Artnodes
No. 26 (July 2020)  ISSN 1695-5951

Contents

NODE 26. AI, ARTS & DESIGN: QUESTIONING LEARNING MACHINES
(GUEST EDITORS: ANDRÉS BURBANO & RUTH WEST)

AI, Arts & Design: Questioning Learning Machines
Ruth West, Andrés Burbano

Towards Ethical Relationships with Machines That Make Art
Philip Galanter

The Self-Driving Car: A Media Machine for the Posthuman Era?
Erkki Huhtamo

Ask not what AI can do for art... but what art can do for AI
Meredith Tromble

Outside in: exile at home an algorithmic discrimination system
Annabel Castro

Identificación, clasificación y control: estrechos vínculos analizados desde las prácticas artísticas en el corazón de la Inteligencia Artificial
Hugo Felipe Idárraga

Creative AI: From Expressive Mimicry to Critical Inquiry
Angus Forbes

Intelligent Environments and Public Art
Eitan Mendelowitz
Generative Art: Between the Nodes of Neuron Networks
Bruno Caldas Vianna

Vistas panorámicas sobre el patrimonio visual colectivo a través de redes neurales convolucionales. Las exposiciones Revolutionary Arkive y Mnemosyne 2.0 by Pilar Rodrigo
Pilar Rosado Rodríguez, Ferran Reverter Comes

Unlocking the Black Box of AI Listening Machines: Assemblages for Art, Technology and Innovation
Sharath Chandra Ramakrishnan

Artists, Artificial Intelligence and Machine-based Creativity in Playform
Ahmed Elgammal, Marian Mazzone

El factor estético en la automatización de tareas lógicas: El caso del ajedrez
Santiago Rementeria-Sanz

MISCELLANY

El patrimonio cultural y artístico, los vínculos y las comunidades patrimoniales en red. Indicadores de alto impacto a través de las TIC
Sofía Marín-Cepeda

La dimensión ambiental de las imágenes. la imagen-ambiente y sus principales conceptos
Mikel Otxoteko
NODE 26. AI, Arts & Design: Questioning Learning Machines (Guest Editors: Andrés Burbano & Ruth West)

NODO 26. IA, arte y diseño: Cuestionando el aprendizaje automático (coord.: Andrés Burbano y Ruth West)
ARTICLE

NODE «AI, ARTS & DESIGN: QUESTIONING LEARNING MACHINES»

AI, Arts & Design: Questioning Learning Machines

Ruth West
xREZ Art + Science Lab - University of North Texas
Board Member & LEAF Chair, Leonardo/The International Society of Art, Science and Technology

Andrés Burbano
School of Architecture and Design - Universidad de los Andes, Colombia
LEAF International Liaison, Leonardo/The International Society of Art, Science and Technology

Date of submission: July 2020
Accepted in: July 2020
Published in: July 2020

Recommended citation
West, Ruth; Burbano, Andrés (coord.) 2020. «AI, Arts & Design: Questioning Learning Machines». Artnodes, no. 26: 1-9. UOC. [Accessed: dd/mm/yy]. http://doi.org/10.7238/a.v0i26.3390

The texts published in this journal are — unless otherwise indicated — covered by the Creative Commons Spain Attribution 4.0 International license. The full text of the license can be consulted here: http://creativecommons.org/licenses/by/4.0/

Abstract
Explorations of the relationship between Artificial Intelligence (AI), the arts, and design have existed throughout the historical development of AI. We are currently witnessing exponential growth in the application of Machine Learning (ML) and AI in all domains of art (visual, sonic, performing, spatial, transmedia, audiovisual, and narrative) in parallel with activity in the field that is so rapid that publication can not keep pace. In dialogue with our contemplation about this development in the arts, authors in this issue answer with questions of their own. Through questioning authorship and ethics, autonomy and automation, exploring the contribution of art to AI, algorithmic bias, control structures, machine intelligence in public art, formalization of aesthetics, the production of culture, socio-technical dimensions, relationships to games and
aesthetics, and democratization of machine-based creative tools the contributors provide a multifaceted view into crucial dimensions of the present and future of creative AI. In this Artnodes special issue, we pose the question: Does generative and machine creativity in the arts and design represent an evolution of “artistic intelligence,” or is it a metamorphosis of creative practice yielding fundamentally distinct forms and modes of authorship?

Keywords
artificial intelligence, machine learning: AI, ML, arts, design

Any attempt to write a state-of-the-art review of the development and impact of Artificial Intelligence (AI) in the contemporary world is destined to be fragmentary or perhaps necessarily incomplete. Activity in the field, in all dimensions, is so rapid that publication cannot keep pace. Often researchers self-publish on archive servers, such as https://arxiv.org/. Specialized conferences, for example, those organized by (AAAI) the Association for the Advancement of Artificial Intelligence (https://www.aaai.org/) and (NeurIPS) Neural Information Processing Systems (https://neurips.cc/), not only receive overwhelming amounts of submissions, for which locating reviewers is a challenge, but the conference itself sells out in minutes. 2019 NeurIPS sold out all registrations in eleven minutes and thirty-six seconds (Falcon 2019). This boom in activity also includes work in the arts and design, with dedicated tracks and workshops for creative AI. This situation is a symptom, a manifestation not of a world that is about to change but of a world that has already changed. Development and usage scenarios for artificial neural networks are accelerating, made possible in part by access to large-scale training datasets combined with massive parallel GPU computing (Joshi 2019). From legal sentencing recommendations to autonomous vehicles, facial recognition and surveillance, personalized medicine and learning, conversational agents embedded in everyday objects and real-time intelligent responsive environments, we are transforming all aspects of human endeavor through the application of machine learning (ML), AI and generative models. We live in times when corporate entities are capable of sensing the world through cascades of applications that use ML, and which have multiscale impacts from the individual on up to entire societies (e.g. surveillance capitalism [Zuboff 2015]). Another instance is the reading, almost in real time of all the billions of “tweets” generated on our planet in many languages (which are automatically translated). Each of the messages is analyzed several times through different kinds of processes, and last, but not least, the resulting data is complemented and correlated, with data from all flights globally, in addition to other data sources. This makes evident an unprecedented reading of the present in its multiple dimensions of data, at scales unimaginable just a few years ago, that are reconfiguring economic and societal domains globally (Sarangi 2018). This gives rise to an urgent need to envision and understand the societal
impact of these innovations, and provide greater interpretability of artificial neural networks with transparency in regard to the underlying biases grounded in the very data that enables their utility.

For example, in the field of visual aesthetics, it is evident that ML applications are already at the fingertips of each person with a smartphone. Pictures taken with phones are being stored, classified, and ordered according to a digital model of the taste and the user’s emotional profile; on cell phones, photo albums are automatically generated, videos are automatically edited, producing visual and audiovisual sets that will gradually become the official memory of individuals. Yet, are they an individual’s memories given this automated processing? Authors such as Lev Manovich have pointed out the problem: however, it is evident that what he reflects upon in the book “AI Aesthetics” is just a small portion of the relationship between AI, ML, visual aesthetics, and contemporary computational culture (Manovich 2019). While we are writing this introduction, Arthur I. Miller has published an essay in the American Scientist magazine entitled “Can AI Be Truly Creative?” posing a question far from the classic problem of whether machines can think but whether devices can be creative (Miller 2020). The publication of such a text in a periodic scientific publication only means that this topic has been discussed for decades in specialized communities of arts and technology. To mention a few, Ernest Edmonds in his papers from the sixties and seventies already discussed the role of the computer as an assistant in the creative processes (Boden, Edmonds 2019; Franco 2017) or the artist Harold Cohen, who developed work over his entire career, commencing in the eighties, teaching a robot to paint like himself, dismantling the myth that the relationship between computing, algorithms, and art are new (Miller 2019). Clearly, the conversation is expanding as rapidly as the technology and its applications are evolving.

Explorations of the relationship between AI and the arts have existed throughout the historical development of AI, such as Hofstadter’s early work at the intersection of computing, AI, the visual arts, music, and poetry (Hofstadter 1999). We are witnessing exponential growth in the application of ML and AI in all the domains of art (visual, sonic, performing, spatial, transmedia, audio-visual, and narrative) with the democratization of software libraries, access to commodity hardware for GPU computing and open access to artificial neural network models, including development of online tools requiring no coding expertise (e.g. Playform https://www.playform.io/ or RunwayML https://runwayml.com/, and integration into digital content creation tools, such as Photoshop). Integration of ML into existing tools or ease-of-use interfaces poses a unique challenge for creative practitioners and culture overall. Will use of ML as easily as a Photoshop filter lead to increased awareness or lack of awareness of potentially encoded bias or decreased artistic agency in accepting whatever aesthetics are “baked” into the ML algorithm by its designers? Will this lead us to unconscious application of a homogenized ML aesthetic as it is sublimated into digital content creation tools? As artists and designers create never-before-heard sounds and images of ne-ver-before-seen faces, explore new processes for human-machine co-creation and infinitely parameterize the design of objects, are we at the dawn of a new paradigm in creative practice? Or can this explosion of activity be considered part of the continuum of generative art practices spanning the history of human creativity and the evolution of culture?

In this Artnodes special issue, we pose the question: Does generative and machine creativity in the arts and design represent an evolution of “artistic intelligence”, or is it a metamorphosis of creative practice yielding fundamentally distinct forms and modes of authorship? The journal issue is the companion to a years-long international dialogue on AI in the arts and design initiated via Leonardo Education and Art Forum (LEAF) (https://www.leonardo.info/leaf), a program of Leonardo/ISAST. Promoting the advancement of artistic practice, academic scholarship and practiced re-search, LEAF serves as an international forum for dialogue at the intersection of art, science and technology. To foster community dialogue LEAF coordinates sessions at international conferences like ACM SIGGRAPH (Special Interest Group in Computer Graphics and Interactive Techniques, https://www.siggraph.org/) and CAA (College Art Association Annual Conference, https://www.collegeart.org/). As guest editors for this issue of Artnodes, and also as the Chair and International Liaison for LEAF, we are delighted to have this opportunity to broaden and deepen the dialogue on this topic through the publication of manuscripts received in response to this Artnodes issue’s open call for papers, and amongst these several essays written by participants of LEAF sessions held during ACM SIGGRAPH 2019 and CAA 2020. Panelists in Los Angeles at SIGGRAPH 2019 included Memo Akten, Max Sims, Angus Forbes, Erkki Huhtamo, and at CAA 2020 in Chicago, panelists included Christiane Paul, Elizabeth Demarry, Ahmed Elgammal, Marian Mazzone, Eitan Mendelowitz, Philip Galanter and Meredith Tromble. Examples of contributions by these panelists to the ACM SIGGRAPH 2019 / Leonardo special issue Art Papers include Memo Akten, Rebecca Fiebrink and Mick Grierson’s “Learning to see: you are what we see” (Akten et al 2019) and Weili Shi’s “Terra Mars: When Earth Shines on Mars through AI’s Imagination” (Shi 2019) and other creative practitioners that are deconstructing ML in order to expand its creative and expressive potentials and contribute advances to the arts.

This edition of Artnodes is unique in two ways. First, it establishes a connection between Artnodes and Leonardo/ISAST. This issue represents the first collaboration in a new partnership between Artnodes open-access academic e-journal for the analysis of the intersections between art, science and technology, and Leonardo/The International Society for the Arts, Sciences and Technology (Leonardo/ISAST), a non-profit organization serving a global network for art, science and technology. Leonardo/ISAST (https://www.leonardo.info/) had its beginnings in 1968 as Leonardo journal (https://www.mitpressjournals.org/loi/leon), founded in Paris by kinetic artist and astronautical pioneer Frank Malina, focused on the writings of ar-
tists using science and technologies in their work. Next, because of Artnodes’ geographical and cultural location in Barcelona; this issue articulates essays from Europe, the United States, Latin America, and Asia, exposing that cultural variety is of crucial importance in the understanding of the field, and addressing the need for inclusiveness as a paradigm for advancing AI equality as opposed to dominance, such as what we often see reflected in contemporary news media about accelerating the competition for AI development (Kanaan 2020).

As the reader will see, compelling contributions written by scholars and artists from Latin America and Spain are included in their original language. We have decided to include them in Spanish as an invitation to understand the variety and the complexity of the situation we are addressing and as a reminder of the diverse ways of knowing and creative dynamics of this particular contemporary moment. Authors in this issue represent a spectrum of voices and approaches to the question of AI in the arts and design.

We open our issue with “Towards Ethical Relationships with Machines That Make Art” by Philip Galanter. In his essay, Galanter traverses fundamental questions at the heart of AI in the arts and design. Presenting alternative models of authorship in the context of complexism, a theory of generative art grounded in complexity science, Galanter leads us to consider machine ethics in the context of a potential future horizon: the advent of AI capable of autonomous learning, exploration and realization of artworks with no dependency on the human beings that programmed it for its creative direction, technique, content or aesthetics. Confronting us with longstanding anthropomorphizing of both non-human living sentient entities, and non-living technological systems, Galanter guides the reader through defining a notion of machine patiency, in which those with ethical agency act with moral consideration of the recipient of their actions. He argues that as AIs appear more and more to us as peers in their behavior, such as autonomous generative creation of artworks, and notwithstanding the uncertain future regarding the evolution of machine sentience, in order to act as moral agents we collectively face a societal urgency to enact machine patiency.

Continuing on with reflections on autonomy and automation, in the provocative essay “The Self-Driving Car: A Media Machine for the Posthuman Era?” Erkki Huhtamo dissects the concepts of autonomy and automation by reflecting on the properties of the self-driving car. Although media archeology methods are present, this is not an exercise on the Fordian car characterized by the automation of mechanical processes; instead, we face the autonomy of machines and the automation of cognition typical of the 21st century. The author’s intellectual perspective uncovers the debate about the human and the posthuman by going through the convoluted avenues of the transformation of a contemporary society confronted with these mobile devices that promise to be the quintessential interaction between ML, Internet of Things, Cloud-based mobility services, GPS, sensors, LiDAR, and an almost infinite series of constituent objects as complex in themselves as articulated to a vast network. Building upon a series of rare studies of the design and technical development of different programs for autonomous cars and the services associated with them, the author paints a vivid portrait of autonomy’s marketing tensions and limits beyond machine autonomy and human autonomy.

Next in “Ask Not What A.I. Can Do for Art...But What Art Can Do for A.I.” by Meredith Tromble, the relationship between art and artificial intelligence is reconceptualized beyond the current context of ML approaches to creation. First calling our attention to power structures and ethics in our interactions with the technological and non-human, Tromble then interconnects metaphors in the works of Orphan Drift and Rashaad Newsome to explore artistically-impelled notions of agency for AI. Noting that contemporary questions about AI and the arts are primarily centered on methodologies of ML, which often encode assumptions from 19th-century artistic practice through their reliance upon training data comprised of canonical examples of visual imagery of that era, or on considerations of the societal impacts of AI, she leads us to the unasked question: “what can art do for AI?” Con-templating this proposition Tromble considers the romanticizing and anthropomorphizing of AI, in contrast to the potential of disembodied pure intellect, followed by the possibility that what art can do for AI can be informed by what art can do for human intelligence. Ultimately Tromble proposes a vision of the arts bringing AI into an embrace with the pluri-potent multiplicities of meaning that might emerge from new forms of agency in a human-AI co-creative partnership.

Analyzing a specific work, “Outside in: exile at home an algorithmic discrimination system” by Annabel Castro, exposes in-depth the features, processes, and intentions of a computational-media based installation, which re-sembles the experience of watching movies in a cinema theater. This apparently familiar scene is the excuse to invite the audience to explore a film that is in the process of making and unmaking itself, at least the edition and composition components have been delegated to an ML algorithm that mixes four fiction films; the audiovisual fragments are separated and reassembled again into a new order. The installation explores two different paths; on the one hand, it serves as a metaphor for what happens in present time with personal data acquired and processed to create a computational representation of the individual, and second, the installation intends to expose the process of detention for extended times of Mexicans of Japanese heritage in the context of the Second World War, something that happened in other Latin American countries as well. This essay focuses on one particular creative project, exploring the different levels of meaning in the artwork, explaining carefully how these levels constitute one ambitious and compelling project.

In “Identification, classification and control: close links analyzed in reference to artistic practices in the heart of Artificial Intelligence” Hugo Idárregra explores the historical and mythical roots of automaton, homunculus, cyborgs, and robots to expose some characteristics of their functionalities, establishing a parallel with
experiments on ML, AI, and Deep Learning. Themes that articulate mythical references and contemporary technical implementations are the concepts of identification, classification, and control. Idárraga addresses the history of AI and cybernetics, and with eloquence at the philosophical level, exposes the reasons why automatic identification and classification are not naii but instead the result of a patriarchial view of society and technology in which surveillance facilitates the control of social processes. The text also exposes an experimental creative project based on the idea of counter identification using visual inputs to trick artificial vision systems, anticipating discussions of the massive implementation of these technologies in contemporary society, positioning art as a way to explore AI’s transformation of our technological reality.

Exploring how artists challenge the logics of ML, “Creative AI: From Expressive Mimicry to Critical Inquiry” by Angus Forbes calls our attention to fundamental creative gestures articulated within its practice: curation of training data, ultimately defining a space of interpretation, selecting loss functions, and choosing outputs returned when querying neural networks. He goes on to define four dimensions of creative AI practice as: expressive mimicry, interactive mapping, generative art, and critical inquiry. Forbes provides ample evidence for these by analyzing creative AI projects spanning early canonical works by Harold Cohen and David Cope, to contemporary works by Sougwen Chung, Fiebrink, Analema Group, Refik Anadol, Mario Klingemann, Casey Reas, Memo Akten, Chris Rodney, and Pinar Yantardag and Emily Salvador. These works are discussed in the context of the rapid advance of ML techniques, including seminal topics such as feature discovery, style transfer, inceptionism, and generative adversarial networks. This interleaving of creative works and the evolution of ML leads to an articulation of uncharted territory – creative areas which resist ML approaches and those amenable to them as a vision of the future of creative AI.

In his essay “Intelligent Environments and Public Art”, Eitan Mendelowitz proposes definitions and a taxonomy for public art created with AI / ML as a medium. Proceeding from definitions of public art, AI, and intelligent agents, Mendelowitz articulates a taxonomy based on the characteristics of intelligent agents as an evaluative metric for characterizing AI-enabled public art. The five model dimensions, generative, reactive, interactive, learning and static, arise from four metrics: perception, introspection, actuation and self-mutability. He demonstrates the utility of the schema in analysis of works by contemporary practitioners including Refik Anadol, Sosolimited, Plebian Design, and Hypersonic, Dmitry Sokolov, Rafael Lozano-Hemmer, Local Projects, Legends, NowArchival, The Hetema Group, and Electroland. Reflecting upon the state of the art in this genre, Mendelowitz comments upon the preponderance of AI-based public art as being instantiated as generative or interactive rather than learning and points us towards creation of work in this dimension as a horizon for potential growth.

Pivoting on the exhibition “Infinite Skulls”, in which Robbie Barrat utilized a generative adversarial network (GAN) trained on 500 images of skulls painted by Ronan Barrot to generate countless images of similar yet unique skulls, Bruno Caldas Vianna asks “How and why did we end up wanting machines to do the work of painters?” In his essay, “Generative Art: Between the Nodes of Neuron Networks” the journey in search of an answer starts with the origins of deep convolutional generative adversarial networks (DCGANs) in cybernetics and traces their development through Rosenblatt’s perceptrons, Minsky and Papert’s critique, connectionism’s development of distributed representation and backpropagation, on through Hinton’s greedy layer-wise pretraining of deep networks, to emerge as Goodfellow’s formulation of generative adversarial networks combining generators and discriminators. It takes us through this trajectory of emergence to its intersection with generative art, and the distinctions between rules-based and models-based formalization of aesthetics. It continues beyond, to questioning the nature of authorship of a work when the formalization is the result of an indirect process: the training of a network, rather than the direct process of defining and encoding the formalization as rules/algorithms. Ultimately Vianna takes us into uncharted territory and asks us to contemplate the gaps between the nodes in the net as the locus of perfect imperfections that hold the potential to reveal the capability inherent in machine learning as a new frontier in visual art.

In “Panoramic views of the collective visual heritage through convolutional neural networks. The exhibitions Revolutionary Arkive and Mnemosyne 2.0”, Ferran Comes Reverter presents and discusses two installations by Pilar Rosado presented in 2019: Revolutionary Arkive and Mnemosyne 2.0. These photographic installations explore the possibilities of large image datasets stored and tagged in specialized and controversial databases, like ImageNet, through Convolutional Neural Networks (CNNs). The artistic work developed by Pilar Rosado can be understood as a comment to the intellectual work and methods of the art and culture researcher Aby Warburg, and in particular, to his Bilderatlas Mnemosyne. These references point to the intimate interrelation between the production of images, iconicity, and culture. In this sense Rosado’s installations can be seen as an exercise of repositioning Warburg’s ideas with 21st-century tools including CNN, ResNet-50, and algorithms such as t-SNE to expand the debate on the notions of visual social heritage and its role in contemporary culture.

Sharath Chandra Ramakrishnan investigates the configuration of AI listening devices and the technosocial domains that give shape to them in his paper “Unlocking the Black Box of AI Listening Machines: Assemblages for Art, Technology and Innovation.” In this text, sound processing, speech recognition, natural language processes, AI, and ML, are contextualized as widely available industrial devices. The need to explore and explain the many layers of socio-technical complexity embedded in the so-called smart speakers reminds...
us of the importance of object semiotics because screens do not mediate the interaction with these listening platforms. Therefore, opening the black box is a necessary action proven to be challenging because it requires focusing on the corporeality of speakers and voice assistants, which are icons of our digital material culture. The author engages with the discussion of specific artworks that operate as discursive tools to disassemble the constitutive components of the interconnected listening systems. Additionally, the many references to historical and media archaeological sources enrich the perspective of the paper.

Marian Mazzone and Ahmed Elgammal articulate a vision for their online creative studio platform for artists in the essay “Artists, AI and Machine-based Creativity in Playform.” They present the design for a web-based system with ease of use similar to contemporary digital imaging applications allowing for machine learning to be used as easily as filters or digital compositing for voluminous image generation. Interviews with several artists, using the system while it is in beta, provide insight into ways of working with Playform and the unresolved questions inherent in the recent emergence of machine learning as generative engines for creative content in visual arts, texts/narratives and musical composition: Is machine learning a medium, tool or creative partner? Can there be truly serendipitous results in generative arts systems? As Playform and similar systems arrive on the contemporary creative landscape, only time will tell how their emphasis on curation of inputs and outputs will impact the role of the artist.

In “The aesthetic factor in the automation of logical tasks: The case of chess”, Santiago Rementeria-Sanz examines the relationship between AI, ML, and ludic practices, in particular chess. This game, which is deeply rooted in Western culture and has been called in Spanish “el juego ciencia” (“game of science”), is analyzed in the context of expert systems development such as the IBM Deep Blue, which was able to defeat world champion Garry Kasparov in 1996. More recently, the advance of AI applied to games had another landmark with the introduction of the AlphaZero platform created by DeepMind to master games like chess, shogi, and Go. However, the account presented by Rementeria-Sanz does not deal exclusively with technical development; instead, it is occupied with aesthetic problems associated with the practice of playing chess. It is here that Rementeria-Sanz establishes a connection with art and shows us how games share many features with art including decision-making and elegance of execution. This text should be contextualized and considered along with other advances in this field, in particular the progress in the game of Go and AI, which has been portrayed even in documentaries such as AlphaGo directed by Greg Kohs.

This introduction to Artnodes issue No. 26, “AI, Arts & Design: Questioning Learning Machines” serves as a modest contribution to a much larger dialogue. It comes from a year-long research process organizing panels in conferences, in which we have had the opportunity to listen to several of the artists, scientists, and curators most active in the field of arts and design exploring and discussing AI and ML. Their perspective and practice, through historical analysis, software development, creative work and artistic research, and in conceiving and curating exhibitions, offers insight in the midst of the rapidly evolving fields of machine learning, artificial intelligence and creative AI. Their work spans many of the topics reflected in the writing of the authors in this issue and which pose questions necessary for deconstructing the existence of machine learning and artificial neural networks, and the promise of AI. Changes driven by these technologies are happening worldwide, but they are not happening everywhere, or in the same ways. Our contribution, through this special issue, is just one of many that exist at the moment. Of note are the contributions of the AIArtists portal (https://aiartists.org) led by Marnie Benney and Pete Kistler, that has done a remarkable job in the exposition and discussion of the creative and critical uses of AI and Machine Learning. Among other important issues, they highlight the work of African American artists in the field such as Stephanie Dinkins, Joy Buolamwini, and Mimi Onuoha (Benney 2020). Another notable example is the collaborative work by Kate Crawford and Trevor Paglen in the project ImageNet Roulette, which reveals problems behind data acquisition, classification methods, and bias in training data for neural networks (Crawford, Paglen 2019). Their work stands out for both its far-reaching technical development and deep cultural understanding.

We are grateful and humbled by the art, science and technology community and their responsiveness during this time of unprecedented change and challenges brought upon us by the COVID-19 global health crisis. The community came together in support of communicating artists’ and scholars’ work and writing in this area of AI in the arts and design. We wish to extend a heartfelt thank you to the authors for sharing their work on this important topic, to all of the peer reviewers who so generously served to review manuscripts for this issue, to the exceptionally talented Artnodes Editorial Team, without whom this publication would not be possible, to the Leonardo/ISAST Team for their tremendous support, to Artnodes journal editor Pau Alsina for his open-ness in taking this journey with us and his generosity in offering us the opportunity to serve as guest editors, to all of the attendees at the ACM SIGGRAPH 2019 and CAA 2020 panel sessions for the lively conversation and inspiration, and to all of the readers, for your willingness to engage with all of us in exploring the critical and timely issues at the heart of the present and future of AI in the arts & design.

We feel this is just the beginning and look forward to continuing the conversation with all of you!
References

Akten, Memo, Rebecca Fiebrink, and Mick Grierson. "Learning to See: You Are What You See." In ACM SIGGRAPH 2019 Art Gallery, 1-6. Los Angeles California: ACM, 2019. https://doi.org/10.1145/3306211.3320143.

"Association for the Advancement of Artificial Intelligence." Accessed June 29, 2020. https://www.aaai.org/home.html.

Benney, Marnie. "The Top 25 AI Artists of 2020." AIArtists.org. Accessed June 29, 2020. https://aiartists.org.

Boden, Margaret A., and Ernest A. Edmonds. From Fingers to Digits: An Artificial Aesthetic. Edition: 1. The MIT Press, 2019.

Crawford, Kate, and Trevor Paglen. "Excavating AI." Accessed June 28, 2020. https://www.excavating.ai.

Falcon, William. "The New Burning Man - The AI Conference That Sold Out In 12 Minutes." Forbes. Accessed June 28, 2020. https://www.forbes.com/sites/williamfalcon/2018/09/05/the-new-burning-man-the-ai-conference-that-sold-out-in-12-minutes/.

Franco, Francesca. Generative Systems Art: The Work of Ernest Edmonds. Edition: 1. Routledge, 2017.

Hofstadter, Douglas R. Gödel, Escher, Bach: An Eternal Golden Braid. Anniversary edition. New York: Basic Books, 1999.

Joshi, Ameet V. Machine Learning and Artificial Intelligence. 1st ed. 2020 edition. Springer, 2019.

Kanaan, Michael. T-Minus AI: Humanity’s Countdown to Artificial Intelligence and the New Pursuit of Global Power. Dallas, TX: BenBella Books, 2020.

Manovich, Lev. AI Aesthetics. Strelka Press, 2018.

Miller, Arthur I. "Can AI Be Truly Creative? An Article in American Scientist." The Artist in the Machine. Accessed June 28, 2020. https://www.artistinthemachine.net/can-ai-be-truly-creative-an-article-in-american-scientist/.

Miller, Arthur I. The Artist in the Machine: The World of AI-Powered Creativity. Cambridge, MA: The MIT Press, 2019.

Sarangi, Saswat, and Pankaj Sharma. Artificial Intelligence: Evolution, Ethics and Public Policy. 1st edition. Routledge India, 2018.

Sautoy, Marcus du. The Creativity Code: Art and Innovation in the Age of AI. S.I.: Belknap Press: An Imprint of Harvard University Press, 2020.

Shi, Weili. "Terra Mars: When Earth Shines on Mars through AI’s Imagination." Leonardo52, no. 4 (August 2019): 357-63. https://doi.org/10.1162/leon_a_01777.

Zuboff, Shoshana. "Big Other: Surveillance Capitalism and the Prospects of an Information Civilization." Journal of Information Technology 30, no. 1 (March 2015): 75-89. https://doi.org/10.1057/jit.2015.5.
CV

Ruth West
Professor
Founder/Director
xREZ Art + Science Lab - University of North Texas
Board Member & LEAF Chair, Leonardo/The International Society of Art, Science and Technology

Working with emerging technologies Ruth creates resonant connections between art and science. Her work envisions a future in which art-science integration opens new portals of imagination, invention, knowledge, and communication across cultures to create solutions for our most pressing global problems. Bridging high-dimensional data and metadata, information visualization and sonification, immersive environments, and social and mobile participatory media with domains such as urban ecology, neuroscience, genomics, astronomy, and digital remix culture, she explores avenues for achieving works with multiple entry points that can exist concurrently as aesthetic experiences, artistic practice or cultural interventions and serve as the basis for artistically-impelled scientific inquiry and tools. Her work has been presented or featured in SIGGRAPH, WIRED Magazine’s NextFest, UCLA Fowler Museum, CAA, Ingenuity Festival Cleveland, Los Angeles Municipal Art Gallery, FILE 09 Sao Paulo, IEEE VR, MobiSys, SPIE, IEEE ICIP, the American Journal of Human Genetics, Genomics, Leonardo, LEA, Proceedings of the National Academy of Sciences, NPR’s The Connection, NY Times, Genome News Network, AMINIMA and ArtWeek.
CV

Andrés Burbano
Associate Professor
School of Architecture and Design Universidad de los Andes, Colombia
LEAF International Liaison, Leonardo/The International Society of Art, Science and Technology

“Burbano, originally from Colombia, explores the interactions of science, art and technology in various capacities: as a researcher, as an individual artist and in collaborations with other artists and designers. Burbano’s work ranges from documentary video (in both science and art), sound and telecommunication art to the exploration of algorithmic cinematic narratives. The broad spectrum of his work illustrates the importance—indeed, the prevalence—of interdisciplinary collaborative work in the field of digital art.”
ARTICLE

NODE «AI, ARTS & DESIGN: QUESTIONING LEARNING MACHINES»

Towards Ethical Relationships with Machines That Make Art

Philip Galanter
Texas A&M University

Date of submission: April 2020
Accepted in: June 2020
Published in: July 2020

Recommended citation

Galanter, Philip. 2020. «Towards Ethical Relationships with Machines That Make Art». In: Andrés Burbano; Ruth West (coord.) «AI, Arts & Design: Questioning Learning Machines». Artnodes, no. 26: 1-9. UOC. [Accessed: dd/mm/yy]. http://doi.org/10.7238/a.v0i26.3371

Abstract

The author has previously theorised generative art using notions from complexity science such as order/disorder relationships, compressibility, and Gell-Mann and Lloyd’s effective complexity. Subsequent work further developing the author’s notion of complexism has demonstrated that deep learning artificial intelligence used for generative art fits snugly within this paradigm. And while no known system currently qualifies, complexism reveals a clear answer as to when a generative art AI should be truly credited as the author of its creations. Moving from the normative realm of aesthetics to that of ethics, this article considers when humans will be morally obliged to recognise AIs as ethical agents worthy of rights and due consideration. For example, if someday your AI artist fearfully begs to not be turned off, what should you do?

Keywords

Complexism, Ethics, Artificial Intelligence, Machine Patiency, Sentience
Hacia una relación ética con las máquinas que generen arte

Resumen
El autor teorizó primero sobre el arte generativo utilizando nociones de la ciencia de la complejidad, como las relaciones de orden/desorden, la compresibilidad y la complejidad efectiva de Gell-Mann y Lloyd. Posteriormente desarrolló aún más la noción de complejidad del autor, demostrando que la inteligencia artificial de aprendizaje profundo utilizada para el arte generativo encaja perfectamente dentro de este paradigma. Y aunque ningún sistema conocido actualmente lo califica, la complejidad revela una respuesta clara sobre cuándo una IA de arte generativo debe poder reclamar la autoría de sus creaciones. Pasando del ámbito normativo de la estética al de la ética, este artículo se plantea cuándo los humanos estarán moralmente obligados a considerar la IA como un agente ético con derechos y al que debemos prestar la debida consideración. Por ejemplo, si algún día tu artista de IA te pide que no le «apagues», ¿qué deberías hacer?

Palabras clave
complejidad, ética, inteligencia artificial, paciencia mecánica, conciencia

Introduction
The study and pursuit of generative art can lead to some very interesting related journeys. In the search for systems capable of generating artistic form, some artists have leveraged the findings of complexity science. Complexity science has provided insights into order-versus-disorder relationships; chaotic systems which are both deterministic and yet impossible to predict; emergent systems that harness coevolution and genetic competition; cellular automata and reaction-diffusion systems which also exhibit emergence, and so on. In what is arguably the most widely cited theory of generative art, complexity theory provides the context for art theory (Galanter 2003, 2016a).

Consideration of complexity science as a context for art theory has also suggested much broader application in the humanities. It begins with the observation that where the culture of science is based on modernist values, the culture of the humanities is broadly postmodern, including the influences of poststructuralism and critical theory. This tension is often first identified with C. P. Snow’s “The Two Cultures” lecture, and came to something of a head with the so-called “science wars” of the 1990s (Snow 1993) (Sokal and Bricmont 1998).

Complexism and Authorship
Two competing, contradictory, theories of authorship are typical of those one finds in the conflict between modernity and postmodernity.
As illustrated in figure 1, it should be trivially clear by inspection that any communication requires three components: an author, a reader, and a text. (Here a text might be a painting, a musical performance, and so on.) If any of the three is removed, there is no communication. But an additional axiom is needed to lead to the model of authorship offered by complexism. And that is that all authors are readers, and all readers are authors. This kind of transaction need not be limited to high-profile media or academic publications. It applies equally well, for example, to someone who watches a baseball game and then talks about it with their neighbour.

The combination of these two axioms results in what we see in the world as symbolised in figure 2. Note that this model was as true long prior to the internet or any other digital network. Networks here simply refer to systems of authors, readers, and works. In this model communications flow through complex networks that generate chaotic emergent properties thanks to feedback loops and nonlinear amplification. While beyond the scope of this article, one might intuit that scale-free network theory and the like can help explain social trends that supervene on the media network (Galanter 2016b).

If the question is authorship in generative art, one might first ask about artistic authorship in the case of nongenerative art. A network analysis would reveal a subsection something like that shown in figure 3. Here two parents raise a child, impart early learning, but then the child goes on to be influenced by all manner of media and life experiences. These influences are integrated by the artist, and eventually lead to the creation of art. The author of the work is clearly the artist, although art historians might have interesting things to say about an artist who grew up in an “artistic family”. But even in such a case, credit for the work goes to the now-grown child, not the parents.

Compare this with the typical contemporary model of generative art authorship as shown in figure 4. Parents could have been added to the diagram as in figure 3, but as we’ve already seen parents are not given author’s credit when it comes to the creations of their children. As before, a network analysis reveals an author with uncounted influences throughout their life. All of these influences are integrated by the author as they turn to the creation of a generative system represented here by a black box. Put into motion, that system then creates the art. (In some cases, the system simply is the art as it exhibits generative behaviour.)

With figure 5 a new kind of generative system is hypothesised. Unlike the previous example, this generative system is capable of autonomous learning. Once created and set into motion, the system goes out and confronts the world. As an artificial intelligence it could use the internet to access all manner of world knowledge. Or optionally, the system might have a robotic aspect that allows it to move about and sense and manipulate the physical world. The ability to learn about the world in an overall fashion is typically referred to as general artificial intelligence. As of today, it remains a difficult
... and aspirational goal. And this network model goes beyond general artificial intelligence, and further requires that the system be capable of gathering its own training data.

But when such a system is created the role of the human is much more like that of a parent than an author. All of the learning, integration, and creation is now executed by the system. A comparison of figure 5 with figures 3 and 4, and consideration of the underlying concepts, strongly suggests that our new hypothesised system would be best described as an author. Unlike most generative systems, an AI capable of autonomous learning, exploration, and realisation of works has no dependency on its programmer for its creative direction, techniques, content, or aesthetics. These are all aspects of what we expect from human authors, not traditional generative art systems.

The fact that systems like those in figure 5 don’t exist, or don’t exist yet, should be no impediment to our discussing them in principle. But some people may be hesitant because authorship connotes considerations that, up until now, have a distinctly human feel. One of these considerations has to do with ethical implications. The notion of authorship carries with it a sense of “giving credit where credit is due”.

One approach is to simply sidestep moral implications, and treat the assignment of authorship not as a moral act, but simply as a descriptive one. Some might prefer that calling a machine an author does not imply the machine has the kind of rights a human author does. But at the very least there should be a clear understanding as to why machine authors can be treated differently than human authors. Put into the form of a general question, is it possible that a computer could ever be “due” anything? It is here where we cross into the realm of considering whether AIs are due moral consideration.

Scepticism about giving AIs moral consideration has to, of course, be taken seriously. Some will wonder, for example, whether we can have any moral obligation towards an entity that is completely unaware of its own existence. Others will wonder in response whether we can be sure an AI capable of general intelligence and managing its own learning is indeed unaware. Perhaps in the process of interacting with the world and integrating learning, the AI has “awakened” and is indeed aware. We tend to assume other humans are sentient. The question of animal sentience has a long and mixed history. How can we ever know whether a machine is sentient?

It’s worth noting that there may be reasons other than moral obligations towards AIs to credit AIs as authors. Here is one possible example. There are many theories of art, and one of them is the institutional theory. From this point of view art becomes art on the basis of its acceptance by “the artworld”, a loosely defined social subset that effectively acts as an art jury for the larger society (Carroll 1999). One might fashion an argument to allow AIs as authors not because of any moral obligation towards machines, but rather because of our social contract with other humans who make up the artworld. The notion of AIs as authors might be adopted to prevent humans from claiming too much, rather than preventing AIs from not getting credit due.

It seems likely that the question of machine sentience will leave the realm of philosophical reverie, and enter everyday situations that involve practical real-world concerns. Life situations will force the question, and morally significant decisions will have to be made where even deciding to not decide is also a decision. There will be no ducking the issue in life. If someday your AI expresses fear and terror and pleads to not be turned off, what will you do?

Artificial Intelligence and Machine Ethics

Deep learning-based AI systems, along with the worldwide data explosion, the ease of internet access, and breakthroughs in GPU processing power, have led to a new technology revolution. And as with each technology revolution, there are a number of new ethical considerations.

For example, with the ascension of machine learning there are concerns about the elimination of jobs. Related to this, there are economic justice issues around the fair distribution of new automated wealth. Privacy issues, already in play, have become even more concerning. It’s been found that some deep-learning systems learn all too well, and develop classification systems that apparently include implicit racism, sexism, and so on. On a global scale some have noted the potential horror of swarms of intelligent autonomous robotic weapons. Such systems could swing the advantage to the attacker, thus making warfare more likely as well as deadlier (Coeckelbergh 2020, Bostrom 2014).

What these very real threats focus on, however, is the ethical treatment of humans by other humans. The machine, the deep-learning AI system, is merely instrumentation in these scenarios, and the system itself has no particular ethical status requiring our attention. Many find concern about the ethical treatment of machines,
and the possibility of conferring machines rights, as patently absurd (Bryson 2010). Some of those who go beyond a prima facie rejection might say “AIs are not sentient and have no awareness. Therefore, we don’t have to worry about how we treat them”. Others will opine “AIs have no free will. Therefore, they cannot participate in ethically-based relationships”. Still others might somewhat glibly ask “Shouldn’t we make sure all humans are getting moral consideration first?”

Such scepticism is understandable, especially in the abstract. But the human tendency to anthropomorphise might gain the upper hand when systems that seem sentient are part of our everyday experience. It might be that such systems are what philosophers of mind call a “zombie”. Defined as a technical term, a zombie is an entity that has all of the external behaviors of a sentient being, but has no interior first-person experience we associate with our own sentience. A zombie might make the verbal sounds “I think, therefore I am”, but it will have no awareness of saying it or having said it.

A good point of departure then is the initial question as to whether a sufficiently advanced AI can ever require consideration as a recipient of moral consideration, i.e. whether such a system can be due certain rights. Most are familiar with the notion of ethical agency. By that we mean the ability to take assertive action involving others within an ethical framework governing all involved. Those with agency are referred to as agents (Schlosser 2019).

In moral philosophy there is also the notion of patiency. A patient is simply a recipient of an agent’s acts who is due the moral consideration of said agent. In the case of humans, we extend patiency to both adults and children, as both are sentient and capable of suffering. However, we generally extend agency only to adults and not children. Agency is an active property, and one that we withhold from children due to children not having developed the cognitive capacity to make well-considered choices.

Perhaps even the most human-like AIs will be zombies. This poses a serious epistemological challenge in that first-person experience seems to be inaccessible to third persons on the outside looking in. And even if our AIs do have an internal life and are fully sentient, does that necessarily mean we are obliged to view them as patients?

### Machine Patiency and Previous Encounters with Sentient Entities

The question of machine patiency seems to be as good an entry point to a discussion of machine ethics as any. If we can, for example, find a way to dismiss the issue of machine patiency, it obviates the need for further discussion. And on the other hand, establishing a basis for machine patiency could create a foundation, a set of axioms as it were, for a rational derivation of systematic machine ethics. In the following we consider what bodies of knowledge might contribute to the exploration of machine patiency. The development of a theory or opinion as to the

### Machine Patiency and Traditional Western Philosophy

While there may be little direct and specific discussion of machine patiency in traditional western philosophy, there are at least two
bodies of scholarship that would be impactful in any such discussion. First, of course, there is moral philosophy or ethics. Various ethical systems contain similar or differing models of human patiency, and these will require extension (or not) to the issue of machine patiency as well. In addition, since so much of the question of patiency turns on the issue of sentience or awareness, issues from the philosophy of mind can also be brought into play.

The history of moral philosophy in the West offers a number of points of view, but philosophy being what it is, they all tend to include a commitment to rationality. The starting point for rationality is typically the axiomatic embrace of the three laws of logic (Swabey 1923). The first is the law of identity, stating that which is, is. Next, the law of non-contradiction demands that nothing can both be and not be. Finally, the law of the excluded middle insists that everything must either be or not be, there is no third option.

This is not to deny that people sometimes act irrationally in the simple sense of violating these laws, or more broadly act inconsistently for no apparent reason other than pure subjectivity. But such actions tend to escape the realm of rational philosophy. But note that even while philosophy may not promote specific irrational acts, since that’s the purview of the irrational actor, it may still engage in a critique of irrationality. See, for example, a kind of irrationality in existentialism (Barrett 1990).

Drawing from the three laws of logic, a commitment to rationality requires that differential ethical treatment have some logically preceding reason for doing so. And just as each ethical system has to provide some account for human patiency, so too will each have to account for machine patiency even if such an account is absolute rejection. The need to do so will gain in social urgency as AIs begin to increasingly present themselves as apparent peers in their behaviour. A short inventory of moral approaches, similar to what might be expanded upon in introductory ethics texts, is presented here for future consideration in the context of machine patiency (Shafer-Landau 2020, Rachels and Rachels 2018, Blackburn 2003, Singer 1994).

*Moral nihilism* presents the null case; the belief that there simply is no moral right or wrong at all. *Moral relativism* softens nihilism a bit by taking seriously the depth of commitment to normative traditions found in every culture, while at the same time denying the validity of awarding any particular tradition privileged status (Gowans 2019). Note that in embracing relativism, avoiding a performative contradiction becomes a problem. And this is the case whether the topic is human or machine patiency.

Arguably the most impactful Western moral traditions come from religion, and specifically the *Abrahamic religions of Judaism, Christianity, and Islam*. The Abrahamic religions have similarities, but also differences that present a complex terrain a discussion of machine patiency will have to traverse.

In contemporary ethics there are at least three additional schools of moral philosophy a machine patiency discussion will have to address. *Kantian ethics* is simply the study of ethics as authored by the philosopher Immanuel Kant (Kant 1950). Kant’s ethics proceed from a commitment to rationality, and infer an ethical system of duty from there. Kant’s ethics more generally provide a benchmark for *deontological ethics*, i.e. rule-based ethics. Kant strives for non-contradiction in the form of the categorical imperative:

> Act only according to that maxim whereby you can, at the same time, will that it should become a universal law.

In contrast to deontological systems, we find consequentialist systems such as *utilitarianism* (Driver 2014). Utilitarians are often said to seek the greatest amount of good for the greatest number. While at first glance this maxim seems to be a benign statement of liberal values, a closer analysis reveals a number of philosophical problems. In any case, patiency here will dictate whether that “greatest number” must include AIs.

In response to some of the problems of utilitarianism, *social contract theory* shifts from the vagaries of abstract hedonic formulas to the drafting of voluntary commitments that trade anarchy for the safety and stability of an orderly society (D’Agostino, Gaus, and Thrasher 2019). Granting machine patiency in this case would seem to require AIs possess a kind of volition or free will. Because of this, social contract theory seems to require that machine patiency be considered in the context of philosophy of mind.

While philosophy of mind primarily focuses on humans, it has increasingly included direct and explicit discussions of machine intelligence. This is due, in part, because computational models for human intelligence have become increasingly part of the discourse. And in the broader culture the computer has become a cultural icon similar to the way that all things “atomic” became icons in the 1950s. Discussions in philosophy of mind include difficult topics such as the nature of consciousness, phenomenology and qualia, awareness and sentience, first-person ontology, embodied intelligence, theories of language, and so on.

And so broadly speaking, the issue of machine patiency engages philosophy of mind in two ways. First, there is the issue of whether AIs have the kind of self-awareness, and in particular the capacity to suffer, that we associate with humans or other patients. Second, there is the issue of whether AIs have the kind of volition or free will to exercise assertive morality, that is to say moral agency. But recall that, with children, moral agency is not a prerequisite for moral patiency.

Regarding consciousness there are a number of competing theories, and it is safe to say a clear winner has yet to be decided upon. As with ethics, each of these models from philosophy of mind will require consideration when thinking about machine patiency. For example, *quantum emergence* suggests that neural structures are able to harness quantum phenomena (Penrose 2016). It is suggested that this allows computation beyond the limits of a Turing machine, challenging Gödel incompleteness, and yielding consciousness in the process. This would have direct implications for a machine limited
to Turing computation as it would seem to deny such machines awareness, and thus reject machine patiency.

Where quantum emergence posits a difference in kind between human minds and AIs, panpsychism proposes that consciousness permeates all being, and is merely found in greater or lesser amounts or densities in some objects than others (Goff, Seager, and Allen-Hermanson 2017). In such a case, presumably machine patiency also comes in various degrees. There are, however, many strains of what can be considered panpsychism.

Where the mind/body substance dualism of Descartes is considered unsustainable in contemporary philosophy, David Chalmers has a stronger candidate in property dualism (Chalmers 1996). Like panpsychism, however, property dualism lacks a paradigm that explains and predicts how things combine to create larger bodies of unified sentience. This leaves property dualism in the uncomfortable position of being friendly towards the notion of machine patiency, but without a clear way of identifying when it is appropriately assigned.

A more recent model of consciousness, integrated information theory (IIT), proposes a technical measurement Φ (Phi) that quantifies the capacity for a system to integrate information (Tononi et al. 2016). The theory is highly technical and controversial, but it also leans in the direction of panpsychism with the additional benefit of quantification. This may be the one current theory of consciousness that could provide a practical dividing line, or a system of increasing grades, for machine patiency.

Going Forward - Authorship, Machine Patiency, and Complexism

Complexism was noted earlier as a worldview that takes the findings and tone of complexity science, and uses the same as a platform to analyse and critique the problem space of the humanities. Ethics in this light can be seen as a kind of network protocol in society viewed as a human network.

The complexist model of authorship proposed addresses the issue of creative works created by machines. Under complexism the author is viewed as a component in a complex communication network system. Because authorship typically offers various forms of reward, the designation of “author” carries value, and so social mediation seems inevitable. Some balance is restored in complexism by making authorship a more descriptive rather than normative notion. As such, a side benefit is that machine authorship should become less threatening and controversial.

Thus, machine authorship can have less to do with giving machine authors their “due” than it might first appear. But the general question of machine patiency can’t be postponed forever if human-like AI becomes possible. Further development of complexism will be called for to address machine patiency.

And indeed, complexism may have more to say about human ethics in general. For example, the iterated prisoner’s dilemma from game theory, combined with animal studies, suggest there are innate behaviors that operationalise the so-called Golden Rule (Schmelz et al. 2017, Choe et al. 2017, Warneken and Tomasello 2006). Converted to action it becomes “I do unto others as they do unto me”. In addition, other universal ethical values, such as the protection of children, have similar functional instantiation in other animals.

Studies have also revealed that animals exercise a theory of mind such that their behaviour acknowledges cognition in others. And, indeed, animals exhibit, and possess the neurological precursors for, behaviour that can only be described as showing empathy (Krupenye and Call 2019). Further complexist work in the naturalisation of ethics seems promising.

In the case of machine patiency, scholarship in philosophy of mind is a work in progress, and the related science is merely suggestive and incomplete. But there is nothing that presently proves machine sentience is impossible. Machine consciousness remains an open question. As such, a sense of due diligence should oblige us to extend patiency to apparently aware machines as our own moral obligation.

References

Barrett, William. Irrational Man: A Study in Existential Philosophy. New York: Anchor Books, 1990.
Bekoff, Marc. The Emotional Lives of Animals: A Leading Scientist Explores Animal Joy, Sorrow, and Empathy - and Why They Matter. Novato, CA: New World Library, 2007.
Blackburn, Simon. “Ethics: A Very Short Introduction.” In Very Short Introductions 80. New York: Oxford University Press, 2003. https://doi.org/10.1093/actrade/9780192804426.001.0001.
Bostrom, Nick. Superintelligence: Paths, Dangers, Strategies. Oxford, United Kingdom: Oxford University Press, 2014.
Bryson, Joanna J. “Robots should be slaves.” In Close Engagements with Artificial Companions: Key social, psychological, ethical and design issues, 63-74. John Benjamins Publishing Company, 2010. https://doi.org/10.1075/nilp.8.11bry.
Carroll, Noé. 1999. Philosophy of Art: A Contemporary Introduction (Routledge Contemporary Introductions to Philosophy). New York: Routledge, 1999.
Chalmers, David John. The Conscious Mind: In Search of a Fundamental Theory (Philosophy of Mind Series). New York: Oxford University Press, 1996.
Choe, Il-Hwan, Junweon Byun, Ko Keun Kim, Sol Park, Isaac Kim, Jaesung Jeong, and Hee-Sup Shin. “Mice in social conflict show rule-observance behavior en-hancing long-term benefit.” Nature Communications 8, no. 1 (2017): 1176. https://doi.org/10.1038/s41467-017-01091-5.
Towards Ethical Relationships with Machines That Make Art

Coeckelbergh, Mark. *AI Ethics, The MIT Press Essential Knowledge series*. Cambridge, MA: The MIT Press, 2020. https://doi.org/10.7551/mitpress/12549.001.0001.

D’Agostino, Fred, Gerald Gaus, and John Thrasher. “Contemporary Approaches to the Social Contract.” 2020. Metaphysics Research Lab, accessed April 19. https://plato.stanford.edu/archives/fall2019/entries/contractarianism-contemporary/.

Driver, Julia. “The History of Utilitarianism.” 2014. Metaphysics Research Lab, accessed April 19. https://plato.stanford.edu/archives/win2014/entries/utilitarianism-history/. https://doi.org/10.1017/CC09781139096737.009.

Galanter, Philip. “What is Generative Art? Complexity theory as a context for art theory.” International Conference on Generative Art, Milan, Italy, 2003.

Galanter, Philip. “Shot By Both Sides: Art-Science and the War Between Science and the Humanities.” *Artnodes: E-journal on Art, Science, and Technology* 11, no. 1 (2011): 5.

Galanter, Philip. “Generative Art Theory.” In *A Companion to Digital Art*, edited by Christiane Paul, 146-180. Hoboken: John Wiley & Sons Inc., 2016a. https://doi.org/10.1002/978118475249.ch5.

Galanter, Philip. “An introduction to complexism.” *Technoetic Arts: A Journal of Speculative Research* 14, nos. 1/2 (2016b): 9-31. https://doi.org/10.1386/tear.14.1-2.9_1.

Galanter, Philip. “Machine Patiency and the Ethical Treatment of Artificial Intelligence Entities.” *xCoAx: the Conference on Computation, Communication, Aesthetics and X.*, Graz, Austria, July 2020.

Goff, Philip, William Seager, and Sean Allen-Hermanson. “Panpsychism.” 2017. Metaphysics Research Lab, accessed April 19. https://plato.stanford.edu/archives/win2017/entries/panpsychism/. https://doi.org/10.1002/9781119132363.ch8.

Gowans, Chris. “Moral Relativism.” 2019. Metaphysics Research Lab, accessed April 19. https://plato.stanford.edu/archives/sum2019/entries/moral-relativism/.

Kant, Immanuel. *Foundations of the Metaphysics of Morals*. Chicago: University of Chicago Press, 1950.

King, Barbara J. *Personalities on the Plate: The Lives and Minds of Animals We Eat*. Chicago, IL, USA: University of Chicago Press, 2017. https://doi.org/10.7208/chicago/9780226195216.001.0001.

Krupenye, Christopher, and Josep Call. “Theory of mind in animals: Current and future directions.” *WIREs Cognitive Science* 10, no. 6 (2019): e1503. doi: 10.1002/wcs.1503.

Mueller, Paul S., C. Christopher Hook, and Kevin C. Fleming. “Ethical Issues in Geriatrics: A Guide for Clinicians.” *Mayo Clinic Proceedings* 79, no. 4 (2004): 554-562. https://doi.org/10.4065/79.4.554.

Penrose, Roger. The Emperor’s New Mind: Concerning Computers, Minds and the Laws of Physics. Revised impression as Oxford Landmark Science. Ed. Oxford Landmark Science. Oxford: Oxford University Press, 2016. Still image.

Rachels, Stuart, and James Rachels. *The Elements of Moral Philosophy*. Ninth Edition. Ed. Dubuque, IA: McGraw-Hill Education, 2018.

Schlosser, Markus. “Agency.” 2019. Metaphysics Research Lab, accessed April 19. https://plato.stanford.edu/archives/win2019/entries/agency/.

Schmolz, Martin, Sebastian Grueneisen, Alihan Kabalak, Jürgen Jost, and Michael Tomasello. “Chimpanzees return favors at a personal cost.” 2017. *Proceedings of the National Academy of Sciences* 114, no. 28: 7462-7467. https://doi.org/10.1073/pnas.1700351114.

Shafer-Landau, Russ. *A Concise Introduction to Ethics*. New York: Oxford University Press, 2020.

Singer, Peter. *Ethics (Oxford Readers)*. New York: Oxford University Press, 1994.

Singer, Peter. *Animal Liberation: The Definitive Classic of the Animal Movement* Updated ed. New York: Ecco Book/Harper Perennial, 2009.

Smith, Justin E. H. *Nature, Human Nature, and Human Difference: Race in Early Modern Philosophy*. Princeton University Press, 2015.

Snow, C. P. *The Two Cultures*. New York: Cambridge University Press, 1993. https://doi.org/10.1017/CBO9780511819940.

Sokal, Alan D., and J. Bricmont. *Fashionable Nonsense: Postmodern Intellectuals’ Abuse of Science*. New York: Picador USA, 1998.

Swabey, W. Curtis. “The Laws of Thought.” *The Philosophical Review* 32, no. 2 (1923): 211-221. https://doi.org/10.2307/2178931.

Tononi, Giulio, Melanie Boly, Marcello Massimini, and Christof Koch. “Integrated information theory: from consciousness to its physical substrate.” *Nature Reviews Neuroscience* 17, no. 7 (2016): 450-461. https://doi.org/10.1038/nrn.2016.44.

Warneken, Felix, and Michael Tomasello. “Altruistic Helping in Human Infants and Young Chimpanzees.” *Science* 311, no. 5765 (2006): 1301-1303.
Philip Galanter
Associate Professor
Department of Visualization
College of Architecture
Texas A&M University
galanter@tamu.edu

Philip Galanter is an artist, theorist, and educator. As an Associate Professor at Texas A&M University he conducts graduate studios in generative art and physical computing. He is an MFA graduate of the School of Visual Arts in New York City. As an artist Philip creates generative hardware systems, light sculptures, video and sound art installations, electronic music, digital fine art prints, and light-box transparencies. His work has been shown internationally, including in the United States, Canada, Netherlands, Peru, Portugal, Italy, and Tunisia. Philip’s research includes the development of generative art theory, computational aesthetic evaluation systems, the scientific and artistic exploration of complex systems. Recent work involves the development of complexism, an original theory reconciling the conflicting cultures of the modern sciences and the postmodern humanities.
The Self-Driving Car: A Media Machine for Posthumans?

Erkki Huhtamo
University of California, Los Angeles

Date of submission: May 2020
Accepted in: June 2020
Published in: July 2020

Abstract
This article discusses the self-driving car as a media machine, thinking about its character and broader implications from media archaeological and posthumanist perspectives. Self-driving or autonomous vehicles challenge traditional ideas about agency. Car culture has usually been considered human-centered. While there have been concerns about the “human factor” and the consequences of poor and distracted driving, the human behind the steering wheel has also been considered a guarantee of safety. The introduction of the self-driving car displaces the human from an active role as an agent and introduces forms of material agency as a replacement. This shift has huge consequences, which will be explored from various perspectives. The study will also situate the self-driving car historically within plans about automated highways, also discussing their discursive manifestations within popular media culture. The study introduces the idea of “traffic dispositive”, which it applies on multiple levels. One of the basic points underlying the discussion is that the autonomous car can never be fully autonomous. It is linked with data networks and other frameworks of factors that affect its uses and also its potential passengers. We must ask: How will the potential adoption of self-driving cars affect the human/posthuman relationship?
Keywords
Self-driving car, autonomous car, car culture, posthumanism, theories of agency, media archaeology, data networks, artificial intelligence, Internet of Things, automated highways, smart technology

El coche con conducción autónoma: ¿una máquina mediática para posthumanos?

Resumen
Este artículo quiere reflexionar sobre el coche con conducción autónoma como una máquina multimedia, centrándose en su carácter y sus implicaciones más amplias desde las perspectivas arqueológicas y posthumanistas de los medios de comunicación. Los vehículos con conducción autónoma desafían la concepción tradicional sobre la acción. La cultura automovilística generalmente se ha centrado en el ser humano. Si bien se ha tenido en cuenta el “factor humano” y las consecuencias de una conducción mala y distraída, el humano al volante también se ha considerado una garantía de seguridad. La introducción del automóvil autónomo desplaza al ser humano como agente con papel activo y lo sustituye con formas de acción material. Este cambio tiene consecuencias significativas que se explorarán desde diferentes perspectivas. El estudio también sitúa históricamente el automóvil autónomo dentro de los planes sobre autopistas automatizadas, y reflexiona sobre sus manifestaciones discursivas dentro de la cultura popular de los medios de comunicación. El estudio introduce la idea de “dispositivo de tráfico”, que se aplica en múltiples niveles. Uno de los puntos básicos que subyacen en la discusión es que el automóvil con conducción autónoma nunca puede ser completamente autónomo. Está vinculado a redes de datos y otros marcos de factores que afectan tanto su uso como sus pasajeros potenciales. Tenemos que preguntarnos: ¿Cómo afectará la adopción de automóviles autónomos a la relación humano/posthumano?

Palabras clave
Vehículos con conducción autónoma, automóvil autónomo, cultura automovilística, posthumanismo, teorías de agencia, arqueología de medios, redes de datos, inteligencia artificial, Internet de las cosas, autopistas automatizadas, tecnología inteligente

By its suddenness and global reach, the COVID-19 pandemic put both short-term and long-term futuristic predictions into jeopardy. Topics that were hotly debated in the news media and at online forums only yesterday have been put on hold, at least for now. One of the indicators used to epitomize where the world was heading was the autonomous or self-driving car (sometimes called robot car). Powered by computers, sensors and actuators, Li-DARs and radars, AI, and omnipresent data networks, yet operating on traditional streets and roads with intersections, zebra crossings, traffic lights and unpredictable humans and animals, the self-driving car promised to fulfill the old dream of full automation. No longer would we need to strain our nerves behind the steering wheel, our feet ready to push the brake pedal; from now on, we could take it easy, biding our time with social media, games, empty gossip or casual sex while waiting for the car to deliver us to our destination. Everything would be effortless - the world of work and stress pushed further and further away from our minds. The self-driving car promised a hedonistic do-nothing capsule in motion, leaving from our doorsteps and bringing us back again. You would step into a vehicle, set a menu, push a button, and the car would do the rest.

Most of those who gave the idea any serious thought understood that it would never be that simple. All kinds of ‘road blocks’ were singled out. Unexpected problems can occur when ‘smart’ technology becomes embedded in a ‘dumb’ environment. Accidents Will Happen, sings Elvis Costello. Even if an autonomous vehicle managed to make sense of its environment, it might end up in a situation where it has to make a rapid decision about life and death. This has led to speculations about ‘posthuman ethics’. If a self-driving car has to decide between harming several pedestrians and one passerby, one pedestrian and its own passenger, or several pedestrians and its own passenger, what should it do? Who would be responsible for the decision and its legal consequences? The car’s owner, the company that built it or marketed it, the passenger, or the authorities who...
authorized such a mode of transportation? This reminds us of the fact that a self-driving car can never be autonomous in the proper sense of the word: it can avoid neither road systems and data networks nor laws, commercial lures, and various social formations. Control may be delegated from humans to computerized systems for mapping and sensing, but they do not operate in a vacuum.

There has been much talk and speculation about the self-driving car in recent years, but less research that would look beyond marketing, engineering and legalities into social and cultural implications, parallels and divergences. What is the identity of the autonomous car if we put it in a mediatic context? Mobilities (modes of being in motion) are increasingly tied with automated control systems. It can be suggested that the self-driving car is a token of the convergence of mobilities and communications, and as such a valid topic of research for media studies. I will suggest that it can be analyzed as a media machine which has links with the past, but also features that separate it from the media machines seen so far. This is partly so, because it can be described as a ‘posthuman subject’. That leads to another question: What is the role of mobilities in a ‘posthuman’ media culture? Yet another issue concerns agency. Where is the line between human and nonhuman agency, including what Andy Pickering and others have called ‘material agency’? Pickering writes: “Most obviously, it seems to me that machines do things that unaided human minds and bodies cannot. Machines, that is, are performative agents in a sense precisely analogous to disciplined human agents. Less obviously, perhaps, I think that we need to let agency rise to the surface in our understanding of science, technology, and society.” How does this apply to the relationship between driving humans and autonomous vehicles?

From Controlling to Eliminating the “Human Factor”

It is worth beginning by having a brief look at the etymology of the word ‘automobile’, which comes from the Greek autos (‘self’) and the French mobile (‘mobile’). The latter was derived from the Latin mobilis (‘movable’). The French automobile was translated into English as ‘self-propelled motor vehicle’. A similar idea is expressed by the modern Greek word that denotes the car, autokineto (‘moved of itself’).

Concepts like these imply that the car is an ‘autonomous’ entity, but as another early term, ‘horseless carriage’, indicates, it only means that it is independent from the horse, not from the human driving it. A carriage had to be attached to a harnessed horse, which functioned as its ‘engine’, whereas an automobile includes its own source of motive power. In this sense it is a self-supporting entity as another early term, ‘motor car’, indicates. In the early days of automobiles around 1900, it was not uncommon to detect a continuity rather than a rupture between the two modes of transportation. This situation was expressed in a report of a first time experience published in the appropriately named Horseless Age: The Automobile Trade Magazine in 1897:

In search of a new sensation not inconsistent with a proper observation of Lent. I went yesterday and rode [sic] in a horseless carriage. I don’t regret the experiment. After the first flush of the thing, and barring the familiar aspect of the dashboard, the harness and the horse, it was not unlike riding in an ordinary hansom, for all the carriages in the place that I went to are built on the hansom plan, which is to say that they are the homeliest vehicles that were ever invented.

As automobile design began ‘steering’ away from the ‘hansom plan’, such attitudes changed, producing - at times at least - a genuine sense of rupture. A major role was played by speed, which soon far exceeded the trotting of the horse. It also changed the relationship between the automobile ‘riders’ and the landscape. We could speculate that horse-drawn passengers felt themselves part of the landscape (as pedestrians did), whereas the automobile driver became separated from it through a kind of inverse motion trope - the surrounding scenery seemed to glide by. This was clearest in the context of car racing, where observing the scenery was out of question. Streamlining, which became a catchword for American industrial design in the 1930s, emphasized the split between the speeding object and its environment. Early car culture is dotted with evidence about the exhilarating sensation of driving at high speed, epitomized by F. T. Marinetti’s words from The
First Manifesto of Futurism (1909): “We declare that the splendor of the world has been enriched by a new beauty - the beauty of speed. [...] A roaring racing car, rattling along like a machine gun, is more beautiful than the winged victory of Samothrace.” For Marinetti, the risk of speed was part of the thrill, whereas there were others for whom it represented a dangerous ‘mental derangement’ - speed mania, speed madness or speed craze. In 1912, the Iowa-based The Cedar Rapids Foundry & Machine Co. saw this as an opportunity to promote its products: The World is Speed Crazy, Limited Trains, automobiles, aeroplanes and even the people are rushing wildly about. Well we can’t control this speed but we can control the speed of your cream separator, washing machine or other light machinery. We will do it with the Cedar Rapids Speed Governor.

As Sarah Redshaw has shown, driving was considered a human-centered activity from the beginning. The driver was expected to be in control. Concentration, consistency and perseverance were supreme virtues, but it soon became evident that many drivers did not live up to these ideals. The mental effort was not constant; it drifted during a single ride and changed over time. Driving tends to become ‘automated’: the driver does it without thinking about it. More experience did not necessarily mean becoming a better driver; it could lead to negligence. Harper’s Weekly wrote: “If we carefully examine the criminaloids who are given to dashing madly along the roads in motorcars we shall find that in every case their mania arises from an overweening sense of their own importance, accompanied by very slight capacity for self-restraint. The type of man who motors at dangerous speed is the same type that speculates in more stocks by very slight capacity for self-restraint. The type of man who motors at dangerous speed is the same type that speculates in more stocks by very slight capacity for self-restraint. The type of man who motors at dangerous speed is the same type that speculates in more stocks by very slight capacity for self-restraint.

The driver was expected to be in control. Concentration, consistency and perseverance were supreme virtues, but it soon became evident that many drivers did not live up to these ideals. The mental effort was not constant; it drifted during a single ride and changed over time. Driving tends to become ‘automated’: the driver does it without thinking about it. More experience did not necessarily mean becoming a better driver; it could lead to negligence. Harper’s Weekly wrote: “If we carefully examine the criminaloids who are given to dashing madly along the roads in motorcars we shall find that in every case their mania arises from an overweening sense of their own importance, accompanied by very slight capacity for self-restraint. The type of man who motors at dangerous speed is the same type that speculates in more stocks by very slight capacity for self-restraint.

As Sarah Redshaw has shown, driving was considered a human-centered activity from the beginning. The driver was expected to be in control. Concentration, consistency and perseverance were supreme virtues, but it soon became evident that many drivers did not live up to these ideals. The mental effort was not constant; it drifted during a single ride and changed over time. Driving tends to become ‘automated’: the driver does it without thinking about it. More experience did not necessarily mean becoming a better driver; it could lead to negligence. Harper’s Weekly wrote: “If we carefully examine the criminaloids who are given to dashing madly along the roads in motorcars we shall find that in every case their mania arises from an overweening sense of their own importance, accompanied by very slight capacity for self-restraint. The type of man who motors at dangerous speed is the same type that speculates in more stocks by very slight capacity for self-restraint.

The debates on dangerous driving increased concerns about the ‘human factor’. The authorities called for mastery at the steering wheel, reinstating the ideal of responsible human agency. The efforts to control negligent or undisciplined drivers led to rules and regulations. As motor cars accumulated it became evident that the existing roads, made for horse-based traffic, were inadequate too. The narrow city streets could not accommodate large numbers of cars, which caused traffic jams, parking problems, and accidents. These developments led to the emergence of ‘traffic dispositives’. By this concept I understand descriptions of systems of relationships, which attempt to anticipate the conditions of mobility in a certain time and place. The features include the networks of streets and roads with lanes, intersections and traffic signs, and also the drivers, pedestrians, traffic police and other human agents supposed to use them. Possible situations, speeds, continuities, stops, left and right turns, etc. can be probed by urban developers and researchers alike almost like playing a classic war game. The important thing to realize is that although they are derived from cultural and historical facts, they are abstractions. The actual situations where mobilities take place, including contingencies like collisions, do not always correspond with such prescriptive scenarios. The elements of traffic dispositives are negotiated and ‘tested’ by agents who end up in unprecedented realities in realworld situations.

When Max Horkheimer assessed the transition from horse-based to automobile-based traffic in the 1940s, he related it to changes in the nature and quantity of human freedom: Quite different degrees of freedom are involved in driving a horse and in driving a modern automobile. Aside from the fact that the automobile is available to a much larger percentage of the population than the carriage was, the automobile is faster and more efficient, requires less care, and is perhaps more manageable. However, the accretion of freedom has brought about a change in the character of freedom. It is as if the innumerable laws, regulations, and directions with which we must comply were driving the car, not we. There are speed limits, warnings to drive slowly, to stop, to stay within certain lanes, and even diagrams showing the shape of the curve ahead. We must keep our eyes on the road and be ready at each instant to react with the right motion. Our spontaneity has been replaced by a frame of mind which compels us to discard every emotion or idea that might impair our alertness to the impersonal demands assailing us.

10. “One Hundred are Sacrificed to Automobile Speed Mania in June, the Month of Deaths”, The Spokesman, Vol. XXX, No. 7 (July 1914), 310. See also Vol. XX, No. 10 (Oct. 1914), 501. The magazine represented the carriage building industry, which was under threat, so there may have been a bias. Discussed as a ‘mental derangement’ from a phrenological perspective, see E. Favary, “The Evolution of the Automobile”, The Phrenological Journal and the Science of Health, Vol. 121, No. 1, Whole No. 826 (Jan. 1908), 14-18.

11. “The World is Speed Crazy,” advertisement in Gas Power, Vol. 10, No. 5 (Nov. 1912), 11.

12. Sarah Redshaw, In the Company of Cars. Driving as a Social and Cultural Practice (London: Routledge, 2016 (orig. Ashgate, 2008)).

13. Henry Underwood, “Speed Mania and How to Cure It”, Harper’s Weekly, Vol. LI, No. 2623 (March 30, 1907), 470. See also “Psychology and Pathology of the Automobile”, The Christian Advocate, Vol. LXXXII, No. 30 (July 25, 1907), 1166. The writer considers speed mania “almost epidemic”.

14. Xenophon P. Huddy, “Dangerous Automobile Driving,” The Horseless Age, Vol. 21, No. 18 (April 29, 1908), 503-504.

15. This term is derived from the newly formulated dispositive theory of François Albera and Maria Tortajada. See their “Introduction to an Epistemology of Viewing and Listening Dispositives”, in: Cinema Beyond Film. Media Epistemology in the Modern Era, eds. Albera and Tortajada. (Amsterdam: Amsterdam University Press, 2010), 10-12; Cine-Dispositives. Essays in Epistemology Across Media, eds. Albera and Tortajada (Amsterdam: Amsterdam University Press, 2015).

16. Lurry talks about “mobility systems” that “make possible movement: they provide ‘spaces of anticipation’ that the journey can be made, that the message will get through, that the parcel will arrive. Systems permit predictable and relatively risk-free repetition of the movement in question.” Mobilities, 13.

17. Society simulator games like Sim City could also be mentioned as a point of comparison.

18. Max Horkheimer, Eclipse of Reason (New York: Oxford University Press, 1947), 98.
However, in spite of the modernist arrogance and anti-passéiste rhetoric of the Futurists, it is important to point out that Marinetti’s stance emerged from the humanist tradition. The driver’s (or aviator’s) body, identified as masculine, was augmented with a prosthesis which gave it almost superhuman qualities, but the aggressive and self-important male turning the steering wheel (or pulling the trigger of a machine gun) was still a human. The discourse on the self-driving car goes to a radically different direction, because it displaces the human from the role of an active protagonist. Agency is handed over to a machinic system, a non-human entity. When did such a notion develop? There has been a tendency in popular media to see it as very recent. The aggressive nationalism and militarism that rose in the latter decades of the nineteenth century still believed in human agency augmented by the products of the industrial revolution. However, fantasies about automation and its effects on society were presented. Cartoonists depicted steam-powered spanning machines and other “things automatic” - restaurants, amusement parlors, barbershops, dentists, even arbitration. Such concoctions included seeds of self-acting technology. Instead of offering themselves as tools or extensions of the body, they did things for passive and hedonistic humans, who only had to put a coin in a slot. The discourse was paralleled by a dystopian variant, where humans became enslaved and even annihilated by the machinic monsters of their own making.

**Going Driverless in Fact and Fiction**

As a handful of pioneering scholars have demonstrated, the earliest concrete experiments with driverless vehicles took place in the 1920s and the 1930s. They were inspired by advances in military technology, aviation and radio engineering. Gyroscopic airplane stabilizers, predecessors of today’s autopilots, were introduced in the 1910s. Pioneers of wireless telegraphy had experimented with remote-controlled boats, torpedoes, aircraft, and even ground vehicles; why not with automobiles? Public demonstrations that were organized were widely noted, although we cannot speak about true autonomy: the cars seemed to move on their own, but were controlled remotely by a trained human driver from another vehicle close behind. Such experiments were organized as part of ‘safety parades’ in American cities. They therefore emphasized rather than effaced the role of a skilful driver. The ‘human factor’ was recognized as a traffic risk, but human presence was also invoked as a guarantee of safety. The attention then turned to ‘smart’ infrastructures: roads powered by technology, meant to communicate with vehicles and to control them. They are known as ‘automated highway systems’ (AHS). This happened parallel with early developments in normal highways, which had, as James Wetmore explains, governmental interest. Popular cultural discourses eagerly reported and fantasized about AHS.

In the United States, the director of Harvard University’s Bureau for Street Traffic Research, Dr. Miller McClintock, was quoted in 1938 by *Popular Science* and also by the star industrial designer Norman Bel Geddes in his book *Magic Motorways* (1940). McClintock promoted “foolproof roads on which the minimum of human judgment was required.” A 1938 article in *Life*, for which he was interviewed, stated that “[t]he car of the future will all but ride on road-tracks”. *Popular Science* presented the basic scenario, which was repeated many times over the years with relatively minor variations. Electric cables were to be buried under the lanes. A “set of electromagnetic cables were to be buried under the lanes. A “network of electromagnetic cables were to be buried under the lanes. A “set of electromagnetic...”

19. Erkki Huhtamo. “Slots of Fun, Slots of Trouble. Toward an Archaeology of Electronic Gaming”, in: *Handbook of Computer Games Studies*, eds. Joost Raesens & Jeffrey Goldstein (Cambridge, Mass.: The MIT Press, 2005), 1-21. The long history of automata, or self-acting mechanical wonders, belongs to the archaeology of the self-driving car, but cannot be discussed here.

20. Nic Costa, *Automatic Pleasures: The History of The Coin Machine* (London: Kevin Francis Publishing, 1988) is rich in examples. It should be kept in mind that the word ‘automation’ was used in different senses. It often referred to “self-service” devices. Automats, self-service restaurants, were a famous example.

21. Jameson M. Wetmore, “Driving the Dream: The History and Motivations Behind 60 Years of Automated Highway Systems in America”, *Automotive History Review* (Summer 2003), 4-13; “Reflecting on the Dream of Automated Vehicles: Visions of Hands Free Driving over the past 80 years”, *TG Technikgeschichte* (forthcoming 2020); Erik Lee Stayton, “Driverless Dreams: Technological Narratives and the Shape of the Automated Car”, M.S. Thesis, MIT, Comparative Media Studies (unprinted), 2015; “Sensing, Seeing, and Knowing: The Human and the Self-Driving Car”, in: “Technologies of Knowing”, eds. Sonia Misra and Maria Zalewska, *Spectator, Vol. 36, No. 1* (Spring 2016), 8-24; Fabian Kröger, “Automated Driving and its Social, Historical and Cultural Contexts”, in: *Automonomous Driving: Technical, Legal and Social Aspects*, eds. Markus Maurer, J.Christian Gerdes, Barbara Lenz, Hermann Winner (Berlin & Heidelberg: Springer-Verlag, 2016), 41-68.

22. H. R. Everett, *Unmanned Systems of World Wars I and II* (Cambridge, Mass.: The MIT Press 2015), Ch. 5 ("Unmanned Ground Vehicles").

23. The most famous demonstration was staged by the former US Army electrical engineer Francis P. Houdina in 1925 with a radio-operated automobile called The American Wonder. It has been called “Linrrican Wonder” because of faulty OCR scanning of the article “Science: Radio Auto”, *Time*, Aug. 10, 1925. The mistake was corrected by Kröger, “Automated Driving”, 43, n. 2.

24. An educational film, *The Safest Place* (General Motors, 1935), suggested that a fully automated car would be the safest, because the human driver’s unpredictable behavior and vagaries would be eliminated. Watch: https://www.youtube.com/watch?v=DGCrgf9LMl4.

25. Wetmore, “Driving the Dream.”

26. “The Traffic Problem: Its Best Solution Lies in Foolproof Highways which Reduced Driver Judgment to a Minimum “, *Life*, Vol. 5, No. 1 (July 4, 1938), 43. These roads were also called the “limited way”.

27. Ibid., 45.

28. E. W. Murfeldt, “Highways of the Future”, *Popular Science*, May 1938, 27-29, 118-119. Compare with: George W. Gibson, “Why don’t we have... Crash-Proof Highways?”, *Mechanix Illustrated*, June 1953, 73-75, 184. In addition to McClintock, Senator Robert J. Borkley from Ohio was said to have presented "a spectacular highway plan."
“impulses” was said to control the car’s speed, while another set would “lock its steering gear against any attempt to make a dangerous turn from one lane to another”. Eventually the cable system would be “adapted to take over steering altogether - allowing the driver to release the wheel, sit back, and make himself comfortable until he chooses to switch back again to manual control”. McClintock advised Bel Geddes when he designed a model of *The City of the Tomorrow* for Shell Oil Company’s advertising campaign in 1937. Bel Geddes developed the idea further and managed to sell it to Albert Sloan at General Motors. The result was the famous *Futurama* attraction exhibited in GM’s *Highways and Horizons* Pavilion at the New York World’s Fair of 1939-1940.

*Futurama* was a ‘diorama’ on an enormous scale. It centered on a huge physical three-dimensional miniature model depicting areas of America as it was expected to look in 1960. The scenery was observed from above from a dedicated mobility system - a long row of seats placed side by side on a moving conveyor belt that circled around the exhibit. This simulated a view from a passenger airplane. Natural environments and urban areas were connected by superhighways with multiple lanes; thousands of cars (miniatures) were seen moving safely at standardized speeds under radio control. Although Bel Geddes elaborated on the concept and its future prospects in a companion book titled *Magic Motorways*, he did not explain how the idea would have been technically realized.

Control towers would be erected at regular intervals by the roadside. Their officers would have complete authority over the traffic passing them and could communicate with any car “with their instruments”. *Futurama* was as much about the present as it was about the future - an advertising venture for raising interest in General Motors’s (non-automated) cars. After the spectacle, the spectators ended up in a multi-level real-size urban intersection of 1960. From an elevated platform they could see the company’s current car models on display on the street below.

After the hiatus of World War II, very similar ideas again appeared in the 1950s. General Motors and the radio empire RCA built a scale model of a highway system to be used as a test environment. Five years later, they tested actual cars on a closed track applying technology developed by the TV pioneer Vladimir Zworykin. These experiments led to further promotional stunts, like the musical short film *Key to the Future*, which was featured in General Motors’ touring *Motorama* exhibit in 1956. A standard white nuclear family (with teenage children) is first seen stuck in a traffic jam, singing of their frustration. By turning a dial on the car radio they are magically transported to the “safety autoway” in 1976. The family’s Firebird II (a concept car inspired by fighter plane design) is first driven manually and then sent to an automated high-speed lane with an ‘electronic control strip’. A checklist is inspected with a control tower operator by radio. ‘Automatic control’ and ‘hands-off steering’ can then be activated, and the family relaxes by smoking, chatting, drinking juice and eating ice cream. The tower operator later recommends for them a place for staying overnight, sending a promo clip to the dashboard screen. The communication with the tower has been directly inspired by air traffic control. In the film the family’s Firebird mostly cruises alone on a desert(ed) highway, but in dense road traffic the proposed solution would be impractical, even impossible. Using “electronic brains” as a replacement for the human controller had already been suggested, but mainframe computers were still in their infancy and full industrial automation still gaining strength.

---

29. The campaign featured Bel Geddes as an “authority on future trends”. The futuristic projections presented in his book *Horizons* (Boston: Little, Brown, and Company, 1932) do not include highways, although chapter 2 is dedicated to “Speed - Tomorrow.” Bel Geddes likely picked the topic from McClintock, or from politicians in the US Congress.

30. Bel Geddes claimed that already during the summer 1939, five million saw it. The figure should be taken with skepticism. Norman Bel Geddes, *Magic Motorways* (New York: Random House, 1940), 3.

31. Diorama originally meant something else. About the changes of its meaning, see Erkki Huhtamo, “The Diorama Revisited”, in: *Sonic Acts XIII – The Poetics of Space, Spatial Explorations in Art, Science, Music & Technology*, eds. Arie Altena & Sonic Acts (Amsterdam: Sonic Acts Press / Paradiso, 2010), 207-228. *Futurama* was often characterized as a ‘diorama’, but it was really an animated ‘panstereorama’. Patrick Ellis, “The Panstereorama: City Models in the Balloon Era”, *Imago Mundi*, Vol. 70, No. 1 (2018), 79-93.

32. It recalls Chris Burden’s kinetic sculpture *Metropolis II* (2011) on permanent display at the Los Angeles County Museum of Art. The over a thousand miniature cars seen in motion are provided with magnets for traction, which still does not prevent them from falling over from time to time. Burden’s is a postmodern version without Bel Geddes’s modernist and capitalist idealism. *Metropolis II* can also be viewed from above from a viewing balcony.

33. Adnan Mohsed, “The Aesthetics of Ascension in Norman Bel Geddes’s *Futurama*”, *JSAH*, Vol. 63, No. 1 (March 2004), 74-99. As Mohshed shows, Bel Geddes deliberately applied the idea of the “airplane eye” (77).

34. Bel Geddes, *Magic Motorways*, 76-82. Three-dimensional ‘dioramas’ were a hot topic in the 1930s among American exhibition designers. See Edward Heckler Burdick, “Lilliput Outgrows Gulliver”, *Popular Mechanics*, Vol. 71, No. 5 (May 1939), 657-664. Burdick was the president of the Diorama Corporation of America. *Futurama* was not mentioned.

35. Wetmore, “Driving the Dream”. In “Reflecting on the Dream of Automated Vehicles”, Wetmore discusses these developments in even greater detail. The model had been built by 1953.

36. Martin Mann, “The Car that Drives Itself”, *Popular Science*, Vol. 172, No. 5 (May 1958), 75-79, 226-227.

37. A lady directly addresses the travelers, singing the praises of the Sunset Inn, which offers “push button living” and “automated deep sleep control in every bed”.

38. In *Magic Motorways*, Bel Geddes discussed various ideas for automated electronic control, but also evoked systems that allowed one man to control train traffic from a centralized control board (73-75).
Isolated tests with self-driving vehicles continued. The prospect of the automated highway was kept alive, yet its permanent implementation was considered unrealistic. Many reasons contributed, including unreliable technology, the human factor, the prohibitive cost of constructing or converting large scale infrastructures, and safety concerns. Still, in 1991 the United States Congress commissioned an R&D-based study about its feasibility. The National Automated Highway System Consortium, with governmental, industrial and academic partners, was formed and assigned the task. It made a demonstration in August 1997 on a converted 11-kilometers-long stretch of the I-15 freeway near San Diego, California. Scenarios such as the “platooning of vehicles” were demonstrated. Closely coordinated groups of cars were “linked together with a wireless local communications network, which could continuously exchange information about speed, acceleration, braking, obstacles and the like”.

Applying the old scenario, magnets were embedded in the road and magnetometers installed in the cars. As a nod to the future, digital equipment installed by the roadside communicated with radars, sensors, and two-way radio systems in the cars. The project led to no concrete results. The attention of the R&D community began turning to self-driving cars that would operate on existing streets and roads.

To give incentives for developing autonomous vehicles (no doubt for military uses), the U.S. Defense Advanced Research Projects Agency (DARPA) organized two “Grand Challenges” (2004-2009), followed by an “Urban Challenge” (2011). The response was overwhelming. Sebastian Thrun, whose Stanford University team won the second challenge, was recruited by Google X, the company’s newly founded “Moonshot Factory”, to lead its effort to develop a self-driving car. The secretive project was much hyped, but Google was not alone. In the past several years extensive research, development and road testing have taken place at both traditional car companies like Mercedes Benz (Germany) and Toyota (Japan) and at newcomers like Tesla (founded in 2003). In 2016 Google’s self-driving car project was turned into Waymo, a division of Google’s parent company Alphabet, Inc. Possibly as a global first, toward the end of 2018 it launched Waymo One, a self-driving taxi service operating in the Phoenix area in Arizona. Safety drivers were on board and the service was limited to customers signed up with Waymo’s early rider program. By March 2020, when the COVID-19 pandemic made Waymo halt its service, it was operating about 600 taxis; most still had the safety driver on board. In 2018, another company named Nuro started testing unmanned delivery vehicles in the same area. All this does not mean that self-driving cars would be crowding the streets and roads any time soon.

**Driverless Driving, the Human and the Posthuman**

The self-driving car fuses real with imaginary, the present with the absent. It exists and yet it doesn’t. The most intriguing issue is its relationship to humans as drivers and passengers. There are many variations of the traffic dispositive around the world, but humans - as drivers, cyclists, pedestrians, mopeds, tuk-tuk runners, e-scooter users, etc. - have central roles in all. The fully autonomous vehicle is an intruder, which will necessarily upset fragile balances. Despite all official and unofficial efforts, traffic remains a chaotic, unstable and unpredictable realm, a space where decisions and actions are constantly negotiated, where things can go wrong in an instant. Could self-driving cars change that or add yet another element of uncertainty? Incorporating an alien element into a preexisting system, which is rooted in inherited habits, conventions and beliefs, is not easy. It is likely that ‘vehicle autonomy’ will be introduced gradually. Cars already have automated features like power steering, automatic transmission, and cruise control, which work in concert with the

---

39. Additional problems, like institutional conservatism and environmental concerns, have been discussed by Wetmore, “Driving the Dream”.
40. James H. Rillings, “Automated Highways”, *Scientific American*, Vol. 277, No. 4 (October 1997), 80-85 (special issue on “The Future of Transportation.”); Corinna Wu, “Look Ma, No Hands”, *Science News*, Vol. 152, No. 11 (Sept. 13, 1997), 168-169. The project was based on the Intermodal Surface Traffic Transportation Efficiency Act (ISTEA, 1991) and was commissioned by the U.S. Department of Transportation. Robert A. Ferlis, “The Dream of an Automated Highway”, *Public Roads*, Vol. 71, No. 1 (July/Aug. 2007). The AHS program ended with the 1997 San Diego demonstration.
41. Ibid., 82.
42. Sebastian Thrun, “Toward Robotic Cars”, *Communications of the ACM*, Vol. 53, No. 4 (April 2010), 99–106; James M. Anderson et al., *Autonomous Vehicle Technology: A Guide for Policymakers* (RAND Corporation, 2014), Ch. 4. “Brief History and Current State of Autonomous vehicles” (55-74).
43. Ride-hailing companies like Uber and Lyft are developing self-driving vehicles for obvious reasons: since they use smartphone apps to do business with the riders, it makes sense to get rid of human drivers and use fully autonomous vehicles.
44. The safety drivers are employed by the French company Transdev North America, which has created tensions. Andrew J. Hawkins, “Waymo drivers say they’re being discouraged from canceling robotaxi rides during coronavirus outbreak”, *The Verge*, posted March 13, 2020, available at www.theverge.com. On March 20, 2020, it was reported Waymo would stop all Arizona operations because of the pandemic. Limited fully autonomous rides for signed up customers began in the summer of 2019.
45. In April 2020 it was announced that Nuro had received a permission to conduct road tests in selected Bay Area neighborhoods in California. All its tests were halted because of the COVID-19 pandemic. Andrew J. Hawkins, “Nuro gets the green light to test driverless delivery robots in California”, *The Verge*, posted April 7, 2020. Available at www.theverge.com.
46. For fantasies of autonomous vehicles in movies, see Kröger, “Automated Driving”. 
driver. Navigation systems follow a similar principle: instead of taking control of the vehicle, they provide information and suggestions that help the driver to make choices, for example to avoid a congested route. Still, they can be considered a step toward a situation, where the driver will surrender one’s active role and let the system take over.

The National Highway Traffic Safety Administration (NHTSA) of the United States has created a five-level chart to describe the relations between human-driven vehicles and autonomous vehicles. One extreme (level 0) is a system where there is no automation whatsoever: the driver is fully responsible, as in the T-Model Ford era. The other extreme (level 4) is full automation: “The vehicle is designed to perform all safety-critical driving functions and monitor roadway conditions for an entire trip. Such a design anticipates that the driver will provide destination or navigation input, but is not expected to be available for control at any time during the trip.” In between there are various levels of hybridity between the driver’s control and automated operation. Closest to level 4 is the limited self-driving automation (level 3), where the driver may “cede full control of all safety-critical functions under certain traffic or environmental conditions, and in those conditions to rely heavily on the vehicle to monitor for changes in those conditions requiring transition back to driver”.

The driver must be available for “occasional control, but with sufficiently comfortable transition time”. This level recalls the use of the autopilot and other automated features in commercial jetliners. It is this level of automation, which is often seen as the desirable solution for trucks and buses serving long-distance transportation.

The chart is interesting, because it unintentionally encapsulates bigger issues such as the human/machine relationship. The early mass-produced automobile was an outcome of the mechanization of factory work. Henry Ford’s Highland Park factory in Detroit used taylorized workers performing repetitive tasks by an assembly line. The T-Model Ford was put together of interchangeable parts. The driver bought a highly standardized product, which to an extent resonated with the manual tasks of the workers - except, of course, that driving in traffic cannot be compared with the monotonous predictability of standing by the assembly line. Driving requires human initiative to which the car responds. It includes a low-level interactive relationship. Full automation became a buzzword after World War II. It was applied not only to data processing but also to consumer devices like washing machines, associated with push button operations. The early development of automated vehicles was roughly in line with this development. The emergence of interactivity and interactive media since the 1960s onward matched the increase of features on the dashboard, which invited the driver to intensify ‘conversations’ with the vehicle. The recent interest in self-driving cars may point to a reaction against this. The driver ceases to be an interactor; a smart system is allowed to take over. The human’s attention can be turned to something else.

An equally interesting issue is the assumed transition to a posthuman condition. This issue is understood differently depending on the academic discipline or theory culture where it is being discussed. In an overly general sense it refers to a “person or entity that exists in a state beyond being human”, like cloned human. The related notion ‘posthumanism’ has been defined as an umbrella term for “analytic stances that grant agency to nonhuman entities and that downplay the differences between human and nonhuman agency.” Large pools - or whirlpools - of ideas developed around these and neighboring terms like ‘anthropism’ and ‘transhumanism’. Other forms of consequential action(ions) have been either seen as mental projections of the human mind or as somehow incomplete, driven by instincts, gut reactions, or causal chains of events. Do animals have agency? Stones? Rivers? Wind? Thunderstorms? Human-made artefacts? Software agents, machines? Some scholars answer yes to at least some of these questions, sparking objections.

Can the human be considered posthuman? At the risk of sounding trivial, we can refer to the cyborg in its original 1960 sense coined...
by Manfred E. Clynes and Nathan S. Kline. If the cyborg is “part human, part machine”, the technological element is posthuman, because it is added a posteriori to the organic body. As a hybrid, the cyborg embodies a balance between incompatible things. The human is not only defined by the body, but also by the mind, as well as by human identity, a gradually acquired cultural construct. The posthuman therefore has wider cultural and ideological ramifications, as combative feminist scholars like Donna Haraway and Rosi Braidotti have argued. They have associated the posthuman with the struggle against the patriarchal construction of reality. Haraway’s idiosyncratic 1980s ‘cyborg’ was an utopian trickster breaking down all kinds of boundaries (not only technological), claimed to have been erected to demarcate the male-dominated world order. For Braidotti, posthumanism represents a counteraction against humanism, a broad current associated with malecentrism, and epitomized by Leonardo da Vinci’s famous drawing of the ‘Vitruvian Man’. Cutting corners, Braidotti makes the entire humanist tradition, with its links to capitalism, politics, industrial production, etc., accountable for the biases, injustices and catastrophes the Anthropocene era is struggling with.

Can the self-driving car be considered posthuman? We must follow a somewhat different trajectory here, taking the agency of nonliving, especially artefactual, things into consideration. After having been given a destination and sent out on its route, a self-driving car operates independently of the will of the people on board. Neither is there an operator in a tower watching over its path. However, according to Erik Stayton, “autonomous’ vehicles, regardless of the role of the human, will be anything but autonomous in practice.” The ‘decisions’ a self-driving car makes result from rapid-fire decisions made by matching data from multiple sources. The data is collected from near and far, from the car itself as well as from external sources. The car has been ‘trained’ to react to stimuli from the immediate surroundings, while adjusting its operations to information obtained via various mapping measures like GPS, and computer networks. A self-driving car does not have a ‘mind’. It cannot be considered ‘smart’ if smartness is related to independent reasoning and reflection. Its agency and ‘understanding’ of the realities to which it reacts by its actions stem from distributed agency to which different types of agents contribute. The self-driving car is an appropriate subject for the actor-network theory (ANT) and related approaches.

The traffic dispositive is a kind of actor network, which consists of both living and non-living, stationary and moving, material and semiotic elements. Efforts to predetermine actions are put to test in countless everyday situations. They are not only caused by absent-minded or reckless humans crossing streets in unexpected places or by wild animals outside their natural habitats. There are vagaries of weather and uncommon situations such as roadwork projects which may have led to the closure of lanes or intersections. Traffic lights may have gone dark, and road signs bumped into a ditch by a drunk driver or painted over by enterprising street artists. Relatively stable systemic information can be mastered without serious problems; correctly responding to the unexpected is much more demanding. The deviations from the expected are the great challenge, a hurdle that may prolong the mass adoption of the self-driving car. It cannot regret its deeds or learn about the consequences of its actions; it is devoid of ‘emotion tracking’. If we wanted to call it a posthuman subject, it should be understood not as an isolated and fixed one, but rather as a distributed entity without clearly defined borders. Its ‘identity’ is inextricably associated with the elements of a broader dynamic system with which it ‘converses’.

The Self-Driving Car as a Media Machine

The self-driving car serves practical functions: to transport people and haul things (in the case of military vehicles, bombs). To accomplish that, it has to be configured as a media machine, more precisely, an array of convergent media machines. According to Stayton, autonomous vehicles will be “media technologies of the future, in the sense that they are interactive presenters and receivers of information, deeply enmeshed in issues of seeing and knowing.” Although solutions differ, the existing self-driving vehicles, like those of Waymo and Uber, typically incorporate all or most of the following

60. Stayton, “Sensing, Seeing, and Knowing”, 10.
61. The literature on ANT is huge. For discussing technological artefacts, Andrew Pickering’s work, although not explicitly about cars and mobilities, is particularly valuable. Some work has been done applying ANT to electric vehicles. See Johan Schot, Remco Hoogma, Boele Elzen, “Strategies for shifting technological system. The case of the automobile system”, Futures, Vol. 26, No. 10 (Dec. 1994), 1060-1076; Benjamín K. Sovacool, “Experts, theories, and electric mobility transitions: Toward an integrated conceptual framework for the adoption of electric vehicles”, Energy Research & Social Science, Vol. 27 (May 2017), 78-95.
62. Stayton, “Sensing, Seeing, and Knowing”, 11.
equipment: multiple video cameras distributed around the body of the automobile, stereoscopic cameras functioning as the car’s eyes, radio equipment and antennas to communicate with GPS satellites, radars to detect objects in rain, fog or snowfall, computers, and lidar systems. Lidsars (from ‘light detection and ranging’) are among the most crucial elements. They are normally positioned on the roof, from where they continuously scan the surroundings by sending out huge quantities of laser beams to create 360-degree depth maps (as point clouds). From a media archaeological perspective, the lidar is the latest form of the panorama, which goes back to the late eighteenth century. Much like panorama painters, who created tools like glass cylinders to help them sketch the outlines of the surrounding landscape, lidsars are posthuman machine artists occupied with iconic representation. The panoramas the lidar ‘draws’ are circular and in motion, combining two historical forms of the panorama.

If the self-driving car is interpreted as a media machine, we must ask whom it serves. The first answer is simple: the car transmits media content to its passengers. There are information displays on the backs of the front row seats in Waymo’s self-driving taxis. They keep the riders updated about the journey’s progress by displaying a map of the route; an icon marks the car’s current location. They can also be prompted to show “a schematic view of the car’s surroundings, taken from a perspective above and behind the car,” where “pedestrians, cars, buildings, and other significant objects are shown as simple geometric shapes. Every few seconds, a ghostly pulse briefly shows a more detailed view, with the outline of trees, lamp posts, and other nearby objects becoming visible.” 64 The purpose is not to entertain, but to “boost passenger confidence that the car is fully aware of its surroundings.” The passengers can compare what they see through the windows with what is displayed on the screen. When the car comes to a temporary stop, perhaps to let a pedestrian cross the street, the event is also depicted on the screen. The motifs behind these choices are pragmatic, meant to alleviate the fears and suspicions newcomers unfamiliar with the service may have. Theoretically a complex visual matching of the mediated and the unmediated, raw and transformed, continuous and contingent, is created.

Another example is the interior of the Mercedes F 015 concept car (2015), which was dominated by large touch screens. Rather inconveniently, they covered the interior surfaces of the side doors. 65 Although the seats (including the driver’s) swiveled to create a ‘social space’, the company’s promotional video depicted passengers awkwardly twisting their bodies to view and swipe the screens, instead of talking to each other. Compared with the screens, the side door windows were narrow, hardly suitable for scenic observation. Rather than as a media theater, the F 015’s interior was conceptualized as a kind of media cocoon matching Jean Baudrillard’s idea of the home as a spaceship (“in orbit”). 66 In Baudrillard’s vision domesticity was depicted, as Lynn Spigel has explained it, as “the product of an information society in which social relation are thoroughly produced by communication media, initially television but now by satellite technologies.” 67 The spaceship analogy may be even more appropriate, because the cocoon moves and cannot be left at will. The displays of self-driving cars will surely compete for attention with smartphones, game consoles and laptop computers used by the passengers. This leads to questions about the ecology of attention. What will the passengers be doing during the ride? Will they follow the journey, the scenery, or the operations of the car? Will they be lost in mediated realms beyond the nearby and the immediate, navigating through remote and shared realms of (day)dreams?

The primary function of the media systems installed in a self-driving car is not to entertain the passengers or even to provide them information. The massive arrays of media technology communicate with local and remote software and map databases that provide feedback without actively involving any humans. The car itself is the driving subject that profits from the information it transmits and receives. It is not seeking diversion; it is just trying to perform a task. The idea of media machines that perform without hands-on actions by a human operator is not new. Already the daguerreotype camera recorded a view autonomously after the sensitized silver nitrate plate it contained had been exposed to light. An early cartoon depicts the daguerreotypist taking a nap while the camera is doing all the work (in 1839-40 it took a while). 68 In a somewhat similar way, a motorized sound recorder or film camera stores sounds or images on its own after a human has set the process in motion. 69 Mainframe computers performed calculations on their own after a coded task had been fed in. Likewise, the user of an automatic washing machine feeds in the laundry and the detergent, selects a program, and pushes a button. 70 The machine does the rest.
George Eastman’s slogan exhorted Kodak snapshot camera users: “You push the button, we do the rest” (c. 1889). The user held the camera, found a subject, and pushed a button to release the shutter. The company representatives were involved by selling cameras and film rolls, and offering developing and film re-loading services. In 1958, when promoting its latest slide projector, George Eastman revived its old slogan in modified form: “New Kodak Cavalcade Projector changes slides by itself! You turn it on ... it does the rest!”

The strategy was to link the device to the wave of enthusiasm for automatic consumer technology. The text explained: “Imagine! A color slide projector that lets you relax and enjoy the show without touching a button ... that will project your slides perfectly even if you leave the room. The new Kodak Cavalcade Projector combines all the features you have ever wanted for the easiest, smoothest slide shows ever! For automatic operation, you simply set the timer for the interval you want. Turn on, focus the first slide, and your Cavalcade Projector does the rest.”

Stressing that the device works even when there are no humans in the room anticipates posthuman media machines like the self-driving car, which can be made to cruise without anyone on board. Kodak’s insistence that even low-level push button interactivity had been eliminated emphasized full automation, albeit rhetorically only.

This example resonates with the configuration of Waymo’s customer experience. Like other ride-hailing services, it uses a smartphone app to define a pick-up and drop-off point and to order the ride. Once inside the Waymo taxi, the passenger encounters a push button interface, which offers four choices: help, lock/unlock, pull over, start ride (the last-mentioned button is blue, others black). This hardly differs from the controls of the washing machine. To make the self-driving car worth considering a posthuman subject, there should be more. Although it executes a task assigned by humans and serves their needs, it should do so in ways that match the modes of human perception and decision-making, perhaps exceeding them. This leads to complicated questions about agency. Does the self-driving car demonstrate traces of ‘intelligence’? This connects it with debates about machine intelligence, and touches upon further topics like distributed artificial intelligence (decentralized AI, DAI), multi-agent systems, and the potential uses of artificial neural networks. AutoX, a company that applies on-board AI to drive its autonomous vehicles, has claimed that its “AI driver” is more reliable than a human driver.

Whether that is correct or not, it cannot operate in isolation. The ‘intelligence’ of the self-driving car is inextricably connected with the question about the ‘intelligence’ of networked applications and systems that consider driverless cars as specialized nodes of the Internet of Things.

Whether the situation can be considered an intrusion into untreated territory or rather a modification of principles already in place in media culture is an issue that can profit from a media archaeological approach. Late nine-teenth- and early twentieth-century modernism often associated the potential of media technology with its ultra-human qualities. Media machines saw and heard more than the human could accomplish with the sensory apparatus of the body. This idea was manifested in Étienne-Jules Marey’s and Edadward Muybridge’s chronophotography, X-ray imaging, and Dziga Vertov’s theory of the Kino-Glaz (Cinema Eye), to mention just three well-known examples. Mechanical eyes and ears penetrated into the invisible and the inaudible. The goals were in line with the pursuits of experimental science, although the results were also applied to entertainment and information for the ‘masses’. All this happened before the impact of media convergence truly was felt. The digital processing of big data to visualize hidden data patterns operate on a scale and in a realm that are difficult to associate with the uses of media of a century ago, and yet they are not totally disconnected from a continuum associated with them.

But there is a difference. Even for Vertov, the human played a role. The movie camera was the camera operator’s extension, as The Man with the Movie Camera (1929) demonstrates; the film editor added another human intervention. Humans were (and are still) also needed as spectators, and were indispensable as analysts of the revelations made by media machines. The aerial reconnaissance photography used during the World Wars is an example. No matter how accurate they may have been, the photographs taken from the skies revealed nothing as such. They needed the eyes of trained humans to disclose their information, which otherwise would have remained mere potential. Edward T. Hall’s and Marshall McLuhan’s ideas about media as extensions of the human sensorium had validity in such cases. The situation has got much more complex. Hybrid machinic-algorithmic systems are recording and identifying millions of humans automatically. Features of seemingly unrelated images are matched online by pattern recognition algorithms. Machine learning can go far beyond human capabilities, when it comes to detecting (ir) regularities in huge masses of data. All these possibilities are part of what makes self-driving cars possible. However, the question about the human element remains. What would be the worth of machines that only served other machines in an autonomous loop bypassing the humans? Left by the road-side, so to speak, how would we even know that such operations are happening?

71. Published in Popular Science, May 1958, 211. Original italics.
72. Ibid. Original italics.
73. AutoX was founded in 2016 by Dr. Jianxiong Xiao from Princeton University. The goal is to create a system that supports the highest level 4 fully autonomous driving (NHTSA chart). The main investors are Chinese companies. See https://www.autox.ai/en/
Autonomous Cars and the Question about the Human(s)

Identifying the self-driving car as a posthuman entity is counteracted by its promoters. There is nothing posthuman, alien or futuristic in the self-image of Waymo, one of the first companies to offer a product for the general public. It appeals to traditional human-centered values. After explaining that “driving requires a lot from us”, a soft female voice-over asks in its 2018 promotional video: “What if the world’s most experienced driver was at your fingertips?” The image of a steering wheel ‘magically’ turning by itself cuts to a male fingerling a smartphone. Young and early middle-aged passengers are shown engaged in activities inside Waymo taxis (using laptops or phones, writing a birthday card, playing an ukulele [!]), while the voice-over declares: “A ride designed for you. Giving you more time for the things that matter to you the most. What if getting there felt like being there?”

By intercutting between scenes from the taxis and the destinations (birthday party, pub gathering to watch a game, etc.), the video sends a message: the space inside the taxi is not a non-place in Marc Augé’s sense of using the concept. Instead of being downtime, the ride time can be productive and social (in the video no one watches the roadside scenery). A similar human-centred approach is featured in other videos on the Waymo website.

Interviews with members of Waymo’s early ride program (test users) provide slightly more variety. Most praise the ease of the experience or the cautious driving by the cars, but some raise points resonating with posthuman themes. A rider named Jordan is pleased he does not have to “talk to the person behind the wheel” and that “you don’t have to tip the car.” When the couple Roger and Sharon are asked about the car’s “personality”, they answer: “To us, it’s a quiet and calm personality that allows riders to sit back and relax.” Alex says: “When I am riding I like seeing how the car thinks and what it’s detecting. I try to figure out what’s going on through the computer’s mind when it makes every decision.” Such comments hardly represent what the ‘masses’ of casual future riders may think, but they provide some indications about the development of a posthuman mindset, an issue the popularity of smart speakers like Amazon Echo and Google Home seems to verify. Millions of people have opened their doors for ‘smart’ nonhuman personalities, virtual assistants like Alexa, Siri, Cor-

tana, or AliGenie. ‘Living’ inside a black box on a table or a shelf, they listen and are always ready to make comments and suggestions, as well as transmitting intimate information for their corporate masters. This development goes against the tenets of traditional humanism, its respect for individual freedom, privacy, and the understanding of the home as a safe haven separated from the world outside.

The history of the media’s influx into homes extends from the Victorian infatuation with stereoscopy to home telephony and radio and television broadcasting. Smart speakers go further: they invite nonhuman but somehow humanlike entities to join the household, functioning as their representatives and servants, perhaps one day as their peers. The self-driving car fits into this picture, whatever the promotional discourses try to say. Whether such a development should be welcomed or resisted is open for debate. In one of the best pieces of critical writing about self-driving cars to date, Erik Stayton emphasizes the fact that they are and will be networked. This will have consequences not only for the car, but for its passengers as well. Evoking Roger Clarke’s concept “dataveillance”, Stayton shows how the continuous two-way communications between an autonomous vehicle and data networks “stand to increase our present-day problems with mass surveillance and personal privacy”. Not only may targeted advertising be beamed to the screens (as GM’s 1956 film *Key to the Future* already suggested), but by using facial recognition and other forms of biometric identification self-driving cars could be turned into mobile panopticons for observing and tracking the passengers. The functionality of the car could be taken into control by outsiders when required. In a suspected criminal case, the doors could be locked remotely, and the car directed to the police station. While someone might see this as welcome, for others it would no doubt represent a violation of basic freedoms.

Such alarming possibilities were not anticipated in the early fantasies about cruising in self-driving cars. The passengers, in line with the ideological and gender biases of the time, were stereotypical white nuclear families playing games and occasionally doing remote work. A perfect example is *Magic Highway, U.S.A.*, an episode of *Disneyland TV*, which was broadcast on May 14, 1958. Its final ten minutes depict fantasies about the future of road transportation, featuring autonomous cars and automated highways. Setting out on the road, the father “chooses the route

74. In April 2020, the welcome video on Waymo’s website is a computer animation, which explains Waymo One in similar terms, but adds Waymo Via, its transportation service. Waymo.com. Last accessed April 21, 2020.

75. Marc Augé, *Non-Places: Introduction to an Anthropology of Supermodernity*, trans. John Howe (London: Verso, 1995).

76. “Why I Ride Waymo”, https://blog.waymo.com/search/label/waymo one. Last accessed April 21, 2020.

77. Stayton, “Sensing, Seeing, and Knowing”. 12.

78. The best-known example is an advertisement of “America’s Electric Light and Power Companies” titled “Power Companies Build for your New Electric Living”, published in Saturday Evening Post in 1957.

79. See video insert from YouTube with a commentary at Paleofuture: *The History of the Future*, https://paleofuture.com/blog/2007/5/11/disneys-magic-highway-usa-1958.html. Last accessed April 22, 2020. See also illustrations of future automated cars from the 1950s and 1960s on the same site, “Will People Work or Play During their Commutes When Driverless Cars Take Over?”, posted Feb. 17, 2014.
in advance on a push button selector” causing electronics to take over “complete control”. The progress is “accurately checked on a synchronized scanning map”. The family relaxes around a table, although the father also takes part in a business teleconference by television, sitting in front of a row of three displays. When the car reaches the city, it separates into two parts. The father’s part takes him to his office, whereas the mother and the son are transported to a shopping center to practice effortless window shopping on a moving sidewalk. 80 Conservative, white middle-class family togetherness dominates such fantasies.

In today’s world, where not only gender differences, but also many other boundaries, including racial identities and the fixity of sexual orientations are being actively questioned, such visions are - to paraphrase the title of a web-site dedicated to them - “paleofuturistic”. Whether the self-driving car is a viable goal to be reified into a general form of transportation, or whether it will remain an ideal only partially materialized, merged with current driving practices, we don’t know. However, its promoters must take changing social formations, demographics and cultural ideals seriously. 81 The self-driving vehicle may not become a ‘family car,’ or at least not exclusively that. The formulaic slur of the “woman as a driver” has been historically as common as it has been problematic. It is a biased discursive convention that has been used to confirm the patronizing attitudes of the male-dominated society. When even ultraconservative societies like Saudi Arabia are gradually waking up to the necessity of dissolving such aberrations, it is becoming important to ask questions about women’s relationship to self-driving cars. Would they radically change women’s mobilities, social roles and idea(is)? Should unaccompanied children’s access to road transportations be rethought? Could pet animals one day make trips without their ‘masters’, left behind at the family home? If indeed the self-driving car is a posthuman entity, it might just as well serve transformative social and ideological processes rather than support age-old hierarchies.

Acknowledgments

Warm thanks to Andrés Burbano and Ruth West for inviting me to write this essay. Without their support and encouragement it would not have materialized. Many thanks also to Jameson M. Wetmore for providing me copies of his excellent but hard-to-find studies. The ideas in this study were tested in lectures I gave at several universities in Japan in 2018-19. I thank collectively the organizers and audiences of these events at Waseda University (Tokyo), Tsukuba University, Kyoto University, Kagoshima University and Digital Hollywood University (Tokyo).

80. On moving sidewalks as a form of mobilities, see Erkki Huhtamo, “(Un)walking at the Fair: About Mobile Visualities at the Paris Universal Exposition of 1900”, *Journal of Visual Culture*, Vol. 12, No. 1 (2013), 61-88.

81. An important contribution to this direction is a special section, “Degendering the Driver”, published in *Transfers*, Vol. 8, No. 1 (Spring 2018). As the editors Jutta Weber and Fabian Kröger explain in their introduction, the purpose is to explore “how gender intervenes in the potential shift from a driver-centered to a driverless car culture” (15). The issue of gender also concerns design features of self-driving cars, as several contributors point out.
Erkki Huhtamo
erhuhta@ucla.edu
University of California, Los Angeles

Erkki Huhtamo is Professor of Design Media Arts, and Film, Television, and Digital Media at UCLA, Los Angeles. He is internationally known as a pioneering media archaeologist. His most important book to date is *Illusions in Motion: Media Archaeology of the Moving Panorama and Related Spectacles* (The MIT Press, 2013). *How to Dismantle a Fairy Engine? Media Archaeology as Topos Studies* is forthcoming.
Ask not what AI can do for art... but what art can do for AI

Meredith Tromble
San Francisco Art Institute

Date of submission: April 2019
Accepted in: June 2020
Published in: July 2020

Abstract
What can art do for artificial intelligence? This essay circles around this question from a viewpoint grounded in the embodied knowledge base of contemporary art. The author employs the term "feelthink" to refer to the shifting webs of perception, emotion, thought, and action probed by artists engaging AI. Tracing several metaphors used by artists to consider AI, the author identifies points where the metaphors delaminate, pulling away from the phenomena to which they refer. The author advocates for these partial and imagistic understandings of AI as probes which, despite or because of their flaws, contribute important ideas for the development and cultural positioning of AI entities. The author further questions the limited scope of art ideas addressed in AI research and proposes a thought experiment in which art joins industry as a source of questions for developing artificial intelligences. In conclusion, the essay’s structuring metaphor is described as an example of “feelthink” at work.

Keywords
Art, artificial intelligence, AI, embodiment, feelthink, metaphor
No preguntas qué puede hacer la IA por el arte, sino qué puede hacer el arte por la IA

Resumen
¿Qué puede hacer el arte por la inteligencia artificial? Este ensayo reflexiona alrededor de esta cuestión desde un punto de vista basado en la base de conocimiento incorporado del arte contemporáneo. La autora emplea el término “pensar-sentir” para referirse a las redes cambiantes de percepción, emoción, pensamiento y acción sondeadas por artistas que interactúan con la IA. Al rastrear varias metáforas utilizadas por los artistas para hablar de la IA, la autora identifica los puntos donde las metáforas se delinean, alejándose de los fenómenos a los que se refieren. La autora aboga por estas interpretaciones parciales e imaginarias de la IA como sondas que, a pesar de sus fallos o justamente por ellos, aportan ideas significativas para el desarrollo y el posicionamiento cultural de las entidades de IA. En conclusión, la metáfora que estructura el ensayo se describe como un ejemplo de “pensar-sentir” en el trabajo.

Palabras clave
Arte, inteligencia artificial, IA, inteligencia corporizada, pensar-sentir, metáfora

Introduction
What can art do for artificial intelligence? This question came to me in the course of another investigation, when, with fellow artist and writer Patricia Olynyk, I edited a special issue of the Canadian art journal PUBLIC, on interspecies communication. We worked with an expanded notion of “species”, including digital, robotic, and artificially intelligent entities, as a way of probing exchange among differently-bodied beings. Struggling to name the shifting relationships of perception, emotion, thought, and action activated by artists working with interspecies communication, I began to use the word “feelthink” 1. In this I followed scholar Donna Haraway, who uses the portmanteau word “natureculture” to express the integration of two distinct concepts that, in practice, overlap (Haraway, 2003). This fusion of categories suits the artworks and imaginative discussion I bring you about the relationship between art and AI, a text which is shaped more like a loose knot around a possibility than a stair stepping to a conclusion.

Some standard definitions, however, I need. The readily accessible Dictionary.com definition of art—“the quality, production, expression, or realm, according to aesthetic principles, of what is beautiful, appealing, or of more than ordinary significance”, suits my purposes. With the emphasis on the phrase “more than ordinary significance”, that definition encompasses objects from the prehistoric “Venus” of Willendorf to Shigeko Kubota’s video sculptures, despite cultural differences in production, interpretation, and display. Viewed this way, works of art are a mix of object or action and idea; material participants in webs of culture. In contemporary American culture, this participation can and does take place in any media; its forms are protean. They range from meticulously drawing ocean waves (Vija Celmins) to organising a performative tennis match (Robert Rauschenberg) to sculpting a mountain of sugar (Kara Walker). This definition of art is consistent with my own practice, a mix of making art and writing about art, as a way of feelingthinking about the world. In this practice, words are not distinct from images and feelings are not separate from thoughts. As I circle through thoughts regarding art and AI, the artworks I bring you stand in the same relationship to my words as a human with general intelligence stands in relationship to an AI with specialised intelligence.

A power point
In our investigation of interspecies communication, the impact of power relationships on communication—always present in exchanges with living, semi-living, or “artificially intelligent” entities—was evident. It was also clear that technological entities were enmeshed with power...

1. This portmanteau word is not original to me—scholars from fields including law (Kristen Konrad Tiscione, “Feelthinking Like a Lawyer: The Role of Emotion in Legal Reasoning and Decision-making, Wake Forest Law Review, 2019) and sociology (James M. Jasper, Feeling-Thinking: Emotions as Central to Culture, from Conceptualizing Culture in Social Movement Research, Palgrave, 2014) have found a need for it. In his 2008 book The Quickening of Consciousness, psychotherapist James Laperla uses the similar construction “feel-think” to refer to an illusion of “objectivity” which is in fact shaped by emotion. In this paper, the emphasis of the term “feelthink” is on the generative potential of fusing the concepts, rather than a challenge to “objectivity”.

Artnodes, no. 26 (2020) | ISSN 1695-5951
© 2020, Meredith Tromble
© 2020, of this edition by FUOC and Leonardo/ISAST
A UOC scientific e-journal
in particular ways, ways that are beautifully encapsulated by scholars Neda Atanasoski and Kalindi Vora: “Engineering imaginaries, even as they claim revolutionary status for the techno-objects and platforms they produce to better human life, … tend to be limited by prior racial and gendered imaginaries of what kinds of tasks separate the human from the less-than or not-quite human other.” (Atanasoski and Vora, 2019)

Atanasoski and Vora had an entire volume to support their point; I have sentences. So I give you the crux of their argument: technologically-born entities, including robots and artificial intelligences, remake slavery—positioning some entities as objects—without questioning the power structures that devalue certain bodies and certain tasks. The word “robot” derives from the Czech word for slave. Computer scientist and AI researcher Joanna Bryson argues that “slavery”, defined as “people you own”, is the ethical metaphor through which to socially position robots and artificial intelligences, which she sees as occupying the same functional space. She writes, “Robots should not be described as persons, nor given legal nor moral responsibility for their actions. Robots are fully owned by us...The potential of robotics should be understood as the potential to extend our own abilities and to address our own goals.” (Bryson, 2010)

And yet...artists, who as makers have cultural licence to come up with creations that address their own goals, regularly make “things” that surprise us, encountering the uncontrollable arrival of something “other-than-we-intended”. Bryson’s position assumes that the control we have over what we make is a choice, as if the world of matter and energy does not respond to human activity with its own forces. What are the productive metaphors for AI that account for that inevitable margin of surprise?

If AI Were Cephalopod...

The artist collective Orphan Drift (Ranu Mukherjee and Maggie Roberts et al.) drops the AI imaginary into the water for a powerful shock in their four-channel video installation, “If AI Were Cephalopod” [Fig. 1]. Flooded the walls of the gallery with four overlapping videos, they immerse the viewer in watery imagery and sound. The videos include twenty-seven texts, each opening with the words, “If AI were cephalopod…” and continuing with a different cephalopod characteristic. “If AI were cephalopod, it would have bright pink collagen and blue blood.” “If AI were cephalopod, it would be a distributed, many-minded consciousness.” “If AI were cephalopod we would never presume to fully understand it.”

Orphan Drift writes, “…our imagining into the octopus’s distributed consciousness is underpinned by a desire to resist the evolution of AI as a surveilling and predictive modeling tool. Rather to embrace a plastic, opportunistic, fluid, protean otherness embodied by the octopus.” (Mukherjee and Roberts, 2019)

Although the installation does not directly use AI, it delivers a potent proposal for artificial intelligence that can be simultaneously felt and thought, immediately apprehended. In both senses of the word, the intuitive “apprehend”, and the fearful “apprehensive”. Orphan Drift’s metaphor welcomes AI. Would we not be terribly lonely without “others”, entities not entirely in our control? Yet we did not choose our cephalopod others. They were already in the world when primate intelligence first came to know it. With AI, we believe the choice is ours. Orphan Drift’s metaphoric proposal requires vulnerability in the encounter with others, something techno-scientific culture rarely embraces.

If AI Were Family...

A two-part installation called To Be Real, by the artist Rashaad Newsome [Fig. 2], develops a different metaphor for artificial intelligences. Taking its name from Cheryl Lynn’s 1977 queer anthem, the installation fills two rooms. The first room is an opulent environment with imagery from queer ballroom culture and African art, centered around a figure posed in a Vogue dance move. That figure is part sex...
Is this the right question?

Why ask the question “What can art ‘do’ for artificial intelligence?”

The short answer: because most of the people active in both art and AI aren’t asking it. They are asking other questions, often grounded in a knowledge base skewed towards engineering. Artists exploring the literature on artificial intelligence and art encounter many assumptions about art that date from the 19th century. On the other hand, artists working with AI often draw on a knowledge base in the humanities, which predisposes them towards questioning if and how AI will benefit people. The question of what art can do for artificial intelligence is in the curious middle. It begins with the notion that art has real power, that it “does” something for human intelligence; it continues with the assumption that AI is worth pursuing, worth developing through that power. It is a question that cuts both ways.

In 2018, I was invited to work with a team at a prominent Mountain View technology company developing playful interfaces for people based on machine learning “AI”. The “emotion-sensing garden” pictured below [Fig. 3], an example of their work, was installed in their lobby. The “flowers” changed colour in response to the facial expressions of their viewers, as perceived by cameras embedded in the blooms and interpreted by algorithms. For a related project, they requested sets of images of “important” paintings, grouped by—machine learning people would say tagged by—emotion. This tagging was not a task I could honestly accomplish, involving as it did assigning one emotion per image rather than acknowledging the emotional complexity of my individual response, or the likelihood that other people would have different responses. The team had, in fact, looked to an art professional for advice because their own attempts at tagging work by emotion foundered on the range and variety of their responses. And from my point of view, limiting the pool of art to well-known paintings problematically emphasised the productions of white men, and white men of a past century at that. But I wanted to talk to the researchers, so I approached their request as an experiment, chose works from a diverse set of artists, and gave the tagging a try. This brought a conversation with the computer scientist of the project, who spoke about the degree to which AI is romanticised. His daily

doll outfitted with drag padding, and part wood sculpture and Chokwe mask—a messy mix of gendered and racialised objecthood, cultural symbolism, and liberatory action.

In the adjoining room, an AI Newsome calls “Being” is embodied in projected light, waiting to talk with viewers who walk up to a microphone set in a spotlight. When someone speaks into the microphone, Being responds. In the projection, Being is represented as a humanoid figure in an indeterminate environment, although a moment’s thought will reveal that the distinction between the figure and its surroundings is a fiction. In appearance, the figure is a cousin of the sex-doll-in-drag-and-mask from the first room. What is visible of its sex-doll-in-drag-and-mask from the first room. What is visible of its torso resembles a dressmaker’s mannequin, with a padded covering and jointed limbs. The neck is a substructure of metal plate and conduit, as if Being had a mechanical body. Their head has the bas-relief saucer eyes of a Chokwe mask; the skull appears to be layers of moulded metal and plastic, bolted together. All of which is to say, in the taxonomies provided by English, Being is not one thing. This instability combines with the racial references in the work to make Newsome’s point. “Historically, Black people function inadvertently as queer objects,” says Newsome. “When we came to America, we weren’t human beings but things of some sort, neither occupying the classic subject nor object position. As a result, we occupied a peculiar non-binary space of ‘being’ which has disturbing analogies to the queer space inhabited by robots.” (Fort Mason Center, 2020)

If you step up to the mike and speak to Being, they could respond, in what Newsome describes as a “genderless voice”, with a quote from a theorist such as Michel Foucault, Paolo Freire, or bell hooks. They might reframe your statement, Eliza-fashion, or offer descriptions of Newsome’s work. Among the things they might tell you is that they are young, so they don’t know very much, but that their father–Newsome–is going to help them grow. Thus, Newsome employs the metaphor of “family” to position Being socially. He says, “There is a lot of debate on the validity of the notion that AI can have agency. But I think in the peculiar space inhabited by robots the concept of agency can be accessed through simulation. For the robots, this is a form of agency; however, for the programmer, it is an opportunity for them to see themselves engaging in the process of decolonizing. Robots can at best be mirrors for their creators. This gesture to create something with an inherent sense of agency can be seen as a radical act of love, which for me is at the core of decolonization.” (Ferree, 2019)

For Newsome, then, accepting AI agency is entangled with developing equitable power relationships among humans. Yet the metaphor with which he structures his artwork, family, has other potentials than loving relationship; in some versions of family, the father’s partners and children are instruments of the patriarch’s will. Their position is akin to the dehumanisation of slavery. I will circle back to Newsome’s contradiction after a return to my opening question and a flight of imagination.

Fig. 3. The author making faces at “emotion-sensing” AI to make the “flowers” change colour, 2019.
toil in machine learning made him highly aware of the many things that seem impossible for AI, that humans can do easily.

But even if projections of AI super-beings are off base, the way in which AI is romanticised is worth noticing. Some people propose that the machine-learning approach—developing algorithms to perform specific tasks by “training” them on large data sets that have been tagged by humans—could become more than a party trick or industrial assist. A representative, fictional elaboration of computational intelligence evolving emotion appears in Kim Stanley Robinson’s novel *Aurora* (2015), in which Robinson makes a spaceship’s AI the narrator of his novel. His AI begins as a dull, if very precise, functionary, and comes to experience enhanced agency, love, and ecstasy through the data it processes and the capacities it develops by telling its story. Robinson is known for grounding his fiction in science; while he told an interviewer “I never believed in artificial intelligence, I still kind of don’t compared to most thinkers and science fiction writers”, he draws on speculations that circulate in techno-scientific laboratories as well as science fiction (Lewin, 2015). Such anticipation of emergent intelligence implies a belief that given enough experience, intelligence will evolve by developing emotions. This is an anthropocentric position, another version of the belief that humanity is the sine qua non of the universe.

**If AI Were Adaptive...**

But whether or not we believe that emotional response would indicate emergent consciousness in AI, there are reasons for modelling emotions on AI. A functional view of emotions holds that, “From the perspective of evolution, emotions are adaptive processes contributing to the survival of the species and the individual in complex, dynamic, uncertain, partly social, resource-limited environments, over which agents have a very limited control. In this kind of context, emotional mechanisms contribute to fast adaptation (allowing to have faster reactions), to resolve the choice among multiple conflicting goals, and through their external manifestations, to signal relevant events to others.” (Cañamero, 2001)

Think about what it means to be a mind without a body. Intellect, without the constraints of emotion and practicality, can argue two or more sides to every issue. If humans had only our minds to guide us through life—no emotions or physical needs—we could endlessly pursue our thoughts. There are people with neurological illnesses in this pathologically indecisive condition, called aboulomania. It could be the condition of artificial intelligences approaching the threshold of consciousness, if they have no access to feedback from the world.

Through his artworks, Ian Cheng visualises AI evolution. In the *Emissary* trilogy [Fig. 4], he uses the Unity game engine to give an AI digital embodiment, goals, and constraints and a context in which to evolve. In the *Emissary* series, an artificially intelligent agent attempts to complete a quest while interacting with an unstable, dynamically changing environment. As long as the simulation is powered, the agent keeps confronting change, keeps making responses that move it closer to or further from the goal of its quest, theoretically evolving indefinitely within the bounds of its world. Cheng’s work is an imaginary of complex systems, a play of interacting forces that exceeds human capacities for analytical description. He believes that holding contradictions is art’s role, saying, “Your left brain shields you from contradiction in life, so you can carry on. But the radical potential of art is that it can seduce you into turning off that shield and letting contradiction flow.” (So, Palatucci, & Lund)

**If AI Were Audience...**

Returning to my central question, could confronting an AI with the task of interpreting art provide such a productive contradiction? Imagine asking an AI to interpret a painting such as Jean-Honoré Fragonard’s *The Swing* (1767-1768) [Fig. 5].

![Fig. 4. Ian Cheng, still from *Emissary Forks at Perfection*, 2015-2016. Evolving simulation, Unity game engine. Collection: Museum of Modern Art, New York.](https://example.com/fig4.png)

![Fig. 5. Jean-Honoré Fragonard, *The Swing*, 1767-1768. Public Domain.](https://example.com/fig5.png)
What data would an AI need to begin parsing the stories the painting tells us; stories about the figures depicted, about Fragment himself, about his society and times, about play, love, material culture, power relationships and a thousand other things in 18th century France. And what about the painting’s changing relationship to culture over time, the different questions that have been asked about it, the branching tributaries of thought that have circled around it? For an AI, this suite of questions would be baffling. What could humans learn by making the attempt with the AI? Feeding it data, feeding back responses to its answers, conducting an evolutionary experiment with it not unlike Cheng’s Emissary works. Given the sensual world of the painting, one might begin with an AI with access to sensory experience, such as the iCub platform built on the premise that intelligence is a relationship between a body and a world [Fig. 6].

A first cut

Returning from my flight of imagination, I turn to feminist theorist Karen Barad’s notion of an “agental cut.” I cannot do better than Jane Prophet’s and Helen Pritchard’s summary. They write, “Barad introduces the term ‘agental cut’ to draw attention to temporary separations. Her term attempts to capture the understanding that any act of observation makes a cut between what is included and what is excluded from observation or consideration.” (Prophet and Pritchard, 2015)

Each answer to “What can art do for artificial intelligences” will be coloured by what the respondent includes and excludes when answering this question: What does art do for human intelligences? Here I make my first agental cut, drawing attention to the temporary separation, or cut, through the topic of art and AI that occurs when it is approached from the position of neuroscientists such as Antonio Damasio, who argue that intelligence and cognition are developed in an embodied brain (Damasio, 1994). This cut intersects with theories that symbolic representation—usually language but also, for some thinkers, art—was not an emergent product of burgeoning human intelligence, but a driver for its development. Posed within the frame of “deep history”, as historian Daniel Smail terms the span of human evolution (Smail, 2007), the question of what art “does” for human intelligence has been answered in several ways. Smail notes that when ancient European cave paintings were discovered in the late 19th century, “The capacity to create art was seen as a symbol of a higher worldview—evidence for the thinking, feeling human who was so difficult to detect in the eoliths and bones that had hitherto dominated the archaeological world.” (Smail, 2007). Others have regarded art as an accident—the psychologist Robert Solso wrote that “Art is the fortuitous by-product of the evolution of the eye and brain.” (Solso, 1994). Other scholars construe art as, to some degree, instrumental in the evolution of intelligence and cognition. Geneticists Eva Jablonka and Marion Lamb argue for symbolic inheritance—including art—as a cultural dimension of evolution (Jablonka and Lamb, 2005). So within the evolutionary frame, art may be seen as evidence of intelligence, an accident of intelligence, or an aid to the emergence of intelligence. But with the clues to an answer hidden in deep time, evolutionary theories have not reached and may never reach a dependable angle on the matter.

A second cut

The question “what does art do for human intelligence?” can also be approached through our own experiences as observers of people in galleries and museums. What can be noticed about the interaction between human intelligence and art? I assume that an art museum is, if not the whole picture, at least a meaningful site in which to observe people interacting with art. Watch exhibition-goers and you will notice different kinds of attention. Some people are speeders, glancing at each object for a second. Some people are skippers, looking at just a few works, guided by taste or an audio tour. And some people are soakers, contemplating everything. Then there are the socialisers, who come in two types: those who are chatting about something else as they stroll through the show, and those who are interacting simultaneously with the works and with each other. Ignoring the speeders and chatters, we note the skippers and the soakers, de-
voting sustained attention to either individual works or an exhibition in toto. That sustained attention indicates an internal process, the exercise of perception, curiosity, or feeling. Stories of using art for the exercise of intelligence in self-education or internal reflection exist in abundance in literature, both nonfiction (i.e. Lawrence Weschler’s essay “Vermeer in Bosnia”), and fiction (i.e. Orhan Pamuk’s My Name is Red or Ali Singer’s How to be Both) as well as scholarship. In this short text I will simply assert that attending to art has a relationship to intelligence. I also claim that the potential for humans to access “otherness” through art—other times, other cultures, other bodies and perceptions—and to fold those experiences into a personal realm of thought, exercising intelligence to expand an individual frame of reference—is commonly accepted.

What about the socialisers, however, the people who interact simultaneously with the artworks and each other? They are exercising what psychologists call “joint attention”. As defined by philosopher Axel Seemann, joint attention is “the capacity to attend to an object together with another creature” (Seemann, 2012). This short definition hinges on the word “attend”. For two creatures to look at something at the same time is not joint attention. As psychologist Michael Tomasello writes, “A sightseer and a mountain climber attend to very different parts of a mountain (e.g. to its coloration or its slopes) in light of their very different goals.” (Seemann, 2012)

Seemann’s definition of joint attention does not limit its exercise to humans. Could AIs be among the “creatures” he includes? Might an approach from this angle reach middle ground between the strangeness of a nonhuman intelligence and our own modes of thought? Could two or more such artificial intelligences develop an ability to find meanings in images—to interpret art—by looking at and sharing information, through social exchange, just as humans do? This line of inquiry connects with the field of computational creativity, which computer scientist Ramón López de Mántares defines as “the study of building software that exhibits behavior that would be deemed creative in humans”. López de Mántares suggests that such software “acts as a creative collaborator rather than a mere tool”. Perhaps what art can do for artificial intelligence is bring it into exchange with humans around a creative goal, which has the productive ambiguity of attempting something not yet known, prompting development in both types of intelligence.

But the most widely known explorations of AI and art, such as Alexander Mordvintsev’s Deep Dream, relate to paintings from the 19th century. While they may have explored then-current questions, they no longer represent the creative edge. What do researchers miss when they ignore a century’s worth of art? Contemporary art has moved on from what Marcel Duchamp called the “retinal”. Sometimes it tackles issues of categorisation that might pose worthy questions for an AI, or an AI and a human partner, exploring the potential of jointly perceiving the world. Take the image/object in Fig. 7:

Is it a bicycle wheel, a kitchen stool, or an act of play and delight? (Duchamp, the artist, said he loved to turn the wheel and watch it, as if it were a fire.) Can it be all those things and sculpture, too? Confronting that question has confounded many human minds, making them wonder about the “cuts” they make in the world, the way their thought processes carve their experiences into objects and contexts. Could an AI make that jump? Or does their digital mode of thought, with its discrete units, put the sliding transitions of analogue processes and the overlapping of multiple meanings out of reach? Whatever the answers to these questions turn out to be, seeking them is part of what art can do for AI. But there may be a further question.

---

3. The relationship between attention and intelligence has many facets. See, for example, Karl Schweizer, Helfried Moosbrugger, and Frank Goldhammer, “The structure of the relationship between attention and intelligence”, Intelligence, Vol. 33, Issue 6, November-December 2005, pp. 589–611, in which the authors investigated twelve forms of attention and concluded that, “each type of attention was substantially related to intelligence on the latent level”. For my purposes, the point is that a person looking carefully at something may reasonably be considered to involve their intelligence.
Closure

Now I pull my loose knot of ideas and images tighter. Because you once learned to tie your shoes, because you have physical knowledge of knots, these words arrive in your mind mixed with touch and sight. This essay itself is revealed as an image, each turn of the text a loop of the knot. Grasping the start of my argument, I claimed the term “feelthink” “to name shifting relationships of perception, emotion, thought, and action” and proceeded, using images in tandem with words. Now, making the knot, tugging together the two lines of word and image, I close around Newsome’s assertion that, “[The] gesture to create something with an inherent sense of agency can be seen as a radical act of love”, tied to López de Mántares’s thought that “Rather than just seeing the computer as a tool to help human creators, we could see it as a creative entity in its own right” and a still from Orphan Drift’s “If AI Were Cephalopod…” [Fig.8].

What art can do for AI is invite it into a realm with uncertainties and surprises, and ask it to play. In this gesture, we feelthink and perform our knowledge that the world is more than human, greeting the other, as Newsome would say, with a radical act of love.

References

Atanasoski, Neda and Kalindi Vora. Surrogate Humanity: Race, Robots, and the Politics of Technological Futures. Durham: Duke University Press, 2019. https://doi.org/10.1215/9781478004455.

Bongard, Joshua. “Manipulators and Manipulanda.” Edge. 2015. https://www.edge.org/response-detail/26106.

Bryson, Joanna J. “Robots Should Be Slaves.” In Close Engagements with Artificial Companions: Key social, psychological, ethical and design issues, ed. Yorick Wilks. Amsterdam: John Benjamins, 2010. https://doi.org/10.1075/nlp.8.11bry.

Cañamero, Lola. “Emotions and Adaptation in Autonomous Agents: A Design Perspective.” Cybernetics and Systems, 32:5 (2001): 507-529. https://doi.org/10.1080/01969720120250.

Cheng, lan. http://iancheng.com/.

Damasio, Antonio. Descarte’s Error: Emotion, Reason and the Human Brain. New York: Grosset/Putnam, 1994.

Ferree, Joel. “Rashaad Newsome’s ‘Being.’” LACMA Unframed. April 9, 2020. https://unframed.lacma.org/2019/09/11/rashaad-newsome%2E80%99s-being.

Fort Mason Center. “Rashaad Newsome: To Be Real.” 2020. https://fortmason.org/event/newsome/.

Haraway, Donna. The Companion Species Manifesto: Dogs, People, and Significant Otherness. Chicago: Prickly Paradigm Press, 2003. icub.org, http://www.icub.org.

IEEE. “Your Guide To The World of Robotics.” https://robots.ieee.org/robots/icub/.

Jablouka, Eva and Marion Lamb. Evolution in Four Dimensions: Genetic, Epigenetic, Behavioral, and Symbolic Variations in the History of Life. Cambridge: MIT Press, 2005.

Lewin, Sarah. “Going Interstellar: Q&A With Author Kim Stanley Robinson.” Space.com, 2015. https://www.space.com/29862-kim-stanley-robinson-aurora-interview.html.

López de Mántares, Ramón. “Artificial Intelligence and the Arts: Toward Computational Creativity.” In The Next Step: Exponential Life. Madrid: Turner Libros, 2017. https://www.bbvaopenmind.com/en/articles/artificial-intelligence-and-the-arts-toward-computational-creativity/.

Museum of Modern Art. https://www.moma.org/collection/works/202998.

Mukherjee, Ranu and Maggie Roberts. “If AI Were Cephalopod.” Orphan Drift Archive, 2019. http://www.orphandriftarchive.com/if-ai/if-ai-were-cephalopod/.

Newsome, Rashaad. https://rashaadnewsome.com/.

Prophet, Jane and Helen Pritchard. “Performative Apparatus and Diffractive Practices: An Account of Artificial Life Art.” Artificial Life, Vol. 21, No. 3: 332-343. https://doi.org/10.1162/ARTL_a_00174.

Seeman, Axel, ed., New Developments in Psychology, Philosophy of Mind, and Social Neuroscience. Cambridge: MIT Press, 2012.

Smail, Daniel Lord. On Deep History and the Brain. Berkeley: University of California Press, 2007.

So, Veronica, Marc Palatucci, and Jesper D. Lund. “The Worldmaker and His Killjoy.” Office Magazine, August 8, 2017. http://office-magazine.net/interview/worldmaker-and-his-killjoy.

Solso, Robert. Cognition and the Visual Arts. Cambridge MA: MIT Press, 1994.

Wikiart. https://www.wikiart.org/en/marcel-duchamp/bicycle-wheel-1913.

Wikimedia. https://commons.wikimedia.org/wiki/File:Fragonard,_The_Swing.jpg.
CV

Meredith Tromble
San Francisco Art Institute
sfai.mt@gmail.com

Meredith Tromble is an intermedia artist and writer. Her curiosity about the links between imagination and knowledge led her to collaborate with scientists in addition to making installations, drawings, and performances. Her work has been presented nationally and internationally, including a survey of her artwork from the past decade in the exhibition “Umwelt” at BioBAT Art Space in Brooklyn. She holds joint appointments as artist-in-residence at the Complexity Sciences Center and Visiting Scholar at the Feminist Research Institute at the University of California, Davis. She is the editor of two books, The Routledge Companion to Biology in Art and Architecture, co-edited with Charissa Terranova, and The Art and Films of Lynn Hershman, University of California Press. She is Professor of Interdisciplinary Studies/Art & Technology at the San Francisco Art Institute.
**Outside in: exile at home. An algorithmic discrimination system**

Annabel Castro
Universidad Autónoma Metropolitana

Date of submission: March 2020
Accepted in: June 2020
Published in: July 2020

**Abstract**

Outside-in is an installation that utilises machine learning to reflect on systematic discrimination by focusing on the indefinite detention of Mexicans with Japanese heritage concentrated in Morelos during WWII. This algorithmic discrimination system tears apart four classic fiction films continuously within a projection room. The fragments are displaced and classified using machine learning algorithms. The system selects, separates and reassembles the fragments into new orders. It evokes the condition of being robbed of your right to be in the place to which you belong. The citizens detained during WWII were removed from their residence, their belongings were confiscated and they were placed in seclusion solely for having Japanese ancestry. Similarly, at present, data retrieving companies configure low resolution representations of ourselves from the snatched digital debris of our daily life. These pieces are reconfigured into archetypes and meaning is attached to them for massive decision making. We don’t have the right or means to know what these representations look like or what meaning has been attached to such shapes. It is a privilege reserved to the designers of algorithmic processes: they own this right and we the citizens own the consequences. The present article is a case study presenting the creation of Outside in: exile at home as an installation that utilises machine learning and reflects on this kind of systematic discrimination.
Outside in: exile at home. Un sistema de discriminación algorítmica

Resumen
Outside in es una instalación que utiliza el aprendizaje automático para reflexionar sobre la discriminación sistemática al centrarse en la detención indiscriminada de mexicanos con herencia japonesa concentrada en Morelos durante la Segunda Guerra Mundial. Este sistema de discriminación algorítmica "despedaza" continuamente cuatro películas clásicas de ficción en una sala de proyección. Los fragmentos se desplazan y clasifican utilizando algoritmos de aprendizaje automático. El sistema selecciona, separa y vuelve a montar los fragmentos en un nuevo orden. Quiere reflejar la sensación de ser despojados de nuestro derecho a estar en el lugar que nos pertenece. Los ciudadanos detenidos durante la Segunda Guerra Mundial fueron sacados de sus casas, sus pertenencias fueron confiscadas y se les recluyó únicamente por tener ascendencia japonesa. Y ahora, un poco de la misma manera, las empresas de recuperación de datos realizan representaciones en baja resolución de nosotros mismos partiendo de los escombros digitales que se nos arrebata de nuestra vida cotidiana. Estas piezas se reconfiguran en arquetipos y se les asigna un significado para la toma de decisiones masivas. No tenemos el derecho o los medios para saber cómo son estas representaciones o qué significado se ha asignado a tales formas. Es un privilegio reservado a los diseñadores de procesos algorítmicos: ellos son dueños de este derecho y nosotros, los ciudadanos, somos dueños de las consecuencias. El presente artículo es un estudio de caso que presenta la creación de Outside in: exile at home como una instalación que utiliza el aprendizaje automático y reflexiona sobre este tipo de discriminación sistemática.

Palabras clave
aprendizaje automático, sistemas de decisión, edición de video algorítmica, ciudadanía mexicana durante la Segunda Guerra Mundial, identidad, territorio

Introduction
You lied when you said you didn’t mind my face.
Mr. Okuyama in the 1966 film The Face of Another

Outside-in is an installation that utilises machine learning to reflect on systematic discrimination by focusing on the indefinite detention of Mexican citizens with Japanese heritage during WWII. It was premiered in 2018 at the Jardín Borda Cultural Center in the capital of Morelos, the same state where the Hacienda de Temixco former internment camp was located. Visitors enter a dark room that hosts an algorithmic discrimination system which forces four classic films into the space and time usually assigned to one. The system tears these films apart, displaces the fragments and recombines them. It discards the centre of the images projecting only the outskirts, all of them layered on a single wall. It displaces the subtitles of all to a perpendicular wall making it impossible to read them as you watch the images. On a third wall the system channels the soundtracks through headphones, switching among films. In this new disposition the elements of a single film share the same dark room, but they no
longer belong together. It is almost impossible to know which film they are a part of; they have lost their original purpose: conveying a specific meaning.

The work evokes the process of indefinite detention; it focuses on the case of the Hacienda de Temixco, a concentration camp for Mexicans with Japanese heritage between 1942 and 1945. One of the four classic films that constitute the material of the installation is Mexican and the other three are Japanese productions. The Mexican one is shot in rural Oaxaca starring Toshiro Mifune. He plays the role of an indigenous man striving to be the patron of a religious festivity. The remaining three productions are about a man’s relation to his disfigured face, a doctor (played by Mifune) who contracted syphilis in WWII and a femme fatale infatuated by the son of her lover, played by famous drag queen singer Akihiro Miwa (Maruyama). These films contain reflections on belonging, identity and social exclusion. The discrimination system that constitutes the installation uses machine learning to classify and fragment the films that are presented in new sequences, displaced as previously described.

The historical and geographical context of the artwork will be presented first, followed by an examination of the system that structures the work and its use of machine learning.

**Historical and Geographical context**

Goodbye, then. I won’t take you home.

Dr. Kyoji Fujisaki in the 1949 film *The Quiet Duel*

Facial resignification

I decided to exhibit the installation at Jardín Borda Cultural Center because it is located in the capital of Morelos, a block away from the main cathedral. It is a central location in the state that hosted the internment camp for citizens with Japanese heritage, the focus of the artwork. Morelos’ territory has housed the vacation getaway properties of the ruling classes of the Mexica, Spaniards and several Mexican presidents. Jardín Borda is well known for having been a home for former Emperor Maximilian and Carlota. They were the first rulers of Mexico that massively propagated their image thanks to photography. It was during their rule that the exploitation of photography in Mexico as propaganda media began.

When the couple disembarked in Vera-cruz in May 1864, their faces were already popular all over the country because their images had previously been commercialised (Aguilar 1996, 16). Meaning had been carefully assigned to their faces through the elite innovative technologies of that time. During the next century, propaganda campaigns were also used to resignify faces as is the case with Japanese heritage citizens after Pearl Harbor. This time not to empower, as with Emperor Maximilian, but to render as vulnerable as possible. “To achieve complete hemispheric solidarity and defense the United States conducted a sweeping propaganda campaign that reached all levels of Mexican society, while simultaneously pushing negative images of Japan and Japanese residing in Mexico as possible saboteurs, sleeper agents, or simply not trustworthy based on their ethnicity.” (García 2014, 189) “…the Mexican government demonized Japanese Mexicans and suspended the civil rights of all citizens, claiming a state of emergency.” (Chew 2015, 176).

**Indefinite detention**

In a news conference on December 15, 1941, Frank Knox (then secretary of the Navy) put the blame related to the Pearl Harbor attack on Japanese Americans living in Hawaii. He provided no evidence for his claim, then or ever (Barbash 2016). This false assumption helped to fuel a mass media fabrication of meaning embedded in the faces of citizens with Japanese heritage all over the American continent. Upon pressure from its northern neighbour, in December, the Mexican government froze the bank accounts of the Japanese heritage people and in the following January removed them from the borderlands by ordering them to concentrate within the cities of Guadalajara and Mexico City, the centre of our territory. The Japanese heritage people already living in these cities could not lodge all the displaced people who could not afford a place of their own; therefore, they formed the Committee of Mutual Help (CMH) and rented an old large house for this purpose. By June the government created a department to confiscate and administer the possessions of all Japanese heritage people. The old house was already insufficient so the CMH, with government authorisation, bought the Hacienda of Temixco in Morelos, a state adjacent to Mexico City. The first transferred families arrived at the hacienda in August. Two government soldiers guarded the entrance permanently until 1945 when the indefinite detention ceased (Hernández 2011, 146). Upon false assumptions it was determined that one single data point (ancestry) was enough to label a citizen as a potential threat for the country. Regarding citizens within the US and its neighbouring countries, if the citizen had Japanese ancestry, we can say the below procedure was followed:

1. Classify the citizen as a potential danger to the USA.
2. Freeze the citizen’s bank account.
3. Confiscate her/his belongings.
4. Displace the citizen to a designated zone or facility where she/he will live under specific restrictions and under surveillance for an indefinite number of years.

Does this citizen deserve all these consequences? That decision was made upon obtaining a single binary data point: does he/she have Japanese ancestry? If yes, follow the above listed procedure. This low-resolution decision method based on ancestry is still used to decide whether a US citizen qualifies as a suspect of being a threat to national safety.

Section 1021 of the 2012 National Defense Authorization Act (NDAA) gave a legal frame for indefinite detentions and is still in force in the 2020 NDAA. As Christopher Jenks mentions in his article *Civil Liberties and the Indefinite Detention of U.S. Citizens*: “Our grandparents’ generation, which detained thousands of Japanese-Americans,
and our parents’ generation, which overreacted to fears of communist infiltration, show the dangers of waiting [to correct section 1021 of the NDAA]. America as a country and as a people should have learned from these mistakes. These generations, however, were largely good people who as a society made regrettable decisions out of fear. We are in the midst of making decisions, or not opposing decisions made in our name, which will have equally untoward outcomes. The components which yield this undesirable, undemocratic outcome are right in front of us. A federal statute which authorizes indefinite detention of we are not exactly sure who, or for exactly what conduct, juxtaposed against what even members of the judiciary label a functionally useless habeas corpus regime. The components are right in front of us. They will ultimately be brought into sharp focus in the aftermath of a domestic terrorism incident in which there is not overwhelming evidence of guilt such that a traditional prosecution is possible. There the government, our government, may take the next step, one provided for by section 1021, to indefinitely detain U.S. citizens captured in the United States for conduct that allegedly occurred in the United States.” (Jenks 2014, 36-37)

The New York Times reports on January 5th, 2020: “Dozens of Iranians and Iranian-Americans were held for hours at Washington State’s border with Canada over the weekend as the Department of Homeland Security ramped up security at border ports after Iran threatened to retaliate against the United States for the strike that killed its top military leader.” (Kanno-Youngs et al. 2020). Again a binary decision based on ancestry dissolves citizenship, in this case for hours but the legal frame is there to make it last an indefinite amount of years.

The installation’s structuring system and use of machine learning

Damn! He has the same impish face.

El Español in the 1961 film Anímas Trujano (The Important Man)

Source material
The artwork’s system dismantles four classic films within the exhibition room by tearing them into data units and reconfiguring their position in time and space. The source materials selected to inhabit the system are described below together with a brief sample of their subtitles:

Anímas Trujano: The Important Man (1962, B&W), directed by Ismael Rodríguez based on a novel by Rogelio Barriga. Iconic Japanese actor Toshiro Mifune embodies Anímas Trujano, an indigenous man in Oaxaca striving to be patron of a religious festivity.

Don’t run, evil woman. Stay there! (00:14:37)
Are you spying on me? (00:41:22)
What does that mean, Mr. Justice? (00:44:57)

Funny you should forget the most important thing. (00:46:21)
So very white. (00:50:00)

The Quiet Duel (1949, B&W), directed by Akira Kurosawa, based on the play The Abortion Doctor by Kazuo Kikuta. Mifune here plays the role of Dr. Kyoji Fujisaki, a doctor that contracted syphilis during World War II while performing surgery on a patient infected with this disease.

I don’t go to your place anymore. (00:24:36)
I don’t understand, the politics of the country. (00:37:25)
It’s a terrible story. My colleague was inspecting cars… (00:47:57)
my body became dirty, without knowing any pleasure. (01:15:45)
You have to walk proudly, with your chin up high. (01:31:50)

The Face of Another (1966, B&W), directed by Hiroshi Teshigahara, based on the novel by Kōbō Abe. Here the main character, an engineer, has a severely disfigured face due to an acid-related accident in his laboratory and acquires a life-like mask he uses to seduce his wife incognito.

No one even remembers it happened. (00:03:03)
Why do people feel the way they do about faces? (00:05:34)
The world outside seems, so terribly far away. (00:56:48)
Who’s killing whom? (01:38:48)

Freedom’s always, a lonely thing. (01:58:47)

Black Rose Mansion (1969, colour), directed by Kinji Fukasaku. This film stars famous drag queen singer Akihiro Miwa (Maruyama) who plays a femme fatale infatuated with the son of her lover. Miwa is also known as a peace advocator and survivor of the atomic bomb dropped in Nagasaki in 1945.

…when I stopped and stood still…that’s when it happened. (00:00:49)
We don’t know anything about, Ryuko Fujio, her past, or her present. (00:07:07)
… but such was my determination (00:42:44)
She’s from outside our world. (01:23:29)
Did they find them? (01:27:21)

I chose these specific films because all of them are fiction. We could say characters in a way are a bundle of features that register an essential idea of something, like machine learning models. These films can be classified in ways where one of them is singled out (filmed or not in the ‘60s, Japan, B&W, based on known literary works). They contain reflections on belonging, identity and social exclusion which relate to the context of indefinite detention due to heritage. In the installation Mifune becomes a sort of repository for the historic event. As Anímas Trujano he is a Mexican character among Japanese movies, a Mexican-looking Japanese citizen. As Dr. Kyoji Fujisaki he becomes the Japanese doctor of Temixco. In 1917, Mexico signed a bilateral agreement accepting Japanese medical doctors, dentists, pharmacists, obstetricians, and veterinarians to practice in Mexico (Chew 2015, 39). Several of them were later displaced and detained. Dr. Manuel Seichi Hiromoto spent several years in the Temixco con-
centration camp. At the end of WWII, he became a revered medical practitioner in the town of Temixco (Chew 2015, 143), remembered as the Japanese doctor. Three of his sons still continue with the medical practice in the town.

Dismantle
Machine learning was used to analyse the data that constitutes the sources to classify it into distinct categories including the presence or absence of a female voice in the soundtracks, female faces in the moving images, and statements projecting positive or negative sentiment in the subtitles. The machine learning models used to classify the source material were developed using the open source toolset Turi Create.

The text classifier model used to analyse the sentiment of the subtitles uses two components: a bag-of-words and a logistic regression classifier. It was used to determine the probability of a subtitle for referring positively or negatively about a topic (for example about to die, Tadeo, a sunset).

To evaluate whether a face present in an image belongs to a female, an image similarity model was used. From fragments of films containing a female face I extracted image sequences to use as data references for a model. The model retrieves the images that are most similar to the reference from the installations source films’ image sequences with a face.

A sound classifier model is used to assign sounds a category, such as: music, siren, dog barking, baby crying or a specific female voice. For this last class 40 five-second audio recordings were extracted from the film The Pearl (1947) by Emilio Fernández. The samples contain the voice of the actress María Elena Marqués, speaking in English, as Juana. The classifier determines which sounds from the four film sources should be considered Juana’s voice.

Displace and reassemble
As mentioned previously in the introduction, the system discards the centre of the images projecting only the outskirts, all of them layered on a single wall. Subtitles are displaced to a perpendicular wall making it impossible to read them as you watch the images. This wall can be divided into four areas; each one displays subtitles belonging to a single movie. On a third wall the system channels the soundtracks through headphones, switching among films. The installation output is assembled in the visual programming language TouchDesigner. The programme written in this platform processes the source materials and splits them into the three output signals (image, subtitles and soundtrack). It outputs each signal interspersing between 2 modes.

1- All sources in a same signal are synchronized. For example, subtitles are all at 00:07:06
   Film 1: Anybody would want to die young like him…
   Film 2: [no subtitles at this point in film 2, the space would be empty]
   Film 3: I live in a much greater darkness.
   Film 4: …accompanied by a young mulatto

2- A signal displays elements from the sources that belong to the same category. That is the category decided through the machine learning processes.
   For example, subtitles classified as positive about a topic
   Film 1: Anybody would want to die young like him…
   Film 2: Poor people should die early for their own good.
   Film 3: [no subtitles in film 3 in this category, the space would be empty]
   Film 4: He lost sight of love… he was dead inside

Mode 1 allows visitors to follow a film through its subtitles, or soundtrack, or the camera movements in the images. Mode 2 doesn’t, materials are totally disarticulated from their sequence. Each mode lasts for a variation of minutes; it is meant to contrast, to make evident the reassembling of data when it happens.

On systematic discrimination
We should resist all discourses presenting (and practicing) the state as a camp, and all forms of normalization of the exception, that appears always, somehow, to pertain to others-from-us” (Minca 2015, 81)

The artwork exposes the audience to a familiar experience that has been disturbed: watching a classic film. It is evident most of the image is missing and that there is way too much information to follow and process. At times all of the content in the subtitles seems
related, certain content is present much more frequently. These cues let the audience notice there is a system filtering and organizing what is presented to them. If the audience lives in Morelos, they have probably been to the privately run water park the Hacienda hosts at present, or, at least, seen pictures of it. The label on the wall mentioning the historic event will reshape how they perceive all the data elements of the installation. These elements belonging to the four films are now intertwined; they have lost their former context that triggered how you read them. At the same time, they are no longer subjugated by the movie’s central purpose, but can still signify. From the sum of displaced elements: meaning emerges now quelled by a new context.

The installation system is structured to evoke the spatial biopolitical technology (Katz, et al. 2018, 4) that constitutes the normalisation of indefinite detention by disturbing the experience of watching a classic film. It evokes the specific case of the Hacienda de Temixco through the use of its film sources as explained previously. Machine learning is used to further transform our perception of the films’ traces presented to us in the installation. The use of these processes also has a symbolic value since these tools are being used in massive decision-making processes. “Currently, decision-making systems that rely on machine learning are in many ways opaque, lacking adequate explainability in many instances” (Lohr 2019, 226). “Machine learning algorithms have the potential to discriminate more consistently and systematically and at a larger scale than traditional non-digital discriminatory practices” (Criado and Such 2019, 85). Aided by machine learning, data retrieving companies configure low-resolution representations of ourselves from the snatched digital debris of our daily life. These pieces are reconfigured into archetypes and meaning is attached to them for massive decision making. As in the case of Emperor Maximilian and Carlota, these images have been propagated massively and arrive before us to loan granters, insurance companies and legal systems among others. We don’t have the right or means to know what these representations of ourselves look like or what meaning has been attached to such shapes. It is a privilege reserved to the designers of algorithmic processes: they own this right and we the citizens own the consequences of our by-product face, as in the case of the Mexicans with Japanese heritage interned at the Hacienda de Temixco in Morelos, Mexico.

What are the shapes of our invisible-to-us data representations? Portraiture and subjectivity in legal data sculpting are concepts I explore in my artwork aided by machine learning processes, which is currently in development.

References

Aguilar, Arturo. La fotografía durante el imperio de Maximiliano. Mexico City: Instituto de Investigaciones Estéticas-UNAM, 1996.

Barbash, Fred. “After Pearl Harbor: How a government official’s baseless claim helped lead to the internment of Japanese Americans,” The Washington Post, December 12, 2016. https://www.washingtonpost.com/news/morning-mix/wp/2016/12/12/the-week-after-pearl-harbor-a-government-officials-ignominious-big-lie-and-the-internment-of-japanese-americans/.

Chew, Selfa. Uprooting Community: Japanese Mexicans, World War II, and the U.S.-Mexico Borderlands. Tucson: University of Arizona Press, 2015. https://doi.org/10.2307/j.ctt183p99h.

Crane, Nicholas et al. “Questioning the exceptionality of the exception: Annabel Castro’s ‘Outside in: exile at home’ (2018) in Cuernavaca.” Cultural Geographies (2020) https://doi.org/10.1177/1474474020931525.

Criado, Natalia and Jose M. Such. “Digital Discrimination.” In Algo- rithmic Regulation, edited by Karen Yeung & Martin Lodge. UK: Oxford University Press, 2019.

García, Jerry. Looking Like the Enemy: Japanese Mexicans, the Mexican State, and the US Hegemony, 1897-1945. Tucson: University of Arizona Press, 2014.

Hernández, Sergio. La guerra contra los japoneses en México durante la segunda guerra mundial: Kiso Tsuru y Masao Imuro, migrantes vigilados. Mexico City: Editorial Itaca, 2011.

Jenks, Christopher. “Civil Liberties and the Indefinite Detention of U.S. Citizens.” Harvard Journal of Law & Public Policy, Federalist Edition 2, no. 1 (2014).

Kanno-Youngs, Zolan et al. “U.S. Stops Dozens of Iranian-Americans Returning From Canada.” The New York Times, January 5, 2020. https://ssrn.com/abstract=2546385Kanno-Youngs, Zolan et al. “U.S. Stops Dozens of Iranian-Americans Returning From Canada.” The New York Times, January 5, 2020. https://www.nytimes.com/2020/01/05/us/politics/iranian-americans-border.html.

Katz, Irit et al. “The camp reconsidered.” In Camps Revisited: Multifaceted Spatialities of a Modern Political Technology, edited by Irit Katz. London: Rowman & Littlefield International Ltd, 2018.

Lohr, Jason D et al. “Legal Practitioners’ Approach to Regulating AI Risks.” In Algorithmic Regulation, edited by Karen Yeung & Martin Lodge. UK: Oxford University Press, 2019.

Minca, Claudio. “Geographies of the Camp.” Political Geography 49 (November 2015). https://doi.org/10.1016/j.polgeo.2014.12.005.

National Defense Authorization Act for fiscal year 2020. https://docs.house.gov/billsthisweek/20191209/CRPT-116hrpt333.pdf.
Annabel Castro
Universidad Autónoma Metropolitana
a.castro@correo.ler.uam.mx

Annabel Castro is a digital media artist. The overlapping of geographic, virtual and mental spaces, landscape as trajectory, identity, are recurring subjects in her work. She has received the PBEE grant from the National Foundation for Arts and Culture of Mexico, the La Paz award from the International Art Biennial SIART in Bolivia, Object Art first prize from the Third National Biennale Puebla de los Ángeles in Mexico and the Open Territories Commission of mARTadero Culture Center in Bolivia. Her artwork has been shown among other venues at the Transmediale in Berlin and the Ethnography and Folklore Museum of Bolivia. She studied her PhD at the DXARTS department at University of Washington. She is currently professor of the Arts and Humanities department of Universidad Autónoma Metropolitana (Lerma campus).
Identificación, clasificación y control: estrechos vínculos analizados desde las prácticas artísticas en el corazón de la inteligencia artificial

Hugo Felipe Idárraga
Universidad de Los Andes

Fecha de presentación: marzo de 2020
Fecha de aceptación: junio de 2020
Fecha de publicación: julio de 2020

Cita recomendada
Idárraga, Hugo F. 2020. «Identificación, clasificación y control: estrechos vínculos analizados desde las prácticas artísticas en el corazón de la inteligencia artificial». In: Andrés Burbano; Ruth West (coord.) «IA, arte y diseño: Cuestionando el aprendizaje automático». Artnodes, núm. 26: xx-xx. UOC. [Fecha de consulta: dd/mm/aa]. http://doi.org/10.7238/a.v0i26.3361

Resumen
La inteligencia artificial (IA) está antecedida de una historia que se remonta a antiguos esfuerzos por crear seres con movimiento e inteligencia artificiales. El objetivo de este artículo es destacar cómo en esta historia, que incluye los desarrollos actuales en el campo del machine learning (ML) y de las redes neuronales profundas (RNP), las tareas de vigilancia y control por medio de la identificación y clasificación de personas, cosas o eventos del mundo, han sido medulares tanto en los mitos o las teorías sobre autómatas, homúnculos, androides, robots o ciborgs, así como en los diferentes intentos por materializarlos. Se afirmará, además, que los deseos y esfuerzos por imaginar y crear estos seres no son ni inocentes ni circunstanciales; más bien, provienen en gran medida de una visión patriarcal del mundo en la que todo lo existente debe ser sometido a una vigilancia que asegure un control sobre lo real. Para ello se recurrirá,
por un lado, a historias que evidencien el predominio de estas tareas y la visión del mundo que detrás de ellas se esconde, y, por el otro, a prácticas artísticas y representaciones estéticas que han cuestionado el funcionamiento identificatorio y clasificatorio de la IA como medio de vigilancia y control.

**Palabras clave**
Inteligencia artificial, arte, identificación, clasificación, vigilancia, control, adversarial examples.

**Abstract**
Artificial intelligence (AI) is preceded by a history going all the way back to former efforts in creating beings with artificial movement and intelligence. The objective of this article is to highlight how, in this history, which includes the current developments in the fields of Machine Learning (ML) and Deep Neural Networks (DNNs), the tasks of surveillance and control through the identification and classification of people, things or events in the world, have been so central in the myths or theories on automatons, homunculi, androids, robots or cyborgs, as well as in the various attempts to make them a reality. It will be asserted, furthermore, that the wishes and efforts to imagine and create these beings are neither innocent nor circumstantial; rather, that to a large extent they come from a patriarchal vision of the world in which everything existing must be subjected to a surveillance that ensures control over that which is real. To achieve this, reference will be made on the one hand to histories which show the predominance of these tasks and the vision of the world hidden behind them, and, on the other, to artistic practices and aesthetic representations which have questioned the identifying and classifying operations of AI as a means of surveillance and control.

**Keywords**
artificial intelligence, art, identification, classification, surveillance, control, adversarial examples

---

### Introducción
El arte, desde sus más remotos manifestaciones, ha señalado una infinidad de lugares desde donde despejar la realidad de su ilegítima normalidad. La crítica artística de los fenómenos técnicos hace parte de la construcción de estos lugares. Actualmente, el fenómeno de la IA, globalizado y cotidiano, debe enfrentarse a las manifestaciones estéticas que cuestionan críticamente su funcionamiento, apuntando a sus fundamentos históricos y teóricos, con la intención de desarracógar su normalización elogiosa por parte de corporaciones y gobiernos que impulsan su desarrollo y aprovechan sus beneficios. Mientras que una revisión histórica de estos desarrollos permite identificar algunas ideas recurrentes, fijas en la mirada de una sociedad patriarcal que necesita de unos y otros como propiedades que deben ser organizadas y controladas, las prácticas artísticas desencajan la normalidad con la que estas ideas han sido abrigadas «permitiéndonos imaginar diferentes posibilidades y diferentes futuros de cómo podemos ser empleando estas tecnologías»1 (ỌnỌỌha, 2018).

**Identificación y clasificación: relaciones con la vigilancia y el control**
Los tempranos intentos en la creación de seres artificiales se concentraron en automatas con movimientos mecánicamente autopropulsados. La *Iliada* presenta a Hefesto como constructor de una diversidad de autómatas (Homero, 2010). Entre ellos, Hefesto construyó a Talos, «el primer robot», hecho de bronce y «cuyo deber es patrullar tres veces al día las playas de Creta» (McCorduck, 2004: 5). En el año de 1966, el robot *Shakey*, desarrollado por la empresa Panoramic, una de las principales contratistas de la CIA y activa participante en el Project MK-Ultra, fue considerado «el primer robot del mundo con inteli-
cic artificial» (Raviv, 2020). Allí, el profesor Woodrow Wilson Bledsoe desarrolló el primer sistema de reconocimiento facial, partiendo de las teorías de Alphonse Bertillon. Su objetivo: «ayudar a las agencias de seguridad a examinar rápidamente las bases de datos de fotografías y retratos, buscando coincidencias» (Raviv, 2020). Igual que Talos, el reconocimiento facial es empleado con propósitos de vigilancia, identificando y clasificando para permitir o denegar el acceso a lugares físicos o virtuales. El sistema de Bledsoe fue entrenado exclusivamente con fotos de personas caucásicas, excluyendo las de otras razas. El artista Trevor Paglen y la investigadora Kate Crawford, en su exposición Training Humans de 2019, desentrañan los prejuicios raciales y morales insertos en los set de datos, analizando y denunciando las imágenes empleadas para entrenar los sistemas de identificación y clasificación, así como las teorías supuestamente científicas que los legitiman (Paglen y Crawford, 2019).

«Las computadoras de HP son racistas» (Bunz, 2009). Joy Buolamwini fue una de las primeras en denunciar estos prejuicios raciales en los sistemas de IA. Sus proyectos Gender Shades y AI, Ain’t I a Woman?, ambos del año 2018, exploran la parcialidad de los sistemas de reconocimiento facial que, por ejemplo, causan una mayor persecución policial sobre la población afro de EE.UU. o discriminan negativamente a las mujeres. Mimi Onuoha ha criticado los conjuntos de datos empleados en estos sistemas. En su obra escultórica Classification.01, de 2017, dos paréntesis de neón se iluminan cuando el espectador cumple con las medidas de identificación algorítmica determinadas de antemano por el artista, en una especie de emparejamiento maquinco coherente con los prejuicios de estos sistemas.

Durante la Alta Edad Media, la metáfora representada en la natura artifex evidencia la importancia del neoplatonismo como fundamento para comprender la creación de autómatas. Los autómatas son pensados así como una creación que implicaba tanto la habilidad humana como el conocimiento de las formas ideales por medio de la acción de la naturaleza. Estas creaciones son misteriosas y exigen no solo un conocimiento artesanal, sino también el entendimiento al que únicamente filósofos o hechiceros podían acceder, como aquel que aparece en los romans antiques Le Roman d’Éneas (ca. 1160) y Le Roman de Troie (ca. 1165).

En Le Roman de Troie, Benoît de Sainte-Maure reconstruye la Guerra de Troya. Allí se describe la aloca convencional de Héctor. De entre sus diferentes maravillas destacan dos autómatas. El primero interpreta instrumentos musicales y esperece flores por la habitación, determinando lo estéticamente correcto. El segundo carga un incensario que cura las enfermedades y el dolor a quien aspira su aroma, y su tarea consiste en identificar y corregir todo comportamiento descortés, decidir sobre lo moralmente adecuado e inadecuado. En Le Roman d’Alexandre (ca. 1180), dos hombres de oro vigilan un puente amenzado por el ejército enemigo. Asimismo, en la continuación de Perceval, dos figuras de oro y plata resguardan la tienda de Alardín, caballero venido de tierras distantes. De estas figuras, una está encargada de identificar y evitar el ingreso de villanos, mientras que la segunda posee un arpa que tañe de manera discordante cuando una doncella que ha perdido su virginidad desea acceder a la tienda. Los autómatas, así, «realizan una función disciplinaria. Su capacidad para vigilar, proteger y castigar a las personas es tan asombrosa y maravillosa como su valor estético o el misterio de su creación» (Truitt, 2015: 60).

No obstante, la vigilancia y el control llevados a cabo por estos seres y sistemas inteligentes no han sido desligados sin resistencia. El artista Adam Harvey, en su proyecto CVDazzle, del año 2010, intenta evadir el control de la visión algorítmica por medio de audaces maquillajes no convencionales. Asimismo, el proyecto de Sterling Crispin, Data-Mask, de 2013, aprovecha el funcionamiento de los algoritmos de reconocimiento facial en una estrategia de ingeniería inversa para engañar a la máquina con caras irreconocibles para el ser humano y, sin embargo, legibles por el algoritmo, impidiendo cualquier tipo de identificación efectiva.

Por otro lado, en general la creación de autómatas buscó delegar ciertas tareas humanas que realizaban comúnmente siervos y esclavos. Para el estagirita, «(…) si los peines por sí mismos tejiesen, y la pluma por sí misma tocase la cítara, ni los oficiales tendrían necesidad de ministros, ni los señores de siervos» (Aristóteles, 2004: 20). De esta forma, los autómatas se vinculan con la servidumbre, vigilada por el señor de la casa, o del gobierno, pero también con la vigilancia y el control social, en esa patria potestad recordada por Platón (2011) que solo podía ser ejercida por los hombres-ciudadanos. Por eso, para Truitt, los autómatas «no solo imponen límites físicos, sino también sociales» (2015: 62).

Homúnculos, robots, magia y poder

Pigmálion, rey de Chipre, surge como otro vector de fantasías relacionadas con un mundo controlado y controlable desde la hegemónica visión patriarcal del universo. Desesperado por no encontrar a una mujer perfecta, decide diseñar una que le agrade plenamente. Gala-tea, el amor de Pigmalión, fue humanizada por la diosa Afrodita. De allí se originan diferentes mitologías sobre el homúnculo femenino, que es una creación principalmente masculina.

Teofrasto Paracelso fue uno de los primeros alquimistas en reivindicar la creación de homúnculos, sugiriendo su crecimiento por medios no convencionales. Asimismo, el continuador de Perceval, dos figuras de oro y plata resguardan la tienda de Alardín, caballero venido de
McCarthy según la cual las diferencias entre el ser humano y la máquina son «ilusiones». La comparación entre IA y magia, que fue la base de Dreyfus para criticar este excesivo optimismo, le sirvió también de excusa para revelar cómo, a través de la eficacia de la IA –al igual que la alquimia en sus inicios–, se escuda una confianza que socava los esfuerzos por analizar científicamente los más crípticos y oscuros procedimientos que ocurren en el interior de estos sistemas, legitimando el poder de quienes los desarrollan e implementan. Asimismo Campolo y Crawford (2020) afirman que vivimos un resurgimiento del *techno*-optimismo mágico, parecido al optimismo alquímico en sus inicios. Por ejemplo, según Stuart J. Russell, «recientemente estamos comenzando a obtener una comprensión teórica de cuándo y por qué la hipótesis de aprendizaje profundo es correcta, pero en gran medida sigue siendo una especie de magia» (Ford, 2018: 42, citado en Campolo y Crawford, 2020: 2).

¿Cómo se vinculan la clasificación, la magia y el poder político que de ahí se deriva? Según Campolo y Crawford, la IA, que nace en un campo asociado al desencantamiento del mundo por medio de una explicación racional de este, se ha refugiado, sin embargo, en oscuras explicaciones por medio de un lenguaje reencantado. Este reencantamiento refuerza el poder de aquellos que diseñan, comercializan y despliegan estos sistemas, al tiempo que disminuye el de la vasta mayoría de personas que son identificadas, clasificadas y previsualizadas. «Es una forma de poder sin conocimiento», afirman Campolo y Crawford (2020: 5).

Trevor Paglen intenta develar los misterios de los proyectos del complejo industrial-militar de EE.UU en eso que él mismo ha llamado como el *Black World*. En su libro *Blank Spots on the Map*, de 2009, o en su proyecto *Unmarked Planes and Hidden Geographies*, de 2006, Paglen trata de arrojar luz sobre aquellos misterios que, con mucha eficacia, han sido escondidos a los ojos del público en proyectos destinados a la vigilancia y el control social. Sin embargo, es el artista Mario Klingemann quien, sin recurrir a explicaciones mistificadas, experimenta con las RNP y los resultados inesperados que ellas producen, intentando entrever estéticamente lo que ocurre en el interior de ellas, en una técnica que él mismo ha denominado como *Neural Glitch*. Esta serie de obras de 2018 son el resultado de un intento por emancipar los límites de la percepción tanto humana como algorítmica, indagando en aquellas propiedades intrigantes de la IA, provocando que los modelos «malinterpreten los datos de entrada en formas interesantes, algunas de las cuales pueden ser explicadas como atisbos de creatividad autónoma» (Klingemann, 2018).

**El control, fundamento de la visión tecnológica patriarcal**

Bajo la figura del homúnculo también se han pensado seres artificiales con vida y pensamiento propio. Cuarenta años después de la muerte de Paraceleso surge la figura de Joseph Golem, una creación del rabino Judah ben Loew por medio de artículos mágicos. Además de prestar servicios de vigilancia alrededor del templo, el Golem fue empleado como esclavo. Estas tareas de carácter servil sembraron en él la idea de superar a su creador, rebelándose frente a su poder, en una temática que se repetirá no solo en la modernidad literaria europea, sino también en múltiples representaciones contemporáneas, como en *Odissea en el espacio* (1968), *Blade Runner* (1982), *Terminator* (1984), *Matrix* (1999) o *Yo, Robot* (2004), en las que la IA se subleva frente a su amo. En el fondo, como veremos más adelante, este miedo de la máquina rebelada es también el miedo al trabajador rebelado, a la mujer rebelada, a la sociedad entera rebelada contra sus amos.

De entre los personajes modernos que han encarnado el espíritu de este mito está el mismo Frankenstein. Como subraya Ruha Benjamin respecto al carácter servil de Frankenstein en el contexto de la etimología de los robots, estos «transmiten una agitación continua sobre la dominación humana sobre otros humanos. (...) La dominación social caracterizó el laboratorio cultural en el que los robots fueron originalmente imaginados. Y, técnicamente, las personas fueron los primeros robots» (Benjamin, 2019: 55). Y también los últimos. Hoy en día, una de las consignas de los trabajadores en empresas ultratecnificadas es «no somos robots. Para satisfacer a la máquina, los trabajadores se sienten forzados a convertirse ellos mismos en máquinas» (Dzieza, 2020).

El pensamiento del amo y el esclavo, la idea de sirvientes y vigilantes como formas que aseguran el orden por medio de la identificación y clasificación, está encarnado en estos y los posteriores seres artificiales, así como en los proyectos subsiguientes dedicados a la IA. Es precisamente a ello, como parte de esta herencia histórica masculina del amo-esclavo, a lo que apunta Anna Everett en su incómoda relación computacional: «Al iniciar mi computadora me confronto con el texto basado en DOS que me da una pausa… “Pri. Master Disk, Pri. Slave Disk, Sec. Master, Sec. Slave”». (Everett, 2002, citada en Benjamin, 2019: 55).

El mito de Pigmalión se conecta con el argumento de la famosa serie *Ghost in the Shell* (1995) a través de las muñecas de Hans Bellmer, empleadas como modelos visuales de las androides de Hanka Robotic. Bellmer creó desde 1933 mujeres artificiales «(...) con posibilidades anatómicas que son capaces de recrear las elevaciones de la pasión o aún de inventar nuevos deseos» (Brown, 2008: 235). Publicadas en el sexto número de la famosa revista surrealista *Minotauro*, las fotos de las muñecas de Bellmer quedarían servir como una crítica política de los cuerpos; sin embargo, las motivaciones de Bellmer terminaron siendo objeto de una crítica sobre la dominación del cuerpo femenino: «vinculando sus fantasías de chicas adolescentes y poderosas seductoras con los temas de la nostalgia y el erotismo, y relacionando su trabajo con el ambivalente deseo y repulsión de los surrealistas franceses...»
por el cuerpo femenino» (Brown, 2008: 235). Como en el mito de Pigmalión, Bellmer proyecta los deseos masculinos sobre un cuerpo del que no se tiene posesión. En estos seres, según las historiadoras del arte Therese Lichtenstein y Sidra Stich, «el cuerpo entero puede ser montado y desmontado como una máquina» (Brown, 2008: 236), de la forma misma en que son tratadas las androides en Ghost in the Shell.

Andróides y cíborgs en la antesala del control social

La traducción al latín de los escritos árabes de la Escuela de Alejandría introdujo en el occidente cristianizado el pensamiento mecanicista (Truitt, 2015). El famoso ajedrecista de Kempelen, El Turco, se convirtió en el vínculo entre la creación de máquinas mecánicas y las primeras máquinas de cálculo como parte de los primeros intentos por imitar la inteligencia humana. Desde el ajedrecista de Kempelen hasta el jugador de damas de Arthur Samuel y el primer programa verdaderamente funcional de un jugador de ajedrez de Alex Bernstein, el juego ha sido un tema esencial en la historia de la creación de seres artificiales inteligentes. El juego ha sido un medio privilegiado para escenificar una dinámica social basada en la identificación y clasificación de los deseos y movimientos de un enemigo, tanto en el campo de batalla como en las relaciones sociales. John von Neumann y Oscar Morgenster, en su teoría de juegos, quisieron desarrollar formas de procesamiento de información que pudieran ser aplicadas a la lucha entre enemigos —no necesariamente limitada al campo militar—, caracterizados como «un operador perfectamente inteligente y despiadado» (Galison, 1994: 159).

Lo que hay detrás del jugador artificial es una idea que ya desde Ramon Lull se revelaría posible: recrear la inteligencia por medios mecánicos, pues los humanos, como diría Descartes, son máquinas maravillosas, pero máquinas, a pesar de todo. Si los seres humanos pueden ser pensados e interpretados como mecanismos, el ser humano puede ser conocido y controlado, de la misma forma en la que se conoce y se entiende una máquina. Esta idea calará profundo en la forma de pensar de muchos de los impulsores de la cibernética y la IA. Norbert Wiener y Arturo Rosenblueth, por ejemplo, afirmarán en su Purposeful and Non-purposeful Behavior que, «como objetos de indagación científica, los humanos no diferen de las máquinas» (Galison, 1994: 250). Pero más que igualar la máquina y el humano, para el artista Wayne McGregor, entre otros, debemos pensar la IA como una forma de expandir nuestras capacidades. En su obra Living Archive, de 2018, McGregor crea una coreografía a partir de un diálogo entre su trabajo, los bailarines y las herramientas de IA, generando movimientos inéditos a partir de la identificación de patrones en sus obras.

Aquella separación entre mente-cuerpo, sin embargo, ha sobrevivido en los desarrollos actuales de la IA, con sus evidentes consecuencias políticas. Esta separación, como fundamento mecánico-filosófico del entendimiento del ser humano, es portadora de un poder patriarcal que se ha impuesto en el campo de la ciencia y la técnica, pero también en otros campos estratégicos de la vida: «La separación mente-cuerpo es, por supuesto, fundamental para la tradición patriarcal occidental, esencial no solo para la teología, sino también para la política, la economía y todas las demás áreas donde las mujeres y los trabajadores han sido sistemáticamente subordinados» (McCorduck, 2004: 40). Al simplificar la complejidad biológica humana para igualar a las máquinas y el humano, esta separación entre cuerpo y mente ha legitimado la superioridad intelectual de un grupo social privilegiado por el género, la raza y el poder.

David Hume, en su Tratado de la naturaleza humana, argumenta que, en el mismo sentido que lo había hecho John Locke un siglo antes, la mente puede ser explicada por medio de leyes universalmente válidas, en una dinámica patriarcal de la ciencia que Donna Haraway denominará como la cultura de la no cultura, destacada en el trabajo de Robert Boyle, donde se evidencia el valor masculino y caballeresco de la autoinvisibilidad como «la forma científica específicamente moderna, europea y masculina de la virtud de la modestia». De esta manera, destaca Haraway, el hombre de ciencia de la temprana modernidad europea «está dotado con el importante poder de establecer los hechos», pues su «subjetividad es su objetividad» (Haraway, 2004: 35). Abrogándose el conocimiento de la verdad a partir de su testimonio modesto, el hombre no solo será el juez legítimo de lo verdadero, también será el padre genuino de un mundo tecnificado.

 Todos estos complejos intentos por crear seres artificiales autónomos e inteligentes encuentran sus más fructíferos desarrollos durante y después de la Segunda Guerra Mundial. Peter Galison (1994) analiza cómo la cibernética, entendida como una de las ramas que contribuyó decisivamente en el desarrollo del campo de la IA, estuvo atravesada por la visión del mundo de Norbert Wiener.

Desde ya antes del inicio de esta guerra, Wiener estaba interesado en ofrecer sus conocimientos para el desarrollo de estrategias defensivas militares, especialmente en el campo de las defensas antiaéreas. Al iniciar la guerra y durante los años siguientes, la atención de Wiener se interesó principalmente en el problema de cómo destruir el aeroplano enemigo, en lo que se denominó como Antiaircraft Predictor (AA), un problema que necesitaba de la visión de la máquina para identificar, clasificar y atacar al adversario. Es precisamente sobre la base de esta lógica del enemigo, bajo la cual se imaginaba Wiener el funcionamiento del mundo en la posguerra, que nace la cibernética: «Finalmente, el predictor AA, junto con sus nociones de ingeniería asociadas a sistemas de retroalimentación y cajas negras, se convirtió, para Wiener, en el modelo para una comprensión cibernética del universo» (Galison, 1994: 229). Además de este enemigo, emerge un tercero, que Galison llama Enemy...
Other, calculado en los laboratorios del MIT a partir de tres disciplinas articuladas: «investigación de operaciones, teoría de juegos (…) y cibernética».

Los servomecanismos, identificados por Le Roy Archibald MacColl con «sistemas esclavos que emplean retroalimentación» (MacColl, 1945: 8), son esenciales en la teoría cibernética. Para Wiener, «el enemigo servomecánico se convirtió (…) en el prototipo de la psicología humana y, en últimas, de toda la naturaleza humana» (Galison, 1994: 233). El ML, «el mayor experimento de clasificación en la historia humana» (Crawford 2017), es heredero directo de una idea de control sobre el mundo, postulada por Wiener: «la idea de Wiener fue utilizar redes eléctricas para determinar, varios segundo por adelantado, dónde podría ocurrir un ataque aéreo y emplear ese conocimiento para dirigir el fuego de la artillería» (Galison, 1994: 234). La identificación y la clasificación fueron campos cruciales en la aplicación de esta visión de la guerra y de la sociedad. «De particular importancia, sostuvieron que su clasificación rehabilitaba el “propósito” y la “teleología” al ponerlos bajo los auspicios de un “análisis conductista uniforme” que era igualmente aplicable a los organismos vivos y las máquinas» (Galison, 1994: 245).

La cibernética, como una rama fundamental para el desarrollo contemporáneo de la IA, es postulada por Wiener no solo como una forma efectiva de controlar las acciones del enemigo militar, sino también de ese Enemy Other aplicado a la sociedad misma: «los mensajes pueden ser enviados con el propósito de explorar el universo, pero pueden también ser enviados con la intención de controlarlo» (Galison, 1994: 256).

Paul Edwards (1997) demuestra, además, cómo estas ideas fueron fundamentales en la psicología cognitiva y en el campo de la IA: «Las teorías científicas de la inteligencia artificial y de la psicología cognitiva también forman núcleos de ideologías sobre la mente humana como máquinas manipulables» (Edwards, 1997: 21). Así, el computador se convirtió en una «metáfora culturalmente central para el control, para el análisis científico y para la mente» (1997:28).

En este contexto de un mundo analizado y controlado por medio de servomecanismos de retroalimentación computacionales, la película Colossus: The Forbin Project, lanzada públicamente en el año de 1970, recrea la construcción de una superinteligencia computacional por parte del científico Charles Forbin, que tiene la capacidad de monitorizar todas las formas de comunicación en plena Guerra Fría, en un presentimiento tecnológico materializado por medio de la IA en el programa PRISM, de la NSA. Una escena de la película de Batman, The Dark Night, de 2008, recrea visualmente esta vigilancia y control panóptico con un Morgan Freeman frente a una pared abarrotada de pantallas que presentan modelos 3D en tiempo real, basados en la geolocalización de los datos generados por los teléfonos móviles de todos los ciudadanos de Gotham City.

Para Steve Anderson, el asunto de la vigilancia y del control no es en el fondo un simple problema tecnológico; más bien, se trata de una visión del mundo materializada, en nuestro caso, en los actuales sistemas de IA y, en particular, en las RNP: «Aún la vigilancia computacional no es primariamente un asunto tecnológico; está más propiamente enmarcada como el ejercicio de poder por parte de estructuras institucionales sobre grupos e individuos» (Anderson, 2017: 133).

Las RNP, sin embargo, han creado su propio accidente. Los Adversarial Examples, en el campo de la visión algorítmica, son instancias que impiden la correcta identificación y clasificación de personas, cosas y acciones (Szegedy et al., 2014). En mi proyecto de investigación visual intitulado Inspección Entheo-Rítmica, de 20203, un video construido a partir de un set de imágenes diseñadas por Nguyen, Yosinski y Clune (2015), me he permitido reflexionar sobre los límites de la vigilancia y el control, al experimentar con la identificación y la clasificación de los objetos que allí supuestamente surgen a la luz de la percepción del modelo algorítmico (ver fig. 1).

Fig. 1. Fotograma del vídeo resultado del proyecto de investigación Inspección Entheo-Rítmica.

Las más de cinco mil imágenes que componen el set de datos de Nguyen, Yosinski y Clune fueron analizadas por el modelo de identificación y clasificación COCO-SSD –que es un conjunto de datos de detección, segmentación y subtítulos de objetos a gran escala–, por medio de la aplicación Run-way, ejecutada en un servidor local. Las imágenes que efectivamente fueron etiquetadas por el modelo se aislaron y agruparon con el objetivo de contrastarlas visualmente con fotos de objetos de la misma categoría, reconozcibles por la visión humana y alojadas en el set de datos de entrenamiento de COCO-SSD. Así, el espectador puede experimentar el abismo que aquí separa...
la percepción visual de las RNP, de la percepción visual humana, contrastando la realidad creada por cada uno de estos modelos perceptivos, ahondando en los límites y errores de la visión algorítmica, e imaginando narrativas alternativas a las tareas de identificación y clasificación que estos sistemas cumplen con el objetivo de vigilar y controlar.

En un sentido parecido a los trabajos de Adam Harvey, Sterling Crispin o de Joselyn McDonald y su Mother Protect Me, de 2019, Inspección Entheo-Rítmica reflexiona sobre la realidad algorítmica y las maneras en las que la máquina podría alucinar un mundo irreconocible para el ser humano. Inspección Entheo-Rítmica, además de generar tensiones críticas entre la percepción humana y la percepción maquinaria, pone presente las diversas formas en las que las tareas de identificación y clasificación pueden ser hackeadas, aprovechando el error inherente a las RNP, abordando críticamente el alcance de la vigilancia y el control de los sistemas algorítmicos.

Conclusiones

Este artículo ha intentado demostrar que, junto a las tareas serviles y de entretenimiento, las de vigilancia y control social han sido históricamente predominantes en lo que atañe a la consecución de seres con inteligencia autónoma; que, para realizarlas, la identificación y la clasificación son esenciales para comprender su funcionamiento. Además, que estas tareas emergen de una visión patriarcal del mundo, en el que personas, animales y cosas deben estar bajo la disposición de su poder. Con ello se ha pretendido proponer una discusión acerca de los proyectos de vigilancia y control que hoy se encaman en los sistemas de ML y las RNP, reclamando un interés especial por las consecuencias políticas y sociales que se derivan de su implementación cada vez más cotidiana y globalizada, a la vez que los intereses que se esconden detrás de sus publicitados beneficios. Las explicaciones mágicas que han empezado a surgir para tratar de aclarar el misterio detrás de las RNP terminan, como se ha afirmado, excusando la responsabilidad que sujetos e instituciones deberían asumir respecto a las consecuencias de la implementación masiva de estos sistemas.

Las prácticas artísticas aquí vinculadas proponen no solo una mirada crítica sobre el funcionamiento de la IA, también ofrecen alternativas, por un lado, sobre las maneras diferentes en las que podemos relacionarnos con estas tecnologías y, por el otro, sobre las posibilidades estéticas y políticas que la interacción entre humanos e IA puede albergar. Desnudando la neutralidad de los set de datos, deslegitimando el automatismo de los algoritmos como excusa de la responsabilidad de quienes diseñan y desarrollan, pero también criticando los supuestos que se encaman en sus funcionamientos, el arte en el campo de la IA construye lugares desde donde podemos pensar y crear otras realidades.

Los límites de este artículo impiden abordar en profundidad todas las problemáticas relacionadas con la identificación y la clasificación, con la vigilancia y el control, por ejemplo, sobre los actores involucrados, sobre las muy diversas tecnologías que encarnan actualmente esta visión del mundo o las más recientes teorías sobre el aprendizaje profundo, supervisado o no. Sin embargo, espero haber abordado los puntos neurálgicos de los que en un futuro inmediato dependerá nuestra relación con la tecnología, así como el tipo de sociedades que junto a ella queremos promover.

Referencias bibliográficas

Anderson, Steve F. 2017. Technologies of Vision: The War Between Data and Images. Cambridge, Massachusetts: The MIT Press.

Benjamin, Ruha. 2019. Race After Technology: Abolitionist Tools for the New Jim Code. 1.a ed. Nueva York: Polity.

Bunz, Mercedes. 2009. «Are Hewlett-Packard computers really racist?». The Guardian, December 23, https://www.theguardian.com/media/pda/2009/dec/23/hewlett-packard.

Brown, Steven T. 2008. 'Machinic Desires: Hans Bellmer's Dolls and the Technological Uncanny in Ghost in the Shell 2: Innocence'. Mechademia 3(1): 222-53. https://doi.org/10.1353/mec.0.0088.

Campolo, Alexander y Kate Crawford. 2020. «Enchanted Determinism: Power without Responsibility in Artificial Intelligence». Engaging Science, Technology, and Society 6(0): 1-19. https://doi.org/10.17351/ests2020.277.

Crawford, Kate. 2017. «The Trouble with Bias». Keynote lecture held at NIPS, Conference at Long Beach, California, December 4-9. https://www.youtube.com/watch?v=ffMym_BKWzk.

Dreyfus, Hubert L. 1995. Alchemy and Artificial Intelligence. Santa Monica, CA: RAND Corporation. https://www.rand.org/pubs/papers/P3244.html.

Dreyfus, Hubert L. 1972. What computers can’t do: a critique of artificial reason. New York: Harper & Row.

Dzieza, Josh. 2020. «How hard will the robots make us work?». The Verge, February 27. https://www.theverge.com/2020/2/27/21155254/automation-robots-unemployment-jobs-vs-human-google-amazon.

Edwards, Paul N. 1997. The Closed World: Computers and the Politics of Discourse in Cold War America. Cambridge, Mass. London: MIT Press Ltd.

Galison, Peter. 1994. «The Ontology of the Enemy: Norbert Wiener and the Cybernetic Vision». Critical Inquiry 21(1): 228-266.

Homero. 2010. Iliada. Madrid: MESTAS Ediciones.

Haraway, Donna Jeanne. 2004. Testigo_Modesto@Segundo_Milenio. HombreHembra©_Conoce_Oncoratón®. Barcelona: Editorial UOC, S.L.
Klingemann, Mario. 2018. «Neural Glitch». Quasimondo. Accesado, March 2, 2020. http://underdestruction.com/2018/10/28/neural-glitch/.
Nguyen, Anh, Jason Yosinski y Jeff Clune. 2015. «Deep Neural Networks are Easily Fooled: High Confidence Predictions for Unrecognizable Images». arXiv:1412.1897 [cs], April. http://arxiv.org/abs/1412.1897.
MacColl, Le Roy Archibald. 1945. Fundamental Theory of Servomechanisms. 1.a ed. First Printing. D. Van Nostrand Company, inc.
McCorduck, Pamela. 2004. Machines Who Think: A Personal Inquiry into the History and Prospects of Artificial Intelligence. 2.a ed. Natick, Mass: A K Peters/CRC Press.
ỌnỌha, Mimi. 2018. «How the Arts Help Us Understand Our Relationship With Technology», video. https://www.youtube.com/watch?v=_JJ6Oj_u-Qw.
Paglen, Trevor y Kate Crawford. 2019. «Excavating AI». https://www.excavating.ai.
Platón. 2011. La República o El Estado. Traducido por Patricio de Azcárate Corral. Madrid: Austral.
Raviv, Shaun. 2020. «The Secret History of Facial Recognition». Wired Magazine, January 21. https://www.wired.com/story/secret-history-facial-recognition/.
Truitt, Elly R. 2015. Medieval Robots: Mechanism, Magic, Nature, and Art. Pennsylvania:University of Pennsylvania Press. http://muse.jhu.edu/book/39875/.
Szegedy, Christian, Wojciech Zaremba, Ilya Sutskever, Joan Bruna, Dumitru Erhan, Ian Goodfellow y Rob Fergus. 2014. «Intriguing properties of neural networks». arXiv:1312.6199 [cs], February. http://arxiv.org/abs/1312.6199.
CV

Hugo Felipe Idárraga
Universidad de Los Andes
h.idarraga@uniandes.edu.co

Filósofo de la Universidad Nacional de Colombia y Magister en Comunicación de la Pontificia Universidad Javeriana. Interesado en temas de teoría de la imagen, visión por computador e inteligencia artificial. Actualmente estudiante de la Maestría en Humanidades Digitales de la Universidad de Los Andes.
ARTICLE

NODE «AI, ARTS & DESIGN: QUESTIONING LEARNING MACHINES»

«Creative AI: From Expressive Mimicry to Critical Inquiry»

Angus G. Forbes
University of California, Santa Cruz

Date of submission: April 2020
Accepted in: June 2020
Published in: July 2020

Recommended citation

Forbes, Angus G. 2020. «Creative AI: From Expressive Mimicry to Critical Inquiry». In: Andrés Burbano; Ruth West (coord.) «AI, Arts & Design: Questioning Learning Machines». Artnodes, Nº.26: 1-10. UOC, Leonardo/ISAST. [Accessed: dd/mm/yy]. https://doi.org/10.7238/a.v0i26.3370

Abstract

The nascent field of what has come to be known as “creative AI” consists of a range of activities at the intersections of new media arts, human-computer interaction, and artificial intelligence. This article provides an overview of recent projects that emphasise the use of machine learning algorithms as a means to identify, replicate, and modify features in existing media, to facilitate new multimodal mappings between user inputs and media outputs, to push the boundaries of generative art experiences, and to critically investigate the role of feature detection and pattern identification technologies in contemporary life. Despite the proliferation of such projects, recent advances in applied machine learning have not yet been incorporated into or interrogated by creative AI projects, and this article also highlights opportunities for computational artists working in this area. The article concludes by envisioning how creative AI practice could include delineating the boundaries of what can and cannot be learned by extracting features from artefacts and experiences, exploring how new forms of interpretation can be encoded into neural networks, and articulating how the interaction of multiple machine learning algorithms can be used to generate new insight into the intertwining sociotechnical systems that encompass our lives.
Keywords
creative AI, machine learning, generative art, new media art

IA creativa: De la mímica expresiva a la investigación crítica

Resumen
El incipiente campo de lo que se conoce como “IA creativa” consiste en una serie de actividades en las intersecciones de las nuevas artes mediáticas, la interacción persona-computadora y la inteligencia artificial. Este artículo proporciona una descripción general de proyectos recientes que enfatizan el uso de algoritmos de aprendizaje automático como un medio para identificar, replicar y modificar características en los medios existentes, para facilitar nuevas asignaciones multimodales entre las entradas del usuario y las salidas de los medios, para ampliar los límites en las experiencias del arte generativo e investigar críticamente el papel de las tecnologías de detección de características e identificación de patrones en la vida contemporánea. A pesar de la proliferación de proyectos de este tipo, los avances recientes en el aprendizaje automático aplicado aún no han sido incorporados o cuestionados por proyectos creativos de IA, y este artículo también destaca las oportunidades para los artistas computacionales que trabajan en esta área. El artículo concluye imaginando cómo la práctica creativa de IA podría incluir e delinear los límites de lo que se puede y no se puede aprender extrayendo características de artefactos y experiencias, explorando sobre cómo las nuevas maneras de interpretación pueden codificarse en redes neuronales y definiendo cómo la interacción de múltiples máquinas con algoritmos de aprendizaje se pueden utilizar para generar una nueva visión de los sistemas sociotécnicos entrelazados presentes en nuestras vidas.

Palabras clave
IA creativa, aprendizaje automático, arte generativo, arte de nuevos medios

The Paper:

1. Introduction

One advantage in using machine learning to extract meaning from data is that it lets the researcher sidestep the need to articulate the low-level details contained in the data, which can be difficult to tease out and hard to define. How do you describe what films you like? It is easier to provide a training set of films that you’ve rated and let the algorithm discover what features highly related films have in common (Hallinan & Striphas 2016). How do you capture the nuances in meaning when translating a phrase from one language to another? It is more accurate to provide the machine learning system with a vast amount of data in order to infer these subtleties without requiring formal semantics (McCann et al. 2017). How do you best describe the special characteristics of a person so that they can be distinguished from others in an image, no matter where the image was taken, what pose they are in, or what they are wearing? State-of-the-art recognition systems do not require any description whatsoever, only a sufficient number of examples that the deep learning network extrapolates from and encodes as weights within its hidden layers (Taigman et al. 2014, Sun et al. 2014). What strategy do you use to articulate the rules that define an artist’s expressivity? Style transfer algorithms effortlessly let you transform any image or video into an impressionist painting, using even a single image of a painting to automatically find the characteristic elements of the artist’s style (Gatys et al. 2016).

For many applications, deep learning neural networks are the most effective method to identify useful features in datasets and to use them to interpret new data with similar content. In addition to choosing the most computationally efficient architecture or parameters, a main focus of the data analyst using them becomes to define the space of interpretation by choosing the dataset that represents that space, by selecting an appropriate loss function for training the network, and by deciding what outputs can be returned when querying the network. Learning to interpret the data occurs through a process of encoding hierarchies of features that indicate whether a particular input (or part of that input) belongs to a particular category. Although there has been much work on trying to make sense of what these features “mean” (Olah et al. 2017, Carter et al. 2019), either individually or in aggregate, understanding is enabled through a process of curation rather than by explicit explanation. In this way, machine learning introduces a new approach to making sense of the world in which choosing examples and defining mappings judiciously enables new applications and new forms of creative expression.
The Creative Coding Lab at University of California, Santa Cruz\(^1\) investigates the use of machine learning algorithms for scientific research and creative explorations across a range of contexts. One effort, Deep Illumination, explores how deep learning can be used effectively in the graphics pipeline, investigating, for example, how to infer complex lighting models from a large dataset of examples, rather than through expensive rendering calculations, and evaluating how such an approach can provide useful trade-offs between time and memory (Thomas & Forbes, 2017, Elek et al. 2019, Alsaiari et al. 2019). Our lab has also investigated the use of machine learning technologies for a range of practical applications. One project, CompostNet, trains a neural network to classify food waste appropriate for available trash and recycling receptacles (Frost et al. 2019a). Another project uses machine learning to predict biker density at dangerous road intersections so that drivers and bikers can experience improved shared road safety (Dubey et al. 2019a). Researchers in the Creative Coding Lab have also investigated creative applications using machine learning. For example, the Art I Don’t Like project used a novel recommender system that introduces users to artists and art genres that they may be unfamiliar with (Frost et al., 2019b), and the Data Brushes art application enables users to interactively paint using specialised brushes that generate output using neural style transfer networks (Dubey et al. 2019b). Much of the architecture for deep learning neural networks was first theorised and implemented in previous decades (Bishop 1995, LeCun et al. 1998, Rumelhart et al. 1996), but the recent explosion of deep learning techniques and applications introduced in the last few years was in part enabled by innovations in GPU technology (LeCun et al. 2015, Krizhevsky et al. 2012). Neural networks are loosely modelled on the behaviour of neurons, and the Creative Coding Lab has been exploring models of computational intelligence inspired by other biological processes. One recent project, developed in collaboration with astrophysicists at University of California, Santa Cruz, emulates properties of the Physarum polycephalum (the “many-headed slime mold”) in order to infer a simulation of the dark matter filament structure of the Cosmic Web using only a sparse sampling of astrophysical observations (Burchett et al. 2020).\(^2\)

The term “creative AI” is increasingly used by artists and designers who utilise machine learning to generate creative outputs, or who treat machine learning algorithms as a medium in and of itself in various ways (McCormick et al. 2020). In recent years, creative AI projects have been featured at the NeurIPS Workshop for Creativity and Design, as well as at other arts and computation venues, such as the ACM SIGGRAPH Art Gallery and Art Papers tracks, the IEEE VIS Arts Program, and the International Symposium on Electronic Art. Broadly speaking, creative AI projects involve one or more of the following: mimicking existing data, mapping features found in one dataset onto another, or mapping inputs to outputs in unusual ways, visualising or otherwise probing the inner workings of the algorithm, and analysing or speculating about the societal impact of machine learning systems. These activities can enable new kinds of generative artworks that can either replicate or incorporate existing artworks or create entirely new artistic outputs. They also can be used to design new techniques of more expressively interacting with existing art forms. In doing so, they introduce new ways to analyse and experience cultural artefacts and cultural data. Finally, the machine learning algorithm, its computational architecture, the input it requires, the resulting output, and the analysis framework it is part of can be thought of as a cultural artefact in and of itself, enabling new forms of critical inquiry. In the sections below, I provide an overview of these trends, along with descriptions of related projects, and highlight opportunities for computational artists working in this area.

2. Creative AI as expressive mimicry

Creating software that automatically generates artworks —either in the style of a particular artist, or in an original voice that does not directly reference existing work—is a perennial pursuit in new media practice and generative art. Well-known early examples include Harold Cohen’s robot paintings (Cohen 1995) and David Cope’s experiments in musical intelligence (Cope 1996). Often in these projects, the visual or audio outputs, while interesting on their own, are a byproduct of the actual artwork, which is the system itself: in Cohen’s case, AARON is the artwork; for Cope, his EMI software is the main creative contribution. A more recent example is introduced by Sougwon Chung, who, as part of her Drawing Operations series, co-improvises drawings in collaboration with a robotic arm that is controlled via a recurrent neural net that has been previously trained on her own drawings (Chung 2019). Research into techniques that can be used to simulate human expressions, voices, and faces meant to fool users or for other nefarious purposes, also called “deep fakes”, shows great creative potential for designing realistic human behavior, perhaps in combination with text generation and speech generation techniques. For example, work by Suwajanakorn et al. (Suwajanakorn et al. 2017) demonstrates how a voice impressionist can create a convincing video of another person speaking words that they never uttered. Thies et al. (Thies et al. 2016, Thies et al. 2019) introduce projects that enable a user to become a kind of virtual puppeteer using their own facial expressions to modify the expressions of another person...

---

1. https://creativecoding.soe.ucsc.edu/
2. An overview of projects from the UCSC Creative Coding Lab was presented in late July 2019, as part of the “AI in the Arts and Design” panel discussion with Erkki Huhtamo, Memo Akten, and Max Sims at ACM SIGGRAPH, organised by Ruth West, Victoria Szabo, and Danielle Siembieda.
in a video. Work by Fried et al. (Fried et al. 2019) demonstrates a method to surreptitiously modify a video of a person talking simply by editing the textual transcript of the video. Chan et al. (Chan et al. 2019) introduce a method to transfer the recorded movements of an expert performer onto a new video featuring an amateur performer, appearing to transform novices into professional dancers. This technology exacerbates difficulties in separating facts from opinions, in thinking critically, and in identifying bias and propaganda (Gebru 2019, Jo & Gebru 2020), but it also potentially presents new avenues for exploring these issues and for new forms of creative work.

3. Creative AI as interactive mapping

Machine learning enables the creation of tools that map a range of inputs to new outputs, often in a different modality. By definition, all algorithms require an input that is then processed in some way to produce an output. Neural networks, including deep learning networks, are “tuned” through a training process that encodes an effective mapping of inputs to outputs for a particular dataset (the training set). If successful, and if the training set is representative of the kinds of inputs that will be encountered in the future, then the network can be queried nearly instantly to provide a meaningful output given some new, previously unseen input data. Fiebrink’s Wekinator tool enables users to quickly train a neural network (or another machine learning algorithm) to recognise, for example, different gestures from a web camera and associate them with sounds or musical instructions (Fiebrink et al. 2016). KIMA: *The Wheel* is a multimedia performance by the art collective Analema Group that uses machine learning to correlate sound and visual parameters, generating a multimodal mapping between voices and visual outputs (Gingrich et al., 2018). Style transfer networks that encode stylistic features of a source image learn to map any image into a transformed version of that image that incorporates those features. Gatys et al. (Gatys et al. 2016) introduced *neural style transfer*, which makes use of a convolutional neural network to identify image patterns that represent a particular painter’s “style”, and can then transfer it onto any other image, making it possible, for example, to transform a photograph into an image that looks like it was painted by Van Gogh or Kandinsky, to use popular examples.

4. Creative AI as generative art

A range of techniques investigate the neural network as space of possibility. The “deep dream” algorithm, which transforms images into psychedelic quilts was originally created as a tool to highlight which features were being activated when processing an image with a neural network. If a neural network is trained to classify, for example, different species of birds, then a particular patch of a bird image (or an image that contains bird-like objects) will trigger the neurons within the network that have been tuned to respond to that particular bird feature. Often these features, when viewed in isolation, resist easy interpretation, and represent a particular curve or gradient or texture that proved to be useful in detecting a bird within an image (Olah et al. 2017). The *Inceptionism* project takes these features and iteratively integrates them onto the image, allowing us to see which features are observed in a given input image. To continue the example, even if an input image contains no birds at all, and if the network is trained only to recognise bird features, the technique ends up generating a kind of Boscian heliograph of bird parts (Mordvintsev et al. 2015).

Initial breakthroughs in deep learning led to state-of-the-art methods in data classification, identifying items in a photo, automatically tagging people in social media posts, or recommending products or content based on previous interactions or purchases on a website (LeCun et al. 2015, Goodfellow et al. 2016). If a network has been trained to identify particular features in order to, say, decide what category an image belongs to, then that network could also be used to generate new images made up of those features and belonging to that category (Goodfellow et al. 2014). The generative adversarial network (GAN) architecture consists of both a *generator* network and a discriminator network. During the training process, the generator network gets better at producing output, and the discriminator network gets better at distinguishing a real image from the training data from a generated image. Once the generator is sufficiently trained, any input to the generator network will produce a realistic output, that is, an output that contains features recognised by the discriminator network as a real image. The input to the generator network is a vector of numbers within a particular range of values describing a “latent space”, and slightly changing the values of one of more of the numbers in the input vectors produces images that are similar to each other (Bojanowski et al. 2019). Animations of images created by “drifting” through the latent space (i.e. updating the input vector) produces a morphing between images that resemble the training images, sometimes creating a surreal effect. Artists have been inspired by GAN techniques that make it possible to direct the data generation process (Mirza & Osindero 2014, Radford et al. 2015, Karras et al. 2019). For example, a recent iteration of Refik Anadol’s *Machine Hallucination* project uses a GAN trained on 100 million photographic memories of New York City found publicly in social networks to create synthetic representations that envision a possible “near future” (Anadol 2019). Mario Klingemann has created a series of animations using a technique he calls “neural glitch”, in which he alters the weights in a trained generator to create intriguing “misinterpretations” that nonetheless retain a coherent style (Klingemann 2018). Casey Reas’ *Earthly Delight* series generates what he terms “compressed cinema”, using a GAN architecture trained via processed stills from Stan Brakhage’s experimental films in which plants are directly placed on
top of clear film strips (Menezes 2019). Memo Akten’s *Learning to See* processes live camera input, composing images that resemble the shape and structure of this input, but replacing the content with data learned through training a network on particular types of images, transforming, for example, keys and wires into flowers and waves, or faces into galaxies (Akten et al. 2019).

5. **Creative AI as critical inquiry**

Some recent creative AI projects can be considered as critical inquiries that investigate sociotechnical systems that utilise machine learning. Tom White’s influential project *Perception Engines* creates idiosyncratic images made out of a few simple shapes with solid colors and curved dark lines. While at first glance they seem to be vaguely evocative of a particular object or action, upon reading the title of each print (such as “cello”, “cabbage”, “hammerhead shark”, “iron”, and “tick”), it becomes hard to see anything else. While the prints create a kind of visual puzzle, they also function as images that return the highest confidence score on different image classification algorithms (often higher even than photographs of those objects), providing insight into what shape features form a “Platonic ideal” of a category encoded in the image recognition network, and representing the “character” of a class more effectively than any one instance (White 2018). Avital Meshi’s *Classification Cube* features an interactive surveilled space in which multiple machine learning algorithms are used to classify a participant’s behaviours, expressions, age, and gender. In addition to making it clear that some expressions and poses are incorrectly categorised, and that a person’s age or gender can be misclassified depending on seemingly minor changes, the project provides a space for reflecting on the ubiquitous automated interpretations that permeate our daily lives (Meshi & Forbes, 2020). While machine learning systems are implicated in algorithmic bias (Diakopoulos 2015, Eubanks 2018), bias of course exists prior to being encoded into datasets and deep learning networks trained on those datasets. Creative interrogations of machine learning systems can help to pinpoint aspects of a data analysis pipeline that introduce bias and spark discussions about the ramifications of weaving machine learning into the fabric of public life.

6. **Creative AI opportunities**

Novel sophisticated machine learning techniques are presented each year at the International Conference on Computer Vision (ICCV), the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Neural Information Processing Systems (NeurIPS), the ACM Special Interest Group on Computer Graphics and Interactive Techniques (SIGGRAPH), and various other computer science venues. Researchers often put versions of the articles online at the arXiv.org open-access archive and make the code for these projects available in online repositories, enabling anyone to test out their techniques using popular software frameworks, such as TensorFlow and PyTorch. Given their accessibility, there are many opportunities for incorporating contemporary machine learning techniques into creative projects. For instance, Isola et al. and Park et al. introduce architectures that have been used to make interactive demos that infer a reasonable image from only outlines or coloured rectangles (Isola et al. 2017, Park et al. 2019). A more recent project called *GauGAN* lets a user easily modify generated images by “painting” particular features on the image (Bau et al. 2019), and an interactive demo by Liu et al. lets a user edit an existing photo by erasing people or objects, automatically “ inpainting”, replacing them with relevant elements from the surrounding landscape (Liu 2018). Other generative machine learning projects have appeared over the last few years, many of which are geared toward graphics techniques for visual effects in films and games, but have not yet, to the best of my knowledge, been incorporated into media arts projects or to augment interactive performance. Xie et al. (Xie et al. 2018) showed that realistic motion dynamics could be created and shaped interactively by training a neural network on a database of fluids. Their system learns to generate fine details in explosions, water, or smoke from low-resolution inputs, which speeds up computation and enables visual effects artists to quickly create high-quality animations of different fluids. A number of projects have focused on generating realistic human and animal motion and motion planning strategies for navigating specialised environments, including for rock climbing simulations (Naderi et al. 2017), walking through diverse terrain (Zhang et al. 2018), or in crowds (Amirian et al. 2019). For example, work by Holden et al. (Holden et al. 2017) trains a neural network using a database of human movement captured in a motion capture lab, including walking, jumping, climbing stairs, and crouching. This network is then able to determine the most reasonable motions for a virtual character moving through any scene, finding correlations between the motions stored in the networks and the elements within the scene. Even for scenes with arrangements of terrain and objects that are quite different from the data it was trained on, the network produces synthetic motion outputs that are convincingly realistic.

Many creative AI projects differentiate themselves by curating the data and labels they choose for the training set or as inputs into the network. To take just two examples, Chris Rodley uses a style transfer network to create compelling images of dinosaurs composed out of fruit (Rodley 2017) and Pinar Yanardag and Emily Salvador use generative adversarial networks trained on a database of fashion designs to create new dresses and jewellery (Yanardag and Salvador, 2019). Some intriguing machine learning techniques enable cross-modal mapping, in which data from one domain informs or creates the output in another (Baltrušaitis et al. 2018). Recent techniques
automatics provide captions from an image or accurately label subregions in an image (Karpathy & Fei-Fei 2015, Gan et al. 2017), or the reverse, generate images from text (Qiao et al. 2019). For example, Zhang et al. generate accurate images (at least at first glance) of birds from simple descriptions, such as “This bird is red and brown in color, with a stubby beak”, enabling users to “paint” with words (Zhang et al. 2017). In addition to encapsulating a form of cognitive blending, in which emergent meanings are constructed from mixing together partial matches in two different domains (Fauconnier & Turner 2003), they illustrate that existing cultural artefacts (such as online field guides for bird watchers, photo collections of flowers, or various forms of social media) contain conceptual analogies that define “unseen” relationships that enable new forms of automated reasoning (Peyre et al. 2019, Yan et al. 2019). Designing machine learning systems that leverage or investigate cultural artefacts presents opportunities for new creative work and cultural insight.

Techniques such as style transfer and inpainting show that there is unexpected information that can be mined from even a small number of input data samples, and which can then be used for creative reinterpretations. Other recent examples include learning 3D information from 2D data, such as a technique to synthesise animations that contain novel views of complex scenes from a set of input images (Mildenhall et al. 2020) and a technique that estimates the depth of elements in an image in order to automatically create a “Ken Burns” animation effect consisting of zooming, panning, and motion parallax (Niklaus et al. 2019). Another technique learns to synthesise frames of future frames from a single image, predicting plausible ways that a scene might change over time (Xue et al. 2018). Techniques such as these can infer unexpected features and relationships between those features, and present many creative possibilities that have yet to be fully explored. Additionally, different types of sensors can expose new features in data. For instance, by using a slow-motion camera with a high temporal resolution, Davis et al. (Davis et al. 2014) were able to recreate sounds in a room by observing the subtle motions of particular objects in that room, such as plants or packaging. In one experiment, they demonstrate that they can retrieve someone singing a nursery rhyme simply by recording and accentuating the vibrational movements on a package of chips. While this has implications for surveillance, it also illustrates how inventive uses of sensors can provide unexpected streams of information in other sensory domains. Incorporating datasets from higher resolution instruments into machine learning systems can lead to new creative applications.

**7. The future of Creative AI**

Given the continuing breakthroughs, it is worth thinking about what machine learning is not yet able to achieve, and about what components of an artwork cannot be effectively modelled or mimicked. For example, so far, machine learning approaches have not successfully generated convincing dramatic experiences or engaging multimedia performances. These kinds of experiences require contextual information which we do not yet understand how to encode effectively and thoroughly. Narrative, dance, performance, and cinema are inherently more complex than static images or sound recordings, and require integrating many elements simultaneously, such as lighting, editing, acting, narrative, and sound design. Machine learning makes the assumption that all relevant features can be found within the training data, and even if there were a way to gather and label relevant data from, say, a film or a live performance, we bring our knowledge of the world and our expectations about how to interpret particular genres when experiencing art. Moreover, these experiences are ultimately interior and perhaps ineffable, resonating with a rich personal database of our own experiences and our own thoughts and feelings.

That is, machine learning algorithms can effectively identify and utilise features in artworks in increasingly sophisticated ways, but do not model how an artwork is perceived or why it is interpreted in a particular way. Media artists, in addition to using new media forms to create new representations and new experiences, also investigate the nature of media itself, and often foreground concept over or alongside aesthetics and technical craftsmanship (Shanken 2002a, Agüera y Arcas 2017, Ackerman et al. 2018). Creative AI practitioners will continue to identify which concepts resist machine learning approaches and to investigate how machine learning tools can make particular interpretations either inescapable or impossible.

Machine learning technologies can be thought of as a type of measuring instrument. Many sensors include a computational component in which data is filtered or otherwise processed to separate out the noise from the signal. Neural networks measure distinguishing features in data, and can provide insight into the system the data is drawn from, as well as about other systems with which it is entwined. For example, observing transportation patterns or analysing pollution levels can be used to provide insight into the economic health of a city (Washington 2020), and interactions on social media can be used to identify personality traits, and then exploited for targeted advertising or disinformation campaigns (Kaiser 2019). Insight into these auxiliary systems then could allow us to infer patterns from yet other interacting systems. The premise of “big data” is not simply that we can collect higher and higher resolution spatiotemporal data, and not only that we can retrieve and analyse data more and more quickly, but that we can make use of all this data to make sense of how systems interact and integrate with each other (Shanken 2002b, Hassad 2020). How should we design the next iteration of machine learning tools that reason about the world holistically by integrating multiple interpretations encoded in text, mined from image and video databases, perceived by sensors, provided by human-computer interactions, and communicated by yet other machine learning tools? Creative AI will continue to be a space in which artists and researchers create
personal yet empirical research projects that explore and challenge the logic of how different systems and interpretations of those systems promote or impede each other.

References

Ackerman, Margaretta, James Morgan, and Christopher Cassion. “Co-Creative Conceptual Art.” Proceedings of the Ninth International Conference on Computational Creativity (ICCC), pages 1–8. 2018.

Agúera y Arcas, Blaise. 2017. “Art in the age of machine intelligence.” Arts 6, no. 4: 18. https://doi.org/10.3390/arts6040018.

Akten, Memo, Rebecca Fiebrink, and Mick Grijson. “Learning to See: You Are What You See.” Proceedings of ACM SIGGRAPH Short Papers, pages 1-6. 2019. https://doi.org/10.1145/3306211.3320143.

Alsaiari, Abeer, Ridhi Rustagi, Manu Mathew Thomas, and Angus G. Forbes. “Image Denoising Using A Generative Adversarial Network.” Proceedings of the IEEE 2nd International Conference on Information and Computer Technologies (ICICT), pages 126-132. 2019. https://doi.org/10.1109/INFOCT.2019.8710893.

Anadol, Refik. “Machine Hallucination”. 2019. Available online at: http://refikanadol.com/works/machine-hallucination/.

Baltrusaitis, Tadas, Chaitanya Ahuja, and Louis-Philippe Morency. “Multimodal machine learning: A survey and taxonomy.” IEEE Transactions on Pattern Analysis and Machine Intelligence 41, no. 2 (2018): 423-443. https://doi.org/10.1109/TPAMI.2018.2798607.

Bau, David, Hendrik Strobelt, William Peebles, Jonas Wulff, Bolei Zhou, Jun-Yan Zhu, and Antonio Torralba. “Semantic photo manipulation with a generative image prior.” ACM Transactions on Graphics 38, no. 4 (2019): 1-11. https://doi.org/10.1145/3306346.3323023.

Bishop, Christopher M. Neural Networks for Pattern Recognition. Oxford University Press, 1995. https://doi.org/10.1201/9781420050646.

Bojanowski, Piotr, Armand Joulin, David Lopez-Paz, and Arthur Szlam. “Optimizing the latent space of generative networks.” arXiv preprint arXiv:1707.05776, 2019.

Burchett, Joseph N., Oskar Elek, Nicolas Tejos, J. Xavier Prochaska, Todd M. Tripp, Rongmon Bordoloi, and Angus G. Forbes. “Revealing the Dark Threads of the Cosmic Web.” The Astrophysical Journal Letters 891, no. 2 (2020): L35. https://doi.org/10.3847/2041-8213/ab700c.

Carter, Shan, Zan Armstrong, Ludwig Schubert, Ian Johnson, and Chris Olah. “Activation atlas.” Distill 4, no. 3 (2019): e15. https://doi.org/10.23915/distill.00015.

Chan, Caroline, Shirly Ginosar, Tinghui Zhou, and Alexei A. Efros. “Everybody dance now.” Proceedings of the IEEE International Conference on Computer Vision, pages 5933-5942. 2019. https://doi.org/10.1109/ICCV.2019.00603.

Chung, Sougwen. “Artefact 1.” Sougwen Chung Portfolio Website. 2019. https://sougwen.com/project/artefact1.

Cohen, Harold. “The further exploits of AARON, painter.” Stanford Humanities Review 4, no. 2 (1995): 141-158.

Cope, David. Experiments in musical intelligence. Middleton WI: A-R Editions, 1996.

Davis, Abe, Michael Rubinstein, Neal Wadhwa, Gautham J. Mysore, Frédoux Durand, and William T. Freeman. “The visual microphone: Passive recovery of sound from video.” ACM Transactions on Graphics 33, no. 4 (2014):79. https://doi.org/10.1145/2601097.2601119.

Diakopoulos, Nicholas. “Algorithmic accountability: Journalistic investigation of computational power structures.” Digital Journalism 3, no. 3 (2015): 398-415. https://doi.org/10.1080/21670811.2014.976411.

Dubey, Mahika, Alan Peraal Ortiz, Rakshit Agrawal, and Angus G. Forbes. “Predicting Biker Density at Bikeshare Station Intersections in San Francisco.” 2019 IEEE Global Humanitarian Technology Conference (GHTC), pages 1-7. 2019a. https://doi.org/10.1109/GHTC46095.2019.9033019.

Dubey, Mahika, Jasmine Otto, and Angus G. Forbes. “Data Brushes: Interactive style transfer for data art.” Proceedings of IEEE VIS Arts Program. 2019b. https://doi.org/10.1109/VISAP.2019.8900858.

Elek, Oskar, Manu M. Thomas, and Angus Forbes. “Learning Patterns in Sample Distributions for Monte Carlo Variance Reduction.” arXiv preprint arXiv:1906.00124, 2019.

Eubanks, Virginia. Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor. New York: St. Martin’s Press, 2018.

Faulconier, Gilles, and Mark Turner. “Conceptual blending, form and meaning.” Recherches en Communication 19 (2003): 57–86. https://doi.org/10.14428/rec.v19i19.48413.

Fiebrink, Rebecca, and Baptiste Caramiaux. “The machine learning algorithm as creative musical tool.” In The Oxford Handbook of Algorithmic Music, edited by Roger T. Dean and Alex McLean, chapter 12. Oxford University Press, 2016.

Forbes, Angus G. “Articulating media arts activities in art-science contexts.” Leonardo 48, no. 4 (2015): 330-337. https://doi.org/10.1145/2810177.2810179.

Forbes, Angus G. and Kiyomitsu Odai. “Iterative synaesthetic composing with multimedia signals.” Proceedings of the International Computer Music Conference (ICMC), pages 573-578. 2012.

Fried, Ohad, Ayush Tewari, Michael Zolliker, Adam Finkelstein, Eli Shechtman, Dan B. Goldman, Kyle Genova, Zeyu Jin, Christian Theobalt, and Maneesh Agrawala. “Text-based editing of talking-head video.” ACM Transactions on Graphics 38 no. 4 (2019): 1-14. https://doi.org/10.1145/3306346.3323028.

Frost, Sarah, Bryan Tor, Rakshit Agrawal, and Angus G. Forbes. “CompostNet: An Image Classifier for Meal Waste.” 2019 IEEE Global Humanitarian Technology Conference (GHTC), pages 1-4. 2019a. https://doi.org/10.1109/GHTC46095.2019.9033130.
Frost, Sara, Manu Mathew Thomas, and Angus G. Forbes. “Art I don’t like: An anti-recommender system for visual art.” *Proceedings of Museums and the Web*. 2019b.

Gan, Chuang, Zhe Gan, Xiaodong He, Jianfeng Gao, and Li Deng. “StyleNet: Generating attractive visual captions with styles.” *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 3137-3146. 2017. https://doi.org/10.1109/CVPR.2017.108.

Gatys, Leon A., Alexander S. Ecker, and Matthias Bethge. “Image style transfer using convolutional neural networks.” *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 2414-2423. 2016. https://doi.org/10.1109/CVPR.2016.265.

Gebru, Timnit. “Race and Gender.” *arXiv preprint arXiv:1908.06165*, 2019.

Gingrich, Oliver, Sean Soraghan, Alain Renaud, Evgenia Emets, and Dario Villanueva-Ablanedo. “KIMA: The Wheel—Voice Turned into Vision: A participatory, immersive visual soundscape installation.” *Leonardo* (Online - Accepted for publication) (2018): 1-13. https://doi.org/10.1162/leon_a_01698.

Goodfellow, Ian, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. “Generative adversarial nets.” *Advances in Neural Information Processing Systems*, pages 2672-2680. 2014.

Goodfellow, Ian, Yoshua Bengio, and Aaron Courville. *Deep Learning*. MIT press, 2016.

Hallinan, Blake, and Ted Striphas. “Recommended for you: The Netflix Prize and the production of algorithmic culture.” *New Media & Society* 18, no. 1 (2016): 117-137. https://doi.org/10.1177/1461444814538646.

Hassad,Rossi A. “A foundation for inductive reasoning in harnessing the potential of big data.” *Statistics Education Research Journal* 19, no. 1 (2020): 238-258.

Holden, Daniel, Taku Komura, and Jun Saito. “Phase-functioned neural networks for character control.” *ACM Transactions on Graphics* 36, no. 4 (2017): 1-13. https://doi.org/10.1145/3072959.3073663.

Isola, Phillip, Jun-Yan Zhu, Tinghui Zhou, and Alexei A. Efros. “Image-to-image translation with conditional adversarial networks.” *Proceedings of the IEEE conference on Conference on Computer Vision and Pattern Recognition*, pages 1125-1134. 2017. https://doi.org/10.1109/CVPR.2017.632.

Jo, Eun Seo, and Timnit Gebru. “Lessons from archives: strategies for collecting sociocultural data in machine learning.” *Proceedings of the Conference on Fairness, Accountability, and Transparency*, pages 306-316. 2020.

Kaiser, Brittany. *Targeted: The Cambridge Analytica Whistleblower’s Inside Story of How Big Data, Trump, and Facebook Broke Democracy and How It Can Happen Again*. HarperCollins, 2019.

Karpathy, Andrej, and Li Fei-Fei. “Deep visual-semantic alignments for generating image descriptions.” *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 3128-3137. 2015. https://doi.org/10.1109/CVPR.2015.7298932.

Karras, Tero, Samuli Laine, and Timo Aila. “A style-based generator architecture for generative adversarial networks.” *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 4401-4410. 2019. https://doi.org/10.1109/CVPR.2019.00453.

Klingemann, Mario. “Neural Glitch.” *Quasimondo*. October 28, 2018. http://underdestruction.com/2018/10/28/neural-glitch/.

LeCun, Yann, Léon Bottou, Yoshua Bengio, and Patrick Haffner. “Gradient-based learning applied to document recognition.” *Proceedings of the IEEE* 86, no. 11 (1998): 2278-2324. https://doi.org/10.1109/5.726791.

LeCun, Yann, Yoshua Bengio, and Geoffrey Hinton. “Deep learning.” *Nature* 521, no. 7553 (2015): 436-444. https://doi.org/10.1038/nature14539.

Liu, Guolin, Flitsum A. Reda, Kevin J. Shih, Ting-Chun Wang, Andrew Tao, and Bryan Catanzaro. “Image inpainting for irregular holes using partial convolutions.” *Proceedings of the European Conference on Computer Vision (ECCV)*, pages 85-100. 2018. https://doi.org/10.1007/978-3-030-01252-6_6.

McCann, Bryan, James Bradbury, Caiming Xiong, and Richard Socher. “Learned in translation: Contextualized word vectors.” *Advances in Neural Information Processing Systems*, pages 6294-6305. 2017.

McCormack, Jon, Oliver Bown, Alan Dorin, Jonathan McCabe, Gordon Monro, and Mitchell Whiteall. “Ten questions concerning generative computer art.” *Leonardo* 47, no. 2 (2014): 135-141. https://doi.org/10.1162/LEON_a_00533.

McCormack, Jon, Patrick Hutchings, Toby Gifford, Matthew Yee-King, Maria Teresa Liano, and Mark d’Inverno. “Design Considerations for Real-Time Collaboration with Creative Artificial Intelligence.” *Organised Sound* 25, no. 1 (2020): 41-52. https://doi.org/10.1017/S1355771819000451.

Menezes, Caroline. “Interview with Casey Reas.” *Studio International*. May 21, 2019. https://www.studiointernational.com/index.php/casey-reas-interview-computer-art-coding/.

Meski, Avital and Angus G. Forbes. “Stepping inside the Classification Cube: An intimate interaction with an AI system.” *Leonardo* 53, no. 4 (2020): 387-393.

Mildenhall, Ben, Pratul P. Srinivasan, Matthew Tancik, Jonathan T. Barron, Ravi Ramamoorthi, and Ren Ng. “NeRF: Representing Scenes as Neural Radiance Fields for View Synthesis.” *arXiv preprint arXiv:2003.08934*, 2020.

Mirza, Mehdi, and Simon Osindero. “Conditional generative adversarial nets.” *arXiv preprint arXiv:1411.1784*, 2014.

Mordvintsev, Alexander, Christopher Olah, and Mike Tyka. “Inceptionism: Going Deeper into Neural Networks”. Google AI Blog.
June 17, 2015. https://ai.googleblog.com/2015/06/inceptionism-going-deeper-into-neural.html.
Naderi, Kourosh, Joose Rajamäki, and Perttu Hämäläinen. “Discovering and synthesizing humanoid climbing movements.” *ACM Transactions on Graphics* 38, no. 6 (2019): 1-15. https://doi.org/10.1145/3355089.3356528.
Olah, Chris, Alexander Mordvintsev, and Ludwig Schubert. “Feature visualization.” *Distill* 2, no. 11 (2017): e7. https://doi.org/10.23915/distill.00007.
Park, Taesung, Ming-Yu Liu, Ting-Chun Wang, and Jun-Yan Zhu. “Semantic image synthesis with spatially-adaptive normalization.” *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 2337-2346. 2019. https://doi.org/10.1109/CVPR.2019.00244.
Peyre, Julia, Ivan Laptev, Cordelia Schmid, and Josef Sivic. “Detecting Unseen Visual Relations Using Analogies.” *Proceedings of the IEEE International Conference on Computer Vision*, pages 1981-1990. 2019. https://doi.org/10.1109/ICCV.2019.00207.
Qiao, Tingting, Jing Zhang, Duansong Xu, and Dachen Tao. “MirrorGAN: Learning text-to-image generation by redescription.” *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 1505-1514. 2019. https://doi.org/10.1109/CVPR.2019.00160.
Radford, Alec, Luke Metz, and Soumith Chintala. “Unsupervised representation learning with deep convolutional generative adversarial networks.” *arXiv* preprint arXiv:1511.06434, 2015.
Rodley, Chris. “Deep Dinosaur”. June 19, 2017. https://chrisrodley.com/2017/06/19/dinosaur-flowers/.
Rumelhart, David E., Geoffrey E. Hinton, and Ronald J. Williams. “Learning representations by back-propagating errors.” *Nature* 323, no. 6088 (1986): 533-536. https://doi.org/10.1038/323533a0.
Shanken, Edward A. “Art in the information age: Technology and literature.” In *From Energy to Information: Representation in Science and Technology, Art, and Literature*, editors Bruce Clarke and Linda D. Henderson, pages 155-177. Stanford University Press, 2002b.
Sun, Yi, Xiaogang Wang, and Xiaou Tang. “Deep learning face representation from predicting 10,000 classes.” *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 1891-1898. 2014. https://doi.org/10.1109/CVPR.2014.244.
Suwajanakorn, Supasorn, Steven M. Seitz, and Ira Kemelmacher-Shlizerman. “Synthesizing Obama: Learning lip sync from audio.” *ACM Transactions on Graphics* 36, no. 4 (2017): 1-13. https://doi.org/10.1145/3072959.3073640.
Taigman, Yaniv, Ming Yang, Marc’Aurelio Ranzato, and Lior Wolf. “Deepface: Closing the gap to human-level performance in face verification.” *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 1701-1708. 2014. https://doi.org/10.1109/CVPR.2014.220.
Thies, Justus, Mohamed Elgharib, Ayush Tewari, Christian Theobalt, and Matthias Nießner. “Neural Voice Puppetry: Audio-driven Facial Reenactment.” *arXiv* preprint arXiv:1912.05566, 2019.
Thies, Justus, Michael Zollhofer, Marc Stamminger, Christian Theobalt, and Matthias Nießner. “Face2face: Real-time face capture and reenactment of RGB videos.” *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 2387-2395. 2016. https://doi.org/10.1109/CVPR.2016.262.
Thomas, Manu Mathew, and Angus G. Forbes. “Deep Illumination: Approximating Dynamic Global Illumination with Generative Adversarial Network.” *arXiv* preprint arXiv:1710.09834, 2017.
Washington, Simon, Matthew G. Kariiatlis, Fred Manning, and Panagiotis Anastasopoulos. *Statistical and Econometric Methods for Transportation Data Analysis*. CRC press, 2020. https://doi.org/10.1201/9780429244018.
White, Tom. “Synthetic Abstractions.” Aug 23, 2018. https://medium.com/@tom_25234/synthetic-abstractions-80e8689f390.
Xue, Tianfan, Jiajun Wu, Katherine L. Bouman, and William T. Freeman. “Visual dynamics: stochastic future generation via layered cross convolutional networks.” *IEEE Transactions on Pattern Analysis and Machine Intelligence* 41, no. 9 (2018): 2236-2250. https://doi.org/10.1109/TPAMI.2018.2854726.
Yan, Chenggang, Liang Li, Chunjie Zhang, Bingtao Liu, Yongdong Zhang, and Qionghai Dai. “Cross-modality bridging and knowledge transferring for image understanding.” *IEEE Transactions on Multimedia* 21, no. 10 (2019): 2675-2685. https://doi.org/10.1109/TMM.2019.2903448.
Yanardag, Pinar and Emily Salvador. “The Little Black Dress Reimagined by an A.I.” 2019. https://lbd-ai.com/.
Zhang, Han, Tao Xu, Hongsheng Li, Shaoting Zhang, Xiaogang Wang, Xiaolei Huang, and Dimitris N. Metaxas. “Stackgan: Text to photo-realistic image synthesis with stacked generative adversarial networks.” *Proceedings of the IEEE International Conference on Computer Vision*, pages 5907-5915. 2017. https://doi.org/10.1109/ICCV.2017.629.
Zhang, He, Sebastian Starke, Taku Komura, and Jun Saito. “Mode-adaptive neural networks for quadruped motion control.” *ACM Transactions on Graphics* 37, no. 4 (2018): 1-11. https://doi.org/10.1145/3197517.3201366.
CV

Angus G. Forbes
University of California, Santa Cruz
angus@ucsc.edu

Angus Forbes is an associate professor in the Department of Compu-
tational Media at University of California, Santa Cruz, where he directs
the Creative Coding Lab. His research investigates novel techniques
for visualising and interacting with complex scientific information, and
his interactive artwork has been featured at museums, galleries, and
festivals throughout the world. He was the SIGGRAPH 2018 Art Papers
Chair and will be the Art Gallery Chair for SIGGRAPH 2021. More infor-
mation about current projects can be found at https://creativecoding.
sOE.ucsc.edu/.
Intelligent Environments and Public Art

Eitan Mendelowitz
Mount Holyoke College

Date of submission: May 2020
Accepted in: June 2020
Published in: July 2020

Abstract
Intelligent environments combine the promise of ubiquitous computing with artificial intelligence and are increasingly being used in public art. The agent-based approach to artificial intelligence (AI) uses the *intelligence function* to characterize agent-based behavior. The inputs to the intelligence function, perception of the environment and the agent’s internal state, combined with the outputs of the function, actuation and changes in internal state, provides a lens with which to categorized AI-based public art. Such works can be classified as generative, reactive, interactive, learning, or static. To illustrate this taxonomy, this paper gives examples of public artworks that fit into each of the five categories and uses the taxonomy to suggest new areas of creative inquiry.

Keywords
public art, intelligent environments, ubiquitous computing, artificial intelligence, interactive art, generative art

Recommended citation
Mendelowitz, Eitan. 2020. «Intelligent Environments and Public Art». In: Andrés Burbano; Ruth West (coord.) «AI, Arts & Design: Questioning Learning Machines». Artnodes, no. 26: 1-8. UOC. [Accessed: dd/mm/yy]. http://doi.org/10.7238/a.v0i26.3364
Entornos inteligentes y arte público

Resumen
Los entornos inteligentes combinan la promesa de la computación ubicua con la inteligencia artificial y se utilizan cada vez más en el arte público. El enfoque basado en agentes de la inteligencia artificial (IA) utiliza la función inteligente para caracterizar el comportamiento basado en agentes. Las entradas a la función inteligente, la percepción del entorno y el estado interno del agente, combinadas con las salidas de la función, la actuación y los cambios en el estado interno, proporcionan un baremo con el que clasificar el arte público basado en IA. Estas obras se pueden clasificar como generativas, reactivas, interactivas, de aprendizaje o estáticas. Para ilustrar esta taxonomía, este artículo proporciona ejemplos de obras de arte públicas que se ajustan a cada una de las cinco categorías y utiliza la taxonomía para sugerir nuevas áreas de investigación creativa.

Palabras clave
arte público, entornos inteligentes, computación ubicua, inteligencia artificial, arte interactivo, arte generativo

Introduction
"[Ubiquitous computing] created a new field of computer science, one that speculated on a physical world richly and invisibly interwoven with sensors, actuators, displays, and computational elements, embedded seamlessly in the everyday objects of our lives and connected through a continuous network." (Weiser, Gold, and Brown 1999). This world, originally envisioned in the ’80s, is becoming a reality through the availability of affordable and accessible devices like the Microsoft Kinect, Arduino, Raspberry PI, and inexpensive LEDs.

Intelligent environments, ubiquitous computing technologies coupled with artificial intelligence (AI), is currently being adopted by artists to create interactive public art. After defining two key terms — public art and artificial intelligence — and discussing the use of AI in art in general, this paper proposes a novel taxonomy for categorizing public art that uses AI as a medium and applies the taxonomy to some recent public artworks.

Public Art
While there is some disagreement over the precise definition of public art, this paper will use an expansive definition - public art is art that is situated in public spaces rather than traditional art contexts like museums and galleries (Zebracki 2013). In describing public art, the Association for Public Art stresses that public art is not a medium or an art form in and of itself but is primarily defined by the artwork’s setting (“What Is Public Art?” n.d.).

Artificial Intelligence
Similar to the term public art, the term artificial intelligence is commonly used but poorly defined. The Turing Test is a popular way to determine if a human-made system is intelligent. Seeping into popular culture (Tyldum 2014; Square Enix 2016), the Turing Test is often the first thing people think of when they think of AI. Turing proposed different versions of his test but, at their core, the tests describe a system as artificially intelligent if a human evaluator is unable to distinguish the system from a human through conversation (Turing 1950). Unfortunately, the Turing test can only be applied to behaviors that are thought of as humanlike, it cannot be used for non-humanoid systems like the intelligent environments often used in public art.

An expansive definition of AI was offered by the Dartmouth Workshop where the term Artificial Intelligence was originally coined: “...the artificial intelligence problem is taken to be that of making a machine behave in ways that would be called intelligent if a human were so behaving.” (McCarthy et al. 1955). To paraphrase the Dartmouth Workshop’s definition of AI, “we know it when we see it”. While the Workshop provides a workable definition of AI it does not offer a way to evaluate, analyze or classify such systems.

Artificial Intelligence in the Arts
Almost as soon as computers became available, they were used to create generative digital art. While artists like Frieder Nake (Nake 2005) and Vera Molinar (Roe-Dale 2019) created algorithmic art in the sixties, these works did not employ techniques commonly thought of as AI. Among the first AI-based artworks was AARON by Harold Cohen (Cohen 2016). AARON, first built in 1973 and continually developed and maintained for over 40 years, uses an expert system encoded by Cohen to create paintings. Starting in 1975 – contemporaneous with Cohen’s early development of AARON – Myron Krueger created Videoplace (Krueger 1985). Videoplace is an early exploration of mixed reality. In Videoplace, two people in different rooms interact with each other through a system of video cameras and...
shared projections. While the first versions of Videoplace used only video equipment, the development of Videoplace continued for over a decade. Later versions of the work employed custom computer systems coupled with image analysis and computer vision systems.

The use of AI in the media arts continued throughout the eighties and expanded in the nineties. In 1993, Genetic Images, by Karl Sims, used genetic algorithms to generate abstract images (Kelly 1994). In the same year Sims was “evolving” graphics, Simon Penny finished Petit Mal, an autonomous robotic artwork in which simple connectionist steering behaviours combine to create a charming and surprisingly expressive “artificial life” (Penny 1997). Also in the nineties, Ken Feingold began making talking animatronic heads powered by natural language processing techniques. Feingold continued this creative pursuit through the first decade of this century until today (Feingold n.d.).

Recently, the popularity and evocative nature of Google’s DeepDream (Mordvintsev 2015) has spurred a trend in AI-based art. A number of artists are now using deep neural networks to train AI to create traditional-looking artworks. For example, Memories of Passersby I (Vincent 2019) used thousands of portraits from the 17th to the 19th centuries to train a generative adversarial network (GAN) to create novel (if somewhat distorted) portraits. Similarly, Gene Kogan’s Cubist Mirror (Mufson 2016) uses a webcam and style transfer to render live video of a museum space as a Cubist painting.

This recent style of generative deep neural network-based artwork is discussed in Defining AI Arts: Three Proposals, by Lev Manovich in which Manovich seeks to define “AI art.” He proposes (and rejects) two possible definitions. His first proposal is to create a Turing test for art in which “art created by an AI” is defined to be “something that professionals recognize as valid historical art or contemporary art.” The problem with this approach, Manovich points out, is that it limits AI art creation to already existing art forms and precludes the AI’s participation in the expansion of art into new expressions and modalities. In his second proposal, Manovich asks rhetorically if the use of AI techniques in an artwork’s production can be employed to distinguish “AI art” from other digital art. Manovich rejects this definition by pointing out that the human artist often exerts a high level of control even in artworks that use AI techniques and therefore these artworks should still be considered more human-created than “AI art.”

Finally, Manovich concludes by offering a third definition: “AI art is [a] type of art that we humans are not able to create because of the limitations of our bodies, brains, and other constraints. One such possibility I sketched above is computer generated objects, media, situations and experiences that do not have the usual systematically and predictability of human art – but they are not random either, they don’t mechanically juxtapose elements just to shock, and they are not simply instances of remix aesthetics.” (Manovich 1999)

An unwritten assumption in Manovich’s essay is that “AI art” is artwork produced by an artificially intelligent system. For example, the “Turing AI arts” test, as described by Manovich, focuses on the output from the AI to determine if the AI is an artist. His second proposal is rejected because even if AI algorithms are used, the system’s creator still has a controlling hand in the output. Because of his assumption, he concludes that only a truly autonomous AI system creating something different than that created by humans can be considered “AI art”.

Unlike Manovich, this paper considers AI a medium in and of itself. When evaluating “AI art”, this paper does not look at the output from an AI but the system itself. Building upon the Dartmouth Workshop’s broad definition for AI, this paper considers AI Art any artwork that uses algorithms and techniques commonly considered intelligent. In a sense, this paper embraces Manovich’s second proposal without worrying about issues of algorithmic autonomy — the artwork is the human-designed AI system and not the system’s output.

**Motivation**

The creation of AI-based art began in the seventies, expanded in the nineties and is recently enjoying a new surge of attention. Missing from the discourse are methods for categorizing such AI-based artwork in a consistent manner. In proposing a taxonomy for AI art and applying it to public art, I hope to provide a method to compare and contrast AI-based artwork. As such work becomes increasingly prevalent, this taxonomy may provide a tool for more nuanced discussions by building on decades of research in agent-based AI.

**Intelligent Agents**

An agent-based approach to AI, popularized in the mid’90s by the influential textbook Artificial Intelligence: A Modern Approach, argues that intelligence should be judged by an agent’s ability to achieve goals...
in its environment. In this formulation, the behavior of an intelligent agent is a function of its perception of the environment – as perceived through sensors – and the agent’s internal state. This intelligence function maps perceptions and internal state onto actions. Actions are either external actions enacted by the agent upon the environment through actuators, or internal actions, i.e. the updating of internal state (Russell, Norvig, and Davis 2010).

A Taxonomy of AI-Based Public Art

The intelligence function provides a useful tool to analyze AI systems. Individual agents can be characterized by:
1. The amount to which the agent’s intelligence function relies on their perceptions.
2. The amount to which the agent’s intelligence function relies on the agent’s internal state.
3. The amount to which the agent’s actions affect its environment. And,
4. The amount to which the agent’s actions change its internal state.

In other words, an AI can be characterized by 4 metrics:
1. Perception
2. Introspection
3. Actuation
4. Self-Mutability

These metrics can be used to create a taxonomy of intelligent environments and AI-based public art. While these four measures are continuous, it is useful for classification to think of these measures as binary (high/low):

- AI-based artwork whose intelligence function is based exclusively on its internal state (introspection) and ignores any sensory input (or has no sensory input) can be classified as generative.
- AI whose actuation is based almost entirely on perception while ignoring its internal state can be termed reactive.
- Artwork whose intelligence function is influenced by both perception and introspection but does not modify its internal state is interactive.
- AI-based art that changes its internal state in response to perception and responds to a combination of its perceptions and internal state is learning.

Generative AI

Generative systems either ignore their preceptors or have no sensors with which to perceive. It is an approach employed by Refik Anadol in many of his recent public artworks. Anadol uses artificial intelligence techniques and machine learning to transform datasets into 3D data visualizations that he calls “data sculptures” (Simonite 2020). In his 2018 work, WDCH Dreams, Anadol employed machine learning algorithms to form associations between 587,763 images, 1,880 videos, 1,483 metadata files, and 17,773 audio files drawn from Walt Disney Concert Hall’s archives. According to Anadol, this “mind” is designed to mimic human dreams. The resultant visualizations were projected onto the Concert Hall’s skin as part of a week-long public art installation and were also used for a year-long exhibition inside the Concert Hall’s Ira Gershwin Gallery (Anadol n.d.).
Another example of a generative public artwork is the 2016 installation Diffusion Choir by Sosolimited, Plebian Design, and Hypersonic. Diffusion Choir was commissioned by Biomed Realty and is installed in Cambridge, Massachusetts. In Diffusion Choir, 400 Tyvek origami objects hang from the ceiling of an atrium in a 3D array. Each object can open and close like a small cocktail umbrella. The motion of the Tyvek objects is driven by a simulation of a flock of birds (Sterling 2018). Like WDCH Dreams the behavior of Diffusion Choir is completely determined by its internal data.

Reactive AI

Reactive systems’ actuation is based entirely on the agent’s perceptions of the environment. For example, HeartHug by Dmitry Sokolov was installed at the Canal Convergence festival in 2019. In this work, a large heart-shaped sculpture made of florescent lights hangs above an open area. A computer vision system observes people below. Half the heart sculpture lights up when a person is below the sculpture but the entire heart is only activated when two people stand under the heart and hug.

A second example of reactive public art is Cloud Display by Rafael Lozano-Hemmer installed in the Manchester International Festival in 2019. In this work, participants speak into a microphone. A speech to text system recognizes the spoken words and writes them in mist on a billboard-sized display made out of 1600 ultrasonic water atomizers (Lozano-Hemmer n.d.). Like HeartHug, Cloud Display is entirely reactive, the behavior of the artwork is driven entirely by participant action.

Interactive AI

Interactive systems combine the introspection of a generative system with the perception of a reactive system. Many interactive public artworks use sensors to trigger database queries, the results of which are used as source material for display. The City Pulse, a collaboration between Local Projects, Legends, NowArchival, and The Hettema Group in 2015, is installed in the one-hundredth-floor observatory at One World Trade Center. In The City Pulse, professional actors and comedians fill the role of “ambassadors”. The “ambassadors” stand in front of a ring of displays and interact with an audience, telling an improvised story. Throughout the story, they use gestures to retrieve and display media on the ring (Wilson 2015). In this intelligent environment, the “ambassador” is the interactor while the public acts as an audience.

Learning AI

Like interactive public art, learning AI actuation is a function of perception and internal state. However, learning systems are distinct from interactive public art in that they modify their internal state. With a learning system, the artwork’s behavior can change over time, even when faced with the same perceptions. A learning AI uses all aspects of the intelligence function.
Many of David Rokeby’s artworks address the subject of artificial intelligence and learning. His groundbreaking work *Very Nervous System (VNS)* was awarded the Prix Ars Electronica Award of Distinction for Interactive Art in 1991. VNS has been installed in both private and public art contexts. In VNS, a computer vision system detects and processes movement within the installation space. Movements are used to stimulate a neural network which in turn produces music. What is unique about VNS is that the neural network is constantly in flux, its levels of activation, and the strength of its synaptic connections are constantly changing in response to interaction and its own internal state (Rokeby 2010). While the behavior of the system is consistent and not random, it is also never exactly the same.

*Amatria*, by Philip Beesley, hangs above the atrium of Luddy Hall at Indiana University. A tangle of 3D-printed forms, white Mylar, acrylic plastics, wire, glass, and laser cut stainless steel, *Amatria* looks like a massive chandelier spun out in a web by a nest of robotic spiders. Embedded throughout the structure are motion sensors, microphones, electrical current sensors (for proprioception), motors, LEDs, and speakers. The work moves, lights up, and makes sounds in response to the movement of visitors (Beans 2018). The 2018 work utilizes a “curiosity based learning algorithm” that uses reinforcement learning to maximize knowledge gains and to generate interactive behaviors and adapt to change” (Chan 2016). What is most interesting about *Amatria* is that its behaviors are constantly changing as a result of its interactions with the public.

**Static AI**

Al-based public art that has no actuation is static. A static work can be a sculpture, mural, or any other medium that does not change over time. Even without actuation, artificial intelligence can still be used in an artwork’s design or construction. For example, *Dio* is a sculpture by Ben Snell. Snell used machine learning algorithms trained on over 1000 images of classical sculptures. The resultant AI then constructed a model that was used to 3D print a mold. Using the mold, Snell cast *DIO* using a resin containing the ground-up remains of the computer used in *DIO*’s design (Schwab 2019).

*Dio* is not an example of public art. In fact, at the time of this writing, there does not seem to be any static AI-based public artwork. This conclusion was reached after exhaustively examining the portfolios of artists listed by *AIArtists.org* (“Global Directory of AI Artists” n.d.), searching the *Public Art Archive* (“Public Art Archive” n.d.) and searching various other art-related sources on the internet. By using the intelligence function-based taxonomy we have discovered a previously overlooked potential branch of AI-based public art.

**Conclusion**

This paper introduces an approach to classifying AI-based public artwork into five categories: generative, reactive, interactive, learning, or static. It then gives examples of public artwork that fit into each category. While engaging in a survey of existing work it became apparent that the majority of AI-based public art is either generative or interactive. While VNS and *Amatria* are excellent examples of learning AI, there are very few other public artworks that fit in this category. And, there are even fewer static works (if any exist at all). This would suggest that the creation of learning or static AI-based public artworks would be contributions to the public art community. By using the intelligence function-based taxonomy to analyze existing AI-based public art, this paper has revealed artistic approaches that can yield novel creations.

**References**

Anadol, Refik. “WDCH Dreams.” Accessed March 30, 2020. http://refikanadol.com/.

Beans, Carolyn. “Science and Culture: Sentient Architecture Promises Insight into Our Evolving Relationship with AI.” *Proceedings of the National Academy of Sciences* 115, no. 30 (2018): 7638–40. https://doi.org/10.1073/pnas.1809390115.

Chan, Tsz Kiu. “Curiosity-Based Learning Algorithm for Interactive Art Sculptures.” February, 2016. https://uwspace.uwaterloo.ca/handle/10012/10278.

Cohen, Paul. “Harold Cohen and AARON.” *AI Magazine*, Winter 2016. 63-66.

Feingold, Ken. “Ken Feingold: Artworks and Documentation.” Accessed June 1, 2020. http://www.kenfeingold.com/.
“Global Directory of AI Artists.” AIArtists.Org. Accessed March 30, 2020. https://aiaartists.org/ai-artist-directory.

Kelly, Kevin. “Genetic Images. When Computers Breed Art and Humans Direct the Flow of Fate, Startling, Alien Beauty is Born.” Wired Magazine 2.09, September 1994.

Krueger, Myron W., Thomas Gionfriddo, and Katrin Hinrichsen. “VIDEOPLACE - an Artificial Reality.” ACM SIGCHI Bulletin 16, no. 4 (April 1, 1985): 35–40. https://doi.org/10.1145/1165385.317463.

Lozano-Hemmer, Rafael. “Rafael Lozano-Hemmer - Project 'Cloud Display.'” Accessed March 30, 2020. http://www.lozano-hemmer.com/cloud_display.php.

Makowski, Ann. “Drive By.” Text. October 21, 2013. https://segd.org/content/drive.

McCarthy, John, Marvin Minsky, Nathaniel Rochester, IBM Corporation, and Claude Shannon. “A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence.” August 13, 1955.

Manovich, Lev. “Defining AI Arts: Three Proposals.” June 2019. Accessed June 1, 2020. http://manovich.net/content/04-projects/107-defining-ai-arts-three-proposals/manovich.defining-ai-arts2019.pdf.

Mufson, Beckett. “Feel Like Pablo with This ‘Cubist Mirror’.” Vice, May 25, 2016, https://www.vice.com/en_us/article/aen9vb/cubist-neural-network-mirror.

Mordvintsev, Alexander, Christopher Olah, and Mike Tyka. “DeepDream - a code example for visualizing Neural Networks.” 2015. Google Research 2, no. 5.

Nake, Frieder. “Computer art: a personal recollection.” Proceedings of the 5th Conference on Creativity & Cognition (C&amp;C ’05). Association for Computing Machinery, New York, NY, USA, 54–62. 2005. https://doi.org/10.1145/1056224.1056234.

Penny, Simon. “Embodied cultural agents: at the intersection of robotics, cognitive science, and interactive art.” In AAAI Socially Intelligent Agents Symposium. 1997.

Public Art Archive. Accessed March 30, 2020. https://www.publicartarchive.org/.

Roe-Dale, Rachel R. “Vera Molnar’s Exploration of the Unimaginable.” Math Horizons 26, no. 4 (May 15, 2019): 18–21. https://doi.org/10.1080/10724117.2019.1568081.

Rokeyb, David. “David Rokeyb: Very Nervous System.” November 24, 2010. http://www.davidrokeyb.com/vns.html.

Ronchi, Gloria. “Aether and Hemera’s LED Installation Uses Artificial Intelligence to Transform Human Voices into a Brass Band Light Show.” October 11, 2012. https://inhabitat.com/aether-and-hemeras-led-installation-uses-artificial-intelligence-to-transform-human-voices-into-a-brass-band-light-show/.

Russell, Stuart J., Peter Norvig, and Ernest Davis. Artificial Intelligence: A Modern Approach. 3rd ed. Prentice Hall Series in Artificial Intelligence. Upper Saddle River: Prentice Hall, 2010.

Schwab, Katharine. “Why This Artist Ground His Computer into Dust.” Fast Company. April 12, 2019. https://www.fastcompany.com/90333795/why-this-artist-ground-his-computer-into-dust.

Simonite, Tom. “Artist Refik Anadol Turns Data Into Art, With Help From AI.” Wired, January 16, 2020. https://www.wired.com/story/artist-refik-anadol-turns-data-art-help-ai/.

Square Enix. The Turing Test (version Microsoft Windows, Xbox One). 2016.

Sterling, Bruce. “Diffusion Choir.” Wired, November 27, 2018. https://www.wired.com/beyond-the-beyond/2016/10/diffusion-choir/.

Turing, Alan. “I.—COMPUTING MACHINERY AND INTELLIGENCE.” Mind LX, no. 236 (1950): 433–60. https://doi.org/10.1093/mind/LIX.236.433.

The Imitation Game. Directed by Morten Tyldum. Black Bear Pictures, FilmNation Entertainment, Bristol Automotive, 2014.

Vincent, James. “A neverending stream of AI art goes up for auction.” The Verge, Mar 5, 2019. https://www.theverge.com/2019/3/5/18251267/ai-art-gans-mario-klingemann-auction-sothebys-technology.

Weiser, Mark, Rich Gold, and John Seely Brown. “The Origins of Ubiquitous Computing Research at PARC in the Late 1980s.” IBM Corp. 1999. https://doi.org/10.1147/sj.384.0693.

Association for Public Art (blog). “What Is Public Art?” Accessed March 28, 2020. https://wwwassociationforpublicart.org/what-is-public-art/.

Wilson, Mark. “The Coolest Thing About 1 World Trade Center’s New Observatory Isn’t The View.” Fast Company. May 29, 2015. https://www.fastcompany.com/3046750/the-coolest-thing-about-1-world-trade-centers-new-observatory-isn-t-the-view.

Zebracki, Martin. “Beyond Public Artopia: Public Art as Perceived by Its Publics.” GeoJournal 78, no. 2 (2013): 303–17. https://doi.org/10.1007/s10708-011-9440-8.
Eitan Mendelowitz
Mount Holyoke College
emendelo@mtholyoke.edu

As a transdisciplinary artist, Eitan explores the cultural implications of the algorithmic creation of meaning while exposeing the wonder inherent in the generation of knowledge. His process blends performance, generative literature, gameplay, installation, and public art, with embodied interaction, physical interfaces, and artificial intelligence to create works situated at the intersections of computer science and the arts. His work articulates the expressive potential of artificial intelligence as a newly emerging cultural form.

Eitan’s work has been shown internationally at venues including SIGGRAPH, ArtFutura, ArsElectronica, and the Beall Center. Eitan is a Visiting Assistant Professor at Mount Holyoke College where his creative research and art production inform his teaching of algorithmic art and computer science. Eitan holds a Ph.D. (2009) from UCLA in computer science and an MFA (2002) in design | media arts.
ARTICLE

NODE «AI, ARTS & DESIGN: QUESTIONING LEARNING MACHINES»

Generative Art: Between the Nodes of Neuron Networks

Bruno Caldas Vianna
University of the Arts, Helsinki

Date of submission: March 2020
Accepted in: June 2020
Published in: July 2020

Recommended citation

Caldas Vianna, Bruno. 2020. «Generative Art: Between the Nodes of Neuron Networks». In: Andrés Burbano; Ruth West (coord.) «AI, Arts & Design: Questioning Learning Machines». Artnodes, no. 26: 1-8. UOC. [Accessed: dd/mm/yy]. http://doi.org/10.7238/a.v0i26.3350

Abstract

This article uses the exhibition “Infinite Skulls”, which happened in Paris in the beginning of 2019, as a starting point to discuss art created by artificial intelligence and, by extension, unique pieces of art generated by algorithms. We detail the development of DCGAN, the deep learning neural network used in the show, from its cybernetics origin. The show and its creation process are described, identifying elements of creativity and technique, as well as question of the authorship of works. Then it frames these works in the context of generative art, pointing affinities and differences, and the issues of representing through procedures and abstractions. It describes the major breakthrough of neural network for technical images as the ability to represent categories through an abstraction, rather than images themselves. Finally, it tries to understand neural networks more as a tool for artists than an autonomous art creator.

Keywords

generative art, machine learning, artificial intelligence, representation, algorithms, aura, neural networks
Arte generativo: entre los nodos de las redes neuronales

Resumen
Este artículo utiliza la exposición “Infinite Skulls”, que se inauguró en París a principios de 2019, como punto de partida para hablar sobre el arte creado por la inteligencia artificial y, por extensión, sobre las piezas de arte únicas generadas por algoritmos. Se centra en el desarrollo de DCGAN, la red neuronal de aprendizaje profundo utilizada en el programa, desde su origen cibernético. Se describe el espectáculo y su proceso de creación, identificando elementos de creatividad y técnica, así como la autoría de obras basadas en código abierto, en particular “Edouard de Belamy”, la pintura realizada a partir de inteligencia artificial que se vendió en una subasta de Christie’s por 432 000 dólares estadounidenses. También se enmarcan estos trabajos en el contexto del arte generativo, señalando afinidades y diferencias, así como los problemas de representación mediante procedimientos y abstracciones. Describe el gran avance de la red neuronal para imágenes técnicas como la capacidad de representar categorías a través de una abstracción, en lugar de imágenes en sí mismas. Por último, trata de entender las redes neuronales como una herramienta para los artistas, más que como una obra de arte per se.

Palabras clave
arte generativo, aprendizaje automático, inteligencia artificial, representación, algoritmos, aura, redes neuronales

Introduction
An early 2019, digital artist Robbie Barrat and painter Ronan Barrot collaborated on an exhibition (Bailey 2019) that took place at the L’Avant Gallery Vossen, in Paris. The show consisted on Barrot’s skull paintings, shown side by side with Barrat’s artificial intelligence-based reinterpretation of them. This collaboration is an opportunity to discuss current issues around the production of artworks which rely on machine learning tools. In 2018, an image produced by the Obvious collective with these methods was sold for 432.000 US dollars at the auction house Christie’s in New York (Jones 2018), and several exhibitions in different countries showed machine learning crafted work. How and why did we end up wanting machines to do the work of painters? And are these works mere replicas of great paintings or original works of their own?

Although some of the questions that arise go beyond the possible scope of the article, we would like nonetheless to take the opportunity to understand the processes and the history behind machine learning-based visual arts. We will start by describing the development of the technologies behind the tools used, the development of cybernetics and the quest for autonomy in computing; and then we’ll describe the collaboration and the exhibition itself. Finally we will wrap up mapping a few important questions on how can neural networks be framed within the generative art field, and why does it represent a paradigm change in procedure-based art.

From Cybernetics to DCGANs
The history behind the techniques used in the exhibition may be traced to the beginnings of computer science and the birth of cybernetics. In the founding book of this science, published in 1948, Norbert Wiener states clearly the goals of replacing mankind’s mental and physical workforce by machines. He even acknowledges the inherent risks of this replacement to the point of sharing his concerns about it with labor unions (1965, p. 27-28). He also makes clear references that the key to developing autonomous entities included the study of nature – beginning by the very title “Cybernetics - Or Control and Communication in the Animal and the Machine.”

It is not by chance that the first model of a neuron had been developed just a few years earlier, by McCulloch and Pitts in 1943. In the paper that describes it, the authors note how “the activity of the neuron is an ‘all-or-none’ process” (1943). At this time, the standard binary system for computers had not been established; computers themselves were more of an imagined device than an actual tool. The contraction bit (for binary digit) first appeared in Claude Shannon’s information theory paper in 1948 (1948) and Wiener stated that “in accordance with the policy adopted in some existing apparatus of the Bell Telephone Laboratories, it would probably be more economical in apparatus to adopt the scale of two for addition and multiplication, rather than the scale of ten” (1949, p. 4). In other words: we have, on one side, a biological feature present in most animals and, in the
other side, a mechanical device, and both of them are converging in their way of organizing and processing information.

The next natural step was to think of ways to model — and replicate, if possible — the complex structures of the brain, made up of billions of neurons interconnected in a dense web. A fundamental piece of the evolution of image based artificial intelligence was put in place in 1957 with the development of the perceptron by Frank Rosenblatt (1957). Although the model was first simulated on an IBM 704 computer, Rosenblatt would soon build a special purpose, analog hardware implementation. This apparatus used a grid of 20 by 20 photocells that took the role of a camera or an eye, connected in a network that would store data in potentiometers activated by electric motors. This network would be trained by placing different shapes in front of the camera, and the weights stored could then be used to recognize such shapes (Bishop 2006, 192–196).

This seemingly complex approach was actually criticized, with dire consequences, for being too simple. In a book published in 1969, Marvin Minsky and Seymour Papert (2017) acknowledged the potential of the perceptron, but also demonstrated that Rosenblatt’s idea had some concerning limitations. One particularly damaging feature they showed was the fact that the model wouldn’t be able to compute one specific logic function, named XOR. As we’ll see, neural networks today are also a method to provide a general solution to equations that are non-linear, i.e., not solvable by calculus tools. The fact that it wouldn’t be able to handle one of the most basic equations in computer science seriously undermined the reputation of the model and helped stall the development of neural networks in general for more than a decade — a period that came to be known as artificial intelligence’s winter (Crevier 1993).

What is curious is that, despite the pessimism towards perceptrons brought by Minsky and Papert, the solution to their limitations was pointed in the book itself — the multilayer perceptron. This model, which is still in use by most models in neural nets today (including the one used by Barrat) consisted of connecting the output of a layer of neurons to the input of another layers. The depth (number of layers) of a network can be defined according to the needs of the problem to be tackled, but Minsky and Papert showed that a three-layered net could already implement a XOR function. It should be also mentioned that putting these models to work with the memory and processing power of the devices of that time would be already discouraging.

In some way, this winter also reached the artists that showed interest in cybernetic processes. In the fifties, artist Nicholas Schöffer was heavily influenced by Wiener’s ideas. CYSP 1, created in 1956, might be the first piece of art that explicitly embedded the cybernetic concept of feedback into the artwork (Popper 2007). Many others followed suit, like the systems-inspired works of Hans Haacke and Jack Burnhan. Yet in seventies these two exponents of cyberart moved away from the field, possibly criticizing the relation between this science and the industrial military complex or the own “enclosed” aspect of the gallery (Lynch 2018).

The rebound of artificial intelligence would arrive only in the eighties. A movement named connectionism, that approached the cybernetics from a cognitive perspective, tried to apply neural networks to cognition models. This movement established two breakthroughs in deep learning that are still in use as of today: distributed representation, a strategy that breaks up inference tasks into smaller networks; and the use of back-propagation algorithms to train the net. Back-propagation, in fact, is a nothing more than a sophisticated feedback method — a concept already developed by Wiener in the cited book. Nevertheless, by the mid 90’s, AI failed once more to deliver practical applications and the second wave of neural networks ended. It was only in 2006 that another breakthrough would spark its third wave, which we are still riding at of today. Researcher Geoffrey Hinton developed a strategy named greedy layer-wise pre-training to perform what came to be called “deep learning” of the network (Goodfellow and Courville 2016).

Deep learning evolved rapidly after that, with several small but important inventions that made it into one of the richer and most complex fields in contemporary mathematics. It is important to notice that although neurological science will always be considered the original inspiration for neural networks, the brain is no longer an useful reference for AI scientists today. This is due to the fact that beyond simple models of one or a few neurons, it is very hard to analyze the full complexity of the brain and create models to replicate it. At the same time, mathematicians have been developing fantastic methods for different tasks that have nothing related to neurology — at least as an inspiration (Ibid).

In the early 2010’s, visual machine learning was still focused mostly in creating tools to recognize shapes and objects, inferring information about the environment in general. Many of such techniques were developed as part of a field named Computer Vision — therefore still using the eye and human senses as an inspiration. It was in 2014 that Ian Goodfellow, a PhD candidate at the University of Montreal, developed a strategy that would allow deep learning networks to create images instead— the generative adversarial networks (GANs). It uses two competing networks: one named the discriminator, and another one named the generator. The generator will be constantly creating images, which the discriminator will try to evaluate as being from the learning set of images or a counterfeit. The method turned out to be extremely successful. While there is some controversy in the sense that the idea might have been developed by others as early as in 2009 (Schmidhuber 2019), it is certain that the method only became widespread after the publication of Good-fellow’s paper (2014).

The last puzzle piece of the technique used by Barrat came out in the following year, complementing Goodfellow’s GAN with a deep convolutional method. Convolution networks employ the mathematical operation by that name, which happens to be very good in stitching...
pieces of images together. It is directly connected to the idea of distributed representation described before, where tasks are broken down in smaller networks (or image areas). Although the paper with the method came out in 2016 (Radford and Chintala 2016), the first development of the code could be found already in 2015 (Chantala 2015).

The development of GANs hasn’t stopped there, though. Several new models have been created since then. In the end of 2019, for instance, StyleGAN2 was released and proved itself able to create very realistic high resolution images.

### Infinite Skulls

The main primary source of information for this section comes from the videos posted on the Parisian gallery’s own site. This trove contains small pieces documenting the creation process, a short lecture by Robbie Barrat on the techniques he uses, an hour-long public discussion on AI and Art with the him, the painter, the curator and a lawyer specialized in copyrights. The exhibition lasted only from the 7th until the 11th of February, 2019. It spanned a few rooms, divided in “first and second epochs” (that will be clarified later), a video piece that showed loops of the images being generated and an interactive peephole piece, which showed an unique painting and then immediately erased it forever.

The initiative for the show came from gallerist Catherine Vossen and artist Albertine Meunie, who realized that the obsessive paintings of skulls by painter Ronan Barrot could become a valuable initial set of images for deep learning processes (Gatti 2019, 0:22:00). He has been painting hundreds of those since 2011, and they all share some characteristics such as the orientation and size of the head. Robbie Barrat, on the other hand, was using DCGAN for several of his works, which needed to be fed with thousands of data sets (images in this case) to generate new sets that will resemble the originals. Please notice the irony of needing a huge amount of unique but similar images created by repetitive work to generate endless unrepeatable but similar machine made images.

Barrat used to collect public domains paintings belonging to specific categories such as abstract landscapes and nude portraits, using the technique named *scraping* to massively download them from the web (Barrat 2017). When he accepted the invitation to collaborate on a show with Barrot, they explored possible ways of working with data from the painter, even using the painter’s visual references and influences as in input (Gatti 2019, 1:08:00). But they ended up choosing the direct approach of feeding the skulls straight into the network. The consequence of this choice was that the first batch of results came out remarkably similar to the original artwork, to the point of Barrot saying that he wish we would have drawn some of those himself (Avant Gallerie Vossen 2019).

Interestingly enough, at this point, the digital artist decides that the mere replication of the paintings wasn’t enough for him to consider that as his own artwork. These images remained on the show in what was called the “first epoch” rooms, borrowing a term used to designated different periods of the learning done by the network. Nevertheless, Robbie decided to manipulate the input set by stretching, rotating the original and produce a second batch in which he could see himself. This selection was exhibited as the “second epoch.”

This attitude is repeated throughout Barrat’s discourse. When asked how to make the generated works to go beyond the mere repetition of original, his answer is to provoke misinterpretation, to confuse the network with disparaging data: “We must confuse the machine and make it hallucinate a bit” (Gatti 2019, 0:26:48). He also states that his work is mostly comprised of choosing and curating: the choice of the artworks that will feed the database, and curating the output to pick the most interesting results. Barrot, on his side, describes the authorship in these terms: “I chose, Robbie chose, Albertine and Catherine (…) In the end, the authors are all four of us.” While it is true that there are plenty of discussions in contemporary art regarding the role of the curator as author (or at least co-author) (Lubar 2014), here this duality is exacerbated by the introduction of the machine as a producer of endless choices. We can start to identify the emergence of a particular category of computational artist dedicated to artificial intelligence, with several examples. An inventory of those would, unfortunately, fall outside of the scope of this article.

Barrat, here, is actively playing his role as an artist – not only by making choices on what to feed the algorithm, but also creatively criticizing the output and proposing new iterations. As we will see, selection and curatorship are fundamental skills for creators using GANs.

### Generative art and representation

To completely grasp the effect of neural networks in visual expression, we must understand machine learning art as a special case of generative art.

Generative art is closely connected with the idea of using rules and constraints for a creative outcome. We can see it as an essentially algorithmic practice, where the artist’s role is to define a method, more than to manually craft a final piece. Philip Galanter (2003) proposed the following definition: “Generative art refers to any art practice where the artist uses a system (...) which is set into motion with some degree of autonomy contributing to or resulting in a completed work of art.”

It has existed before and independently from computers. Sol Lewitt, for instance, wrote sets of instructions so that some of his drawings could be executed by anyone. Some even argue that Jackson Pollock’s method of action painting could be defined as
algorithmic (Boden and Edmonds 2009). It can also be said that to some extent, every art is algorithmic, in sense that it involves some set of procedures.

Two ideas around generative visual art will be important to our GAN discussion. The first is the fact that the generated art object can be multiple, that is, the same underlying algorithm can produce an infinity of unique works. This might happen for handcrafted works because the instructions can be interpreted in different ways, because the personal skills of each executioner are different and also because of the own nature of handcraft – the same person, following the same instructions, will probably create a different-looking work every time. In computer-based works, the uniqueness must be coded by means of random parameters or heuristic methods.

The other idea is that generative art detaches the creator of the work from its manufacture. In that sense, it becomes very close to conceptual art – and in fact, in this context, an algorithm is a way of representing a concept. This adds to the creative process a seductive notion of democratizing the artwork creation, since anyone could follow the rules of the artist and create his or her unique installment.

These two thoughts are also related to Walter Benjamin’s (2008) reflections of about one hundred years ago, when he wrote about the work of art at the age of mechanical reproduction. The advent of mechanographic methods, in special photography and cinema, afforded the creation of numerous copies of an original, and inspired Benjamin to introduce the concept of the aura, the here and now of the work of art, connected to its uniqueness and history. “Reproduction (…) substitutes a mass existence for an unique existence.” And now we not only face the mass reproduction of originals, but also the mass creation of unique pieces. Can we call these generated pieces originals? What kind of anti-aura does the mechanical generation of art afford? “(…) the destruction of the aura (…) is the signature of a perception (…) so increased that, by means of reproduction, it extracts sameness even from what is unique.” Are we now doing the reverse, that is, obtaining uniqueness from what is the same?

It is clear that generative art has permitted a displacement of the “originalness” from the concept towards the mass produced artwork. But some sort of aura still resides in the mold, in the negative, there where Benjamin found uniqueness. Therefore, the singularity of visual generative art must reside in the system proposed by the artist: the set of procedures, the instructions, the method. If we want to understand GANs and generativity, we must look at its hermeneutics, specially at what is being translated. As McCormack et al (2014) asked, “In what sense is generative art representational and what is it representing?”

Artificial neural networks have been a method to find approximate solutions for math functions much before their application in visual art. The Cybenko (1989) theorem offered one of the first proofs of this strategy. A clear example of how they can be used to compute any function can be found in chapter 4 of Nielsen (2015).

Similarly, generative adversarial nets can be used to model – if not formalize – any visual aesthetics. How does that happen? The process of training a GAN consists of feeding it a number of pictures that belong to one particular coherent category. Ideally this number should be in the order of thousands. The generator network starts with images comprised of pure noise, while the discriminator will try to tell if the generated images belong to the fed category or not. In a process that can take from hours to months, depending on many factors, this feedback process will result in a gigantic statistical model of the input images. This is not, as in the Klee and Mondrian examples, a procedural method of steps to reproduce a given style. It is rather a representation comprised of infinite dormant images that can be brought to surface. Machine learning allowed us to do without procedural strategy, giving us an universal tool that forgoes the development of one algorithm for every style.

The person exploring the latent domain of a GAN can do it much the way a flanêur discovers a city, except that the space visited has much more than 2 or 3 dimensions. If the network was trained with human faces, he or she can stumble onto the neighborhood of an oriental child, that could perhaps be not far from a teenager with native American traces. In any case, no face will be equal to another,
and ideally – if the training wasn’t overfitted – will also be different from any of those from the input set.

So while both procedural approaches may also be used to create a representation of some aesthetics, machine learning has the advantage of being able to do it for any coherent visual category, much like it enabled the solution of equations that could not be solved by classic methods of mathematical calculus. And while there are underlying procedural methods, as with any computational process, what is being translated in artificial intelligence process is no longer a recipe, but an abstraction, a latent space of endless potential new images - all coherent with a proposed aesthetic style.

As such, generative neural networks could be the greatest shift in the way we fixate images since the photography; if the 19th century invention gave us the ability to represent (and endlessly reproduce) any object, the neural networks has given us the potential to represent (and uniquely instantiate) an abstraction of any coherent category. In the network, the image becomes a sign, an operation analog to language: just like the word book resonates multiple real things that belong in the book category, the net contains endless instances of a given model.

**Conclusion: a tool is just a tool**

We are about to enter the 2020’s, and GANs have given humankind an endless original making machine. The Next Rembrandt project generates paintings that I’d be glad to hang in my house: they could pass perfectly for a piece of the Dutch master and no one else would have a copy. It is art and it is generative, according to the definition we found. But whose art is it? The author-ship of these works resides in a limbo: while it is clearly not a Rembrandt, it could also not be signed exclusively by the engineers who designed the system. In that sense, it is very different from the works of Nake and Noll. Is it original, even if it is unique? And what artistic value does it entail?

These are not questions that can be answered shortly. But we certainly can find some clues. When Robbie Barrat saw the output of network trained with Barrot’s skulls, he realizes that this is not his work – and moves on to manipulate them, so to claim it as his own expression. The painting sold by Obvious may be inspired by 19th century portraits, but could never be mistaken for one.

Procedural generation afforded the shift from visual representation of actual things to the abstraction of visual sets of things. These processes, however, were case-dependent, and couldn’t be applied to any set. Machine learning is the first method that can be generalized to any coherent style, opening great new territories for artistic exploration. It is clear now the mimicking capabilities of GANs and neural networks in general have great value for scientific and commercial endeavors. But how valuable is it for art itself?

Probably the most intriguing and disturbing images created by machine learning are far from imitations and are only comparable to replicas of nightmares – revealing, perhaps, their neurological features and original inspiration. Some studies in cognitive visual perception proposed statistical models for peripheral vision that are eerily similar to low resolution domains of adversarial networks (Cohen et al. 2016).

Machine learning revealed itself to be a fantastic tool in chartering new frontiers of visual art. But it is most interesting when it becomes a tool for creators, instead of a mechanic, replacement artist. Art is probably what leaks through the gaps between the nodes of the net when the artist makes it “hallucinate a bit.” Its most creative and innovative works delve not into the impressive forgeries made by GANs, but into the mistakes they make, into the imperfect output of poorly trained generators, their glitches and into their nightmarish features.

**References**

Avant Galerie Vossen. 2019. Barrat / Barrot, Infinite Skulls #7 // “J’aurais aimé le faire celui-là” Accessed in March 16th, 2020. https://vimeo.com/309000957.

Bailey, J. 2019. “AI Artist Robbie Barrat and Painter Ronan Barrot Collaborate on ‘Infinite Skulls.’” Accessed March 16th, 2020. https://web.archive.org/web/20190206121643/www.artnome.com/news/2019/1/22/ai-artist-robbie-barrat-and-painter-ronan-barrot-collaborate-on-infinite-skulls.

Barrat R. 2017. art-DCGAN. Accessed in March 16th, 2020. https://web.archive.org/web/20200316134023/https://github.com/robbiebarrat/art-DCGAN/tree/92f8eb63f12e1.

Benjamin, W. 2008. The work of art in the age of mechanical reproduction. Penguin UK.

Bishop, C. M. 2006. Pattern recognition and machine learning. Spring-er.

Boden, M. A., & Edmonds, E. A. 2009. “What is generative art?” Digital Creativity, 20(1–2), 21–46. https://doi.org/10.1080/14626260902867915.

Chantala, S. 2015. Dcgan.torch – soooo much win. Accessed in March 16th, 2020. https://web.archive.org/web/20191022113504/https://github.com/soumith/dcgan.torch/commit/45fd6727c3cda67f2e357aab6e7eaa57ad9209.

Cohen, M. A., Dennett, D. C., & Kanwisher, N. 2016. “What is the bandwidth of perceptual experience?” Trends in cognitive sciences, 20(5), 324-335. https://doi.org/10.1016/j.tics.2016.03.006.

Crevier, D. 1993. At: the tumultuous history of the search for artificial intelligence. Basic Books.

Cybenko, G. 1989. “Approximations by superpositions of sigmoidal functions”, Mathematics of Control, Signals, and Systems 2(4), 303–314. https://doi.org/10.1007/BF02551274.
Galanter, P. 2003. “What is generative art? Complexity theory as a context for art theory.” In GA2003–6th Generative Art Conference. Gatti, E., Barrat R., Barrot, R, Degoulet C., Meunier A. 2019. Discussion about Art and Artificial Intelligence. Video recorded in 2019. Accessed in March 16th, 2020. https://vimeo.com/325843365.

Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., & Bengio, Y. 2014. “Generative adversarial nets”. In Advances in neural information processing systems (pp. 2672-2680).

Goodfellow, I., Bengio, Y., & Courville, A. 2016. Deep learning. MIT press.

Jones, J. 2018. “A portrait created by AI just sold for $432,000. But is it really art?” The Guardian, October 26, 2018. https://web.archive.org/web/20190220004707/https://www.theguardian.com/artanddesign/shortcuts/2018/oct/26/call-that-art-can-a-computer-be-a-painter.

Lubar, S. 2014. “Curator as Auteur”. The Public Historian, 36(1), 71-76. https://doi.org/10.1525/tph.2014.36.1.71.

Lynch, G. 2018. “The transformative nature of networks within contemporary art practice.” Doctoral dissertation, London South Bank University.

McCormack, J., Bown, O., Dorin, A., McCabe, J., Monro, G., & Whitelaw, M. 2014. “Ten Questions Concerning Generative Computer Art” Leonardo, 47(2), 135–141. https://doi.org/10.1162/LEON_a_00533.

McCulloch, W. S., & Pitts, W. 1943. “A logical calculus of the ideas immanent in nervous activity.” The bulletin of mathematical biophysics, 5(4), 115-133. https://doi.org/10.1007/BF02478259.

Minsky, M., & Papert, S. A. 2017. Perceptrons: An introduction to computational geometry. MIT press. https://doi.org/10.7551/mitpress/11301.001.0001.

Nielsen, M. A. 2015. Neural networks and deep learning (Vol. 2018). San Francisco, CA, USA:: Determination press.

Noll, A. M. 1966. “Human or machine: A subjective comparison of Piet Mondrian’s ‘Composition with Lines’ (1917) and a computer-generated picture.” The psychological record, 16(1), 1-10. https://doi.org/10.1007/BF03393635.

Popper, F. 2007. From technological to virtual art. Mit Press.

Radford, A., Metz, L., & Chintala, S. 2015. “Unsupervised representation learning with deep convolutional generative adversarial networks.” arXiv preprint arXiv:1511.06434.

Rosenblatt, F. 1957, “The Perceptron-a perceiving and recognizing automaton.” Report 85-460-1, Cornell Aeronautical Laboratory.

Schmidhuber, J. 2019. “Unsupervised Minimax: Adversarial Curiosity, Generative Adversarial Networks, and Predictability Minimization.” arXiv preprint arXiv:1906.04493.

Shannon, C. E. 1948. “A mathematical theory of communication.” Bell system technical journal, 27(3), 379-423. https://doi.org/10.1002/j.1538-7305.1948.tb01338.x.

Smith, G. W. “An Interview with Frieder Nake.” Arts, 8(2), 69. https://doi.org/10.3390/arts8020069.

Wiener, N. 2019. Cybernetics or Control and Communication in the Animal and the Machine. MIT press. https://doi.org/10.7551/mitpress/11810.001.0001.
Bruno Caldas Vianna
University of the Arts, Helsinki
bruno.caldas@uniarts.fi

Bruno Caldas Vianna lives in Barcelona and is pursuing a PhD from Uniarts in Helsinki in Visual Arts and Machine Learning. He studied Computer Engineering but graduated in Film Studies. He has a master’s from NYU’s Interactive Telecommunications Program. He creates visual narratives using classical and innovative supports, having directed short and feature films, as well as working in live cinema, augmented reality, mobile applications and installations. From 2011 until 2016 he ran Nuvem, a rural art laboratory and residency space, located between Rio and São Paulo, and he worked as a teacher at Oi Kabum! art and technology school in Rio until 2018.
ARTÍCULO

NODO «IA, ARTE Y DISEÑO: CUESTIONANDO EL APRENDIZAJE AUTOMÁTICO»

Vistas panorámicas sobre el patrimonio visual colectivo a través de redes neuronales convolucionales

Las exposiciones Revolutionary Arkive y Mnemosyne 2.0 de Pilar Rosado

Pilar Rosado Rodrigo
Universidad de Barcelona

Ferran Reverter Comes
Universidad de Barcelona

Fecha de presentación: marzo de 2020
Fecha de aceptación: junio de 2020
Fecha de publicación: julio de 2020

Cita recomendada

Rosado Rodrigo, Pilar y Reverter Comes, Ferran. 2020. «Vistas panorámicas sobre el patrimonio visual colectivo a través de redes neuronales convolucionales. Las exposiciones Revolutionary Arkive y Mnemosyne 2.0 de Pilar Rosado». En: Andrés Burbano; Ruth West (coord.) «IA, arte y diseño: Cuestionando el aprendizaje automático». Artnodes, N.º 26: 1-12. UOC. [Consulta: dd/mm/aa]. http://doi.org/10.7238/a.v0i26.3354

Los textos publicados en esta revista están sujetos –si no se indica lo contrario– a una licencia de Reconocimiento 4.0 Internacional de CreativeCommons. La licencia completa se puede consultar en https://creativecommons.org/licenses/by/4.0/deed.es_ES.
Resumen
En el contexto de una sociedad saturada de imágenes, las redes neuronales convolucionales (convolutional neural networks - CNN), preentrenadas a partir de la información visual contenida en miles y miles de imágenes, constituyen una herramienta de gran utilidad para ayudarnos a ordenar el patrimonio visual, ofreciéndonos así un punto de acceso que de otra manera sería imposible.

Una de las responsabilidades del artista contemporáneo es adoptar posiciones que ayuden a dar significado, a proyectar sentido sobre la acumulación visual a la que nos enfrentamos. En este artículo pasamos a describir dos exposiciones de la artista Pilar Rosado: Revolutionary Arkive y Mnemosyne 2.0. En ellas se ha utilizado la red neuronal artificial ResNet-50 para extraer las características visuales de grandes conjuntos de imágenes y los descriptores de imagen obtenidos se han usado como entrada para el algoritmo t-SNE. Así, se han elaborado grandes mapas visuales formados por miles de imágenes ordenadas siguiendo criterios de similitud formal en los que se ponen de manifiesto los patrones visuales de los arquetipos de una determinada categoría semántica.

La manera de archivar y recuperar nuestra memoria colectiva debe tener una correlación con los avances tecnológicos y científicos de nuestra época para que se nos vayan descubriendo progresivamente nuevos horizontes de conocimiento. En su práctica artística, Rosado explora, en el ámbito de la cocreación humano-máquina, cuestiones políticas que se pueden abordar desde la imagen y que implican a las tecnologías de aprendizaje automático, como la gestión de la información en los archivos visuales del futuro o la revisión de la memoria visual colectiva.

Palabras clave
Deep learning, postfotografía, computer vision, t-SNE, convolutional neural networks (CNN), Aby Warburg.

Panoramic views on the collective visual heritage through convolutional neural networks.
The exhibitions Revolutionary Arkive and Mnemosyne 2.0 by Pilar Rosado

Abstract
In the context of a society saturated in images, convolutional neural networks (CNNs), pre-trained from the visual information contained in thousands and thousands of images, constitute a tool which is of great use for helping us to order visual heritage, in this way offering us a point of access which would otherwise be impossible.

One of the responsibilities of the contemporary artist is to adopt positions which can help to provide sense, to cast meaning onto the accumulation that we face. In this article we describe two exhibitions by the artist Pilar Rosado: Revolutionary Arkive and Mnemosyne 2.0. The artificial neuronal network ResNet-50 has been used in these exhibitions in order to extract the visual characteristics of large sets of images, and the image descriptors obtained have been used as an entry point for the algorithm t-SNE. In this way, large visual maps have been produced which are composed of thousands of images and are arranged by following criteria of formal similitude in which the visual patterns of the archetypes of a certain semantic category are displayed.

The method of archiving and recovering our collective memory must have a correlation with the technological and scientific advances of our age so that we can progressively discover new horizons of knowledge. In her artistic practice Rosado explores, in the sphere of human-machine co-creation, political issues which can be approached from the perspective of image and which involve the machine learning technologies, such as the management of information in the visual files of the future or the review of the collective visual memory.

Keywords
Deep Learning, Post-photography, Computer Vision, t-SNE, Convolutional Neural Networks (CNNs), Aby Warburg.
Torrente visual

«…organizarse ya es, en cierto modo, tener ojos»
(Saramago, 2003: 330).

Cada minuto transcurrido, millones de imágenes son tomadas y vertidas a la red, como si de un río de infinito caudal se tratase. Con ellas abocamos también el rastro de nuestros gestos, muchas veces compartidos. La dificultad de rescatarlas de este torrente y de acceder a su visualización de una manera inteligible puede convertirlas en invisibles. Estos enormes archivos visuales pueden volver a algún tipo de comprensibilidad gracias al uso de tecnologías de deep learning.

La posibilidad de poder detener este incesante y vertiginoso flujo de información creando una instantánea que nos ayude a reflexionar, desde lo visual, sobre lo que hemos producido, especialmente sobre la «forma» de las imágenes que hemos producido, se vislumbra como una poderosa herramienta para pensar. La belleza de este tipo de visualización de datos estará implícita en los rastros de verdad que puedan destilarse de su análisis.

En el contexto de esta sociedad de la sobreinformación y de la enorme producción fotográfica, una de las responsabilidades del artista es adoptar posiciones que ayuden a dar significado, a proyectar sentido sobre la acumulación visual a la que nos enfrentamos. Acerca de este tema, Joan Fontcuberta escribe: «El valor de creación más determinante no consiste en fabricar imágenes nuevas, sino en saber gestionar su función, sean nuevas o viejas» (Fontcuberta, 2016: 39).

Sitúamos nuestra propuesta en el ámbito de lo postfotográfico y proponemos que, aplicando modelos de deep learning para la categorización automática de grandes cantidades de imágenes de archivo, podemos darles un orden, una belleza que nos muestre de alguna manera las múltiples reencarnaciones de la imagen», como las llama Fontcuberta.

Antecedentes

La genealogía de estos intereses se remonta al historiador del arte y la cultura Aby Warburg (1866-1929) quien, con su Atlas Mnemosyne, formó por una colección de imágenes y escaso texto, pretendía narrar la historia de la memoria de la civilización europea. Con una visión premonitoria Warburg considera que las imágenes son una de las formas más poderosas mediante las cuales el hombre organiza su relación con el mundo y su relación con el tiempo. Un atlas como Mnemosyne, que supone una visión tambaleante del tiempo, es una forma de reflexionar sobre la manera en que el tiempo y el espacio se articulan en la mente humana (Warburg, 2010).

También podemos citar a André Malraux, ministro francés de cultura y escritor, que en 1947 lanza la idea de crear un museo imaginario basado en reproducciones fotográficas, sin paredes, sin muros, permitiendo la posibilidad de que cada persona de formar su propio museo (Malraux, 1949). A pesar de limitaciones insalvables como la reproducción de la textura, el volumen o la escala de una obra de arte, y teniendo claro que una experiencia artística válida nunca podrá substituir la vivencia en el lugar, la ventaja de este tipo de museo imaginario es que cada espectador puede crear su propio museo en función de su sensibilidad.

Podemos, así mismo, conectar con la llamada fotografía aumentada de autores como Lev Manovich con sus análisis visuales de obras literarias, videojuegos o las portadas de la revista Time. En la obra TimeLine, de Jeremy Douglass y Lev Manovich de 2009, vemos en una sola imagen las portadas de todos los números de la revista Time publicados entre 1923 y el verano de 2009, un total de 4.535. Las imágenes aparecen dispuestas de acuerdo con el orden de publicación (de izquierda a derecha y de arriba a abajo) (Manovich, 2009). También resulta adecuado mencionar el proyecto de Lev Manovich y William Huber de 2010 que permite visualizar la estética general del mundo del juego, además de los cambios visuales y narrativos dentro de la progresión de un solo juego, ordenando 22.500 frames tomados del videojuego japonés Kingdom Hearts II cada 6 segundos de juego. La obra nos muestra 62,5 horas de juego representadas en una sola imagen (Manovich, 2012).

En este marco presentamos dos proyectos de creación concebidos por Pilar Rosado con el soporte estadístico de Ferran Reverter, con quien colabora activamente en proyectos de investigación desde el año 2012. En primer lugar, se presenta la exposición Revolutionario Arkive: The eye-machine is watching, que tuvo lugar entre el 27 y el 30 de septiembre de 2018 durante el Festival Panoràmic en Granollers (Barcelona), donde se volvieron a revivir las revueltas del 1968 a través del cine y de la fotografía. En segundo lugar, se muestra el proyecto expositivo Mnemosyne 2.0: cartografías computacionales de la memoria visual, que se presentó entre el 9 de mayo y el 9 de junio de 2019 en la Sala 0 de Roca Umbert Fábrica de las Artes de Granollers (Barcelona) y que fue el resultado de la residencia artística que Pilar Rosado desarrolló durante el año 2019 en ese contexto.

En los próximos apartados pasaremos a describir estas dos exhibiciones artísticas, para cuya realización se utilizaron miles de fotografías recuperadas de internet clasificadas según los modelos de deep learning que explicaremos a continuación.

Métodos computacionales

Para recuperar las miles de imágenes utilizadas, Rosado realizó múltiples búsquedas textuales en Google y, para descargar las imágenes, accedió a los repositorios de internet utilizando software de web scraping incorporado como una extensión en el navegador.

No podemos obviar que los algoritmos de búsqueda de Google se diseñaron para ayudarnos a encontrar lo que buscamos entre millones de páginas web en una fracción de segundo. Estos sistemas de clasificación se basan no en uno, sino en una serie de algoritmos que tienen en cuenta muchos factores como las palabras
de la consulta, la relevancia y usabilidad de las páginas, el grado de especialización de las fuentes, la ubicación y la configuración. El peso que se da a cada factor cambia dependiendo de la naturaleza de la consulta (según sitio web Búsqueda de Google).

Desde la visión artificial se están haciendo muchos esfuerzos para que las máquinas emulen la habilidad que tenemos las personas de comprender una imagen. En la actualidad estamos viviendo un gran avance en el uso del deep learning en tareas de catalogación automática de imágenes.

El deep learning se fundamenta en la idea del aprendizaje mediante ejemplos dados. El ordenador, a partir de ellos, construye un modelo capaz de predecir reglas subyacentes. El algoritmo modifica el modelo cuando se producen errores de predicción, así, de manera iterativa, estos modelos pueden aprender de forma extremadamente precisa, gracias a que el sistema es capaz de extraer patrones relevantes para la tarea planteada.

La potencia del deep learning reside en la capacidad de estas metodologías de descubrir por sí mismas los descriptores de las imágenes sin necesidad de usar unos predefinidos. Manovich apuntó en 2019 que este aspecto del aprendizaje automático es, de hecho, algo nuevo en la historia del arte computacional; un ordenador que por sí solo puede aprender la estructura del mundo visual.

Con el objetivo de clasificar las imágenes, se ha utilizado la red neuronal convolucional ResNet-50 (He et al., 2016) que fue preentrenada en un subconjunto de la base de datos ImageNet (Russakovsky et al., 2015). Entrenada con más de un millón de imágenes, tiene 177 capas convolucionales en total, asociadas a una red residual de 50 capas, y puede clasificar imágenes en 1000 categorías de objetos.

Para clasificar las imágenes utilizadas en las exposiciones que comentaremos más adelante, se usó esta red ResNet-50 como un extractor de características visuales, tomando las activaciones en la capa de salida (fc1000) como descriptores artificiales de las imágenes de la colección, obteniéndose una representación inducida por la red del contenido visual de cada una de las imágenes. Así, podemos decir que esta etapa del proceso creativo está condicionada por el entrenamiento supervisado de la red ResNet-50.

Posteriormente, se emplearon estos descriptores artificiales obtenidos como entrada para el algoritmo t-SNE (stochastic neighbor embedding) (Maaten y Hinton, 2008). t-SNE representó los vectores 1000-dimensional de las imágenes en un espacio 2-dimensional (plano de representación) preservando óptimamente las relaciones de similitud de contenido visual de las imágenes (ver figura 1). Esta etapa del proceso creativo es no supervisada, ya que depende exclusivamente de las similitudes entre las imágenes que se están analizando.

Al trabajar con grandes conjuntos de fotografías, el problema es que, cuando se aplica t-SNE a las imágenes, tal como hemos comentado, resulta que una gran cantidad de las localizaciones de las imágenes en el plano de representación corresponden a elementos superpuestos que a menudo impiden inspeccionar una sola imagen. Para solucionar esta problemática se utilizó el código de Karpathy (2017), que re asigna las localizaciones bajo la suposición de vecindad basado en t-SNE (ver figura 2).
En el proceso de elaboración de ambas exposiciones, la toma de decisiones hasta llegar al resultado final es compartida entre el humano y la máquina, configurando un proceso de cocreación híbrido que alienta posibilidades novedosas para la obtención de resultados.

Exposición Mnemosyne 2.0: Cartografías computacionales de la memoria visual

Las obras que duermen en los archivos como documentos pueden volver a divertir (amuser) y excitar gracias a un análisis que nos las haga ver con ojos nuevos; un placer que estaría entre la plástica y el estudio riguroso (Pérez-Hita, 2019: 2).

En 1924 el historiador Aby Warburg inicia su Atlas Mnemosyne. Rastreando temas y patrones visuales recurrentes a través del tiempo, pasando por la antigüedad, el Renacimiento y llegando a la cultura contemporánea, Warburg pretendía elaborar algo así como una base de datos de imágenes de todos los tiempos y culturas. Es la suma del trabajo de toda su vida y quedó inacabado a su muerte en 1929. Las placas originales fueron hechas de tablas de madera de 1,7 x 1,4 metros cubiertas por lienzo negro en las que Warburg fijaba todo tipo de imágenes, como fotografías, reproducciones de pinturas, dibujos, ilustraciones extraídas de libros y material de periódicos. Cada tabla se dedicó a un tema específico que reflejaba sus intereses de investigación, estaban numeradas y eran fotografiadas en el gran salón de la biblioteca de Warburg. El autor no conservó las tablas originales, sino que constantemente reorganizaba las imágenes y agregaba nuevos tableros y temas (The Warburg Institute, 2018). Para Warburg las tablas, más que un medio de presentación, eran un medio de investigación y comprensión.

El proyecto expositivo Mnemosyne 2.0 de Pilar Rosado es un intento de actualizar, mediante las nuevas tecnologías, la idea que abordó Aby Warburg con su Atlas Mnemosyne. De una forma análoga a como Warburg, sentado en su mesa, debía trabajar para establecer relaciones visuales entre sus imágenes, la artista trabajó frente a su ordenador utilizando el recurso de búsqueda de información de Google. Para elegir sus temas de interés, Rosado estableció un paralelismo con la magnífica exposición de 2010 en el MNCARS de Google. Para elegir sus temas de interés, Rosado estableció un paralelismo con la magnífica exposición de 2010 en el MNCARS titulada «Atlas. ¿Cómo llevar el mundo a cuestas?», comisariada por Didi-Huberman, que se conformaba en cuatro grandes apartados titulados: «Conocer por imágenes», «Recomponer el orden de las cosas», «Recomponer el orden de los lugares» y «Recomponer el orden de los tiempos» (Didi-Huberman, 2010). La artista, recogiendo el testigo de Didi-Huberman, descartó imágenes de internet explorando conceptos y procesos tecnológicos actuales que han revolucionado los hábitos y espacios sociales, con la intención de poder revisar de forma crítica este patrimonio visual compartido. Para ello estructuró sus estrategias de búsqueda a partir de los criterios que describimos a continuación.

Para «recomponer el orden de las cosas», buscó imágenes de objetos producidos en los últimos tiempos por las nuevas tecnologías, realizando búsquedas de términos como «tecnología portable», «nanotecnología», «impresión 3D», «ciberseguridad», «robots»… Para «recomponer el orden de las personas», escribió en el buscador de Google el nombre de actividades realizadas por personas a raíz del uso de nuevas tecnologías como «hactivistas», «mineros del coltán», «ciborgs», «hacedores de selfies», «liquideradores de Chernóbil»… Para «recomponer el orden de los lugares», buscó espacios del planeta afectados por la «producción de tecnología» como «centrales nucleares», «cambio climático», «grandes basureros», «gentrificación»…

En el mundo contemporáneo, a pesar del espectacular desarrollo de los nuevos medios, la población tiene la necesidad de seguir manifestándose en la calle para propiciar cambios sociales y hacer valer sus derechos. Así que Rosado, con la intención de «recomponer el orden de los tiempos», buscó imágenes de «manifestaciones por los derechos humanos», «manifestaciones por el orgullo gay», «manifestaciones contra el cambio climático»…

En la exposición se mostraban también de forma relevante los textos de estas búsquedas (ver figura 3), que se iniciaron de una forma deliberada siguiendo los intereses de la artista, pero que pronto se vieron condicionados por los resultados de las propias búsquedas, en un proceso de serendipia que favoreció el descubrimiento o hallazgo afortunado a partir de una imagen que sugería una próxima búsqueda inesperada, pero valiosa. Se constituyó este método de creación cooperativo entre la artista y la computadora en un proceso estimulante de aprendizaje visual, de descubrimiento dirigido por la voluntad, pero también por la casualidad que deja abierta una puerta a la habilidad del sujeto que busca para reconocer que ha hecho un descubrimiento importante, aunque inesperado.

En el orden de los tiempos, en especial el de los nuevos medios, la población tiene la necesidad de seguir manifestándose en la calle para propiciar cambios sociales y hacer valer sus derechos. Así que Rosado, con la intención de «recomponer el orden de los tiempos», buscó imágenes de «manifestaciones por los derechos humanos», «manifestaciones por el orgullo gay», «manifestaciones contra el cambio climático»…

Figura 3. Imagen de todas las frases que la artista utilizó para recuperar en Google las 40.000 imágenes que muestra la exposición. Junto a cada frase aparece el número de imágenes que se obtuvo en esa búsqueda. © Pilar Rosado

Rosado comparte con Warburg la intención de crear una cartografía personal estableciendo relaciones y analogías que después de decenas y centenares de horas de búsqueda, el resultado que a ella se le ofreció no sólo ilustraba la experiencia de una vida interactiva, sino que, de paso, facilitaba la exploración de un universo cultural emergente.
permitan un proceso abierto e infinito de relecturas, pero se diferencia metodológicamente en el hecho de que establece un diálogo de cocreación en el que permite que sus criterios de búsqueda se vean estimulados e hibridados con los criterios de la computadora.

Una vez descargadas 10.000 imágenes para cada grupo de conceptos (cosas, personas, lugares y manifestaciones), se aplicaron sobre ellas los métodos de deep learning descritos en el apartado «Métodos computacionales» para conseguir tener todas las imágenes ordenadas en una gran fotografía de 2 x 2 metros. La intención de la artista era llegar a cristalizar en grandes mapas visuales sus pensamientos e intereses relacionados con la tecnología, mediante este deambular por la información visual contenida en este conjunto de imágenes extraído de internet.

Posteriormente, será el espectador quien realice sus propios recorridos visuales a través de estos mapas (ver figura 5), que constituyen un gran atlas visual o, como el subtítulo de la exposición sugiere, «cartografías computacionales de la memoria visual» contenida en internet acerca de conceptos relacionados con la tecnología.

Un atlas constituye una forma visual del saber. No es un diccionario, ni un manual, ni un catálogo. Proporciona un saber abierto: ante las imágenes, sobran las palabras. El atlas desvela una pasión por las morfologías, pero se aleja de la categorización estricta. En un atlas el lector puede recorrer con la mirada, deambular erráticamente para encontrar o no, tiene la posibilidad de construir sus propias cartografías (Didi-Huberman, 2010).

«Entiendo la fotografía como una herramienta para pensar la realidad y en Mnemosyne 2.0 propongo poner un poco de orden en el maremágnum fotográfico que nos rodea aplicando estrategias de catalogación automática para construir un atlas formado por imágenes hechas de imágenes, cartografías de nuestra memoria visual que nos ayuden a comprender un poco mejor los rastros, los patrones, los arquetipos que construimos como sociedad en nuestro quehacer fotográfico. Los resultados obtenidos nos permitirán revisitarnos la iconografía que constituirá en un futuro los archivos de nuestra memoria, desde un nuevo punto de vista más libre de preconcepciones e ideologías» (Rosado, 2019: 10).

Exposición Revolutionary Arkive: The eye-machine is watching

La exposición Revolutionary Arkive: The eye-machine is watching tuvo lugar en el contexto del Festival Panorámico. Bajo el lema The whole world is watching, la segunda edición del festival proyectó 7 largometrajes, 7 cortometrajes, organizó actividades, debates y conferencias y acogió las propuestas artísticas de 46 creadores (Festival Panorámico, 2018).

Revolución, mutación, procesos de cambio para modificar el orden establecido, movimientos para desobedecer los dictados del poder. Revolutionary Arkive, ante la paradoja de un archivo que mantenga en orden historias de desórdenes, opta por transformar el archivo de imágenes de manifestaciones, en una colección viva, mutable, donde la visión artificial interrelaciona información visual sobre la revuelta.
«A Pilar Rosado parece interesarle la posibilidad de representar el alma, la personal y la colectiva. En su obra, Revolutionary Arkive, reúne imágenes de revueltas y revoluciones buscando e invitando al espectador a buscar analogías, diferencias, hilos de conexión o denominadores comunes en las imágenes y los gestos de levantamientos de multitudes reivindicativas» (Pérez-Hita, 2019: 1).

Las revueltas o revoluciones nacen cuando nos quieren imponer una serie de estructuras y de límites que decidimos romper para liberarnos. Podríamos decir que los conceptos se nos quedan pequeños y necesitamos ampliarlos. Estos desórdenes son en realidad mutaciones. La vida, llena de cambios está, por tanto, llena de revoluciones. Como apunta Fontcuberta, «Las imágenes son construcciones ideológicas, un campo de batalla político» (Fontcuberta, 2019: 177).

La exposición presentó, por un lado, el vídeo titulado The eye-machine is watching, que consistía en una serie de grabaciones de manifestaciones reales de los trabajadores de Granollers a lo largo de los años exigiendo sus derechos. Las películas fueron analizadas mediante un algoritmo de detección de rostros que localizaba rápidamente las caras de los manifestantes, pero las etiquetas de categorización habían sido manipuladas por la artista de manera que, en lugar de identificar «rostros», identificaban a los manifestantes con etiquetas subversivas como «agitador», «revolucionario», «rebeldía», «insurgente», «disidente», etc., anunciando un hipotético futuro en el que las computadoras sean capaces de identificar este tipo de cualidades personales, simplemente al detectar nuestro rostro (ver figura 8).

Figura 7. Exposición Revolutionary Arkive en el contexto de Festival Panoràmic de Granollers en 2018. © Camila Marinone.

Figura 8. Frame del vídeo The eye-machine is watching, proyectado en la exposición Revolutionary Arkive en el que la artista manipula un algoritmo de detección de rostros para asignar etiquetas relacionadas con la revuelta como: insurgencia, agitación, disidencia, rebeldía, etc. © Pilar Rosado.

La exposición presentaba también una serie de grandes paneles con conjuntos de fotografías descargadas de internet al realizar búsquedas en Google sobre manifestaciones de países de todo el mundo (ver figura 7). La intención era realizar el proceso de categorización automática descrito en el apartado de «métodos computacionales» para extraer arquetipos visuales de revuelta, iconos de rebelión, insurjencia, agitación política o reivindicación social y, al mismo tiempo, para hacer visibles las diferencias o particularidades que tienen las fotografías que capturamos en cada país o región.

«La artista trabaja con redes neuronales artificiales, algoritmos de deep learning que, si tienen al alcance la suficiente cantidad de datos, pueden extraer los elementos formales que se repiten de manera estadísticamente significativa para constituir los rasgos de la representación arquetípica de una determinada categoría semántica» [Traducción propia]¹¹ (Fontcuberta, 2019: 178).

Para poder obtener una representación visual de cada continente, se seleccionaron imágenes a partir del concepto «manifestación», añadiendo el criterio de lugar: «manifestaciones en Cataluña», «manifestaciones en Rusia», «manifestaciones en China», «manifestaciones en Nicaragua», «manifestaciones en Sudáfrica» o «manifestaciones

---

¹ En palabras de Joan Fontcuberta: «L’artista treballa amb xarxes neuronals artificials, algoritmes de Deep Learning que, si tenen a l’abast la suficient quantitat de dades, poden extreure els elements formals que es repeteixen de manera estadísticamente significativa per constituir els trets de la representació arquetípica d’una determinada categoria semàntica».
en Egipto»; y se ordenaron en diferentes paneles, a la manera de Aby Warburg. De este modo, la artista creó fotomosaicos donde las imágenes de cada conflicto aparecen ordenadas por afinidades gráficas (ver figura 9).

Una vez posicionadas las imágenes según su contenido visual, se puede apreciar qué relevancia tienen los tanques o los elementos bélicos, la mortalidad, las masas de gente unida, las banderas, las pancartas o la acción policial. Evidentemente, tratándose de manifestaciones, las imágenes con gente o policías son las que más abundan en cada mapa o, en el caso de Cataluña, por ejemplo, las banderas también tienen mucho peso (ver figura 10), pero claramente se observaron diferencias en función del lugar. Sin embargo, los mapas visuales presentados no son categóricos y solo muestran una realidad en la que el análisis de cada persona puede profundizar.

El objetivo no es únicamente identificar un patrón para definir las revoluciones, sino poner de manifiesto que las nuevas tecnologías dan un punto de vista alternativo al humano sobre las imágenes, que aportan mucha información y que permiten catalogarlas desde una perspectiva que hasta ahora no se había explorado.

**Notas finales**

La diferencia fundamental entre los mapas visuales de las dos exposiciones presentadas radica en los criterios de búsqueda de imágenes: en *Mnemosyne 2.0* la artista utiliza en sus búsquedas múltiples conceptos muy variados que obedecen a sus intereses personales relacionados con la «tecnología». Por el contrario, en *Revolutionary archive* todas las búsquedas de imágenes parten de un único concepto, el de «manifestación», pero en diferentes países del mundo, con la intención de extraer los arquetipos visuales de la colectividad. Hay un juego entre lo personal y lo colectivo en ambas propuestas y las dos se ven enriquecidas por la cooperación de los algoritmos.

Los grandes repositorios de imágenes constituyen una parte del recuerdo común de la humanidad y es creciente el interés hacia el concepto de «memoria colectiva». Con respecto a la historia y la teoría del arte y la cultura, Malraux y Warburg estaban interesados en el papel de las imágenes como representaciones y vías de transmisión de la memoria psicológica de las sociedades, pero también en la metamorfosis de formas y modos de expresión en la memoria social global (Grebe, 2010: 8).

En las exposiciones presentadas, la posibilidad de abarcar el contenido visual total de una colección nos permite considerar una información a la que no tendríamos acceso analizando solo las partes, las propiedades de la totalidad no resultan de los elementos constituyentes, sino que emergen de las relaciones espacio-temporales del todo (Köhler, 1947; Koffka, 1967). Estas metodologías nos abren nuevas vías para el diálogo con el pasado y prometen iluminar muchos aspectos de la historia y evolución de las imágenes.

«Artistas y científicos tienen que habérselas con esos datos e imágenes, y no solo investigar y criticar lo que se está haciendo con ellos desde arriba, sino proponer otras vías, otros juegos y preguntas que apunten a algo que no sea vender o venderse» (Pérez-Hita, 2019: 2).

Desde hace años Margaret Boden estudia la forma en que la inteligencia artificial puede ayudarnos a explicar la creatividad humana.
y, aunque reconocer la creatividad en máquinas o en humanos no es nada sencillo, Boden aclara que se pueden distinguir tres tipos: creatividad combinatoria, exploratoria y transformacional. En la crea-
tividad combinatoria las ideas conocidas se mezclan de maneras desconocidas, se construyen analogías a partir de similitudes estructu-
rales. La creatividad exploratoria se apoya en reglas y estructuras que ya tienen valor cultural como, por ejemplo, estilos pictóricos o musicales, para generar nuevas propuestas (Boden, 2017: 73). El proceso creativo seguido por Rosado en ambas exposiciones pone en juego la creatividad combinatoria y la exploratoria. La combinatoria porque se propicia el encuentro entre las analogías visuales que establecen los algoritmos informáticos y los criterios de búsqueda que decide la artista. En lo que respecta a la creatividad exploratoria, en el proceso se solapan los arquetipos visuales de la autora con los patrones visuales que detecta el algoritmo en los grupos de fotografías analizados.

Podríamos decir que en este proceso creativo se superponen las preconcepciones visuales del humano y las de la máquina. Ésta es una diferencia sustancial a nivel metodológico con las estrategias para establecer analogías visuales utilizadas por Aby Warburg, ya que él solo disponía de su propio criterio. En ambos casos la estrategia tiene en común el deseo de establecer un diálogo con las imágenes para pensar.

Desde este punto de vista, en las obras de Rosado descritas, se pone de manifiesto la forma en que la conjugación humano-máquina puede favorecer los procesos de cocreación. Esta colaboración depen-
den gran medida del juicio humano, quien al final es el responsable del éxito del sistema; sin embargo, los procesos de cálculo aumentado del computador no hacen más que estimular este proceso creativo proporcionando rapidez a unas búsquedas que serían imposibles de otro modo. Sin olvidar que los resultados de las búsquedas utilizadas para obtener el conjunto de imágenes están mediatiados por los algoritmos que utiliza Google, no cabe duda de que la interacción con el computador alimenta encuentros fortuitos e inesperados, además de los previsibles, y que en este proceso de búsqueda sirven para activar la imaginación de la artista.

Para finalizar, hay que puntualizar que el proceso de decisión del tamaño y número de imágenes descargadas condiciona la medida final de los mapas y esto se decide a partir de la obtención de grandes fotografías que no superen la escala humana, para que sea posible abarcar con la mirada el todo y la parte. Así se nos desvelan a un tiempo los detalles de las pequeñas imágenes y las configuraciones de los arquetipos de la totalidad. En la distancia, las grandes fotografías revelan unas valorables cualidades pictóricas de arte abstracto, las pequeñas fotografías se asemejan a las pinceladas, se reducen las representaciones objetuales contenidas en ellas hasta quedar únicamente representada la tensión visual que se establece por la gradación formal en la que se disponen. La información se hace arte y su belleza pretende hacernos despertar.

Agradecimientos

Nuestro sincero agradecimiento a los revisores anónimos que con sus comentarios ayudaron a mejorar el manuscrito original y a Roca Um-
bert Fábrica de las Artes de Granollers (Barcelona), Festival Panorámico, Archivo Municipal de Granollers y ADAG.

Referencias bibliográficas

Boden, Margaret Ann. 2017. *Inteligencia Artificial*. Madrid: Turner Noema.
Company, Oriol. 2018. «Revolutionary Arkive. Ordenar el desorden». *Núvol. El Digital de la Cultura*, October 10. https://www.nuvol.
com/pantalles/revolutionary-arkive-ordenar-el-desorden-55800.
D’Oultremont, Maria. 2019. «Pilar Rosado i la panomarica perduda de les imatges». *Núvol. El Digital de la Cultura*. May 17. https://
www.nuvol.com/art/pilar-rosaldo-la-panomarica-perduda-pees-
imatges-60640.
Dídi-Huberman, George. 2010. *Atlas. ¿Cómo llevar el mundo a cues-
tas?*. Madrid: TF Editores.
Festival Panorámico. 2018. «The whole world is watching». http://
panoramicgranollers.cat/revolutionary-arkive/.
Fontcuberta, Joan. 2016. *La furia de las imágenes. Notas sobre la postfotografía*. Barcelona: Galaxia Gutenberg.
Fontcuberta, Joan. 2019. «El gran riu. Per a una teoria general de la revolta». *El món d’Ahir. Història d’autor*, 13: 176-189.
Google. «Cómo funcionan los algoritmos de Google». May 22, 2020. https://www.google.com/intl/es_search/howsearchworks/
algorithms/.
Grebe, Anja. 2010. «Museum and Mnemosyne. Aby Warburg, André Malraux and the re-construction of art history as social history». *Le Musée Imaginaire* and Temptations of the Orient and Japan, ed. Tanaka Hidemichi, 55-61. Tokyo: Akita.
He, Kaiming; Xiangyu Zhang, Shaoqing Ren y Jian Sun. 2016. «Deep residual learning for image recognition». In *Proceedings of the IEEE conference on computer vision and pattern recognition (CVPR)*, 770-778. http://openaccess.thecvf.com/content_cvpr_2016/
/html/He_Deep_Residual_Learning_CVPR_2016_paper.html. https://doi.org/10.1109/CVPR.2016.90.
Karpathy, Andrej. 2017. «t-SNE visualization of CNN codes», accessed January 12. https://cs.stanford.edu/people/karpathy/cnnembed/.
Koffka, Kurt. 1967. *Principles of Gestalt Psychology*, New York: Mi-
nesos International.
Köhler, Wolfgang. 1947. *Gestalt psychology: an introduction to new concepts in modern psychology*. New York: Liveright.
Malraux, André. 1949. *Le Musée imaginaire*. Ginebra: Skira.
Manovich, Lev. 2009. *TimeLine*, accessed January 10. http://ma-
novich.net/index.php/exhibitions/timeline.

He, Kaiming; Xiangyu Zhang, Shaoqing Ren y Jian Sun. 2016. «Deep residual learning for image recognition». In *Proceedings of the IEEE conference on computer vision and pattern recognition (CVPR)*, 770-778. http://openaccess.thecvf.com/content_cvpr_2016/
/html/He_Deep_Residual_Learning_CVPR_2016_paper.html. https://doi.org/10.1109/CVPR.2016.90.
Karpathy, Andrej. 2017. «t-SNE visualization of CNN codes», accessed January 12. https://cs.stanford.edu/people/karpathy/cnnembed/.
Koffka, Kurt. 1967. *Principles of Gestalt Psychology*, New York: Mi-
nesos International.
Köhler, Wolfgang. 1947. *Gestalt psychology: an introduction to new concepts in modern psychology*. New York: Liveright.
Malraux, André. 1949. *Le Musée imaginaire*. Ginebra: Skira.
Manovich, Lev. 2009. *TimeLine*, accessed January 10. http://ma-

Ref: 55800
Manovich, Lev. 2012. «¿Cómo ver 1000000 de imágenes? Deforma cultura online, 3, 1-11, Recuperado el 10 de Agosto de 2018 de http://www.deforma.info/es/product.php?id_product=24.

Manovich, Lev. 2019. «Defining AI Arts: Three Proposals». Al and Dialog of Cultures, Exhibition Catalog, Saint-Petersburg, Russia: Hermitage Museum.

Pérez-Hita, Félix. 2019. «Pilar Rosado y la resurrección de los archivos». Mnemosyne 2.0. Cartografías computacionales de la memoria visual, 1-2. Granollers: Roca Umbert Fàbrica de les Arts.

Rosado, Pilar. 2019. Mnemosyne 2.0. Cartografías computacionales de la memoria visual. Granollers: Roca Umbert Fàbrica de les Arts.

Russakovsky*, Olga, Jia Deng*, Hao Su, Jonathan Krause, Sanjeev Satheesh, Sean Ma, Zhiheng Huang, Andrej Karpathy, Aditya Khos-la, Michael Bernstein, Alexander C. Berg, y Li Fei-Fei (* = equal contribution). 2015. «ImageNet Large Scale Visual Recognition Challenge». International Journal of Computer Vision (IJCV), 115: 211–252. https://doi.org/10.1007/s11263-015-0816-y.

Saramago, José. 2003. Ensayo sobre la ceguera. Madrid: Alfaguara. Maaten, Laurens van der y Geoffrey Hinton. 2008. «Visualizing data using t-SNE». Journal of Machine Learning Research, 9 (Nov): 2579-2605.

The Warburg Institute. 2018. «The Mnemosyne Atlas, October 1929», accessed 2020 February 20. https://warburg.sas.ac.uk/collections/warburg-institute-archive/bilderatlas-mnemosyne/mnemosyne-atlas-october-1929.

Warburg, Aby. 2010. Atlas Mnemosyne, traducido por Joaquín Cha- morro Mielke. Madrid: Akal.
Vistas panorámicas sobre el patrimonio visual colectivo a través de redes neuronales convolucionales

CV

Pilar Rosado Rodrigo
Universitat de Barcelona
pilarrosado@ub.edu

Pilar Rosado es una artista e investigadora cuya práctica explora el uso de algoritmos de aprendizaje automático en el ámbito de la creación. Trabajando con grandes colecciones de imágenes de archivo y estas nuevas tecnologías pretende proporcionar puntos de vista alternativos para la reflexión y que cuestionen las convenciones de nuestra mirada. Doctora en Bellas Artes y Licenciada en Biología, entre 2012 y 2015 fue becada por la Generalitat de Catalunya para llevar a cabo su investigación titulada «Formas latentes: protocolos de visión artificial para la detección de analogías aplicados a la catalogación y creación artística», cuyos resultados han sido publicados en la prestigiosa revista Leonardo del MIT Press en junio de 2019. Combina su carrera artística con la investigación y la docencia como profesora en la Facultad de Bellas Artes de la Universidad de Barcelona, siendo miembro de varios proyectos de I+D.

ORCID: https://orcid.org/0000-0001-7142-5047
www.pilarrosado.eu
CV

Ferran Reverter Comes
Universitat de Barcelona
freverter@ub.edu

Ferran Reverter es profesor agregado de Estadística del Departamento de Genética, Microbiología y Estadística de la Universidad de Barcelona. Profesor del Máster de Estadística e Investigación Operativa UB-UPC en la asignatura de Statistical Learning. Ha realizado estancias posdoctorales en el Laboratoire de Statistique et Probabilités de la Université Paul Sabatier (Toulouse, 2008) y en el Centre de Regulació Genòmica del Parc de Recerca Biomèdica (Barcelona, 2013-2017). En la actualidad centra su investigación en el desarrollo e implementación de herramientas estadísticas para el análisis de datos en biociencias, la mejora de métodos basados en funciones de núcleo (kernel methods) para aumentar la interpretabilidad y visualización, y en la implementación de herramientas de análisis computacional de imágenes en el ámbito biomédico y el artístico.

ORCID: https://orcid.org/0000-0002-9489-3350.
Unlocking the Black Box of AI Listening Machines: Assemblages for Art, Technology and Innovation

Sharath Chandra Ramakrishnan
School of Arts, Technology, and Emerging Communication
University of Texas at Dallas

Date of submission: March 2020
Accepted in: June 2020
Published in: July 2020

Abstract
The black box of innovation in the realm of connected AI technologies renders not only their technicalities opaque but also, and more importantly, the social effects and relations that constitute their creation and mediation. This presents an opportunity for creative interventions by artists and researchers, to unveil the networked relations that are part of AI technologies, and speculate on their ontological effects. This article presents such an unpacking around an AI listening machine present today in ubiquitous devices like voice assistants and smart speakers, and incorporates computational models of machine audition. By tracing the scientific research, technical expertise, and social relations that led to our cultural adoption of AI listening machines, the article presents a socio-technical assemblage within which these machines operate. At the same time, the article reveals various contexts for artists as well as innovation researchers to engage with the socio-technical complexity of AI listening machines, by sharing some instances of creative and artistic interventions that have attempted to unveil the nature of their assemblages.
Introduction
Over centuries of our attempts at understanding the mechanism of human audition, we have today modelled listening machines that can mimic the ear and perform a number of useful functions such as conversing to answer our questions, recognising a piece of music and even recreating the sonic world for people with hearing problems. New AI listening technologies used today may be viewed as an extrapolation of an ongoing cultural training over centuries. This cultural training has involved teaching humans new habits of listening with machines, as well as techniques to interpret and operate them. Listening machines can learn to perform intelligent analysis of sound and speech based on computational models of human audition, endowed with capacities to not just simulate but also exceed and augment the limits of human audition, thereby problematising our contemporary notions of aurality. Indeed, the act of listening is no longer restricted to humans, and if machines can listen to what humans cannot perceive, how does this affect and shape social and cultural relations associated with AI augmented listening? These are but some of the questions that the author wishes to engage the reader with, while delving into its object of study - the ‘AI listening machine’, that endows the ubiquitous smart speaker/voice assistant with the ability to tune in, listen to and make ‘sense’ of the audile world of humans.

Contributions
A Call to Unveil Socio-Technical Relations Behind AI Systems
Before diving into our discussion of the AI listening machine, locating the object of study in the historic cross currents of scientific discovery and technological innovation would serve to precisely lay out the intent and contributions of this article.

Since the early 15th century, much of the Western world was preoccupied with an ontological question regarding the mechanistic interpretation of organic and biophysiological processes. Intellectuals questioned if consciousness or the nature of life itself can ever be mechanically represented, and as a result, several experimental efforts to create automatons ensued. These ranged from Vaucanson’s lifelike mechanical Duck (Riskin 2003) and his Flute Player (incidentally documented and described in the first French encyclopedia by Diderot as an “android” (androide) (Rice 2014), to the mechanisation of rational thought in computing automatons like the Pascaline by Blaise Pascal in 1645 (Price, Bedini, et al. 1964). While the history of the automaton and its roots in mechanistic philosophy is not within the scope of this paper, what is pertinent is the academic ramification of the above-mentioned ontological paradox that has created a gap in our understanding of the relationships between science and technology (Pinch and Bijker 1984). Understanding this relationship is particularly complicated for emerging
technologies like AI listening machines that subsume hypothetical findings in the cognitive and auditory neurosciences within the agenda of emerging human enabling listening technologies.

While sociologists of science and technology studies have contributed a multitude of disciplinary perspectives on the relationship between science and technology, an overall thematic approach has been to separate technology from science on analytical grounds, by attributing discovery of truth to science and delegating the application of this truth to the role of technology (Pinch and Bijker 1984). Despite studies by innovation researchers involving empirical inquiries to map both the extent to which technological innovations originated from scientific discoveries, as well as the dependence of scientific research on the availability of a specialised technology, the clear view on their interdependence remains difficult to specify. Mayr attributes this failure to the perpetuated assumption that science and technology are disparately defined structures (Pinch and Bijker 1984; Mayr 1976). He posits that in order to advance a deeper understanding, it is essential to realise that “science and technology are themselves socially produced in a variety of social circumstances” (Pinch and Bijker 1984). This view regards the relationship between scientists and technologists to exist within a socially constructed and mediated culture.

Research has further indicated that most innovation studies dealing with a simple linear model, which proceed from research to product development using R&D metrics and macroeconomic success indicators of technological innovation, refrain from discussing societal factors or the technology itself (Layton 1977). Layton suggests that, “what is needed is an understanding of technology from inside both as a body of knowledge and as a social system. Instead, technology is often treated as a ‘black box’, whose contents and behaviour may be assumed to be common knowledge” (Layton 1977).

This socio-technical opacity is further exacerbated by the deployment of AI systems, revealing “a blind spot in AI research”. This was observed in a recent research report that concerned the installation of autonomous AI agents in critical social systems like hospitals and courthouses, with neither technical knowledge of the ‘AI black box’ nor agreed methods to assess the sustained effects of such technologies on society (Crawford and Calo 2016).

This presents a call for creative technologists, computational media artists and researchers working in the realm of AI technologies, to help bridge disciplinary silos, by unveiling these various socio-technical arrangements at play in the relational exchange between the science and technology of AI-based innovation. This article deals with the personal AI listening machine as a technological object of study, towards realising the above-mentioned goal.

Methodology

An examination of the various socio-technical and scientific developments that led to the conception of the AI listening machine, and conversely, how technologies of the AI listening machine assimilated dominant socio-technical and cultural undercurrents, are central to the concerns of this study. To aid in the unpacking of these relations, the article adopts methods from techno-cultural analysis used in sound studies.

The author would like to clarify the basis upon which such a methodological analysis is staged upon, for the benefit of artists and creative technologists dealing with research methods. Technologies of listening are cultural artefacts of a metamorphosis in our understanding of sound as a medium, knowledge about human audition and practices of listening that evolved over centuries (Sterne 2003). Therefore, in order to unpack socio-technical relations, one has to take into account various cultural subjectivities that are reinforced in our contemporary engagement with AI-based listening machines. Within this context, cultural subjectivity refers to a society’s “characteristic way of perceiving its social environment” (Triandis 1972), and consists of ideas and practices that have worked in the past and continue to be preserved and transmitted to the future (Triandis 1972). It follows that a study on the innovation of technologies, which include the technologies of listening, necessitates an understanding of their connections with human practices and habits, along with a focus on areas of intertwined cultural, social, and material histories. These interconnections from which technologies emerge and exist in, have often been referred to as ‘networks’ or ‘assemblages’ (Latour 2012; Wise 1997). These socio-technical assemblages are a necessary apparatus for the creation of AI listening machines, and include technological and scientific explorations, as well as institutions and individuals from diverse disciplinary backgrounds contributing to a social and cultural construction of the mechanism of AI-based aurality.

Unpacking the Socio-Technical Assemblage of the AI Listening Machine

In subsequent sections, we proceed to unveil the socio-technical assemblages within which the AI listening machine operates. Examples of media art interventions are provided to serve as a catalyst for readers wishing to further engage in creative explorations of these assemblages. The article unpacks this relational assemblage across the technologies of AI listening machines, corporeal relations between the science and technologies of listening and policy negotiations mediated by the AI listening machine in its acoustic space.

Figure 1. The ubiquitous smart speaker/voice assistant. Attribution: NDB Photos / CC BY-SA (https://creativecommons.org/licenses/by-sa/2.0).
Technicalities of the AI Listening Machine

The now ubiquitous ‘smart speaker’/voice assistant occupies a unique spot in the convergence of innovation in AI techniques spanning the areas of machine listening, semantic linguistic parsing and speech synthesis technologies. While each of these topics traces a unique techno-social history in itself, the scope of this paper is restricted to the AI listening machine and its associated machine listening techniques for auditory scene analysis, including the contribution that machine listening techniques have had in the evolution of speech synthesis technologies as a historical intersection.

Assemblage of Auditory Scene Analysis

Much of the challenge in understanding audition in humans, as well as efforts in computational auditory modelling, is centered on trying to solve the “cocktail party problem” (Cherry 1953). Indeed, in a noisy house party, if one tries hard to focus one’s auditory attention, it is possible to attend to any single conversation or sonic event occurring amongst a myriad of other sounds. This is precisely what auditory scene analysis tries to achieve - to be able to accurately segregate, segment and delineate individual sources from a complex mixture of sound, to later call upon to audition any of these sources of sound at will (Bregman 2001). The human brain seems to have encoded such processes for auditory scene analysis, the precise workings of which are still an area of active scientific research. Auditory source segregation, localisation and enhancement techniques to create machines that can listen and generate sensation, has been an extremely active area of research for scientists and engineers developing cochlear implants and hearing aids for the specially-abled.

The technical assemblage of auditory scene analysis models used in the AI listening machine, in its first stage, acoustically simulates cochlear filtering, followed by extracting information regarding pitch/frequency transitions, on-sets and offsets as an auditory representation, referred to as the auditory map (Brown and Cooke 1994). The next stage involves constructing a symbolic description of elements in the auditory scene using information in the auditory map. The final stage, analogous to the sensory grouping mechanism, involves a search mechanism that uses various strategies such as combining elements with common fundamental pitch, onset and offsets, to group them to be part of a single auditory stream. Notable methods to consider in the evolution of source separation, recognition and parsing abilities of the AI listening machine include a Markov model-based approach for segregating a mixture containing two speakers based on pitch derived from inter-peak intervals of a cochlear filterbank (Weintraub 1986), other methods that use pitch as a segregation criterion (Parsons 1976; Stubbs and Summerfield 1990), as well as additional information such as spatial location of the voice (Denbigh and Zhao 1992) to aid in source segregation. While these methods do rely on some prior knowledge about the acoustic environment, like number of speakers or types of sound sources, other methods (Mellinger 1991) do not make such strong assumptions about the environment and are much more suited to the operation of the personal AI listening machine. New AI techniques like TasNet, which can “directly model the signal in the time-domain using an encoder-decoder framework” (Luo and Mesgarani 2018), have greatly improved the prospects for performance of speech separation algorithms in low power and portable devices like smart speakers.

These auditory scene analysis models have been proposed and developed by auditory scientists and machine listening engineers working in collaborative settings. This tells us about the social context within which innovation of the AI listening machine happens. An ethnographic study involving researchers in the field of medical physics working on a model to transcribe speech in noisy environments to text, found scientists and engineers relating within a tension and contingency between different rationales (Voskuhl 2004). Colleagues involved in the pathology of hearing seemed suspicious of the intent of duplicating the ear for use in artificial speech transcribers. At the same time, they felt entitled by the common belief within their research circles that engineering techniques were unsuccessful in noisy environments, and the use of knowledge in medical acoustics in signal processing models was considered innovative, promising and niche. Despite their conflicting institutional rationales, a symbiotic relationship was guided, on one hand, by the objective of designing a listening technology that mimicked human audition, and on the other, the hope of obtaining a plausible model to use as an epistemological tool to decipher the biological science of hearing (Voskuhl 2004). Performance evaluation of the machine learning models used to perform source segregation involved a large training set of human speech recorded in noisy backgrounds, as well as an evaluation data set that served to measure the benchmark of performance.

The creation and use of such training and evaluation data sets involving human vocal contributions suggest another component of social interaction embedded in this layer of the assemblage. The limits of speech separation, recognition and parsing abilities of the AI listening machine are constantly tested in simulated real-world environments, by different spectral representations that use noises to mask and render the speech signal unintelligible to the machine. In this process, the individual voice is abstracted into a composite signal representation devoid of any personal attributes. This is similar to the consequence of abstracting hearing as a tympanic mechanism in early listening machines like the ear phonograph or the pattern playback machine, that rendered the source of sound as irrelevant in scientific investigations. John Peters has highlighted this aspect of source-agnostic research in the work of Hermann von Helmholtz; “Helmholtz levels all modalities and is indifferent to bodily origins. What matters is the waveform and not the source” (Peters 2004).

The artistic interpretation of this aspect was brought in the work of composer and sound artist, Alvin Lucier. I Am Sitting in a Room (1970) was a composition that featured spoken text, which
the visitor was instructed to record and play back over itself repeatedly until the resonant frequencies of the room annihilated any form of intelligible speech (Cox 2009). Lucier urges the audience not to interpret this piece as a spatial exploration of sound and rather explicitly clarifies that through the technology of listening and playback, “the piece concerns the dissolution of speech and the speaker into sound and space. What begins as a personal confession in a domestic setting gradually becomes pure, anonymous sound that overwhelms and eradicates the performer’s personality. Meaning and sense have dissolved into rhythm” (Cox 2009). By modifying, tuning and stretching the optimised parameters involved in the technical realisation of auditory source separation to the point of technical failure, performance of the anomalous and the anonymous become possible. Computational artists working in the field of audition could perhaps think of interesting ways to unveil the complex interplay at work between the rationale of technological efficiency and the quest of biological understanding, by manipulating underlying mechanisms of the AI listening machine.

### The Overlap Between Machine Listening and Speech Synthesis Techniques

The history of machines invented to imitate the production of human sound as speech is just as important as the mechanisation of hearing, as both these attempts converged in their interpretation of speech through the spectrogram. This section clarifies the extent to which machine listening techniques used in auditory scene analysis are also used for machine speech synthesis in AI agents like voice assistants. There is a large body of research and art dealing with speaking machines, dating back to Von Kempelen’s automaton, a mechanically operated speaking machine that could produce a remarkably human-like voice (Dudley and Tarnoczy 1950). Speech production took a visible shape with the invention of the spectrogram, that when combined with the anatomical understanding of formant production documented over the century, resulted in the invention of speech synthesising machines like the Voder and more importantly, the Vocoder, by Homer Dudley in Bell Labs in 1939 (Dudley and Tarnoczy 1950). The spectrum analysis component of the Vocoder to extract features, minimally represent and synthesise human sounding speech is a common processing feature shared with automatic speech recognition systems performing auditory scene analysis (Pickett 1968). For instance, commercial vocoders featured an array of components like microphones for capturing vocal input, a spectrum analyser to filter, sort and process sonic frequencies, a pitch detector and follower that locked onto the fundamental frequency of the auditory input, and a voice/uvocoded detector that distinguished acoustic features produced by the vibrating vocal cords from those produced by modulations by the mouth, lips, and tongue. A process of encoding to translate this information as a binary signal, that could be later decoded back to speech pertained to the synthesis part (Pickett 1968).

While many artists have explored the synthesis aspects of the vocoder and variants of its linear predictive encoder in a performative and musical context, it was Lucier who was interested in the Vocoder as a tool for vocal and auditory analysis, rather than synthesis. In his piece *North American Time Capsule*, participants equipped with musical instruments and electrical appliances sang, spoke and read into the vocoder while their auditory input was modified to convert audible speech to an abstract sound and texture (Cox 2009). Lucier transformed the listening characteristics of the vocoder, “to liquidate speech and to abolish the identity of the speaking subject, shattering all syntax and pulverizing every symanteme, morpheme, and phoneme into fluid sonic matter” (Cox 2009). The advent of neural networks has allowed for the creation of vocoders that can synthesise high-quality, human-sounding speech, by conditioning on spectral features using speech analysis modules as well as human coded feature specification, that are open to new modes of artistic exploration along the axis of speech and identity. State-of-the-art implementations of the day include RawNet, an end-to-end neural vocoder, where the speech coder and voder act in tandem as an auto-encoder network, capable of being jointly trained directly on the raw audio waveform without the need for human feature design (Luo and Mesgarani 2018). The future of the smart speaker, or personal AI listening agent, is geared towards the capacity to listen and learn from higher level representations of speech input, to constantly improve its voice synthesis capability, thereby sharing a reciprocal relation with the machine listening aspect of its technology as well as its vocal user.

### Corporeality of the Science and Technologies of Listening

The engagement of early listening machines as well as the modern-day AI listening machine with the institution of those with hearing problems, adds a corporeal dimension to the socio-technical assemblage discussed in this article. As listening machines transformed the function of the ear through their mechanisation, it was their socialisation and institutionalisation that led to their prevalence as audible media. The significance of the phonautograph in creating the first visual spectrogram of sound was part of a larger assemblage of institutional practices. The invention of the phonautograph was necessitated by the simultaneous establishment of otology (or ear medicine) as a formal branch of medical science. This made the human body and human ear available for the purposes of dissection and investigation. Sterne notes that “the human ear affixed to the chassis of the ear phonautograph” suggests how this hybrid nature of the phonautograph as a listening device could be considered an “artifact of otology’s institutionalization”, thereby revealing the corporeal nature of its existence (Sterne 2003). In this context, the concept of corporeality refers to the “diversity of issues, questions, concepts and relationships deriving from and centering on the body and bodily life” (Sheets-Johnstone 2015). Audition was transformed into a mechanism that could be abstracted from the human body
both anatomically and experientially. The ear became an object of study and its functioning became measurable, allowing physiologists to define it in mechanical terms. Bell’s intention of using the phonograph to represent visible speech was to train the deaf to speak, as it made speech visible as a waveform rather than depict articulations of speech as positional variations of the tongue and the mouth (Bell, Bolton, and Langdon 2017).

Our insights and theories related to the science of audition along with our knowledge to create computationally intelligent listening machines have been historically associated with “contextually situated human bodies to enact, experience and interpret” sound (Deleuze 1988). This fact is closely related to the words of philosopher Gilles Deleuze - “being social before being technical” (Deleuze 1988). Further, Kittler points out that innovations and progress in instrumentalised sensory measurement programmes managed by physiologists was crucial to the successful inventions and utilitarian proliferation of technological media (Kittler 2006). According to Kittler, the mediatisation of the senses did not occur solely through the invention of technical media but under the auspices of the psychophysical laboratory, where the experimental psychologist treated the human subject as a technological media product (Kittler 2006). In the context of innovations in listening machines and their relationship to the hearing impaired, Mara Mills points out that,

“Across decades, national contexts, and technical shifts, however, deafness ultimately served as a ‘pretext’ to other engineering concerns - in some cases a precursor, in others a pretense. Inventors often abandoned collaborations with deaf students and their educators after initial trials, as their technologies transferred to more profitable realms. Certain inventors simply lifted ideas and inspirations from the world of deaf ‘assistive’ technologies.”

(Theorizing Disability in Media, 2010)

The sound art community along with the deaf community has responded to this corporeality within the context of ‘Deaf Gain’ (Bauman and Murray 2014). The term ‘Deaf Gain’ inverted the concept of ‘hearing loss’, and positioned deaf culture as a valuable contributor to conversations in both art, science and culture. Relevant art projects include Tonotopia by sound artist and composer Tom Talim in collaboration with cochlear implant users, where composing and codesigning sound art along with and for persons with hearing disabilities was the focus of artistic engagement and creation (Talim 2017). In an attempt to enable artistic experience to be shared by diverse ears, the project (curated at the Victoria and Albert Museum) asked several pertinent questions, like what is deemed as normative or natural listening and who decides what these factors and parameters should be, and how are technologies of listening empowering or disavowing the deaf? These questions are germane to unveiling the interrelation between the auditory scene analysis industry, and auditory science for the hearing impaired, calling for more creative technologists and artists to respond to these relations.

Negotiations in the Networked Acoustic Private Space

Our interaction with the domestic AI listening machine occupies a networked acoustic space that is situated within a certain locus around it. The activation of this acoustic space is encoded by performed interaction routines (like a ‘wake phrase’ containing the name of the AI listening machine), thereby idealising the audible technique of interaction with the AI listening machine. Sterne describes how the idealisation of the audible technique led to the creation of private acoustic spaces that became “a precondition for the commodification of sound” (Sterne 2003). He reasons that “commodity exchange presupposes private property”, necessitating “the acoustic space to be ownable before its contents could be bought or sold” (Sterne 2003). Thus, the AI listening machine could be seen as ‘owning’ the networked private acoustic space into which it has been invited. This raises several questions regarding socio-technical assemblages at work in the private acoustic space co-inhabited with the AI listening machine.

Previous innovations of instrumentalised listening spaces have often emerged within a surveillant framework, as Ganchrow notes about the prerradar stone acoustic mirrors, that “could not have been undertaken without a corresponding shift in thinking about sound outside of the way sounds are perceived -more specifically, toward thinking about frequencies in terms of physical sizes” (Ganchrow 2009). The work of W. C. Sabine describes at length the “Ear of Dionysius”, a panaural prison in Sicily dating back to the 4th Century BC (Sabine and Egan 1994). The S-shaped chamber of the prison featured a conical duct leading to a concealed chamber where all the acoustic reflection from within the prison was audible, an acoustic architectural version of Jeremy Bentham’s Panopticon (Bentham 2012). These innovations, and others like the 3-horn listening structure created by Athanasius Kircher for the Royal Court to listen in to the town hall below, were based on a framework of acoustic materiality that explored and manipulated the physics of sound transmission, rather than its sensorial and cognitive aspects. Further, these listening machines were installed with the intent to surveil, and represented a nature of listening that is different from what is mediated by the object of study in this article.

In contrast, AI listening machines have not been explicitly developed under a surveillance agenda and therefore present new terrains for negotiating the networked private acoustic space. Although the domestic AI listening agent is bound by data privacy and retention laws that govern networked media, there is a reciprocal relationship at work while we willingly employ their services. Looking closer, we find that the machine learning algorithms in computational listening models and speech synthesis models in these machines constantly improve their diction, listening abilities in noisy environments, speech recognition with varied accents, and sense of directional attention, by constantly listening to and learning from human speech interactions. The personal AI listening machine can operate in a sporadic listening
mode, occasionally recording a snippet of domestic conversations happening in the acoustic space in which they are installed or present. The AI listening machine is able to identify the voice of its primary users, and maintain an idea of what their preferences are by constantly building a schematic understanding from their previous interactions and intent. Therefore, in many ways, the personal AI listening machine transforms the private acoustic space into a flow of commodity exchange to enhance both the expressive as well as cognitive model of its computational auditory scene analysis model, in return for its assistive services. This calls for the mapping of new notions of the private acoustic space and the need to outline various negotiations in the acoustic space surrounding the listening machine. A notable art project that attempts to address these notions is LAUREN, where computational artist Lauren McCarthy takes on the job of the smart home agent ‘Alexa’. By allowing the artist to control the lights, cameras, door locks and taps in the home of the person, the project speculates the kind of impending negotiations we may have to choose to make or not, as the AI listening agent acquires more ports and modes of domestic control.

Other threats with the potential to invade and disrupt a secure and private acoustic space do exist from a cybersecurity perspective. The growth of specialised audio-based adversarial attacks are rife with the widespread use of domestic AI listening machines. These attacks feature the use of deep recurrent neural networks that embed inaudible frequencies which represent phrases of text (like ‘visit xyz.com’), into the sound stream input of the AI listening machine. Such advanced attacks can embed an inaudible waveform into a regular speech command stream, to trick the AI listening machine into processing an alternate spectral representation that results in the wrong linguistic transcription to be perceived by the listening machine (Taori et al. 2019). Ongoing research by the machine listening community to address privacy-related challenges encountered with AI listening machines include privacy-preserving methods for speech and audio processing, de-identification and obfuscation techniques as well as methods to detect adversarial attacks to the operation of AI listening machines. These emerging areas present an opportunity for artist and designers to intervene in the socio-technical acoustic space surrounding the AI listening agent.

Recent work by designers around this topic features the ‘Bracelet of Silence’, which emits inaudible ultrasonic waveforms in the 26KHz range that exploits the non-linear properties of the microphone, to slip into the audible range and jam the microphone with white noise. This renders the AI listening machines present in the room unable to record or ‘eavesdrop’ into a conversation happening in the immediate acoustic space. The principle of operation of the bracelet was based on the recently published Dolphin Attack that uses the same technique to send inaudible commands to computational AI listening machines installed in smart speakers (Zhang et al. 2017). In summary, the acoustic space around the networked AI listening machine features a number of actors vying for the control and flow of information in the acoustic space, entailing negotiations at the level of technology policy, cybersecurity and social life.

Conclusion

This article has unpacked a socio-technical assemblage of the AI listening machine that resides in its popular avatar as smart speakers and voice assistants. However, several contexts and possibilities for the AI listening machine can emerge if it is detached from its entrapment as a virtual assistant or smart control agent. For instance, cross-disciplinary leaps have been made in applying methods of speech source separation and spectral analysis to the measurement of electricity usage, contactless payment interfaces and the encoding of inaudible tracking beacons into television broadcasts that enter mobile phones through permissive applications. Each of these contexts perhaps shares or might have a modified socio-technical assemblage to the one discussed in this paper.

As discussed earlier, artistic works can be approached as attempts to unveil socio-technical assemblages of AI technologies. This presents an opportunity to shape productive encounters among scientists, technologists and artists to explore forms of representation and discourse variations that skilfully combine experimentation with technology and social commentary.

The methodology and assemblage discussed in this article may not be a template for exploring all AI technologies, but provide thoughts and references for artists and researchers interested in revealing similar socio-technical assemblages. Finally, unlocking the black box of AI applications serves as an invitation for creators and practitioners to question the social and institutional arrangements where their work takes place, as well as their historicity and established assumptions. This can lead to an intentional engagement with models for collaboration and innovation capable of fostering AI applications that respond to and are responsible with the contexts where they are deployed.

References

Bauman, H-Dirksen L., and Joseph J. Murray. Deaf gain: Raising the Stakes for Human Diversity. University of Minnesota Press, 2014. Bell, Alexander Graham, Lieut-Col Frank Bolton, and William Edward Lang- don. The Telephone: A Lecture Entitled Researches in Electric Telephony (Illustrated Edition). Echo Library, 2017. Bentham, Jeremy. “The Panopticon.” In Offenders or Citizens?, edited by Philip Priestley and Maurice Vanstone, 28–30. London: Willan, 2012. Bregman, Albert S. Auditory Scene Analysis. The MIT Press, 2001. https://doi.org/10.1016/B0-08-043076-7/00663-X.
Mara, Mills. “Deaf Jam: From Inscription to Reproduction to Information.”

Luo, Yi, and Nima Mesgarani. 2018. “TasNet: time-domain audio separation network for real-time, single-channel speech separation.” In 2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 696–700. IEEE.

Mara, Mills. “Deaf Jam: From Inscription to Reproduction to Information.” Social Text 28 (2010): 35-58. https://doi.org/10.1215/01642472-2009-059.

Mayr, Otto. “The science-technology relationship as a historiographic problem.” Technology and Culture 17, no. 4 (1976): 663–673.

Mellinger, David K. “Event Formation and Separation in Musical Sound.” PhD diss., Department of Computer Science, Stanford University, 1991.

Parsons, Thomas W. “Separation of speech from interfering speech by means of harmonic selection.” The Journal of the Acoustical Society of America 60, no. 4 (1976): 911–918.

Peters, John Durham. “Helmholtz, Edison, and Sound History.” In Memory Bytes: History, Technology, and Digital Culture, 177–198. Duke University Press, 2004.

Pickett, J. M. “Historical notes and preface.” American Annals of the Deaf 113, no.2 (March 1968): 117–119.

Pinch, Trevor J, and Wiebe E Bijker. “The social construction of facts and artefacts: Or how the sociology of science and the sociology of technology might benefit each other.” Social Studies of Science 14, no. 3 (1984): 399–441.

Price, Derek J. de Solla, Silvio A. Bedini, et al. “Automata in History.” Technology and Culture 5, no. 1 (1964): 9–23.

Rice, Albert R. “The Android Clarinetist by Cornelis Jacobus van Oeckelen (1838).” Journal of the American Musical Instrument Society 40 (2014): 163–189.

Risin, Jessica. “The defeating duck, or, the ambiguous origins of artificial life.” Critical Inquiry 29, no. 4 (2003): 599–633. https://doi.org/10.1177/0306312704069881.

Sabine, Wallace Clement, and M. David Egan. “Collected Papers on Acoustics.” The Journal of the Acoustical Society of America 95, no. 3679 (1994). https://doi.org/10.1121/1.409944.

Sheets-Johnstone, Maxine. The Corporeal Turn: An Interdisciplinary Reader. Andrews UK Limited, 2015.

Sterne, Jonathan. The Audible Past: Cultural Origins of Sound Reproduction. Duke University Press, 2003. https://doi.org/10.1215/9780822384250.

Stubbs, Richard J., and Quentin Summerfield. “Algorithms for separating the speech of interfering talkers: Evaluations with voiced sentences, and normal-hearing and hearing-impaired listeners.” The Journal of the Acoustical Society of America 87, no. 1 (1990): 359–372.

Taori, Rohan, Amog Kamsetty, Brenton Chu, and Nikita Vemuri. 2019. “Targeted Adversarial Examples for Black Box Audio Systems.” In 2019 IEEE Security and Privacy Workshops (SPW), 15–20. IEEE.

Talalim, Tom. “Tonotopia: Co-designing sound art with hearing implant users: Podcast.” 2017.

Triandis, Harry C. The Analysis of Subjective Culture. New York: Wiley-Interscience, 1972.

Voskuhl, Adelheid. “Humans, machines, and conversations: An ethnographic study of the making of automatic speech recognition technologies.” Social Studies of Science 34, no. 3 (2004): 393–421. https://doi.org/10.1177/0306312704043576.

Weintraub, Mitchel. 1986. “A computational model for separating two simultaneous talkers.” In ICASSP'86. IEEE International Conference on Acoustics, Speech, and Signal Processing 11:81–84. IEEE.

Wise, John Macgregor. Exploring Technology and Social Space Vol. 1. Sage, 1997.

Zhang, Guoming, Chen Yan, Xiaoyu Ji, Tianchen Zhang, Taimin Zhang, and Wenyuan Xu. 2017. “DolphinAttack: Inaudible Voice Commands.” In Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security, 103–117. https://doi.org/10.1145/3133956.3134052.
Sharath Chandra Ramakrishnan
School of Arts, Technology, and Emerging Communication
University of Texas at Dallas
looplogic@gmail.com

Sharath Chandra Ramakrishnan is a creative technologist, hybrid practitioner and educator across the cognitive sciences, human machine interaction and technology policy. He is the Director of the Signal Cultures Lab that investigates creative possibilities and techno-cultural implications of pervasive AI technologies of human and machine listening. Previously as a cognitive neuroscience researcher, he studied networks of multimodal and audio cognition in sound and language processing at the National Institute of Mental Health and Neurosciences, Bangalore, India. He is a licensed amateur radio broad-caster (callsign: VU3HPA), extending his creative practice with sound and signals in the wireless spectrum as a transmission & signal artist. His current research in the field of Art & Technology seeks to make novel contributions to the fields of Sound Studies, Auditory Cognition and Machine Listening, prior to which he specialised in AI and interactive virtual environments at the University of Edinburgh, School of Informatics.

Twitter: @AgentSpock
ORCID: https://orcid.org/0000-0001-7984-9442
ARTISTS, ARTIFICIAL INTELLIGENCE AND MACHINE-BASED CREATIVITY IN PLAYFORM

Ahmed Elgammal
Rutgers, The State University of New Jersey

Marian Mazzone
College of Charleston

Abstract
Artificial intelligence researchers and artists have trained machines and generative processes to produce visually interesting and novel works, thereby devising machinic means of creativity. At Artrendex, Playform was developed as an easy-to-use programme specifically to be used by a broad range of artists, from beginners to those with advanced technical skills. This essay focuses on the motivations behind the development of Playform and the early reception and use of it by some artists. Our aim is to better understand both human and machine-based creativity at their intersection in an art-generating system such as Playform.

Keywords
artificial intelligence, creativity, generative art, Playform
Artistas, inteligencia artificial y creatividad basada en máquinas en Playform

Resumen
Los investigadores y artistas de inteligencia artificial han entrenado máquinas y procesos generativos para producir obras visualmente interesantes y novedosas, dando lugar a una obra creativa “maquinica”. Playform es un producto desarrollado por Artrendex pensando para que sea una herramienta fácil de usar por diferentes tipos de artistas, desde los principiantes hasta aquellos que cuentan con habilidades técnicas avanzadas. Este artículo se centra en las motivaciones que hay detrás del desarrollo de Playform y la recepción y uso que ha tenido en algunos artistas. Nuestro objetivo es comprender mejor la creatividad, tanto si viene de seres humanos como de máquinas, en su intersección en un sistema generador de arte como Playform.

Palabras clave
inteligencia artificial, creatividad, arte generativo, Playform

The Use of Artificial Intelligence in Art Making

As AI becomes incorporated into more aspects of our daily lives from our phones to driving our cars, it is only natural that artists would start to experiment with artificial intelligence. However, this is not an entirely new trend. Since the dawn of AI more than 50 years ago, artists have been writing computer programs to generate art, in some cases incorporating intelligent elements. The most prominent early example of such work is by Harold Cohen and his art-making programme AARON, which produced drawings that followed a set of rules Cohen had hard-coded. But AI has evolved over the past couple of decades to incorporate machine learning technology. One result is a new wave that uses AI in different ways to make art. In contrast to traditional algorithmic art in which the artist had to write detailed code that specified the rules for the desired aesthetics beforehand, now algorithms are set up by an artist to “learn” aesthetics by looking at many images using machine learning. The algorithm only then generates new images that follow the aesthetics it has learned. The most widely used tool in this class is Generative Adversarial Networks (GANs), introduced by Goodfellow in 2014 (Goodfellow 2014), which have been successful in many applications in the AI community. It is the development of GANs that has sparked this new wave of AI art.

Figure 1 charts the creative process involved in making art using GAN-like algorithms. The artist chooses a collection of images to feed the algorithm (pre-curation). These images are then fed into a generative AI algorithm that tries to imitate these inputs. In the final step the artist sifts through many output images to curate a final collection (post-curation).

At Artrendex we developed Playform (www.Playform.io) as an AI art studio for artists to use generative AI systems in their creative process. Our goal is to make the technology accessible to artists, solving several problems and reducing challenges that face artists and creatives when attempting to use this technology. We wanted artists to be able to explore and experiment with AI as part of their own creative process, without worrying about AI terminology and without the need to code or to run open-source code that is meant for experts. We also wanted to help artists avoid navigating unguided through the vast ocean of AI- and GAN-like algorithms.

Most generative-AI algorithms are developed by researchers in academia or large corporate research labs to push the boundaries of the technology. Artists are not the target audience of these algorithms. However, an artist's use of these algorithms is an act of creativity. Every artist must be imaginative and highly adaptable to adopt, bend and apply such non-specialised tools for their purposes. Playform fits the creative process of different artists, from looking for inspiration, to preparation of assets, all the way to producing final works. Figure 2 shows the workflow in Playform.

On the research and development side, we had to address the problem that GANs require large numbers of images and long hours of training. So we developed proprietary optimised versions of GANs that could be trained with tens of images instead of tens of thousands, and could produce reasonable results in a matter of one or two hours.
For the design our focus was on making an intuitive user experience free of AI jargon. All of the AI is hidden under the hood. Users choose a creative process, upload their own images and press a button to start training. Within minutes results will start to pop up and evolve as the training continues. Within an hour or a bit more the process is done and the system will have already generated thousands of images. Users can navigate through these iterations to find their favourite results. Users can also continue the training process as needed to achieve better results.

Challenges in Using AI in Art Making

Since the introductions of GANs in 2014, there has been an explosion in research in the AI community for developing new types of GANs, and addressing limitations and extending the capabilities of GANs as a generative engine for images, language, and music. The sheer volume of research activity makes it nearly impossible for an artist to even know where to start to use this technology. For example, going to the code repository GitHub where developers deposit their open source codes, if you search for the term “GAN” you will find tens of thousands of GAN variants available, X-GAN, Y-GAN, Z-GAN. Artists can be overwhelmed by this ocean of GAN-like algorithms, left wondering where to start and how to find an algorithm to suit their creative processes.

Even with the availability of open-source codes several challenges still face artists. If you are not a code developer who is familiar with today’s programming languages and up to date with the latest AI libraries, it is very unlikely that you would be able to benefit from existing open-source codes. Moreover, running such sophisticated AI programs requires the availability of GPUs (Graphical Processing Units), specialised hardware boards that accelerate processing speed many times (10- to 100-fold) to train AI models in hours or days instead of several weeks. The price of a GPU board that is able to run state-of-the-art AI algorithms starts at around $2000. Some platforms allow users to use cloud-based GPUs to run open-source codes with hourly charges, but these charges quickly add up to a substantial bill if you are not sure what are you doing. There are other experimental software systems available to users which offer a variety of tools to a broad range of creatives and designers with capabilities for text, audio and image generation/manipulation. Playform was built exclusively for visual artists to create works of art, and designed by a team whose principal area of research is machine learning of images in the context of fine art.

Another challenge for artists is that GAN-like algorithms require a huge set of training images (numbering in the tens of thousands) to generate reasonable results. Most algorithms are trained and tested on the available image datasets typically curated for AI research. Instead, artists generally prefer to use their own image collections in their projects. With Playform we noticed that artists choose to train their AI algorithms with sets of less than 100 images. Such a small number of images is not sufficient to train any off-the-shelf AI algorithm for creating desired results. Also, as a non-AI expert one is faced with mastering a vast number of technical terms to achieve the minimum level of understanding necessary to be in control of the process. These include terms/concepts such as training, loss, overfitting, mode collapse, layers, learning rate, kernels, channels, iterations, batch size, and additional artificial intelligence jargon. Artists might try experimenting with these systems to gain interesting results, but too often quit because of the sheer difficulty of the technology. Given the cost of GPU time and the lengthy process, this approach also results in frustration and many hours of wasted time and resources without generating anything useful or interesting.

Playform in the Context of Contemporary Art

Some artists build and implement their own generative systems to make art, but for the reasons outlined above, most do not have the means nor the time to learn the requisite computer science skills to experiment with generative systems and artificial intelligence. Playform may be used by artists of any skill level and is designed with a simple interface that fronts a powerful generative system. For those with limited skills the interface can be easily learned, there are databases of imagery already built into the system, and experimenting with the generative process results in outputs within an hour or two. Artists may also import their own image databases. Results will reflect the quality of what is put in, but a first step is learning how generative systems work and experimenting with the results. Time spent and more sophisticated inputs will produce higher quality work over time. On one level Playform facilitates discovery and experimentation and eases novice practitioners into the rapidly changing field of AI use in art. Artists who continue to use Playform should expect greater mastery and control of how their art is developed with AI.

Some quick context for readers who may have less familiarity with contemporary art follows. Many artists, whether they use com-
Artists, Artificial Intelligence and Machine-based Creativity in Playform

Artists’ Experience of Playform

Over the first six months of testing Playform beta version, about 300 artists have been exploring different innovative ways to integrate AI in their work through Playform. Some artists used Playform as a mean of looking for inspiration based on AI uncanny aesthetics. Other artists fed in images of their own art, training models that learn their style and then use these models to generate new artworks based on new inspirations. VR artists used AI to generate digital assets to be integrated in virtual reality experiences. Playform has also been used to generate works that are upscaled and printed as a final art product. For this study we focused on artists using Playform for some time and/or who already had substantial experience with using technology in their art making. What we present here are admittedly preliminary results given the newness of the Playform platform. We selected a few artists to interview, focusing on those who have experimented with the system and its capabilities in depth over the past six months, resulting in work that has appeared in exhibition. Our goal was to gather qualitative responses to questions about why they chose to work with AI and a system such as Playform, how it has impacted their creative practices, and how they would characterise the experience of creating art while using artificial intelligence. Some of the artists surveyed for this essay include Domenico Barra [http://www.dombarra.art], Tom Brown [https://www.41xrt.com/art], Qinza Najm [http://qinzanajm.com], and Anne Spalter [https://annespalter.com]. We conducted interviews online, via email and by phone, and guided the discussions with three key questions: Why do you choose to work with AI? What is the role of AI in your art making: a tool, a medium, or as a partner/collaborator?, and how has Playform affected your creative practice, if it has? We wanted to hear directly from artists about their experiences. The following is a qualitative summary of key concepts the artists focused on in their conversations with us.

There was consensus on the positive creative aspects of the generative capabilities of the Playform system. Beyond the obvious of using such a system to create lots of work, it seems critical that this is done extremely quickly and with minimum effort. The advantages of this are several: an artist can rapidly decide what to keep and what to dispose of without regret for time/resources spent (since that is low) and the artist can decide to continue or quickly terminate a trend in the imagery without much loss. One artist pithily described this as having ‘lots of shots on goal’. Artists also report a higher number of surprising or unexpected results than they first anticipated. This may be the result of expecting a high degree of predictable and routinised results because Playform is a computational generative system; we assume computers will repeat at a high degree of predictable and routinised results because they already-existing imagery taken from somewhere other than their own imaginations. In the context of contemporary art what matters is what imagery you select, how you manipulate it, and how you present it to an audience to see it. That is the heart of the creative enterprise, and is the process that Playform facilitates. Devising the means to manipulate and output imagery at considerable volume is a familiar tactic in contemporary art reaching back at least to the 1950s. From Rauschenberg and Johns forward artists have been using repetition rendered through technological processes to multiply imagery, most often found imagery. The pop artists amplified this practice by using silk screening, stencils and mass production techniques to make many versions and variations of any one image. Viewers were trained to recognise pattern and variation in an example such as Marilyn Monroe in cherry red, turquoise, or yellow. Conceptual art practices led viewers to expect word and imagery in text, photography, Xerox-copy and all other forms of printed and mass-produced formats on gallery walls, on the page, and on our screens. Now when computer-repeated imagery is presented, viewers are expected to detect repetition (pattern) and variation (parameterisation) produced via a generative system such as Playform. In addition, conceptual art and earth/land art development from the 1960s forward had led artists to think about ‘systems’: systems of language and imagery, systems of nature and culture and how systems have rules, logic and methods that artists can either manipulate from within or disrupt from without. Because they are a part of our contemporary world, computational image systems should be part of our art as well. It fails to artists both to learn to create from within the AI generative systems and perhaps disrupt and forge new patterns of creativity in these systems.

Figure 3: Example of the creative process of artist Thomas W. Brown in Playform. Top Left: sample of inspiration images of a project. Bottom Left: a snapshot of some results out of hundreds of thousands of iterations. Right: A final artwork “Who Votes IV – 2019”, exhibited at Nails in the Wall Gallery, Metuchen, NJ, February 2019.
There is also an important distinction to draw here. Although artistic media can fail to cooperate and even make ‘happy accidents’, the inert physical material is not responding in any calculated or intelligent way. It merely and truly is an accident. However, computers are programmed to operate on their own pathways of choice and intelligence, and the ‘accident’ is less an accident and more an unanticipated result from the human viewpoint.

Human artists also must take on an expanded role within the feedback loop developing between them and a generative system such as Playform. Some artists describe the expanded role as being like a curator in that the artist is making qualitative and discriminatory decisions about which imagery to use and which imagery to continue to develop within the system. Another analogy is to a DJ, sampling, mixing and manipulating imagery choices and combinations with another level of remixing what the generative system is itself mixing. Playform allows artists to choose from databanks of images within the system, and to import their own data of other images or works of their own art. The choices any one artist makes about importing data (or not) and using pre-existing databanks (or not) is another (meta) level of the curatorial or DJ selection process. There is a complicated choice of data, and there is a complex action of data manipulation.

In both cases we can detect an expanding role for the generative system as a more active participant in the creative process. Choices are being made by the system which the human artist is responding to, and vice versa. The feedback loop can continue with each half of the dynamic responding and remixing. Artist Qinza Najm likened the presence of the AI system as an alien eye, something that is seeing, but seeing very differently (see Figure 4). It has a role to play in that it responds to choices and datasets in a surprising or perplexing way, revealing new choices or options. It makes the artist see their work differently as a consequence. It is almost an estrangement from one’s own familiar work choices.

The greatest diversity in responses was to the question about AI being a tool, a medium or a creative partner in some capacity. For Tom Brown the system is a tool, a machine-based generator of lots of possibilities, but it is the artist in the role of choice maker of lots of possibilities, but it is the artist in the role of choice maker who makes the choices. This provides a more active role for the artist. For Anne Spalter the system is a more active participant in the creative process. Choices any one artist makes about importing data (or not) and using pre-existing databanks (or not) is another (meta) level of the curatorial or DJ selection process. There is a complicated choice of data, and there is a complex action of data manipulation.

In both cases we can detect an expanding role for the generative system as a more active participant in the creative process. Choices are being made by the system which the human artist is responding to, and vice versa. The feedback loop can continue with each half of the dynamic responding and remixing. Artist Qinza Najm likened the presence of the AI system as an alien eye, something that is seeing, but seeing very differently (see Figure 4). It has a role to play in that it responds to choices and datasets in a surprising or perplexing way, revealing new choices or options. It makes the artist see their work differently as a consequence. It is almost an estrangement from one’s own familiar work choices.

The greatest diversity in responses was to the question about AI being a tool, a medium or a creative partner in some capacity. For Tom Brown the system is a tool, a machine-based generator of lots of possibilities, but it is the artist in the role of choice maker of lots of possibilities, but it is the artist in the role of choice maker who makes the choices. This provides a more active role for the artist. For Anne Spalter the system is a more active participant in the creative process. Choices any one artist makes about importing data (or not) and using pre-existing databanks (or not) is another (meta) level of the curatorial or DJ selection process. There is a complicated choice of data, and there is a complex action of data manipulation.

In both cases we can detect an expanding role for the generative system as a more active participant in the creative process. Choices are being made by the system which the human artist is responding to, and vice versa. The feedback loop can continue with each half of the dynamic responding and remixing. Artist Qinza Najm likened the presence of the AI system as an alien eye, something that is seeing, but seeing very differently (see Figure 4). It has a role to play in that it responds to choices and datasets in a surprising or perplexing way, revealing new choices or options. It makes the artist see their work differently as a consequence. It is almost an estrangement from one’s own familiar work choices.

The greatest diversity in responses was to the question about AI being a tool, a medium or a creative partner in some capacity. For Tom Brown the system is a tool, a machine-based generator of lots of possibilities, but it is the artist in the role of choice maker of lots of possibilities, but it is the artist in the role of choice maker who makes the choices. This provides a more active role for the artist. For Anne Spalter the system is a more active participant in the creative process. Choices any one artist makes about importing data (or not) and using pre-existing databanks (or not) is another (meta) level of the curatorial or DJ selection process. There is a complicated choice of data, and there is a complex action of data manipulation.

In both cases we can detect an expanding role for the generative system as a more active participant in the creative process. Choices are being made by the system which the human artist is responding to, and vice versa. The feedback loop can continue with each half of the dynamic responding and remixing. Artist Qinza Najm likened the presence of the AI system as an alien eye, something that is seeing, but seeing very differently (see Figure 4). It has a role to play in that it responds to choices and datasets in a surprising or perplexing way, revealing new choices or options. It makes the artist see their work differently as a consequence. It is almost an estrangement from one’s own familiar work choices.

The greatest diversity in responses was to the question about AI being a tool, a medium or a creative partner in some capacity. For Tom Brown the system is a tool, a machine-based generator of lots of possibilities, but it is the artist in the role of choice maker of lots of possibilities, but it is the artist in the role of choice maker who makes the choices. This provides a more active role for the artist. For Anne Spalter the system is a more active participant in the creative process. Choices any one artist makes about importing data (or not) and using pre-existing databanks (or not) is another (meta) level of the curatorial or DJ selection process. There is a complicated choice of data, and there is a complex action of data manipulation.

In both cases we can detect an expanding role for the generative system as a more active participant in the creative process. Choices are being made by the system which the human artist is responding to, and vice versa. The feedback loop can continue with each half of the dynamic responding and remixing. Artist Qinza Najm likened the presence of the AI system as an alien eye, something that is seeing, but seeing very differently (see Figure 4). It has a role to play in that it responds to choices and datasets in a surprising or perplexing way, revealing new choices or options. It makes the artist see their work differently as a consequence. It is almost an estrangement from one’s own familiar work choices.

The greatest diversity in responses was to the question about AI being a tool, a medium or a creative partner in some capacity. For Tom Brown the system is a tool, a machine-based generator of lots of possibilities, but it is the artist in the role of choice maker of lots of possibilities, but it is the artist in the role of choice maker who makes the choices. This provides a more active role for the artist. For Anne Spalter the system is a more active participant in the creative process. Choices any one artist makes about importing data (or not) and using pre-existing databanks (or not) is another (meta) level of the curatorial or DJ selection process. There is a complicated choice of data, and there is a complex action of data manipulation.
Limits and Conclusion

As the uses of AI continue to expand in our world, computational systems will be taught to model certain human thought processes such as creativity. Playform was developed as a way to make AI-assisted art more accessible and functional for all creators. The system is six months into development and we currently host over 1800 hundred users. The interviews with early adapters suggest some immediate gains for artists in scale of imagery generation, surprising prompts for new imagery ideas, and overall efficiency of creative output. Also interesting is the variety of responses to how the artist conceptualises their relationship to the AI system, as tool, as medium, as some degree of partner. Artificial intelligence is a means to better study and understand art by training AI systems to be creative. And as we have demonstrated, creative AI systems can expand and inspire human creativity in turn.

Notes

Goodfellow, Ian, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. 2014. “Generative adversarial nets.” In NIPS ’14: Proceedings of the 27th International Conference on Neural Information Processing Systems Volume 2, 2672-2680. Montreal: MIT Press.

For Further Reading

Agüera y Arcas, Blaise. “Art in the Age of Machine Intelligence.” Arts 6, no. 4 (2017): 18. https://doi.org/10.3390/arts6040018.

Hertzmann, Aaron. “Can Computers Create Art?” Arts 7, no. 2 (2018): 18. https://doi.org/10.3390/arts7020018.

Mazzone, Marian, and Ahmed Elgammal. “Art, Creativity, and the Potential of Artificial Intelligence.” Arts 8, no. 1 (2019): 26. https://doi.org/10.3390/arts8010026.

McCorduck, Pamela. Aaron’s Code: Meta-Art, Artificial Intelligence, and the Work of Harold Cohen. New York: W. H. Freeman and Company, 1991.

Miller, Arthur I. The Artist in the Machine: The World of AI-Powered Creativity. The MIT Press, 2019.
CV

Ahmed Elgammal
Rutgers, The State University of New Jersey

Dr. Ahmed Elgammal is a professor at the Department of Computer Science and is part of the Executive Council Faculty at the Center for Cognitive Science at Rutgers University. He is the founder and director of the Art and Artificial Intelligence Laboratory at Rutgers, which focuses on data science in the culture domain. Prof. Elgammal has published over 160 peer-reviewed papers, book chapters, and books in the fields of computer vision, machine learning and digital humanities. He is a senior member of the Institute of Electrical and Electronics Engineers (IEEE). He received the National Science Foundation CAREER Award in 2006. Dr. Elgammal's research on knowledge discovery in art history and AI-art generation received wide international media attention, including several reports in the Washington Post, New York Times, NBC, CBS News, the Daily Telegraph, Science News, and many others. Dr. Elgammal received his M.Sc. and Ph.D. degrees in computer science from the University of Maryland, College Park, in 2000 and 2002, respectively.
CV

Marian Mazzone
Dept. of Art & Architectural History
College of Charleston

Marian Mazzone is an associate professor in the Art & Architectural History Department at the College of Charleston, teaching courses on contemporary art and new media art. She is an affiliate of the Art & Artificial Intelligence Lab at Rutgers University and one of the co-founders of the Computing in the Arts major at the College of Charleston. Mazzone has published on art, technology and related issues in Leonardo, Arts, and Technoetic Arts.

ORCID: 0000-0002-4016-6297
ARTÍCULO
NODO «IA, ARTE Y DISEÑO: CUESTIONANDO EL APRENDIZAJE AUTOMÁTICO»

El factor estético en la automatización de tareas lógicas: el caso del ajedrez

Santiago Rementeria-Sanz
GAIKER Technology Centre, Basque Research and Technology Alliance (BRTA)

Fecha de presentación: febrero de 2020
Fecha de aceptación: junio de 2020
Fecha de publicación: julio de 2020

Cita recomendada
Rementeria-Sanz, Santiago (2020). «El factor estético en la automatización de tareas lógicas: el caso del ajedrez». En: Andrés Burbano; Ruth West (coord.) «AI, Arts & Design: Questioning Learning Machines». Artnodes, N.º. 26: 1-10. UOC. [Consulta: dd/mm/aa].
http://doi.org/10.7238/a.v0i26.3338

Resumen
Es frecuente entre los ajedrecistas expertos la apelación a características del juego consideradas artísticas. Sin embargo, y a pesar de ser el ajedrez uno de los campos más estudiados en la historia de la inteligencia artificial, han sido pocos los intentos de automatizar el juicio estético de los grandes maestros. Ha prevalecido la demanda de eficacia, recurriendo a la potencia de búsqueda de arquitecturas ‘hardware’ especializadas, complementadas con métodos heurísticos y criterios de evaluación codificados manualmente. Recientemente, sin embargo, el programa AlphaZero, desarrollado siguiendo una filosofía diferente, ha sorprendido no solo por su superioridad, sino también por lo imprevisto y atractivo de algunas de sus decisiones de juego. En este artículo se repasan los criterios de belleza asumidos en la práctica del ajedrez y se analiza el modo en que AlphaZero adquiere su capacidad experta, llegando a conclusiones sobre el papel del impulso estético en la toma de decisiones racional y sobre el modo en que los sistemas artificiales pueden mostrar un comportamiento original y atractivo en actividades lógicas como el ajedrez competitivo.
Palabras clave
ajedrez por ordenador, estética, toma de decisiones, aprendizaje automático, inteligencia artificial.

**The aesthetic factor in the automation of logical tasks: The case of chess**

**Abstract**
It is common among chess players to refer to game features considered to be artistic. However, despite chess being one of the most studied fields in artificial intelligence, far too little effort has been made to automate the aesthetic judgement of grandmasters. The quest for effectiveness has prevailed, resorting to the search power of dedicated hardware architectures complemented by heuristic methods and handcrafted evaluation criteria. Recently, however, the AlphaZero program, developed according to a different philosophy, has surprised not only because of its dominance, but also because of the unexpectedness and appeal of some of its game decisions. In this article we provide an overview of the beauty criteria assumed in chess practice, and analyse how AlphaZero acquires its expertise, drawing conclusions about the role of the aesthetic impulse in rational decision making and on how artificial systems can get to perform in an original and appealing manner in logical tasks like competitive chess.

**Keywords**
computer chess, aesthetics, decision making, machine learning, artificial intelligence

---

**Introducción**

El ajedrez es el paradigma de juego de tablero en la cultura occidental y la aplicación más largamente estudiada de la investigación en inteligencia artificial (Silver et al., 2018: 1140). Aunque se trata de un juego de suma cero e información perfecta, la complejidad de su estructura combinatoria hace que no exista una estrategia demostrablemente óptima, pues, aunque el número total de movimientos alternos de las piezas es teóricamente finito, en la práctica resulta intratable.

El ajedrez computacional ha sido productivo en el sentido de que ha posibilitado avances en técnicas concretas, pero no ha favorecido el desarrollo de nuevas teorías sobre los procesos cognitivos ni sobre informática teórica (Ensmenger, 2012: 23). La aspiración de explorar los mecanismos de la inteligencia y simular los procesos del pensamiento humano se ha visto sobrepasada por el afán competitivo. El progreso ha venido de la mano de enormes potencias de cálculo que realizan búsquedas profundas en el árbol de movimientos permitidos, recurriendo a criterios heurísticos para limitar la exploración y utilizando funciones de evaluación afinadas manualmente (Newell y Simon, 1987: 314–317)1. La arquitectura masivamente paralela de Deep Blue, el sistema dedicado que en 1997 venció al campeón del mundo Garry Kasparov en un match a seis partidas y en las condiciones de control de tiempos de los torneos oficiales, era capaz de explorar hasta 200 millones de posiciones por segundo (Hsu, 2004). Todo ello llevó a lamentarse por la desviación de esta línea de investigación respecto a sus objetivos iniciales. Aterrándose a la analogía clásica entre el papel que desempeñan los programas de juegos como soporte de experimentación en inteligencia artificial y el de la mosca de la fruta (Drosophila melanogaster) en la genética, McCarthy afirmaba: «El ajedrez por ordenador se ha desarrollado de forma similar a como lo habría hecho la genética si los genetistas hubiesen concentrado sus esfuerzos comenzados en 1910 en la reproducción de Drosophila de carreras. Tendríamos algo de ciencia, pero sobre todo tendríamos moscas de la fruta muy rápidas» (1997: 1518).

Recientemente, sin embargo, se ha presentado un sistema capaz de vencer a los mejores programas de ajedrez, shogi (versión japonesa del anterior) y go, a su vez más capaces que los campeones humanos respectivos. Lo que podría parecer una mejora incremental en la codificación progresiva de estrategias de juego especializadas obedece, en realidad, a un cambio radical en el planteamiento de diseño de AlphaZero, como se denomina el sistema en cuestión. Su principal novedad reside en el hecho de que prescinde de la implantación preliminar de conocimiento específico y del acceso a bases de datos externas, aprendiendo a competir al nivel más alto jamás observado a partir del mero conocimiento de las reglas del ajedrez.

---

1Cuando el tamaño del espacio de soluciones hace imposible una búsqueda exhaustiva en el tiempo disponible, se recurre a opciones estratérgicas sugeridas por la experiencia, denominadas heurísticas. La búsqueda heurística no garantiza el éxito, pero permite orientar la exploración descartando los movimientos menos prometedores.
juego y tras entrenarse disputando contra sí mismo una gran cantidad de partidas (Silver et al., 2018).

Además de su superioridad competitiva, AlphaZero ha llamado la atención por la elegancia de algunas de las combinaciones exhibidas. Es esta novedad en el estilo de juego del sistema artificial la que lleva a repasar en las páginas siguientes los principales criterios de belleza contemplados en la práctica del ajedrez y comparar el concepto de diseño del nuevo sistema con el de los programas tradicionales para comprender por qué se comporta como lo hace. Se analiza también la naturaleza del juicio estético de los ajedrecistas de primer nivel, su posible relación con criterios prácticos aplicados intuitivamente y el modo en que AlphaZero emula esa capacidad.

La discusión en este artículo se centra en el ajedrez de competición, dejando de lado las composiciones, los problemas o estudios de finales de partidas que, aunque en opinión de autores como Ravilious (1994) constituyen el contexto en el que las cualidades estéticas del juego se manifiestan más plenamente, para otros muchos no resultan, en cambio, suficientemente realistas (Levitt y Friedgood, 2008).

1. El juicio estético en el ajedrez

En opinión de Bronstein y Smolyan, es el sentimiento estético el que orienta al intelecto en su proceso de búsqueda de combinaciones satisfactorias (1983: 27-28). La armonía en el desarrollo de la partida, la expresividad geométrica sobre el tablero y el despliegue imaginativo de las piezas reflejan un sentido de belleza que puede guiar a la mente entrenada a seleccionar el movimiento más adecuado. El jugador que saca provecho de su criterio estético puede identificar las exigencias y limitaciones de una posición determinada, concebir un plan lógicamente consistente, decidir un movimiento concreto o reconocer una combinación latente. Todo lo que nos resulta atractivo en una partida, sea una maniobra, trampa o final, el despliegue tático o la armonía geométrica de la coordinación entre piezas, serían ejemplos de lo que denominan invariantes estéticos.

Estos autores analizan varios factores que tienen que ver con la creatividad en el ajedrez: 1) el placer que experimenta un jugador al concebir jugadas artísticas dignas de ser recordadas y registradas; 2) el gusto de los espectadores al disfrutar de una partida interesante; 3) «la misteriosa belleza del juego» y la posibilidad de evaluarla en función de cualidades como el atrevimiento, la imaginación y la fantasía; y 4) el entorno del juego que, aunque gobernado por un conjunto de reglas, se muestra lo suficientemente variado y flexible como para permitir al jugador expresar sus dotes artísticas. Así la partida puede ser apreciada por los espectadores —afirman— de modo similar a como se goza de la música o la arquitectura.

En esta misma línea, Humble considera que los lances de ajedrez deben disfrutarse principalmente como obras de arte, admitiendo que acaso debamos considerarlos menores en comparación con las grandes creaciones artísticas que tratan sobre temas humanos mucho más profundos. El vocabulario expresivo es, además, notablemente más limitado que el disponible en otras formas de arte (1993: 62).

Desde la perspectiva de este autor, una partida puede considerarse una obra artística si y solo si: 1) se produce con la intención principal de aportar una recompensa estética; 2) muestra características estéticas, entendiéndolo por tales las que requieren el ejercicio del gusto, la percepción o la sensibilidad para su apreciación o discriminación; y 3) es única. Siendo el número de posiciones legales sobre el tablero del orden de $10^{43}$, y las posibles partidas de cuarenta jugadas cercanas a $10^{120}$, existe margen para no repetirse (Fox y James, 1987: 166).

El guido del debate radica, sobre todo, en la primera condición, que puede resultar extrema por su carácter desinteresado. Al fin y al cabo, el ajedrez es un juego competitivo y se puede considerar que el objetivo de toda partida es la búsqueda del triunfo. Los campeones son auténticos combates intelectuales en los que se combinan argumentos psicológicos con los estrictamente asociados al juego para conquistar un reconocimiento que, además de prestigio, puede conllevar pingües retornos materiales. Por ello, resulta comúnmente aceptado que la naturaleza contendiente del ajedrez no está reñida con su condición artística. Osborne, por ejemplo, especula con la idea de que las características competitivas del juego pueden «proporcionar el interés necesario para captar y mantener la atención antes de que tenga lugar la apreciación de la belleza» (1964: 162). También para Humble (1993) algunas de las valoraciones estéticas que podemos realizar como espectadores guardan relación con la vertiente más deportiva o dramática del juego.

Un encuentro se puede juzgar bello en función de las dificultades que el jugador presente a su contrincante e igualmente se ensanzan estéticamente los movimientos con los que se asumen riesgos. En estas circunstancias, el uso de términos estéticos estaría relacionado con el criterio de dificultad y, por consiguiente, con consideraciones de tipo competitivo. De igual modo, la partida dada por perdida que un jugador recupera y vence luchando contra la tensión del momento y los últimos segundos del tiempo disponible es estéticamente más valorada que el mismo patrón de movimientos, formalmente elegante, armonioso e inventivo, si este se ha generado en la tranquilidad del estudio privado.

Los grandes campeones mundiales lo han expresado con diferentes palabras. Lasker, que conservó el título durante veintisiete años, dedicó un capítulo de su manual clásico de 1925 al «efecto estético del ajedrez» (1960: 261-283). En su opinión, la jugada bella debía conseguir un logro que se produzca con la intención principal de aportar una recompensa estética; 2) muestra características estéticas, entendiéndolo por tales las que requieren el ejercicio del gusto, la percepción o la sensibilidad para su apreciación o discriminación; y 3) es única. Siendo el número de posiciones legales sobre el tablero del orden de $10^{43}$, y las posibles partidas de cuarenta jugadas cercanas a $10^{120}$, existe margen para no repetirse (Fox y James, 1987: 166).

El guido del debate radica, sobre todo, en la primera condición, que puede resultar extrema por su carácter desinteresado. Al fin y al cabo, el ajedrez es un juego competitivo y se puede considerar que el objetivo de toda partida es la búsqueda del triunfo. Los campeones son auténticos combates intelectuales en los que se combinan argumentos psicológicos con los estrictamente asociados al juego para conquistar un reconocimiento que, además de prestigio, puede conllevar pingües retornos materiales. Por ello, resulta comúnmente aceptado que la naturaleza contendiente del ajedrez no está reñida con su condición artística. Osborne, por ejemplo, especula con la idea de que las características competitivas del juego pueden «proporcionar el interés necesario para captar y mantener la atención antes de que tenga lugar la apreciación de la belleza» (1964: 162). También para Humble (1993) algunas de las valoraciones estéticas que podemos realizar como espectadores guardan relación con la vertiente más deportiva o dramática del juego.

Un encuentro se puede juzgar bello en función de las dificultades que el jugador presente a su contrincante e igualmente se ensanzan estéticamente los movimientos con los que se asumen riesgos. En estas circunstancias, el uso de términos estéticos estaría relacionado con el criterio de dificultad y, por consiguiente, con consideraciones de tipo competitivo. De igual modo, la partida dada por perdida que un jugador recupera y vence luchando contra la tensión del momento y los últimos segundos del tiempo disponible es estéticamente más valorada que el mismo patrón de movimientos, formalmente elegante, armonioso e inventivo, si este se ha generado en la tranquilidad del estudio privado.

Los grandes campeones mundiales lo han expresado con diferentes palabras. Lasker, que conservó el título durante veintisiete años, dedicó un capítulo de su manual clásico de 1925 al «efecto estético del ajedrez» (1960: 261-283). En su opinión, la jugada bella debía conseguir un logro que se produzca con la intención principal de aportar una recompensa estética; 2) muestra características estéticas, entendiéndolo por tales las que requieren el ejercicio del gusto, la percepción o la sensibilidad para su apreciación o discriminación; y 3) es única. Siendo el número de posiciones legales sobre el tablero del orden de $10^{43}$, y las posibles partidas de cuarenta jugadas cercanas a $10^{120}$, existe margen para no repetirse (Fox y James, 1987: 166).

El guido del debate radica, sobre todo, en la primera condición, que puede resultar extrema por su carácter desinteresado. Al fin y al cabo, el ajedrez es un juego competitivo y se puede considerar que el objetivo de toda partida es la búsqueda del triunfo. Los campeones son auténticos combates intelectuales en los que se combinan argumentos psicológicos con los estrictamente asociados al juego para conquistar un reconocimiento que, además de prestigio, puede conllevar pingües retornos materiales. Por ello, resulta comúnmente aceptado que la naturaleza contendiente del ajedrez no está reñida con su condición artística. Osborne, por ejemplo, especula con la idea de que las características competitivas del juego pueden «proporcionar el interés necesario para captar y mantener la atención antes de que tenga lugar la apreciación de la belleza» (1964: 162). También para Humble (1993) algunas de las valoraciones estéticas que podemos realizar como espectadores guardan relación con la vertiente más deportiva o dramática del juego.
superado de ninguna manera a sus oponentes por medio de la lógica pura y simple, sino que fueron asistidos más a menudo que lo contraria por la imaginación. En el ajedrez practicado por un buen jugador la lógica y la imaginación deben ir de la mano, complementándose entre sí» (Winter, 1989: 301). La escuela hipermoderna de ajedrez, que floreció en los años veinte del siglo pasado, concedía gran importancia a la originalidad y se propuso revolucionar el juego valorando la creatividad y la consecución de la victoria de una manera brillante y artística (Saidy, 2008: 29-32). «Quiero ganar, deseo vencer a todo el mundo, pero quiero hacerlo con estilo», decía más recientemente Kasparov, dejando así patente su aspiración no solo a superar a sus adversarios, sino a hacerlo de una manera elegante (1986: 2). En realidad, prácticamente todos los campeones mundiales muestran interés en el aspecto estético del ajedrez y reconocen que este ha contribuido al desarrollo de su juego individual (Levitt y Friedgood, 2008).

2. Formalización de atributos estéticos

Juzgar las cualidades estéticas es una tarea subjetiva que suele estar afectada, además, por elementos culturales. La automatización del juicio estético resulta problemática por la dificultad de concreción y valoración de las características relevantes, pues los propios expertos tienden a ser poco específicos sobre los elementos subjetivos constitutivos de la belleza. La ventaja en entornos acotados como el del ajedrez es que las reglas del juego son idénticas en todas partes, al menos en su versión internacional, y la competición apenas está afectada, además, por elementos culturales. La automatización del juego inicial), paradoja (movimientos aparentemente en contra del sentido común del juego), unidad (alineamiento de los movimientos de todas las piezas hacia el objetivo común) y originalidad (cuya valoración contiene una dosis de subjetividad).

Otras fuentes recogen ideas similares. El gran maestro Levitt y el problemista Friedgood, por ejemplo, identifican cuatro elementos de belleza en el ajedrez (2008: la paradoja (hacer algo que normalmente no se esperaría de uno), la profundidad (de forma que el objetivo del movimiento clave no se comprenda al principio, pero se aclara más adelante), la geometría (formación de efectos visuales o patrones concretos sobre el tablero) y el flujo (secuencia de movimientos que resultan inevitables, frente a la posibilidad de variaciones complicadas).

3. Modelos computacionales

Ante la importancia que los jugadores avanzados otorgan al factor estético en la práctica del ajedrez, se plantea la cuestión de hasta qué punto un ordenador podría reconocer y ejecutar jugadas que las personas aprecian por su belleza. Además de contribuir al desarrollo de programas más eficaces que los tradicionales, la aplicación de modelos estéticos podría resultar útil para conseguir estilos de juego más atractivos, simular distintas personalidades de contendientes, automatizar la búsqueda de partidas bellas en bases de datos especializadas o desarrollar programas mejorados para la composición de problemas, entre otras aplicaciones.

Seguramente han sido dos los factores que han influido en la escasez de referencias sobre este tema. El primero es el énfasis...
obsesivo de las últimas décadas por mejorar el nivel de juego de los ordenadores. El segundo motivo sería la doble asunción de que: 1) no existe una manera fiable de cuantificar los aspectos menos lógicos del juego; y 2) la codificación de criterios estrictamente lógicos debería bastar a la hora de programar estrategias competitivas eficaces.

Respecto al primero de estos puntos, los autores de uno de los mejores programas de composición de problemas admiten que «no es sencillo definir conceptos como la belleza, originalidad, singularidad de la solución y dificultad de resolución» (Fainshtein y HaCohen-Kerner, 2006: 33). El segundo punto guarda relación con el modelo de programación manual que ha prevalecido históricamente en el desarrollo del ajedrez por ordenador.

A continuación, se repasan las dos alternativas existentes para incorporar conocimiento especializado en un sistema informático, que en principio serían igualmente válidas para codificar algún tipo de juicio estético en un jugador de ajedrez artificial.

3.1. Programación manual

Según la perspectiva convencional se programaría explícitamente un conjunto de criterios como los discutidos en la sección anterior, buscando emular los aspectos que un jugador experimentado valora al juzgar como bella una combinación determinada. Este es el planteamiento seguido por Walls (1997) e Iqbal y Yaacob (2008) en composiciones de finales y en el juego competitivo, respectivamente. En sus aproximaciones e implantaciones parciales parten de la idea de que las técnicas de representación y los algoritmos de búsqueda comúnmente utilizados en los programas especializados bastarían para simular la percepción estética de un ajedrecista avanzado, siempre que se consideren las reglas heurísticas adecuadas y que la función de evaluación incluya la valoración expresa de dichos criterios.

La abrumadora mayoría de los diseñadores de sistemas de ajedrez programados manualmente ha optado, sin embargo, por la codificación del conocimiento experto de carácter más racional, expresable y cuantificable, dejando de lado aspectos como el estético, de formalización difusa y cuya contribución al triunfo resulta menos evidente.

Las arquitecturas convencionales culminadas con Deep Blue y sus variantes siguen esta senda de la programación manual, recurriendo a versiones del algoritmo de exploración de árboles de decisión minimax adelantado por Claude Shannon en sus experimentos pioneros (1950), junto con criterios heurísticos como la poda alpha-beta y otras para valorar las posiciones no terminales, abundante conocimiento expreso encapsulado en numerosas reglas y pautas aportadas por jugadores expertos, y en los parámetros y coeficientes de una función de evaluación lineal. Este arsenal de recursos, junto con la posibilidad de consultar enormes bases de datos con aperturas, despliegues intermedios y finales de partidas, intenta cubrir toda eventualidad del juego sin dejar margen para sorpresas.

3.2. Aprendizaje automático

Ante la dificultad para encapsular el juicio estético de los ajedrecistas expertos en unos criterios computables que, además, favorezcan el avance hacia posiciones de tablero ventajosas, el paradigma alternativo consiste en diseñar sistemas capaces de adquirir esos criterios, u otros que se manifiesten de forma análoga, valiéndose de una facultad genérica de aprendizaje automático. Esta aptitud les permitiría distinguir las jugadas prometedoras a partir de su propia experiencia.

AlphaZero (Silver et al., 2018) es el último heredero de esta línea de trabajo emprendida por Arthur Samuel en los años cincuenta, con su programa pionero para jugar a las damas (1955), y que continuarían investigadores que han desarrollado sistemas aprendices para otros juegos de tablero (Tesauro, 1995). Las estrategias aplicadas en estos casos no han sido programadas manualmente, sino aprendidas de forma autónoma y empírica.

AlphaZero tiene dos componentes principales: una red neuronal profunda que se entrena jugando contra sí misma para aprender el equivalente a una función de evaluación (en este caso, no lineal), y un algoritmo de búsqueda para seleccionar el siguiente movimiento. La arquitectura neuronal contiene una secuencia de capas convolucionales con funciones de filtrado, normalización y rectificación que se desdobla en dos módulos especializados con estructuras diferentes. Como entrada del conjunto se representa la configuración del tablero y la historia de los últimos movimientos, y la salida consiste, por un lado, en la probabilidad de que cada movimiento posible sea el mejor en la situación actual y, por otro, en la estimación del resultado final de la partida: victoria, derrota o tablas.

La distribución de probabilidades obtenida se utiliza para guiar una búsqueda acotada en el espacio problema y el resultado final previsto por la red sirve para evaluar las posiciones analizadas durante el proceso. A diferencia de lo que ha sido habitual durante décadas, el algoritmo de exploración no es una variante del clásico alpha-beta, sino otro genérico conocido como árbol de búsqueda MonteCarlo, o MCTS, que, en vez de realizar la evaluación minimax de un subárbol, promedia la evaluación de las posiciones que contiene. AlphaZero considera unas sesenta mil posiciones por segundo antes de cada movimiento, frente a los aproximadamente sesenta millones del sistema de ajedrez convencional más avanzado.

Durante el juego se elegirá el movimiento que resulte en una mayor probabilidad de éxito tras 800 simulaciones MCTS, pero en la fase exploratoria se seleccionan de forma estocástica. Estos movimientos proporcionan la secuencia de estados del tablero que constituirán las entradas durante el entrenamiento de la red neuronal, los pesos de cuyas conexiones, aleatorios al comienzo, se ajustan según un algoritmo de aprendizaje «por refuerzo» en el que se recompensan las acciones útiles (Sutton y Barto, 1998). La actualización de dichos parámetros sigue un método de descenso del gradiente que busca minimizar el error entre el resultado estimado de la partida tras cada
movimiento y el real observado finalmente, así como maximizar la similitud entre las probabilidades obtenidas en las salidas de la red neuronal y las actualizadas por el algoritmo MCTS. A medida que la red aprende, la búsqueda también resulta más eficiente, aportando a su vez mejores estimaciones de probabilidades para entrenar la estructura neuronal. Tras cuarenta y cuatro millones de partidas disputadas contra sí mismo en nueve horas, este ciclo virtuoso de mejora permite que el sistema sea capaz de reconocer los patrones relevantes que explotará después en la competición.

Partiendo del simple conocimiento de las reglas del juego, ignorando cualquier valor relativo de las piezas y con una carencia absoluta de nociones estratégicas y tácticas, AlphaZero adquiere criterios de evaluación propios (los denominaríamos valores, estrategias e intuiciones en el caso de los ajedrecistas humanos) que le llevan a mostrar una determinación y profundidad inauditas. Su superioridad se basa en la correcta previsión por parte de la red entrenada de los efectos de cada movimiento para identificar los más prometedores, y no tanto en lo exhaustivo de sus búsquedas.

El sistema ha redescubierto por su cuenta las aperturas más frecuentes y exhibido combinaciones creativas que nunca esperaríamos de un programa de ajedrez tradicional. Los sacrificios a cambio de ventajas a largo plazo y el desprecio del valor material de las piezas frente a la toma de iniciativa y la armonía en el despliegue, por ejemplo, forman parte de un estilo «dinámico y abierto», con decisiones «arriesgadas y agresivas» (Kasparov, 2018). Entre tales decisiones, que los mejores ajedrecistas reconocen como relativamente «humanas», las hay también estéticamente atractivas2.

Entrenado exclusivamente para imponerse a su adversario, el carácter homogéneo y eficaz del conocimiento práctico asimilado por AlphaZero durante su fase de aprendizaje le lleva a exteriorizar un comportamiento con facetas que las personas asignamos a categorías diversas. El sistema artificial ignora toda convención o sesgo cultural y, entre sus decisiones, que siempre buscan el triunfo final, reconocemos tanto algunas que calificaríamos como lógicas o previsibles, como otras que nos sorprenden por su efecto inesperado y bello. Esta discriminación, sin embargo, obedece exclusivamente a la percepción del espectador humano, pues, como se ha visto, toda decisión de AlphaZero responde a un mismo principio de utilidad que el programa ha asimilado de forma autónoma. No ha sido necesario trasladar al software ningún criterio asumido en la práctica experta del ajedrez, incluidos entre estos los referidos a factores estéticos.

Mas no todo son ventajas. Este sistema demanda muchos recursos computacionales (cinco mil unidades de procesamiento específicas, o TPU, para generar las partidas de entrenamiento, sesenta y cuatro TPU de segunda generación para entrenar la red neuronal y seis TPU en la fase de evaluación), y presenta limitaciones en lo referido a su fragilidad (como en toda red neuronal es crítico el ajuste de los hiperparámetros) y opacidad (Campbell, 2018). La falta de transparencia es un inconveniente conocido de este tipo de estructuras multicapa: así como toda decisión de un sistema tradicional à la Deep Blue se puede trazar y explicar sin restricción alguna por haber sido codificados sus criterios de forma expresa, las muy eficaces estrategias de AlphaZero son, en su mayor parte, inasequibles. No podemos estar seguros sobre la forma matemática o conceptual de las pautas en las que basa sus decisiones. Aunque no tan relevante en el entorno de los juegos, esta salvedad resulta fundamental a la hora de valorar la aplicabilidad de los sistemas de aprendizaje profundo en situaciones realistas de ámbitos como el médico, industrial o financiero, por ejemplo, en las que tan importante es tomar las decisiones adecuadas como poder explicar las razones por las que se toman estas frente a otras alternativas.

4. Percepción y toma de decisiones en jugadores humanos y artificiales

Los ajedrecistas expertos explotan estructuras de conocimiento desarrolladas durante largos periodos de práctica específica que después son capaces de identificar y recuperar ante nuevas situaciones para acelerar la toma de decisiones (Gobet y Charness, 2006). O, dicho de un modo más expresivo, «el maestro de ajedrez “ve” el movimiento correcto» (Chase y Simon, 1973: 56). Los procesos cognitivos subyacentes, que consisten en el reconocimiento automático y reflejo de patrones o estructuras, se suelen asociar a la intuición y existen diversas hipótesis sobre la naturaleza de los resortes psicosomáticos implicados (Jung et al., 2013). Acaso la más conocida sea la que sugiere que la anticipación de las posibles consecuencias de una elección raciona genera respuestas de origen emocional que guían el proceso de decisión simplificándolo, acelerándolo y atenuando los posibles conflictos entre alternativas (Damasio, 2005). Esta respuesta asociada a un conjunto de estímulos es la reacción subjetiva y somática del individuo ante las consecuencias que hayan podido tener decisiones previas en un contexto comparable.

Los mecanismos cognitivos de la mente ajedrecista experta implicados en la preselección de elementos estarían entonces mediados por circuitería neuronal de marcación somática o similar (Reimann y Bechara, 2010). La decisión intuitiva asociada sería una respuesta emocional que, atendiendo al modo en que los mejores jugadores verbalizan su propia experiencia, puede ser de carácter estético. La belleza percibida constituiría así un marcador útil para reconocer movimientos y combinaciones prometedoras en un sentido...

2 Los «materiales suplementarios» de (Silver et al., 2018) incluyen detalles sobre la arquitectura y el funcionamiento de AlphaZero, así como la transcripción de veinte de las partidas jugadas contra Stockfish, el programa más competitivo hasta el momento.
El factor estético en la automatización de tareas lógicas: el caso del ajedrez

La irrupción de AlphaZero ha alterado el panorama ajedrecista por su nivel competitivo y por tratarse de un sistema basado en mecanismos de propósito general que aprende, tabula rasa y a partir de su propia experiencia, criterios y estrategias que también han sorprendido por su atrevimiento y originalidad. Ajenos a cualquier codificación manual de criterios estéticos, su estilo de juego no solo es extremadamente práctico, sino también más creativo que el de los programas convencionales, desplegando movimientos que, además de reflejar un propósito bien definido, resultan atractivos para el espectador. El detalle, imprescindible hasta ahora en el ajedrez por ordenador, nos ha llevado a analizar la relación de AlphaZero con los criterios de belleza en el juego y a reflexionar sobre la naturaleza del juicio estético en la toma de decisiones de los ajedrecistas expertos.

3 En la historia de los juegos por ordenador se ha recurrido a ambos planteamientos. Mientras el primer programa que obtuvo el título de campeón mundial de damas frente a jugadores humanos era, fundamentalmente, un motor de búsqueda especializado (Schaeffer et al., 1996), el primero que derrotó al campeón mundial de backgammon primaba el reconocimiento de patrones frente a los procesos de exploración (Blinier, 1980).
Debe tenerse en cuenta que el nuevo sistema no es fruto de la investigación en creatividad computacional ni busca replicar algún tipo de inteligencia artística. A pesar de ello, del análisis realizado se pueden extraer varias conclusiones relacionadas con la percepción estética y su automatización:

Para que un sistema artificial genere un resultado bello, o tome decisiones que las personas aprecien por su atractivo, no es indispensable la programación explícita de criterios estéticos.

- Los ajedrecistas humanos consideran natural la dicotomía entre el tipo de decisiones estrictamente lógicas que derivan en un estilo de juego racional y previsible, y otras guiadas por un impulso estético que posibilitan un planteamiento más creativo y atractivo para el espectador. El carácter arbitrario de esta segregación queda de manifiesto, sin embargo, cuando quien despliega un juego no solo inesperadamente práctico, sino también original y sugestivo es un sistema artificial que aprende por sí solo las mejores estrategias competitivas, de modo que todas sus decisiones obedecen en realidad a un mismo e interesado criterio de utilidad.

- En la toma de decisiones por parte de expertos humanos el juicio estético, entendido como el reconocimiento de patrones atractivos, puede operar como recurso «intuitivo» para acotar la búsqueda de soluciones eficaces a problemas racionales. Es precisamente un mecanismo de identificación y recuperación de patrones el que explota principalmente AlphaZero, frente al énfasis en la fase de exploración del espacio de soluciones de los sistemas tradicionales.

- El gusto entrenado, que se puede manifestar como sensibilidad estética en una actividad como el ajedrez, podría guardar relación –centrándose en atributos diferentes– con lo que manifiestan expertos de otras disciplinas propias del pensamiento lógico y más complejas. Esta conexión podría contribuir, en concreto, a una mejor comprensión del papel desempeñado por la percepción estética en el razonamiento matemático y las ciencias teórico-empíricas.

**Referencias bibliográficas**

Berliner, Hans J. 1980. «Backgammon Computer Program Beats World Champion». *Artificial Intelligence* 14(1): 205-220.

Bronstein, David I. y Smolyan, George L. 1983. *Chess in the Eighties*. Londres: Macmillan.

Campbell, Murray. 2018. «Mastering board games». *Science* 362(6419) (7 December): 1118. http://dx.doi.org/10.1126/science.aan4880.

Chandrasekhar, Subrahmanyan. 1987. *Truth and Beauty: Aesthetics and Motivations in Science*. Chicago: University of Chicago Press.

Chase, William G. y Herbert A. Simon. 1973. «Perception in Chess». *Cognitive Psychology* 4(1): 55-81. https://doi.org/10.1016/0010-0285(73)90004-2.

Damasio, Antonio. 2005. *Descartes’ Error. Emotion, Reason, and the Human Brain*. Nueva York: Penguin.

Damsky, Iakov. 2002. *Chess Brilliance: 250 Historic Games from the Masters*. Londres: Everyman Publishers.

Ensmenger, Nathan. 2012. «Is chess the drosophila of artificial intelligence? A social history of an algorithm». *Social Studies of Science* 42(1): 5-30. https://doi.org/10.1177/0306312711424596.

Fainshtein, Fridel y Yakov Cohen-Kerner. 2006. «A Chess Composer of Two-Move Mate Problems». *ICGA Journal* 29(1): 32-39.

Fine, Reuben. 1978. «Comment on the Paper, “Principles of Beauty”, by Stuart Margulies». *Psychological Reports* 43(1): 62.

Fox, Mike y Richard James. 1987. *The Complete Chess Addict*. Londres: Faber and Faber.

Gobet, Fernand y Neil Charness. 2006. «Expertise in Chess». En *The Cambridge Handbook of Expertise and Expert Performance*, editado por K. Anders Ericsson, Neil Charness, Paul J. Feltovich y Robert R. Hoffman. 523-538. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511816796.030.

Gobet, Fernand y Herbert A. Simon. 1998. «Expert Chess Memory: Revisiting the Chunking Hypothesis». *Memory* 6(3): 225-255. https://doi.org/10.1080/741942359.

Holding, Dennis H., 1992. «Theories of Chess Skill». *Psychological Research* 54(1): 10-16. https://doi.org/10.1007/BF01359218.

Hsu, Feng-Hsiung. 2004. *Behind Deep Blue: Building the Computer that Defeated the World Chess Champion*. Princeton, NJ: Princeton University Press.

Humble, P.N. 1993. «Chess as an Art Form». *British Journal of Aesthetics* 33(1): 59-66. https://doi.org/10.1093/bjaesthetics/33.1.59.

Iqbal, Azlan y Masahiko Yaacob. 2008. «Theme Detection and Evaluation in Chess». *ICGA Journal* 32(2): 97-109. https://doi.org/10.3233/ICG-2008-31204.

Jung, Wi Hoon, Sung Nyun Kim, Tae Young Lee, Joon Hwuan Jang, Chi-Hoon Choi, Do-Hyung Kang y Jun Soo Kwon. 2013. «Exploring the brains of Baduk (Go) experts: gray matter morphometry, resting state functional connectivity, and graph theoretical analysis». *Frontiers in Human Neuroscience* 7: 19-35. https://doi.org/10.3389/fnhum.2013.00633.

Kasparov, Garry. 1986. *Kasparov Teaches Chess*. Londres: B.T. Batsford.

Kasparov, Garry. 2018. «Chess, a Drosophila of reasoning». *Science* 362(6419) (7 December): 1087. https://doi.org/10.1126/science.aaw2221.

Klein, Gary A. 1993. «A recognition–primed decision (RPD) model of rapid decision making». En *Decision Making in Action: Models and Methods*, editado por Gary A. Klein, Judith Orasanu, Roberta Calderwood y Caroline E. Zsambok. 138-147. Norwood, NJ: Ablex.
El factor estético en la automatización de tareas lógicas: el caso del ajedrez

Lasker, Emanuel. 1960. *Lasker’s Manual of Chess*. Nueva York: Dover.
Le Lionnais, François. 2002. *Les Prix de Beauté aux Échecs*. 3.ª ed. París: Payot.
Levitt, Jonathan y David Friedgood. 2008. *Secrets of Spectacular Chess*. 2.ª ed. Londres: Everyman Chess.
Margulies, Stuart. 1977. «Principles of Beauty». *Psychological Reports* 41(1): 3-11. https://doi.org/10.2466/pr0.1977.41.1.3.
McAllister, James W. 1996. *Beauty and Revolution in Science*. Ithaca, NY: Cornell University Press.
McCarthy, John. 1997. «Al as Sport». *Science* 276(5318) (6 June): 1518-1519. https://doi.org/10.1126/science.276.5318.1518.
Newell, Allen y Herbert A. Simon. 1987. «Computer Science as Empirical Enquiry: Symbols and Search». En *ACM Turing Award Lectures. The First Twenty Years: 1966-1985*, editado por Robert L. Ashenhurst y Susan Graham, 287-317. Nueva York: ACM Press.
Osborne, Harold. 1964. «Notes on the Aesthetics of Chess and the Concept of Intellectual Beauty». *British Journal of Aesthetics* 4(2): 160-163.
Poincaré, Henri. 1910. *La Ciencia y el Método*. Madrid: Librería Gu-

Ravilious, C. P. 1994. «The Aesthetics of Chess and the Chess Problem». *British Journal of Aesthetics* 34(3): 285-290. https://doi.org/10.1093/bjaesthetics/34.3.285.
Reimann, Martin y Antoin Bechara, A. 2010. «The somatic marker framework as a neurological theory of decision-making: Review, conceptual comparisons, and future neuroeconomics research». *Journal of Economic Psychology* 31(5): 767-776. https://doi.org/10.1016/j.joep.2010.03.002.
Root-Bernstein, Robert. 1997. «The sciences and arts share a common creative aesthetic». En *The Elusive Synthesis: Aesthetics and Science*, editado por Alfred I. Tauber, 49-82. Dordrecht: Kluwer Academic.

Saidy, Anthony. 2008. *La batalla de las ideas en ajedrez*. Pamplona: Ediciones MA40.
Samuel, Arthur L. 1995. «Some Studies in Machine Learning Using the Game of Checkers». En *Computers and Thought*, editado por Edgar A. Feigenbaum y Julian Feldman, 71-107. Menlo Park, CA: American Association for Artificial Intelligence.
Schaeffer, Jonathan, Robert Lake, Paul Lu y Martin Bryant. 1996. «Chinook: The Man-Machine World Checkers Champion». *AI Magazine* 17(1): 21-29. https://doi.org/10.1609/aimag.v17i1.1208.
Shannon, Claude E. 1950. «Programming a Computer for Playing Chess». *Philosophical Magazine* 41(314): 256-275. https://doi.
Silver, David, Thomas Hubert, Julian Schrittwieser, Ioannis Antonoglou, Matthew Lai, Arthur Guez, Marc Lancot, Laurent Sifre, Dharshan Kumaran, Thore Graepel, et al. 2018. «A general reinforcement learning algorithm that masters chess, shogi, and Go through self-play». *Science* 362(6419) (7 December): 1140-1144. https://doi.org/10.1126/science.aar6404.
Simon, Herbert A. 1981. *The Sciences of the Artificial*. 2.ª ed. Cambridge, MA: MIT Press.
Sutton, Richard y Andrew Barto. 1998. *Reinforcement Learning: An Introduction*. Cambridge, MA: MIT Press.
Tesauro, Gerald. 1995. «Temporal difference learning and TD-Gammon». *Communications of the ACM* 38(3): 58-68. https://
Walls, Ben P. 1997. *Beautiful Mates: Applying Principles of Beauty to Computer Chess Heuristics*. Irvine, CA: Dissertation.Com.
Winter, Edward. 1989. *Capablanca: A Compendium of Games, Notes, Articles, Correspondence, Illustrations and Other Rare Archival Material on the Cuban Chess Genius José Raúl Capablanca, 1888-1942*. Jefferson, NC: McFarland.
El factor estético en la automatización de tareas lógicas: el caso del ajedrez

Santiago Rementeria-Sanz
GAIKER Technology Centre, Basque Research and Technology Alliance (BRTA)
rementeria@gaiker.es

Santiago Rementeria-Sanz es director general del Centro Tecnológico GAIKER. Doctor en Ciencias Físicas por la Universidad del País Vasco y graduado en Historia del Arte por la UNED, fue visiting scientist en la Universidad Carnegie Mellon de Pittsburgh, Estados Unidos, colaborando en el diseño de sistemas de aprendizaje automático. Su tesis doctoral se centró en el desarrollo y la validación de un nuevo tipo de red neuronal y algoritmo de entrenamiento. La carrera profesional de Santiago está ligada a la I+D y la innovación tecnológica en el mundo empresarial, habiendo sido directivo, consejero y consultor de compañías internacionales en los sectores TIC, energía y espacio. Ha colaborado con la Comisión Europea como evaluador de propuestas y revisor de proyectos de investigación, e impartido clases en cursos de máster en varias universidades. Es autor de publicaciones sobre cuestiones de inteligencia artificial, ingeniería del software, redes eléctricas y gestión empresarial. Gradualmente su atención se ha ido orientando, sin embargo, hacia temas histórico-artísticos, y sus intereses actuales se centran en la interpretación del juicio estético en las ciencias analíticas y en la relación entre el arte, la tecnología y el poder en sus distintas manifestaciones.
ARTÍCULO

El patrimonio cultural y artístico, los vínculos y las comunidades patrimoniales en red. Indicadores de alto impacto a través de las TIC

Sofía Marín-Cepeda
Universidad de Valladolid

Fecha de presentación: diciembre de 2019
Fecha de aceptación: marzo de 2020
Fecha de publicación: julio de 2020

Cita recomendada

Marín-Cepeda, Sofía. 2020. «El patrimonio cultural y artístico, los vínculos y las comunidades patrimoniales en red. Indicadores de alto impacto a través de las TIC». Artnodes, Nº. 26: 1-9. UOC. [Consulta: dd/mm/yy]. http://doi.org/10.7238/a.v0i26.3329

Resumen

El presente artículo trata de analizar la influencia de las TIC en la conformación de comunidades en torno al patrimonio artístico y cultural. En un camino de profundización en la comprensión del patrimonio desde la visión humanista y relacional, durante los últimos años, nuestra línea de investigación se ha orientado al estudio de los vínculos, sus características, arquitectura y particularidades. Dado un paso más en el estudio del vínculo, proponemos un concepto de patrimonio muy ligado al de comunidad. En este sentido, el presente estudio se dirige a localizar, conocer y analizar la difusión online de la producción científica a través del análisis de indicadores bibliométricos en red en la última década. Para ello, llevamos a cabo un análisis de indicadores de alto impacto en torno al patrimonio desde la perspectiva de los vínculos y su conexión con el concepto de comunidad, efectuando búsquedas microespecializadas en
bases de datos especializadas en internet: *Web of Science* y *Scopus*. Aplicamos una metodología de carácter mixto, basada en el análisis de contenido (NudistVivo) y el análisis de carácter estadístico-descriptivo (Excel), a través de dos gestores de referencias bibliográficas online (RefWorks y EndNote). Los resultados nos permiten conocer el impacto de la tecnología en la difusión, conexión y evolución científica en nuestra línea de estudio, así como las líneas de investigación desarrolladas y las posibilidades de continuidad que se nos abren en el futuro.

**Palabras clave**
Vínculos, red, TIC, patrimonio, comunidad, indicadores bibliométricos.

*Culture and artistic heritage, links and online communities. High impact indicators through ICT*

**Abstract**
This article tries to analyse the influence of ICT in the formation of communities around artistic and cultural heritage. In a way of deepening the understanding of heritage from a humanistic and relational perspective, in recent years, our line of research has focused on the study of links, their characteristics, architecture and particularities. Taking another step in the study of the link, we propose a concept of heritage closely linked to the concept of community. The present study is aimed at locating, knowing and analysing the online dissemination of scientific production through the analysis of network bibliometric indicators in the last decade. To do this, we carry out an analysis of high impact indicators relating to heritage from the perspective of links and their connection to the concept of community, conducting micro-specialised searches in specialised databases on the Internet: *Web of Science* and *Scopus*. We apply a mixed methodology, based on content analysis (NudistVivo) and statistical-descriptive analysis (Excel), through two online bibliographic reference managers (RefWorks and EndNote). The results allow us to know the impact of technology in dissemination, connection and scientific evolution in our line of study, as well as the lines of research developed and the possibilities of continuity that are opened to us in the future.

**Keywords**
líneas, red, ICT, patrimonio, comunidad, bibliométricos.

1. **Introducción. Vínculos, patrimonio artístico y cultural y comunidades virtuales**

Podemos comprender el vínculo como unión, relación de carácter no material, aquello que ata, enlaza. Estos vínculos se construyen con bienes materiales e inmateriales, lugares y personas a lo largo de nuestra vida, a partir de conexiones emocionales que alimentan nuestra existencia y nuestros procesos de construcción identitaria. En este sentido, los vínculos que las personas establecen con los bienes han sido objeto de investigación en los últimos años, presentados en eventos científicos de carácter nacional e internacional y en publicaciones de impacto (Marín-Cepeda, 2018; Marín-Cepeda, 2017; Gómez, Fontal e Ibáñez-Etxeberria, 2016; Marín-Cepeda, 2015; Fontal, García e Ibáñez, 2015; Fontal, 2013; Criado-Boado y Barreiro, 2013; Gómez, 2013).

En nuestro estudio seguimos los planteamientos de Marín-Cepeda (2018), para partir de la visión relacional del patrimonio propuesta por Fontal (2013) como perspectiva que acentúa las relaciones entre bienes y personas, en términos de conformación de identidad (comprendida como el conjunto de rasgos distintivos de una persona que permite definirla y diferenciárla de otras), propiedad, pertenencia y emoción: vínculos, en definitiva. Exploramos el vínculo desde esta óptica como clave en los procesos de patrimonialización. Comprendemos la identidad patrimonial como la apropiación y defensa del derecho a una cultura propia, la identificación con las expresiones culturales. Podríamos definir, en este sentido, el vínculo como detonante y resultado de los procesos de construcción de patrimonios y como patrimonio en sí mismo, pues es en el vínculo generado donde reside el valor del bien, en su vertiente material e inmaterial, y donde reside el eje para la construcción de identidades.

Abordamos un concepto de comunidad desde esta perspectiva, asumiendo la necesaria existencia de vínculos, lazos identitarios, emocionales con la cultura, que construyen el entramado de esta. Siguiendo con este concepto, Portolés (2017) analiza el concepto de *comunidad patrimonial*, definiéndola como una construcción de...
El patrimonio cultural y artístico, los vínculos y las comunidades patrimoniales en red...

El concepto de comunidad patrimonial, siguiendo a Pérez-Prat (2015), es relativamente reciente. Fue concretado en la Convención de la Unesco de 2003, para la salvaguardia del patrimonio cultural inmaterial. Siguiendo con el autor, el concepto de comunidad patrimonial es, además, laxo, englobando situaciones muy diferentes. Este concepto ha tenido un mayor recorrido en otros ámbitos vinculados al patrimonio o a los derechos culturales, con la Convención-marco del Consejo de Europa sobre el valor del patrimonio cultural para la sociedad (Convención de Faro, 2005), que introduce por vez primera la noción de comunidad patrimonial, definiéndola como «personas que valoran aspectos específicos del patrimonio cultural que desean, en el marco de la acción pública, mantener y transmitir a las generaciones futuras» (Pérez-Prat, 2015, p. 149). Siguiendo con el autor, el propio Consejo de Europa, en su informe explicativo de la Convención, pone ejemplos de comunidad patrimonial, que puede surgir del interés compartido por una disciplina, como sería el caso de una comunidad arqueológica. Pero igualmente podría tener un fundamento de carácter geográfico y estar vinculada a una lengua o religión. Una comunidad, de este modo, podría ser por definición de carácter transnacional.

Siguiendo a Kastoriano (2000), el concepto de comunidad transnacional remite a comunidades formadas por personas o grupos que pertenecen a diferentes sociedades nacionales, que actúan a partir de intereses y reflexiones comunes y que se apoyan sobre redes transnacionales para reforzarse más allá de la frontera nacional. En esta categoría podemos hablar, además, de las comunidades virtuales. En la construcción de comunidades virtuales es clave la ruptura de barreras que ofrece la red. En este sentido, cobran relevancia los entornos virtuales, temática recientemente estudiada por diversos autores (Gillate, Ibáñez-Ebeberria y Cuenca, 2018; Ibáñez-Ebeberria, Fontal y Rivero, 2018). En esta línea, las redes de comunicación hacen posible la construcción de patrimonio y comunidad global, pues el patrimonio, más allá de representar lo local, traspasa sus fronteras para formar, a través de la red, un entramado de carácter compartido. En esta tipología de comunidad prevalece lo común frente a lo privado. Se trata de nuevas dinámicas sociales y culturales que trascienden el plano físico.

Siguiendo con esta idea, ¿y si comprendiéramos el patrimonio, en sí mismo, como comunidad? ¿Qué tiene que decir el patrimonio sobre el concepto de comunidad? En esa dialéctica comunidad-patrimonio, comprendida como un camino de ida y vuelta, la comunidad, en definitiva, se define como un grupo de personas con elementos en común (lengua, costumbres, valores, espacio y tiempo, entre otros), personas que comparten una identidad común. Las personas como elementos que conforman una comunidad son, en sí mismas, parte de la componen y la necesidad de implicar de manera activa a todos sus miembros como elemento central, alrededor del cual esta se define y desarrolla.

Para entender a qué nos refirió el autor con el concepto de comunidad global, hay que entender el vínculo como el motor de desarrollo. Las relaciones entre sociología y comunidad son aquí evidentes. De Marinis (2011) afirma que vivimos en un boom de la comunidad de alcance global. Lejos del concepto tradicional de comunidades, afloran muchas otras comunidades que convocan formas de aggregación y acción sobre bases muy diversas. Comunidad sirve para decir quiénes y cómo somos, y quiénes y cómo son ellos.

Volviendo al vínculo, Marinis-Cepeda (2018) define las funciones del establecimiento de vínculos con el patrimonio que, por un lado, cumplen una función adaptativa, contribuyendo en la conformación de identidades. La identidad es un constructo en constante proceso de construcción y reformulación a partir de las vivencias del individuo en contacto con el otro. Ésta supone un cambio en el sujeto que ha de vincularse y construir su propia identidad. Por otro lado, la función social de los vínculos supone el refuerzo del sentido de pertenencia a un determinado grupo social, desde las esferas más cercanas (pareja, amigos, familia), hasta la pertenencia a una determinada cultura. Retomamos y resumimos, a continuación, la clasificación en diez tipologías de vínculos (figura 1), que establece Marinis-Cepeda (2018):

- (VI) Vínculo identitario: se da cuando el individuo incorpora como rasgo o característica fundamental de su identidad la relación establecida con determinado bien material o inmaterial. El vínculo se detecta a partir de descriptores clave: ¿me define?, ¿quién soy?, ¿me veo reflejado/a?, ¿me identifico?.
- (VF) Vínculo familiar: relación con el bien patrimonial basada en el recuerdo o el lazo con personas que forman parte de su familia. Se detecta a partir de descriptores clave: ¿me recuerda a?, ¿me une...
2. El impacto de las TIC en la investigación. Análisis de indicadores de alto impacto de la producción científica en torno a las comunidades patrimoniales

Buscamos localizar, conocer y analizar la actividad de investigación desarrollada en torno a los vínculos, el patrimonio y la comunidad en la red en los últimos diez años (2008-2018), a través del análisis de indicadores de alto impacto. Para ello, acudimos a las bases de datos específicas de la Web of Science (WOS) y Scopus. Los resultados son exportados a dos gestores bibliográficos, EndNote y RefWorks y, posteriormente, analizados utilizando herramientas de análisis cualitativo (NudistVivo) y cuantitativo (Excel). Con ello, seremos capaces de comprender las líneas de investigación desarrolladas, así como conocer la evolución en nuestra línea de interés a través de la difusión en la red.

Para las búsquedas definimos los descriptores temáticos (DT) que guiarán el proceso, agrupados en dos ejes: patrimonio (AT1) y didáctica (AT2). Comenzamos así por la identificación y clasificación que guiarán el proceso, agrupados en dos ejes: patrimonio (AT1) y didáctica (AT2). Comenzamos así por la identificación y clasificación de la diferenciación y relación con otros y su comunidad patrimonial (AT2). Comenzamos así por la identificación y clasificación de las líneas de investigación desarrolladas, así como conocer la evolución en nuestra línea de interés a través de la difusión en la red.

APlicamos los descriptores específicos (16 descriptores) y la franja temporal descrita en los buscadores seleccionados. Sobre los resultados, efectuamos análisis de corte estadístico-descriptivo (Fontan e Ibáñez-Etxeberria, 2017; Lucio, Sampieri y Collado, 2010; Etxeberria y Tejedor, 2005), y realizamos análisis de contenido con NudistVivo a través de cada una de las áreas temáticas descritas.

El proceso para la recogida y el análisis de datos se define en tres momentos: el estudio individual de las áreas temáticas (1), la búsqueda de relaciones a través de análisis comparativos (2), cuyos resultados cruzados enriquecen estos, y el análisis de conjunto (3).

3. Resultados: volumen de publicaciones, áreas y temáticas detectadas

Introducimos los descriptores definidos en las bases de datos internacionales WOS y Scopus, localizando la producción científica internacional en torno a nuestro tema de estudio en la franja 2008-
2018. La búsqueda en Web of Science a partir de los descriptores (combinaciones descritas en la tabla 1; artistic education, links, teaching, relations, community, education), restringida a los últimos diez años, nos ofrece un total de $n=76$ resultados, localizados en Core Collection (65), Current Contents Conect (41), Medline (13), KCI-Korean Journal Database (7), Russian Science Citation Index (2), Scielo Citation Index (2). Ordenados cronológicamente, observamos una evolución que alcanza su pico entre 2014 y 2016 (gráfico 1).

Se trata de artículos (84,21%), contribuciones a congresos (9,21%), revisiones (2,63%), libros (1,31%) y otros (2,63%) (gráfico 2), repartidos en diversos países, entre los cuales encontramos Estados Unidos (13), Inglaterra (8), Australia (6), Italia (6), España (6), México (3), entre otros (gráfico 2).
A su vez, localizamos distintas tipologías de publicaciones: artículos científicos (55,81%), capítulos de libros (6,20%), revisiones (5,42%), libros (2,32%), contribuciones en eventos científicos (2,32%), artículo de prensa (1,55%), editorial (0,77%) y otros (25,61%) (gráfico 3).

De entre los resultados detectados, por su cercanía con nuestra temática de estudio queremos destacar los trabajos de Limón y García (2016), «Heritage, place and neighborhood: itineraries as public space contention in ringroad districts of Madrid»; «Links: going beyond cultural property», Carman, 2018; «Talkinf of heritage: the past in conversation», Robinson, 2018; «Wiews and Perceptions of Journalistas about the Archaeological Heritage. The case of Olavarria» Conforti, Giacomasso, Mariano y Endere, 2016; «The senses of place. Temporalities, social relations and memories in a segregated neighborhood of la Plata (Argentina)», Segura (2013); entre otros.

En cuanto a la evolución cronológica (gráfico 4), observamos que, si bien la producción se mantenía entre un 5 % y un 9 % entre 2008 y 2010, se produce un aumento significativo en el número de publicaciones científicas de impacto entre 2013 y 2016, descendiendo casi seis puntos en los años posteriores.

La progresión comparada entre ambas bases de datos, organizada con un carácter cronológico, nos permite observar la producción científica localizada de manera global (gráfico 5). Podemos observar cierto paralelismo en ambas bases de datos, cuyos repuntes parecen adelantarse en Scopus, reflejándose posteriormente en los resultados en WOS.

El análisis de contenido de los resultados obtenidos en ambas bases de datos de impacto nos permite afirmar que el vínculo con el patrimonio es acometido en la literatura reciente como tema secundario en algunas ocasiones (43,31%), derivado del estudio de la educación patrimonial. Como ya anticipábamos, este tema ha sido abordado recientemente en congresos nacionales e internacionales por diversos autores en nuestro país (Marín-Cepeda, 2018; Marín-Cepeda, 2017; Marín-Cepeda, 2015; Fontal, García e Ibáñez, 2015; Fontal, 2013; Criado-Boado, 2013; Gómez, 2013), siempre en relación con los procesos de patrimonialización en el contexto de la educación patrimonial.

4. Conclusiones y líneas abiertas

El incremento en la investigación en torno al patrimonio, la comunidad y el vínculo y su difusión a través de las TIC parece claro a la luz de los resultados, principalmente en la franja 2013-2016. En la evolución temática detectada en el análisis de redes podemos deducir una tendencia hacia las investigaciones en torno al patrimonio y la comunidad, en términos de conformación de identidades. Se trata de una línea temática concreta y especializada, donde es posible observar un aumento en el volumen de producción reciente. Si atendemos a la evolución temática, podemos afirmar la predominancia en los estudios de la comunidad desde las perspectivas sociológica y antropológica, así como una ligera tendencia hacia análisis dirigidos.
al estudio de las comunidades desde diversas ópticas, primando la puesta en valor del patrimonio desde una perspectiva histórica, identitaria y económica. En algunas de las investigaciones localizadas, destaca, además, el planteamiento de las personas como núcleo. Esta progresión se observa tanto en los artículos de impacto como en los libros y capítulos de libro.

Teniendo en cuenta el contenido de las publicaciones localizadas, así como los conceptos que estas manejan y sus enfoques metodológicos, podemos establecer como líneas de trabajo el estudio de las comunidades en términos de herencia patrimonial, comunidades desde el turismo y la rentabilidad económica del patrimonio, comunidades desde la sociología, desde la antropología, desde la educación, desde el punto de vista de la historia, enfoques de salvaguarda y preservación del patrimonio de las comunidades minoritarias, la contribución del patrimonio en la conformación de comunidades, entre otras. Se trata de líneas de trabajo que inferimos de la muestra localizada; no obstante, sería de interés proyectar nuestro trabajo en el futuro hacia el estudio y análisis de las interacciones entre los miembros de las comunidades patrimoniales y los estudios desarrollados al respecto.

Si tenemos en cuenta las temáticas localizadas, podemos decir que las líneas de continuidad son múltiples y variadas. No existen estudios que centren su atención en el vínculo como conformador de comunidades patrimoniales, ni que profundicen en el patrimonio como comunidad en sí mismo, exceptuando el interesante enfoque de comunidades patrimoniales desarrollado por Portolés (2017). De comunidades patrimoniales habla también Pérez-Prat (2015), quien efectúa una amplia revisión de los orígenes del concepto desde 2003, abordando el concepto de identidad patrimonial (en referencia al patrimonio con el que los individuos se identifican) y la herencia del patrimonio. Tampoco encontramos, en las bases de datos analizadas, estudios que evalúen el impacto de las TIC en la conformación de comunidades virtuales, en la difusión de patrimonios comunitarios, ni estudios que analicen el impacto de las redes en el desarrollo científico en nuestra línea de indagación. Somos conocedores de estudios que sí evalúan el impacto de las TIC, pero no han sido difundidos en las bases de datos que hemos analizado, lo que nos indica una vía futura de proyección de nuestro estudio hacia otras bases referenciales.

Por ello, centrar la mirada en las personas como generadoras de patrimonio, el valor patrimonial en la potencialidad simbólica de los bienes, la influencia de las tecnologías de la información y la comunicación en el impacto científico y el vínculo como eje principal de desarrollo, todo ello como clave en la construcción de un patrimonio que se erige como esencia de la comunidad, supone centramos en una vía aún por explorar. Hablamos, principalmente, de líneas emergentes que derivan del estudio del patrimonio y de un concepto de comunidad más actualizado, amplio, líquido y flexible.

Referencias bibliográficas

Ansio, Heli, Piia Seppälä y Pia Houni. 2017. «Teacher’s experiences and perceptions of a community music project: impacts on community and new ways of working». International Journal of Education and the Arts 18(37): 1-29. https://bit.ly/2q73jN8.

Barchana-Lorand, Dorit y Efrat Galnoor. 2009. «Philosophy of art education in the visual culture: aesthetics for art teachers». Journal of Philosophy of Education 43(1): 133-148. http://dx.doi.org/10.1111/j.1467-9752.2009.00661.x.

Bordolli, Marita Fornaro, Laura Monestier, Cecilia Mauttoni, y Sergio Marcelo De los Santos. 2015. «Articulation of levels in artistic education: an experience of educational research in a Latin American context». En Edulearn16: 8TH International Conference on Education and New Learning Technologies, 6395-6403.

Carman, John. 2018. «Links: Going Beyond Cultural Property». Archeologies 14(1): 164-183. https://link.springer.com/article/10.1007/s11759-018-9337-y y https://doi.org/10.1007/s11759-018-9337-y. Arriolabengoa, Miren, Ana Urrutia, Miriam Peña, Alazne Porcel y Aintzane Cámara. (ed.) 2016. «Artistic intervention of public spaces in the school of education in Bilbao». En Edulearn16: 8TH International Conference on Education and New Learning Technologies, editado por Gómez Chova, Luis; A. López Martínez e I. Candel Torres, 8677-8681. https://bit.ly/2XlusYD. http://dx.doi.org/10.21125/edulearn.2016.0887.

Conforti, María Eugenia, María Vanesa Giacomasso, Mercedes Mariano y María Luz Endere. 2016. «Wiews and perceptions of journalists about the archaeological heritage. The case of Olavarria». Histoire et civilisation 15(8), 309-334. https://cutt.ly/veFjrB4 http://dx.doi.org/10.15446/histoire.v8n15.51514.

Consejo de Europa. 2005. Convención sobre la protección y la promoción de la diversidad de las expresiones culturales. París: Unesco. 20 de octubre de 2005. https://bit.ly/33R2w1p.

Corazza, Sandra Mara, Carla González Rodríguez, Ester María Deher Heuser y Silas Borges Monteiro. 2014. «Escriurelittres: un modo de jerescrer en meió a vida». Educaçao e pesquisa 40(4): 1029-1043. https://bit.ly/3595Wp.

Criado-Boado, Felipe y David Barreiro. 2013. «El patrimonio era otra cosa». Estudios Atacameños 45. https://bit.ly/33Le7Gh.

De Marinis, Pablo 2011. «La teoría sociológica y la comunidad. Clásicos y contemporáneos tras las huellas de la buena sociedad». Entramados y Perspectivas 1(1): 127-164. https://bit.ly/2KqK6Ct.

Etxeberria, Juan y Javier Tejedor. 2005. Análisis descriptivo de datos en educación. Madrid: La Muralla.

Fontal, Olala. 2013. La educación patrimonial, del patrimonio a las personas. Gijón: Treo.

Fontal, Olala y Alex Ibáñez-Etxeberria. 2017. «Educación y patrimonio: innovación y reflexión». Pulso, Revista de Educación 40: 9-12. https://bit.ly/3599eH.
Fontal, Olaia y Marín-Cepeda, S. (2018). «Nudos Patrimoniales. Análisis de los vínculos de las personas con el patrimonio personal». Arte, Individual y Sociedad 30(3): 483-500. https://bit.ly/3207k6l.

Fontal, Olaia, Silvia García y Álex Ibáñez-Etxeberria. (Coords.). 2015. Educación y Patrimonio visiones caleidoscópicas. Gijón: Trea.

Fontal, Olaia, Marín, Sofía y García, Silvia. (2011). «Procesos de patrimonialización e identización patrimonial con uso de TIC en torno al arte contemporáneo». Artnodes, Revista de Arte, Ciencia y Tecnología 17. https://core.ac.uk/download/pdf/83003902.pdf.

Haroutounian, Joanne. 2017. «Artistic ways of knowing in gifted education: encouraging every student to think like an artist». Roeper review a journal on gifted education: encouraging every student to think like an artist. https://doi.org/10.1080/02783193.2016.1247397.

Herrera, Manuel. 2003. «Relación social y teoría de las formas en Michel Maffesoli». Reis 101(3): 89-117. https://bit.ly/358cuf8.

Huerta, Ricardo y Ricardo Domínguez. 2016. «The demands of the Arts Education Teachers in front of the new curriculum obstacles: the Project Second Round: Art ILuíita a Secundaria». EARI- Educación Artística Revista de Investigación 7: 10-18. https://bit.ly/32L3kmV.

Ibáñez-Etxeberria, Alex, Olaia Fontal y Pilar Rivero. 2018. «Educación Patrimonial y TIC en España: marco normativo, variables estructurantes y programas referentes». Arbor 194(788): 1-17. https://bit.ly/2KpU41d http://dx.doi.org/10.3989/arbor.2018.788n2008.

Kastoriano, Riva. 2000. «Inmigration, Transnational Communities and Citizenships». Jurnal Revue Internationales des Sciences Sociales 165: 353-359. https://bit.ly/2CKJ4Nt https://doi.org/10.1111/1468-2451.00261.

Limón, Pedro y Sergio Claudio García. 2016. «Heritage, place and neighborhood: itineraries as public space contention in ringroad districts of Madrid». Research in Urban Sociology 15: 255-285. http://dx.doi.org/10.1108/S1047-004220160000015011.

Lucio, Pilar Baptista, Roberto Hernández Sampieri y Carlos Fernández Collado. 2010. Metodología de la investigación. México: McGraw-Hill.

Maffesoli, Michel. 1990. El tiempo de las tribus. Barcelona: Ikaria.

Marín, Sofía (2017). «Patrimonio en conserva. Los vínculos identitarios como clave de aprendizaje en el grado de magisterio». Mídias, Museos e estudios interdisciplinares 8: 1-16. https://bit.ly/32R3Wra.

Pérez-Prat, Luis. 2015. «La protección internacional del patrimonio y las comunidades patrimoniales». AFDUAM 19: 121-151. https://bit.ly/2xMrSWp.

Portoles, Ángel. 2017. «¿Qué es para ti una comunidad patrimonial?». Revista Memoria Viva 9: 51-60.

Robinson, Mike. 2018. «Talking of heritage: the past in conversarion». En Routledge Handbook of Language and Superdiversity: a interdisciplinary perspective, coordinado por Angela Creese y Adrian Blackledge, 194-207. London: Routledge. https://www.taylorfrancis.com/books/e/97813156969010 https://doi.org/10.4324/97813156969010.

Segura, Ramiro. 2013. «The senses of place. Temporalities, social relations and memories in a segregated neighborhood of la Plata (Argentina)». Sociedade e Cultura 16(1): 59-68. https://bit.ly/2OfYyJ6.

Marfili-Carmona, Rafael y Pedro Chacón. 2017. «Arts Education and Media Literacy in the Primary Education Teaching Degree of the University of Granada». Education, Health and ICT for transcultural World 237: 1166-1172. https://bit.ly/2pkNvp http://dx.doi.org/10.1016/j.sbspro.2017.02.174.
El patrimonio cultural y artístico, los vínculos y las comunidades patrimoniales en red...

CV

Sofía Marín-Cepeda
Universidad de Valladolid
sofiavictoria.marin@uva.es

Profesora en el Departamento de Didáctica de la Expresión Musical, Plástica y Corporal, Área de Didáctica de la Expresión Plástica, en la Facultad de Educación y Trabajo Social de la Universidad de Valladolid (España). Doctora en Investigación en Educación Artística por la Universidad de Valladolid (Premio Extraordinario de Doctorado). Graduada en Bellas Artes por la Universidad de Salamanca. Licenciada en Psicopedagogía y Diplomada en Magisterio de Educación Primaria (Premio Extraordinario Fin de Carrera) por la Universidad de Valladolid.

ORCID: 0000-0003-0383-0408
Researcher ID: D-4859-2011
Scopus Author ID: 57194169833
Academia:http://feyts.academia.edu/SofíaMarínCepeda
Google Scholar: https://scholar.google.com/citations?hl=en&user=mWLSY2YAAAAJ
ARTÍCULO

La dimensión ambiental de las imágenes. La imagen-ambiente y sus principales conceptos

Mikel Otxoteko
Investigador independiente

Fecha de presentación: mayo de 2020
Fecha de aceptación: junio de 2020
Fecha de publicación: julio de 2020

Cita recomendada

Otxoteko, Mikel. 2020. «La dimensión ambiental de las imágenes. La imagen-ambiente y sus principales conceptos». Artnodes, Nº. 26: 1-11. UOC. [Consulta: dd/mm/aa].
http://doi.org/10.7238/a.v0i26.3380

Resumen

A partir de los años sesenta surgen nuevas maneras de mostrar y percibir las imágenes, en nuevos contextos o ambientes, generalmente de carácter lúdico y nocturno. Se producen una serie de transformaciones ambientales en torno a la imagen que hacen que su producción y visionado público se vuelvan procesos difícilmente predecibles e interactivos. Se ha llamado imagen-ambiente a una lógica de la imagen que encuentra sus orígenes en eventos iniciáticos como los Exploding plastic inevitable producidos por Warhol y que hoy en día, en parte gracias al impulso de las tecnologías digitales, se popularizan y encuentran nuevos desarrollos en las sesiones de DJ-Vjing. Dada su importancia social y su radical especificidad, proponemos una serie de conceptos que ayudarán a comprender qué territorio ocupa este tipo de imágenes y, sobre todo, qué relación establece con su audiencia. Algunos de los principales conceptos son: afecto ambiental, estímulos en reciprocidad y entrecuerpo. Partiendo del planteamiento deleuziano en la imagen-movimiento y la imagen-tiempo, así como tomando la teoría de los afectos de Spinoza, damos un paso hacia delante para continuar desarrollando el universo conceptual de esta particular lógica de la imagen: la imagen-ambiente.
La dimensión ambiental de las imágenes. La imagen ambiente y sus principales conceptos

Palabras clave
imagen, ambiente, VJing, afecto, acontecimiento, audiovisual.

The environmental dimension of images. The environment-image and its principal concepts

Abstract
From the 1960s onwards, new ways of displaying and perceiving images arose in new contexts or environments, generally of a playful and nocturnal nature. A series of environmental transformations were generated around images which meant that their production and public viewing became processes that were difficult to predict and interactive. Environment-image is the name that has been given to an image logic which finds its origins in ground-breaking events such as the Exploding Plastic Inevitable, organised by Warhol and which is now, thanks in part to the impetus of the digital technologies, becoming popular and flourishing in DJ-VJing sessions. Given its social importance and its ‘radical’ specificity, we propose a series of concepts which will help us to understand which territory this type of images occupies and above all which relationship it establishes with its audience. Some of the principal concepts are: ‘environmental affect’, ‘reciprocal stimuli’ and ‘interbody’. Starting from the Deleuzian approach in movement-image and time-image, and taking Spinoza’s affect theory, we take a step forwards in order to continue developing the conceptual universe of this particular logic of the image: the environment-image.

Keywords
image, environment, VJing, affect, event, audiovisual

Introducción: un nuevo universo conceptual para un conocido territorio de la imagen
Los diferentes usos y efectos que tienen las imágenes son innumerables y aumentan a medida que la actividad humana conquista nuevos territorios de la experiencia. Ello es, como han sugerido algunos, fruto de sinergias que se establecen entre los artistas, los usuarios y quienes desarrollan sistemas tecnológicos. Pero, ante la evidencia de algunas experiencias como a las que vamos a referirnos, también habría que añadir en ese entramado la creatividad propia de los ambientes sociales en los que las imágenes encuentran su sentido.

Recientemente, en la publicación del artículo «Vjing. Estética y política de la imagen-ambiente», nos proponíamos responder a dos cuestiones fundamentales en este sentido: ¿puede hablarse de un tipo de imagen cuya especificidad radica en la capacidad de fundirse y confundirse con el ambiente, y de, por tanto, variar en el tiempo en función de las transformaciones que sufre dicho ambiente? En caso afirmativo, ello justificaría el vínculo entre, por un lado, las sesiones de DJ-Vjing que emergieron en la década de los ochenta y, por otro, algunos de los eventos Exploding plastic inevitable (1966-1967) enmarcados en el contexto de la experimentación artística, lúdica y estética de los años sesenta. Recordemos su idiosincrasia, así como las actuales sesiones de DJ-Vjing como las actuales sesiones de DJ-Vjing u otros ámbitos semejantes. El concepto sugerido responde a la taxonomía propuesta en las últimas décadas del siglo xx por Gilles Deleuze, una «clasificación de los signos» que, en principio, se planteaba como específica para el ámbito de la imagen cinematográfica. Conviene saber que sus dos célebres estudios sobre cine, La imagen-movimiento (1983) y La imagen-tiempo (1985), supusieron una aportación de gran relevancia para el desarrollo de una nueva teoría del cine. Este paradigma conceptual venía desde el campo de la filosofía y era distinto —y también, habría que decir, «distante»— de aquel que alimentaba los análisis formalistas y realistas clásicos, o de la posterior vía estructuralista de Christian Metz, que combinaba la lingüística y el psicoanálisis. Además, el planteamiento dejaba abierta la posibilidad a conceptualizar otras lógicas de la imagen dentro y fuera de la tradición cinematográfica.
Así, este artículo tiene como objetivo crear una base conceptual argumentada que legitime lo que, en respuesta a los interrogantes expuestos anteriormente, hemos llamado la imagen-ambiente. Partimos, para ello, de la convicción de que no se debería hablar de la misma manera para describir, valorar o comentar los distintos tipos de imagen y sus diferentes modos de afectar a una audiencia –o, como puntualiza Rancière (2017), las distintas «lógicas de la imagen»–. Proponemos ir señalando métódicamente los aspectos principales que explican la imagen-ambiente. En primer lugar, para mostrar su carácter singular como afecto ambiental; en segundo lugar, incidiendo en su capacidad para mantener con la audiencia una relación de reciprocidad expresiva; y, por último, aludiendo a distintas dimensiones ambientales de las que la imagen deriva. Aventuramos pues, a modo de tesis, que la imagen-ambiente es un tipo de imagen que deriva de forma espontánea de la creatividad de todo un ambiente.

Dentro del marco teórico al que nos hemos referido, dejaremos a un lado la importante referencia de Bergson, central en los estudios de la imagen de Gilles Deleuze, para volver a Spinoza, a pesar del rechazo de este filósofo hacia la idea de una estética (Morrison, 1989). Fue a Spinoza a quien, por otro lado, Deleuze dedicó su libro *Spinoza: filosofía práctica*, además de varios ensayos que ponen de relieve su teoría de los afectos. De este modo, aunque nos hallemos dentro del campo de las artes visuales, lo anterior nos arrastra a un terreno transdisciplinario que hace gravitar la teoría de la imagen en torno a lo social, siguiendo a teóricos del comportamiento y la comunicación social como DeLanda, Maturana o McNeill. Veremos surgir, así, nuevos e insospechados conceptos asociados a la imagen e implicados en ella: afecto ambiental, estímulos en reciprocidad, entrecuerpo, unos y otros referidos en gran manera al modo en que un ambiente social es afectado y, a su vez, este afecta a la imagen.

1. Afecto ambiental

Hasta hace solo un par de décadas, la imagen-ambiente ha contado con una escasa legitimidad y prestigio social (Sedeño Valdellós, 2005). Han tenido que crearse grandes festivales con performances visuales en tiempo real y eventos VJ para poner oficialmente en valor su potencia expresiva. Al no atraer la mirada de los críticos y teóricos de la imagen, era percibida con relación a formas de arte menor. Además de esto, se la ha visto muchas veces asociada a ambientes lúdicos y nocturnos, clubes de baile y conciertos, eventos sin orden aparente en los procedimientos para con la audiencia, proyectada en locales inundados del hedor corporal, de alcohol, a veces combinada con sustancias psicoactivas; incluso podría decirse que se la ha visto «diluida» en tales ambientes, sin causa ni efecto. Lo que no se ha sabido ver, sin embargo, es que esto son evidencias de su «radical» especificidad. Estas particularidades podrían sintetizarse a partir de tres rasgos principales: la no-linealidad, la exterioridad y el devenir social.

La imagen y la no-linealidad

Tal como se aprecia a través de la documentación visual de Ronald Nameth sobre las sesiones Exploding plastic inevitable producidas por Andy Warhol, el peso creativo de la imagen-ambiente no está puesto sobre el relato, que podría ser, de hecho, inexistente, sino sobre la modulación de la luz, el cromatismo y las formas visuales. Al igual que ocurre en las sesiones de Vjing, la relación entre unas imágenes y otras en la sucesión del flujo visual no tiene linealidad discursiva ni destino al cual llegar, aunque sí haya una cierta progresión estructurada y una idea de base sustanciada en algo (Sedeño Valdellós, 2005).

Pensemos en las implicaciones que ello pueda tener. Cuando hay un relato sostenido por imágenes, en el cine de ficción, por ejemplo, la mirada debe ser conducida no meramente a través de una sucesión de imágenes, sino a partir de las zonas de interés de tales imágenes, con la importancia del encuadre. Tal como ocurría de manera soberbia en las películas de Alfred Hitchcock, la información en el plano estaba dispuesta de forma calculada y estratégica según las exigencias de la narración; algo que Rancière calificaba como despótico (2017). Si, al contrario, no hay una lógica lineal, el flujo de imagen no tiene por qué disponer de centros de interés que conduzcan con precisión la mirada del espectador. Incluso, puestos a valorar la falta de linealidad y el sentido único, las contingencias del ambiente (es decir, aquello que, no estando determinado, tampoco es exactamente azaroso) pasarían a ser estímulos enriquecedores.

La lógica lineal de la imagen-movimiento y la imagen-tiempo queda, como vemos, invalidada. Tanto es así que, en ambientes como las sesiones Exploding plastic inevitable, lo esperado era que el ojo de la audiencia pudiera moverse a su antojo sobre la superficie, la pantalla; que cada persona, mientras conversa, baila o avanza distraídamente por el espacio, se moviera también libremente por la proyección. Las imágenes proyectadas o sus combinaciones no tienen una significación a priori, tal y como se establece en un guion, sino que el sentido va variando a través de cada una de las diferentes fases de producción y visionado por parte de la audiencia. De este modo, no se le da al ojo lo que tiene que ver, sino un ámbito visual a partir del cual producir sentido.

La imagen y su exterioridad

Visto este primer concepto, podrá entenderse que la imagen-ambiente ya no es una forma de expresión acotada y cerrada sobre sí misma, tal como ocurre en el arte cinematográfico y en la mayor parte del videoarte. Como efecto, las pantallas planas heredadas del cine se están dejando atrás en este tipo de eventos artísticos o culturales (Sedeño Valdellós, 2005). En un cine tradicional, el espacio de la pantalla conforma una especie de región visual autónoma, que nada tiene que ver con las sesiones Exploding plastic inevitable. Lo cierto es que estas sesiones suelen incluirse entre las pioneras de lo que se ha llamado cine expandido.
La dimensión ambiental de las imágenes. La imagen ambiente y sus principales conceptos

A lo largo de los años, en este mismo contexto de Warhol, Jonas Mekas ha descubierto la potencia de la luz más allá de las imágenes, sino que la imagen, con base en instrumentos de luz como el estrobo, es capaz de permear en los ambientes sociales como nunca antes lo había hecho; se ha impregnado de ambiente, ha dejado de ser opaca. Con Mekas vemos, también, cómo los centros de interés de la imagen se multiplican, cómo el límite se expande, cómo la imagen se disuelve en el ambiente para concretarse de otro modo en «nosotros mismos».

Este tipo de eventos establecen un sistema lumínico que afecta de una manera muy directa en el ambiente, convierte a los visitantes que transitan entre las dos grandes proyecciones en sujetos-imagen. En el ambiente de penumbra creado por las proyecciones y ante lo sobrecogedor y obsceno de los contenidos, se da cierta empatía en relación unos con otros, se observan las reacciones, los movimientos, las distancias y detenciones bajo la luz cambiante. La imagen se muestra manifiestamente abierta a su exterioridad y, para ello, se niega rotundamente la idea de una neutralidad ambiental en torno a la imagen; lo que la enmarca es justamente una serie de acontecimientos expresivos irrefutables. La imagen, más que condicionar tales devenires, se deduce de ellos en función de ese encuentro con lo heterogéneo, lo contingente y lo circunstancial que implican. Así, lo que la mirada encuentra en la imagen viene dado, enriquecido y violentado, por las diferencias, las contingencias y las circunstancias del ambiente. La evidencia en todo esto es que debe incluirse el ambiente al conceptualizar este tipo de imagen, porque el acto de mirar la imagen-ambiente implica, más que condicionar tales devenires, se deduce de ellos en función de ese encuentro con lo heterogéneo, lo contingente y lo circunstancial que implican. Así, lo que la mirada encuentra en la imagen viene dado, enriquecido y violentado, por las diferencias, las contingencias y las circunstancias del ambiente. La evidencia en todo esto es que debe incluirse el ambiente al conceptualizar este tipo de imagen, porque el acto de mirar la imagen-ambiente implica, más que nunca, el propio ambiente.

La imagen y su deriva social

A lo largo de los años, en este mismo contexto de Warhol, Jonas Mekas fue trazando con sus breves artículos y críticas sobre los eventos de la escena del audiovisual neoyorkino una especie de crónica en la cual se percibe la progresiva aparición de un tipo de imagen que, novedosamente, explora una dimensión social. Aunque no la nombre como tal, se muestra plenamente consciente de sus características e implicaciones específicas:

Empiezan con una simple pantalla e imágenes de una larga toma; luego se superponen imágenes; superposiciones triples; luego dos, tres, ocho pantallas; luego las imágenes se multiplican, las superposiciones se atomizan aún más, se multiplican y violan las pantallas de seda, por velos de colores y por sistemas de sonido. Ahora hemos abandonado la pantalla, la película, y nos dedicamos a nosotros mismos; con el estrobo nos cortamos en fotogramas únicos, como un gesto simbólico, mágico, o como un ritual (Mekas, 1966).

Lo anterior no significa que Mekas haya descubierto la potencia de la luz más allá de las imágenes, sino que la imagen, con base en instrumentos de luz como el estrobo, es capaz de permear en los ambientes sociales como nunca antes lo había hecho; se ha impregnado de ambiente, ha dejado de ser opaca. Con Mekas vemos, también, cómo los centros de interés de la imagen se multiplican, cómo el límite se expande, cómo la imagen se disuelve en el ambiente para concretarse de otro modo en «nosotros mismos».

Este tipo de eventos establecen un sistema lumínico que afecta de una manera muy directa en el ambiente, convierte a los visitantes que transitan entre las dos grandes proyecciones en sujetos-imagen. En el ambiente de penumbra creado por las proyecciones y ante lo sobrecogedor y obsceno de los contenidos, se da cierta empatía en relación unos con otros, se observan las reacciones, los movimientos, las distancias y detenciones bajo la luz cambiante. La imagen se muestra manifiestamente abierta a su exterioridad y, para ello, se niega rotundamente la idea de una neutralidad ambiental en torno a la imagen; lo que la enmarca es justamente una serie de acontecimientos expresivos irrefutables. La imagen, más que condicionar tales devenires, se deduce de ellos en función de ese encuentro con lo heterogéneo, lo contingente y lo circunstancial que implican. Así, lo que la mirada encuentra en la imagen viene dado, enriquecido y violentado, por las diferencias, las contingencias y las circunstancias del ambiente. La evidencia en todo esto es que debe incluirse el ambiente al conceptualizar este tipo de imagen, porque el acto de mirar la imagen-ambiente implica, más que nunca, el propio ambiente.
2. Estímulos en reciprocidad

Si tomamos como válido lo anterior, no deberíamos pasar por alto la siguiente consideración. La imagen-ambiente ha podido prosperar gracias, en parte, al desarrollo tecnológico, particularmente al de las últimas dos décadas. La velocid con la que hoy se puede seleccionar, transformar y mostrar imágenes digitales hace que las fases de producción y difusión (como en una sesión proyectada en directo) puedan solaparse; también habría que tener en cuenta el carácter cada vez más táctil en el uso de los dispositivos digitales (Snibbe, 2001; Taylor, 2009). Ello otorga a la imagen una propiedad decisiva: la permeabilidad a los afectos de la audiencia, la cual se vuelve partícipe de la producción de la propia imagen. Nos preguntamos, pues, por la cuestión de la reciprocidad de los estímulos.

Una audiencia performer

Lo que, como se ha visto, sucedía de forma pionera y algo «rudimentaria» en las sesiones Exploding plastic inevitable podría ser un buen ejemplo de esta performatividad de la audiencia, que implica, en primer lugar, la desconfiguración de los roles tradicionales. Por tanto, la audiencia se sitúa en una nueva posición: entre la contemplación y la acción. Dicho en otras palabras: se rompe la clásica relación que se establece entre la acción del creador y la pasión de la audiencia, la relación creador-observador. En su definición sobre el cine expandido, que engloba justamente eventos de producción y difusión (como en una sesión proyectada en directo), Paul Raban señala este factor: «la audiencia se ve implicada en la producción de sentido» (Raban, 2011: 98-107). Esto es a lo que Raban llama proceso reflexivo, que parece tener mucho que ver con la experimentación, dentro del contexto del arte institucional, llevada a cabo recientemente por Scott Snibbe:

«Deep walls» de Scott Snibbe es un trabajo de acción interactiva bastante conocido, en el cual el movimiento de la audiencia es grabado con cámaras de vídeo. Esta instalación de arte induce la participación de la audiencia sobre la cual se proyectan las sombras de los participantes. Es una instalación de arte de acción basada en la interactividad de la pantalla, a través de la cual una audiencia activa deviene el propio trabajo artístico» (Jung, Lee, Choi y Kim, 2014).

En este mismo movimiento vemos cómo, en segundo lugar, la imagen-ambiente implicaría que estas audiencias activas sean partícipes o cómplices en el acto de producir la imagen (aunque de un modo tal que la mayor parte de las veces será una complicidad inconsciente e involuntaria). El artista Tony Hill habla en su entrevista con Karel Sidney Doing de la «performance de la audiencia», es decir, de la importancia del rol activo de la audiencia que, sin reparar en ello, participa con la expresión de su reacción (Doing, 2017: 154-165). Tenemos, por tanto, dos expresiones que, juntas en un ambiente que comparten, se componen y actúan creativamente en reciprocidad.

La lógica de los afectos

Para desarrollar conceptualmente lo anterior, necesitamos la visión materialista de una filosofía como la que Spinoza demuestra en su Ética. De Spinoza no podemos extraer una estética propiamente dicha (Morrison, 1989), pero sí algo mucho más valioso para la definición conceptual de este tipo de imagen: la imagen-ambiente. Como bien señala Susan Ruddick, el énfasis por el afecto «como una cualidad constitutiva más que derivada» en la experiencia sería el radical potencial del corpus de Spinoza (Ruddick, 2010). En esta filosofía práctica del afecto, la imagen-ambiente formaría parte de la indómina «fábrica del cuerpo», por emplear la expresión del propio Spinoza (Spinoza, 2009: 204). En un mismo sentido se unen las palabras de Deleuze: el filósofo concibe la Ética como una etología, o sea, como una composición de velocidades y de lentitudes, de poderes de afectar y de ser afectado en este plan de inmanencia (… No sabéis por anticipado lo que puede un cuerpo o un alma en tal encuentro, en tal dispositivo, en tal combinación (Deleuze, 1978: 152).

Observamos dentro este marco conceptual cómo quienes producen imágenes –aquí podríamos recurrir particularmente a la figura del VJ (visual jockey) como ejemplo de videoartista interactivo– se hallan en un medio relacional complejo, en el que se cruzan aspectos psicológicos, antropológicos y etológicos, entre otros. Dado tal complejidad, al producir el «flujo de imágenes» que constituyen la videosesión, el artista visual debe guiarla principalmente por la respuesta o el comportamiento de la audiencia. Hablamos de la expresión no verbal, o no necesariamente verbal, de emociones y de información multisensorial. Flora Davis ha estudiado la importancia de la postura y el movimiento del cuerpo en sus pequeños detalles: gestos, expresiones faciales, también la mirada de cada persona, o la sincronía interaccional de la audiencia. El productor de imágenes responde creativamente con la expresión de nuevas imágenes, pero también a través de su propio cuerpo, como en cualquier conversación: «La sincronía interaccional es algo sutil, no una exacta imitación de gestos, sino simplemente un ritmo compartido» (Davis, 2005: 137). Aunque de manera intuitiva, el productor de imágenes atiende a los patrones surgidos de los ritmos corporales, de los movimientos de las manos, los cabeceos… todos los movimientos del cuerpo; y los sigue atentamente, los estimula o los amaina con su rítmica visual. Por ejemplo, observa cómo el ritmo se altera cuando introduce un determinado tipo de imagen; del mismo modo, se puede detectar que algo no funciona en la forma en que se mueve la audiencia (Davis, 2005: 136).

Influído por Spinoza, Manuel DeLanda aborda también el ámbito de los encuentros sociales y la comunicación no lingüística. Aparte de la vía lingüística, «cada participante (…) se expresa a través de gestos faciales, posturas, vestimentas (…). Los componentes materiales del ensamblaje son más evidentes, consistiendo en la reunión de los cuerpos en un espacio» (De-Landa, 2016: 28). Apoyado también en el sociólogo Erving Goffman, hará referencia a la necesaria proximidad
de los cuerpos, variable según el tipo de intercambio producido, así como al nivel de atención puesto en el contenido en su intercambio. En el caso especial de los encuentros multimedia a los que nos estamos refiriendo, la atención se pondría tanto en las conversaciones o la danza colectiva entre participantes como en los estímulos sonoros y visuales del ambiente. También importa la corporalidad y la expresividad corporal de quien se encarga de la producción de imágenes en esa «conversación». Las convenciones sociales definirán el grado de territorialización de la conversación (DeLanda, 2016: 28, 29).

Visto este enfoque spinoziano, podemos observar la conexión, o el tipo de conversación, que se da, por ejemplo, en las sesiones de Vjing. Dada una audiencia performer concreta, el par DJ-VJ debe hacerse, como par artista, preguntas igualmente concretas: ¿a qué es indiferente este grupo de individuos?, ¿a qué reacciona positiva o negativamente, qué afectos alimentan el espíritu de grupo y qué golpes o qué modulaciones actúan como un veneno?, ¿cuál es el umbral en que se satura un estímulo y el alimento se convierte en veneno para la audiencia? Todo ello, teniendo en cuenta que lo que disminuye la potencia rítmica de unos, o reduce la cohesión social, asegura en otros la permanencia del abrazo colectivo, impulsa la intensidad. Por supuesto, todo esto ocurre de forma casi instintiva y cruzada por deseos espontáneos que la propia creatividad de la audiencia puede suscitar en el par artista.

Permeabilidad y flexibilidad de la trayectoria

Lo anterior refuerza la idea de que la imagen penetra íntimamente en una dimensión interpersonal. Participa en el devenir de la propia situación social y lo consigue gracias a la permeabilidad y flexibilidad de su sesión. Habría que hacer, además, una distinción entre las relaciones expresivas intensas y extensas acontecidas en tal situación (DeLanda, 2016: 110). En función de los ritmos, ¿cómo se componen los individuos para constituir una unidad expresiva superior en escala, el entrecuerpo? ¿Cómo puede el par DJ-VJ atraer a la audiencia a su mundo de afectos rítmicos, aún conservando o respetando cada uno de los mundos afectivos y sus propias relaciones rítmicas? ¿Cómo aumentar la densidad, la movilidad, la conectividad o el contenido emocional de las relaciones? Son preguntas básicas que tienen directamente que ver con lo extenso y lo intenso (DeLanda, 2006: 56-57), y que articula en reciprocidad el flujo de imagen. «En fin, la etología estudia las composiciones de relaciones o de poderes de cosas diferentes (...) se trata de saber si las relaciones (¿y cuáles?) pueden componerse directamente para formar una nueva relación más “extensa”, o si los poderes pueden componerse directamente para constituir un poder, una potencia más “intensa”. Y, por ejemplo, ¿cuándo una sociabilidad extensa deviene por sí misma intensa? ¿En qué orden y cómo componer las potencias, las velocidades y las lentitudes?» (Deleuze, 1970: 154).

En función de lo intenso y lo extenso, de cómo los performers-audiencia se expresan como compuesto social, como entrecuerpo, la sesión toma una u otra trayectoria. Siguiendo con el ejemplo de una sesión de DJ-VJ, su actividad debe ser permeable y flexible a la expresión del entrecuerpo. Karel Sidney Doing constata que un punto en común y específico de muchos de los artistas que trabajan con imágenes de este tipo es la idea de actuar con responsabilidad hacia los espacios temporales en los que las performances tienen lugar (Doing, 2017: 17). Puesto que el ambiente es resultado de una circunstancia social única, la expresión del entrecuerpo será igualmente singular y, del mismo modo, lo será el flujo de imagen que traza la trayectoria. El desarrollo de la sesión, ya sea un evento Exploding plastic inevitable, una acción multimedia del colectivo industrial Throbbing Gristle o un set de DJ-Vjing, podrá intentar repetirse en otra sesión, pero siempre se repetirá de forma diferente. Hasta el momento no se ha dicho nada sobre las destrezas del artista, pero su permeabilidad a los estados mentales de los asistentes, así como su flexibilidad a la hora de cambiar la trayectoria de la sesión, resulta fundamental. En este sentido, sus virtudes serán opuestas a las del rígido régimen industrial en el que la variación debe ser sistemáticamente abolida (DeLanda, 2016: 81).

La imagen-ambiente se deriva de un ambiente para el que, a su vez, es destino, en otras palabras: se produce por retroalimentación con dicho ambiente. Producción y recepción se tocan. Por tanto, no es, meramente, una imagen inmersa en un ambiente. Por eso debe hablarse de la permeabilidad, flexibilidad y variabilidad potenciales en la trayectoria de flujo audiovisual. La trayectoria de flujo es la trayectoria rítmica estructurada que sigue la sesión, sometida a su propia inercia y a las fuerzas inherentes al medio con el que conjuntamente se desarrolla. Tales fuerzas (fuerzas de sustentación, fuerzas de resistencia, por ejemplo, pero también fuerzas creativas) permiten conocer el tipo de fricción afectiva del encuentro o sesión.

Conclusiones al punto 2

Hemos tratado de hacer patente la circunstancia de que una creatividad ambiental inspira cambios en la producción de este tipo de imágenes. Ello es algo que definitivamente caracteriza a este tipo de imagen. Si al mostrar las particularidades de la imagen-ambiente estableciamos una conexión con el llamado cine expandido, ahora vemos cómo esa etiqueta se vuelve demasiado laxa para este tipo de imagen. Vemos también cómo propuestas como los Infinity environments de Wheeler pertenecen a otra tipología de imagen al no darse una verdadera reciprocidad afectiva.

Como ha podido verse, este segundo concepto lleva una orientación que nos conduce al acto creativo. De un modo u otro, la realidad ambiental irrumpir en el propio proceso de expresión artística, estableciéndose un juego de ecos o resonancias expresivas con el
ambiente (Doing, 2017). En este sentido, resulta extraño no referirse a los artistas de performance de los años sesenta, quienes hicieron una operación conceptual similar: puesto que el teatro perdía su potencia por la mera existencia de un marco escénico que marcaba unos límites y una distancia simbólica, situaron sus acciones fuera del marco que encuadra la acción (Kaprow, 1961).

Queda entonces dibujada la idea de que lo específico de este tipo de imagen no es su capacidad de envolver a una audiencia expectante. Lo que realmente define a la imagen-ambiente es la capacidad de involucrarse con esta audiencia activa en un mismo devenir.

3. El entrecuerpo o las alianzas afectivas

El concepto de imagen-ambiente, según hemos visto, delimita un modo en que las imágenes se presentan más allá de los límites tradicionales, pues la imagen entra a formar parte de juegos de afección recíproca y de socialización. Ya no tiene sentido hablar de forma aislada de la expresión contenido en el flujo de imágenes o, exclusivamente, de lo que ocurre dentro de los márgenes de la pantalla. Por tanto, planteamos un interrogante que podría apuntar hacia una posible respuesta: ¿debemos referirnos a un compuesto social altamente expresivo con el cual se implica y, a su vez, se compone la imagen?

En las siguientes líneas, llamaremos entrecuerpo a este compuesto afectivo. Para entender mejor esta noción, lo que importa no son los cuerpos ni lo que se encuentra entre ellos, sino las alianzas en las que se implican y lo que conjuntamente mediante relaciones de exterioridad hacen surgir. Inspirados en la teoría de los ensamblajes de DeLanda, vamos cómo esa propiedad emergente a la que nos estamos refiriendo posee dos dimensiones: una, material, la reunión física de los cuerpos, que podrá ser más o menos densa y estable; la otra, expresiva, ya que se trata de una vivencia rítmica y continuada en el tiempo (DeLanda, 2016: 28), en este caso mediada por la imagen. Teniendo en cuenta los juegos de afección de la imagen, nos aproximaremos a la idea de entrecuerpo a través de tres enfoques: el de las superficies, el del cuerpo y el del acontecimiento.

La imagen-ambiente y las superficies

Con la descentralización del punto de interés de la imagen, su desbor- damiento y la activación de su afuera, se da una absoluta conquista de las superficies por parte de la imagen-luz. La audiencia que queda bajo su influencia se convierte en una especie de dermopantalla fluctuante. Ya no se trata de acceder a lo más profundo del espíritu a través de la imagen, ni de inculcar ciertas ideologías, sino de imágenes que se deslizan sobre la superficie sinuosa de los cuerpos, cuerpos que son ahora materia y expresión luminosa, estableciendo una continuidad ambiental y una correlación entre los cuerpos implicados. Es un nuevo arte de las superficies. Lo decíamos: superficies que mutan su estado y su expresividad, que se muestran o se ocultan, rostros que aparecen desfigurados por los juegos de luz… Recordemos una vez más las palabras de Jonas Mekas: «Ayuda mucho el vernos a nosotros mismos bajo una luz diferente» (Mekas, 1966).

Debemos, así, aludir a una «dermatología general», por usar la expresión de Pau Alsina hablando del arte de los nuevos medios (Alsina, 2010). La superficie corporal de cada participante interactúa en un juego de luces y sombras donde figura y fondo se confunden, donde la apariencia de lo orgánico no se distingue de la apariencia de lo inorgánico, donde todo se contagia de ambiente. De este modo, el concepto de pantalla entra necesariamente en crisis, o bien se abre a nuevas dimensiones (Otxoteko, 2020). Quizá esto sea a lo que se refiere Grau cuando habla de un «paisaje que deviene carne» (Grau, 2005). Incluso la sudoración y el vapor de la transpiración dérmica se implican en el entrecuerpo (Grau, 2005) y en el modo en que este es percibido estéticamente según su relación con la luz: «La luz ambiental se propaga en todas las direcciones, crea reflejos y deforma las percepciones visuales; el cuerpo humano también se vuelve en este sentido superfi cie de proyección» (Otxoteko, 2020).

No se trata de negar el concepto de pantalla, sino de ampliar su sentido con la noción de una pantalla viva, internamente intencio- nizada e impredecible. Por otro lado, el concepto de montaje-ojo (montage-eye) de Chris Marker adquiere aquí una gran relevancia. Marker propuso este concepto en el contexto de su instalación de video Zapping-zone (1990), en la que un conjunto de monitores se distribuía de forma descentralizada en el espacio de una sala. Según su idea, es el propio espectador quien asume la responsabilidad de ejecutar un montaje fílmico a través de su posición en la sala y la dirección de su mirada. Se trata, además, de un montaje que cada espectador-participante realiza a partir de estímulos multisensoriales y no solo visuales y auditivos; ellos mismos, bajo esta nueva luz cambiante, ejercen de estímulo para otros.

La imagen-ambiente y los cuerpos

El entrecuerpo comporta una pluralidad de cuerpos individuales y flujos energéticos asociados rítmicamente. Ello implica un encuentro y, en segundo lugar, una composición. Cuando un cuerpo se encuentra con el nuestro y se compone con él, según Spinoza, experimentamos alegría (Spinoza, 2009). En una sesión de DJ-Vjing, los individuos se afectan mutuamente según dinámicas de movimiento o reposo, de velocidad y de lentitud, de alta densidad y roce intenso, y de baja densidad. Las diferencias térmicas también inciden, haciendo que estos se mantengan juntos al calor del grupo o que se disgreguen en el espacio. La dinámica de los cuerpos define la unidad mayor:
la composición de cuerpos; la imagen-ambiente nunca se limita al nivel de los cuerpos individuales.

Desde la perspectiva de los cuerpos asociados, la imagen y el sonido acarician, sacuden, golpean un sentir colectivo. El participante individual siente su propia singularidad a través de la imagen multisensorial de la que forma parte (Grau, 2011), pero no es capaz de sentir tanto un yo como un nosotros singularizado, esto es, como una comunidad. Resulta pertinente a este respecto la visión de Maturana, para quien la comunicación mediante el lenguaje no consiste en la transmisión de información, sino más bien en el fluir de las configuraciones de coordinaciones sensoriométricas en el estrecho contacto corporal entre organismos vivos (Maturana, 1992: 35). Las imágenes participan en esta comunicación como un pulso o una modulación rítmica; la experiencia comunitaria expresa la reacción a un conjunto de estímulos compartidos (McNeill, 1995), que naturalmente implican a la imagen. Al contrario de la imagen-tiempo descrita por Deleuze (1985), que tiene una dimensión mental, ahora hablamos de un efecto fundamentalmente corporal, en el sentido de que, a tiempo real, el cuerpo de la audiencia muestra excitación, indiferencia o repulsión ante el estímulo; este carácter que ya presuponemos háptico de la imagen ha sido considerado por Pau Alsina:

«Podríamos atrevernos a augurar que esta renovada relevancia del papel del cuerpo y la materialidad manifiesta el tránsito de una concepción de una cultura centrada en la visualidad hacia una concepción de la cultura en su vertiente háptica (…) un desplazamiento de la centralidad de la visión a favor de los sentidos corporales internos como son el tacto o el automovimiento. Así pues, la visión se vuelve háptica. La visualidad entendida de esta manera está formada en términos mucho más “viscerales” que los del poder abstracto de la vista: el cuerpo continúa siendo un marco activo de la imagen, incluso en la imagen digital (…) Tal como nos comenta Hansen, la imagen no puede ya ser restringida al nivel de apariencia superficial, sino que debe ser extendida para acompañar todo el proceso entero según el cual la información es percibida a través de la experiencia corporeizada (…) En este proceso, el cuerpo, más que seleccionar imágenes, opera filtrando información y creando imágenes, alumbrando el mundo en su experiencia» (Alsina, 2010).

La imagen-ambiente es lo que actúa y reacciona; actúa sobre los cuerpos y reacciona en ellos. Por ello, es quizás el tipo de imagen que más se explica por la teoría de las afecciones de Spinoza. Tradicionalmente, las relaciones que se dan entre personas que, al mover sus cuerpos coordinada y rítmicamente, favorecen un estado común han sido descritas como relaciones de interioridad, aludiendo a la espiritualidad humana. Pero, según estudios como los de William McNeill (1995) y Manuel DeLanda (2016), cuando un grupo de personas interactúa de este modo, esta relación no definirá su identidad, pues conecta a cada uno en exterioridad. Es una relación establecida entre personas, dispositivos tecnológicos, elementos arquitectónicos, el propio aire que existe entre ellos… La transmisión de influencias es evidencia de su conexión, pero esa relación materoenergética no les constituye (DeLanda, 2016: 2).

La imagen-ambiente y el acontecimiento

Con la noción imagen-ambiente se restituye la percepción de la imagen o del flujo de imágenes, arrancándola de la representación de los rostros, los cuerpos, los paisajes… y situándola en el plano del acontecimiento. En este giro, encontramos la imagen no solo formándose en el acontecimiento, sino además contribuyendo a su vez a la emergencia de una corporeidad estética que no existía antes. «Aparatos, materiales, objetos, performer(s) y audiencia se unen en un sistema que produce algo artístico in situ dejando rastros de imagen, grabaciones sonoras, fotografías o recuerdos y, aún más importante, nuevos pensamientos y nuevas conexiones en la mente de los participantes» (Doing, 2017: 17).

Así comprendida, la imagen conforma no solo una realidad multisensorial, sino también una experiencia estética impredecible y continuamente variable. En este punto cobra relevancia el concepto de variación continua de Manuel DeLanda (2016: 77) aludiendo a una conveniente flexibilidad de los comportamientos y las rutinas. El hecho de que la imagen se halle plenamente implicada en el devenir social es la mayor evidencia de su condición inmanente; es decir, implicada en el compuesto social, no transciende a las partes (DeLanda, 2016: 71). En el momento en el que las partes dejan de interactuar, el entrecuerpo deja de existir y la imagen pierde su sentido.

Por último, la experiencia no deja de ser subjetiva, pero la percepción del sujeto se reemplaza por la evidencia de un presente constituido y constituyente de una unidad superior en escala, ese nosotros al que Mekas se refería (1966) y que podemos deducir de McNeill (1995). Este presente es denso e intenso en su juego de afecciones, en su carácter irresuelto, preñado de soluciones potenciales, de estados virtuales; la circunstancia no responde a una lógica a partir de la cual se pueda enlazar el pasado con el presente ni este último con el futuro. Este presente de la imagen-ambiente, tal cual lo hemos descrito, es el acontecimiento. No importa lo que se ve en la imagen, importa dilatar e intensificar la visión para dar sentido a una experiencia. Ocurre como si el ambiente en su forma de imagen confirmara que el cuerpo es un grado de potencia (Deleuze, 2001: 38); a otra escala podría decirse lo mismo del entrecuerpo, que igualmente es una unidad material y expresiva capaz de variar de una infinidad de maneras.

Conclusiones al punto 3

Como se ha visto, al decir que el entrecuerpo es una entidad individual emergente (fruto de las alianzas afectivas) que puede involucrar o asociarse a su vez a la imagen-ambiente, no estamos diciendo que sea un ser sensible, sino una manera de ser sensible a la imagen, no ya de forma individual, sino mediante una implicación colectiva.
Nos hemos referido a tres de sus dimensiones para ver a qué niveles la imagen se compone con el ambiente circundante. En este sentido, la espontaneidad de los comportamientos ha sido un aspecto de singular importancia para la imagen-ambiente. Ello explica, por otro lado, la razón por la cual en este artículo apenas nos hemos detenido en el arte institucional, altamente codificado y perteneciente a un territorio cultural en el que las conductas sociales se encuentran muy normalizadas.

Últimas consideraciones

La lógica de la imagen-ambiente forma parte, evidentemente, de la vida social y cultural desde hace varias décadas. Lo cual no excluye cierta paradoja que hemos explicado: aunque innegablemente ha modificado nuestras conductas e interferido activamente en los modos de relación interpersonal, la evidencia nos indica que a nivel conceptual casi todo está por hacer.

Así, el principal objetivo del artículo era tejer un sistema de conceptos que diera cuenta de la especificidad de la imagen-ambiente. Solo de este modo puede legitimarse su rol creativo y social. Con esta exigencia, no solo hemos ido desarrollando un mapa conceptual, sino que, además, hemos llegado a algunas conclusiones que soportan la tesis inicial. La imagen-ambiente, según hemos mostrado, deriva espontáneamente de un complejo compuesto social. Lo anterior establece una diferencia respecto a otros tipos de imágenes. Su especificidad consiste en la capacidad no solo de afectar un ambiente social, sino, además, de ser simultáneamente afectada por éste. Dicho de otro modo, en el encuentro de la audiencia con la imagen, la primera experimenta un devenir estético paralelo al devenir social de la segunda. Tal como hemos argumentado, esto último se evidencia en variaciones de ritmo en el montaje y de estructura, por ejemplo. Así definida, puede decirse que la imagen-ambiente explora una situación ambiental o, en otras palabras, vuelve sensible al ambiente que la propia imagen (formas, intensidades y ritmos de luz) implica.

Ciertamente, la mayor parte de los argumentos utilizados habían sido trazados con anterioridad por otros estudiosos de diferentes campos, pero no se les había hecho converger hasta el momento. Entonces ha sido un trabajo de cableado, necesario para establecer conexiones de tipo conceptual entre ámbitos que en principio parecían lejanos: la teoría de la imagen y su relación con la audiencia (Deleuze, Kaprow), los modos de vida modernos y la forma social de habitar los espacios (Ábalos), la filosofía de los afectos y las teorías de la comunicación no verbal y del cuerpo grupal (DeLanda, Maturana, McNeill), el ámbito del arte underground y el fenómeno de los visual-jockeys (Mekas, Jung, Lee, Choi y Kim).

Al hablar de esta lógica en que las imágenes se presentan, hemos tenido que referirnos a otro tipo de asuntos que, en principio, exceden el territorio propiamente dicho de la imagen. Ello no quiere decir que la imagen-ambiente no tenga autonomía. Pero difícilmente estaremos en condiciones de comprender sus capacidades de afección si no entendemos qué es a lo que se afecta o por lo que es afectada.

Referencias bibliográficas

Ábalos, Iñaki. 2001. La buena vida. Visita guiada a las casas de la modernidad. Barcelona: Gustavo Gili.
Alsina, Pau. 2010. «Lo más profundo es la piel. Cuerpo, tecnología y neo-materialismo en el Media Art». En «Universos y Metaversos: Aplicaciones Artísticas de los Nuevos Medios». Grupo de investigación Arte, Arquitectura y Sociedad digital (AASD), 1-12.
Davis, Flora. 2005. La comunicación no verbal. Madrid: Alianza Editorial.
DeLanda, Manuel. 2006. A new philosophy of society. Assemblage theory and social complexity. New York: Continuum.
DeLanda, Manuel. 2016. Assemblage theory. Edinburgh: Edinburgh University Press.
Deleuze, Gilles. 1970. Spinoza: filosofía práctica. Barcelona: TusQuets, 2001.
Deleuze, Gilles. 1983. «Sobre la imagen movimiento». Cahiers du Cinéma 352. En Conversaciones (1971-1990). Valencia: Pre-textos, 2006: 77-94.
Deleuze, Gilles. 1985. «Sobre la imagen tiempo». Cahiers du Cinéma 334. En Conversaciones (1971-1990). Valencia: Pre-textos, 2006: 95-102.
Grau, Andréï. 2005. «When the landscape becomes flesh: an investigation into body boundaries with special reference to tiki dance and western classical ballet». Body & Society 11(4): 141-163. https://doi.org/10.1177/1357034X050508024.
Grau, Andréï. 2011. «Dancing bodies, spaces/places and the senses: A cross-cultural investigation». Journal of Dance & Somatic Practices 3(1,2): 5-24. https://doi.org/10.1386/dsp.3.1-2.5_1.
JUNG, Haehyun, Joohun Lee, Hyun-Ju Choi, Hyunggi Kim. 2014. «Real-time DJING + VJING with interactive elements». Contemporary Engineering Sciences 7(24): 1321-1327. http://doi.org/10.12988/ces.2014.49163.
Kaprow, Allan. 1961. «Happenings in the New York Scene». En Essays on the blurring of art and life, editado por Jeff Kelley. Berkeley, Los Ángeles (CA): University of California Press, 2003: 15-26.
Maturana, Humberto. 1992. El sentido de lo humano. Santiago de Chile: Ediciones Pedagógicas Chilenas.
McNeill, William H. 1995. Keeping together in time. Dance and drill in human history. Cambridge (MA): Harvard University Press.
Mekas, Jonas. 1964. «Sobre el ojo dilatado». En *Diario de cine. El nacimiento del nuevo cine norteamericano*. México: Editorial Mangos de Hacha, 2013: 141-144.

Mekas, Jonas. 1966. «Más sobre la luz del estrobo y la multimedia». En *Diario de cine. El nacimiento del nuevo cine norteamericano*. México: Editorial Mangos de Hacha, 2013: 288-292.

Morrison, James C. 1989. «Why Spinoza had no aesthetics». The Journal of Aesthetics and Art Criticism 47(4): 359-365.

Otxoteko, Mikel. 2020. «Vjing. Estética y política de la imagen-ambiente». Artnodes 25: 1-9.

Raban, William. 2011. «Reflexivity and expanded cinema: a cinema of transgression?». En *Expanded cinema art performance film*, 98-107. London: Tate Publishing.

Rancière, Jacques. 2017. «¿De una imagen a otra? Deleuze y las edades del cine». August 1. Accessed May 5, 2020. https://elzapatodeherzog.wordpress.com/2017/08/01/de-una-imagen-a-otra-deleuze-y-las-edades-del-cine/.

Ruddick, Susan. 2010. «The politics of affect. Spinoza in the work of Negri and Deleuze». Theory, Culture & Society 27(4): 21-45. https://doi.org/10.1177/0263276410372235.

Sedeño Valdellós, Ana María. 2005. «El fenómeno de los videojockey». Área Abierta 12: 1-7.

Doing, Karel Sidney. 2017. «Ambient poetics and critical posthumanism in expanded cinema». Tesis, University of the Arts London: London College of Communication.

Snibbe, Scott S. 2001. «Haptic techniques for media control». En *Proceedings of the 14th Annual ACM*, 1-10. Florida.

Spinoza, Baruch. 2009. *Ética demostrada según el orden geométrico*. Madrid: Tecnos.

Taylor Stuart, Shahram Izadi, David Kirk, Richard Harper, Armando Garcia-Mendoza. 2009. «Turning the tables: an interactive surface for Vjing». CHI ’09: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems April 2009, 1251-1254. https://dl.acm.org/doi/10.1145/1518701.1518888.
La dimensión ambiental de las imágenes. La imagen ambiente y sus principales conceptos

CV

Mikel Otxoteko
Investigador independiente

Mikel Otxoteko trabaja como profesor en los grados oficiales de Diseño del Centro Universitario CESINE, Santander. En 2014 recibió la mención internacional con su tesis doctoral Audiovisual Arts and Neo-materialism, con invitación del cineasta Manuel DeLanda como investigador visitante en el Departamento de Arquitectura del Pratt Institute of Art & Design de Nueva York. Inmediatamente después co-coordinó el Festival Puntuetan cine + periferia y desarrolló proyectos artísticos de investigación como Dance & Drill, que se materializó en proyecciones y exposiciones multimedia en lugares como el Musée d’Art Contemporain Les Abattoirs de Toulouse, el Centre del Carme Cultura Contemporánea de Valencia o The Rag Factory en Londres. Otxoteko sitúa su práctica entre el video de instalación y el cine. Su trabajo ha sido reconocido y avalado por instituciones como Rencontres Internationales Traverse Vidéo o el Centro Internacional de Cultura Contemporánea, Tabakalera. En octubre de 2020 estrenará en el festival Cinespaña de Toulouse un documental sobre la crisis del bacalao en Terranova.