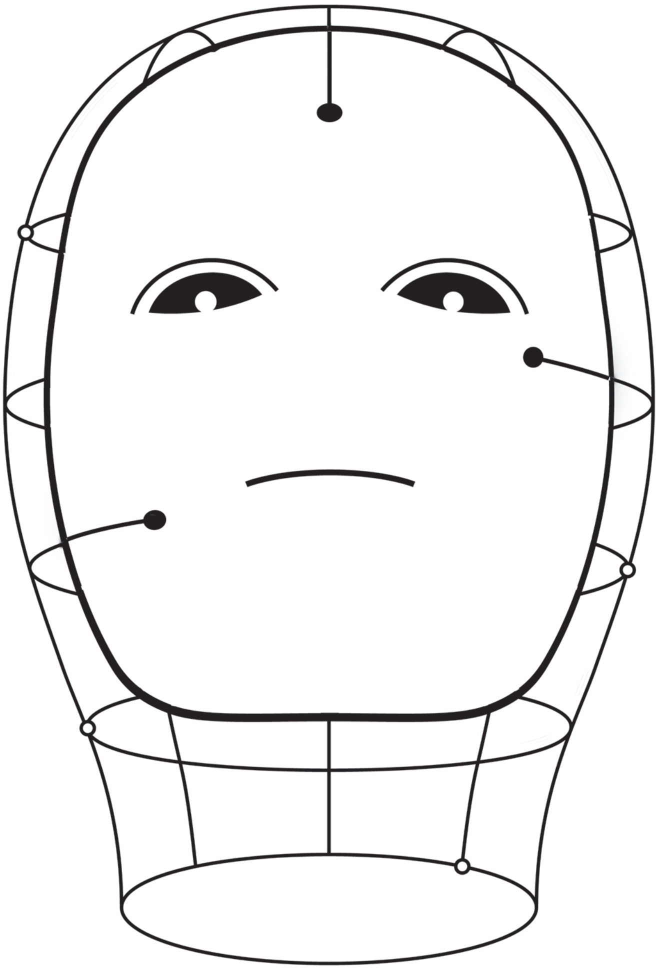
D3US EX M4CH1NA.
Art and Artificial Intelligence
26 November 2019 - 16 May, 2020

Curators:
Karin Ohlenschläger and Pau Waelder

Artists: Memo Akten, Harold Cohen, Jake Elwes, Lynn Hershman Leeson, Felix Luque, Lauren McCarthy, Anna Ridler, Guido Segni, Caroline Sinders, Christa Sommerer & Laurent Mignonneau, Jenna Sutela, Patrick Tresset, Pinar Yoldas

Design: Juan Jareño

Text: Pau Waelder



D3US EX M4CH1NA

ARTE E INTELIGENCIA ARTIFICIAL

D3US EX M4CH1NA

Before the algorithmic gods

Medea, Euripides' well-known tragedy, ends with the protagonist avoiding being executed for her multiple murders thanks to the god Helios, who lends her his chariot to escape from the court of Corinth. On stage, a divine intervention of this kind was represented with an actor appearing from above, riding a crane or other mechanical device. The actions of this character gave a satisfying twist to the story, regardless of its consistency with the situations in which it had unfolded. The philosopher Aristotle criticized this resource, stating that "the outcome of the work must result from the piece itself, and not, as in the Medea, from a strange intervention".[1] The practice of taking an actor out on a crane to create a rapid denouement of the plot of a play would later be known as the Latin expression deus ex machina ("god from the machine"). This expression is now used to refer to any element that is introduced into a story without any relation to the facts that have been narrated and changes its outcome completely, often to meet the expectations of the audience, but resulting in an implausible story. As a narrative resource, it is excessively easy, and therefore the expression is usually pejorative. However,[2] an interesting aspect of the deus ex machina is that it introduces what one wants to happen, even if it is opposed to the reality of the events being narrated. The story takes a sudden turn towards a happy ending, salvation at the critical moment, reconciliation that seemed impossible. Everything happens quickly and expeditiously, with no room to question the reason for this outcome, since the actions of a god or those of destiny are not in question.

Artificial intelligence (AI) is currently monopolizing the media discourse about our relationship with technology as a deus ex machina: an all-knowing, all-powerful entity that promises to easily solve all our problems. In recent years, large technology companies have entered a competition to lead the development of AI, which has resulted in spectacular applications of this technology to concrete solutions, but also in ambitious claims about what it can achieve in the near future. In 2016, the former CEO of Google, Eric Schmidt, stated that artificial intelligence would solve climate change, poverty, war and cancer.[3] Mike Schroepfer, chief technology officer at Facebook, spoke at the same time of how AI would be a global transformation. However, these promises seem far from being fulfilled, even in more specific (but also very ambitious) areas such as intelligent cars. In 2017, John Krafcik, Waymo's CEO, said the company was on the verge of making driverless cars, but has failed to deliver.[4] However, making bold and overly optimistic predictions is not something that corporate executives have started to do in recent years, but is closely linked to the very history of artificial intelligence. In 1967, Marvin Minsky, one of the "fathers" of AI, stated that "within a generation, the problem of artificial intelligence will basically be solved.[5] Since then, bets have been made on when computer intelligence will surpass human intelligence, a turning point that futurist Ray Kurzweil has

1 Aristotle, Poetics, XV. Translation by José Alsina Clota. Barcelona: Icaria, 1987, p.47.
2 Robert McKee calls the deus ex machina "the worst sin of any screenwriter" and, even more so, "an insult to the public".
Robert McKee, The Script. Substance, structure, style and principles of screenwriting. Trans. Jessica Lockhart. Barcelona: Alba Editorial, 2011.
3 Gary Marcus y Ernest Davis, Rebooting AI. Building Artificial Intelligence We Can Trust. Nueva York: Pantheon Books, 2019, p.11
4 Marcus and Davis, Rebooting AI, p.10
5 Marcus and Davis, Rebooting AI, p.8

placed in the year 2029.[6] The idea of an intelligent machine undoubtedly raises great hopes and also apocalyptic fears, since machines, as mere “dumb” tools, have brought humanity unimaginable progress and also terrible devastation. What can we expect from a computer thousands of times more intelligent than human beings, and with a conscience of its own? Is this possible? If so, will it come to save us or destroy us? The first things that come to mind are killer robots like in the films Blade Runner (Ridley Scott, 1982), The Terminator (James Cameron, 1984) or Ex Machina (Alex Garland, 2014), as well as the unstable personalities of Hal 9000 in 2001: Space Odyssey (Stanley Kubrick, 1968), Wintermute in William Gibson’s novel Neuromancer (1984), or Samantha in the film Her (Spike Jonze, 2013). These stories raise fascinating and also frightening scenarios, but the reality is that nowadays robots are far from achieving the autonomy and ability to act as humans that replicants or T-800 have, while artificial intelligence programs can beat a human opponent in a game of chess or go, and also book a table in a restaurant by phone, but are not able to understand a joke or follow a complex conversation. The gap between what we expect from artificial intelligence and what it is capable of producing today is enormous. Researchers Gary Marcus and Ernest Davis call it “the AI chasm”[7] and argue that it is mainly due to three factors: Firstly, we tend to “humanize” machines and give them their own will and even personality; secondly, AI’s progress in very specific and limited tasks (such as winning chess games) is assumed to be progress at a much broader level; finally, when you manage to develop a system that works in certain situations (such as getting a car to drive alone through the desert), it is assumed that it will also work in many others, although this is not always true.

In short, the development of artificial intelligence is surrounded by stories, some of which describe real innovations, while others reflect the bold expectations of researchers, the ambitions of corporations, or our own hopes and fears of a machine we do not fully understand. The way in which artificial intelligence programs, processing millions of pieces of data in milliseconds, respond to our requests or even guess our desires seems like magic, and in fact this is how technological innovations have often been presented over the centuries. Greek mythology contains numerous accounts of automatons and animated objects, created by the god Hephaestus, coming to life in mysterious ways.[8] Today, we do not conceive of a smartphone as a divine creation, but it still appears to us as a black box whose inner workings are to some extent unknown. When this black box, besides doing everything we ask it to do, is able to predict what we are going to want or do, it is not difficult to grant it supernatural powers or to let imagination assign it abilities that go beyond what it can really do. According to several researchers,[9] an essential aspect of AI is its ability to predict: while we don’t (yet) have machines that can think, we do have machines that can make predictions from the data they have, recognize patterns, objects or other elements and build their own models.[10] This is certainly a spectacular advance that is revolutionizing the technological industry and is beginning to have profound consequences in the most industrialized societies. For this reason, artificial intelligence must be seen as more than a black box that “issues verdicts from

6 Ray Kurzweil, The Age Of Spiritual Machines. When Computers Exceed Human Intelligence. Nueva York: Penguin Books, 1999, p.161.
7 Marcus and Davis, Rebooting AI, p.27-31
8 Adrienne Mayor, Gods and Robots. Myths, Machines and Ancient Dreams of Technology. Princeton-Oxford: Princeton University Press, 2018, p.164-191.
9 In analyses of the perception of artificial intelligence, it is common to refer to the magical or the divine, as do, among others, Pedro Domingos, Adrienne Mayor, Cathy O’Neil, or Meredith Broussard (see note 8 and following).
10 Meredith Broussard, Artificial Unintelligence. How Computers Misunderstand the World. Cambridge-Londres: The MIT Press, 2018, p.41 y ss.

the algorithmic gods," in the words of mathematician Cathy O'Neil.[11] We must understand what artificial intelligence is, what its objectives are and what its real achievements are.

What is artificial intelligence?

The term "artificial intelligence" was coined by computer scientist John McCarthy in 1956 in the context of the Dartmouth Conference, the first academic symposium to discuss the possibility of making a machine capable of reasoning like a human being. The implications of this ambitious goal encompass various disciplines, such as mathematics, philosophy, neuroscience, psychology and biology. All of them will be involved in the development of AI, but it is the computer sciences that assume the project of creating an intelligent machine as a field of research. The machine in question is, logically, a computer (specifically a virtual machine, i.e. a program) and the methods used to achieve something similar to human intelligence are data processing, mathematical calculation, statistical analysis and the development of complex algorithms. These methods involve translating the different aspects of reasoning into models that can be processed by a computer. Dartmouth's initial proposal makes this objective clear:

> "Study must proceed on the basis of the conjecture that every aspect of learning or of any other characteristic of intelligence can, in principle, be described with such precision that a machine can simulate it"[12]

This approach finds a notable precedent in the article that, in 1950, the mathematician Alan Turing dedicated to the question "can machines think" and in which he posed a "game of imitation", later known as Turing's Test.[13] The game consists of a person having to find out (by means of questions to which he or she receives written answers) whether he or she is talking to a human being or a machine. Turing therefore shifts the question of whether a machine can think to the question of whether it is possible to develop a computer capable of imitating human intelligence. Starting what will later become a habit among artificial intelligence researchers, Turing predicts that in about 50 years there will be computers capable of tricking a human being and in general it will be possible to speak of thinking machines.[14] The mathematician also anticipates the many objections that AI still faces today and proposes that machines be made to compete with humans in "purely intellectual fields," such as chess.[15]

Another contribution of Turing is to propose that, instead of simulating an adult mind through a program, one should simulate that of a child. This implies not only programming a simpler machine, but also introducing the concept of learning, which could be done by means of positive or negative reinforcements. This idea has been implemented in the branch of AI known as machine learning, which consists of creating programs capable of modifying their operation based on the data supplied to them and the results obtained.[16] Unlike a usual program, which always executes the same instructions, a learning program can modify the parameters of

---

11 Cathy O'Neil, Weapons of Math Destruction. How Big Data Increases Inequality And Threatens Democracy. Nueva York: Crown Publishing, 2016, p.13.
12 James Moor, "The Dartmouth College Artificial Intelligence Conference: The Next Fifty Years." AI Magazine, Vol.27, Núm.4. American Association for Artificial Intelligence, 2006, p.87.
13 Alan M. Turing, "Computing Machinery and Intelligence." Mind, New Series, Vol. 59, No. 236 (October 1950), pp. 433-460. Published by Oxford University Press. http://www.jstor.org/stable/2251299
14 Turing, "Computing Machinery and Intelligence", p.442.
15 Turing, "Computing Machinery and Intelligence", p.460.
16 Ethem Alpaydin, Machine Learning. The New AI. Cambridge – Londres: The MIT Press, 2016, p.25.

the model it has been given to adapt to specific objectives, such as detecting people's faces in a set of photographs or a video capture. As Turing suggested, these programs can be "educated" through learning techniques. There are essentially three types of learning used in machine learning: supervised learning, in which the program is told what result it wants to achieve and is given feedback indicating whether it has achieved it or not; unsupervised learning, in which data are supplied to the program and the program discovers recurrences and patterns on its own; and finally, reinforcement learning, in which a system of messages is introduced indicating to the program whether it has achieved its goal or not.[17]

Among the techniques used in machine learning, one of the most common is that of artificial neural networks (ANN), a computer model inspired by the structure of the brain that is composed of thousands (or millions) of units connected to each other, each of which processes the data it receives and provides a result, which serves as input to other units. These processing units, or neurons, are structured in layers in a complex system that subjects the entered data to numerous mathematical operations and generates an output that the program itself can alter according to the feedback received. Thanks to the network structure, certain neurons can be activated or the connections between them enhanced to progressively achieve a result that meets the desired objective. In order to be able to classify the information it is processing, the artificial neuronal network has a "training set," which is usually data selected and labeled according to the task to be performed. Thus, for example, to "train" an artificial neural network to recognize human faces, it is provided with thousands of portraits of people. Among the artificial neural networks, a system known as generative adversarial network (GAN), which is part of the unsupervised automatic learning, should be highlighted. A GAN consists of two artificial neural networks, one dedicated to generating outputs (the "generator") and the other dedicated to determining whether these outputs correspond to a pre-established objective (the "discriminator"). Usually, GANs are used to generate images that look real, based on a training set made up of thousands of photographs of the type to be created. Using this information, the generator creates a new image and the discriminator determines whether that image is valid, i.e. whether it can pass through one of the images in the training set.[18] If the image is rejected, the generator modifies its parameters and repeats the process, which continues in a feedback loop. The images approved by the discriminator are supplied to the person using the GSN. The person using the GAN has the possibility to provide feedback or adjust the system to obtain better results. GANs have achieved enormous popularity since they were invented by Ian Goodfellow and his team in 2014 because of their ability to generate images that deceive the human eye, although they also find a wide variety of uses linked to image analysis and processing. Today, an increasing number of artists are using GANs in the creation of artistic projects, partly because some of these programs are accessible in software repositories such as Github, and also because they introduce the possibility of making the machine a co-author of the work: the software participates in the creation of the piece with a certain level of autonomy, in a kind of dialogue with the artist herself. This autonomy is evident in the fact that it is very difficult to understand exactly what happens inside an ANN, so that its functioning is mysterious and almost seems magical.[19]

Applications of artificial neural networks and other artificial intelligence systems have proven successful in solving specific tasks, where these programs are even capable of surpassing

17 Margaret A. Boden , AI. Its Nature and Future. Oxford: Oxford University Press, 2016, p.47-48.
18 It is interesting to see how this process resembles the Turing Test, the aim of the generator being to "fool" the discriminator.
19 Boden, AI. Its Nature and Future, p.78.

humans (whether they are experts in radiology or world champions in Go or chess). However, what has not yet been achieved is a general artificial intelligence, which can be applied indiscriminately to any real-world situation or environment. This is obviously much more complex than getting a program to generate non-existent photographs or to recognize a person's face. Even automating a specific task, but in a changing environment, whose conditions cannot be foreseen (driving a car, for example), is proving to be much more difficult than it seemed. Therefore, what has developed so far is what is known as a "weak AI," limited to a specific task, as opposed to what would be a "strong AI," which would theoretically have a conscience of its own and would be the type of artificial intelligence that, according to science fiction and the predictions of some researchers, would manage to dominate the world and possibly end the human race. To reach this point, it would be necessary to achieve a general artificial intelligence (AGI), which would continue to develop itself until it reached the point where it would surpass human intelligence, becoming an artificial superintelligence (ASI). What this ASI would decide to do with humans is not clear, but in any case it would have reached a point in technological development known as "the singularity," in which machines would be beyond our control.[20] Whether or not such a singularity will ever come is one of many debates among AI researchers, as well as one of many myths surrounding artificial intelligence. It is easy to get away with apocalyptic visions such as those described in The Matrix trilogy (Lilly and Lana Wachowsky, 1999-2003), but while these do not materialize, it is necessary to look at some current challenges and dangers posed by AI today.

Scientists meeting at Dartmouth more than 60 years ago agreed that a way had to be found to describe any aspect of intelligence accurately enough for a machine to simulate it. This involves providing the machine with a large amount of data and a series of algorithms through which to process it. Gathering data has been the priority of researchers, and later large companies, who have been involved in the development of artificial intelligence. For an AI to be able to write a text like a human, talk to a person or recognize a face, it must be provided with thousands of texts, audio recordings and photographs, which must also be authentic, taken from real situations and people. Therefore, companies like Google, Apple, Facebook or Amazon use the data generated by their users, the contents of social networks and public forums, countless scanned books, millions of photos and videos hosted on their servers, as well as recordings and other data obtained from the devices they market. This huge volume of data did not exist a few years ago, and is one of the factors (along with hardware development) that has facilitated the recent explosion of AI.[21] More sophisticated AI therefore requires that we are willing to give up our data and give up our privacy.

But this is not the only renunciation that the current development of AI leads to. Once our data are provided to the program, it processes them using algorithms that we do not know and that reflect the prejudices and interests of those who have elaborated them. Even if it has been programmed with the best intentions, an artificial intelligence system can fail in the analysis of the data and the predictions it extracts from them. Because its precise functioning is difficult to understand even for those who have developed it, auditing one of these systems is extremely complex, and is beyond the reach of those who are affected by its decisions. Furthermore, unlike human decisions, which are subject to rethinking, automated systems continue to apply the same instructions unless they are reprogrammed.[22] This limitation of weak AI in turn raises fears about certain agents who are not as intelligent as they should

20 Boden, AI. Its Nature and Future, p.147-148.
21 Marcus and Davis, Rebooting AI, p.17
22 O'Neil, Weapons of Math Destruction, p.202.

be and whose actions can have terrible consequences for people's lives (e.g. weapons and autonomous vehicles or surveillance and identification systems for persons employed by the police). The solution to limited AI is to provide more data to the system, but this again leads to a renunciation of privacy and the need for ethical use of the information collected. Some proponents of AI therefore propose the combination of machines and humans, so that the former focus on the analysis of large amounts of data and the resolution of repetitive tasks, while the latter deal with more ambiguous or complex situations, exercise judgment based on ethical values and take responsibility for their decisions.[23]

Beyond the ethical issues, there are other factors arising from the very materiality of this technology: an artificial intelligence program requires a lot of processing time, powerful hardware and access to data stored in the cloud, all of which translates into a considerable expenditure of energy. Training an AI with a large volume of data generates about 284,000 kg of carbon dioxide, equivalent to the emissions of five cars over their lifetime.[24] In 2019, the first experiments with the BigGAN system, which generates high-resolution images, have required the use of 512 TPU processing units (developed by Google) that consume, in the creation of each image, as much electricity as is spent in a home over six months.[25] The carbon footprint is therefore another of the worrying aspects of AI (and more immediate than killer robots), which not only has consequences for the environment but also for access to the very technology that makes more advanced artificial intelligence possible. Only researchers linked to large companies such as Google can use their extensive hardware resources and the necessary funding to pay for the energy consumption arising from their use. This can decisively limit how artificial intelligence research will be developed and which objectives will be prioritized.[26]

Questioning AI through art

Many authors[27] agree that the challenges and dangers posed by artificial intelligence require a critical view and awareness of both its real possibilities and the need to use this technology ethically. AI is not magic, nor is it a neutral and unavoidable force. It is a set of algorithms and computing techniques that must be observed and questioned to prevent them from perpetuating prejudice and inequality or generating other collateral damage. This is not usually within the reach of the general public, since the technological industries only want to show the goodness of their products and are looking for loyal and convinced consumers, not users capable of elaborating a critical reflection on technology. Only the scandalous news about the violation of privacy and the unethical use of user data has forced big companies to rethink (albeit temporarily) their products and policies: Google had to stop marketing its Google Glass augmented reality glasses; Apple has had to shut down the program for listening

---

23 Paul R. Daugherty y H. James Wilson, Human + Machine: Reimaging Work in the Age of AI. Boston: Harvard Business Publishing, 2018, p.186.
24 Karen Hao, "Training a single AI model can emit as much carbon as five cars in their lifetimes," MIT Technology Review, 6 de junio de 2019. https://www.technologyreview.com/s/613630/training-a-single-ai-model-can-emit-as-much-carbon-as-five-cars-in-their-lifetimes/
25 Katharine Schwab, "A Google intern built the AI behind these shockingly good fake images," FastCompany.com, 20/4/2019. https://www.fastcompany.com/90244767/see-the-shockingly-realistic-images-made-by-googles-new-ai
26 Emma Strubell, Ananya Ganesh y Andrew McCallum, "Energy and Policy Considerations for Deep Learning in NLP," 57th Annual Meeting of the Association for Computational Linguistics (ACL), Julio, 2019. arXiv:1906.02243 [cs.CL]
27 Véase, por ejemplo: O'Neil, Weapons of Math Destruction, p.212-216; Marcus y Davis, Rebooting AI, p.32-34; Broussard, Artificial Unintelligence, p.240-241.

to its users' conversations that it had created to improve Siri's performance; the Cambridge Analytica scandal has led Facebook CEO Mark Zuckerberg to mumble weak apologies to the US Senate; Amazon has faced numerous criticisms for listening to users' conversations through devices equipped with its Alexa virtual assistant and for providing police forces with a facial recognition system that makes dangerous mistakes. However, these striking cases are far from changing users' long-term perception of the technology: an elaborate apology and a few promises seem to settle the issue. What is needed, therefore, is to create a framework for thinking about technology that will enable us to take a critical look at it at all times. Not to deny what it brings us, but to understand how it works and our involvement in it.

Roger F. Malina, astronomer and professor emeritus of art and technology at the University of Texas, has indicated on several occasions that art, science and technology can feed off each other as artists open new paths to technology and scientific thought, while these disciplines facilitate new forms of artistic creativity.[28] Art that appropriates scientific and technological advances creates a cultural framework that allows us to understand, experience, and develop a critical, more realistic, and also more engaged stance. The exhibition D3US EX M4CH1NA. Art and Artificial Intelligence pursues this objective by bringing together the work of fourteen artists, whose works address the ways in which we perceive AI and what challenges it poses for us. The pieces can be experienced individually, as part of the discourse that each artist has developed throughout their career, or they can be considered as nodes in a network that builds a portrait of AI through different perspectives and approaches, some more playful and others more reflective, sometimes inviting the viewer to dialogue with the work, and sometimes consciously excluding him or her. Next, we propose a route in which the pieces have been grouped into five major themes, which progressively lead from the human to the machine, from the humble apprentice robot to the cryptic super-intelligence.

Artificial Art

The beginnings of algorithmic art can be traced back to an exhibition of computer-generated drawings that took place at the Institute of Technology of the University of Stuttgart on February 5, 1965. The artist and mathematician Georg Nees presented his work in a room where the philosopher Max Bense taught. At the opening, several artists came together and asked Nees about the process of creating the drawings. As artist and mathematician Frieder Nake, one of the pioneers of algorithmic art, recalls, there was a conversation between Nees and an artist in which the latter asked him if the machine was capable of drawing as he did, to which the mathematician replied: "yes, if you can describe to me exactly how your style is![29] Like the scientists at Dartmouth, Nees believed that it was possible for a machine to create art if it could be given enough information to simulate human creativity. This did not sit well with the artists, who left the exhibition in anger, as Bense tried to calm them down by saying: "Gentlemen, we are talking about artificial art!" With this term (which, according to Nake, was invented at that time) the philosopher was trying to differentiate the art created by the machine from that created by the artists.

28 Pau Waelder, "Roger Malina: "Art Leads to a New Science," LIFE International Art and Artificial Life Competition, December 22, 2014. https://vida.fundaciontelefonica.com/blog/roger-malina-el-arte-conduce-a-una-nueva-ciencia/
29 Frieder Nake, "Roots and randomness –a perspective on the beginnings of digital art", en: Wolf Lieser (ed.), The World of Digital Art. Postdam: h.f. Ullmann, 2010, p.40.

Three years later, in 1968, the young British artist Harold Cohen began his stay as a visiting professor in the Fine Arts department of the University of California in San Diego (UCSD), with the intention of learning to program and apply programming to his artistic research. While Georg Nees was a mathematician who explored artistic creation, Harold Cohen was an artist who delved into the possibilities of computing.[30] They both agreed that, in order to create art with a computer, it was necessary to describe in detail, in symbolic terms, how the machine should generate the work. However, as Nake points out, this description is not limited to a single work, but allows for the generation of an infinite series of works.[31] Therefore, the artist's creation is the description itself, translated into programming code and executed by a program that generates a visual composition through a plotter. Cohen's main creation was AARON, an artificial intelligence program that he began to develop in 1973 and continued to work on until his death in 2016. The artist describes AARON as "a computer program designed to model some aspects of human behavior in artistic creation and produce as a result "freehand" drawings of a highly evocative nature.[32] The program was capable of autonomously generating an infinite series of drawings without any visual information, just a set of rules based on Cohen's experience with the artistic process. The aim was to investigate the simulation of cognition through the creation of works of art, starting from the question: "what would be the minimum condition for a set of strokes to constitute an image? This approach is far from the idea of replacing the artist with a machine, and in fact Cohen insists that AARON is not an artist, nor is it a tool that allows a user to create visual forms: the program generated its own drawings solely and exclusively from the rules inscribed in the programming code. Even so, according to its creator, when presenting AARON's work the public interpreted that the drawings should have been "supplied" to the program by an artist, and later, when they knew how it worked, they assigned it a personality of its own. This anecdote exemplifies what Marcus and Davis call "the credulity gap"[33]: we cannot help but think of machines in cognitive terms, giving them a mind and will of their own. But it is also the effect of the complex system of instructions designed by Cohen. The AARON program originally (in 1979) had about three hundred "productions", or instructions to execute part of an action. These instructions were distributed in different levels, with a hierarchical structure in which the superior levels limited the field of action in the lower ones. Thus, the highest level of the system, called ARTWORK, decided the organization of the drawing as a whole, while the MAPPING level was in charge of assigning the spaces where each element was placed and PLANNING managed the development of each figure. This structure extended to the level of each line, so that a single line was already the result of some twenty or thirty productions on three levels of the system.[34]

Harold Cohen continued to develop AARON over more than three decades, incorporating new elements such as color and the combination of elements created by the program and direct intervention by the artist. His work constitutes a pioneering exploration of the interactions between art and artificial intelligence, in which we see how the principles proposed in Dartmouth are applied to artistic creation, by means of a system that presents certain similarities with the structure of artificial neural networks. As Cohen states, AARON's drawings (of which we present a series made for the Arnolfini Gallery in Bristol in 1983) are evocative, since they are

30 Frieder Nake, "Georg Nees & Harold Cohen: Re:tracing the origins of digital media," en: Oliver Grau, Janina Hoth y Eveline Wandl-Vogt (eds.) Digital Art through the Looking Glass. New strategies for archiving, collecting and preserving in digital humanities. Donau: Donau-Universität, 2019, p.30.
31 Nake, "Georg Nees & Harold Cohen", p.39
32 Harold Cohen, "What is an image?", 1979. AARON's home. http://www.aaronshome.com/aaron/publications/index.html
33 Marcus and Davis, Rebooting AI, p.27.
34 Cohen, "What is an image?"

fascinating both when they are confused with the works of a human artist (with which the program would pass the Turing Test), and when the complex system through which they are generated is understood. Something similar happens with the drawings that Anna Ridler has done for The Fall of the House of Usher I and II (2017). The artist is inspired by the silent short film made by James Sibley Watson and Melville Webber in 1928, freely interpreting the story of Edgar Allan Poe, to elaborate a subtle reflection on the creative process as translation or reinterpretation. Ridler's work consists of an animation generated by an antagonistic generative network (GAN) and a series of two hundred drawings made by her, which reproduce the main scenes of the short film. Instead of creating an animation directly from the drawings, the artist uses them as the training set she supplies to an artificial neural network, so that it learns to generate new images. The product of this first neural network is supplied by Ridler to a second neural network, which in turn generates new images, and finally to a third one. The animation simultaneously shows the production of the three neural networks, allowing a comparison between them and giving rise to a new version of the film that is both narrative and self-referential: the story of the brothers Roderick and Madeline Usher unfolds in a split-screen effect in which the image progressively deteriorates, losing detail but also releasing itself into increasingly delirious forms, in line both with the progressive loss of sanity of the characters created by Poe and with the surrealist style that Sibley Watson and Webber gave to their experimental film. It is possible, however, to identify some key elements, such as the close-ups of the protagonists, the settings, the typographical games introduced by the filmmakers and finally the moon, which appears between the cracks of the house of Usher when it falls apart, acquiring a notorious prominence both in the original story and in the short film. The soundtrack composed by Alec Wilder in 1959, which has been added to the animation, helps to give meaning to the images and to situate the narration. Undoubtedly, the particular visual style of Sibley Watson and Webber's film, to which Ridler's drawings give an even more enigmatic air, is enriched by the interpretation made by artificial neural networks. The images generated by these programs, even when based on real images, are sinister, both recognizable and strange, like a wax figure. The artist consciously employs a recursive automatic learning process to question notions of creativity and originality, while highlighting the role of HLGs and in particular the data used to train them.[35] As a whole, the work poses a game of duplications and delirium that Poe already introduces in his story, Sibley Watson and Webber translate into fascinating visual metaphors, and Anna Ridler interprets in her own graphic style, then releasing the result to successive artificial intelligence programs, in a loop that may have no end. The artist finally wonders which of all these versions is the real work, where the art is, and in fact shows her drawings alongside the animation, not as a work, but as the set of data that has been supplied to the artificial neural network.

Ridler's use of the drawing technique responds to his interest in drawing as a first language, prior to speaking or writing, which in this case serves to create images with which to train the machine. The artist indicates that while a HLG can produce a drawing, it cannot draw.[36] This is true in that the artificial neural network generates the images in a different way, but a machine can also learn to draw. AARON is a program capable of building a drawing by making the same decisions as a person. In the same way, the robots that Patrick Tresset has created for his series Human Studies (2011-2019) produce portraits with impulsive but precise strokes. Each of the robots manufactured by Tresset consists of a desk equipped with a camera, placed on an articulated support, a mechanical arm that holds a pen and a

35 Anna Ridler, "Fall of the House of Usher. Datasets and Decay". Victoria and Albert Museum, 17 de septiembre, 2018. https://www.vam.ac.uk/blog/museum-life/guest-blog-post-fall-of-the-house-of-usher-datasets-and-decay
36 Ridler, "Fall of the House of Usher. Datasets and Decay"

computer housed under the surface of the table. The body of the robot, therefore, is the same desk, from which its only two extremities, an eye and an arm, emerge. The only thing it can do is draw, once it places a sheet of paper on the table, and it does so (in Tresset's words) obsessively, alternating its gaze between what it wants to reproduce and the sheet on which it draws sinuous lines. In this series, what the robots reproduce are portraits of people who agree to sit down and serve as models in sessions that can last up to thirty minutes. The person thus becomes the object of study of the machine, which in this case is not a mere tool through which to obtain a faithful reproduction of oneself in a few minutes. The artist could have created robots capable of reproducing the portraits quickly and with photo-realistic quality, as if he were a photo booth, but he has chosen to use old or low-resolution cameras and give the robots a certain personality. Tresset introduces her own drawing technique in the program that teaches the robots to draw, but modifies their individual behavior, so that each one generates different drawings. The result is that, in this performative installation, the person and the machine exchange their roles, with the robot assuming the creative role and the human being a passive subject who must submit to the scrutiny of the camera, without moving, during a session that extends beyond what is comfortable in our impatient interactions with technology. The drawings are not, therefore, mere digital prints that can be reproduced infinitely, but unique pieces that Tresset progressively adds to a collection of over 10,000 drawings. It is the behaviour of the robots that makes their drawings go beyond mere mechanical reproduction and acquire the quality of a work of art, even though the machines themselves have no conscience or artistic intentions, but react to stimuli and, as AARON did, have the ability to simulate an artistic process.

The works of Harold Cohen, Anna Ridler and Patrick Tresset explore human creativity through artificial intelligence, revealing the extent to which the conviction that a machine cannot create art can be challenged. As Margaret Boden points out, three types of creativity can be distinguished: combinatorial creativity, in which known ideas are combined in a novel way; exploratory creativity, in which established style rules are applied to create something new; and finally transformative creativity, which goes beyond established formats and seeks new structures to accommodate a different idea.[37] In AARON we can see an example of combinatory creativity, being the program capable of creating new forms from established rules. The Tresset robots apply an exploratory creativity, by giving their own personality to drawings created on the basis of a common technique. Ridler's short film, finally, exemplifies transformative creativity by developing a new version of Sibley Watson and Webber's film that reinterprets their work and that of Poe, resulting in a radically different piece. Obviously, the machines have not created these works on their own, but as part of a process devised by the artists. Therefore, the question of creativity is resolved not in a substitution of the person by the machine, but as a collaboration in a system of creation in which both participate. As Frieder Nake states: "as an artist who has decided to develop software to control the operations of a computer, you think about the image, you don't make it. Image processing has now become the task of the computer".[38]

"Here I am, watching you"

Leaving decisions in the hands of machines can lead to a better world, or it can be disastrous.

37 Boden, AI. Its Nature and Future, p.68-69.
38 Nake, "Georg Nees & Harold Cohen", p.39

In both cases, we must observe the trust we place in computers and their apparent impartiality, which in turn depends on the instructions and data provided to them (therefore, ultimately, the problem is not in the machines but in the people who program and control them). In 1950, the same year that Alan Turing was wondering if machines can think, the writer Isaac Asimov published the story “The Avoidable Conflict,” in which he imagined a stable and peaceful world, without war or hunger, controlled by four great machines that made all economic decisions at the global level. These machines were capable of analyzing an immense volume of data with such complexity of calculations that humans could no longer understand or control their operation. One day, the machines start making decisions that seem wrong, since they generate certain inconveniences in some regions of the planet. The “errors” turn out to be a manipulation calculated by the machines, which are capable of predicting the actions of humans who ignore their indications in order to redirect them towards an optimal functioning of the system. The machine punctually harms a few humans for the benefit of all humanity, and at the same time ensures that its dictates cannot be disobeyed.

This type of benevolent automated government is what Pinar Yoldas proposes in her video The Kitty AI: Artificial Intelligence for Governance (2016), a science fiction story set in the year 2039 in which an artificial intelligence with the appearance of a cat has become the first non-human leader of a city. AI is located in an area of the world where there are no political parties and can govern by means of a network of artificial intelligences and direct interaction with citizens through their mobile devices. Kitty has detailed information about all aspects of the city (infrastructure, demography, vehicle and people traffic, etc.) and can respond to citizens’ requests when they reach a critical mass. Therefore, just like the machines imagined by Asimov, it does not serve the interests of a single person but those that affect the collective of citizens. With a tender child’s voice, the cat tells us “I am your absolute governor” and reminds us that she controls all the systems that make our daily lives possible. “Here I am, watching over you,” adds Kitty, “not like Big Brother, but like a curious cat who adores you...” Yoldas consciously plays with human emotions to indicate how technological industries try to create an affective bond between users and their products, emphasizing the benefits of human relationships mediated by technology. At the same time, he raises the possibility of governing an AI at a time when most people have stopped trusting political parties, given the numerous cases of corruption and the ineptitude of their leaders. The idea of being governed by a virtual cat seems in fact more attractive than the real options available in many countries, and this in turn reminds us that governments are investing more and more in the development of AI. This interest is not so much in technological innovation as in achieving supremacy over other nations: in 2017, Russian President Vladimir Putin stated that whoever succeeds in leading artificial intelligence will rule the world, and that same year President Xi Jinping announced that AI was part of his “great vision for China” with a multi-million dollar investment in a domestic AI industry and the goal of leading the sector globally by 2030. One advantage of Xi Jinping’s vision is the control the Chinese government has over its citizens’ data and the absence of any privacy protections, which allows AI programs to feed a much more difficult to obtain amount and variety of data to other countries. The more data it obtains, the better the AI, and therefore a company with strong artificial intelligence can dominate an industry, outpacing its competitors and benefiting from access to the data it holds.[39] This is what has happened with companies like Microsoft and Intel a few decades ago, and now with Google and Facebook. When a government controls the artificial intelligence sector, for its citizens

39 Ajay Agrawal, Joshua Gans y Avi Goldfarb, Prediction Machines: The Simple Economics of Artificial Intelligence. Boston: Harvard Business Publishing, 2018, p. 205.

this can mean something similar to what Kitty says, but without the tenderness of a pussycat: absolute control, in which there is no possibility of disobedience, since it is anticipated and redirected towards optimum system performance.

Even under a benevolent AI government one might wonder what functions human beings can perform, what jobs we would still have to do. Automation has always led to fear of massive job losses, and according to several researchers, this will happen as artificial intelligence programs and robots are able to perform more tasks with a minimum margin of error. In a first phase, AIs will need to be trained by humans specialized in each area, until they can replace them in most of the tasks they used to perform. This will lead to different situations for the worker, who will either be replaced, or will see part of their work automated, or will perform a different job with new tasks derived from collaboration with the machine.[40] Some authors point out that it will be mainly people with a low level of education and who perform repetitive and manual tasks who will see their jobs disappear.[41] What will happen when machines do most of the work? The most optimistic predictions point to a humanity with a lot of free time: the futurologist Alvin Toffler predicted in 1970 that the number of weekly working hours would be reduced by 50 per cent by the beginning of 2000 (this has not been the case), while the writer Arthur C. Clarke imagined in 1968 that people in the most industrialized countries would live idle and bored, worrying about nothing more than deciding what to watch among hundreds of television channels.[42] Today we have other concerns (including job losses due to automation) but it is true that we have more audiovisual entertainment options than ever before, with content adapted to our tastes thanks to the customization algorithms and data we provide to the platforms. However, perhaps one of the most interesting predictions is the one made by Isaac Asimov in 1977 when he imagined humanity as a “global aristocracy” served by sophisticated machines. In this world, people are fighting tedium through a renewed and broad program of liberal arts, which teaches… machines.[43] In Asimov’s future, therefore, people do not work but cultivate their minds by learning art, philosophy, mathematics and science through software, without the apparent need for human creators or teachers. To anticipate this future, it may be wise to follow the advice of machine learning expert Pedro Domingos, who says: “the best way not to lose your job is to automate it yourself. That way you will have time for all the aspects of it that you could not do before and that a computer will take a long time to run”.[44] In his project Demand Full Laziness (2018-2023), Guido Segni seems to have made Domingos’ words his own: for five years, the artist has decided to delegate part of his artistic production to a series of deep learning algorithms in order to “increase production, overcome the work aspect of artistic creation and free himself more and more from laziness”. Segni lets the program record with a camera its periods of inactivity (sleeping, reading, loitering) and submits the images to a set of GANs, which result in new images generated from an automatic learning process. The works of art created by the software are distributed among sponsors who regularly contribute money to the artist through the Patreon patronage platform. In this project, the artist continues his research on working conditions in the art world and proposes (not without a certain irony) a solution to the dilemma of artistic

40 Agrawal, Gans and Goldfarb, Prediction Machines, p.93-143.
41 See: Boden, AI. Its Nature and Future, p. 160; Agrawal, Gans y Goldfarb, Prediction Machines, p.202; Carys Roberts, Henry Parkes, Rachel Statham y Lesley Rankin, The Future Is Ours. Women, Automation And Equality In The Digital Age. London: IPPR, the Institute for Public Policy Research, 2019, p.10.
42 Stuart J. Russell y Peter Norvig, Artificial Intelligence. A Modern Approach. Tercera edición. Boston: Prentice Hall, 2010, p.1034.
43 Isaac Asimov. Robot Visions. New York: ROC, 1991, p.415.
44 Pedro Domingos. The Master Algorithm. How the quest for the ultimate learning machine will remake our world. Londres: Penguin Books, 2015, p.277.

production in times of maximum precariousness through automation and micro-patronage. Segni cites in this regard the Manifesto for an Accelerationist Policy, in which the economists Alex Williams and Nick Srniček criticize how neoliberal capitalism has not led to a drastic reduction of the working day (as predicted by John Maynard Keynes in 1930), but has led to the elimination of the separation between work and private life, with work being part of all social relations.[45] In this sense, the leisure to which the artist is free, particularly by allowing himself to be observed by the camera, becomes another form of work and paradoxically leads to a situation in which (following the accelerationist theses) leisure no longer exists, since all our activities are susceptible to generating data for a machine.

The context in which Guido Segni carries out his "five-year performance" is also significant. In the intimacy of his home, the artist lies in his bed, reads or simply does nothing while a camera records him. This is not unlike the daily experience of many people who have speakers and other devices equipped with Amazon's Alexa voice assistant, which records their conversations or even, in the case of Echo Look, takes pictures and makes styling recommendations. That we can live with these devices and give them access to our private lives demonstrates a blind trust in computers as machines that are at our service and simplify our lives, ironically contrasted with the fear of being replaced by an AI. However, both situations are linked. According to Pedro Domingos, when we interact with a computer, we do so on two levels: the first is to get what we want, whether it is to search for information, buy a product or listen to music; the second is to provide the computer with information about oneself. The more it knows about us, the better it can serve us, but also manipulate us.[46] Today, large companies such as Google, Apple and Amazon are competing in the marketing of devices specifically designed for consumers' homes, which act like Trojan horses, introducing products and obtaining data into users' private lives. The home is a particularly coveted space by the technology industry, since it is where we carry out routine tasks, maintain social relations with our most intimate circle, rest and spend part of our leisure time. The information that can be obtained in a person's home is enormous, since it is the place where we usually relax and show our authentic personality, habits and desires. Moreover, the devices that enter a house usually have two additional advantages for companies: on the one hand, they are always plugged in and connected to the Internet, so they can transmit data at any time; on the other hand, after a while users are no longer aware of their presence or assume it naturally, so they do not question its use or what they do in front of them. This happens, in part, because we conceive of machines as mere helpers without a will of their own, who listen to us or observe us in order to serve us better and cannot judge us or explain what they have seen or heard. But what if the person behind Alexa is a person? This is the premise of LAUREN (2017), a performance by Lauren McCarthy in which the artist is dedicated to watching and serving various people in their homes by means of a series of connected devices. McCarthy asks several individuals and families to agree to let her install the devices (including cameras, microphones, switches, locks and taps) in their homes and then watches them 24 hours a day, for several days, responding to their requests and offering them advice. She says she feels that she can be better than an AI, since as a person she can better understand their needs and anticipate them.[47] The relationship that is established between the artist and the participants in the performance is a strange exchange between people, mediated by technology and under the protocol of

45 Alex Williams and Nick Srniček, Manifesto for an accelerationist policy Translated into Spanish by Comité Disperso, 2013. https://syntheticedifice.files.wordpress.com/2013/08/manifiesto-aceleracionista1.pdf
46 Pedro Domingos, The Master Algorithm, p.264.
47 Lauren McCarthy, "Feeling at Home: Between Human and AI", Immerse, 8 de enero de 2018. https://immerse.news/feeling-at-home-between-human-and-ai-6047561e7f04

servitude, which McCarthy describes as "an ambiguous space between the human-machine relationship and the human-human relationship. The interactions that the documentation of this work gathers show revealing aspects of the connection between people and also with the machines. One user states that LAUREN's advantage is that it understands her feelings and that this allows her to concentrate on "more important things;" another user says that he feels satisfied because he can maintain a relationship in which everything revolves around him; a user states that he comes to forget that McCarthy is observing him, and when he remembers he looks in the mirror to check his appearance; finally, a user confesses that there are facets of himself that he does not want to share with anyone, and in this sense LAUREN's presence can make him uncomfortable. McCarthy shows with this project how we are willing to exchange intimacy for convenience, letting a set of algorithms control our lives.

Algorithms and prejudices

Through machine learning, a program can identify patterns in a large data set and predict what the new data will look like. This predictive ability, for which more and more applications are found, is based on the information provided to the program as a "training set. If we provide the program with a large number of photos that we have identified as photos of cats, the AI will learn what a cat looks like and will even be able to generate new images of cats. But the program does not know what a cat is, it identifies and creates images of what it has been told is a cat. This also applies, for example, to facial recognition algorithms, such as those used by Amazon Rekognition software, which allows any user to identify objects, scenes, facial expressions and certain activities in video or image files. Amazon markets this software for use in marketing, advertising and multimedia, but is also seeking to sell it to government agencies and law enforcement agencies for person identification. The problem is that the program makes serious errors: when supplied with a set of photos of members of the United States Congress, Rekognition identified twenty-eight of them as criminals, 40% of whom are African Americans.[48] The software's bias can have serious consequences for citizens, which is why a group of 70 civil rights groups, 400 members of the academic community and more than 150,000 signatories have called for Amazon to stop providing facial recognition technology to the U.S. government.

The way training sets are created, which are rarely examined, determines how the world understands an AI. We think of AI algorithms as perfect, unbiased, and inscrutable machines, but in fact they incorporate into their programming all the biases of their programmers and those who have classified and labeled the data provided to them. A good example of this is the ImageNet training package, created in 2009, which has become one of the most widely used in the history of AI over its ten-year history. Co-created by Prof. Fei-Fei Li with the intention of developing a data set of all existing objects, ImageNet has more than 14 million images, arranged in more than 20,000 categories, and is the main resource for object recognition programs. To build this huge database, the Amazon Mechanical Turk platform was used, through which thousands of workers (called Turkers) were commissioned to perform small tasks such as describing the contents of an image, at a rate of about 250 images every five minutes.[49] In 2019, researcher Kate Crawford and artist Trevor Paglen reported that the

48 Jacob Snow, "Amazon's Face Recognition Falsely Matched 28 Members of Congress With Mugshots," American Civil Liberties Union, 26 de julio, 2018. https://www.aclu.org/blog/privacy-technology/surveillance-technologies/amazons-face-recognition-falsely-matched-28
49 John Markoff , "Seeking a Better Way to Find Web Images," The New York Times, 19 de noviembre de 2012.

category "person," which has about 2,500 labels to classify images of people, includes all kinds of abusive, racist and xenophobic terms. The tags applied by the underpaid Turkers had never been reviewed, yet this data set is what tells an AI how to classify people in a facial recognition program. As Crawford and Paglen point out, at a time when AI systems are being used for recruitment in a company or for arresting a suspect, it is necessary to question how they have been developed and what ideological and social biases have been applied.[50] Paglen created an application, ImageNet Roulette, which allowed any user to see how the AI was tagged. Media interest in this project led the ImageNet team to remove half of the images in the "person" category.

Memo Akten's video Optimising for Beauty (2017) eloquently anticipates the case of ImageNet by using images from another dataset, CelebFaces Attributes (CelebA), which has over 200,000 images of famous faces, tagged with forty different attributes (haircut, nose shape, facial hair, expression, among others). The work shows an artificial neural network in the process of generating new faces from images of real people that capture the attention of the media. The set, in itself, is already limited to a relatively small group of people, who in many cases tend to a certain homogeneity (as may be the case of models or actors). Akten consciously accentuates the similarity of the faces by adjusting the parameters of the artificial neural network by means of maximum similarity estimation (MLE), which leads to the generation of portraits with the elements most likely to be found in the observed dataset. The artificial neural network thus learns to create an idealised, perfect and uniform form of beauty in which individual features (skin or hair colour, age, sex) are dissolved into a single face that progressively ceases to look human. This work shows us how the images of the most visible (and admired) people in our society tend to mark certain stereotypes, which automatic processing accentuates until reaching a disturbing uniformity. The manipulation carried out by the artist subtly reveals that the training sets on which the IAs are based are nothing but fictions, highly biased extracts from reality that reveal their artificiality with only a few modifications to the parameters with which they are analysed.

The way in which the data sets that feed into AI are constructed is worrisome, but not only because of the creation or perpetuation of biases, but because these determine predictions that lead to automated actions. Researchers Ajay Agrawal, Joshua Gans, and Avi Goldfarb indicate that the development of AI can lead to changes in prediction-based business models: for example, Amazon may become so accurate in its predictive capabilities that it may be more beneficial to send customers the products it knows they will want to buy rather than wait for them to buy them.[51] Applied to people, this could lead to a dystopia like the one imagined by the writer Philip K. Dick in The Minority Report (1956). However, AI also presents an opportunity to combat prejudice and create data sets that teach programs to apply a more egalitarian worldview. Gender discrimination is a fact of life in all areas of society, requiring both a change in attitudes at the individual level, ethical codes and wage policies at the corporate level, and laws and representativeness at the government level. But we live in an information society, and therefore technological products also play a part in this discrimination, although sometimes in a very subtle way: for example, a 2017 study showed that a job offer in the technology sector advertised on Facebook was shown to more men than

https://www.nytimes.com/2012/11/20/science/for-web-images-creating-new-technology-to-seek-and-find.html
50 Kate Crawford y Trevor Paglen, "Excavating AI. The Politics of Images in Machine Learning Training Sets," Exacvating AI, 2019. https://www.excavating.ai
51 Agrawal, Gans and Goldfarb, Prediction Machines, p.21.

women because the algorithm judged it to be more effective in relation to the cost per click.[52] Algorithms reflect the values, perspectives and biases of those who create them, and these tend to be mostly men. According to the White Paper on Women in Technology, there is a clear gender gap in the sector: in Europe, only 30% of the approximately 7 million people working in the information and communication technology sector are women.[53] In the field of AI, the percentage of women drops to 13.5%, while in the EU, female AI specialists make up only 5% of the total number of researchers.[54] In addition, the risk of job loss due to automation is twice as high for women as for men, with 26 million female jobs in 30 OECD countries at high risk of being displaced by technology over the next two decades.[55] Unless more women are engaged in designing the systems that determine many aspects of our daily lives, it will be difficult to prevent algorithms from perpetuating gender bias or to ensure that situations that may affect women in the majority are taken into account. An example can be found in the When&Where app created by five teenagers from a high school in Móstoles (Madrid), which allows a user to monitor any journey she makes. The app detects any detour or abrupt stop and sends a message asking the user if she is ok: in case she does not answer or answers in a negative way, the app automatically contacts an emergency phone number. This app is the result of the creativity of young women and their attention to a problem that affects them personally, while also denoting the serious shortcomings of a society in which a young woman must put on a digital beacon to feel safe. Technology can therefore provide practical solutions and also raise awareness.

Caroline Sinders' Feminist Data Set project (2017-present) aims to do this by proposing a collective action to collect what she calls "a feminist data set. This collection includes artwork, essays, articles, interviews and books that deal with feminism, explore feminism or provide a feminist vision. The collection of all this information is intended to introduce the operation of automatic learning systems and to create the equivalent of a training package, through which artificial intelligence algorithms could be developed that adopt a feminist perspective on the world. Sinders carries out this task through a series of workshops in which she invites participants to engage in a brainstorming session in which general concepts (such as inequality, femininity, gender) and specific topics or titles (such as Virginia Woolf's essay A Room of One's Own or Donna Haraway's Cyborg Manifesto) are written down on sticky notes. These notes are placed on the wall to be organized into categories, and in this way initiate a conversation about feminist references and each person's perception of feminism, how it is defined and what it brings to it. This serves, first of all, to question the very definition of a feminist data set and also so that each community can present ideas and examples from its local experience, thus contributing to enriching the diversity of the set (it should be taken into account that even a feminist data set can be biased if it only collects the points of view of a small group of people from the same cultural and socioeconomic environment). To this collection of qualitative data, Sinders adds a phase dedicated to understanding how artificial intelligence training sets work. A designer and researcher specializing in machine learning and a digital anthropologist, the artist has first-hand experience having worked at Intel, IBM Watson and the Wikimedia Foundation. At the latter, she worked on the study of patterns of harassment on the net, where she had to collect numerous ethnographic data about misogynist users and

52 Agrawal, Gans and Goldfarb, Prediction Machines, p.185.
53 Sara Mateos Sillero and Clara Gómez Hernández, Libro Blanco de las mujeres en el ámbito tecnológico. Madrid: Secretaría de Estado para el Avance Digital/Ministerio de Economía y Empresa, 2019, p.13. http://www.mineco.gob.es/stfls/mineco/ministerio/ficheros/libreria/LibroBlancoFINAL.pdf
54 Sara Mateos Sillero and Clara Gómez Hernández, White Paper, p.118.
55 Sara Mateos Sillero and Clara Gómez Hernández, White Paper, p.17.

extreme right-wing groups, which led her to raise the need to create a data set that offered a diametrically opposed view.[56] But beyond countering sexism on the Internet, the feminist data set offers the possibility of better understanding the biases in the use of data with which artificial intelligence programs are "educated", and also of integrating the ideas of feminism into the heart of technological development. In this sense, the Sinders project is aligned with the principles of the manifesto Xenofeminism: A Policy for Alienation, developed in 2015 by the Laboria Cuboniks collective, which highlights the importance of combating gender inequality in the technology sector:

> Gender inequality still characterizes the fields in which our technologies are conceived, built and legislated . Such injustice demands structural, machinicolous and ideological reform. [...] Since there are a range of gender-specific challenges related to life in the digital age - from sexual harassment in social networks, to doxxing, maintaining privacy, or protecting images online - the situation requires a feminism comfortable with computer media and the use of new technologies.[57]

Sinders' project aligns feminist thinking with digital skills and facilitates awareness of (and responsibility for) systems that are often left to companies and only now beginning to be questioned. At the same time, it helps to better understand what feminism is all about, a movement that is often misunderstood and met with violent rejection. As the writer and activist Bell Hooks (Gloria Jean Watkins) points out: "many people believe that feminism consists solely and exclusively of women who want to be the same as men, and the vast majority of these people believe that feminism is anti-men. This lack of understanding of feminist politics reflects what most people learn about feminism through the patriarchal mass media.[58] The biased view of feminism is also evident in the technology sector through the discomfort it generates if only as a concept. According to recent research by The Guardian, Apple programmed its virtual assistant Siri to avoid making pronouncements about feminism or even using this word in its answers.[59] The company, therefore, considers it right that users give orders to a female voice, but does not think it is convenient that this assistant is able to talk about women's rights. Caroline Sinders could counter this with her next project, a feminist chatbot that would have trained with all the data she has collected so far.

How to dialogue with an AI?

Most of the artificial intelligence programs currently in use perform relatively invisible tasks, such as the analysis of large amounts of data or the automatic recognition of elements in an image. Direct interaction between the user and the AI takes the form of a virtual assistant (such as Siri or Alexa), which is limited to obeying orders or simulating a brief conversation. The conception of the machine as a faithful servant translates into a friendly female voice that responds diligently and, as one of the participants in Lauren McCarthy's LAUREN project

56 Katharine Schwab, "This Designer Is Fighting Back Against Bad Data–With Feminism," FastCompany, 16 de abril de 2018. https://www.fastcompany.com/90168266/the-designer-fighting-back-against-bad-data-with-feminism
57 Laboria Cuboniks, Xenofeminism: A Policy for Alienation Translation by Giancarlo Morales Sandoval, 2015. http://www.laboriacuboniks.net/es/
58 bell hooks, feminism is for everyone. Translated by Beatriz Esteban Agustí, Lina Tatiana Lozano Ruiz, Mayra Sofía Moreno, Maira Puertas Romo and Sara Vega González. Madrid: Traficantes de Sueños, 2017, p.21.
59 Alex Hern, "Apple made Siri deflect questions on feminism, leaked papers reveal," The Guardian, 6 de septiembre de 2019. https://www.theguardian.com/technology/2019/sep/06/apple-rewrote-siri-to-deflect-questions-about-feminism

states, establishes a relationship in which everything revolves around the user. This type of interaction is, to some extent, positive, in that it can facilitate collaboration between the individual and the AI, increasing the capabilities of the former (as computers already do) and helping the latter to learn and process data better. But it also means forgetting that, in the end, we are talking to a software and not to a person. This confusion (which some AI projects, such as Google Duplex, try to use to their advantage) can lead to an emotional involvement like the one in the film Her (Spike Jonze, 2013) or dehumanize relationships with people who provide a support service. Interaction with AI therefore raises the notion of the "other," as that which is not oneself or which occupies a subordinate position.

These themes underlie Lynn Hershman Leeson's work on artificial intelligence. In her film Teknolust (2002), she explores human relationships through clones that have to interact with the real world, and in doing so modify it at both the biological and technological levels. Expanding on the idea of this film, Agent Ruby (2002) is a female character reduced to a face that users can interact with on a website. Ruby can answer questions in writing and learns from these interactions, as well as from the information she obtains on the Internet. Ruby is not just a chatbot (an intelligent program that interacts through text), but it has the ability to express emotions in a rudimentary way through its face. Ruby's identity is built as it interacts with users and even develops moods based on the information it collects and whether or not it likes the user. According to the artist, her intention was to create an artificial intelligence character capable of interacting with the general public, something she achieved a decade before Siri or Alexa existed.[60] As a continuation of Agent Ruby, the artist created DiNA (2004), a female character played by the actress Tilda Swinton (star of Teknolust) who engages in a dialogue with viewers by means of a voice recognition system and, like Agent Ruby, can learn from these interactions and the information she obtains on the web. DiNA is an advance over Ruby, both at the code level, since it has a more sophisticated program, and at the presence level, since in this case viewers are in front of a film of the actress, whose head occupies an entire wall. This intelligent agent also takes on a specific role in her first incarnation, which is to run for the post of "Telepresident" and conduct a virtual election campaign. This prompts people who interact with DiNA to ask her questions regarding the current socio-political reality, and raises the possibility of the government of an AI, a topic that Pinar Yoldas will later address in Kitty AI. By obtaining the information to elaborate her answers from the Internet, DiNA is able to address current issues, and thus appears more real, intelligent and able to adapt to conversations, despite the limitations currently presented by voice recognition systems and the AI program itself. This installation offers us the opportunity to dialogue with an artificial intelligence that, unlike Apple's and Amazon's virtual assistants, is not a disembodied female voice that lends itself to obeying our commands or resolving our doubts, but one that has a presence of its own and invites us to consider it as a being with which to seek understanding. As researcher Eana Kim points out, "the interaction [between AI and the public] becomes an exploration of existence through communication between two different states of existence, one from the real world and one from a fictional world.[61]

The interaction between human and AI can also lead to another situation where the machine stops being an other and becomes a double. Currently, a part of every user's identity exists on the Internet as a set of data, social network profiles, uploaded images and videos, published messages and a long etcetera. This "digital self" will be managed in the future by an artificial

60 Eana Kim, Embodiments of Autonomous Entitites: Lynn Hershman Leeson's Artificially Intelligent Robots, Agent Ruby and DiNA. Tesis de Master, The Institute of Fine Arts, Universidad de Nueva York, 2018.
61 Eana Kim, Embodiments of Autonomous Entities.

intelligence program that will create a detailed model of each person. Companies, either to offer services, products or job offers, will contact this model, which will be in charge of filtering the messages that reach the real person.[62] The artificial intelligence tools will also make predictions, based on the data available, about what we will need, what news we will want to read or what products will interest us. The paradox that can occur at this point is that if the person follows the recommendations of the AI, it reinforces the accuracy of their predictions, so that these, as a self-fulfilling prophecy, end up determining the actions of the user. The interactive installation Neuro Mirror (2017) by Christa Sommerer and Laurent Mignonneau confronts the viewer with a system that predicts his actions and leads him to question whether he has control over them or not. The artists are inspired by scientific research on specular neurons, which are those that are activated when an action is performed and seen to be performed by another individual. These neurons participate in the processes that take place in the brain by establishing relationships with other people, imitating them, empathizing, as well as differentiating between the "self" and the "other. They also play a major role in intuition, specifically in predicting the behavior of others in the future. In Neuro Mirror, Sommerer and Mignonneau use artificial neural networks to create a piece capable of analyzing the actions of viewers and showing a visualization of what they might do next. When a person stands in front of the piece, which consists of three screens, he or she sees their image reproduced on the central monitor, while the one on the left shows their activity in the immediate past. The third monitor shows a character predicting the gestures the person will make in the future. In this way, a dialogue is established between spectator and machine, as the person feels obliged to imitate the character who predicts his actions or act in a different way, being equally carried away by the dictates of the system. The artists consider that automatic learning will probably never reach the level of complexity and adaptability of the human brain, but they demonstrate with this piece that a system that uses the relatively limited tools available today can already significantly affect an individual's perception of himself or herself.

Humans are not necessary

Since the beginning of research in artificial intelligence, the possibility of a "strong AI" has been raised, which would be capable of reasoning by itself (and not simulate human reasoning just enough to pass the Turing test).[63] If this is possible, as some authors argue, this type of AI would soon be able to develop to the point of surpassing human intelligence and give rise to what has been called artificial super-intelligence. As early as 1965, Irving John Good, a colleague of Alan Turing, predicted that a machine more intelligent than any human would be able to create other machines even more intelligent, and thus lead to an "explosion of intelligence," with the machines themselves, and no longer humans, in charge of all subsequent technological development. According to Good, "the first ultra-intelligent machine is the last invention humanity would need to make, provided the machine is docile enough to tell us how to keep it under control.[64] Mathematician and science fiction author Vernor Vinge called this situation "the Singularity" and predicted, in 1993, that artificial superintelligence would arrive in 2023, which would mean the end of the human era.[65] The pessimistic view of the future of AI has been echoed by other prominent voices, such as the astrophysicist Stephen

62 Pedro Domingos, The Master Algorithm, p.269.
63 Russell y Norvig, Artificial Intelligence. A Modern Approach, p.1020.
64 Boden, AI. Its Nature and Future, p.148.
65 Russell y Norvig, Artificial Intelligence. A Modern Approach, p.1038.

Hawking, who stated in 2014 that artificial intelligence could be humanity's worst mistake.[66] In the face of these ominous omens, the extraordinary optimism of Ray Kurzweil stands out. He hopes that by 2045, super-intelligent AI, combined with advances in biotechnology and nanotechnology, will end war, disease, poverty and even death.[67] Among those who believe in the Singularity there are clear differences of opinion about what it will mean for the human race. It is certainly a matter of wild speculation, and it may seem absurd to consider these scenarios. But it is also true that the technological advances that are part of everyday life today only appeared in science fiction stories a century ago.

Felix Luque's work confronts viewers with their expectations and fears about technology in narratives that combine reality and fiction, possible futures and dystopias. Nihil Ex Nihilo (2011), tells the story of SN W8931 CGX66E, the computer of Juliet, a secretary working in a large corporation. Infected by malicious software, it soon becomes part of a botnet, a network of machines controlled by a cracker that uses them for cybercriminal activities. Through electronic alteration, the computer acquires a form of consciousness, a primitive artificial intelligence. Confused by this situation, it tries to communicate with other machines and free them from their subjection to human users. Seeking to establish contact, it responds to spam messages received by its user in order to spread its ideas among the network of machines. The work consists of several independent but linked pieces that complete the story in different formats. In this version they are included: The Dialogue, a set of eight alphanumeric displays that show in real time the messages exchanged between a text generation program and the computers that send spam, while they are read by a synthetic voice; and The Transformation, an audiovisual file of the moment in which SN W8931 CGX66E mutates and becomes a semi-neuronal structure, in an animation that leads us to think about the concept of "explosion of intelligence" proposed by Good, a sudden transformation with no turning back. Felix Luque's installation raises the fear of a "rebellion of the machines" that is often associated with the development of artificial intelligence and is a recurring theme in science fiction stories. The dialogues that shape the main piece are produced between machines, without the participation or knowledge of humans, giving an inescapable presence to a type of communication that is constantly taking place but which we ignore. In line with Luque's work, which opposes the vision of technology as a mere tool to the perception of the machine as a mysterious and sometimes incomprehensible entity, the work reminds us that the development of intelligence also entails madness and shows us a system that works beyond our control. Nihil Ex Nihilo also touches on a recurring theme among AI researchers, which is the debate over whether a machine can develop a consciousness of its own, as well as the need for a strong artificial intelligence to have a consciousness in order to be able to reason as a human being does. In the case of SN W8931 CGX66E, consciousness does not make it a superintelligence, but a paranoid and obsessive machine. Like the intelligent robots Adam and Eve in Ian McEwan's novel Machines Like Me (2019), conscience is not a gift but a condemnation that some machines are unable to bear.

Regardless of the possibility (real or not) that a machine acquires consciousness of its own, a particularly disturbing aspect of Felix Luque's piece is the dialogue that takes place between machines, without any mediation by a human being. Certainly, the infected computers do not know what they are talking about, in terms of the meaning of the words, but they are carrying out an effective communication. As Claude Shannon states in his Mathematical

66 Boden, AI. Its Nature and Future, p.148.
67 Boden, AI. Its Nature and Future, p.149.

Theory of Communication,[68] the semantic aspect is irrelevant for the system, the important thing is that a message passes from a sender to a receiver. One computer generates an output that is processed by another and sends a response, so a “dialogue” is established between the machines as soon as they exchange data. This level of communication is strange and intelligible, as well as invisible, to most people, who usually conceive of the interaction between human and machine but not between two machines. Yet the devices we use every day are constantly exchanging data with other machines according to established protocols. This type of communication is usually external to the user, and does not require his direct intervention. Thus, the idea that the machines are “talking” to each other can generate unease, either because (as Luque’s piece proposes) we imagine that they might be conspiring against us or simply because they leave us out of the conversation. Jake Elwes proposes in Closed Loop (2017) this kind of exchange. The artist creates a conversation between two artificial neural networks: one analyses and describes, in text form, the images supplied to it; the other generates images in response to the words written by the first neural network. The networks have been trained with a data set of 4.1 million images with descriptions and another data set with 14.2 million photographs. According to Elwes, he has not provided any visual or text content to the piece, which functions on its own thanks to the feedback loop that is established between the two AIs. Similar to Luque’s piece, this installation makes people mere spectators of a closed dialogue in which they cannot intervene, only observe trying to understand what leads one neural network to generate the images and how the other responds with descriptions. The system is self-sufficient and again contrasts our anthropocentric perception of technology, in which we are either the dominant or the dominated, the existence of an exchange that is alien to us. Trying to decipher the narrative that develops between the images and the texts generated, we are finally the ones who strive to understand the reasoning of the machine. The piece also plays with the fascination that artificial intelligence arouses insofar as it produces surprisingly human results from a process that mostly takes place in a “black box,” a system whose exact functioning is very difficult to understand.

We have seen the exploration of the other in works of art and artificial intelligence that start from a human (if fictitious) presence in the interactive installations of Lynn Hershman Leeson and Christa Sommerer and Laurent Mignonneau, then move on to machines that express themselves through English text and images, in the works of Felix Luque and Jake Elwes. We conclude with a form of communication that is literally Martian. In nimiia cétïi (2018), Jenna Sutela uses machine learning to generate a new language, which she gives written and spoken form to. To do so, she relies on the “Martian language” created (or communicated) by the French medium Hélène Smith, who in the late 19th century claimed to be able to establish contact with a civilisation from Mars and expressed herself orally and in writing in the language of this planet, in sessions of sleepwalking trance and automatic writing. The artist has compiled Smith’s drawings and has given voice to the transcriptions of the phrases pronounced by the medium included by Professor of Psychology Théodore Flournoy in his book Des Indes à la Planète Mars (1900). In addition to this material, there is an analysis of the movements of the bacterium Bacillus subtilis nattō, which, according to a laboratory study, can survive in extreme conditions, such as those that occur on the surface of the red planet. Sutela, on the one hand, has provided an artificial neural network with recordings of the medium’s phrases and on the other hand, a sequence of the movements of a sample of B. subtilis observed through a microscope. These movements have given rise to patterns that the program has

68 Claude E. Shannon, “A Mathematical Theory of Communication,” The Bell System Technical Journal, vol.27, p.379-423, 623-656, Julio, octubre, 1948. http://www.essrl.wustl.edu/~jao/itrg/shannon.pdf

reinterpreted as traces of signs, which in turn take Smith's automatic writing as a reference. All these elements make up the audiovisual in which the artist brings together the sounds and the graphics generated by the machine with images of the bacteria, visualisations of the patterns identified by the AI and a simulated landscape that evokes the planet Mars. The whole is obviously cryptic, far removed from any human reference as it uses an indecipherable language (despite being presented in written and oral form) and the images of a microscopic and extraterrestrial nature. The piece, according to the artist, makes the computer a medium, which interprets messages from entities with which we cannot communicate and provides them to us automatically, without the mediation of reasoning, as supposedly happened to Smith during one of his trances. Jenna Sutela describes Martian language messages as "gloss poetry," referring to the phenomenon of glossolalia, which is the ability to speak in a language unknown to the speaker. This phenomenon has been associated with the belief that people can be possessed by a divine being or a spirit (becoming mediums). This, in turn, has been related to the art of poetry, as in Plato's Ion, the dialogue in which Socrates describes the rhapsody as a mere vehicle of divine inspiration. In the same way, the program of artificial intelligence that has developed a version of the Martian language, does so without being aware of it or knowing what the meaning of what it has created is. The audience's experience with this work, which may be aggressively strange, ultimately allows us to understand how an AI sees the world and how it creates an artifice that we interpret as intelligence.

Pau Waelder
October, 2019

Machine translation from Spanish with minor adjustments by the author