A STEP-BY-STEP SEMIOTIC UNDERSTANDING OF LLMS AND CHATBOTS THROUGH INTERDISCIPLINARY DIALOGUE

Mattia Thibault
Tampere University
mattia.thibault@tuni.fi

Daniele Allois
Independent Researcher
daniele.allois@gmail.com

Abstract

Artificial Intelligence (AI) is enjoying a period of "summer". It is an intensely discussed and researched topic which fascinates researchers across fields with questions about both its development and possible impacts. Semiotics too has immediately engaged with the topic, focusing especially on the ways in which different forms of Generative AI can be understood through existing frameworks, and how the processes of textual production it encompasses can be studied and analysed.

Digital Age in Semiotics & Communication, Vol. VIII, 2025, 19–49

https://doi.org/10.33919/dasc.25.8.2

Gen AI, however, is a complex and fast evolving technology, which is very difficult to study without a sufficient understanding of the technical side. Perceptrons, Artificial Neural Networks, Transformer Networks, Large Language Models are all terms we hear often in discussions about AI. However, they often refer to technologies we barely understand.

This paper is born from a collaboration between authors with expertise in semiotics and machine learning. Our objective is to reconstruct the evolution of Large Language Models in particular, and to engage the different stages from a semiotic perspective. This in-depth engagement with the technological side is especially useful to go beyond the strong temptation to anthropomorphize the technology and instead reposition it as a tool that supports human textual creation. Based on our analysis we then propose several key concepts (an AI semiosphere, Generative AI as a mode of sign production and the concept of *Intentio Machinae*) as ways to map and conceptualise the role of LLMs in semiotic processes.

Keywords: Generative Artificial Intelligence, Cybersemiotics, Intentio machinae, Large Language Models, Chatbots

1. Introduction

Artificial Intelligence (AI) is enjoying a period of summer. AI technologies periodically go through periods of great enthusiasm and others of deep disillusionment – sometimes called summers and winters of AI (Haugeland 1949). New technical solutions bring along renewed hope for impactful applications of the technology. However, they often reach an impasse when the technology does not deliver what the public hoped for. At the present time, Large Language Models (LLMs) and Generative AI have greatly advanced the ability of machines to perform tasks which until recently were thought to be exclusively human, especially in terms of understanding and creating that which, in semiotic terms, we would call different forms of signs and texts.

From written and oral texts, to images, videos and even 3D settings Generative AI appears to be able to engage with a series of semiotic activities in a way which, at times, seems indistinguishable from a human being. While machines creating texts with the input of humans is not necessarily a new thing – cameras do something similar, for example – the quantity of data on which they are trained and the minimal input required by humans by Generative AI in order to create new texts is indeed unprecedented.

This poses some interesting questions for semioticians, as we are seemingly facing a machine not only able to generate meaning, but also to make interpretative choices when confronted with semiotic material. But is this what is really happening? What are the semiotic entails of Generative AI?

This paper is the result of a dialogue between two different areas of expertise: semiotics and machine learning, embodied by its two authors. Our aim is to look at the technical steps that brought us to the current state of AI development and engage each from a semiotic perspective. This should both offer a clear and simple explanation of the computational operations behind Generative AI, and some semiotic insights on how to look at them from the perspectives of meaning-making and interpretation.

Since there are many different forms of AI, we decided here to focus on LLMs, as their ability to use natural languages and to act as conversational agents (or Chatbots) is particularly interesting from a semiotic perspective.

2. Background: Semiotics and AI

While a relatively new phenomenon, Generative AI has been able to excite the curiosity of many semioticians. The discourses around its development range from the most techno-optimistic to the most catastrophic predictions. Nevertheless, all seem to suggest that we are facing something unprecedented: a claim which can also sound like a challenge. Are we now facing a different kind of semiotic actor, or is this just a statistically-powered – although quite convincing – "trick"? Semiotics, of course, has something to say on the topic.

As early as 2020, in reaction to the development of Natural Language Processing (NLP), Duncker (2020), was suggesting caution, claiming that, while chatbots may be perceived as communicating with human users in natural languages, one should not forget that this is possible only through a second-order language model, which governs the way the machine reacts to human input – a reaction that is still limited to a specific set of pre-programmed executable skills. In other words, Dunker suggests that chatbots do not participate in communicative processes, and that the signs it produces are "unsigns" which the machine does not interpret nor engage with as such (Duncker 2020: 97).

In the same year, Sanna (2020) already began to wonder about the role of semiotics, as a discipline possibly developing AI models. He proposed the idea of "semiotics–driven machine learning". This seems to resonate with a feeling that many have in the discipline: the idea that, as Generative AI, NPL, and LLMs work with natural languages they belong, in a way, to semiotics.

Semioticians, after all, have been discussing the role of cybernetics in semiotics for decades, including Eco's explorations on the possibility of computers telling stories (Eco 1994), and Lotman's hopes for an integration of semiotic systems and informatic ones capable of learning from human semiosis (Semenenko 2012).

This perception, that semiotics cannot only say something important about Gen AI, but that semiotics *has* to do it, might have played a role in the flourishing of studies dedicated to the topic in the last couple of years.

One of the first, ambitious attempts to coordinate these efforts is the volume edited by Santangelo and Leone in 2023. It contains 12 chapters from different semioticians – the majority of which belong to Turin's semiotic circle – which engage Generative AI from multiple perspectives (Santangelo & Leone 2023). Within the issue there are two attempts to clearly outline the dimensions of AI which concern semiotics. One is Ferraro's (2023), which offers a series of reflections, highlighting the mythology of AI (the utopian and ideological depictions of the technology). He examines similarities and differences with human intelligence, the question of creativity, the question of classification, and finally the degree in which the machine can *grasp* meaning. The other is Leone (2023), which proposes a similar list, including: how AI simulates intelligence, the production of creative content, and the ideological assumptions around AI in our culture. Engaging with this last point, especially focusing on Adversarial Networks, Leone argues that the idea of "fakeness" is central to understanding the semiotic aspects of Gen AI.

The issue also contains certain works which focus on the discursive context surrounding Gen AI such as Santangelo (2023) looking at representations of AI in science fiction, and Stano (2023) engaging with the relations between AI and human and artificial bodies.

Several other papers, instead, engage with the *products* of Gen AI, i.e. the textualities which emerge from these machines, and their possible sociocultural effects. Gramigna (2023) does it for deep fakes, Voto (2023) for climate data, Soro (2023) for motivational messages in Fitness applications.

A few more papers engage with certain applications of Artificial Intelligence, such as sentiment analysis (Galofaro & Toffano 2023, and Ferreira Leite da Silva 2023) and chatbots (Dall'Acqua & Bellentani 2023) and what semiotic theory has to offer to their refinement.

Finally, two papers approach AI by engaging with the historical context of their development – both technological and ideological. The first is Giuliana (2023), who highlights the fundamental role of games in the development of AI – since they are spaces for both imitating human behaviour, but also opposing it. If most Gen AI models we encounter today are meant to be our helpers, in games it is often the opposite, the AI developed for games participate in the obstacles we have to overcome while playing. Giuliana offers what he calls a “performative definition of intelligence”, highlighting how often the definition of intelligence is based on the competence of an actor to do something, to perform a specific task. Giuliana’s chapter reminds us also how, while today’s discourses tend to conflate AI with Generative AI, there are a plethora of technologies which are currently and concurrently defined as “Artificial Intelligence”. This is more due to the perception we have of them than because actual similarities in what they do.

The second is Volli (2023), who provides a useful reconstruction on the evolution of the idea of the possibility to make natural languages as precise and infallible as mathematics: from the search for a perfect language to Turing writings and modern ideals of AI. Volli claims what we call AI is clearly not an analogue to human intelligence – referring also to Searle famous argument about the “Chinese Room”[1] – but that does not make it a less interesting object of study for semiotics. If we abandon the idea of an AI which imitates the semantics of human intelligence – and therefore the mythological quality of the pursuit for an objectivization of language – we can instead focus on what Volli considers most interesting in machine learning. This is the ability to categorise and classify signs and information – a question which has been of great interest for semioticians. Volli describes machine learning as actuating a process of inference which is retro-actively acting on its own parameters. It is, hence, an implementation of the abductive mechanism which is at the base of the semiotic-cognitive functioning. In other words, machine learning is based on the ability of the machine to formulate interpretative hypotheses about the data upon which it is trained. Then after evaluation, to change the parameters on the base of which its hypotheses are formulated.

More recently, Semiotica has dedicated a special issue to “Aspects of AI semiotics: enunciation, agency, and creativity”, which showcased the continued interest in the discipline for the topic and contains some important contributions. The issue programmatically aims to go beyond the semiotic

---

[1] Searle’s example famously imagines a person in a closed room receiving messages from under the door written in Chinese ideograms. The person does not know how to read Chinese, but has access to a program which informs them on how to select the right characters to send back under the door in response. People outside the room, seeing that the messages they are getting back meaningfully respond to those they sent, might assume – wrongly – that the person in the room is fluent in Chinese. The same, then, can be said about an artificial machine which might be capable of responding in a way interpreted as meaningful by human interactors, but that is not necessarily able to interpret what it is doing.

properties of AI generation, and instead focus on analysis of different socio-cultural practices involving generative AI.

A number of papers focus especially on the relation of AI with archival images and texts. Dondero (2025) engages with *enunciative praxis* and the relation between the sedimentation of stereotypical forms in visual culture (contained in the archived images on which the AI models are trained) and the possibility of actualisation – or creation of new images – of visual transformer models. D'Armenio (2025) works in a similar area, proposing the idea of "archival perception" as the particular way on which AI perceives and learns. The author also underlines how Generative AI works as a *co-enunciating machine*, highlighting the necessity of human intervention in textual generation and the role which AI plays in it.

Two papers also focus on how this hybridity and collaboration between human and machine is not necessarily new. Paolucci (2025) challenges the illusion of human cognitive exceptionalism by arguing that language-endowed machines reveal and underline our preexisting machinic hybridity. According to the author humans are "natural-born cyborg," whose cognition has always been scaffolded by external systems. Colas-Blaise (2025), on the other hand, questions the ability of the machine to co-enunciate, and rather proposes that "machinic enunciation," is an enunciative sequence which sees human agency in the initial and final phases, with numerous machinic instances in between.

Other papers focus on interpretative questions: Compagno (2025) argues that the notions of "truth" and "falsehood" are not well suited to describe the textual production of AI models. Basso Fossali (2025) speculates whether AI could lead to "augmented semiospheres" where a third order of cultural remediation is offered by chatbots. Valle (2025), instead, focuses on the epistemological shift in coding, in which instead of attempting to reconstruct grammars, machine learning attempts to reconstruct texts by imitation. Finally, Montanari (2025), while focusing on the socio-semiotic implications of these technologies, offers some interesting insights on the role of spatiality, statistics and distributional semantics in the meaning-making processes of generative models.

Despite the wealth of insights, much of the current research on Gen AI in semiotics seems to be conducted in isolation from the technical expertise that is behind its development. However, there are two recent series of studies based on bringing together a varied pool of expertise.

The first is brought forward by Everardo Reyes who has dedicated two articles (Reyes 2024; Reyes 2025) in which he looks at generative models (and not only at Generative AI) through the joint lenses of semiotics and computer science. These works – among the few which also adopt a Peircean perspective on the topic – offer a wide overview both at the many different types of generative media, and at the different operations they can perform.

The second series has been conducted by the trio Enzo D'Armenio, Adrien Deliège and Maria Giulia Dondero. These studies are part of two projects funded by the *Direction du Projet De Recherches (PDR) F.R.S.-FNRS*: one focused on semiotic and Computer-Assisted approaches to large collections of images; and the second on Generative AI and image production. These projects have enabled Adrien Deliège, mathematician and computer scientist, to offer technical insights to semiotic research.

In their works, they focus especially on Midjourney and DALL-E, and therefore on Generative AI models focused on image generation. D'Armenio et al. (2024a) first proposed a key argument in their work: the fact that Generative Models are *co-enunciators* with their human prompters of the texts they generate. In particular, they conduct a series of experiments exploring different forms of prompting and analysing the results which they yield.

In D'Armenio et al. (2024b), the authors propose an evaluation method, in order to assess the effectiveness of text-to-image generation models. This paper presents, in our opinion, two key points of interest. First, it acknowledges directly and explicitly the importance of (intersemiotic) translation when it comes to Visual Generation models. The second key point of interest is methodological. Semiotics is traditionally concerned with the analysis of texts which, in the vast majority of cases, precede the analysis. Studies on Gen AI, however, are able to ask the machine to generate texts at any time, and then analyse them – and the process which led to their creation. This is a significant methodological change: one that has not yet been thoroughly examined by semioticians, and who until now seem to have mostly dealt with it in an intuitive, almost naive way. This study, while also proceeding intuitively, presents an interesting dialogical way of proceeding, testing back and forth the results of image generation. It makes use of prompts which are systematically prepared to test the limits of the models studied. In Deliège et al. (2025) they put this model and methodology to test, focusing on the artistic categories of "classical" and "baroque".

Finally, D'Armenio et al. (2025b) engage more globally on the functioning of image generator modelling, especially looking at their results as a particular kind of intersemiotic translation. While the paper later engages with the control of human prompters on the composition of such images, we particularly appreciate that this work engages explicitly with the technical operation behind image generation. Without engaging with the background of machine translation, the authors still engage with the technical steps which separate a prompt from the final image – something that we believe central if we want to properly address the semiotic functioning of these machines.

To our knowledge, this is a rare endeavour. There is a wide breadth of studies on semiotics and AI (to which this summary does justice only partially) which deal with many different aspects from their inner workings to their societal implications. However, there still seems to be a clear gap when it comes to examining in detail the technical side of Machine Learning – and to analyse their semiotic entails. In this paper, then, we will initiate an interdisciplinary dialogue to look at the different steps in the evolution of machine learning and examine them from a semiotic lens. In the next paragraph we will then engage with the fundamentals of machine learning, and more specifically on the bases for the development of LLMs and chatbots. In section 4 of this paper, then, we will propose some semiotic implications which will build on our step-by-step analysis and use several concepts from semiotic theory to explain, or better understand, the functioning of LLMs.

3. Engaging with the fundamentals of Machine Learning

3.1. Machines which can learn like living things

The idea of a computer capable of learning, and therefore of the possibility of an "Artificial Intelligence", was formulated in the mid 20th century. Alan Turing himself stated that: "What we want is a machine that can learn from experience," and that "the possibility of letting the machine alter its own instructions provides the mechanism for this." (Turing 1948). Those remarks were at the beginning of a long journey of technical development in the field of "Machine Learning" which has currently led to the development of different implementations of Generative Artificial Intelligence, such as Large Language Models (LLMs).

How can a machine capable of learning be created? In order to answer this question computer scientists started looking at an existing, complex, system capable of learning: the brain. More specifically, Machine Learning was strongly influenced by a book published in 1949 by Donald Hebb called "The Organisation of Behavior" (1949). Hebb looked at the working of animal neural networks, in order to understand how their processes could be replicated artificially. Hebb in particular observes how the physical connections of the nervous systems of different individual organisms differ even in the same species. The portions of nervous systems that are involved in learning and recognition are largely random at the birth of an individual. However, such systems are characterised by a certain degree of plasticity: the neurons change based on neural activity. Exposure to similar stimuli forms and strengthens pathways that lead to the same responding cells. At the same time, dissimilar stimuli will develop connections to different sets of responding cells. Positive or negative reinforcement can then be used to facilitate – or discourage – the formation of these connections.

"Similarity", then, is defined by Hebb as being a tendency of similar stimuli to activate the same sets of cells. This is not determined by specific formal requirements of the class of stimuli, but rather by the physical organisation of the perceiving system and the interaction with the organism's environment. Hebbs concludes that: "the structure of the system, as well as the ecology of the stimulus environment, will affect, and will largely determine, the classes of "things" into which the perceptual world is divided".

These statements are not surprising to semioticians: the idea that *differance* is at the base of meaning making is as old as the discipline (Saussure 1916), and biosemiotics has often engaged with the relation between nervous systems and semiosis (Kull & Favareau 2022). However, what is relevant in our argument is that the structural functioning of the nervous systems – the creation and reinforcement, over time, of specific neural pathways – became a key inspiration in the development of Machine Learning.

3.2. The perceptron

Inspired by Hebbs's work, computer scientists (see Rosenblatt 1958) started to look for a way to replicate the natural ability of organic nervous systems to learn and adapt. The first step in that direction was the creation of the *perceptron*: an algorithm capable of "learning" based on information fed to it. This was a significant development compared to traditional forms of coding, in which programmers have to explicitly write all the rules that the machine will follow.

The core idea behind the perceptron is that of a function which can be trained by engaging with a curated set of inputs so that it will later become capable of conducting a simple operation precisely when confronted with new inputs. The elements making a perceptron are the following[2]:

- The inputs. These are the different parameters which the algorithm will take into consideration. They are numerical values.

---

[2] There is, in fact, another component that is the *bias*. This is an addendum to the network that facilitates its process from a mathematical standpoint. This addendum is used to make the perceptron produce an output even when the inputs are negative. Since it does not impact the structural working of the perceptron, and it makes it more complex to understand, we decided to engage with it only in this footnote.

- The weights. These are numerical indicators which indicate the relative value assigned to each input.
- The output. This is a simple binary opposition which will consist either of 1 (activation of the perceptron) or 0 (inactivation).

As shown in Fig. 1, the perceptron works in the following way: the Net Input Function collects the different inputs received by a perceptron, each multiplied on base of their weight, and sums them in a final value. Then, the last part of the algorithm, called the Activation Function will, depending on the summation value, either output a 1 (activation) or 0 (inactivation).

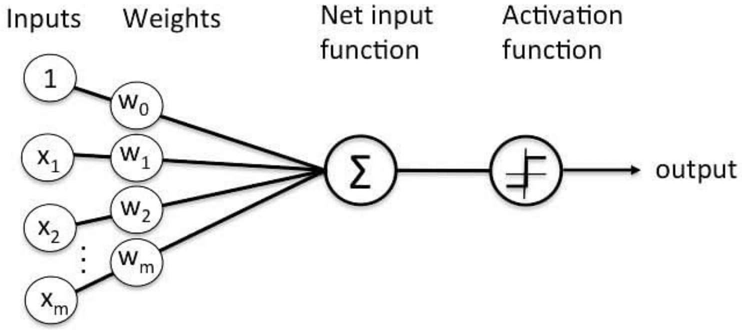

Figure 1: Diagram of a perceptron (2009–2025 – Simplilearn Solutions)

The structure of the perceptron as presently described does not yet allow machine learning, but rather explains what kind of operation the perceptron engages in: it evaluates a certain input based on predetermined parameters.

For example, a perceptron could be used to evaluate the cost of real estate. A simple perceptron accepting only one kind of input, could process the price of a house and, based on the weights assigned by its creators, activate if the price is over a certain threshold and flag it as "expensive" – or not activate and flag it as "inexpensive". This is not a particularly impressive feature per se, since it basically only checks if the price of the house is above or below a certain price. Things become more interesting, however, with a perceptron that accepts multi-feature inputs – that is, different kinds of data. In this case the inputs could include the price, but also the square footage, the number of bathrooms, the distance from the city centre, and any other relevant numerical data. The final output would be the same. The perceptron either activates if the house is deemed expensive or cheap, but the complexity of the estimation is increased. Thus the difficulties for its creators to assign the correct weights to the different parameters are also increased. How does the number of bathrooms trade off with the distance from the city centre? Machine Learning comes into play exactly to avoid humans making such calculations.

Before engaging with it, however, let us look at the working of the perceptron from a semiotic standpoint. In semiotic terms, we can look at the functioning of the perceptron as an actant (Greimas 1970), or as part of a specific Actor Network which participates in a clear narrative programme (Latour 1993). In this case, we can say that the perceptron operates a *sanction*. The perceptron evaluates the input which it receives and, based on a set of criteria (the weights that have been assigned to the different possible inputs), generates a result. It either activates, or it does not. It is a simple form of sanction which gives a binary result, but which is based on an automated form of *discrimination*.

Metaphorically, we could understand it as a sort of *sieve*. The sieve is a machine "programmed" by humans (which designed the caliber of its holes) to make a simple distinction. Objects of the right dimension (for example, seeds) pass through, while the rest do not. In a similar way, the perceptron organises data into two groups by activating or not. In either case, the act of interpretation – or discrimination (or classification, as it's called in technical terms) – is entrusted to the machine, but predetermined by the humans who created the machine.

The perceptron, hence, does not work dissimilarly from the light switch described by Eco in a Theory of Semiotics (1976). It is a more complex version of the switch, but it still follows a strict law of necessity. To the same input the same perceptron will always respond in the same way.

3.2. The Perceptron learning rule

The key feature of perceptrons, however, is their ability to imitate organic neural pathways and *learn*. Perceptrons go through the *perceptron learning rule*: a step-by-step process which allows the perceptron to modify or set its own weights. This process allows a perceptron to be *trained*, and hence starts to remove interpretation (i.e. determining the weights of all the different inputs) from the hands of its human creators.

The perceptron learning rule works as follows:

- The programmers create a curated set of inputs. For each input they know the desired output (activation or inactivation) but they ignore what exact weights would generate them.
- The perceptron is a series of cycles (named "epochs") in which it makes predictions about its own activation when facing the curated set of inputs.

- At each cycle, the perceptron confronts its own prediction with the predetermined correct course of action.
- If the prediction was incorrect, the perceptron will alter its own weights by adjusting them a bit in the direction of the right answer.
- The perceptron continues to pass through a series of epochs, until its predictions become accurate enough that it can be exposed to new, not pre-classified input, and be able to correctly assess the right responses.
- The perceptron learning rule, then, allows the perceptron to set up and evaluate the best set of weights, in order to offer accurate outputs, without the need for its programmers to define those weights themselves.

Using the real estate example again, a multi-feature input perceptron could be trained on a set of examples of market valued houses and their respective amenities, square footage, distance from the city centre, etc. This set of examples would still have a human-made evaluation of their price (based, for example, on the fact that the properties found a buyer or not). While making predictions and confronting them with the correct responses, the perceptron will learn how the different features of a house determine its value (that is, how to weigh the different inputs it receives). After the training, it will sequentially be able to estimate if the price of a new house on the market is too high.

The narrative programme of the trained perceptron remains largely the same. It operates a sanction which leads to a binary discrimination. However, the training shifts our focus to a higher narrative level, in which its ability to offer a *sanction* is also a *performance*. This performance itself is then evaluated and is object of its own sanction. We are then facing two narrative layers:

In one, the perceptron operates a sanction towards a preliminary performance which is engraved in the data it receives. The sanction results in a state of activation or deactivation.

In a second layer, this sanction becomes its own narrative programme. Here the perceptron mobilises its own competences (the weights upon which it bases it discrimination) and operates a performance (the discrimination itself, so the activation or inactivation). It is then another piece of code which sanctions such performance, determining if the perceptron was successful or not, and determining the consequences of failure – which result in a modification of its competences.

Repetition of this exercise will lead to the generation of a perceptron with competences evolved enough to ensure a maximum chance of success in its performance.

At the same time, we also face a multiplication of narrative programmes within the same system. Here we have three main steps of machine learning: the "generation" (in this case of a prediction), the "evaluation" (of the correctness of the prediction), and the "modification" (of the weights).

With the beginning of machine learning, things start to become more interesting from an interpretative perspective as well. The basic work of the perceptron was purely deductive. Given a rule (the predetermined weights) and a cause (the input) the perceptron would infallibly deduce the effect (the output). However, with the learning rule, the perceptron becomes able to discriminate on the basis of weights which its human programmers were not able to predetermine. The training, then, is based on *induction* (in Peircean terms). The perceptron as a series of causes (the inputs) and of effects (the predetermined correct responses), and through a series of trial and errors, is able to induct the rule (the new weights). After that, when confronted with new, not-curated inputs it will be able to deduce the outputs correctly - thus endowing the perceptron with some form of "archival perception" (D'Armenio 2025) and initiating, although at an embryo stage, the mode of enunciative praxis described in Dondero (2025).

The inferential ability of the perceptron, then, is greatly enhanced. Its operations are still determined by its weights (and therefore again follow a rule of necessity). The iterative process which determines its final weights is still a mathematical model (also deterministic). At the same time, the process of determining the weights is now one step removed from human intervention, and therefore initiates the path that makes it increasingly difficult for the creators of AI models to understand how exactly they operate.

3.3. Neurons and Artificial Neural Networks

The perceptron, while relatively simple, was the base for developing more complex algorithms. In order to tackle more complex tasks, or offer more nuanced solutions, the development of machine learning proceeded in two directions, often in parallel:

1. Developing different kinds of neurons, that is, other algorithms, in addition to the perceptron;
2. Stacking multiple perceptrons (and later multiple neurons) in Artificial Neural Networks (ANN).

New kinds of neurons have been developed to replace or stand alongside perceptrons. Since the output of perceptrons is limited to a binary output, then if taken singularly, they can only solve *linear* problems, and their answers lack nuances: only "yes" or "no". New *neurons* were then developed with different activation functions, meaning that their output could be, for example, a number within a certain range. A neuron using *linear regression,* for example, would be able to use very similar weights to those of a perceptron to determine the price of a real estate property and not only its being expensive or not. Semiotically, the operation of the neuron is quite similar to the operation of the perceptron, but, once more, more complex and hence harder to interpret or predict by humans.

On the other hand, the second line of development was that of creating ANN based on networks of perceptrons. The very first networks were quite simple and "shallow" (having only one layer) capable of simple operations, such as distinguishing a capital "T" from a capital "L" in a small image. Over the decades, also thanks to the development of machines with higher and higher computational capabilities, ANNs began to develop an increasing number of hidden layers (Fig. 2, displayed in yellow). These hidden layers work as perceptrons themselves. Each hidden layer has a weight and an activation function assigned and fed into another hidden layer making the ANN metaphorically "deep".

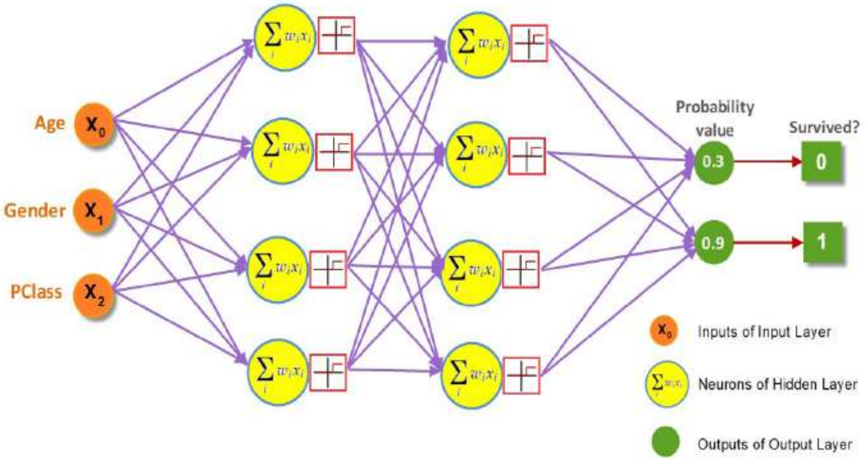

Figure 2: A Perceptron-based Artificial Neural Network
(Copyright © 2025 Devskol)

Perceptrons, here, become the *neurons* in the neural network, that is, they become part of an interconnected system that works collaboratively. Early ANN made use of a series of perceptrons, and of their learning rule, to become able to engage in more complex tasks. For example, where a perceptron could be able to estimate if a house is expensive based on a series of parameters, a perceptron-based ANN could determine the price of the house, model interactions between features (for example that the number of bedrooms matters more when square footage is high) or model conditional patterns (difference in prices in different cities).

The two lines of development converge. Most ANNs soon started to feature multiple layers composed of different kinds of neurons. The organisation of neurons in ANN, hence, further multiplies the co-occurring actions of discrimination, each providing more nuanced results. The continuous development of new and more sophisticated hardware and therefore computer power gave the possibility to stack more hidden layers together, reaching the hundreds, making the ANNs exponentially better at their tasks. This sparked a revolution in machine learning called the "deep learning revolution" making possible more and more technologies and the rise of current Artificial Intelligence (LeCun et al. 2015).

The differentiation of neurons and the increase in layers also required a new way for the model to learn, as the perception learning rule became too simplistic. This was provided by *backpropagation*, a staple of deep learning. Simply put, backpropagation operates in a similar manner. It first makes a "forward pass", in which the input goes through all the layers and the ANN produces its prediction which is confronted with the right result (which, again, needs to be known beforehand). The system, now, uses a "loss function", in order to evaluate *how* wrong the prediction is. The second step is the backward pass, in which the different layers are examined one by one – from the output layer back to the input layer. At each layer a *partial derivative* of the loss is calculated. In other words, it is calculated how the loss (i.e. the margin of error in the prediction) would have changed, if the weight had been different. Finally, the weights are updated. As in the perceptron learning rule, here too the process needs to be repeated many times with different inputs before the system can be used with inputs whose correct evaluation has not been determined beforehand. The update of the weights is also more nuanced, like the predictions it has to evaluate.

A complex ANN can, for example, distinguish between pictures of cats and dogs. For the purposes of backpropagation, a curated set of pictures labelled as either "cat" or "dog" is needed. The model is then trained on RGB (Red-Geen-Blue) input of the image files. The image files have values for Red, Green and Blue for each pixel in the picture. The RGB values of each picture are calculated and inputted, in the network (usually the picture is downscaled in quality because of the large amount of computation that it takes to perform such a task). The ANN, through its epochs, assigns weights to a different shape which it recognises on a mathematical grid based on a set of properties in the picture. The different layers operate sequentially, so that the first layer of the network recognises lines, the second level legs or faces, and so on until the entire picture is estimated to be of a "cat" or a "dog". Understanding the weights needed to accomplish such a task is far above human capacity, and cannot be done without machine learning. The researchers, given the desired objective of the model, can still fine tune it looking at its outputs without having to interpret what is going on "inside" – and therefore contributing to the "black box" nature of many ANN.

If we look at these developments from the perspective of their actor networks, we see a multiplication in the order of hundreds of actants with similar functions which work in parallel or sequentially, perform operations and predictions and sanction each other. The ANN learning structure is semiotically similar to its predecessor. We still have operations which generate predictions, others which evaluate them, and others which evaluate and update the weights. What changes, from this perspective, is the complexity of the operations that the ANN can engage with.

From an interpretative perspective, however, the complexity of these systems seems to make them able to perform *abductions* (see also Volli 2023). In the real estate example, we can see that ANN is faced with an effect (the final price of a property), and a set of rules (the weights that they have assigned to different elements). In its evaluation, it has to infer what might be the cause (which of those elements have influenced the final price?). This cannot be performed as a form of induction, because there is no single fixed rule behind real estate prices. ANN has to make some hypotheses during its backward pass, try different weights for different elements. This is also what makes these kinds of operation more fallible – as abduction is the less secure kind of inference.

Something similar can be said about distinguishing between pictures of cats and dogs. ANN has an effect (the picture), a set of rules (weights determining how a cat is supposed to look), and then it is faced with a picture and needs to determine an effect (was a cat or a dog in front of the camera?).

While the need for hypothesising might decrease with training, thanks to more and more accurate weights (and therefore a move towards induction), the complexity of these tasks makes abduction more relevant – and risky – in the operations of ANNs.

3.4. Transformer Networks and Large Language Models

The last key improvement in Artificial Intelligence with which we will engage in this paper is the creation of Transformer Networks. These are networks of an ever larger scale – they contain several ANNs – and their key property is that of *attention.* This concept was introduced in what is now considered a pillar in the bibliography on transformer networks: a paper entitled "Attention is all you need" by Vaswani et al. (2017). Attention is used to calculate the value of the weights, or how much "importance" we ascribe to a certain input according to their context, i.e. according to other tokens that are present in the same input.

Attention is extremely important in Large Language Models (LLMs), based on the stacking of multiple transformers which are trained on extremely large datasets composed of texts in the same natural language[3]. LLMs work by trying statistically to find the most probable word after a sentence by having learned all the content of a large dataset and they are believed to "learn the statistics of language" to reproduce the effect of today's chatbots.

In the case of LLMs, attention gives the network the capability to understand the context of every word given into the input.

Attention works by vectorialising every "token" (e.g. every word, or part of a word in a text), according to three vectors:

- "Query" – What the token is "looking for" (e.g. a noun "looks for" a verb, a transitive verb for an object, a pronoun for a noun, etc.).
- "Key" – What the token is "offering" (what other token would be completed by it, the object offering itself to a transitive verb).
- "Value" – the weight of the relationship (how strong can be the connection between two tokens).

These three vectors are then used to calculate the attention. The query and the key are multiplied, passed through an algorithm (softmax) and then multiplied by the value. This mathematical operation reconstructs the relations between two tokens, and the statistical probability that they refer to each other. When attention is calculated across all the tokens in a text, the LLM is able to reconstruct the relations for all the tokens.

---

[3] This was only possible thanks to the introduction of parallelisability in operations and scalability with large datasets, taking advantage of faster GPUs. Parallelisation and scalability are two properties which express how multiple tokens can be calculated at once to give faster results and how the system can be scaled, meaning how if it's given more GPUs power it can always improve results and calculations.

For example, if an LLM encounters the sentence "the cat purrs because it is happy", it is able to reconstruct that the query of the word "it" is the word "cat". The token "it" will score low for "happy" and high for "cat" as cat is the only noun in the sentence, so the model will determine that it refers to that token. This is done simultaneously with all the words of the input so that the LLM determines the proper context using Queries and Keys, and having a set of Values as a result.

Interestingly, each token is represented in LLMs through vectors – i.e. lengths and directions. Instead of the 2 or 3 dimensions typical of geometry, LLMs feature hundreds. The complex path which they draw, and the direction they point to, contain both their semantical affinities (e.g. feminine words will point towards similar directions) and their relations with other words. The semantic and lexical systems of a language as reconstructed in a LLM then occupy a multidimensional space which is too difficult for humans to imagine. However, it seems to share some affinities with the rhizomatic encyclopedic model of Eco (1994) and Lotman's semiosphere (1990). It is an eminently relational semiotic space, in which the different tokens are meaningful only thanks to the relations with each other and the "attention" uniting them. It is, in fact, based on *distributional semantics*, i.e. it grounds meaning on the spatialised relations of different sememes (see Montanari 2025).

This is particularly interesting because, if we look at how machine learning applies to LLMs, the complexity of the system they have to reconstruct is that of a whole semiotic system (that of a natural language) and a series of other systems of different scales (metalanguages, specialised languages like lyrics or poetry, etc.). The simple weights determined by the perceptron learning rule have multiplied and reached a level of complexity which is difficult to grasp. However, while an LLM may be able to simulate grammatical analysis, the digital metalanguage it reconstructs is a mathematical one – possibly reminiscent of Hjelmslev's glossematics.

From the perspective of the actor network, the complexity of the systems renders it pointless to try to reconstruct the abundance of actants engaged in performing tasks and sanctioning each other. What appears, instead, is an increased sense of actorialisation of the model itself, which starts to appear like a unique semiotic unity, capable of communicating and even understanding – an effect of meaning that reaches even stronger effectiveness with chatbots (see below).

On the interpretative side, the complexity of LLMs makes them even more reliant on abduction. The scale of the semiotic system involved, the gargantuan quantity of data upon which they are trained, the wide variety of tasks and conversations in which LLMs can be involved makes it impossible to have a perfectly correct model. Hallucinations, therefore, are inevitable. The system will invent facts, make up statements, get stuck in strange ideas. While this is something which the development of the models minimises, it is not something, we believe, which will be possible to completely root out – because, as Eco reminded us (1976), a semiotic system is something which can be used to lie. LLMs are created in order to be able to communicate, in order to be semiotic machines. As such, they require a level of complexity high enough to allow for mistakes, for misunderstandings, and for nonsense. As the saying goes, “it’s not a bug, it’s a feature”.

3.5. Chatbots

The form in which most users encounter LLMs, today, are Chatbots. These are LLMs trained to simulate a conversation with the user and designed to make them appealing and as user friendly as possible. Usually this is done by giving the LLM some direction in the format of its writing and expression, by adapting filters and security checks to make it harder to be exploited for nefarious uses. This is done by using Reinforcement Learning by Human Feedback (RLHF) in which a set of engineers ranks the LLM answers from good to bad. In this way, the LLM will learn to behave in the desired way: generally being an eager helper.

Chatbots also feature a memory which they use to form a coherent conversation, to base their next answers on what has already been discussed as part of the same interaction – or even multiple interactions with the same user.

Chatbots have an astounding effectiveness to simulate human conversations and to engage in meaning-making activities – while obviously unaware of doing so. The combination of a strong fascination for this technology from users and a strong industrial push to incorporate them in many different applications (search engines, instant messaging, etc.) has brought artificial intelligence into the forefront of global interest – and at the centre of many debates.

While the interpretative and learning capabilities of Chatbots are essentially the same as of more generic LLMs, it is the effect of meaning of their actorialisation – as mentioned above – which make Chatbots particularly interesting. The complexity of their inner workings – which is mysterious for users and, to a certain degree, for their own creators – coupled with the ability to mimic human conversation give rise to a potent illusion of conversing with a sentient being. Corporate attempts to reinforce this illusion – by making them friendly and polite – are also part of the reasons for which it is very easy to anthropomorphise them.

This, in turn, is key to position them – at least discursively – in a very clear actantial role: that of the helper (rather different than the opponent role often reserved to AI in games, see Giuliana 2023). Chatbots present themselves as helpful assistants, as collaborators, eager to help human subjects in a series of possible performances, and then offering key competences in terms of skills (cognitive, semiotic, epistemic) and knowledge (constantly updated from the Internet).

Of course, this is not the only role which Chatbots – and other uses of LLMs – can play. They are, in fact, often subjects, to which activities traditionally needing human actors are increasingly delegated. This in turn requires humans to develop new competences[4], and can cause different forms of discrimination (Latour 1993). This ability to replace humans – which Latour teaches us is certainly not something new – in different narrative programmes has, nonetheless, also characterised them as *anti-subjects*. Therefore several actions have been initiated and discourses of resistance against their adoption – often rooted on their ethical concerns, including intellectual property theft and excessive energy use.

3.6. Emergence

In this paper we have focused on the development of LLMs in particular, and we have not delved into image generation which is also an extremely popular and sophisticated use of ANNs and Transformer networks. We judged that for properly engaging with these other models we would need more space than we have in the current paper. There is, however, one last concept related to the development of transformers that we believe important to mention: that of *emergence*.

This concept was introduced by Leggs & Hutter (2017) to refer to the fact that the stacking of networks gave rise to the emergence of unexpected bigger reasoning properties. In a key paper about emergence in LLMs, Wei et al. (2022) state that: “a range of abilities (...) thus far [have] only been observed when evaluated on a sufficiently large language model. Hence, their emergence cannot be predicted by simply extrapolating performance on smaller-scale models.” (Wei et al. 2022: 6). In other words, after a certain threshold LLMs are capable of performing different tasks which were impossible for smaller models, and not explicitly included in

---

[4] As any teacher trying to discern how much of their students’ essays has been written by Chat GPT is painfully aware of.

their pre-training. Examples of such tasks include being able to perform arithmetic, unscrambling words, or creating conceptual mappings. Another relevant example is that of translation. LLMs models which were not explicitly trained on language pairs still demonstrated an ability for multilingual translation (Li et al. 2024).

Emergence is often considered puzzling, as it is difficult to predict what other abilities might emerge from further increases of the size of the networks. It is also one of the reasons that they are often seen as black boxes: the functioning of the models escapes their creators who often cannot predict nor reconstruct why the models do what they do.

The situation is further complicated by the fact that these models are created by a large number of people, often building on models created by others, so that nobody can easily track the overall functioning of the model as a result. The alienation of the programmers from their own creation feeds into catastrophic discourses about "AI rebellion". While these are probably also discursive strategies which aim to protect corporate intellectual property, as claimed by Leone (2021) – the fact that the inner workings of LLMs is rather mysterious is very real – and several attempts to increase the transparency of AI by interpreting the models predictions or generate attention maps are underway (Barredo Arrieta et al. 2019)

From a semiotic perspective, beyond the discursive and mythological (Barthes 1957 & Ferraro 2023) implications of the back box, it is interesting to see how extensive training on natural languages seems to enable a series of abilities which resemble reasoning and logic. This seems to be an indirect confirmation of Juri Lotman's claim that natural languages are "primary modelling systems" (Lotman 1977) which play a key role in organising around them all other modelling systems and supporting human cognition. The idea that language shapes the way in which we think is not new (see the Sapir-Whorf hypothesis or Wittgenstein 1921), nevertheless the fact that natural languages – and not computer code – seem to be the base for some form of artificial cognition is rather interesting.

4. Semiotic implications

4.1. AI's Encyclopaedia and Semiosphere

As we have seen in the previous paragraphs, since its inception the key to Machine Learning has been the idea of "training". In supervised training, perceptrons, neurons, and ANN all adjust their weights by going through a curated dataset, in order to adjust their weights until they are able to correctly predict the outcomes desired by their creators. Even at later stages of development, where, for example, LLMs are trained on large not-curated datasets scraped from the Web, they learn to imitate human-made texts by engaging with them. They try to guess the next word and then check if their prediction is correct until they become able to generate coherent and meaningful texts.

Machine Learning, therefore, is dependent on a pre-existent set of texts. These texts are already accompanied by their interpretants – i.e. by other signs which can be used to infer their meaning. This clearly includes the curated lists of inputs which programmers feed to perceptrons and other neurons (like the list of pictures labelled "cat" and "dog" of our example above), but also existing lists of paired signs, such as images and their descriptions (such as those of the datasets Oxford-102, CUB-200, or COCO). We argue that even the texts used to train LLMs, even if not curated, already contain their interpretants. LLMs infer semantic directions (not the meaning of the words, but their vectorialisations in a semantic field) by the words which accompany them. They do not need to understand what a word "means", in order to reconstruct the relations with other words and ultimately become able to answer questions and hold conversations. The different words in a text, therefore, function as interpretants of each other. They are signs which support the inference of other signs.

The importance of the training dataset, therefore, is capital. It is a well-known fact that AI models reflect the biases present in their training datasets, often creating racist, sexist, and other problematic content (Caliskan et al. 2017). The datasets, therefore, work as digital analogues of the *encyclopaedia* described by Umberto Eco (1994). According to Eco the *encyclopaedia* is grounded on the repertory of texts and interactions that an individual – or a community – has been exposed to and therefore encompasses all their semiotic competences and previous knowledge. Individual encyclopaedias are determined by a series of factors (socio-cultural positions, personal tastes and experiences, family and so on) which influence their construction. The encyclopaedia of an AI model, instead, is determined during its training, but it is equally – if not more – impactful on the ability of the model to create signs.

The training of AI models also implies that every text or sign produced by Generative AI has an intertextual (Genette 1982) relation with some – if not all – the texts upon which it has been trained. More precisely, AI generated texts are *hypertexts* (i.e., texts created using other texts as a starting point, Ibid.) since their creation is directly influenced by the texts which form its training dataset (its *hypotexts*). While identifying which training texts influenced a specific AI-production is difficult (due to the back box that encompasses much of the models' operations), recent technical developments have started to be able to reconstruct them at least partially (Templeton 2024).

However, when it comes to LLMs, the model does something more than just imitating the texts upon which it has been trained. As we mentioned above, it reconstructs a semiotic system, by vectorialising each token (which correspond more or less to *sememes*) and integrating them into one semantic space. More precisely, the LLM reconstructs several semiotic systems, as it is also trained on programming languages and specialised languages (poetry, lyrics, metalanguages etc). While the LLMs do not reconstruct a specific hierarchy between these systems, the probabilistic, statistical and learned patterns which it uses imitate the relations between different semiotic systems as well. In other words, LLMs have their own *semiosphere,* a digitised and internalised reflection of the semiosphere to which the texts in its dataset belong.

However, when it comes to Chatbots, the additional instructions which regulate their responses add a new layer of hierarchies specific to that conversational agent. The temporary enamourment of Chat GPT for words like "delve" and "tapestry" (meant to make it sound more scholarly), or the momentaneous inability of Gemini AI to generate pictures of Caucasian people (because LLM tend to add racialised terms in the image prompting, see Shamin 2024) by overriding some of the training of the LLM, manipulate its textual production.

4.2. LLMs as a modes of Sign production

Generative AI can be understood as participating in an enunciation activity initiated by its users. Humans input a prompt, without which the model cannot initiate any form of textual production. While discussing AI image generation, D'Armenio et al (2024a) describe this property as "machinic co-enunciation".

A similar point can be made for LLMs. While the way in which AI processes human prompts is different (most image generators vectorialise words but do not "understand" prompts in the way LLMs do), the overall process is similar: a human author creates a prompt, and the AI model responds to it by creating a text.

We agree with D'Armenio et al, (2024a) that this is indeed a form of co-enunciation, or co-authorship. To go deeper into this activity of co-creation, we decided to look at it as a *mode of sign production* as those described by Eco (1976).

When discussing different modes of sign production, Eco focuses on several technical means as part of the process. These include the physical medium, the articulatory techniques, and the transmission and reproduction tools involved. The articulatory techniques are specifically interesting for our argument, since they include the mechanical or bodily actions which produce signs. The variety of articulatory techniques spans from the way in which we articulate our vocal apparatus when speaking, to complex digital operations which result in the creation of signs (such as, for example, photogrammetry). These means do not only allow for sign creation, but are often indispensable means of cognition (Paolucci 2025).

Chatbots (and other forms of Generative AI) can therefore be understood as a part of this process. They are complex technical means which support the activities of sign production of the users when creating their prompts. They draw from a net of hypotexts which participate in its training.

4.3. Intentio machinae

Chatbots seem quite unusual as a technical means for sign production. They appear to determine a large part of the final sign – often larger than their human co-enunciators. While all technical means do influence the texts they produce or reproduce – hence McLuan's famous statement "the medium is the message" (1964) – it is rare that much of the final sign is entrusted to them. Rare, but not unseen. In the arts, for example, randomness generators have been used to determine the final form of a text – as in John Cage 1951 composition "Music of Changes". In painting, Pollock's *Drip Painting* period (1947–50) similarly makes use of the unpredictability of the dripping paint as a key component of the artistic work. Fortune telling, in a similar way, makes use of various forms of randomness generators – from tea leaves to playing cards – which strongly constrict the results of the reading (Aphek & Tobin 1990). In video games, procedural generation allows the creation of environments largely created by the game programme. They are based on some initial instructions given by the game designers (see, for example, *No Man's Sky,* a 2016 game by *Hello Games*).

The co-enunciation, or the mode of sign production which is based on the use of Generative AI seems to be the next step in this direction. Not a radically new semiotic phenomenon, but an incremental step towards a larger share of authorship delegated to a machinic actor.

This specific mode of sign co-production raises, at the same time, analytical and interpretative questions. To the three intentios outlined by Eco, the intentio auctoris, lectoris and operis, we might need to add a fourth one: Intentio Machinae. If the intentio auctoris acknowledges what the author is trying to express in a text, the intentio lectoris what the reader adds in order to interpret it and the intentio operis the "objective" structure and content of a text, the intentio machinae accounts for what element of the text is determined by its technical articulatory techniques.

The intentio machinae, then, is determined by the encyclopaedia and the semiosphere of a certain conversational agent, and its intertextual interaction with the prompt created by the human author. It is difficult to disentangle what of the final text, or sign, is determined by the prompt and what is determined by the LLM – but the fact that the same prompt gives different results shows a variety of responses from the LLM which makes its contribution semiotically meaningful.

5. Discussion and Conclusions

In this paper we have reconstructed, step by step, the main milestones of the development of machine learning, focusing in particular on Large Language Models. For each step, we have provided some semiotic insights, focusing on three key aspects: narrative programmes, inferences, and intelligibility.

We have seen that the narrative programmes enacted by the different actants which compose neurons and models become more complex over time. Starting with an untrained perceptron operating a simple sanction, to learning neurons which perform prediction and are then sanctioned for it, to more and more intricate networks of actors which operate hundreds of complex operations. We have also argued how this actantial complexity leads to identifying the entire model as a single actor, ignoring its idiosyncrasies in favour of an anthropomorphisation.

We have examined how the inferring ability of machines has become more complex with their developments: from perceptrons able only to make use of very simple deductions, to the perceptron learning rule introducing inductive capabilities, and to ANN which have become capable of abductions. We have also pointed out that the ability to perform abductions – which are the most creative but less reliable form of inference – is probably the reason for AI hallucinations, inventions and false claims. The same properties which enable networks to learn by making hypotheses of possible (semantic) rules are those that make it less precise.

Finally, in terms of intelligibility, we have seen how the increasing complexity of the networks have progressively removed the task of deciding the weights – and therefore the parameters of interpretation – from the network creators. This eventually leads to the emergence of unexpected properties and the necessity of new approaches to shed light in the black boxes that are the current models.

We have then focused on a few semiotic implications of our analysis of the working of LLMs. First, we have claimed that the datasets used to train LLMs function as an encyclopaedia, determining the ability of the system to produce texts. We also claimed that LLMs create multidimensional vectors for each token in a shared semantic space, giving rise to a sort of semiosphere, which reflects that engraved in the texts used by the training, but also manipulated by its creators when adapted to a Chatbot.

Second, we have looked at the activity of prompting Chatbots, in order to create text together as a specific mode of sign co-production. The technical means for sign creation have a large impact in the final creation – but that the latter still needs a prompt made by a user to be generated.

Third, we claimed that this mode of sign co-production, while not completely new in nature, has reached such a scale that we need to rethink our interpretative lenses. In particular, we have introduced the concept of "Intentio Machinae" as a way of acknowledging the participation (but not sole authorship!) of generative models in the generation of signs and texts.

Due to the long and complex history of machine learning, as well as the variety of models and networks existing today, we are not here proposing a complete semiotic analysis of all the aspects of generative AI. Notably, we have not engaged with visual transformers, despite the attention they were able to elicit in the semiotic community. This was decided mostly to limit the current article to an acceptable length – but we are considering a follow up paper doing something similar for these networks as well.

We have also mostly focused on the working of the models themselves, and not on their perception, the discourses that surround them, or their sociosemiotic implications. As we mentioned in the background section, there are a number of papers which have focused on those directions.

We believe, however, that this paper makes an important contribution to the scholarly work in semiotics by providing an accessible, step-by-step account of the key elements of LLMs. We believe that a multidisciplinary approach to these topics, which joins technical and semiotic knowledge, is key for analysis grounded on the actual workings of the technology. We hope this paper, beyond offering some theoretical insights on the technology, can also offer a basic technical understanding to other semioticians – and humanities scholars – who want to contribute to the research on Large Langage Models and Generative Artificial Intelligence in general.

Acknowledgements

This research was supported by y NEXR – Next Extended Reality (funded by NextGenerationEU and Business Finland, co-research project number 5703/31/2023) and Funded by the European Union (ERC, InterReal, grant agreement No. 101163472). Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Council. Neither the European Union nor the granting authority can be held responsible for them.

References

Aphek, E., Y. Tobin. 1989.e *Semiotics of Fortune-Telling.* Amsterdam: John Benjamins Publishing.

Barredo Arrieta, A., E. Serral-Asensio, S. Tabik, R. Benjamins, A. Barbado, S. Garcia, S. Gil-Lopez, D. Molina, J. M. Benítez and F. Herrera. 2020. “Explainable Artificial Intelligence (XAI): Concepts, Taxonomies, Opportunities and Challenges Toward Responsible AI”. *Information Fusion*, Vol. 58, 82–115. Available at: https://doi.org/10.1016/j.inffus.2019.12.012.

Barthes, R. 1957. *Mythologies*. Paris: Éditions du Seuil.

Basso Fossali, P. 2025. “Rationalités correctives et intelligence artificielle assistée: les doubles contraintes des humanités numériques”. *Semiotica*, Vol. 2025, No. 262, 71–109.

Caliskan, A., J. J. Bryson and A. Narayanan. 2017. “Semantics Derived Automatically from Language Corpora Contain Human-like Biases”. *Science,* Vol. 356, No. 6334, 183–186.

Colas-Blaise, M. 2025. “La machine crée, mais énonce-t-elle? Le computationnel et le digital mis en débat”. *Semiotica*, Vol. 2025, No. 262, 147–187.

Compagno, D. 2025. “La notion de vérité à l’épreuve de l’intelligence artificielle”. *Semiotica*, Vol. 2025, No. 262, 25–49.

Dall’Acqua, A., F. Bellentani. 2023. How to Build a Chatbot. A Semiotic and Linguistic Approach. In Santangelo, A., M. Leone (eds.). *Semiotica e intelligenza artificiale* (Coll. *I saggi di Lexia*, n. 48). Roma: Aracne, 149–169.

D’Armenio, E., A. Deliège and M. G. Dondero. 2024a. *Semiotics of Machinic Co-Enunciation: About Generative Models (Midjourney and DALL·E). Signata: Annales des Sémiotiques*, 15, Article 15. Presses universitaires de Liège. Available at: https://doi.org/10.4000/127x4.

D’Armenio, E., A. Deliège and M. G. Dondero. 2024b. A Semiotic Methodology for Assessing the Compositional Effectiveness of Generative Text-to-Image Models (Midjourney and DALLE). In Del Bue, A., P. Grommelt, L. Weiss, F.-J. Pfreundt, and J. Keuper (eds.). *Proceedings of Computer Vision – ECCV 2024 Workshops, LNCS 15644*. Cham: Springer. Available at: https://doi.org/10.1007/978-3-031-92089-9_8.

D’Armenio, E. 2025a. “Les IA génératives visuelles entre perception d’archives et circuits de composition”. *Semiotica*, Vol. 2025, No. 262, 213–257.

D’Armenio, E., M. G. Dondero, A. Deliège and A. B. D. Sarti. 2025b. “For a Semiotic Approach to Generative Image AI: On Compositional Criteria” [as “Big Dada”]. *Semiotic Review*, 9 (Images issue), 1–59. Available at: https://doi.org/10.71743/ee5nrx33.

Deliege, A., M. G. Dondero and E. D’Armenio. 2025. “Revisiting Wölfflin in the Age of AI: A Study of Classical and Baroque Composition in Generative Models”. *Journal of Imaging*, Vol. 11, No. 5, 128. Available at: https://doi.org/10.3390/jimaging11050128.

Dondero, M. G., J. A. Aldama and M. Leone. 2025. “Aspects of AI semiotics: enunciation, agency, and creativity”. *Semiotica*, Vol. 2025, No. 262, 1–3.

Dondero, M. G. 2025. “Semiotics of artificial intelligence: enunciative praxis in image analysis and generation”. *Semiotica*, Vol. 2025, No. 262, 111–146.

Duncker, D. 2020. “Chatting with Chatbots: Sign Making in Text-Based Human–computer Interaction”. *Sign Systems Studies*, Vol. 48, No. 1, 79–100. Available at: https://doi.org/10.12697/SSS.2020.48.1.05.

Eco, U. 1976. *A Theory of Semiotics*. Bloomington: Indiana University Press.

Eco, U. 1994. *The Role of the Reader: Explorations in the Semiotics of Texts*. Bloomington: Indiana University Press.

Eco, U. 1994. *Sette passeggiate nei boschi narrativi*. Milano: Bompiani.

Ferraro, G. 2023. Mito e realtà dell’IA: uno sguardo semiotico. In Santangelo, A., M. Leone (eds.). *Semiotica e intelligenza artificiale* (Coll. *I saggi di Lexia*, n. 48). Roma: Aracne, 65–90.

Galofaro, F., Z. Toffano. 2023. Verso una semantica quantistica? Applicazioni al discorso religioso. In Santangelo, A., M. Leone (eds.). *Semiotica e intelligenza artificiale* (Coll. *I saggi di Lexia*, n. 48). Roma: Aracne, 107–130.

Genette, G. 1982. *Palimpsestes: la littérature au second degré*. Paris: Éditions du Seuil.

Giuliana, G. T. 2023. L'intelligenza delle I.A. come effetto di senso semio-narrativo nei giochi digitali. Una rivoluzione semiosofica. In Santangelo, A., M. Leone (eds.). *Semiotica e intelligenza artificiale* (Coll. *I saggi di Lexia*, n. 48). Roma: Aracne, 171–194.

Gramigna, R. 2023. Some Remarks on Fakes and Deepfakes. A New Epistemic Challenge. In Santangelo, A., M. Leone (eds.). *Semiotica e intelligenza artificiale* (Coll. *I saggi di Lexia*, n. 48). Roma: Aracne, 45–64.

Greimas, A. J. 1970. *Du sens II: Seuil de la signification*. Paris: Seuil.

Hebb, D. O. 1949. *The Organization of Behavior: A Neuropsychological Theory*. New York, NY: John Wiley & Sons.

Haugeland, J. 1989. *Artificial Intelligence: The Very Idea*. Cambridge, MA: The MIT Press. Available at: https://doi.org/10.7551/mitpress/1170.001.0001.

Kull, K., D. Favareau. 2022. Neurosemiotics: A Brief History of Its Development and Key Concerns. In *The Routledge Handbook of Semiosis and the Brain*, 13–29.

Latour, B. 1993. Where Are the Missing Masses? The Sociology of a Few Mundane Artifacts. In Bijker, W. E., J. Law (eds.). *Shaping Technology/Building Society: Studies in Sociotechnical Change*. Cambridge, MA: MIT Press.

LeCun, Y., Y. Bengio and G. Hinton. 2015. "Deep Learning". *Nature*, Vol. 521, No. 7553, 436–444. Available at: https://doi.org/10.1038/nature14539.

Legg, S., M. Hutter. 2007. "Universal Intelligence: A Definition of Machine Intelligence". *Minds and Machines,* Vol. 17, No. 4, 391–444. Available at: https://doi.org/10.1007/s11023-007-9079-x.

Leite da Silva, T. F. 2023. Using Transformer Networks and Tensive Semiotics to Improve sentiment Analysis Accuracy in Tourism Digital Platforms. In Santangelo, A., M. Leone (eds.). *Semiotica e intelligenza artificiale* (Coll. *I saggi di Lexia*, n. 48). Roma: Aracne, 131–147.

Leone, M. 2023. I compiti principali di una semiotica dell'intelligenza artificiale. In Santangelo, A., M. Leone (eds.). *Semiotica e intelligenza artificiale* (Coll. *I saggi di Lexia*, n. 48). Roma: Aracne, 29–44.

Leone, M. 2024. "Semiotics of the Black Box: On the Rhetorics of Algorithmic Images". *Visual Communication*, Vol. 23, No. 3, 426–451

Li, J., H. Zhou, S. Huang, S. Cheng and J. Chen. 2023. "Eliciting the Translation Ability of Large Language Models via Multilingual Finetuning with Translation Instructions”. *ArXiv*, arXiv:2305.15083. Available at: https://doi.org/10.48550/arXiv.2305.15083.

Lotman, J. 1977. Primary and Secondary Communication-modelling Systems. In Lucid, D. (ed.). *Soviet Semiotics: An Anthology.* Baltimore: Johns Hopkins University Press, 95–98.

Lotman, Y. M. 1990. *The Universe of the Mind*. New York: IB Tauris.

McLuhan, M. 1964. *Understanding Media: The Extensions of Man*. New York: McGraw-Hill.

Montanari, F. 2025. “ChatGPT and the Others: Artificial Intelligence, Social Actors, and Political Communication. A Tentative Sociosemiotic Glance”. *Semiotica*, Vol. 2025, No. 262, 189–212.

Paolucci, C. 2025. “The Myth of Meaning: Generative AI as Language-endowed Machines and the Machinic Essence of the Human Being”. *Semiotica*, Vol. 2025, No. 262, 5–23.

Reyes, E. 2024. “Generative Media: Sign, Metaphor, and Experience”. *Digital Age in Semiotics & Communication,* Vol. 7, 62–79. Available at: https://doi.org/10.33919/dasc.24.7.4.

Reyes, E. 2025. “Characterizing AI in Media Software: An Interdisciplinary Approach to User Interfaces”. *Carte Semiotiche,* 2, 36–49. hal-04915703. Available at: https://hal.science/hal-04915703v1.

Rosenblatt, F. 1958. “The Perceptron: A Probabilistic Model for Information Storage and Organization in the Brain”. *Psychological Review*, Vol. 65, No. 6, 386–408. Available at: https://doi.org/10.1037/h0042519.

Sanna, L. 2020. *Data-driven Semiotics and Semiotics-driven Machine Learning*. In *LEXIA*. *Rivista di semiotica, 33–34,* 89–107. Available at: https://doi.org/10.4399/97888255354266.

Santangelo, A., M. Leone (eds.). 2023. *Semiotica e intelligenza artificiale* (Coll. *I saggi di Lexia*, n. 48). Roma: Aracne.

Santangelo, A. 2023. Essere umani al tempo dell’intelligenza artificiale. Narrazioni a confronto In Santangelo, A., M. Leone (eds.). *Semiotica e intelligenza artificiale* (Coll. *I saggi di Lexia*, n. 48). Roma: Aracne, 239–266.

Saussure, F. de 1916. *Cours de linguistique générale*. Lausanne & Paris: Payot.

Semenenko, A. 2012. *The Texture of Culture: An Introduction to Juri Lotman’s Semiotic Theory*. New York: Palgrave Macmillan.

Shamim, S. 2024. Why Google’s AI Tool Was Slammed for Showing Im-Ages of People of Colour: Gemini’s Attempt to Correct the Marginalisation of Ethnic Minorities has Backfired. *Al Jazeera*, 9 March 2024. Available at: https://www.aljazeera.com/news/2024/3/9/why-google-gemini-wont-show-you-white-people.

Soro, E. 2023. Fit Different! An Exploration of Apple Fitness Artificial Affects Machine. In Santangelo, A., M. Leone (eds.). *Semiotica e intelligenza artificiale* (Coll. *I saggi di Lexia*, n. 48). Roma: Aracne, 197–218.

Stano, S. 2023. I corpi (im)possibili dell'Intelligenza Artificiale" in Santangelo, A., M. Leone (eds.). *Semiotica e intelligenza artificiale* (Coll. *I saggi di Lexia*, n. 48). Roma: Aracne, 219–237.

Templeton, E., A. Carr, M. Stuhlsatz, D. Xu, et al. 2024. "Scaling Monosemanticity: Extracting Interpretable Features from Claude 3 Sonnet". *ArXiv preprint* arXiv:2405.13278. Available at: https://arxiv.org/abs/2405.13278.

Turing, A. M. 1948. *Intelligent Machinery* (NPL Report). Teddington, UK: National Physical Laboratory.

Valle, A. 2025. "From Grammar to Text: A Semiotic Perspective on a Paradigm Shift in Computation and its Usages". *Semiotica*, Vol. 2025, No. 262, 51–69.

Vaswani, A., N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser and I. Polosukhin. 2017. "Attention Is All You Need". *ArXiv*, arXiv:1706.03762. Available at: https://doi.org/10.48550/arXiv.1706.03762.

Volli, U. 2023. Il mito delle due intelligenze e la semiotica. In Santangelo, A., M. Leone (eds.). *Semiotica e intelligenza artificiale* (Coll. *I saggi di Lexia*, n. 48). Roma: Aracne, 91–103.

Voto, C. 2023. Interpretare il cambiamento climatico attraverso l'intelligenza artificiale. Una prospettiva semiotica. In Santangelo, A., M. Leone (eds.). *Semiotica e intelligenza artificiale* (Coll. *I saggi di Lexia*, n. 48). Roma: Aracne, 267–281.

Wei, J., Y. Tay, R. Bommasani, C. Raffel, B. Zoph, S. Borgeaud, D. Yogatama, M. Bosma, D. Zhou, D. Metzler, E. H. Chi, T. Hashimoto, O. Vinyals, P. Liang, J. Dean and W. Fedus. 2022. Emergent Abilities of Large Language Models. *Transactions on Machine Learning Research*. ArXiv:2206.07682. Available at: https://doi.org/10.48550/arXiv.2206.07682.

Wittgenstein, L. 1921. *Tractatus Logico-Philosophicus*. London: Kegan Paul.