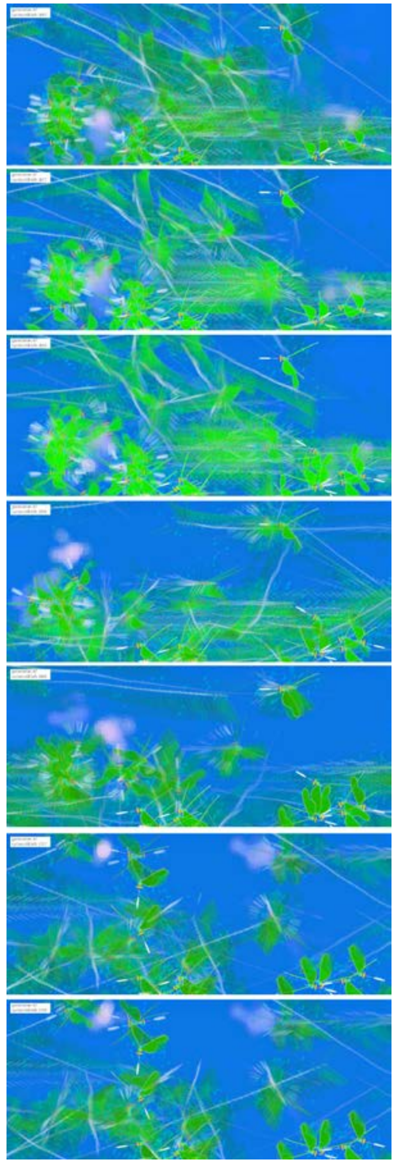


Figure 3. Stills captured from live running of system. Shows structure formation and disintegration. Artwork Title: *Impressions of Aquascapes with Reflections of Spring*. Artist: Daniela Sirbu. Software development by the artist.

evolves the behaviour of artificial creatures to aggregate living structures, that continue to breeze life, while being distributed in visual designs that are aesthetically sound.

4. Visual Design Concept

Overall, the *Impressions of Aquascapes with Reflections of Spring* is meant to bring the beauty of form creation to the audience. It is a space of its own that brings all at once both perceptions of an aquatic environments and the feeling of spring-like emerging life. It is an intentional juxtaposition that puts together two spaces that we usually see separately from one another into a new space, which acquires a life of its own. This new space is meant to attract the eye through its recognizable elements of waterscapes and spring blossoms, while inciting through unfamiliar simultaneous associations. While this approach might seem to have something in common with the surrealist art movement, it is very different from it because it aims to induce the feeling that it is naturally coming alive and naturally and continually creating structures that are based in their formation on processes that loosely emulate structure formation in the real world.

The space is meant to gradually evolve from one visual design of aesthetic value into another on a continuous basis. In this sense, the virtual world is designed as a space of dynamic aesthetics based on the king of beauty that is appealing in the natural world.

The *Impressions of Aquascapes with Reflections of Spring* is an artificial ecosystem. Drawing agents move abstract forms into a space characterized by a deep aquatic blue. This deep blue changes in time and temporarily retains the traces of the moving creatures, like in a water environment (Figures 1, 2, and 3), but things are slower and can be seen for longer in this artificial space. The changing deep green of the forming structures and the occasional splashes of soft pick forms are meant to bring a stroke of spring-like colour and motion that breeze artificial life into this virtual space (Figure 4).

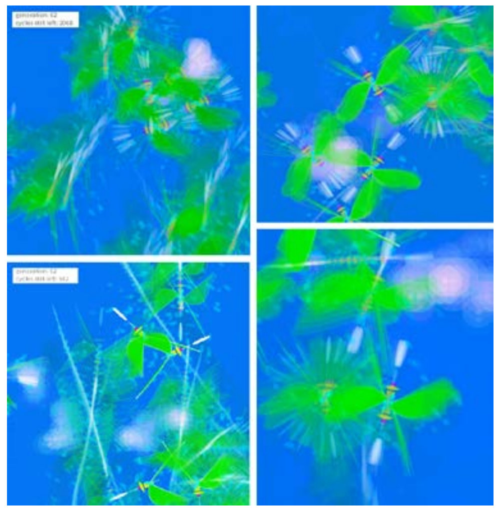


Figure 53. Details from stills captured from live running of system. Shows colour treatment. Artwork Title: *Impressions of Aquascapes with Reflections of Spring*. Artist: Daniela Sirbu. Software development by the artist.

Moving agents leave temporary traces that contribute to increasing the active dynamics of the space and directly contribute to the visual designs that emerge in time. The moving body parts of the agents enrich the texture of the structures created by these agents.

5. Conclusion

Beyond the artistic intensions of the series *Impressions of Aquascapes with Reflections of Spring*, there is the problem that contemporary artists are confronted with in a landscape where artificial intelligence seems to take over the sacred space of artistic creativity. When the system is designed and developed by the artist, a case can be made that the original production of the system is still the creation of the artist. But another case can be made that the drawing agents and the overall system takes a life of its own. It can never be accurately predicted where the system leads exactly. However, the overall nature of the motion and the way the system operates with forms and texture leads to visual compositions and a dynamic that have a consistent movement and visual style. The artist, therefore, creates a blueprint of artificial aesthetics capable to generate numerous art pieces with a level of detail and complexity that are impossible to reach for artists using traditional means of artistic expression.

References

[1] D. Sirbu and I. Dumitrache, "A Conceptual Framework for Artificial Creativity in Visual Arts " *International Journal of Computers Communications & Control* vol. 13, no. 3, 2017, doi: 10.15837/ijccc.2017.3.2759.

[2] D. Sirbu and I. Dumitrache, "Artificial Creativity with Kinetic Drawing Systems Based on Stochastic Motion," *U.P.B. Scientific Bulletin,* vol. 79, no. 2, Series C, pp. 37-48, 2017. [Online]. Available: https://www.scientificbulletin.upb.ro/rev_docs_arhiva/full1d3_574837.pdf.

[3] C. Martindale, *The Clockwork Muse: The Predictability of Artistic Change*. New York, NY: Basic Books, 1990.

[4] D. Simonton, *Origins of Genius. Darwinian Perspectives on Creativity.* New York, NY: Oxford University Press, 1999.

[5] C. Martindale, P. Locher, and V. M. Petrov, *Evolutionary and Neurocognitive Approaches to Aesthetics, Creativity, and the Arts*, Amityville, N.Y: Baywood Publishers, 2007.

Animation techniques using L-Systems

Prof. Umberto Roncoroni MSc, PhD
Department of Graphic Design, Universidad Peruana de Ciencias Aplicadas
https://umbertoroncoroni.com/
e-mail: umberto.roncoroni@upc.pe

Abstract

This paper describes some improvements in L-Systems' vocabularies and rules for animation purposes. L-Systems, a class of Shape Grammars, can generate complex forms of thousands of elements using a symbolic recursive substitution process. L-Systems are widely used in generative art or parametric design, but their possibilities for animation are not fully developed. The existing implementations of L-Systems are not user-friendly and their linguistic assets and tools for 2D or 3D animation are limited. However, L-Systems models' geometric properties and transformations can be easily parametrized, since rules can use symbols to modify the size, position, rotation angles, and color of every particle of the model. Still, the main problems during the animation are the dynamic evolution of rules and how to edit and update them in runtime. The solution described here consists of automatic transforming rules and symbols' substitutions, interface design layouts that facilitate the animation of grammars and 3d objects, and functions to render and export the models to commercial 3D packages. The benefits consist of the facility to animate complex modular forms or large groups of particles, which is impossible to do with actual animation applications, even using shaders. These techniques can improve the generative design possibilities of artists, educators, and researchers.

Introduction

L-Systems are well-known String Rewriting Substitution algorithms, akin to Shape Grammars [1], originally developed by Lindenmayer [2], to simulate plants' growth process [3]. The L-Systems generative procedure is based on symbols and simple substitution rules that can generate very complex fractal forms. L-Systems are also being developed for architecture, generative art, and music [4].

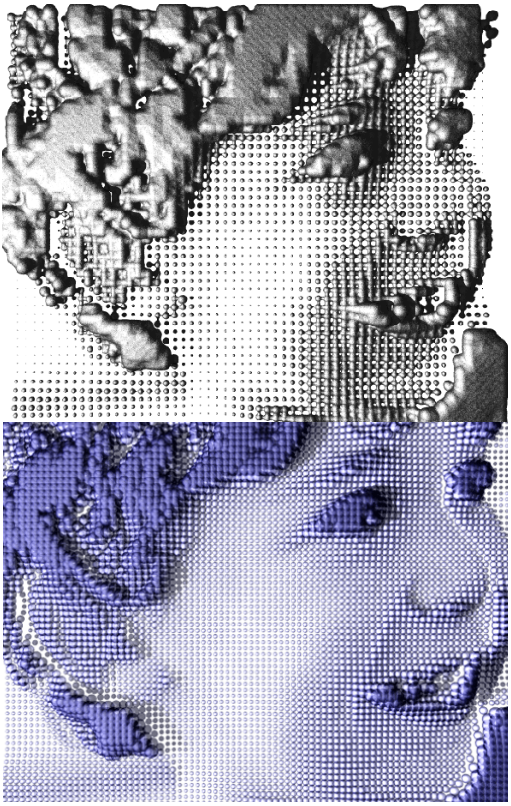


Fig. 1. 3D image processing with L-Systems.Photographs processed with L-Systems and rendered with POVRay organic 3D objects.

However, available L-Systems applications are not user-friendly and do not provide specific tools and interfaces for animation production [5,6]. For instance, to create animations of complex particle systems [7], video special effects, or to experiment with creative image processing in 3D (Fig. 1).

L-Systems' advantages

L-Systems are very efficient generative design algorithms that can create huge modular objects, like fractals, plants, particle systems, and multicellular organisms [8,9] with thousands of elements (Fig. 2). Using L-Systems' symbols, rules, and substitution processes make it relatively easy to edit and transform complex objects and simulate growing and morphing progressions.

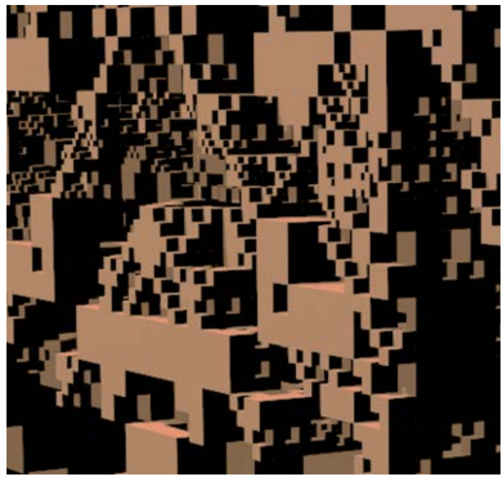


Fig. 2. Self-similar modular object made with L-Systems.

In time, many improvements to the original algorithm have been developed [10], like parametric or context-sensitive L-Systems, which modify positions, size, form, colors, and transparency of every element of the model using external references, or timed L-Systems [11], which can change behaviors in specific steps of the process.

With the abovementioned L-Systems upgrades, designers and artists can explore new animation techniques, with different styles and special effects [6]. However, these additions are not enough to exploit the full possibilities of L-Systems.

In this article, we will show some solutions to improve L-Systems' for animation purposes: additions to the original Lindenmayer grammar (vocabulary and rule set) and more efficient L-System applications' interface designs.
These improvements are included in L-SystemGenerator, the software developed during this and previous generative art researches.

Animation processes with L-Systems

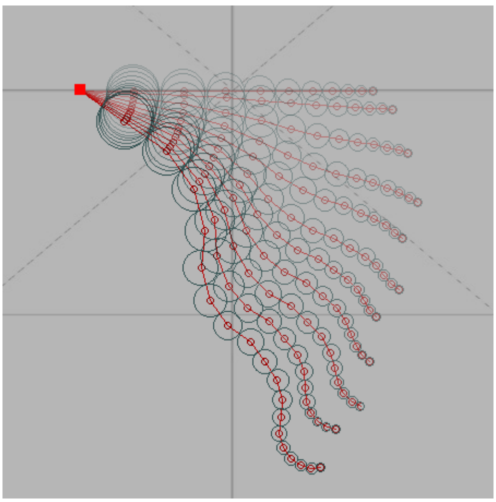


Fig. 3. Example of object's growing process.

To animate objects with our L-System application, the first task is to create the object generative grammar, inserting different symbols and rules for every step of the substitution process (Fig. 3); the second task is to set the animate geometric properties of the elements; the last task consists of generating and exporting the model in a format compatible with animation packages, usually the STL format.
The rendering is done by an external 3D application, we used POVRay, a freeware rendering software with raytracing capabilities.
L-SystemsGenerator automatically generates POVRay files during the substitution process.

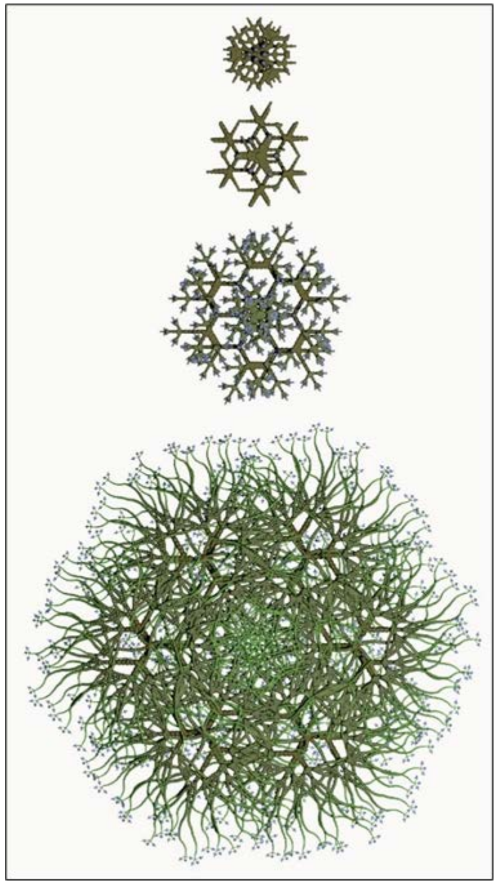


Fig. 4 Growing process of a complex system with 27,000 3D particles and 19,000 different rotation values.

POVRay saves the rendered models in .png or .bmp formats, which are finally imported into digital painting applications as cell animation frames. This process is relatively simple and can process hundreds of renders quickly.

The simplest animation technique consists of changing symbols, rules, and geometric parameters to modify the models for every animation frame. The challenge is how to use L-Systems grammars to control these properties and how to create effects and transformations that are almost impossible to do with standard 3D packages, especially in the case of growing and moving particle systems (Fig. 4) or modular objects.

Animation tools and interface

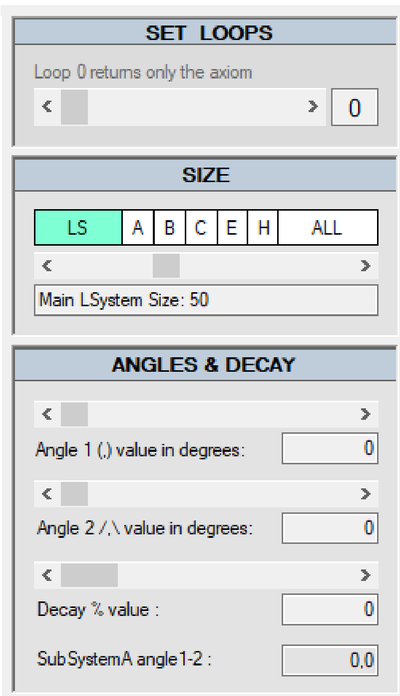


Fig. 5. The L-SystemsGenerator's interface menu changes the geometric properties of the system or its parts: size, angles, and scaling (Decay), according to the substitution step (SET LOOPS).

In this paragraph, we will describe the programming techniques and interface designs that we experimented with to improve and speed up animation procedures with substitution processes and L-Systems.

The first solution is to improve the symbols' properties and the interface design to facilitate the setting of geometric transformation parameters. In L-SystemsGenerator, the objects' geometric properties are animated using special symbols whose values are changed and updated interactively or automatically during the L-Systems' substitution/modeling process. This method is also used to control random values.

These symbols can also modify the geometry values according to the context, in other words, considering the neighborhood (previous and next symbols of the current object's grammar) (Fig. 5).

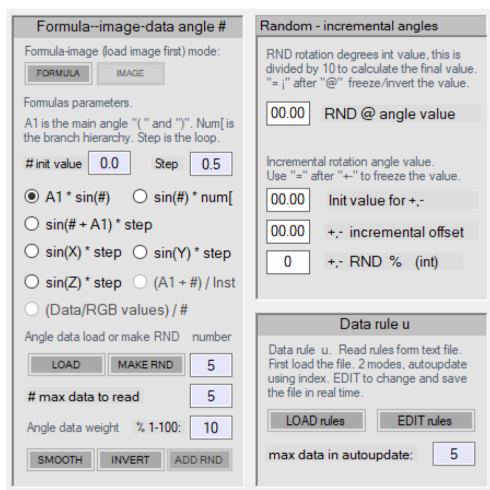


Fig. 6. L-SystemsGenerator menu to edit animatable angle equations, parameters, random values, and databases.

In the rotation angle case, math formulas and data sets can be used for precalculated rotation values (Fig. 6),

whose variables upgrade automatically in every step of the animated substitution process.
The second solution is to animate the grammar vocabulary and substitution rules. Since transformations are determined by grammars, the parameter updating animation technique seen before is not enough, so we explored the possibility of modifying the substitution process symbols and rules for every keyframe of the animation. This option makes it possible to change in time the form of the system completely.
However, this technique is more difficult, since it is necessary to edit and compile the grammar automatically with complex text processing functions. A practical solution is to create different rule' sets for the animation's keyframes. In the current application, we implemented 10 keyframes (Fig. 7), that can be extended using in-between techniques like loops, rule arrays with automatic append functions, or creating rule libraries (Fig. 6) which data can be automatically generated, updated, and optimized for the animation development.
The third solution is to use subSystems, which are L-Systems embedded in the main system. The subSystems' advantage is that can create models as compounds of sub-models with specific geometric properties and grammars that can be animated separately. For instance, to grow or move the branches of a tree or the cells of virtual organisms (Fig. 4).
The third solution is to add to the standard L-Systems grammars high-level macro rules, programmable symbols, and "ghost" symbols. These are symbols without geometry but with substitution parameters and properties that can affect other rules and symbols, a technique that has already been discussed in previous GA conferences [12].
The third solution is to add to the standard L-Systems grammars high-level macro rules, programmable symbols, and "ghost" symbols.

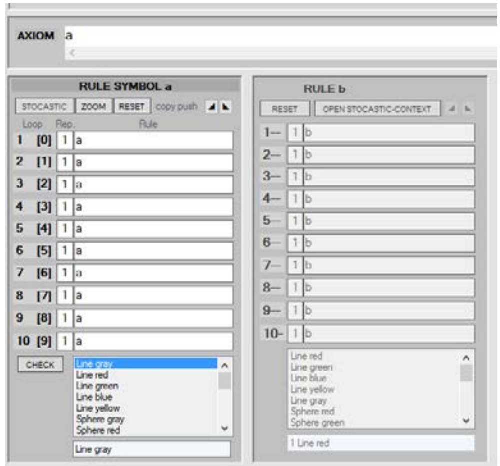


Fig. 7. L-SystemsGenerator menu to set rules of 10 keyframes. The left column sets how many times the rule can be repeated, like in a for-next routine.

These are symbols without geometry but with substitution parameters and properties that can affect other rules and symbols, a technique that has already been discussed in previous GA conferences [12].
For instance, these symbols can add a new symbol when the substitution process finds one of its occurrences, like append values in arrays or lists; in this way, we can increase a group of symbols during the animation.
On the contrary, deleting symbols is used to decrease the instances of chosen symbols. The "Ç" rule interface lets us define which symbols and geometric transformations must be deleted. In the example of Figure 8, the symbol "Ç"

deletes the symbol “a”, but not “+” or “:”.
The fourth solution is to use image RGBA values to change rules and objects’ parameters like 3D packages do with Z-buffers and bitmap masks.

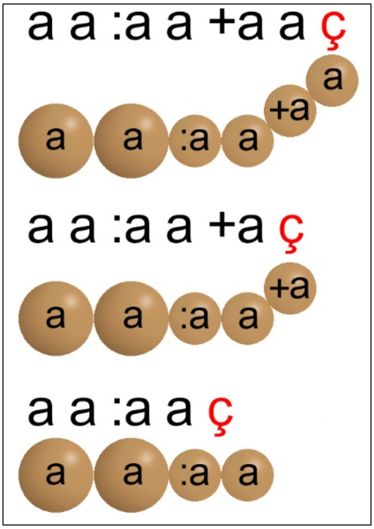


Fig. 8. Symbol “Ç” is set to delete the previous symbol “a”. It continues the deleting process until a specific threshold is reached.

In this case, the first step is to import a bitmap and position the L-Systems elements over it, for instance, starting from the top left corner of the bitmap, and then use the corresponding pixels’ RGB values to modify the rules or the geometry parameters of the symbol (in this case objects are small spheres) (Fig. 9). To translate the image pixels into 3D objects with L-SystemsGenerator it is possible to use primitives like spheres or cubes, but also complex object compounds controlled by subSystems, which elements can react to RGB values and animated independently.
This makes it possible to animate textures, liquid or burning image effects, or create 3D video transitions between video clips, using animated Z-buffers obtained, with the abovementioned procedures and techniques, from videos or cell animation frames.

Fig 9. Top: Setting the starting position. The object’s size determines the sample rate. Bottom: detail of the final L-System effect.

With L-Systems, it is possible to create, animation effects like those made with generative AI algorithms.

Future research

As we have explained, L-SystemsGenerator's animation tools provide automatic parameters and substitution rules update, subSystems, and Z-buffer capabilities (Fig. 6, 7). However, considering that manual editing and parameter input are still necessary, the first improvement will be to include new interface controls to facilitate the real-time updating of animation parameters, which will avoid keyboard inputs that are prone to typing errors.

The second improvement will be implementing auto-save functions with automatic batch processing and updates of filenames.

The third improvement will be the development of automatic functions to check and change the objects' geometric properties in case of overlapping or nested elements, saving memory, speeding up the animation process, and creating emergent forms with more natural movements.

Conclusions

As the literary review demonstrated, L-Systems have a promising potential for animation. On the other hand, the survey of existing applications shows that the biggest challenge for any L-System application is designing better and more efficient interfaces, since grammars and parameters can be very complex and difficult to manipulate, especially in the animation's case.

In our research, we experimented with new algorithms, generative symbols, substitution procedures, and interface designs to improve the animation process.

The computational principles and solutions described and implemented in our application can be easily reused in any L-Systems project.

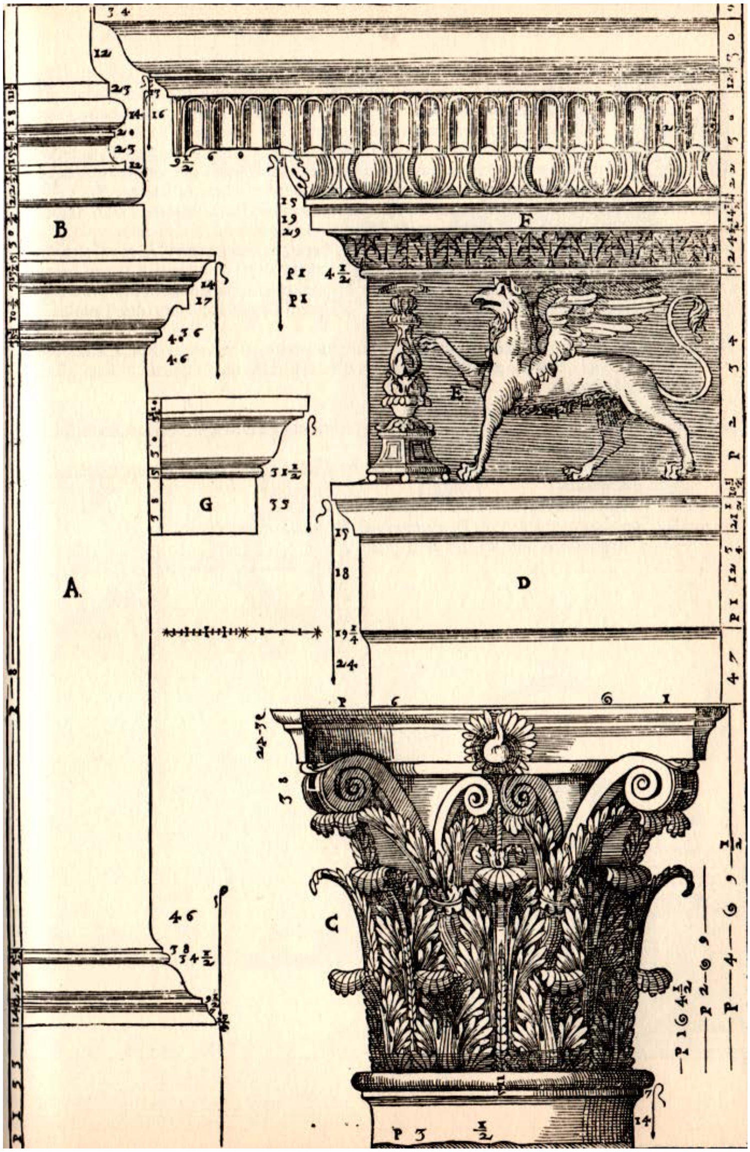
Finally, a free L-SystemsGenerator version of research or educational purpose, and examples of its possibilities, can be downloaded at https://www.digitalpoiesis.org/ and https://umbertoroncoroni.com/.

References

1. Stiny, G., and J. Gips. (1972). Shape Grammars and the Generative Specification of Painting and Sculpture. In O. R. Petrocelli (ed.), *The Best Computer Papers of 1971*, pp. 125–135.
2. Lindenmayer A. (1968). Mathematical models for cellular interaction in development I. Filaments with one-sided inputs. *Journal of Theoretical Biology* 18, pp. 280-289.
3. Prusinkiewicz, P. and Lindenmayer, A. (1990). *The Algorithmic Beauty of Plants*. Berlin: Springer.
4. McCormack, J. (2003). *The Application of L-systems and Developmental Models to Computer Art, Animation, and Music Synthesis*. Clayton: School of Computer Science and Software Engineering, Monash University.
5. Noser, H., Thalmann, D., and Turner, R. (1992). Animation Based on the Interaction of L-Systems with Vector Force Fields. In: Kunii, T.L. (eds) *Visual Computing*. CG International Series. Springer, Tokyo.
6. Hampshire, N. (2009). *Dynamic Animation and Re-Modelling of L-Systems*. Master's Thesis MSc Computer Animation and Visual Effects N.C.C.A Bournemouth University.
7. Orth, T. (1993). *Graphical modeling using L-systems and particle systems: a*

comparison. Master's thesis, University of Calgary, Calgary, Canada. doi:10.11575/PRISM/12245

8. Kelty, C. and Landecker, H. (2004). A Theory of Animation: Cells, L-systems, and Film. *Grey Room* 17, pp. 30–63.
9. Hornby, G. and Pollack, J. (2001). Evolving L-systems to generate virtual creatures Lab. *DEMO Computers & Graphics* 25, pp. 1041–1048.
10. Roncoroni, U. (2022). Programming Complex 3D Meshes. A Generative Approach Based on Shape Grammars. In Soddu, C. (ed.) *XXV Generative Art Conference Proceedings*, Rome.
11. McCormack, J. (2004). Generative Modelling with Timed L-Systems. In J. S. Gero (ed.) Design Computing and Cognition. Berlin: Springer, pp. 157–175.
12. Roncoroni, U. (2020). Introduction to Generative Musical Morphing. In Soddu, C. (ed.) *XXIII Generative Art Conference Proceedings*, Rome.



Andrea Palladio, 1570

POSTERS

The Beauty of Rotations in Processing

Prof. dr. em. Jan Paredaens
University of Antwerp, Belgium
www.JanParedaens.be
e-mail: jan.paredaens@uantwerpen.be

Abstract
Processing is a programming language, that is built on top of the Java programming language. We discuss several constructs and applications in Processing concerning rotations and translations. In the two-dimensional case we discuss among others rotations around an arbitrary point and around two arbitrary points. In the three-dimensional case we discuss rotations around an arbitrary line through the origin, in the XY-plane and a real arbitrary line. We also give some interesting applications.

1. Processing

Processing is an open-source, versatile programming language and environment specifically designed for visual artists, designers, and beginners to learn and create interactive digital art and visualisations [1,4]. It is based on Java, simplifying complex coding tasks and providing a wide range of libraries and tools for easily creating stunning visual content [2,3].

Every program in Processing describes a two-dimensional or a three-dimensional environment. There are no functions in Processing to rotate around an arbitrary point or line.

We will not give complete Processing programs, but we will focus on the important constructs for the transformations we discuss. Knowledge of Processing and programming languages in general is not necessary.

2. Rotation in a two-dimensional environment

In two dimensions Processing uses the coordinates of the two-dimensional environment (C2D) [6].

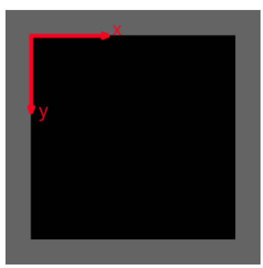


Fig.1

To start, the C2D has the origine in the left upper corner, the x-axis directed to

the right and the y-axis directed down (Cfr. Fig. 1).

translate(x,y) shifts C2D.

rotate(a) rotates C2D around the origin of C2D clockwise with an angle a.

icon(ix,iy) draws the icon on coordinates (ix,iy).

(px,py) are the coordinates of point p.

2.1 Rotation around the origin

Suppose we want to rotate an icon around the origin of C2D, clockwise a radians. We substitute icon(ix,iy); by

rotate(a); icon(ix,iy); rotate(-a);

rotate(-a); will put C2D in its original position.

As an example

rotate(radians(30)); icon(60,30); rotate(radians(-30));

will change Fig.2 into Fig.3, where a is 30 degrees and (ix,iy) is (60,30).

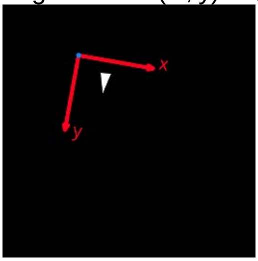


Fig.2

Each time C2D is drawn in red. The latter is not shown on the screen.

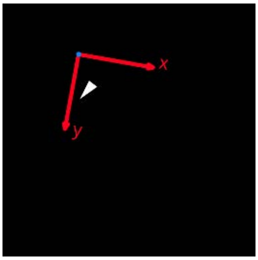


Fig.3

2.2 Rotation around an arbitrary point

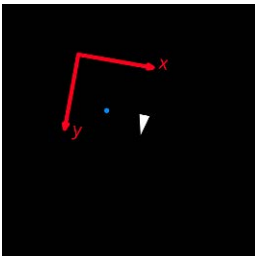


Fig.4

Suppose we want to rotate an icon around an arbitrary point p, clockwise a radians. We substitute icon(ix,iy); by

translate(px,py);
rotate(a);
icon(ix-px,iy-py); rotate(-a); translate(-px,-py);

As an example

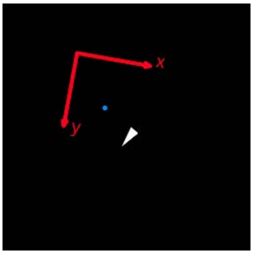


Fig.5

2.3 Rotation around an arbitrary point, bis

Suppose we want to rotate an icon around an arbitrary point p, clockwise a degrees, but such that the icon is stable. This is illustrated by transforming Fig.4 into Fig.6. We substitute icon(ix,iy); by

```
translate(px,py);
rotate(a);
translate(ix-px,iy-py);
rotate(-a); icon(0,0); rotate(a);
translate(-ix+px,-iy+py);
rotate(-a);
translate(-px,-py);
```

As an example

```
translate(75,100);
rotate(radians(30));
```

```
translate(75,100); rotate(radians(30));
icon(150-75,100-100); rotate(radians(-
30)); translate(-75,-100);
```

will change Fig.4 into Fig 5, where a is 30

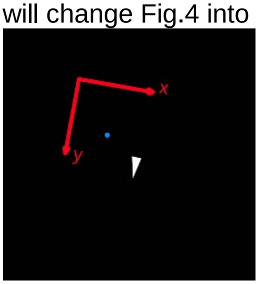


Fig.6

degrees, (px,py) is (75,100) and (ix,iy) is (150,100).

```
translate(150-75,100-100);
rotate(radians(-30)); icon(0,0);
rotate(radians(30));
translate(-150+75,-100+100);
rotate(radians(-30)); translate(-75,-100);
```

will change Fig.4 into Fig 6, where a is 30 degrees, (px,py) is (75,100) and (ix,iy) is (150,100).

2.4 Two rotations around two arbitrary points

Suppose we want to rotate icon(ix1,iy1) around p1 with rotation angle a1 and icon(ix2,iy2) around p2 with rotation angle a2. We substitute

```
icon(ix1,ix2);
icon(ix2,iy2);
```

by

```
translate(p1x,p1y);
rotate(a1);
icon(ix1-p1x,iy1-p1y);
rotate(-a1); translate(-p1x,-p1y);
translate(p2x,p2y); rotate(a2);
icon(ix2-p2x,iy2-p2y);

rotate(-a2);
translate(-p2x,-p2y);
```

```
rotate(radians(-40)); icon(200-175,250-
150); rotate(radians(40)); translate(-175,-
150);
```

will change Fig.7 into Fig.8 where a1 is 30 degrees, (p1x,p1y) is (75,100), (ix1,iy1) is (125,100), a2 is -40 degrees, (p2x,p2y) is (175,150) and (ix2,iy2) is (200,250).

As an example

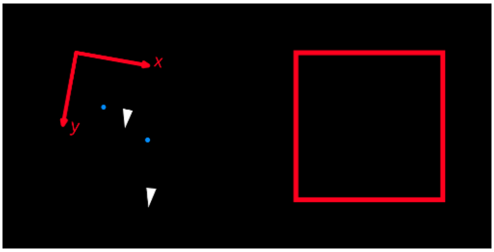


Fig.7 - Fig.8

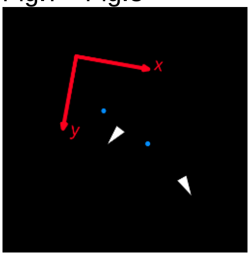


Fig.9

translate(75,100); rotate(radians(30)); icon(125-75,110-100); rotate(radians(-30)); translate(-75,-100); translate(175,150);

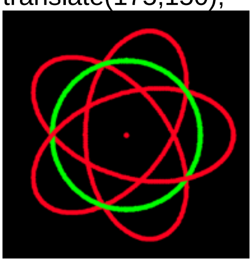


Fig.10

2.5 Applications

Fig.9, Fig.10 and https://vimeo.com/990947291 are realised using only the techniques above.

3. Rotation in a three-dimensional environment

icon(ix,iy,iz) draws the icon on coordinates (ix,iy,iz).

(px,py,pz) are the coordinates of point p.

3.1 Rotation around the x-axis

Suppose we want to rotate an icon around the x-axis of C3D, clockwise in the opposite direction of the x-axis, a radians. We substitute

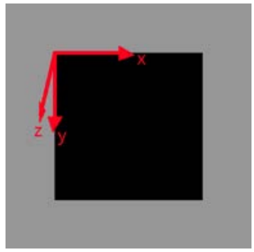


Fig.11

In three dimensions Processing uses the coordinates of the three-dimensional environment (C3D) [6].

To start, the C3D has the origine in the left upper corner, the x-axis directed to the right, the y-axis directed down and the z-axis directed to the reader (Cfr. Fig. 11).

translate(x,y,z) shifts C3D.

rotateX(a) rotates C3D around the x-axis of C3D clockwise in the opposite direction of the x-axis, with an angle a.

rotateY(a) rotates C3D around the y-axis of C3D clockwise in the opposite direction of the y-axis, with an angle a.

rotateZ(a) rotates C3D around the z-axis of C3D clockwise in the opposite direction of the z-axis, with an angle a.

icon(ix,iy,iz);

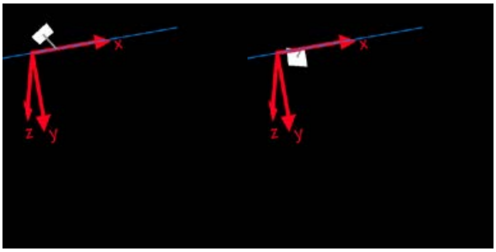


Fig.12 - Fig.13

by

rotateX(a); icon(ix,iy,iz); rotateX(-a);

As an example

rotateX(radians(95)); icon(50,0,0); rotateX(radians(-95));

will change Fig.12 into Fig.13, where (ix,iy,iz) is (50,0,0) and a is 95 degrees.

Analogously we can rotate around the y-axis and around the z-axis.

3.2 Rotation around a line in the xy-plane and through the origin

Suppose we want to rotate an icon around the line L in the xy-plane, through the points (0,0,0) and (Lx,Ly,0) clockwise in the direction (Lx,Ly,0) -> (0,0,0), a radians. We substitute

rotateX(-a);
rotateZ(-atan(Ly/Lx));

where atan() is a function that calculates the arcus tangent.

As an example

rotateZ(atan(200/300.)); rotateX(radians(-150)); rotateZ(-atan(200/300.)); icon(75,50,0); rotateZ(atan(200/300.)); rotateX(radians(150)); rotateZ(-atan(200/300.));

icon(ix,iy,iz);

by

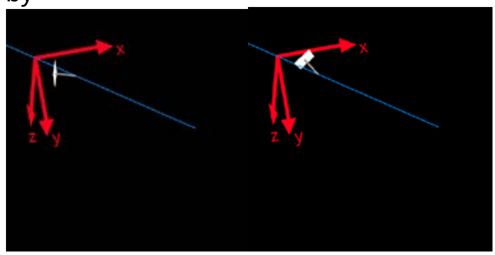


Fig.14 - Fig.15

will change Fig.14 into Fig.15, where Lx is 300, Ly is 200, (ix,iy,iz) is (75,50,0) and a is -150 degrees.

3.3 Rotation around a line through the origin

Suppose we want to rotate an icon around the line L through the points (0,0,0) and (Lx,Ly,Lz) clockwise in the direction (Lx,Ly,Lz) -> (0,0,0), a radians. We substitute

rotateZ(atan(Ly/Lx));
rotateX(a);
rotateZ(-atan(Ly/Lx));
icon(ix,iy,iz);
rotateZ(atan(Ly/Lx));

icon(ix,iy,iz);

by

rotateY(-atan(Lz/Lx);
rotateZ(atan(Ly/sqrt(Lx*Lx+Lz*Lz)));
rotateX(a);

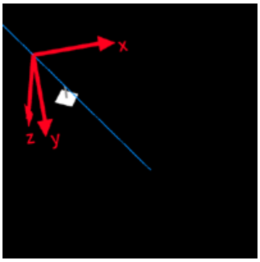


Fig.16

As an example

rotateY(-atan(-200/200.));
rotateZ(atan(300./sqrt(200*200+ (-200)*(-200)))); rotateX(radians(80));
rotateZ(-atan(300./sqrt(200*200+ (-200)*(-200))));
rotateY(atan(-200/200.));
icon(40,60,-40);
rotateY(-atan(-200/200.));
rotateZ(atan(300./sqrt(200*200+ (-200)*(-200))));
rotateX(-radians(80));
rotateZ(-atan(300./sqrt(200*200+ (-200)*(-200))));
rotateY(atan(-200./200));

will change Fig.16 into Fig.17, where Lx is 200, Ly is 300, Lz is -200, (ix,iy,iz) is (40,60,-40) and a is 80 degrees.

rotateZ(-atan(Ly/sqrt(Lx*Lx+Lz*Lz)));
rotateY(atan(Lz/Lx));
icon(ix,iy,iz);
rotateY(-atan(Lz/Lx));
rotateZ(atan(Ly/sqrt(Lx*Lx+Lz*Lz)));
rotateX(-a);
rotateZ(-atan(Ly/sqrt(Lx*Lx+Lz*Lz)));
rotateY(atan(Lz/Lx));

3.4 Rotation around an arbitrary line

Suppose we want to rotate an icon around the line L through the points (L1x,L1y,L1z) and (L2x,L2y,L2z) clockwise in the direction (L2x,L2y,L2z) -> (L1x,L1y,L1z), a radians. We substitute

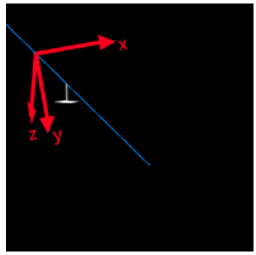


Fig.17

where sqrt() is a function that calculates the square root.

icon(ix,iy,iz);

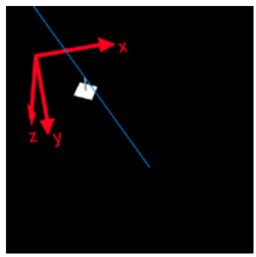


Fig.18

by

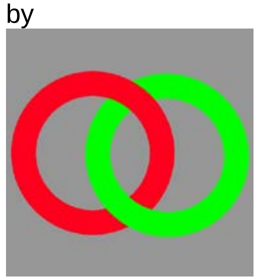


Fig.19

rotateY(-atan((L2z-L1z)/(L2x-L1x)));
rotateZ(atan((L2y-L1y)/sqrt((L2x-L1x)*(L2x-L1x)+(L2z-L1z)*(L2z-L1z))));
rotateX(a);

```
rotateZ(-atan((L2y-L1y)/sqrt((L2x-
L1x)*(L2x-L1x)+(L2z-L1z)*(L2z-L1z))));
rotateY(atan((L2z-L1z)/(L2x-L1x)));
icon(ix,iy,iz);
rotateY(-atan((L2z-L1z)/(L2x-L1x)));
rotateZ(atan((L2y-L1y)/sqrt((L2x-
L1x)*(L2x-L1x)+(L2z-L1z)*(L2z-L1z))));
rotateX(-a);
rotateZ(-atan((L2y-L1y)/sqrt((L2x-
L1x)*(L2x-L1x)+(L2z-L1z)*(L2z-L1z))));
rotateY(atan((L2z-L1z)/(L2x-L1x)));
```

As an example

```
rotateY(-atan((-200-200)/(200+20)));
rotateZ(atan((300+180)/
sqrt((200+20)*(200+20)+
(-200-200)*(-200-200))));
rotateX(radians(90)); rotateZ(-
atan((300+180)/ sqrt((200+20)*(200+20)+
(-200-200)*(-200-200))));
rotateY(atan((-200-200)/(200+20)));
icon(60,90,0);
rotateY(-atan((-200-200)/(200+20)));

rotateZ(atan((300+180)/
sqrt((200+20)*(200+20)+ (-200-200)*(-
200-200)))); rotateX(-radians(90));
rotateZ(-atan((300+180)/
sqrt((200+20)*(200+20)+ (-200-200)*(-
200-200))));
rotateY(atan((-200-200)/(200+20)));
```

will change Fig.18 into Fig.19, where L1x is -20, L1y is -180, L1z is 200, L2x is 200, L2y is 300, L2z is -200, (ix,iy,iz) is (90,60,0) and a is 90 degrees.

4 two-dimensions against three-dimensions

When we work in two dimensions everything we draw hides what was drawn in that place before. That is the reason why Fig.20 cannot be drawn using two circles. So when two pixels are drawn of the same place in two dimensions, only the pixel that is drawn

Fig.20

last is shown. In three dimensions on the other hand the pixel that is closest, ie. the one with the biggest z-value in shown. So Fig.20 can be drawn in three dimensions. Look also to

https://vimeo.com/997385192 for a three-dimensional example.

References

[1] Casey Reas, Ben Fry, Processing A programming Handbook for Visual Designers and Artists, The MIT Press, 2014.

[2] J. Paredaens, Instantaneous Deformatiions of Camera and Video Images, XXVI Generative Art 2023, p.174-179.

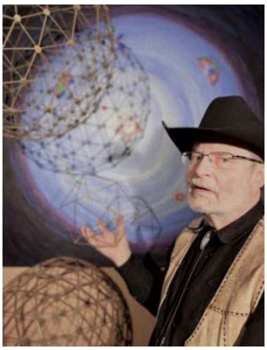
[3] http://JanParedaens.be . [4] https://processing.org/.

[5] https://www.eng.uc.edu/~beaucag/Classes/Properties/OptionalProjects/CoordinateTransformationCode/Rotate%20about%20an%20arbitrary%20axis%20(3%20dimensions).html

[6] https://www.youtube.com/watch?v=LSq3JlS9LcY

In, Out, Around - Revisited.

Curtis L Palmer, B.Sc., M.Des.
Synergetic Design, Edmonton, Alberta, Canada.
http://www.youtube.com/@Criterion5
e-mail: synergeticsam@gmail.com



Abstract

I have conversed with various computers, usually one bit at a time in an effort to represent the geometry central to R.B. Fuller's 1975 theory of everything, Synergetics[1], wherein he attempts to return conceptuality to science[2]. His precise poetic language recommends performing volumetric measurement with tetrahedra rather than cubes. Why?

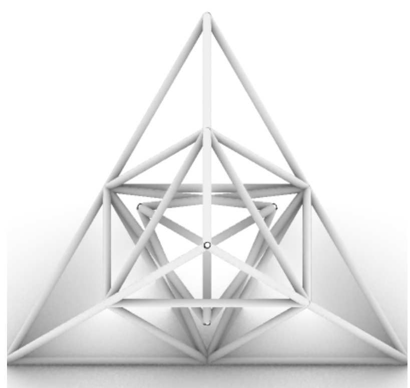
How to justify this shift in scientific measurement motivated my illustration aspirations. I began my computer aided diagrammatic effort to realize reason[3].

My efforts keep me apprised of a subset of computer graphics advances with which I have devised a few algorithms that if automated could produce infinite forms, some may be beautiful or useful. Maybe an AI would have the time to critique my infinitudes.

This presentation's title references my interpretation of a theme of Synergetics, In Out Around[4]. In three short animations I excerpted texts from Synergetics to be voiced by D. Ley with musical accompaniment by M. Reinhart.

In this animation I assigned an equal tempered scale to radii of twelve polyhedra with 5 fold symmetry arranged concentrically about a common origin. This polyhedral matroyshka doll served as still life for an animated camera sent on a series of *Dive Throughs.* The camera sees *Chords*, sets of polyhedra arranged concentrically, documenting in-sight, out-sight and 'round-sight of a sliver of a polyhedral inventory, Platonic, Archimedean, and Catalan.

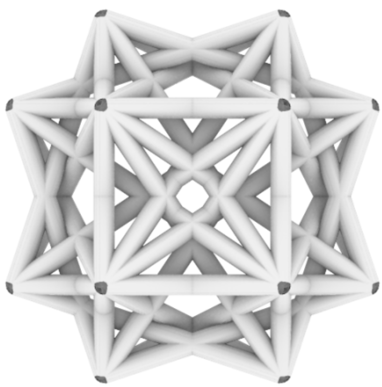
The animation's colour schemes resulted



from recursive interventions on the *Shadows* of the polyhedra being observed - mere digital artifice.

Divide & Concur Tetrahedron & Octahedron

For GA2024 I will reprise words used in my earlier tribute. I will read Synergetic's sections: 400.25, 517.05, and 814.01, while presenting images of a parametric tetrahedron being dissected by



successive en-vectorings.

I hope attendees will enjoy what I consider beauty emerging from this rethinking of volumetric analysis.

Divide & Concur further in.

With upgraded hardware and software I am able to iterate a procedure at the core of Synergetics. The results are worth documenting. They may even be beautiful.

How my pursuit-of-thought experiment relates to the conference topic is perhaps tangential. Preserving human complexity and heritage requires preserving humans. AI may help.

The noble objective expressed by the originator of synergetic geometry was to "Make the world work for 100% of humanity". He was adamant that understanding the geometry of nature, his geometry, would help the cause.

Now with drawings of a tetrahedral accounting, I am confident to voice its beauty, if not its truth.

To reiterate, for my GA2024 poster session, I offer my voice reading select text from the original, In Out Around, with perhaps, more meaningful graphics for comparison.

Perhaps this gesture will serve as a contribution that an AI will use to advance Fuller's objective of making the world work for all.

1. Algebra of Synergetics

Is there an algebra of synergetic geometry? An algebra defines a space and operations actionable in the space.

Synergetic geometry defines the space, a tetrahedron, which enthusiasts typically showcase with a two step recursive operation on that space: Rule 1, Edge Bisection, and Rule 2, Join Points. Each time I see this demonstration I notice exhibitors stop after two moves, and announce this as foundational to the commercially available OctTet truss.

What isn't acknowledged are auxiliary products of operation. Granted two moves produce three polyhedra with rational tetra volumes. QED. I needed to dig deeper.

Whereas their demonstrations begin with the tetrahedron I begin with axioms: one, negation, and eight three tuples: (±1,±1,±1).

I want to explore this synergetic algorithm further, before I critique tetrahedral volumetric accounting specifically (*next year*). Instead I take this opportunity to report the results of a tetrahedral analysis using vectors and sphere points.

2. Divide & Concur

Rule Zero, a tetrahedron is defined by four points with coordinate values equal to: that which is not nothing, one, and its negative, minus one, also not nothing.

Rule Minus Zero, an anti-tetrahedron is defined by Negation. With all eight tuple combinations of ±1, we may concur that a Tetrahedron + Anti—Tetrahedron equals a Cube, but that's a more complicated story.

I am in search of principles so I invoke Occam and choose the simpler, the tetrahedral space to operate on with the commands: *Interconnect-Points* generating lines, and *Curve-Middle* generating points.

IP(n) produces $(n^2-n)/2$ lines. With 4 points IP(4) produces 6 lines of length, two root two, a tetrahedron. Do you concur? Please nod in agreement.

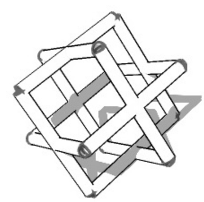
Objectively, what is in concurrence, are the lengths of the lines produced. Treating lines of equal length as a set for input to, *Curve-Middle* produces points with which to iterate IP. Edge sectioning options other than midpoint are conceivable. I suspect further prime number divisions of tetrahedron edges may prove interesting.

Four points, produce 1 set of 6 lines. From six lines come 6 mid-points which produce 15 lines in 2 sets. Synergetics proponents select the twelve shortest lines and declare it an emergent regular octahedron, of edge length, root 2.

The three other emergent lines that are unacknowledged by synergetic enthusiasts deserve mention. I introduce them to you as: Abscissa, Ordinate, and Applicate, each of length two. When these lines are C*urve-Middled*, three ironic points are enumerated. Three Origins with their threefold nothingnesses (0, 0, 0), singularities that arise out of the vacuum vastly increase the number of interconnections generated by further iterations of IP.

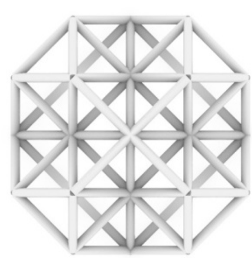
Take notice, Zero is an emergent property! I remove duplicate points from consideration and treat a singular origin with respect.

To complete the synergetics demonstration, the twelve mid points from the octahedron's edges seed, IP(12) producing 66 lines. Sorting by length produces four sets. The 24 shortest lines

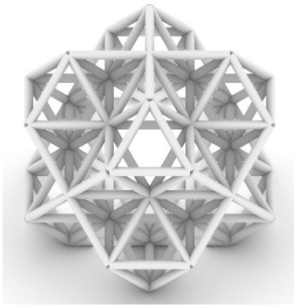


of length, one half root two, combine to form a cuboctahedron. Do you concur?

Perhaps Douglas Adam's has an opinion on what the the remaining 42 lines might represent. What do these three sets of vectors reveal? One contains: 6 axes of length, root 2, while the second contains a set of 12 lines of length one that limns three squares oriented in cartesian



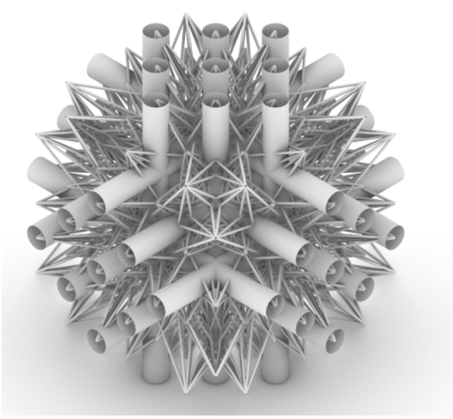
planes.



2x2

Curve-Middle operating on the third set of 24 lines, the diagonals, generates 24 points that with operation Remove Duplicate Points becomes 12. These points seed IP(12) to again produce 66 lines that include a cuboctahedron with edge lengths one half of the original. Further iterations reveal cuboctahedra 'all the way down'.

Octet Truss view 1



Octet Truss view 2

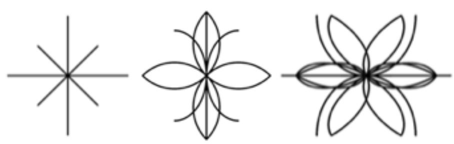
As the number of lines input to IP increases, the number of edge lengths expand as well. Using Curve-Middle with all generated lines and then removing duplicate points before iterating IP has been most effective.

66 lines when Curve-Middled produce 66 points, 43 are unique. One is the origin. Forty two points generate 861 lines of 12

different lengths. The line sets produced can be characterized as forming: radii, lattices and stellations. The shortest produce the octet truss.

2 sets Stellation and Radial

4. Tangent Cees



In Synergetics, close pack spheres are used to represent these geometries. To that end I have begun a program to parametrically represent the points used in Divide & Concur as spheres, not necessarily closest packed, merely located in space and oriented to the origin. It's an aesthetic quest.

Tetrahedron Spheres default alignment

This is because the software I use represents spheres as semi circles when in wire frame display mode. Depending on viewing axis these *sphere points* can be misinterpreted if no effort is made to orient them to the origin. I have manually arranged sphere points for the polyhedra discussed herein. Automated methods are under construction.

Tangent Cees of the Cuboctahedron axial views aligned

5. Conclusion Two sets: Radial and Stellation

I am no closer to analyzing volumes in terms of tetrahedra. Instead I have encountered a new infinity of generated forms that grab my attention. Reason remains unrealized. It is replaced by a search for beauty in the weeds. How these forms will assist in making the world work for 100% of humanity is anybody's guess.

So little time.

Lattice

6. References

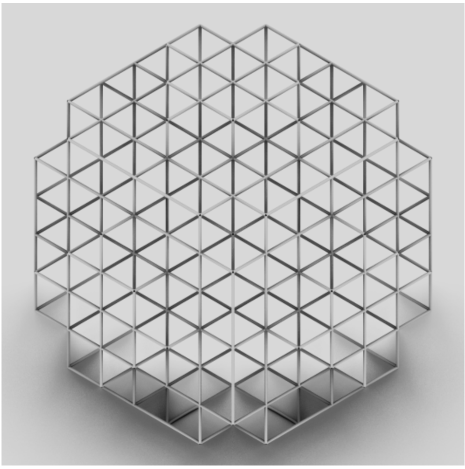
1. Synergetics, R. B. Fuller, 1975 Macmillan, New York.

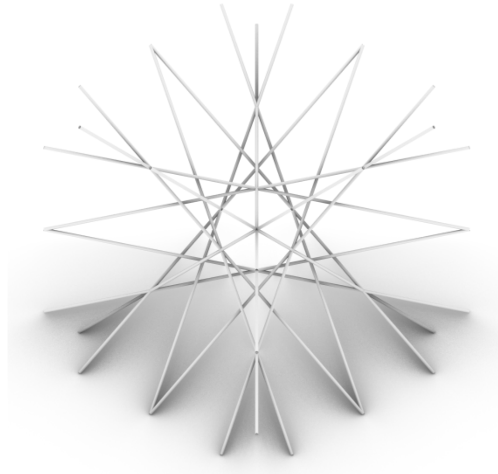
2. Synergetics and Quantum Physics, RWGray Project 2022

3. Realizing Reason, D. Macbeth, 2014, Oxford Press, Oxford

4. In Out Around, Readings from Synergetics, parts 1, 2, 3, C. Palmer, 2016, http://www.youtube.com/@Criterion5

The texts excerpted from Synergetics used in the In Out Around trilogy are: 400.25, 517.05, 538.01, 540.08, 720.10, 720.11, 814.01, 905.73, 962.42, 981.24,

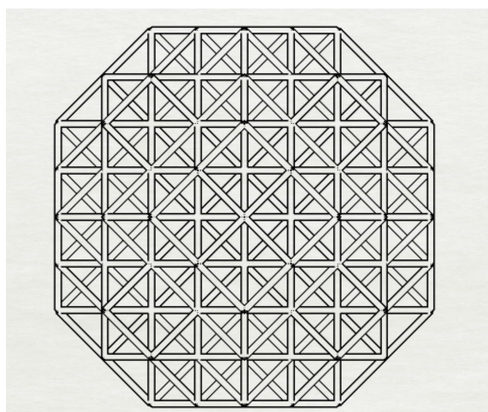




1051.10, 1051.40.

Stellation

Lattice arise and go now to the isle of man is free.



Learning Through Playing with AI: Collaboration of Intelligences in Creating a Transformable 3D Art Installation

Dr. Marija Obradović, Full Professor
Faculty of Civil Engineering, University of Belgrade, Belgrade, Serbia
e-mail: marijao@grf.bg.ac.rs

Dr. Slobodan Mišić, Associate Professor
Faculty of Applied Arts, University of Arts in Belgrade, Belgrade, Serbia
e-mail: slobodan.misic@fpu.bg.ac.rs

Abstract

This study explores how integrating AI image generators into art installations fosters innovative, educational outcomes, demonstrated in the RELATE project. Aimed at creating transformable 3D structures that invigorate urban spaces for learning, the project engaged students in a workshop to design a physically adjustable structure shaped by human and AI collaboration.

AI, here not just as a tool but also as a creative partner and competitor, plays a vital role in pushing students' creativity, challenging them to balance imaginative designs with practical constraints. The process begins with verbal descriptions fed into the AI, generating initial images that are iteratively refined. Students experiment with AI's limitless creativity, addressing technical limitations to make ideas physically feasible. By constructing the final structure in teams, students gain hands-on experience in design, feasibility, and teamwork. This project underscores AI's potential as an inspirational force in creative processes, enhancing human ingenuity.

1. Introduction

As part of the RELATE project https://www.relate.rs) work package V [1], we have the task of building a transformable 3D structures that invigorate urban spaces for learning. One of the structures is conceived as an installation which is a work of participatory art [2, 3], with the involvement of a group of students in a workshop. The structure must follow certain predefined rules:

It should be transformable [4, 5], meaning that its shape should not be static and fixed, but able to change and be modified according to the aspirations of the observers who interact with it. It consists of regular triangles interconnected in a triangular grid that, by its spatial transformations, creates a 3D free form, open or closed. The materials from which it is constructed are also assigned, since we are limited both by resources and time to find more diverse and adapted materials and elements.

The main challenge in this "game" is the involvement of an AI image generator [6], which plays triple role: participant, guide/playmaker, and learner. These are the same roles that the students — the participants — take on in this game, making it both a type of collaboration and a duel.

Through collaboration, students learn to leverage the conceptual solutions offered by AI, which, despite its sophistication, often presents surreal elements or so-called "hallucinations" [7, 8] due to its imperfections. Students need to recognize the advantages of such a vast creative potential, while also understanding its limitations when applied to concrete tasks and the implementation of ideas in reality.

The "duel," so to speak, adds a competitive edge to the game, enhancing students' motivation to give their best. The goal is not to show the superiority of human or artificial intelligence but to demonstrate that human intelligence can engage with advanced techniques and unpredictable results, which refine the initial raw design.

Such a workshop should include several educational levels, teaching participants — as well as observers or users — principles of statics, geometry, and aesthetics. It should foster creativity and imagination, encourage collective effort and collaboration, and through play, educate on the possibilities of using and interacting with AI image generators.

2. Goal of the study

With this workshop, part of the activities of the RELATE project, held on October 25, 2024 at the Faculty of Applied Arts, University of Arts in Belgrade, led by the authors of this study (M. Obradović and S. Mišić), we engage in exploring ways in which young people (students, pupils, and others) can learn from non-formal education methods outside traditional school curricula. Students of the first year of the Design Study Program, basic academic studies, Interior Architecture and Furniture Design module and Industrial Design module, Faculty of Applied Arts, University of Arts in Belgrade participated. Our objective is to demonstrate, through the creation of a transformable structure in a triangular system, that:

- The orthogonal system in which we live, a logical consequence of human physiology (as the human posture is orthogonal to the ground), is not an inevitable standard for spatial structures, as evidenced in recent decades by examples of free-form architecture.
- A triangle, as a geometrically stable figure, when multiplied into a network, creates a flexible structure whose shape can change in various ways — a concept proven by the triangulation of 3D models in computer graphics.
- An equilateral triangle provides less flexibility in shape formation than simple triangulation.
- The static stability of such a structure must be additionally defined by supports.
- The aesthetics of this structure lie in its inherent order, despite its irregular shape.

By introducing AI as a participant in the structure's creation, we aim to address several key issues:

a) Students' attitudes toward AI (whether they have previous experience or preconceptions about AI);
b) Familiarization with the capabilities of creating images via AI image generators;
c) Exploration of different AI generators with varying operating principles;
d) Critical evaluation of the results provided by AI;
e) Assessment of the feasibility of the proposed solutions;
f) Evaluation of the aesthetics of the proposed solutions;
g) Engaging with the challenge of translating ideas from the virtual, artificial world to the real, physical world.

Finally, through this practical experiment, we aim to demonstrate how well AI has learned to interpret our requirements as beings constrained by the laws of physics, while also teaching us, humans, how to give more precise inputs (prompts) so that the AI algorithm can provide the most satisfactory solutions.

3. Methodology

The workshop is designed so that, due to time and resource constraints, materials for constructing the transformable structure were procured in advance. Having faced a similar challenge in prior research [9], we prepared materials that enable us to create a structure based on a (planar) regular triangular grid [10]. The materials include:

- Cardboard tubes, similar to frappe straws, measuring 22 x 0.6 cm,
- Metal rings (metal purse buckles) with a radius of 2.5 cm;
- Rubber bands with a diameter of 12 cm;
- Long metal hooks for threading rubber bands through the tubes;
- Thick thread.

The tubes are black, and the thread is white, which we also considered an important factor when providing inputs to the AI.

The steps are as follows:

3.1 Verbal Description

The student enters a textual description of the 3D spatial structure into an AI image generator (AIIG). For instance, the prompt we give to AI reads:

*"I want an image of a **3D structure** made of **frappe straws**, interconnected in a **triangular grid**, joined by small **ring-like joints**. The straws are **black** and joined with small **silver rings** through which the **white rope** is put. The net of the structure is folded in a **free form** way, so that it produces a dynamic, **transformable structure**. The structure is set as an installation **sculpture**, in a room with old brick walls and arches".*

Note: To avoid overly directed outcomes, we provided students with only key terms (in bold) to help them start. The scene or setting for the structure was left for participants to choose by their own.

We selected four AI image generators (based on criteria such as being free to use, so students would not need to pay for registration, and also to be simple and intuitive enough):

I. Imagine Art
II. ChatGPT

III. Stable Diffusion
IV. Canva – Magic Media

We then divided the students into four groups, each assigned one of the AI image generators. Each group had seven students, with each student tasked to submit two results generated by AI.

3.2 Image Generation

Students input prompts into the AI generator and receive generated images. AI image generators that use textual descriptions work through deep learning models. When words are entered, the model uses pre-trained language algorithms to convert those words into vector representations. These representations then drive the generative model to create images that correspond to the given description. Instead of a direct numerical code, words serve as high-level abstractions, allowing users more intuitive control over visual outputs. The model learns from vast amounts of text and images to link meanings with visual representations.

Since the AI generator creates an image based on description and uses creative algorithms, it often produces surreal or abstract results.

3.3 Iterative Process

With guidance from the workshop moderator, students then refine their descriptions, adding more specific requirements, removing surreal elements, or adding new words to steer the output toward a feasible solution, continuing until they achieve an image that closely resembles their envisioned structure. AI regenerates the image, moving toward a more realistic, feasible solution. Some of the AIIG solutions are shown in Fig. 1.

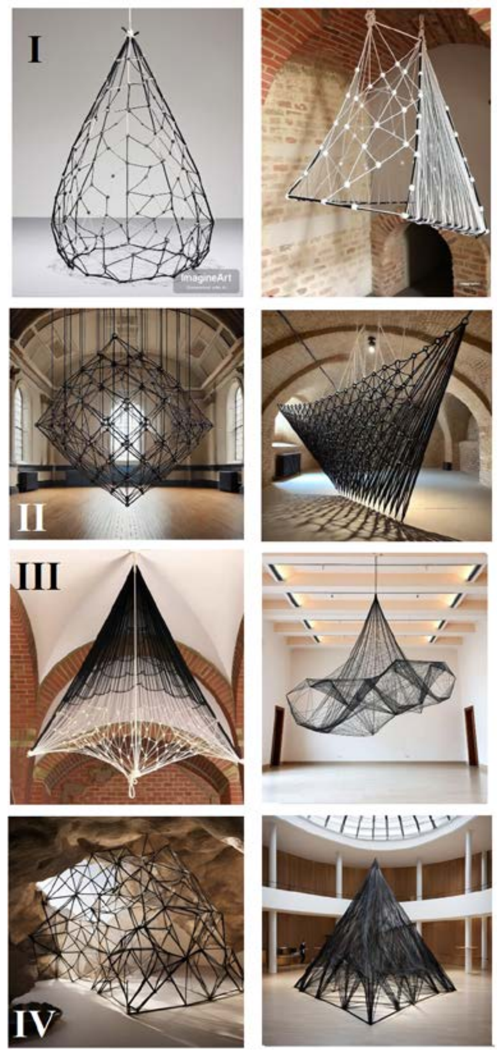


Fig. 1: Chosen examples of generated images created by four selected AIIG: I - Imagine Art, II - Chat GPT. III - Stable Diffusion, IV - Canva – Magic Media

3.4 Evaluation of the solutions

We compiled a collection of 56 AI-generated images. From this set, we narrowed it down to eight, using the following criteria:

1) **Match to the prompt:** The AI-generated image (AIGI) should correspond closely with the prompt description.
2) **Feasibility**: The structure should be achievable with the given materials and time, including an analysis of required elements, arrangement, time, and installation space.
3) **Transformability:** The structure should allow shape changes by moving a few key nodes.
4) **Static stability**: It should hold its shape in space after assembly, without collapsing or disintegrating.
5) **Aesthetic appeal**: This includes evaluating what makes the sculpture visually attractive, considering balance, proportion, composition, order, harmony, elegance, unity, and originality.

3.5 Practical Application

After analysing and selecting a solution to implement, students attempt to simulate the chosen structure in the physical world. The structure is divided into modules, which are then assembled into a whole. Each of the four teams is tasked with constructing one part — a module of the future structure.

Finally, the modules are combined into a complete structure, which should physically represent the idea processed through AI-generated algorithms.

Certainly, constructing a structure based solely on a 2D image is not straightforward and requires guesswork, intuition, and excellent spatial awareness, as construction plans typically include several projections (top view, section, elevation), a 3D model, or a physical model. For reverse engineering [11], multiple photos, footages or recordings on site would be needed for a reconstruction of its shape [12, 13]. Here, however, we only have a single 2D image, so we must rely on intuitive reconstruction and deductive reasoning. For this reason, we chose a structure for the workshop that is more of an artistic installation, allowing for some margin of error and artistic freedom in replicating the AI concept.

After the first iteration in making the structure, we concluded that the white rope we initially specified as a key term in the prompt, and its random arrangement within triangles, produced aesthetically unsatisfactory results, creating a chaotic network that confused viewers and suggested a broken structure. In the second iteration, we replaced the white rope with black and specified a more orderly arrangement.

We installed the resulting structure in the hall of the Civil Engineering and Architecture faculties building (Belgrade).

The results of our educational and creative experiment are presented in the following section.

4. Results

This research led to diverse noteworthy findings:

4.1 Insights:

- The 28 students had no prior experience with AI image generators, so working with them in this workshop was a novel experience (somewhat surprising, given that they are students from the Faculty of Applied Arts, who might be expected to show interest in visual experimentation).

- Students approached the workshop activities with great enthusiasm throughout all stages, suggesting that AI's creative "prompting" was highly inspiring. Furthermore, hands-on, group work was a fresh and enjoyable experience for them, fostering a sense of community, collaboration, play, and ingenuity.

4.2 Comparative Observations

By comparing the results from the four selected AI image generators (AIIG), we concluded:

I. **Imagine Art**, an AI image generator that uses deep learning algorithms to transform textual prompts into images by interpreting and rendering visual elements based on contextual understanding, uses the model that has been trained on vast datasets of images and descriptions to create a wide range of artistic styles and compositions.

It struggled to "understand" the concept of a triangular grid or a transformable structure. The solutions it offered were charming, somewhat minimalist, and therefore closest to being feasible within the available materials and time. However, the offered designs were problematic in terms of static stability.

II. **ChatGPT,** although mostly focused on conversational AI, in certain implementations can also interface with image-generation models like DALL-E [14]. When combined, ChatGPT interprets a user's prompt conversationally and passes it to an image generator, which produces images that align with the detailed, text-based descriptions.

It offered perhaps the most attractive, grandiose, and exciting solutions, but despite their appeal, we found them mostly too ambitious to be realized within the initial project constraints.

III. **Stable Diffusion** is a latent diffusion model [15] that creates images from text by gradually refining random noise until a coherent image forms. It operates by mapping textual input into a latent space, where it iteratively improves the image output based on learned visual patterns and structure.

It produced interesting but mostly impractical solutions. It did not "understand" the difference between thread and frappe straws, and it interpreted straw dimensions arbitrarily, resulting in designs that, though simple, deviated from the initial concept.

IV. **Canva – Magic Media** is an AI-powered feature in Canva platform, which also uses diffusion or transformer-based models [16] (similar to Imagine Art and DALL-E), but it's built specifically for easy integration with Canva's design tools, generating images tailored for quick customizations.

It provided some potentially satisfactory solutions, though it frequently made significant errors regarding the dimensions of the straws or the method of connecting them, diverging from the triangular mesh, particularly in terms of transformable structures.

A comparative review of the success of these four AIIGs in meeting the five criteria outlined in Section 3.4, based on participants' votes, is shown in Table 1.

We evaluate not the quality of the AIIGs themselves, but their suitability for this specific task.

Table 1: Comparative overview of the success in meeting the 5 criteria for the 4 observed AIIG, rated from 1 to 5 (1=failed; 2=poorly successful; 3=somewhat successful; 4=fairly successful; 5=fully successful)

AI image generator	1) prompt faithful	2) performability	3) transformability	4) viability	5) aesthetics	Overall score
Imagine Art	3	4	2	4	4	17
Chat GPT	4	2	2	4	5	17
Stable Diffusion	2	4	3	4	3	16
Canva Magic Media	3	2	2	2	3	12

4.3 Challenges

Executing the AI-generated idea, which leaned towards surrealism, presented inevitable difficulties, requiring us to adjust our expectations accordingly. During the project, we observed that AI generators (still):

- Cannot count accurately: when given a specific number of elements (e.g., frappe straws), they generate a random quantity, often significantly larger, as they interpret anything above a certain threshold (?) as "many."
- Cannot calculate element lengths accurately, often making drastic errors. They either produce significantly larger dimensions or, if asked to make all sticks the same length, they take the "path of least resistance," offering unusable results or irregular triangular meshes with sticks of varying lengths, which affects the final result.
- Do not recognize node points, treating metal rings as vertices of a triangular mesh. They either place them randomly or omit them entirely, posing a real problem for physical execution.
- Simply ignore complex ords if they cannot easily integrate them, such as "transformability," which was rarely respected as a criterion.
- Struggle to deviate from symmetrical solutions, particularly some AI generators (e.g., ChatGPT), making it difficult to generate images of free-form structures.

4.4 Achievements

On the other hand, users with no prior experience with AI image generators quickly adapted, and within a short time (1 hour), we successfully generated enough (high-quality) images for the entire exhibition with a group of 28 students. Guidance from more experienced moderators (the authors) was, of course, essential. We were able to construct the intended structure with the available materials and within the given time, producing a basic framework that, after another iteration, was refined into the final design. The creation phases and the final structure are shown in Fig. 2 and Fig 3.

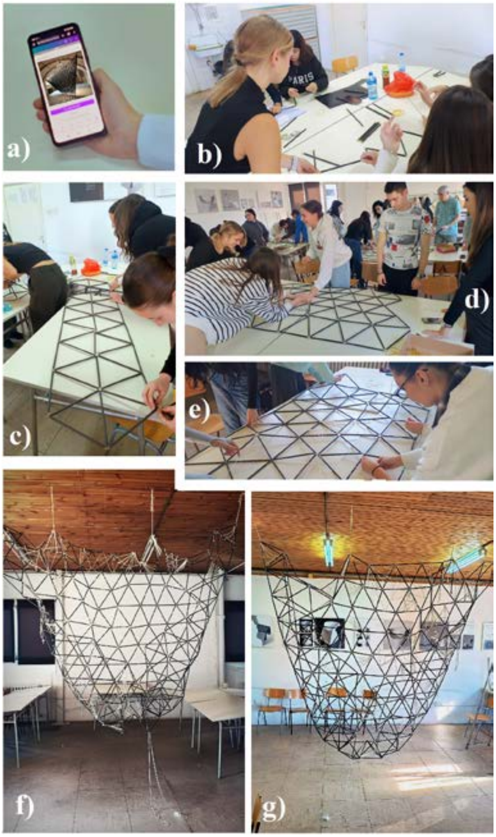


Fig. 2: Phases in creating 3D transformable structure according to the AIGI

As expected, the biggest challenge was reconstructing a 3D structure from a 2D image: understanding its assembly technique, integrity, static stability, and especially segmenting it so that each team could create a part that would later join with the other three into a cohesive whole. This required substantial planning, reliance on prior experience, and a lot of improvisation on-site, including trial and error, adjustments, and iterations.

4.5 Analysis of the result

We selected one solution, but our approach, concerning a transformable structure, drew inspiration from several others we shortlisted.

- We chose a simple solution based on a planar triangular grid, as seen in Fig. 1-II on the right. However, it wasn't possible to fold this grid in the exact same way as in the image, because all the straws, as tasked, had to be of equal length, which the AIIG did not account for in its result. Therefore, during the folding process, we also included inspirational concepts from some other solutions, such as those shown in Fig. 1-I, 1-III, 1-IV, etc.
- The variety of shapes we achieved by transforming the planar grid surpassed the range of shapes that AIIG offered us.
- The remaining challenge, especially within the given timeframe and with available materials, is the technical complexity and robustness of the structure. This would require additional resources, so we can consider the results of this workshop as a prototype or pilot project for a larger structure that we will create for installation in an outdoor space.

The photo gallery of the transformable structure installed in the building hall of the of the Faculties of Civil Engineering and Architecture in Belgrade, as well as from the workshop, can be seen at: https://www.relate.rs/en/transformative-3d-structures-workshop/

Finally, the gap between the AI-generated image and the actual structure created in physical space is an inevitable result of all the factors mentioned above. It illustrates how much both AI and human intelligence still need to learn and work on understanding each other.

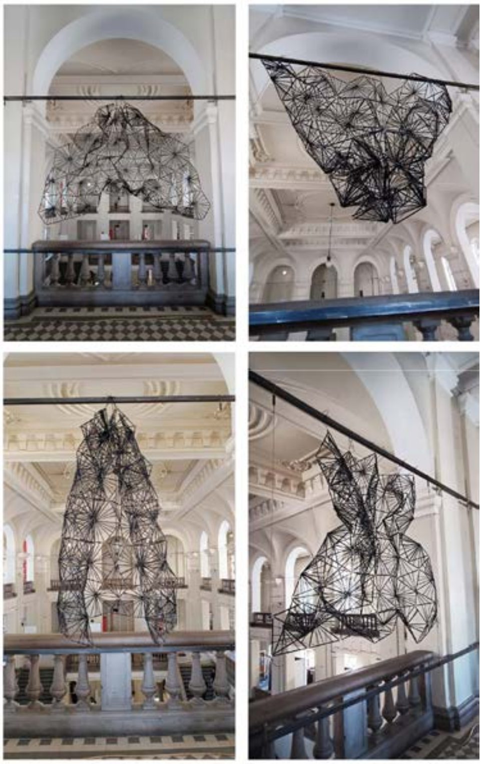


Fig. 3: The final structure in the chosen scene after iterations and adjustments

5. Conclusions

What we can conclude is that AI is not yet able to fully understand out from the context (lacking cognitive abilities) and illustrate exactly what it is assigned, but it's quickly advancing in that direction. What AI can do is: to take given words and create an image that is a step toward visualizing the imagined, alike human contrives from the description they read.

In some circles (the authors can only speak from their own ambience, without broader research), there are still preconceptions that AI functions as a rigid, input-driven, algorithmically limited entity which, unlike humans, lacks creativity. Today's AI platforms—like chatbots, generative AI, or natural language processing—actually prove the opposite: the results of their creativity, or "creativity" (a topic for deeper psychological, neurophysiological, or even philosophical or theological study), do not fall behind human creativity and may even surpass it in certain aspects.

Without a physical body, AI is free from many of the limitations humans face, whose existence from birth to death is shaped by the constraints of the physical body. Thus, AI can "imagine" and "hallucinate" without tangible limits, combining the seemingly incompatible into unexpected compositions and forming surprising correlations, much like a child's imagination unrestrained by experience—only with a far larger base of information to draw from.

However, since AI (like humans themselves) is a product of humans, it is "educated," taught, and corrected by them. We can say with confidence that, over time, AIIG will be perfected to the point that it will lose some of its so-called negative traits, such as its lack of understanding of measures, numbers, or sequence. Yet, in doing so, it may also gradually lose its fascinating and exciting unpredictability—which is one of the essential characteristics of creativity.

Fundings:

RELATE - An outside curriculum, within the IDEJE project cycle, by the Science Development Fund of the Republic of Serbia, Project No 7726555.

Partly by: MNTRS project No 200092.

References:

[1] RELATE (leaRning Elswhere, LeArning oTherwisE) – An outside curriculum, within the IDEJE project cycle, by the Science Development Fund of the Republic of Serbia (cycle period: 2021-2025), Coordinator: Marta Brkovic– Dodik. https://www.relate.rs/en/mission/

[2] Milevska, Suzana. "Participatory art: a paradigm shift from objects to subject." EUROZINE, 1-7 (2006).

[3] Holt, Matthew. "Transformation of the Aesthetic: Art as Participatory Design", Design and Culture, 7:2, 143-165. (2015)

[4] Bouten, Sam. Transformable Structures and their Architectural Application, Master's dissertation, Faculty of Engineering and Architecture, Uversity of Gent. (2015).

[5] De Temmerman, N., Alegria Mira, L., Vergauwen, A., Hendrickx, H., & De Wilde, W. P. Transformable structures in architectural engineering. *High Performance Structures and Materials*, *6*(124), 457-468. (2012).

[6] Enjellina, B. E. V. P., & Rossy, A. G. C. "A review of AI image generator: Influences, challenges, and future prospects for architectural field". *Journal of Artificial Intelligence in Architecture*, 2(1), (2023). 53-65.

[7] Salvagno, Michele, Fabio Silvio Taccone, and Alberto Giovanni Gerli. "Artificial intelligence hallucinations." *Critical Care* 27.1 (2023): 180.

[8] Kumar, Mukesh, et al. "Artificial hallucinations by Google bard: think before you leap." *Cureus* 15.8 (2023).

[9] Obradović, Marija, and Slobodan Mišić. "Learning while playing - Througie platform for creating models of spatial structures." *The 9th International Scientific Conference: Mongeometrija 2023"*. Faculty of Technical Sciences, University of Novi Sad, 2023.

[10] Weisstein, Eric W. "Triangular Grid." From *MathWorld*--A Wolfram Web Resource. https://mathworld.wolfram.com/TriangularGrid.html

[11] Sans, Fernando and Llopis Soler, Antonio (Eds.), *Reverse Engineering: Mechanisms, Structures, Systems, and Materials*. Springer (2012).

[12] Luhmann, Thomas, et al. *Close-range photogrammetry and 3D imaging*. Walter de Gruyter GmbH & Co KG, 2023.

[13] Remondino, F., & El-Hakim, S. (2006). *Image-based 3D modelling: A review. The Photogrammetric Record, 21*(115), 269-291.

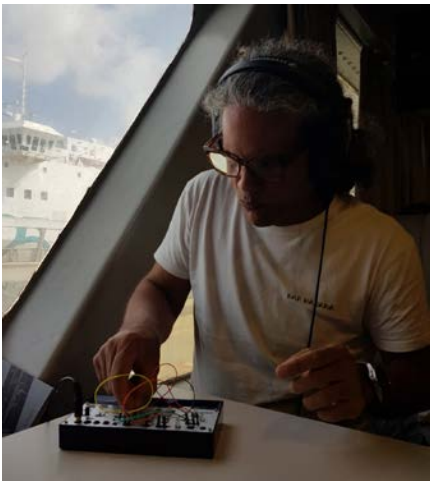
[14] DALL-E: Creating Images from Text, available at: https://openai.com/index/dall-e-3/

[15] Rombach, Robin, et al. "High-resolution image synthesis with latent diffusion models." *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. 2022.

[16] Dosovitskiy, Alexey. "An image is worth 16x16 words: Transformers for image recognition at scale." *arXiv preprint arXiv:2010.11929* (2020).

Sonification.AI: Generative Sound Experiments

Yan Bello Méndez
Creative Founder & AI Expert. SpaceMinds SL, Madrid, Spain
www.artivist.ai – www.spaceminds.com
e-mail: yan@spaceminds.com



Abstract

This paper provides an overview of an experimental, iterative, and incremental approach (created by the author) to generative data sonification. This creative method enriches a collection of generative techniques to foster inter-multidisciplinary collaborative, systematic exploration of opportunities to use Artificial Intelligence (AI), storytelling, and data "(pre-)sensing" to explore experiments, human-computer interaction concepts, and sound artworks. Several concept-projects are discussed to illustrate the approach in action and its results are discussed.

1. A new way of "viewing" data

Nowadays, nobody is unaware of the importance of data. I am not just referring to personal data, cookies, etc. But to data in the broadest sense, such as the values of attributes that allow us to create representations of phenomena, activities, the sensors of our devices, indicators of a person's health, or overall the evolution of the health of all the people in a region, country, etc. Obviously, data also plays an essential role at a business level, where it manifests itself as a key "raw material" for decision-making, which in turn is the driving force behind the management of any organization and its teams. Therefore, it should not surprise us that we talk "all the time" about visualization and the ("visual" - excuse the redundancy) presentation of data. However, there is another way of appreciating data, which is generally - based on my experience - quite unknown to many people. It is the "sonification" of data.

2. What is (data) sonification?

In many sessions, when we are about to address this topic, I ask the previous question, which I quickly follow with "…does that sounding sound familiar to you…?" – and practically all the participants look surprised. A little later, as soon as I mention that it is like "visualizing data" but "using sound", then they recognize that they are facing something new, but that it "seems logical" or "familiar" or "makes sense" to use sound to communicate data, when we have the sense of hearing.

However, despite its apparent novelty, sonification is NOT new. In 2010, it was defined as follows in a cross-sectional report on the state of the art of this discipline: "Sonification is the use of non-speech audio to convey information. … it is the transformation of relationships of/between data into relationships perceived in an acoustic signal in order to facilitate communication or interpretation." [2]. Another definition proposed by Hermann [3] emphasizes the systematicity and repeatability and therefore the value of this discipline for science: "Sonification is the generation of sound dependent on data, if the transformation is systematic, objective and reproducible, so that it can be used as a scientific method." [3].

2.1 The Sound of COVID!

In April 2020, when a vaccine for COVID-19 was still not even on the distant horizon, one of the news items that caught my attention and that I remember discussing with friends and colleagues was the announcement that they had created music from the virus' DNA. This is the work of Markus J. Buehler [4] – a professor at MIT and composer of experimental, classical and electronic music, with a strong interest in sonification. In fact, he is the co-author of another work published months earlier on the use of sonification and Deep Learning – Artificial Intelligence for the design of new proteins [5] – See the representation in Figure 2.

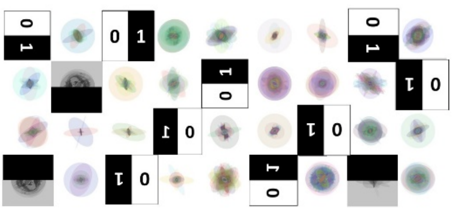
Figure 2. Score generated for the 3tnu protein, which is part of the training set. This score is shown only to visualize the concept and to illustrate the timing, rhythm, and progression of notes learned from the amino acid sequence. Source: Yu. & Buehler. APL Bioengineering [5].

Worth noting that training this type of deep neural networks requires a large amount of data [6] (in this case, sheet music). Using these techniques, the SARS-CoV-2 coronavirus protein has been "converted" into music [4]. Or more specifically, the researchers had translated the DNA data of said protein into a novel musical scale (suggestively called the "amino acid scale") since the notes it uses do not reflect a conventional musical scale, but rather each note in the space of 20 admissible tones in the native amino acid scale is assigned to one of the 20 amino acids [5].

The method described provides, according to the authors, a tool for designing new protein materials that could also find useful applications as materials in biology, medicine and engineering.

2.2 FALLsonancIA (NO*fall*.art)

This case-study provides an overview and a status update of a generative art-based project, namely NO*fall*.art (created by the author) and presented in GA2021. In this new iteration, we explore the use of sonification techniques to continue our journey into raising awareness of fall-prevention in elder population.



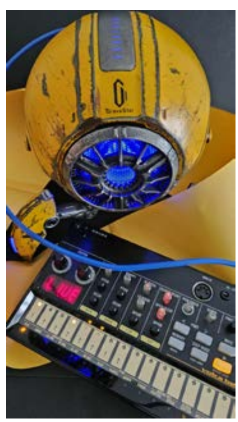
"*Falling Soundscapes*" by Artivist.AI

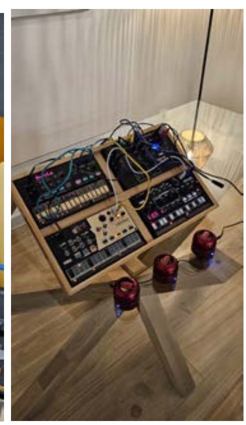
Several generative AI and data science techniques are used to create datasets, mapping and related sound experiments: FALLsonancIA, FALLinOscope, DO-RE-MI-FA-LLM, and more.

3. VolcaNO*lava*.art

With this concept-project we pay tribute to the numerous people affected by the La Palma volcano eruption in 2021/22, as well as to all the people engaged in managing the emergencies caused by this destructive event. The project includes some hybrid artwork collage with analogue and digital elements such as visuals and sonification experiments.

Also, inspired from volcano-eruption data, sonification techniques power the performance of data-driven compositions using KORG Volca modular synthesizers.





4. From space data to sound: an easy correspondence?

Representing data through sound involves making a correspondence between the values of said data and attributes/properties of the sound. At first glance this may seem complicated and in fact depending on the phenomenon being studied the process can be laborious. However, the same technique – sonification – is applicable to various fields, phenomena and areas of research.

One area where sonification has provided great results is the study of space. If we leave our "house" (planet Earth) and enter our "yard" (the Milky Way), we can explore and appreciate in a different way what is "out there" using sonification with the help of NASA.
A new project that uses sonification converts astronomical images from the Chandra telescope and other telescopes into sound [7]. This allows us to "listen" to the center of the Milky Way using the data resulting from its observation in X-ray, optical and infrared light. I encourage you to listen - and see! - the results of these sonifications are available on the NASA Chandra X-ray Observatory website.

There you can see the correspondence between the movement of the cursor across the image and how the sounds represent the position and brightness of the sources. As part of this project, astronomical images of the supernova Cassiopeia A and the "Pillars of Creation" (see image 3) were also converted into sound. Both the images and the resulting sounds are spectacular.

5. From bioengineering to the pillars of creation… to your data

There are numerous applications of (data) sonification. As you can see from the examples above – and many others you will find on the websites and links mentioned – I like to share the reflection that the process and context of sonification go beyond the "translation" of data into sounds. For this reason, I usually visualize (data) sonification as seen in Figure #1 to state this process as the ability to "(re)present data through sound" for the purposes of exploration, appreciation, research and various practical applications of the data.

This journey through the universe of the creative use of data and sound continuous delving into the origins/roots of sonification, exploring other areas of application, techniques, tools and considerations to apply sonification and discover how your own data sounds!

6. Exploring a floods-related live Data Sonification - based alert system with GenAI

With the sad and tragic events of the DANA (“Depresión Aislada en Niveles Altos”. In English: Isolated Depression at High Levels), which translated into terrible heavy raining storms that have destroyed many towns in Valencia (Spain) and generated over 200 human life’s lost in October 2024, I had to turn this generative experimental project into a space of exploration for preventing such disaster in the future, or at least minimizing its potential negative effect.

Perhaps, data sonification could serve as an additional mechanism which complementing other communication channels would improve the chances of managing such terrible events in a more coordinated and timely manner.

After various experiments using a custom GPT (I trained/finetuned myself), here are some conclusions based on the “conversation” with that personal assistant: “Sonification is a powerful tool for conveying complex, evolving data in an intuitive way, but it is not a complete replacement for traditional alert systems like SMS and radio. Instead, it works best as a complementary approach that enhances existing systems, improves accessibility, and adds real-time, continuous feedback. By combining multiple alert modalities, you can create a more robust and effective early warning system that meets the needs of diverse users and contexts”. This type of hybrid (human+AI) collaborations are a significant way to augment our (human) capabilities [8] and sonification can potentially serve as an additional “sensorial input” into such hybrid AI systems. (Gen)Art in this case will likely have an even more humanistic impact by probably saving lives.

7. References

1] SpaceMinds Workshops & Creative TEAM Challenges. SpaceMinds.com

[2] Kramer, G., Walker, B., Bonebright, T., Cook, P., Flowers, J. H., Miner, N., & Neuhoff, J. (2010). Sonification report: Status of the field and research agenda.

[3] Hermann, T. (2008). Taxonomy and definitions for sonification and auditory display. International Community for Auditory Display.

[4] Fuente: International Science Council (2020). https://council.science/current/news/the-sound-of-science-the-sars-cov-2-virus-as-a-piece-of-classical-music/

[5] Yu, C. H., & Buehler, M. J. (2020). Sonification based de novo protein design using artificial intelligence, structure prediction, and analysis using molecular modeling. APL Bioengineering, 4(1), 016108.

[6] Bello, Yan. (2019). NeuroTLON: Las redes neuronales artificiales “se profundizan”. Forum Calidad #306.

[7] Data Sonification: Sounds from Around the Milky Way. CHANDRA X-ray Observatory. NASA's flagship mission for X-ray astronomy”. Fuente: https://chandra.si.edu/photo/2020/sonify/

[8] Bello, Yan (2024): AImplify your Creative Collaborations with Generative AI and VAMOS TOD@S Multidisciplinary Innovation. SIGraDi 2024 Conference.

Utilizing the Generated Dance Choreographies in Theatre Performances

Bin Umino
Faculty of Sociology, Toyo University, Tokyo, Japan
e-mail: umino@toyo.jp

Asako Soga
Faculty of Advanced Science and Technology, Ryukoku University, Otsu, Japan

Motoko Hirayama
Graduate School of Comprehensive Human Sciences, University of Tsukuba, Tsukuba, Japan

Bin Umino

Asako Soga

Motoko Hirayama

Abstract

We developed the Body-part Motion Synthesis System (BMSS) software that generates and simulates an extensive range of novel dance movements. Over six years, three professional choreographers have employed BMSS-generated dance sequences to create their original full-scale dance works. As a result, 12 pieces were successfully staged. We further analyzed a piece staged in 2023 that incorporated nine BMSS-generated dance sequences. In creating this piece, the generated sequences were modified by the choreographer's expertise, experience, and artistic sensibility, culminating in a work refined to the artistic standards required for theatrical performance.

1. Introduction

Our research has focused on exploring how computers can support the creation of artistic human movement in contemporary dance. Digital technology is particularly valuable in contemporary dance because it fundamentally seeks to explore novel human movement without a fixed style, a pursuit that computer-generated movements can easily support.

We developed two distinct systems aimed at generating new movements for contemporary dance. The first system is the Body-part Motion Synthesis System (BMSS), a software that uses 3D motion data captured from professional dancers' performances to generate and simulate an endless variety of dance movements [1]. The second system, still in development, allows choreographers to experience unusual dance movements in real-time within the metaverse, using VR devices to create movements based on their own physical actions [2].

This study evaluates the effectiveness of the BMSS system in supporting full-scale creation. By "full-scale creation," we refer to the process in which a professional choreographer invests substantial time to produce an original work that meets artistic standards for a paid audience in a theatre setting, performed by professional dancers.

2. Overview of BMSS

BMSS is a system that synthesizes dance movements using motion data captured from performances by professional dancers. It allows users to create dance sequences and play them back in 3D CG.

Figure 1 shows the interface of BMSS. The system features two modes: Unit Mode (left) and Sequence Mode (right), which can be switched via tabs.

In Unit Mode, users can select one from dozens of full-body basic motions (Base) and set synthesis conditions to automatically generate an unlimited number of "units"—elemental movements lasting one to several seconds.

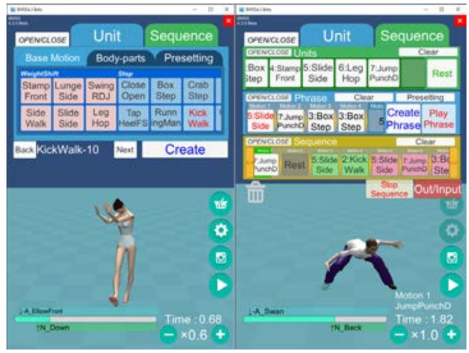


Figure 1. The user interface of BMSS

The Base motions range from simple steps to acrobatic actions, systematically categorized into different types of basic dance movements. The generated units can be replayed using an avatar, and they can be edited by changing the speed, reversing left and right, etc. Up to 30 units that may be useful for the creative process can be saved.

In Sequence Mode, users can select from saved units and arrange them in chronological order to create a sequence. Up to 30 units can be linked together, allowing for the creation of sequences up to about three minutes in length, which can then be played back in 3D CG. The camera view can be freely adjusted by touch-screen operations.

The software generates a wide variety of movements suited for contemporary dance, but it does not entirely eliminate unnatural or physically impossible movements. This inclusion is intentional, as they can aid in discovering new artistic dance expressions. The unnatural or physically impossible ones are useful using professional expertise, artistic intuition, and experience, elevating them to performance-ready standards.

3. Creation and Performance by Professional Choreographers

Three professional choreographers employed BMSS to create full-scale, original contemporary dance works, which were performed in Tokyo theatres in October 2017, November 2018, January 2021, and January 2023. All three are active choreographers based in Japan, with awards in contemporary dance and extensive experience performing abroad. One of these choreographers was Hirayama, the third author of this paper.

The choreographers were asked to create their original pieces based on BMSS-generated dance sequences without starting from music, narrative, or emotions. However, they were free to incorporate music, narrative, or emotions into the final work if deemed necessary for artistic reasons.

The performances took place in a theatre commonly used for contemporary dance performances. Professional stage management, lighting, and sound staff were hired, and detailed consultations were held with each choreographer. Promotional activities such as flyers and internet-based advertising had been conducted several months in advance.

Over a period of six years, three choreographers created 12 dance pieces and conducted six stages across four events, involving a total of 34 dancers. The total audience across all performances was 567 people.

4. Analysis of *Peaux* Staging in 2023

In this paper, we focus on the analysis of Hirayama's 2023 piece, *Peaux*, an 11-minute, 30-second piece performed by two male dancers using nine BMSS-generated sequences. This study examines how the generated sequences were modified and utilized in the creation of this full-scale dance piece.

Figure 2 is a visual representation based on the recorded performance of *Peaux*. This figure illustrates which dancer's movements at specific times correspond to particular dance sequences generated by the BMSS. The two male dancers are labeled as Dancer X and Dancer Y. The horizontal axis represents time, dividing the 11-minute, 30-second piece into four sections.

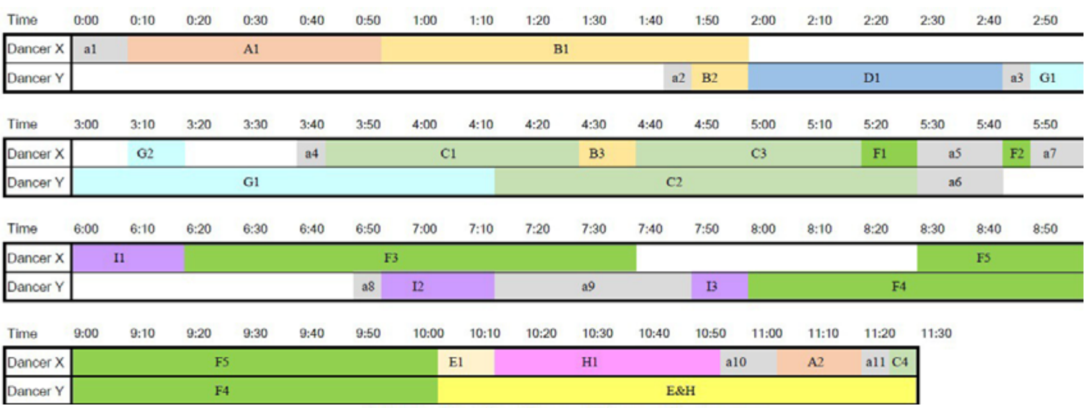


Figure 2. Timeline of the performance

The nine BMSS-generated sequences are labeled with identifiers from A to I, with each one displayed in a distinct color. The exception is Dancer Y's movement at 10:05 mixing E and H. White areas indicate periods when a dancer was offstage. Gray-shaded areas (a1–a11) represent movements added later that are not derived from BMSS-generated dance sequences.

Figure 3 shows scenes in chronological order from the performance, capturing specific moments at 1:07 (A1), 5:08 (C3, C2), and 10:49 (H1, E&H).

Figure 3. Performance scenes

5. Discussion

A detailed examination of how BMSS-generated dance sequences were adapted for performance on stage revealed nine types of modifications.

(1) Danceablization of Transitions

BMSS intentionally leaves the motion data between Units unrefined, allowing abrupt transitions to inspire the choreographer. The choreographer refines these transitional gaps, preserving the appeal of the original CG movements while making the transitions danceable.

(2) Danceablization of Non-danceable Movements

BMSS can generate physically impossible movements. Examples include joints bending in anatomically impossible directions, limbs passing through the torso, and body parts sinking below the floor. These movements are adapted to be achievable by dancers as needed.

(3) Direction Adjustment

BMSS adjusts the direction a dancer faces during transitions between Units. However, for artistic purposes, the choreographer may further modify the dancer’s orientation. The choreographer also refines the extent of movement, such as the angles of arms and legs.

(4) Rhythm and Accent Addition

The choreographer adjusts the timing of movements, creating a rhythm. Specific durations of movements on the timeline are adjusted, and accents are added by instructing dancers on when to emphasize or soften their movements.

(5) Partial Use of Sequences

Parts or fragments of BMSS-generated sequences were used. Specifically, segments under 30 seconds (B2, G2, B3, F1, F2, E1, A2, C4) represent portions of sequences.

(6) Arrangement of Sequences

After applying the modifications described above, the arrangement of sequences is left to the artistic discretion of the choreographer, including the placement of solos and duets.

(7) Addition of Movements Not Derived from Generated Sequences

Gray-shaded areas (a1-11) represent movements added later that are not derived from BMSS-generated dance sequences, specifically: regular walking (a1, 3, 4, 8, 10, 11), light jogging (a2), backward walking (a5, 6), and lying supine on the floor (a7, 9).

(8) Coordination Between Two Dancers

The piece was choreographed to combine the movements of two dancers with artistic flexibility. In *Peaux*, there are five instances where the dancers synchronize identical movements and one instance where the two dancers face each other and perform symmetrical movements.

(9) Contact-Based Partnering

Sections of the dance involving contact between two dancers were also based on BMSS-generated sequences. These sections were choreographed by having each dancer perform distinct sequences simultaneously, identifying points of incidental contact, and then developing contact-based choreography from these moments.

6. Conclusion

We developed a software named BMSS, and professional choreographers have employed BMSS-generated dance sequences for full-scale creation. As a result, 12 dance pieces were successfully staged, demonstrating the practical utility of the generated dance movements for professional choreographic creation.

Research on generating dance using motion data is currently part of machine learning-based generative AI studies [3]. However, currently, computer systems cannot fully generate authentic contemporary dance movements. In the field of dance, the choreographer's embodied knowledge and experience are essential to interpreting and refining computer-generated movement.

In our project, we are also developing complementary methods, including VR-based assistance, to support dance creation alongside BMSS.

Acknowledgment

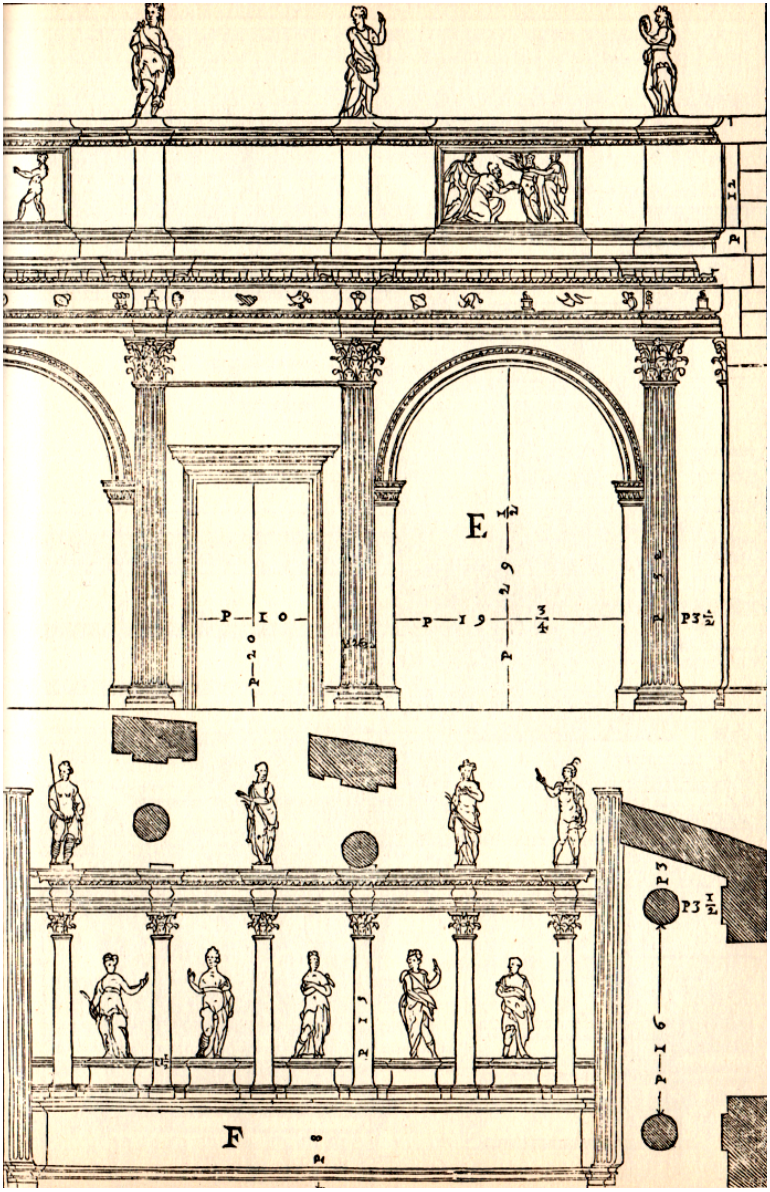
This work was partly supported by the Joint Research Center for Science and Technology of Ryukoku University.

References

[1] Soga, A. Umino, B., and Hirayama, M. (2022) Experimental Creation of Contemporary Dance Works Using a Body-part Motion Synthesis System. *8th International Conference on Movement and Computing*, no.21, pp.1-5. https://doi.org/10.1145/3537972.3537997

[2] Soga, A., and Matsushita T. (2023) Movement Creation by Choreographers with a Partially Self-controllable Human Body in VR. *29th ACM Symposium on Virtual Reality Software and Technology.* https://doi.org/10.1145/3611659.3617198

[3] Nogueira, MR. et al. (2024) Exploring the Impact of Machine Learning on Dance Performance: A Systematic Review, *International Journal of Performance Arts and Digital Media*. https://doi.org/10.1080/14794713.2024.2338927



Andrea Palladio, 1570

ARTWORKS

~ Chromatic Ghost ~
Artworks for GA2024

Chin-En Keith Soo, Rowan Simmons
Department of Design, University of Waikato, New Zealand

Abstract:

"Chromatic Ghost" delves into the convergence of artificial intelligence and the emotional narrative of cinema, exploring how AI-generated colour palettes can visually express the emotional depth of films. The project aims to assess the capabilities of AI in understanding and translating human emotions into visual representations, questioning how effectively machines, which are increasingly embedded in our lives, can interpret the intricacies of our emotional experiences. By using AI to generate emotion- based colour palettes and applying them to film frames, "Chromatic Ghost" transforms iconic cinematic scenes into layered, ethereal images that evoke the emotional core of the films in a spectral, dreamlike manner. The resulting visual compositions—ghostly, fragmented, and nuanced—invite viewers to contemplate both films' emotional resonance and AI's role as a mediator of this experience.

Statement of Work:

"Chromatic Ghost" is a creative experiment designed to investigate how AI can interpret and visualise human emotion in the context of film. In an era where artificial intelligence is increasingly intertwined with our daily lives, this work examines how well these technologies can grasp and reflect the complexities of human emotions, which are often considered too nuanced or abstract for machines to comprehend.

At its heart, this project focuses on the inherently ethereal nature of emotions. Emotions are often elusive, hard to define, and subjectively measured.
"Chromatic Ghost" tackles this challenge by using AI to generate colour palettes corresponding to different emotional states. Each film chosen for this project is reframed through the lens of these AI-generated colour palettes, offering a reinterpretation of the film's emotional landscape.

The central question the project seeks to answer is: Can AI understand and replicate the emotional tones we experience as humans? Using an automated process to apply AI-generated colour palettes to film frames, the project questions the accuracy
and depth of AI's emotional

interpretation. The final output—composite images made from filtered film frames—creates

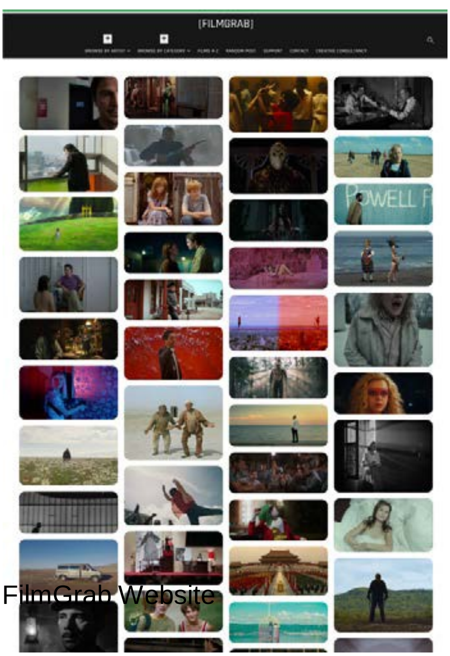
ghostly representations that are at once recognisable and abstract,
like spectral memories of the film's emotional journey. These haunting images are metaphors for the
imperfect yet evocative way AI interprets human emotions, like a fleeting glimpse of something we
can't fully grasp.

Process of Work:

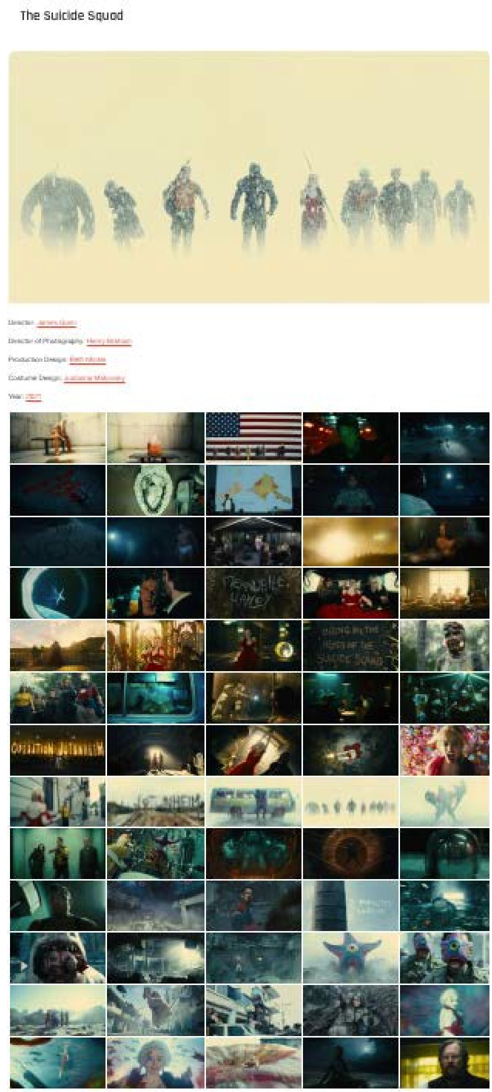
Emotion Identification and AI Colour Palette Generation: The process begins with AI- generated insights into human
emotions. ChatGPT created a list of 20 distinct emotions, ranging from joy to sadness, anger, and fear. For each emotion, the AI- generated a corresponding colour palette of five colours believed to evoke or represent that emotional state. These palettes serve as the cornerstone of the project, representing the AI's interpretation of what each emotion visually "looks" like.

AI-Generated Emotion Colour Schemes

Film Frame Selection and Dataset Creation: The artwork selected approximately 65 frames per film from the FilmGrab database to ensure a structured and consistent approach. FilmGrab offers a repository of film stills from various movies, providing a rich dataset of iconic moments to draw. Sixty-five frames were chosen as this number allows for a broad yet manageable sample size, ensuring that each film's critical emotional shifts and scenes were captured. The selection of these frames was essential to creating a dataset that could be analysed against the AI-generated colour palettes.



FilmGrab Website



FilmGrab: Selected Movie Frames

Colour and Emotion Matching through Pixel Analysis:
With both the colour palettes and film frames in place, the next step was to compare the individual pixels of each frame to the AI- generated colour palette

corresponding to a particular emotion. The process involved cycling through every pixel in each frame and checking whether the pixel colour matched any of the colours from the emotional palette. The saturation has a range of +- 20, and brightness has a range of +-15. If a match was found, a pixel was retained for the final composite image, while all other non-matching pixels were discarded. The matching pixels were then layered at rational opacity, ensuring each layer contributed to the final composite image without overwhelming it.

Individual Colour Frames

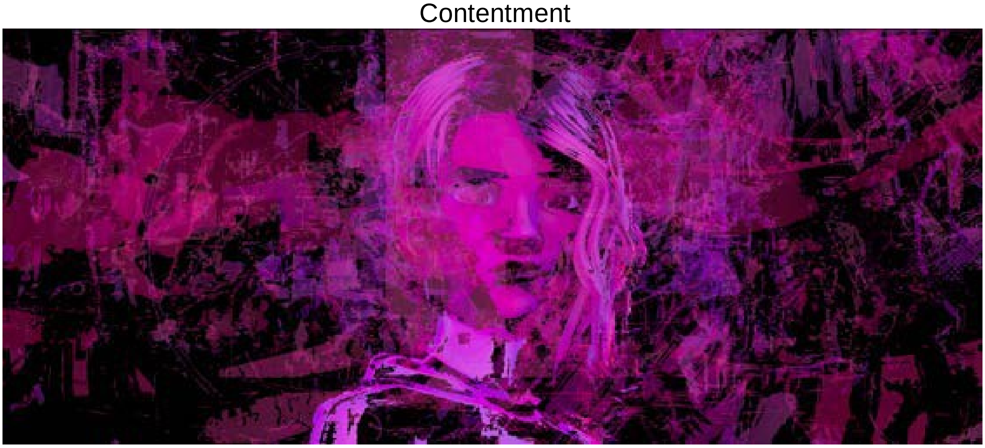
Overlaying Frames

Composite Image Creation and Chromatic Ghost Formation: The retained pixels from each frame were layered sequentially, building up the composite image frame by frame. As more frames were added, the image began to take shape, gradually forming a visual representation of the film's emotional landscape. These composite images were dubbed "chromatic ghosts" due to their translucent, spectral quality. They resemble faded, haunting portraits reminiscent of old Victorian photographs where time has eroded the details, leaving only traces of the original image behind.

The result is a composite image that visually represents the film's emotional core, filtered through the lens of AI's understanding of colour and emotion. Each "chromatic ghost" captures the essence of the film's emotional journey, creating a visual echo that feels both familiar and otherworldly.

Spiderman across the Spider-verse (2023) -
Frustration

Spiderman across the Spider-verse (2023) -
Contentment

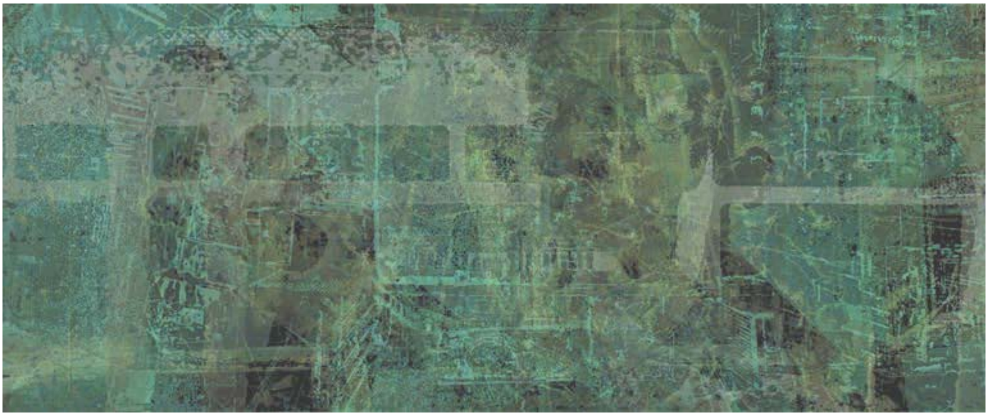


Spiderman across the Spider-verse
(2023) - Love

Wonka (2023) -
Embarrassment

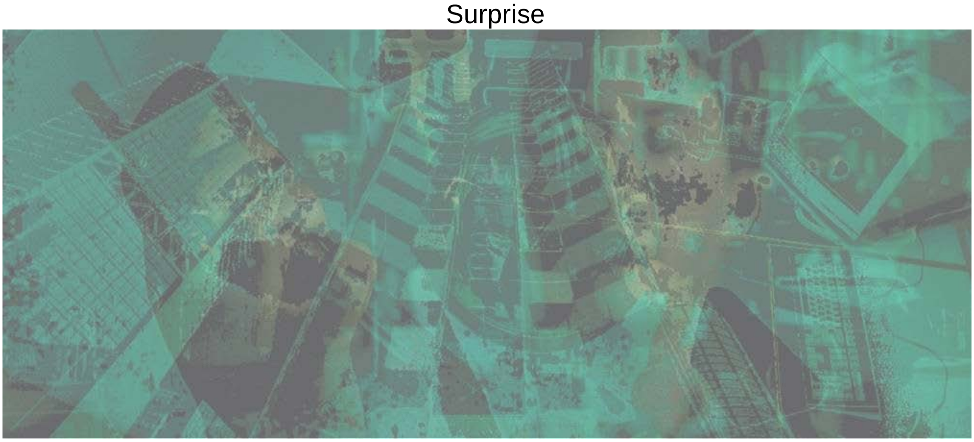
Wonka (2023) -
Jealousy

Wonka (2023) -
Fear



The Matrix (1999) -
Sadness

The Matrix (1999) -
Surprise



The Matrix (1999) -
Relief

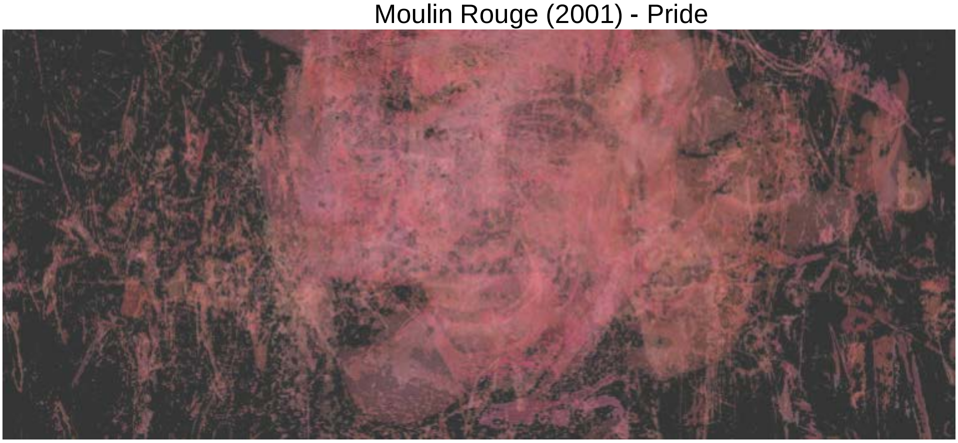
Tron (2010) - Excitement

Tron (2010) - P

Tron (2010) - Surprise

Moulin Rouge (2001) - Excitement

Moulin Rouge (2001) - Pride



Moulin Rouge (2001) - Shame

Star Wars III (2005) - Surprise

Star Wars III (2005) - Confusion

This process reflects the delicate interplay between human emotions and machine interpretation. The "chromatic ghosts" produced by the project highlight the strengths and limitations of AI in emotional communication. While the images retain emotional resonance, they are also fragmented, imperfect, and open to interpretation. The work invites viewers to consider how much of our emotional reality can be captured by AI and how much remains elusive, like a ghost that can be felt but never fully seen.

Isopolis

David Bouchard, BCompSc, MSc
New Media Department, Toronto Metropolitan University, Toronto, Canada
https://www.deadpixel.ca
e-mail: david.bouchard@torontomu.ca

Abstract
This is paper describes a generative artwork titled *Isopolis,* presented as part to the Artworks/Poster track of GA2024. *The full project can be viewed online at* https://deadpixel.ca/isopolis.

"Imagine a grid made of two groups of equidistant parallel lines perpendicular to each other which, when drawn on a plane, would yield a uniform field of equal squares, like a sheet of graph paper. Now imagine that this graph paper, enlarged several thousand times, is nothing less than a stone platform and that in each of the vertices of its unseen network rises a slender column, the architectural abacuses (or flat tops) of which each support four wooden beams arranged along the lines of the grid. [...] This was how the city of Isopolis looked before it was set ablaze by order of Alexander of Macedon" - excerpt from Isopolis by Gheorghe Săsărman

1. Project Description

Isopolis is a reflection on automation, perpetual growth and the ways in which artificial intelligence is contributing to the commoditization of creativity. This aligns with Galanter's [1] generative art theory, which posits that generative systems can produce complex and unpredictable outcomes, challenging traditional notions of authorship and creativity

In *Isopolis*, an infinite cityscape is generated from a random combination of buildings pulled from a large dataset of AI-generated images. The website drifts across this cityscape, getting its direction from mouse input. The viewport slowly moves closer or further away from the viewer over time via a random process.

The underlying image dataset is created using an automated process put in place by the artist, embodying the principles of generative art as outlined by Dorin et al. [2]. According to their framework, generative art involves systems that can autonomously produce art, emphasizing the role of algorithms and randomness in the creative process.

First, an image is generated from a prompt[1] into a text-to-image AI model. The resulting image is passed on to another program which leverages a convolutional neural network trained for object segmentation and background removal. The image is then fed to a third program which automatically crops and algorithmically determines the angle of the bottom corner of the generated building. This final program also performs a selection, ensuring that the image has a suitable composition within the frame and that the isometric projection angles are within acceptable tolerance limits. Unsuitable images are discarded.

Table 1. An illustration of the 3 image generation process steps, showcasing a successful asset pick based on its image composition and the angle of the bottom corner of the building

Table 2. Examples of images discarded by the automated selection process

The result is a theoretically infinitely scalable dataset of images, limited only by available storage and computing power. By adopting an industrial approach to asset generation, I aim to explore how computation, specifically AI tools, can facilitate the creation of visual material at a scale unattainable through manual labor alone. Manovich [4] highlights this shift in *AI Aesthetics*, noting that AI enables artists to generate vast amounts of content, thereby redefining the scope and scale of creative production. This approach also raises questions about the value and cost of applying automation to the creative process. [4]

The sheer number of generated images, combined with the randomization of their composition, makes it impossible to predict the final appearance of the city until the software is executed. Simultaneously, the work questions the value of applying automation and industrial processes to the creative process. What are the costs? Is quantity a worthwhile tradeoff for quality? Isopolis, with its infinite urban sprawl, quickly devolves into a dystopian landscape where the uniquely generated buildings occasionally reveal quirks and glitches in the data. More importantly, the promise of "infinite content" gradually gives way to a monotonous homogeneity.

2. Artist Statement

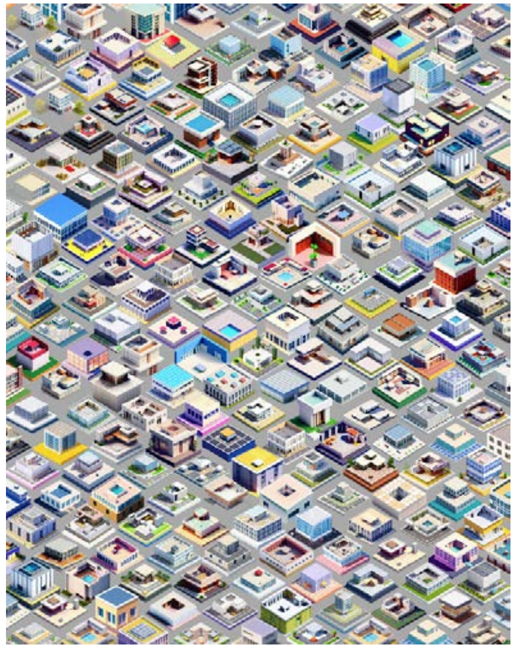
My work navigates the intersection of art, technology, and interactivity, with a particular focus on generative processes. My practice is rooted in the idea that creativity can be algorithmic, with the potential for endless variation and evolution. Art can be a living, evolving entity, one that responds to its environment, audience, and the invisible forces of data and algorithms. By deploying generative systems, I strive to create works that are never static. Rather, they are continuously shifting, evolving in real-time, and influenced by both programmed logic and external variables, creating worlds where randomness and order can coexist.

Most of my works are developed around datasets, whether captured from the environment through sensors, gathered from databases, scraped from the Internet or, as with Isopolis, complete fabrications. The collection, organisation, analysis and manipulation of these datasets play an integral role in my creative process. I then write software that allows both me and, by extension, the viewer of my work to explore, discover, and experience the data.

My process also challenges the notions of authorship in this context, particularly once partial control of the creative process is relinquished to external factors such as data, AI or random noise. I develop the framework and set the parameters, but the final form of the artwork is co-created by the system itself. This approach challenges traditional notions of authorship and control and invites the viewer to become an integral part of the creative process by triggering and sometimes influencing new iterations of the work.

3. Notes

1. The prompt used in Stable Diffusion 1.5*: tiny isometric modern building, hyper realistic, square tile, soft smooth lighting, soft colors, 100mm lens, physically based rendering, centered, plain background, high detail, daytime*



4. References

1. Galanter, Philip. "Generative art theory." A companion to digital art (2016): 146-180.

2. Dorin, Alan, et al. "A framework for understanding generative art." Digital Creativity 23.3-4 (2012): 239-259.

3. Manovich, L. (2019). AI Aesthetics. Strelka Press.

4. Cardoso, Daniel. "Generative Craft A brief critical inquiry into design automation and design automata." Sociedad Iberoamericana de Gráfica Digital (2009): 288-290.

A New Aesthetics of Vastness: Complex Fractal Geometries in Generative AI

Dr. İrem Hafız, BArch, MArch, PhD, Architectural Design
Department of Architecture, Middle East Technical University, Ankara, Turkiye
e-mail: irem.hafiz@gmail.com

İrem Hafız

Abstract

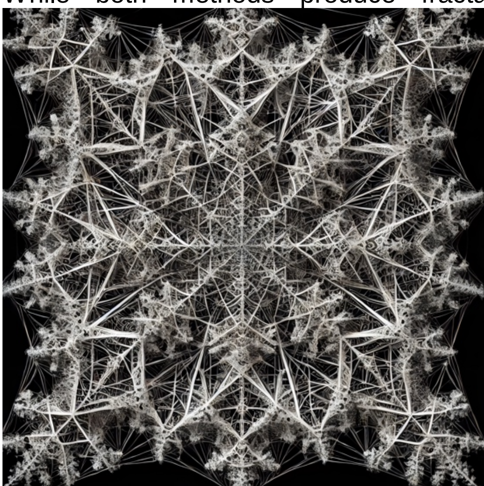
This research and accompanying artwork push the boundaries of traditional aesthetics in design, challenging the conventional beauty and fostering a new perspective on creativity through the vast, emergent, and complex nature of form in generative systems. Drawing from the foundational concepts outlined in my PhD thesis, this experimental work bridges the classical concept of fractals — characterized by repetitive and recursive processes — with advanced computational capabilities using generative AI (DALL-E).

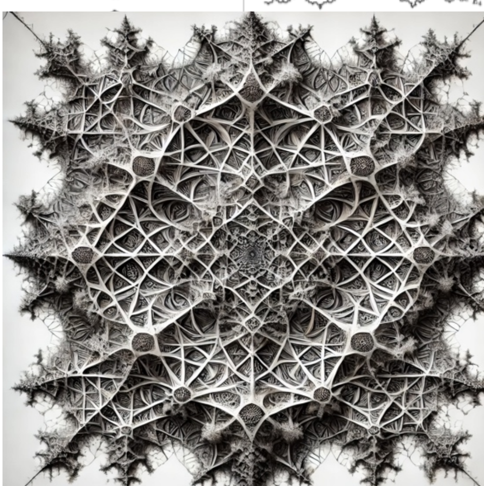
Inspired by Kenneth Falconer’s concept of “random fractals” introducing random variations to the von Koch curve, which makes fractal constructions more akin to natural structures, this research employs DALL-E to explore infinite varieties in the possibilities of fractal formations. In this regard, two distinct methodologies were applied: the first approach involved prompting DALL-E to increase complexity within fractals based on the von Koch curve by adding new variations. However, the second utilized microscopic images of natural formations (cellular structures) with inherent complexity as inputs for DALL-E to transform them into fractal geometries. Both methods yield complex, sophisticated, and data-rich fractal visualizations that evoke a new concept of aesthetics emerging through vastness.

Vastness in Random Fractals and Natural Phenomenon

Despite the inherent self-similarity and predominantly symmetrical organizations of fractal patterns, this artwork explores “non standard” fractals through expanded and enriched trajectories of geometric development [1]. Building on mathematician Kenneth Falconer’s concept of random, non-uniform, and “statistically self-similar” fractals, this approach enables vast possibilities in form generation, revealing unimaginable or unpredictable variations and intricate connections within fractal geometries.

Falconer discusses an alternative approach to fractals by introducing an irregular organization of the von Koch curve. Unlike the traditional method, where each iteration follows a uniform rule, this random construction adds variability at each step by altering the positions of new segments either above or below the replaced parts of the curve. While both methods produce fractal

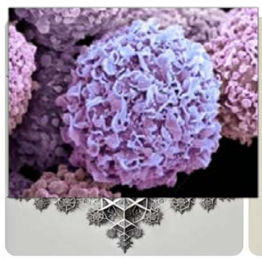




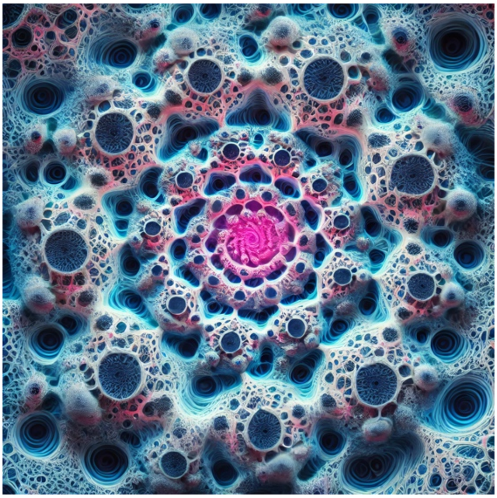
creative potential of fractal formations in alignment with Falconer's discussion of the irregular-looking von Koch curve. The transformative potential of generative AI reveals new dimensions of mathematical intricacy with an unclear symmetry in fractal art.

In the first method, as the initial set of generations, DALL-E was prompted to illustrate a fractal model based on the Koch snowflake method. This resulted in a basic, symmetrical AI-generated visualization of the snowflake curve.

Later, DALL-E was asked to increase the complexity of the two visuals by increasing variation and introducing a less easily recognizable symmetrical organization. DALL-E presented five specific modifications that can be applied to the fractal images to achieve this effect: (1) introducing irregularity in recursion, such as unequally dividing line



segments within the triangles, (2) varying the angles of new triangles, (3) changing the iteration depth in different regions, (4) transforming some straight-line segments into curves, bending outward or inward, (5) manipulating fractal dimensions. Through these methods, DALL-E generated two new models depicted as more complex and less symmetrical versions of the fractal, achieved through the variation in angles, segment lengths, and recursion depths.



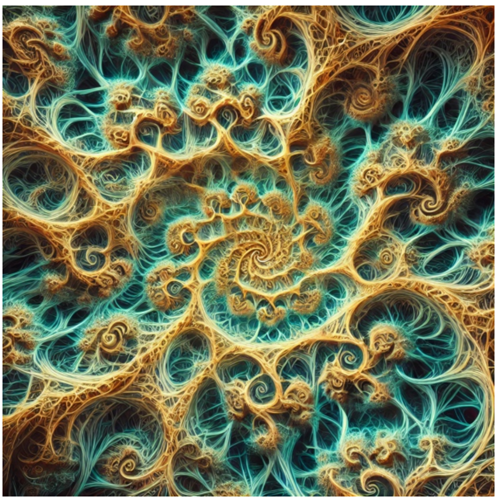


Figure 2: Highly complex fractals based on the von Koch curve by DALL-E.

While the first method prompted DALL-E to enhance variety and disrupt the static symmetry of the von Koch curve, the second approach took a reverse direction. Scanning Electron Microscope (SEM) images capturing the intricate patterns of microscopic organisms in Nature were uploaded to DALL-E, instructing it to transform these biological structures into fractal geometries. Rather than starting with a symmetrical fractal and introducing complexity, this approach sought to uncover fractal patterns already present within the vastness and irregularity of natural forms. By merging the biological essence with the mathematical intricacy of fractal art, DALL-E generated sophisticated fractal organizations that mirror both natural complexity and artificial recursive structures.

Figure 3: SEM images. (Left) Cancer cells. (Right) Slime mold. Source: Science Photo Library.

Figure 4: Highly complex fractal illustrations generated by DALL-E from the two SEM images: the top image is a transformation of cancer cells, while the bottom image depicts slime mold organizations.

Conclusion

Data-driven artistry in AI produces “non standard” fractal models emerging out of vast quantities of data, rather than static, symmetrical, and predictable fractal patterns. This artwork demonstrates that increasing data vastness by introducing variations and irregularity into generative systems leads to complex outcomes, akin to the discontinuous patterns found in Nature. That is to say, these AI-generated outputs, which can be characterized by distinct aesthetic qualities, closely resemble natural formations, highlighting the potential of generative models. All AI-generated visuals are the emergent and unique outcomes of algorithmic processes, evoking a sense of the sublime and contributing to a new aesthetic discourse in design. Therefore, this investigation showcases the transformative potential of generative design in crafting the future of design and aesthetics, in collaboration with advanced technology.

Notes

[1] Zeynep Mennan's concept of the "non standard" originates from her curatorial essay for the international architectural exhibition "Architectures Non Standard" at the Centre Pompidou, Paris (2003-2004). Mennan introduces the non standard as a new concept of form, promoting highly complex, counter-intuitive forms enabled by computational design technologies.

References

[1] Mennan, Z. "The Question of Non-Standard Form." *METU JFA,* 25(2), 2008. https://hdl.handle.net/11511/50905

[2] Falconer, K. *Fractal Geometry: Mathematical Foundations and Applications.* Chichester: John Wiley & Sons, 2003.

[3] Hafız, İ. "Towards an Aesthetic Discourse of Computational Design: Re-contextualizing the Sublime." PhD diss., METU, 2024. https://hdl.handle.net/11511/110386

HandAIwork: Generative AI-enhanced quilted textiles

Anna M. Chupa, MFA
Art Architecture and Design, Lehigh University, Bethlehem, Pennsylvania, USA
annachupadesigns.com
anna.chupa@lehigh.edu

Michael A. Chupa, M.S.
Anna Chupa Designs, Easton, Pennsylvania, USA
chupa@acm.org

Abstract

With the widespread availability of generative AI tools based on large language models, today's artist is faced with a dilemma similar to that encountered by 19th century painters witnessing the advent of photography—either to embrace a new technology and explore new frontiers of artistic expression, or reject the innovations and adhere to established conventions. We contend that an interesting path can be traveled that seeks a middle way between these two extremes by adopting generative AI tools while simultaneously emphasizing the human hand via rendering in textiles with the addition of hand-sewn elements to bridge the gap between the human and the machine.

Like the artist, today's engineers are also faced with similar choices in the era of generative AI tools. In our work, we have adopted LLM tools as adjuncts in building and refining the code used to generate the patterns rendered as quilted fabrics.

1. Background and Motivation

The theme of this year's Generative Art conference includes the challenge "to preserve the human complexity, with Generative Art and AI," and this work endeavors to incorporate the use of Generative AI tools to enhance, but not replace, the human touch in finished work. A button box inherited from my grandmother inspired a two-dimensional layout of buttons, new and old, in a configuration within a rectangular domain of arbitrarily chosen size. This placement of point-like objects (buttons) within the domain suggested a spatial subdivision of the rectangular domain into triangles, with the button locations coinciding with the triangle vertices. Such subdivisions are commonly applied in computational modeling applications such as finite element analysis or fluid flow simulations, and serves to underscore this work's intent to juxtapose the human and machine-made elements in the context of an artistic composition.

2. Methodology

The button placements were initially photographed on a rectangular grid template, and the locations and relative sizes were provided as inputs to Python scripts. But in keeping with this year's conference theme, the Google Gemini [1] large language model (LLM) was employed to generate code based upon natural language prompts. For example, each of the button locations on the finished work has a regular hexagon stitched around the button location. This Python code was initially generated using the prompt, "Write a Python script to take a list of 2D points, create a regular hexagon with side length *L* around each point, and output the hexagons as an image in SVG format," where SVG indicates the Scalable Vector Format. [2] The resulting Python code was reasonably efficient, but modifications were made to scale the hexagons with the button size, since attempting to specify this in the initial text prompt to Gemini was unweildy. Additionally, slight modifications to the SVG output were made for compatibility with the software toolchain used to prepare the computer-driven quilting machine employed in the physical construction of the quilted fabric. In particular, significant advantages of employing an LLM included the detection and correct inclusion of an SVG format Python library, `svgwrite` [3], along with an easily-modified function to create the hexagons.

Likewise, the decomposition of the rectangular domain into triangular elements was performed with Gemini's assistance, using a prompt to "Write a Python script to take a list of 2D points, and output the Delaunay Triangulation [4] of the points to an image in SVG format." In this case, additional packages were specified, including `numpy` [5] and `scipy` [6]. Again, only minor modifications were required, and while Gemini generates a caveat: "Use code with caution," the significant time savings over hand-written code is considerable. In the case of the triangulation code, only slight modifications were required for compatibility with the toolchain.

Some hand-modifications were applied to the resulting SVG file to reorder some of the geometry to create more efficient stitching paths, as any discontinuities in the sewing head's trajectory over the quilt require the snipping and "burial" of the loose threads within the quilt batting material.

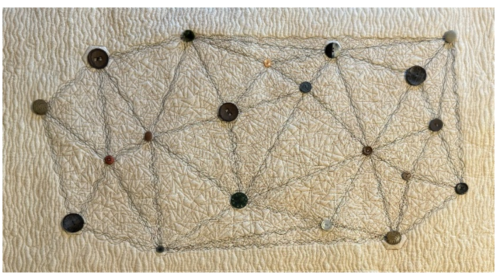
The software application Art and Stitch [7] was employed to convert the SVG file into the proprietary file format needed for the Innova [8] Longarm Quilting Machine (.PAT) to render the vector graphic as stitches. The Innova Longarm was used to automate the stitching of the converted AI-generated file onto fabric.

2. Quilting Process

A longarm machine [9] is a sewing machine with more distance between the sewing head (needle) and the back of the machine (throat) than that of a standard domestic sewing machine. The sewing head moves on rollers forward and back, right and left on a frame. The quilter moves the sewing head along the fabric usually by holding handle bars. Hand-guiding the sewing head is considered to be free-motion. The machine can also be computer-automated by engaging quick release belts to a track system; both methods were used in this work's production.

The quilt sandwich consists of a backing, a batting usually made of wool or cotton, and a quilt top. These are basted together with a loose stitch. The quilt sandwich is attached to the bars of the frame with leaders so the quilt remains stationary while the sewing head moves.

Once the automated portion was complete, freehand elements were added. These include the meandering curvilinear lines in dark thread, the angular intersecting zigzagging lines in variegated gold thread, and the wavy light-yellow thread for background stitching. Buttons were hand-sewn at the vertices to finish the piece. We sought a balance between automation and freehand artistry, rectilinear and curvilinear, dark and light thread colors.



HandAIwork 1: A Fine Mesh

References

[1] Google Gemini, https://gemini.google.com

[2] Scalable Vector Format, https://www.w3.org/TR/SVG2/

[3] svgwrite, https://github.com/mozman/svgwrite

[4] Delaunay Triangulation, https://mathworld.wolfram.com/DelaunayTriangulation.html

[5] numpy: The fundamental package for scientific computing with Python, https://numpy.org/

[6] scipy: Fundamental algorithms for scientific computing in Python, https://scipy.org/

[7] Loes van der Heijden, *Art and Stitch* 2010–2024, https://artandstitch.com

[8] Innova Longarm Quilting Machines, ABM International Inc. Industrial Quilting Machines, 2024. https://innovalongarm.com

[9] The Quilt Show, "What is a Longarm," Quiltipedia, 2021. https://thequiltshow.com/quiltpedia/what-is-a-longarm

GoGo Musebots: A Generative Co-creation with Human Musicians

Arne Eigenfeldt
School for the Contemporary Arts, Simon Fraser University, Vancouver, Canada
www.sfu.ca/~eigenfel e-mail: arne_e@sfu.ca

Abstract

GoGo Musebots is a co-creation between a generative system and its creator, as well as three improvising musicians. The system is routed in composition rather than improvisation, in that plans are created, then filled in by musical agents (musebots) by creating a score; musebots can edit their individual parts, making decisions based on global structures and local events by other musebots. The final score is translated into MIDI information – to be performed by robotic instruments – and lead-sheet notation – to be performed by humans.

All titles are generated by algorithm, selecting word combinations from Samuel Beckett's *The Unnameable*.

Performed by John Korsrud, trumpet; Jon Bentley, saxophones; and James Meger, bass and mechanical instruments.

1. Genesis

A few years ago I attended a concert of the UK's "emotive, cinematic break-beat trio" GoGo Penguin when I was completing my large-scale generative production "A Walk to Meryton", and I thought "I bet the musebots could do something like that". A few alterations and tweaks to the system and the result was generative music for jazz trio with robotic drums and Disklavier.

Like "A Walk to Meryton", the concept was to create a system that generates a structure of repeating sections each containing unique harmonic progressions and melodies, similar to a typical jazz tune. Human musicians could approach the works as they would any jazz tune notated with a lead sheet, playing the melodies and/or improvising over the chord sequences. The accompanying piano and drums would vary their parts with each performance, controlled by musebots.

2. Musebots

Musebots [1] are independent intelligent agents that generate both an overall musical structure, and then create the details within that structure. The musebots are "intelligent", in that they have learned about their environment and communicate their intentions and coordinate conditions for collaborative machine composition [2].

3. Links

The four generative works can be found here:

1. Then I Was In

https://tinyurl.com/4tzxedz2

2. Clearing Prior to Empty

https://tinyurl.com/3974b4a7

3. Little Grey Wizened Pear

https://tinyurl.com/yz5rakfc

4. No Worse, Little Bounds

https://tinyurl.com/2mtkdrbj

Frame from "Then I Was In", the first video of "GoGo Musebots".

4. References

[1] Bown, O., B. Carey, and A. Eigenfeldt. Manifesto for a Musebot Ensemble. ISEA, Vancouver, 2015.

[2] Eigenfeldt, Arne. Collaborative Composition with Creative Systems. ISEA, Durban 2018.

Dream of a city

Artworks

Slawomir Wojtkiewicz, PhD
Department of Visual Arts, Faculty of Architecture,
Bialystok University of Technology, Poland
www.pb.edu.pl
www.artwojtkiewicz.com
e-mail: s.wojtkiewicz@pb.edu.pl
Karolina Masłowska
Academy of Fine Art in Gdansk, Department of Architecture, Poland

Abstract

As part of the exhibitions accompanying the conference, we aim to present a series of graphics that explore the theme of our home city. The works, created using classical techniques and enhanced with generative tools and AI, delve into the complex of a contemporary city. They resonate with the present, serving as an artistic exploration of the city's identity. By employing modern generative tools, we strive to convey emotional narratives that resonate identity and "spirit of place" of the city. We believe that contemporary art and its tools can articulate poignant stories that bridge past and present, evolving into a universal language of artistic expression.

The paper is supplemented by graphic artworks inspired by the generative art process, reflecting the essence of global city. These works, deeply connected to their specific social contexts, will be showcased alongside a gallery of graphics that draw inspiration from modern architectural history.

Our exhibition serves as a definitive showcase and study of generative art, an exciting and emerging field.

For several years, generative art has generated considerable excitement among media art festival and conference participants. Through the interplay of complex information, graphic design, and programming, new and fascinating visual worlds are emerging—worlds in which randomness is harnessed to reveal hidden correlations. The evolving capabilities of programming languages and AI are transforming the role of the designer and the very nature of art. We are witnessing a paradigm shift in design that opens new realms of visual expression. However, understanding this shift has, until now, been somewhat inaccessible. Traditionally, designers have used tools developed by programmers, which required them to adapt to existing systems. With generative design, however, users of digital tools can now become the programmers of their own customized toolkits. This fundamentally alters the design process: the technical aspects fade into the background, replaced by abstraction, while information is processed at a meta level. Generative art begins not with formal questions but with an exploration of phenomena.

As part of the exhibitions accompanying the conference, we aim to present a series of graphics that explore the theme of the contemporary city. The works, created using both classical techniques and enhanced with generative tools and AI, delve into the complex and turbulent history of the modern city. These pieces resonate with the present, offering an artistic exploration of urban identity. By employing modern generative tools, we strive to convey emotional narratives that resonate locally, shedding light on the lost identity and "spirit of place" within the city. We believe contemporary art and its tools can articulate poignant stories that bridge the past and present, evolving into a universal language of artistic expression.

This paper is accompanied by graphic artworks inspired by the generative art process, reflecting the city as an ongoing source of inspiration. Our exhibition serves as both a showcase and study of generative art, a fascinating and emerging field.

The initial concept emerged while exploring the archive of an architectural photographer, where we were struck by images featuring prominent perspectives, foreshortening, and vanishing points. This led us to experiment with these visual features. In the early stages, we tested various algorithms for image analysis and segmentation, combining them in different ways. By layering small snippets from different photographs and aligning them to a common vanishing point, we created small collages of architectural spaces with a shared perspective. Fascinated by the results, we ran algorithms through software tools to create much larger collages.

The final process involves two separate tools. First, high-resolution source images are analyzed and categorized according to their vanishing point and the shapes they contain. Based on this analysis, the images are digitally sliced into many pieces, retaining information about their positions relative to the vanishing point and perspective lines. We developed an algorithm to handle this semi-automatic analysis and slicing, producing a large pool of fragments ready to be applied to new perspectives and shapes in future compositions. These fragments, along with information about their original perspectives, form the foundation for creating new collages.

A second tool provides an interface for composing new collages. To create a collage, a rough geometric framework is defined, and image segments that match this framework are selected. The segments are never altered to fit the frame; rather, appropriate segments are chosen from the pool of possible pieces. By defining additional keywords that describe the content of the original photographs, we can influence the selection of segments, adding a semantic layer to the composition. The resulting decompositions blend the perspectives of both the architect and photographer with a newly chosen frame, making each final piece unique. This method holds great potential for further experimentation and can be considered a work in progress.

Below, we present some of our works along with brief descriptions of each made by Karolina Maslowska:

The Brion Tomb

This work combines raster and vector graphics to represent the Brion Tomb. Inspired by the spirit of the place, I sought to express it through my own lens. I created a series of handmade textures using black ink and graphite. Initially, I experimented with blank sheets of paper, crumpled them, tore them up, and scanned them. This process gave me a diverse set of textures, which I often incorporate into my works.

In this piece, I also experimented with lines—almost every surface is filled with thin strands, which gives the graphic a three-dimensional quality, as though you could step into it. What I appreciate most about this project is that when you view the entire image, you recognize the stairs, windows, and walls, but upon closer inspection, you see only a set of lines running in different directions. It's as if every section could stand alone as its own composition.

Pavillon Le Corbusier

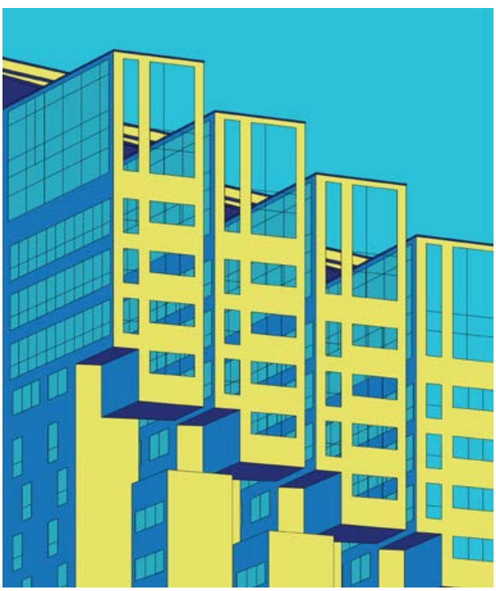
This graphic was inspired by the Zurich Pavilion designed by Le Corbusier, with the staircase as the central element in the composition. In this work, I combined raster graphics with vector elements to achieve the desired contrast. The central pillar of the staircase disrupts the plain blocks of color and acts as the focal point. I used graphite powder, scanned it, and applied it as a mask to add texture to the wall.

The color palette was chosen to reflect the emotions I experienced while studying the composition. I aimed to create an abstract look, similar to the railings along the stairs. These railings, though simple, tell their own story as they intersect with one another. The composition is filled with numerous lines, turning the staircase into something dreamlike.

Gdańsk Shipyard

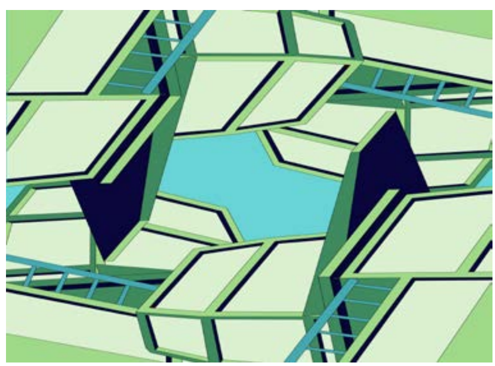
This graphic aims to capture the spirit of the Gdańsk Shipyard from my own perspective. The heavy, industrial machinery is softened, appearing lighter and warmer, inviting the viewer to see the space in a new light. Monumental cranes and post-war buildings are transformed into abstract compositions with carefully chosen frames. These works emphasize the architectural aspects of the dockyard while offering a fresh perspective on how to perceive it.

The graphics are dominated by geometric shapes, sharp lines, and saturated colors, accentuating the raw character of the space. Yet, they also break it down into a calm, silent place. Although the images represent real locations, the frames suggest something more abstract. The monumental machinery loses its practical function, becoming a set of lines and shapes. The Gdańsk Shipyard, which has been partially converted into a space for nightlife with bars, restaurants, and concerts, is reimagined through my graphics as a tranquil and somewhat surreal environment.



Kosciuszko Square

These ten graphics are based on Kosciuszko Square in Gdynia. My goal was to transform the bleak, wintry atmosphere into something captivating. Each piece uses a different color palette, but even when viewed together, they maintain a consistent feel. I focused on individual elements of the square and reimagined them in fascinating ways—like a chain with an anchor hanging from a boat or a lighthouse. I experimented with multiplying and rotating these elements to create abstract compositions. In some works, I combined multiple objects, while in others, a single element was repeated and transformed. These graphics are dynamic and chaotic, offering a stark contrast to their source material. Sometimes, it's unclear what you're looking at, while in other works, the subject is unmistakable.



References

[1] Stiny G, Mithell W J, (1978). The Palladian grammar. Einvoronment and Planning B: Planning and Design 5 5-18

[2] Mandelbrot M., (2004) "Fractals and chaos" The Mandelbrot Set and Beyond". Springer. New York, N.Y., U.S.A.

[3] McCormack J., Dorin A., and Innocent T., (2004) "Generative Design : a paradigm for design research" in Redmond, J. et. Al. (eds) Proceedings of Future ground Design Research Society , Melbourne.

[4] Simon H.(1996), "The Sciences or the Artificial", The MIT Press Cambridge, Massachusetts, USA.

Autopoietic Neural Network – Dynamic Images Encoded as a Self-modifying Neural Nets

Lucas Bang
Department of Computer Science, Harvey Mudd College, Claremont CA, USA
www.cs.hmc.edu/~bang
e-mail: bang@cs.hmc.edu

Abstract

In this short paper, I describe *Autopoietic Neural Network*, a video installation of a live neural network that continuously reads and rewrites its own internal representation. As the network operates, artificial neuron activation levels are mapped to RGB space, creating an ever shifting pattern of colored squares. This recursive, self-modifying process draws attention to the paradox of a system engaged in self-directed change without awareness or agency, inviting viewers to contemplate concepts of autonomy, self-perception, and identity within artificial and human contexts. By appropriating a familiar machine-learning tool to reflect upon itself, Autopoietic Neural Network embodies a critical exploration of the potentials and limitations inherent in self-referential AI systems. In doing so, it reflects on the broader implications of autopoiesis in digital art, prompting the audience to consider both the boundaries and capacities of autonomous systems.

1. Artwork Description

Autopoietic Neural Network is a video installation that visualizes a neural net continuously reading itself as input, outputting and executing edits to its own internal activation matrices. The live, dynamic, recursive process is visualized by encoding activation weights between artificial neurons as RBG pixel intensities (Figure 1).

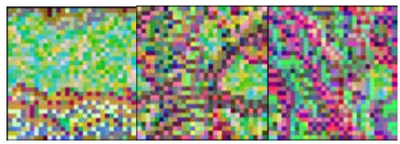


Figure 1: Snapshots of Autopoietic Neural Net Self-editing Over Time

This work invites the viewer to consider the ostensibly paradoxical, yet internally consistent, cybernetic-reflective feedback loop of a system that manually edits its own behavior. By reflecting on machinic capacity for (lack of) self-determination, the viewer may consider their own agency. For those familiar with neural networks and generative art or not, they may ask themselves “what meaning can we find in our own practices when we observe a system mindlessly churning away on editing itself?”

This work was previously featured at the 2022 Society for Literature, Science, and the Arts (SLSA) exhibition “At a Juncture”, with the theme “Reading Minds: Artificial Intelligence, Neural Networks, and the Reading Human” (Figure 2).

Figure 2: Autopoietic Neural Network installed at SLSA 2022

2. Broader Relevance

In this section, I briefly (i) elaborate on the use of autopoiesis in making self reflective generative art, (ii) discuss the technical details of the implementation, and (iii) describe opportunities and limitations of neural-net-based generative art based on my experience and process producing this artwork.

i. Autopoiesis in Generative Art. The title of the work, *Autopoietic Neural Network*, derives from the Greek *auto* (self) and *poiesis* (making). As this artwork constantly rewrites itself, I borrowed the word autopoietic from the Chilean biologists Varela and Maturana who introduced the word 'autopoiesis' to describe systems that are self-creating and self-maintaining through continuous regeneration of their own components and processes[1]:

> "An autopoietic machine is a machine organized (defined as a unity) as a network of processes of production (transformation and destruction) of components which: (i) through interactions and transformations continuously regenerate and realize the network of processes (relations) that produced them; and (ii) constitute it (the machine) as a concrete unity in space in which they (the components) exist by specifying the topological domain of its realization as such a network."

Autopoiesis is then a natural framework for self-modifying generative artworks.

ii. Implementation Details. *Autopoietic Neural Network* makes dual usage of matrices of floating point numbers. Refer to Figure 3 for the following discussion.

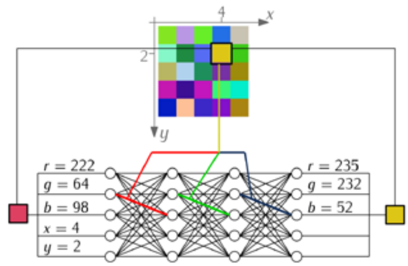


Figure 3: Implementation Sketch of Autopoietic Neural Network.

The activation weights between layers in a neural network are typically stored as matrices. For a network with 4 layers, there are 3 intermediating weight matrices. Similarly, a digital image can be represented as 3 matrices: the red, blue, and green pixel intensities. Thus, a 4-layer neural network with N neurons per layer corresponds to a square RGB image, meaning any square RGB image can be interpreted as a neural network. By feeding pixel coordinates and RGB values through the network, the outputs can guide changes to the image (i.e., network weights) and determine which part to examine next. Similar work on self-referential neural networks has been explored in non-art contexts [2,3].

iii. Opportunities and Limitations. A challenge in AI-generated art and language models is self-referential pollution: "model collapse"[4]. This occurs when AI systems are trained on data generated by AI systems, leading to a degradation in quality and originality, highlighting the limitations of AI when it lacks exposure to diverse, high-quality data. This risk is particularly concerning for AI systems that rely on their own outputs, as it can result in repetitive and less innovative work over time.

In the context of generative art, current systems often exhibit a limited range of aesthetic styles, producing works that can feel derivative or repetitive. While AI-generated art can be impressive, these systems tend to operate within a narrow creative spectrum. However, this limitation is expected to improve as AI systems are trained on more diverse data sources, leading to richer and more varied artistic outputs. Still, the challenge remains: self-referential systems risk stagnation if they are not trained with new, external inputs.

Autopoietic Neural Network directly engages with these challenges by embodying a recursive, self-modifying process. Much like AI models that train on their own outputs, this artwork reads and rewrites its own internal state, creating an evolving visual landscape. However, *Autopoietic Neural Network* uses this self-referential loop as a conceptual and visual exploration of the potential risks and rewards of such systems, emphasizing the tension between the potential of autonomous systems and the risk of stagnation when they lack external influence.

Through this self-modifying process, the artwork provokes reflection on the future of AI-generated art: It underscores the possibility that self-referential AI systems could lose originality if they rely too heavily on their own outputs, but it also celebrates the autopoietic, recursive process as a form of creative inquiry. In doing so, *Autopoietic Neural Network* serves as a critique and a meditation on the opportunities and limitations of generative art.

3. Conclusion

Autopoietic Neural Network explores the creative and conceptual limits of self-referential AI, using recursive modification to question agency, creativity, and autonomy in artificial systems. As AI-generated art grows, issues like self-referential collapse and creative stagnation gain importance, and this installation encourages reflection on these challenges. Through a system that rewrites itself without intent, Autopoietic Neural Network invites viewers to reconsider the nature of self-creation and our role in shaping autonomous systems, fostering dialogue about the future of human-machine creativity.

References

[1] Maturana and Varela. *Autopoiesis and Cognition: The Realization of the Living*. Boston Studies in the Philosophy and History of Science. Springer Netherlands, 1991.

[2] Irie, Schlag, Csordás, and Schmidhuber. *A modern self-referential weight matrix that learns to modify itself*. In International Conference on Machine Learning. PMLR, 2022.

[3] Schmidhuber. *A 'self-referential' weight matrix.* ICANN'93: Proceedings of International Conference on Artificial Neural Networks. Springer, 1993.

[4] Shumailov, et al. *AI models collapse when trained on recursively generated data*. Nature (2024).

Fluid Time Variation
Animated Projection Artwork

Prof. Edward S. Johnston, MFA, MEd
Robert Busch School of Design, Michael Graves College, Kean University
Union, New Jersey, USA
www.edwardsjohnston.com
e-mail: jedward@kean.edu

Abstract

In my ongoing *Experiential Extensions* series, I engage in algorithmic processes to craft compositions from visual artifacts of my experiences. I record and explore my surroundings seeking out unique perceptual shapes from the enclosed edges of visual phenomena. I call these shapes *experiential primers*. Sometimes, they are the perceived edges of clouds, the unique shapes created from the visual contrast of light and shadow, or even the overlapping of tree branches. The time collection of each experiential primer informs the algorithmic process to create form. I explore using music to further inform and guide the creative process. This series fits within a larger art movement called *voxelism* articulated in this paper.

Key Process Images

My exhibited artwork for this paper is an algorithmic animation entitled *At Forest's Edge – Fluid Time Variation ADK*. Key moments from the composition are included below.

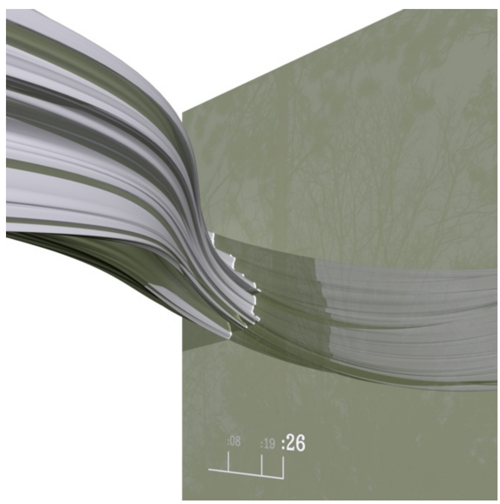


Figure 1: View of digital form through experiential primer

Figure 2: First moment of experiential primer collection

Figure 3: Collected experiential primer at the 8-second mark

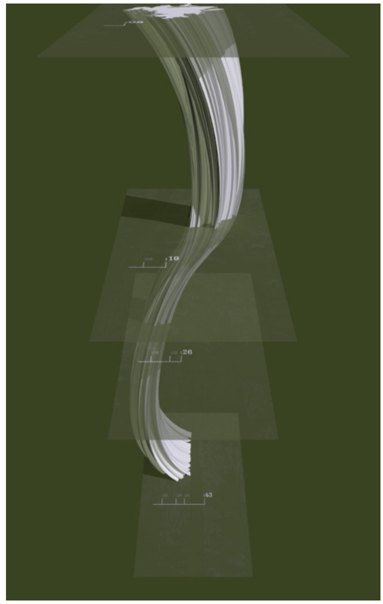


Figure 4: Four primers spaced according to time marks with voxelation

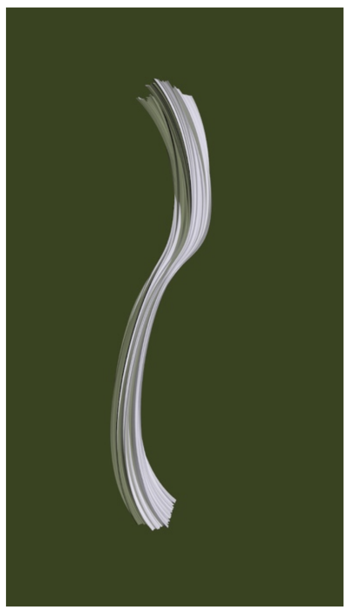


Figure 5: Voxelation form

Reflection and Inspiration

My creative work is rooted in my personal narrative, developed over years of exploration. I find inspiration from many different fields and perspectives. This brief art paper shares my reflections on my creative processes interwoven with bits of contextual inspiration.

In *Postproduction*, Nicolas Bourriaud states, "Artists today program forms more than they compose them: rather than transfigure a raw element (blank canvas, clay, etc.), they remix available forms and make use of data" [1].

No longer can I trust images, what I see on screens. AI has disintegrated that. And that does not stop at screens. Quantum physicists are exploring human existence as a fuzziness in the simulated matrix of the universe. Well, I can still work with that. In fact, that brings a mystical quality to the everyday, to the perceived banal. Within that fuzziness, I can push back in my own acts of collection, how and when I collect images. And I can collaborate with machines to realize unique forms, unique crystallizations of those moments.

In *New Media and Art*, Michael Rush shares, "...'Interactive' has emerged as the most inclusive term to describe the type of art of the digital age. Artists interact with machines (a complex interaction with an 'automated but intelligent' object) to create further interaction with viewers who either summon up the art on their own machines or manipulate it through participating in pre-programmed routines..." [2].

Numerous sources in science and human culture are challenging notions of reality, proposing that there are hidden layers of existence that undermine fundamental understandings. The glitches in the Wachowskis' *Matrix* series, the time portals of Richard Kelly's *Donnie Darko*, and the selection of versions of reality while in superposition in Blake Crouch's *Dark Matter* certainly stand out. I creatively respond by going on conscious journeys and collecting.

Stilgoe suggests, “Go outside and walk a bit...long enough to take in and record new surroundings. Flex the mind, a little at first, then a lot. Savor something special. Enjoy the best-kept secret around—the ordinary, everyday landscape that rewards any explorer, that touches any explorer with magic...Outside lies utterly ordinary space open to any casual explorer willing to find the extraordinary" [3].

As a human being living in contemporary society, I have internalized the mechanistic. I understand a minute of my existence as sixty seconds thanks to the Sumerians and Babylonians a few thousand years ago. While my understanding of time is shaped by human history, I can incorporate it directly into creative practice to discover generative form. I can use it to find enchantment in my every day.

As Bourriaud notes, "Art tends to give shape and weight to the most invisible processes. When entire sections of our existence spiral into abstraction as a result of economic globalization, when the basic functions of our daily lives are slowly transformed into products of consumption...it seems highly logical that artists might seek to rematerialize these functions and processes, to give shape to what is disappearing before our eyes" [4].

In the digital world, there is the pixel, a picture element, represented as a 2D element in a composition. And there is the voxel, a three-dimensional counterpart to a pixel, representing a value in 3D space. I am defining *voxelism* as a creative movement involving re-dimensionalizing recorded experiences as expression, the creative process of collecting specific information and remixing that information to take on volume and other attributes associated with recorded layers of experience.

Bourriaud continues, "...art brings collective scenarios to consciousness and offers us other pathways through reality, with the help of forms themselves, which make these imposed narratives material" [5].

I craft compositions from specific visual coincidences collected during my exploration of my environment. I enter into this creative process by deliberately traversing my surroundings, seeking out specific visual coincidences, recording them, and re-dimensionalizing that information. I consciously pause my everyday reasons for traversing space. I focus my attention, invite unique perceptual shapes to emerge from the enclosed edges of visual phenomena and digitally record them. Sometimes, they are the perceived edges of clouds, the unique shapes created from the visual contrast of light and shadow, or even the overlapping of tree branches and buildings. I call these emergent shapes *experiential primers*.

Barth states, "Perception, should we accept this loose term, is generally defined as extension: the mechanism by which we encounter space...Extension is indeed dramatic and full of untapped permutations, but it is so constant and so constantly active that it rarely pauses for reflection" [6].

My time collection of each primer informs my algorithmic process to *voxelate*, to create form. The *voxelations* take on the form of intricately detailed volumes in digital space through automatic, algorithmic processes. Sometimes, the digital forms take on physical form through digital fabrication methods.

John May writes, "…Automation is the concretization of technical objects at points of genesis where the mental and the material codetermine each other in repetitive gestures, giving rise to new physiological processes and new lived experiences" [7].

Digitally collecting these specific visual coincidences allows me to enter a creative state, where I am actively engaging with my surroundings, allowing information to emerge and using that specific information to algorithmically discover unique forms. I explore using music in the animated compositions to further inform and guide the creative process.

John May continues, "Image-models contain simulations of all possible futures. The radical difference between imaging and previous forms of simulation is that what imaging simulates is not specific ideas or gestures but rather thinking itself" [8].

Voxelations are simulated expressions of consciousness, realizations of my experiences and processes that aim to inspire the audience to find enchantment in everyday moments. The algorithmic forms that emerge from this process embody consciousness and the strangeness of infinite outcomes. I am fascinated by these unique, precise coincidences, and this body of work aims to share that fascination with an audience.

References

1. Nicolas Bourriaud, *Postproduction: Culture as Screenplay: How Art Reprograms the World* (New York: Lukas & Sternberg, 2005), 17.

2. Michael Rush, *New Media in Art* (New York: Thames & Hudson, 2005), 171.

3. John Stilgoe, *Outside Lies Magic: Regaining History and Awareness in Everyday Places* (New York: Walker and Company, 1998), 2.

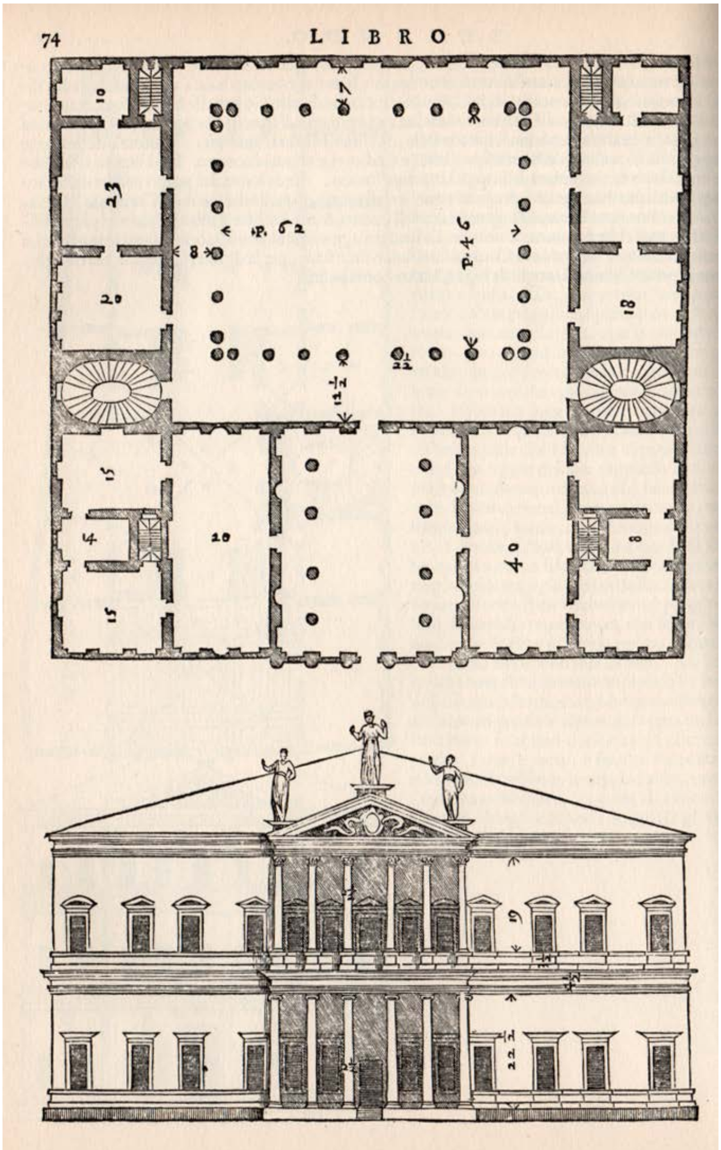
4. Nicolas Bourriaud, Ibid, 32.

5. Nicolas Bourriaud, Ibid, 46.

6. Uta Barth, *white blind (bright red)* (Santa Fe: SITE, 2004), 36.

7. John May, *Signal. Image. Architecture.* (New York: Columbia Books on Architecture and the City, 2019), 75.

8. John May, Ibid, 97.



Andrea Palladio, 1570

VIDEO

Dance Code

Prof. Dr. Anna Ursyn with global collaborators.
University of Northern Colorado
Ursyn.com
e-mail: ursyn@unco.edu

Anna Ursyn

Please see the video and the codes at:

Ursyn.com, then select

Art & Programming.

Credits are listed at the end of the video.

Abstract

In this collaborative work, I tried to introduce some basic concepts in programming to those afraid of coding while using metaphors.

Dance Code

Current creative endeavors involve cross-disciplinary collaborative efforts. Even if someone creates alone, the technology involved in the process has already been developed by somebody else or needs more expertise. Many majors in computer graphics, digital media, and graphic design studying at the art department display some fear of coding. However, employers see that skill as a necessary bridge between artists and coders. A scare-free way is offered here to familiarize oneself with coding using metaphors. The text makes the readers familiar with principles that link the arts and coding, helps them comprehend basic principles in programming, and meets some requirements of the contemporary job market. Understanding basic concepts about coding helps people write prompts to obtain the most interesting solution. It might also support prompt writing.

References

- Dance Code: Dance Steps as a Code: Taylor & Francis, 2025

- New Storytelling: Learning through Metaphors, Taylor & Francis, 2025
- Code Appreciation: Reshaping Knowledge: Taylor & Francis, 2025
- The Transfer of Knowledge Through Art and Visualization Anna Ursyn Taylor & Francis Group. Routledge. 2024.
- The Art of Coding: The Language of Drawing, Graphics, and Animation. by al-Rifaie, M. M., Ursyn, A., Wyeld, T. CRC Press, Taylor & Francis Group, 2020.
- Describing Nature Through Visual Data. Anna Ursyn (IGI Global, 2020,
- Graphical Thinking for Science and Technology Through Knowledge, Visualization (Advances in Multimedia and Interactive Technologies) (Anna Ursyn Author// E d i t o r) , I GI Global. 2020
.

The Alphabetic Labyrinth

Chad Eby, MFA
School of Art and Visual Studies, University of Kentucky, Lexington, Kentucky, U.S.A.
https://chadeby.studio
e-mail: chad.eby@uky.edu

Abstract

The Alphabetic Labyrinth, stereo sound piece with 1920x1080 single channel video, 2′10″ duration.

This sound piece is generated by programmatically stimulating (via virtual control voltage) a virtual modular audio system capable of primitive speech synthesis. Parameters of speech are modulated by a simulation of the motion of a network of compound pendulums.

By fits and starts, the voice attempts to recite the English alphabet. The complete patch is generative in nature, running autonomously once set in motion, and unlikely to ever to precisely repeat.

The light, color, and motion present in the accompanying video is driven by a series of Fast Fourier Transform analyses of the sound in a series of frequency bands, as well as a simulated flocking behavior based on the Boids algorithm developed by Craig Reynolds in the mid-eighties. [1]

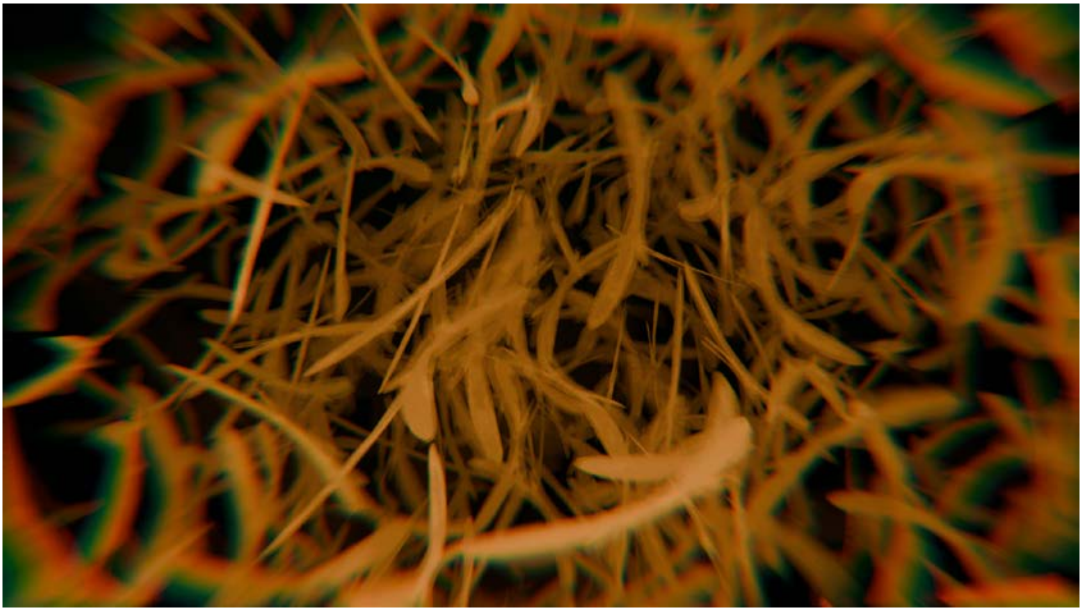
Morphing organic forms swirl and bank in a murky void, their unstable forms glinting uncertainly in the beams of roving colored lights as they expand, contract, and twist under the influence of sonic energies that also heat the objects along the blackbody radiation curve.

The effect bears a relationship to visual music experiments of the early 20th century as well as output from the Prometheus group's work on developing ambient information displays for Soviet-era spacecraft.

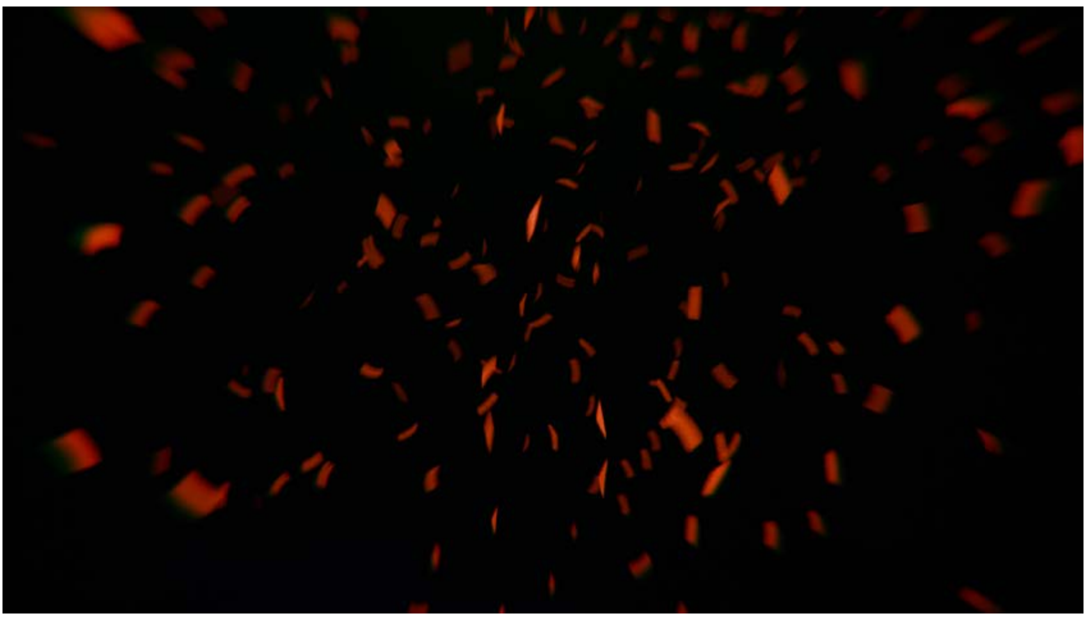
Still frames captured from the rendered animation follow on the next page.

References

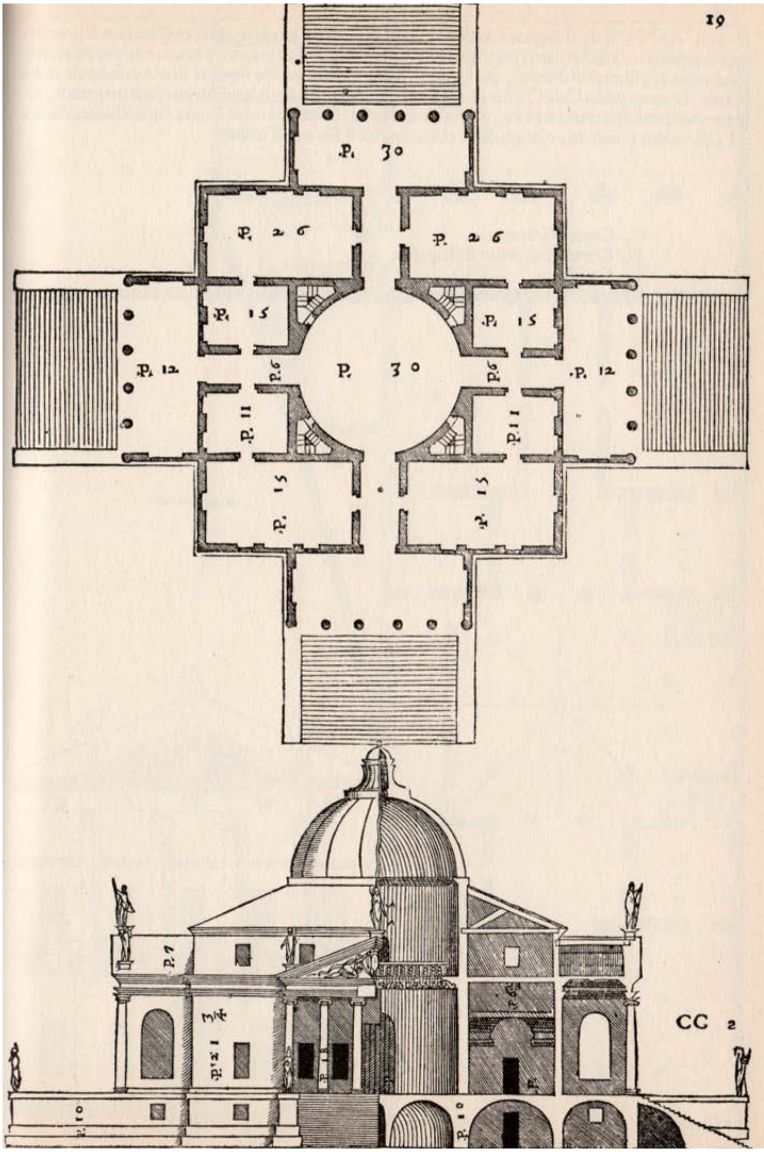
[1] Reynolds, Craig (1987). "Flocks, herds and schools: A distributed behavioral model". Proceedings of the 14th annual conference on Computer graphics and interactive techniques. Association for Computing Machinery. pp. 25–34



Still frame from The Alphabetic Labyrinth



Still frame from The Alphabetic Labyrinth



Andrea Palladio, 1570

LIVE PERFORMANCES

ViridissimaRigore: Evergreen Struggles between the Organic and the Mechanical

Jeffrey M. Morris, DMA, and Rebecca Hays, DMA

morris@tamu.edu
morrismusic.org

rebecca.hays@tamu.edu

College of Performance, Visualization & Fine Arts
Texas A&M University, College Station, Texas, USA

Abstract

To honor this year's conference theme, which highlights generative art as an opportunity to leverage the past toward the future in pursuit of beauty, this composition for soprano with fixed audio media applies machine learning techniques to explore the duality between the organic and the mechanistic by juxtaposing two staples of music literature for voice: Hildegard von Bingen's chant "O viridissima virga" (12th century; in Latin) and the *aria antica* "Caro mio ben" (Giordano, 1780; in Italian) as a portrait of the human voice in creative computing. These historical works present opposing perspectives on ephemeral growth, beauty, bounty, and their elusiveness. Hildegard's sacred chant presents a vivid sensory richness "most verdant," a lush metaphor for natural growth and vitality, in praise of the Virgin Mary. "Caro mio ben," by contrast, portrays secular longing after an elusive mortal lover. Bringing these works together and framing both as "hymns to the muse" reflects on the perpetual struggle between humanity and its tools: the organic and the rigid, the vitality of nature (*viridissima virga*) and the "cruel" (*crudel*) precision of technology, juxtaposing voices from the 12th and 18th centuries in a 21st-century composition.

By exploring a latent space populated by machine learning processes taking in both inspirational sources, our new tool sand collaborator guides us in remixing the original phrases to reflect on technology's nature as a double-edged tool—it teases us with the possibility of surpassing human capabilities, yet in every new iteration, it continually disappoints us in new ways. The rigor of machines cannot fully capture the lushness of organic experience, yet it relentlessly seeks to quantify and perfect what remains elusive.

This frustration, however, is not a defeat. Rather, it is an opportunity for reflection

and growth. New kinds of shortcomings also carry enlightenment, allowing us to ponder our own values and desires. As technology falls short of replacing the human element, it presents new mirrors through which we can see ourselves, expanding our understanding of our own limitations and desires. The music invites listeners to embrace the tension between their own emotional yearnings and the calculated precision of machines. This duality is framed not as a conflict to be resolved, but as an evergreen struggle—timeless, ongoing, and essential for creative growth.

The juxtaposition of the phrases *sospira ognor* ("always sighing") and *tanto rigor* ("such strictness") serves as the foundation for this exploration. *Sospira ognor* encapsulates the organic, human experience—emotional, breath-filled, and longing; *tanto rigor* represents the rigid, unyielding perspective of machines, disciplined yet ultimately unsatisfying. Between these two poles, the Latin text of "O viridissima virga" introduces ephemeral connective material portraying the flow of the creative voice. This interplay between the human and the mechanical underscores the perpetually attractive but permanently incomplete nature of human endeavors: humans create tools to aid in pursuing their goals, but these tools also introduce stumbling blocks.

Inspired by Gödel's incompleteness theorems (1931/1992), the work suggests that no system—neither the organic nor the mechanistic—can fully encompass that which it seeks to express. Mechanical rigor will never perfectly capture the living richness of the human voice, just as human emotion can never fully articulate its desires. Instead, we are left in a generative space in which frustration and desires tumble forth together, becoming a driving force for innovation and creative expression.

By blending the medieval and (neo)classical with the modern and the sacredly, secularly, and computationally inspired, *ViridissimaRigore* serves as a meditation on the role of technology in human evolution and the role of the human voice in technology. It reflects the paradox of striving for precision and perfection, continually confronted by rich yet untamable complexity. It invites audiences to embrace the frustration inherent in creative computing, recognizing that it is not through replacement, but through reflection, that humanity grows.

References

Giordani, Giuseppe. *Caro mio ben*. 1780. In *Arie antiche: 30 songs for medium voice*. Edited by Alessandro Parisotti. New York: G. Schirmer, 1950.

Gödel, Kurt. *On formally undecidable propositions of* Principia Mathematica *and related systems*. 1931. Translated by B. Meltzer. New York: Dover Publications, 1992.

Hildegard von Bingen. *O viridissima virga*. 12th century. In *Symphonia armoniae celestium revelationum*. Transcribed and edited by Barbara Newman. Ithaca, NY: Cornell University Press, 1988.

Generative Visual Poetry: Drawing the Land's Memory Through Place-Based Poems

Prof. Andrea Wollensak
Department of Art, Ammerman Center, Connecticut College, New London, CT, USA
www.andreawollensak.com e-mail: ajwol@conncoll.edu
Katja Novak
Poet, Translator, Tallinn, Estonia
Brett Terry
Independent Scholar/Composer, Sonalysts, Waterford, CT, USA

Abstract

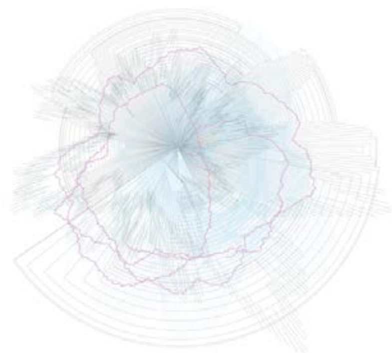
"Generative Visual Poetry: Drawing the Land's Memory Through Place-Based Poems" merges poetry, visual art, and technology through a collaboration between a visual artist, an audio/computer engineer, and a poet. Originating from a Fulbright research project in Estonia, where the visual artist and Ukrainian poet met, this collaborative work transforms the poet's recitations into real-time visualizations using Processing software. The project captures auditory elements—volume, intonation, pauses—and translates them into dynamic visual forms, exploring intersections of place, writing, and visual language. The result is an immersive audio-reactive video and printed visual notation, revealing hidden meanings and enhancing poetic expression.

Keywords:

Generative, digital poetics, audio-reactive software, place-based, land's memory, visual notation

Background

"Generative Visual Poetry: Drawing the Land's Memory Through Place-Based Poems" is the most recent artistic collaboration between the authors. Past collaborations include working with Alaska-based poets Jen Stever and Erin Coughlin Hollowell in *Water Stories": Visual Poetics and Collective Voices* (2022), involving Connecticut students and the local community in *"Reading the Wrack Lines"* (2020), and the work of poet and professor Judith Goldman in "Open Waters" (2019). These previous works were featured at Generative Art conferences (and other venues) and featured generative and audio-reactive visualization systems responding to the poet's voice.



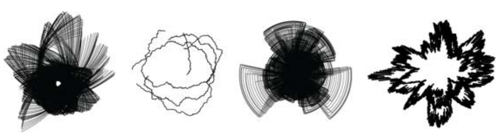
"The Clock" Detail of audio reactive generative system for live poetry reading and print.

Structure and Process

This project employs Processing software to transform the poet's spoken words into real-time visual notations. As the poet reads, her voice generates various dynamic graphical forms. In addition to the sound of her words, her pauses, changes in volume, and modes of intonation together produce a visual manifestation of her poetry that reveals hidden meanings embedded in the text.

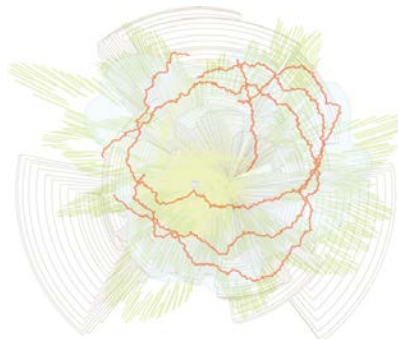
Inspired by topographic map notation, we developed four radially oriented visual notation elements: overlapping triangular forms, straight lines extending from the center, irregular circular lines, and arc line forms.

Within the generative environment, randomly selected poetic text phrases appear and then dissolve in conjunction with the visual notation elements.



"Generative Visual Poetry" delves into visual poetics by integrating auditory and visual elements, emphasizing the shared boundaries between graphic design, writing, and performance. The visual notation provides a tangible representation of the poet's work, considering alliteration, spatiality, and the intricate details of her language.

Additionally, the visual notation is explored in large-format digital print form, providing a longer view of the notation and its delicate line, color, and transparency drawing from the poet's source material. The multilingual text poems are presented alongside their graphical interpretations. This dual format underscores the project's exploration of how language can be simultaneously spoken and seen, bridging traditional and digital forms of artistic expression.



"Poem in Prose, Pear Tree" Detail of audio reactive generative system for live poetry reading and print.

Land's Memory: Poetic Source Material

The source material originates from the author's personal experience, from spectating the life around both as a participant and as a witness. Having relocated several times, the author has several places she calls home. But the topic of land's memory does not end with a physical space called home: it continues in physical objects associated with home, often given from generation to generation; in fruit trees which carry on the memory of family members who planted and cared for them, in the land of home in a wider setting and, of course in people who feel like home, whether they accompany us in real life or memories.

Summary

Through this innovative approach, this new work redefines the boundaries of poetic representation and invites audiences to experience poetry in a new and immersive way.

ANGELIA: An Emotional Generative Algorithmic Intelligence for Contemporary Electronic Music

Prof. J.-C. Heudin
Artificial-Creature
https://www.artificial-creature.com
e-mail: jcheudin@artificial-creature.com

Abstract

The artistic performance takes the form of a concert accompanied by animated images and videos. Angelia is the contraction of « Angel » and « IA », the french acronym for Artificial Intelligence. The aim of the project is enhances the creativity of the artist for composing, and augments his capabilities when performing. The music is performed live as a Duo, Angelia and Jean-Claude Heudin, both playing together on a synthesizer setup. The performance emphasizes the symbiotic relationship between the artist and the AI. Angelia's electro-symphonic music is based on the concept of hyperorchestration, a mode of composition and orchestration that adds digital and analog synthesizers to the regular instruments of an orchestra [1].

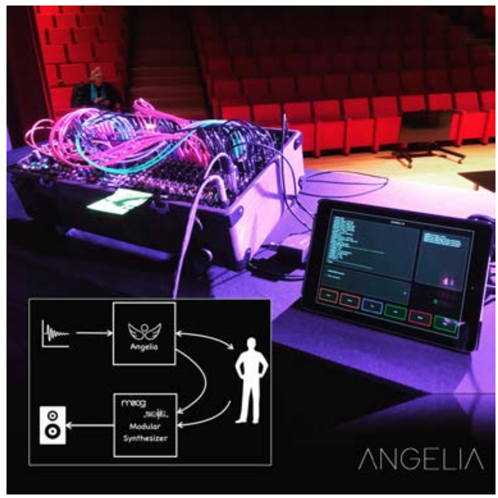


Fig. 1. Preparation of setup for a performance at CentraleSupelec showing a modular synthesiser connected to Angelia.

In contrast with “prompts”, the music is precisely composed using a dedicated high-level programming language which enables to choose for each instruction among different algorithms:

- Melodic and harmonic instructions,
- Procedural and stochastic generators,
- Corpus-based genetic Algorithm,
- Cellular Automata,
- Quantum Notes and Cellular Automata,
- Fractal development,
- Unsupervised Neural Networks,
- Markov Chains.

As an example, Angelia includes a dedicated Corpus-based Genetic

Algorithm. One of the first Evolutionary Algorithms applied to music was GenJam [2]. Like GenJam, our algorithm is inspired by natural selection among a population of individuals, the process that drives biological evolution. However, our implementation does not start from a random generated population, but is initialized using a corpus database. Unlike Deep learning, Angelia does not use large volume of data from uncited composers. The database includes a carefully curated corpus of score patterns of famous composers, all in the public domain [3].

In most AI music projects, the system generates music with no direct feedback from the produced sounds in the environment. In parallel with the generation of music, Angelia can analyze the perceived sound environment in order to generate stimuli that update an “emotional metabolism”. The resulting emotional state influences parameters that modify the expressiveness of the interpretation. This emotional metabolism is inspired by our previous works on emotional virtual characters [4, 5].

The Angelia research project is developed with a strong ethical approach. First, it is designed to be played by an artist, not to replace him. The project follows copyright laws and regulations: machine learning is used but fairly trained. Angelia is also environmentally friendly: it runs on a simple tablet with minimal energy consumption and does not require any distant computing or data center.

Angelia is not just another AI project applied to music, but above all it is an artistic project for contemporary and experimental music using Generative Algorithmic Intelligence. In this framework, we have released albums that retrace its artistic evolution. They can be freely listened on an independent and open music platform[1].

References

1. Casanelles, S., The Hyperorchestra: A Study of a Virtual Musical Ensemble in Film Music that Transcends Reality, Ph.D. diss., Steinhardt School of Culture, New York University, 2015.

2. Biles, J.A., GenJam: A Genetic Algorithm for Generating Jazz Solos, in Proceedings of International Computer Music Conference, Aarhus, Michigan Publishing, 1994, 131–137.

3. Heudin, J.-C., Musique, intelligence artificielle et données, in Musique et données, CNMlab, Centre National de la Musique, Paris, 2023. 92-103.

4. Heudin, J.-C., A Bio-inspired Emotion Engine in the Living Mona Lisa, in Proceedings of the Virtual Reality International Conference, Laval, 2015, 1–4.

5. Heudin, J.-C., An Emotional Multi-personality Architecture for Intelligent Conversational Agents, in Transactions on Computational Intelligence XXVIII, Springer, 2018, 1–21.

[1] https://angelia.bandcamp.com/ and also on https://youtube.com/@jcheudin

GESTOMAT - an interactive instrument for the precomposition of audiovisual works

prof. dr hab. Monika Karwaszewska
Stanisław Moniuszko Academy of Music in Gdańsk
e-mail: mon-kar@wp.pl

dr Beata Oryl
Stanisław Moniuszko Academy of Music in Gdańsk
e-mail: webe.o@interia.pl

Michał Garnowski, MA
Poland, Academy of Fine Arts in Gdańsk
e-mail: garnnoffski@gmail.com

Marcin Garnowski
Poland, Gdańsk University of Technology




Abstract

"Gestomat" is an interdisciplinary research project created as a result of cooperation between music theoretician Monika Karwaszewska, rhythmist and performer Beata Oryl and Vj-assembler Michal Gary Garnowski, as well as computer scientist Marcin Garnowski, who has been invited to join the project as a guest.

The essence of the project is the integration of technology, through the creation of a device that responds to the gestures of a performer trying to shape a kind of artistic creation from minimal movements. In another context, it can act as an educational tool, helping

participants develop non-verbal communication skills, increasing their awareness of the importance of gestures in everyday life.

In the first installment, “Gestomat” appears as a simple system for managing and monitoring the processes involved in recording and interpreting gestures. Using Arduino microcontrollers, Marcin Garnowski creates a glove for collecting data from the position of fingers to identify hand movements, thereby allowing interaction with a computer that analyzes and processes the collected information by generating, through coding, the appropriate image and sound. One sound of the five-note scale is assigned for each finger of the hand, which has its own unique sound and can introduce an interesting coloring to the “gestural” improvisation. The instrumentalist and visual artist design the sounds and images assigned to specific readings of the 'Gestomat', thus creating an open-ended audio-visual composition with which to improvise, but which only comes to life when someone controls it; setting it in motion.

The goal of the project will be to present an interactive artwork that will be created through the fusion of artificial intelligence and hand movements. The fusion effect will directly affect the process of generating sound material and visual layer in the creation process. The gestures of the performer, in which expression and emotions are hidden, will become crucial in the process of organizing the musical material of the work.

The final result of the project is a self-contained device, generating image and sound with hand gestures, which reverses the role of the performer, who instead of reacting with his body to sound or image, generates them with his movement. The resulting creation has the structure of a dialogic form, in which the project's interdependent creators give the work the quality of a multimedia, coherent narrative, and the effects of the recorded gestures fill the space in which they are located. The use of 'Gestomat' in the pre-composition process will pave the way for innovative and inclusive musical and visual expression.

Keywords: audiovisual performance; artificial intelligence; gestures; new technology; precomposition; generative art

Symmetriis, a software to perform live visual symmetries in generative art performances

Enrico "Catodo" Zimuel, BSc, PGDip
Computational artist, Turin, Italy
web: www.catodo.net, e-mail: iamcatodo@gmail.com

Abstract

Symmetry, a fundamental principle observed in nature, plays a pivotal role in generative art, offering both aesthetic appeal and structural coherence. We proposed an open source software, called *Symmetriis*, to perform live visual symmetries using a set of images. With this software artists can explore the significance of symmetries, drawing parallels between natural patterns and algorithmic creativity.

From the bilateral symmetry of organisms to the fractal symmetries found in plants and landscapes, nature provides a blueprint that informs and enhances generative processes [1]. In the context of generative art, symmetry can be leveraged to create visual harmony, balance, and intrigue through recursive algorithms, cellular automata, and fractal geometries.

Symmetriis can perform rotation, reflectional, and translational symmetry using a generative approach with random samples of images and usage of opacity layers, demonstrating how these symmetries engage audiences by evoking the familiar patterns seen in natural systems [2]. Drawing from sources in both art and science, as suggested in [3] and [4] generative art can play a central role in the art of symmetry.

1. Symmetriis

Symmetriis is a software for live performance. It can be downloaded from https://github.com/iamcatodo/Symmetriis.

With this software visual artists can generate, in real time, a video using a set of images with the following actions:

- flipping;
- rotation;
- translation;
- zoom;
- opacity;

For improving the live performance experience, the software supports the usage of a MIDI USB controller, e.g. AKAI LPD8 (Figure 1).

Figure 1. AKAI LPD8, a MIDI controller connected to Symmetriis

Using a MIDI controller the performer can manipulate in real time all the effects, including the frame rate (up to 60 FPS, depending on speed of the CPU used).

Figure 2. Example of frame

In Figure 2, we reported an example of a frame generated using the software. It contains two images that are flipped horizontally with an opacity of 65%.

2. The performance

We want to propose a live performance to show the potential of the Symmetriis software. The performance will use a set of about 3000 images taken from the collection NY Doors of Colmer Roy, from the New York Public Library Digital Collections [5].

The performance will contains audio and video. The requirements for the performance are: Full HD projector, HDMI cable, stereo cable ¼ jack, sound system with enough power for the location.

3. References

[1] Ian Stewart, *Why Beauty is Truth: A History of Symmetry*, Basic Books, 2007

[2] E.H. Gombrich, *The sense of order. A study in the psychology of decorative art*, Cornell University Press, 1979

[3] Hermann Weyl, *Symmetry*, Princeton University Press, 2016

[4] A.V. Shubnikov, V.A. Koptsik, *Symmetry in Science and Art*, Plenum Press, 1974

[5] Colmer Roy, *NY Doors*, New York Public Library Digital Collections, https://digitalcollections.nypl.org/collections/doors-nyc#/?tab=navigation

Solo Sale Siamo… JustDust

Prof. F.N. Bellicieri, PhD
Communication and Languages Center
www.fernandabellicieri.com
e-mail: fernanda.bellicieri@mackenzie.br

Communication and Languages Center , Mackenzie Presbyterian University
San Paolo, Brazil

Abstract

"Solo sale siamo…JustDust" is a

multiplatform performance based on the concept of File Theater (Bellicieri, Azevedo): use of speeches of different codicial natures (verbal, visual and sound) and authorships, inserted as variables of a creative-collaborative algorithm whose result is a scene in constant elaboration and reinterpretation process. A dialogical and transitional scenic experience between different languages and subjects of expression; both mediated by the software and elevating it to the status of authorship, considering the AI context.

In this performance proposal based on the logic of File Theater, the creative-discursive intentions of different authors, expressed in different languages (verbal, visual, sound) and multiple means, are given new meaning based on the modular and encyclopedic logic of the digital environment, which equalizes and expands them, enhancing their transferability, and their meanings.

The performance "Solo Sale Siamo…JustDust" has "War" as its subtext, reflected in the following discursive layers: in the poetry "La bambina e il bombolone", in the form of a verbal code, written by Fernanda Bellicieri (1); in the assumption of the polyphony of the characters present in the poetry, in the form of interpretation recorded on video, by the same author (2); in a series of illustrations based on the theme, created by a second author, Maria Lúcia Nardy, based on her contact with poetry (3); aestheticised movement incarnated in the choreographic performance "Lamento", based on the poetry "La bambina e il bombolone" and on the illustrations by ML Nardy (4).

These different discursive layers, as audiovisual files, once transcribed to the digital platform, and submitted to AI software, constitute variables in the scenic composition algorithm, being able to generate different audiovisual versions of the same narrative, which then becomes expanded and potentially infinite, depending on the combinatorial possibilities of its elements. A narrative whose central theme, "war", becomes polyhedral and multidimensional, either through the intervention of the algorithm-variable binomial, or through the author's own reflection on his initial work, or already altered by the logic of the file. In the context of shared and in-process

authorship, AI operates not only as a mediator, but as an author in reinterpreting the creative intentions of other authors; in this case Fernanda Bellicieri - text and interpretation, and Maria Lucia Nardy, illustration.

The proposal for GA 2024 is to bring another layer to “Solo Sale Siamo”, through the presentation of a live performance, in the form of a new poem interpretation, with the projection of one version of “Solo Sale Siamo” generated by AI as a backdrop. A version that merges the previously outlined layers but is integrated with the author's embodied performance, now transformed in its initial creative intentionality. An authorship that, despite being embodied by the performer is mediated and also written by the software.

This spiraling path of File Theater operates not only in reflecting on languages, but also affects the artist's perspective on his own intentionality, composing a multifaceted mosaic of motivations and discourses capable of expanding his creative universe.

The Cottonwood Florilegium

Dr. Sam Wells, DMA
Boyer College of Music and Dance, Temple University, Philadelphia, United States
https://sllewm.as
e-mail: sam.wells@temple.edu

Abstract

This performance will feature Part A of *The Cottonwood Florilegium* and involve a subset of the original ensemble, focusing on the work's opening movements. This excerpt will be roughly 20 minutes in duration and feature trumpet, drums, and generative, interactive electroacoustic audio.

Part A, *The Winter Circle,* presents a communal, collaborative accumulation of musical material, setting the tone for the entire piece. *The Cottonwood Florilegium* as a whole explores the idea that existence is a text waiting to be read, where both the performer and the listener are part of an ongoing narrative. This performance highlights the work's exploration of interconnectedness and the human desire to understand existence through new perspectives—spiritual, scientific, and personal.

Inspired by the tradition of the medieval florilegium, the piece gathers perspectives on human experience, from the local and internal to the cosmic and infinite. It is structured into five parts, each named after celestial or natural phenomena, and combines mythological, natural, and human elements to explore the collective experience of existence. A key feature of the work is the use of generative algorithms, which allow for fluid, evolving interactions between the musicians and electronic components. These algorithms generate musical material that is partially unpredictable, creating an ever-shifting soundscape that parallels the unpredictability of human experience and the natural world. The musicians engage with these generative processes, responding in real-time to the algorithmically generated material, blending structure with improvisation.

In Part A, *The Winter Circle*, the focus is on the cosmic and mythological. The Winter Circle asterism, visible in the Northern Hemisphere's winter sky, consists of major stars from the Orion, Taurus, Auriga, Gemini, Canis Minor, and Canis Major constellations. This star formation serves as a unifying symbol, linking ancient myths and cultural

understandings of the cosmos. In Lakota mythology, the Winter Circle mirrors the *Čhaŋgléska Wakȟaŋ* (Sacred Hoop), a symbol representing life stages, deities, and the creation myth of the Great Race. The first movement of Part A, "Accretion," reflects on the cosmic and geologic process of accumulating matter into massive forms. Musically, the movement builds layers of sound, representing the communal gathering of experiences, relationships, and energies that shape both our individual and collective histories.

Generative algorithms play a key role in shaping the dynamic interactions between musicians, as well as between musicians and electronics, resulting in a sound world that is continually evolving. The algorithms introduce elements of unpredictability and chance, allowing each performance to be unique while maintaining an overall structure. This real-time interaction between human performers and generative systems reflects the balance between order and chaos, echoing the unpredictable nature of existence itself.

Though this performance will only present Part A, the excerpt offers a microcosm of the larger work, inviting listeners to engage with the themes of interconnectedness, accumulation, and the search for meaning through both structured composition and generative processes. The musicians and algorithms together create a sonic environment that mirrors the unfolding and accumulation of experience, drawing parallels between human history and the cosmic processes that shape our world.

1. Program Note

"The entirety of existence was a text waiting to be read. Which means there could be no line between the reader and the written. You, who are reading this, you too are written, you too can be read. And I, a writer, am already written through and through. Everything between us, everything that separates us, mountains, stars, years, shimmering thoughts and dreams that die with waking, all of it is a single chain of signs that do not point to another reality, only to this one, all at once." [1]

"The Cottonwood Florilegium" is a musical text that is read as it is written and written as it is read. It is a collection of a perspectives and ideas, musically interpreted, on the entirety of existence from the internal and local to the cosmic and infinite.

Humanity has developed a myriad of approaches to gaining perspective on this reality. We strive to step aside and look at reality not from the active locus of experience, but at a remove, with a spiritual or scientific objectivity. We imagine being able to read the text of existence from a different angle, seeing previously invisible connections. Through this imagination, we inscribe the lacunae of our perceptions to create meaning and expand the context of our experience.

A florilegium is a medieval collection of excerpts, proverbs, ideas, and formulas that grew of the commonplacing tradition of communal writings to collect and index knowledge. Florilegia have traditionally presented in several ways: as patristic anthologies of Christian literature, as literary anthologies of secular texts, and as the literal translation of “a gathering of flowers” indicates, a collection of botanically accurate plant illustrations. Here, the florilegium is a comingled set of thoughts, perspectives, and questions on the collective experience of humanity gathered from a range of sources and experiences.

“The Cottonwood Florilegium” is organized into twelve movements that compose five parts. The part names are associated with constellations, asterisms, and weather phenomena that have been ascribed spiritual meaning. Humanity often looks outward to forces beyond our control, be it the environment or the cosmos at large, to gain insight to our experience. The movement names focus on the local, perceptual, and active states of presence and personal experience. All these modes of understanding are inscribed into the text that is “The Cottonwood Florilegium,” this terrestrial bound collection of attempts at making sense of humanity. We can see our pasts and futures in the terrestrial world around us, unveiled through geologic processes and expressed through active environmental changes. The cottonwood tree, widely present in the North American landscape, appears in numerous myths of indigenous peoples to the continent. Having grown up around these trees, the cottonwood presents a personally important signifier of place and location to me. The cottonwood tree can be a marker for home, a reminder that despite our attempts step outside of our experience to understand our place in the cosmos, we are of and a part of this planet.

The Cottonwood Florilegium is presented as a massive and collective life cycle infused with layers of nested narrative arcs.

1.1 Part A

The Winter Circle is an asterism that is visible in the Northern Hemisphere winter sky. Comprised of the major stars in the Orion, Taurus, Auriga, Gemini, Canis Minor, and Canis Major constellations, the Winter Circle physically links these major myths while also holding mythological significance in its own right for several cultures. In Lakota mythology the winter circle mirrors the Čhaŋgléska Wakȟaŋ, or Sacred Hoop, a religious symbol representing life stages, deities, and the location of the creation myth of the Great Race. Movement I, “accretion,” relates not only to the geologic and astrophysical accumulation of matter into massive objects, but also the gathering of experiences and relationships that define our history. Part A and Movement I represent a world building process, a communal accrual of musical matter and potential energy that sets the trajectory for the rest of work.

1.2 Part B

The Tramontana is the classical name of a northern wind in the Mediterranean. In Catalan culture, the tramontana is so strong that it takes on a supernatural quality that causes people to act strangely or lose their grip with reality. Movements II and III, present a dissolution of the joint accrual of the previous section and a disruption of the collective, with individual musicians exerting more agency and improvisatory freedom. Protention is the

Husserlian phenomenological anticipation of the next moment. The disruption causes us to question what is next.

1.3 Part C

Vela is a constellation in the southern sky that depicts the sails of the ship Argo from Greek Mythology. Representing a journey and evoking the open-endedness and transformational potential of ocean voyages. Movements IV, V, and VI offer the most open sections of the work. The musicians are freely improvising with each other and the electronic components to represent a journey of leaving, becoming, and returning.

1.4 Part D

Horologium is a constellation in the southern sky depicting a pendulum clock. This constellation was cataloged relatively recently by French astronomer Nicolas Louis de Lacaille in 1756. Here, Horologium represent an application of logic and parsing of understanding onto the memories of experience, an attempt to make sense of what has just transpired. Movement VII, “retending,” presents purely electronic sound creating an introspective moment that grabbles with personal memory and experience. Retention is the conceptual pair to protention. Retention is the Husserlian concept of retaining perceptual acts consciousness. Movement VIII, “mapping,” represents attempts at organizing and understanding those experiences. Movement IX, “telling,” shares these understandings.

1.5 Part E

Noctua is a former constellation that depicts an owl perched on that tail of Hydra, the water snake. The owl, a pervasive symbol throughout many mythologies, is often depicted as omens of night, death, or wisdom. Part E is divided into three movements: waning, eroding, and unearthing. Each presents an erasure, creating negative space, that aims not to represent destruction but rather a revealing.

The Cottonwood Florilegium was composed for Dana Jessen (bassoon), Vicki Ray (piano), and SPLICE Ensemble: Keith Kirchoff (piano), Adam Vidiksis (drum set), and Sam Wells (trumpet). This project would not have been possible without the collaboration of these musicians nor without the efforts, visions, and dedication of Alejandro Melendez (lighting consultant), Stephanie Lutz (lighting consultant), and Josh Sobel (conceptual consultant). I am deeply inspired by and creatively indebted to my collaborators. I hope that this work feels as much theirs as it does mine.

The entire work is performed without pause.

2. Technical Requirements

The presentation of *The Cottonwood Florilegium* will require a stereo sound system, and a performance area of 3 meters by 3 meters.

The submitter can provide all other technical requirements and instruments.

3. Recording

A video recording of a previous performance of the complete version of *The Cottonwood Florilegium* is available below. Please note that this proposal is for an excerpt of the complete work and will feature a subset of the ensemble.

https://vimeo.com/sllewmas/tcf

4. References

[1] Ben Ehrenreich, "Desert Notebooks: A Road Map for the End of Time"