



Federico Montanari*

ChatGPT and the others: artificial intelligence, social actors, and political communication. A tentative sociosemiotic glance

<https://doi.org/10.1515/sem-2024-0210>

Received December 5, 2024; accepted December 26, 2024; published online February 19, 2025

Abstract: The aim of this article is, on the one hand, to take up and discuss some key categories and concepts in semiotics, in an attempt to analyze the mechanisms underlying current artificial intelligence (AI) models, with a focus on ChatGPT. Although many of these concepts are already being debated, they remain crucial in relation to semiotic and sociosemiotic categories. Concepts such as generativity, perception, textuality, and the effects of meaning, as well as the notion of language itself, require a new semiotic evaluation, including in relation to a cultural history of AI, and the metaphors associated with it. In addition, this article aims to propose more of a general observation and review of some theoretical problems and sociosemiotic issues related to the advent of AIs with some correlated examples, such as in the field of political communication and conflict.

Keywords: artificial intelligence; sociosemiotics; history of AI; STS; AI and political communication

1 Introduction: fears and opportunities

The aim of this paper is twofold: first, it seeks to take up and discuss some key categories and concepts that are critical, in relation to semiotics, for understanding the mechanisms underlying various forms of artificial intelligence (AI), with a particular focus on ChatGPT. Although many of these ideas have already been debated, they remain crucial in connection to semiotic and socio-semiotic categories. Concepts such as generativity, perception, textuality, and the effects of meaning, as well as the very notion of language, require careful evaluation. This includes addressing narrative and figurative dimensions. So, this paper aims to propose more

*Corresponding author: Federico Montanari, University of Modena and Reggio Emilia, Modena, Italy, E-mail: federico.montanari@unimore.it

of a general observation and review of some theoretical problems and related socio-semiotic issues, with some examples, than to present a particular case study.

Within the limits of length of this paper, I would also like to explore some aspects related to the semio-political dimension of AI, both through examples of its use and in a broader context, examining the axiologies, ideologies, and narratives inside discourses concerning AI systems such as ChatGPT. In this regard, the use of metaphors in AI definition and public discourse provides a valuable framework for this discussion. Let us also think of concepts such as AIs as “mediating objects” and “distributed machine perceptions.” Or concepts like “assistant,” “helper,” all the way down to the recent “Transformer models” in LLMs: they work, at the same time, as labels, technical functions, but also in form of narrative functions, as well as “catacrezised metaphors” (cf. on this particularly discussions from Fontanille and Bertrand¹). We will come back on some of these points below.

In the past two or three years, the topic of artificial intelligence – particularly the various forms and projects it encompasses – has seen a remarkable surge in prominence within social and cultural debates. This trend has been catalyzed by the successive releases of ChatGPT and similar technologies. The discourse around AI has quickly evolved into a global phenomenon, cutting across disciplines and societies, and intertwining with broader concerns about big data – its usage, analysis, and accessibility – as well as the overarching role of algorithms in shaping contemporary life.

This debate has unfolded across multiple dimensions. At one level, it encompasses existential concerns, such as the potential risk of an “end of humanity” or the disruption of human civilization, as highlighted by thinkers like Harari, along with widespread calls for moratoria. At another level, it raises pressing questions about freedoms and the future of democracy, particularly regarding who will make critical decisions – whether legal, economic, or otherwise. Education is another area of concern, as AI technologies challenge traditional learning and teaching paradigms. Meanwhile, the media has experienced profound impacts, particularly in the realm of political communication. The rise of fake news, deepfakes, and “super” fake news has exacerbated misinformation, posing significant risks to public opinion and audience manipulation, especially in the volatile context of ongoing crises and warfare.

Moreover, there are concerns about risks (but also opportunities) related to teaching, academic research, and scientific production: who will write the articles?

¹ Regarding the issue of the use of metaphors and catacrezised metaphors in the more or less technical definitions and debate on AIs, cf. the contributions and discussions, in particular by Jacques Fontanille and questions by Denis Bertrand, during the recent edition of the International Semiotics Seminar, Paris, 2023–2024, coordinated by Maria Giulia Dondero and Juan Alonso Aldama, and from which I drew many insights for this article, and whom I would like to thank here very much. All sessions can be found on YouTube (see: <https://www.youtube.com/watch?v=RmMn0sOE4FQ>).

Or the students' theses? In this context, we also find interesting examples and experiments: as recently proposed by some scholars (see the note below, but we believe this is a fairly widespread experience), they attempted to have ChatGPT write a scientific research project, or at least parts of it, obtaining excellent results. This, however, also indirectly highlights all the flaws and shortcomings of the funding systems, application processes, and selection mechanisms for research projects. According to these scholars, "The fact that artificial intelligence is able to perform much of the work ridicules the process. It is time to make it easier for scientists to apply for research funding."²

More broadly, a range of concerns – some more justified than others – regarding the use of AI have emerged, occasionally reviving stereotypes, clichés, and even metaphors. These anxieties often lead us back to the longstanding debate surrounding media and cultural practices. We find ourselves once again grappling with the classic opposition – what we might now term polarization – between "apocalyptic" and "integrated" perspectives, as concepts famously explored by Umberto Eco. This dichotomy tends to resurface whenever society encounters a critical juncture, an epochal moment of crisis or transition, whether technological, socio-cultural, or media-related. However, this familiar opposition has resurfaced with a new complexity, partially diverging from the form Eco originally described. Today, it intensifies and proliferates, giving rise to a spectrum of intermediate and varied positions. On one end, we see dystopian visions reminiscent of a science fiction anthology – from the ever-cited (and perhaps over-cited) HAL 9000 in *2001: A Space Odyssey* to the concept of an alien or infectious entity, whether from outer or inner space, "taking over" and seizing control of humanity. These narratives draw from classic sci-fi films from the 1950s to contemporary series like *The Last of Us* (albeit with a fungal epidemic as the catalyst for takeover).

On the other end, we encounter "techno-enthusiastic" discourses, often espoused by scholars and philosophers who, by emphasizing a sense of "continuity" within the evolution of technology – including digital advancements – slip into a kind of positivist scientism or uncritical faith in engineering, envisioning a techno-vitalist rebirth through AI. Amid these extremes, however, there are many scholars who aim to engage with AI in a more balanced manner, neither uncritically embracing it nor viewing it through an apocalyptic lens.³

² "ChatGPT use shows that the grant-application system is broken"; *Nature*, October 2023: <https://www.nature.com/articles/d41586-023-03238-5> (accessed 26 December 2024).

³ In this regard, for a discussion concerning these forms of uncritical bias in AI discourse, see an interview with Luciano Floridi, in *Link*, edited by Luca Barra, October 2023: <https://www.linkideeperlatv.it/intervista-a-luciano-floridi/> (accessed 30 November 2024).

What impact might these emerging issues, discussed at various levels and already initiating real transformations in different discursive and social practices, have on political communication and public debate? Could the rapid spread of artificial intelligences prompt us to rethink not only the forms of political communication and discourse but also the very ways in which we observe and analyze them?

Allow me a brief digression. When discussing artificial intelligences, it is important to use the plural, as we must recognize the diversity of AI models and types, particularly in terms of their applications. These AIs have been developed and widely adopted with remarkable success and rapid dissemination. They range from “specific and limited” systems – such as those used for voice recognition, translation, writing automation, or assistance in various tasks, including autonomous driving – to various “assistants” like Alexa or Siri. There are also more advanced systems that integrate the analysis of large datasets with heuristic-predictive capabilities, useful for companies and research centers. We then arrive at the more recent “general purpose” or quasi-universal models, such as those based on deep learning and large language models with statistical-generative capabilities, like Google Bard or ChatGPT, which may even anticipate the advent of General AI systems (AGI). As we know, the introduction of neural networks and deep learning models, such as the so-called “Transformers,”⁴ has revolutionized the way AI processes data. These innovations have gradually been integrated into various productive, economic, and intellectual activities, sparking both interest and enthusiasm (consider their applications in medicine, molecular and genetic research, and even agriculture). However, these advancements also raise concerns, particularly regarding the potential loss of jobs and activities, including those that require intellectual and cognitive skills, fueling fears of “thousands of jobs being lost.” This has led to ongoing discussions, initially among experts but now engaging increasingly broader audiences.

2 Generative language models and transformers: intelligent machines or just “clever” and “plagiarizing”?

These issues have become integral to public debate and, in many ways, part of a shared societal sentiment – one that blends perceptions of risks, fears, and concerns

⁴ See, in this regard for a first description of the origins and rapid development of these “Transformer models”: <https://blogs.nvidia.com/blog/2022/03/25/what-is-a-transformer-model/> (accessed 26 December 2024).

with expectations, hopes, and the pursuit of opportunities. For instance, when considering transformations in the workplace, AI offers significant potential to assist in complex and challenging tasks. It can process vast amounts of data in fields like medicine and pharmacological research or take over repetitive, tedious, and exhausting duties, freeing human workers to focus on more creative and strategic endeavors.

In this sense, the well-known critique by Noam Chomsky, very close to the “stochastic parrot” metaphor, for using a definition coined by the linguist Emily Bender – sounds paradoxical, though interesting in its own way, as Chomsky believes that ChatGPT is nothing more than a

clumsy statistical machine for recognizing patterns that ingests hundreds of terabytes of data and extrapolates the most plausible response for a conversation or the most likely one for a scientific question. In contrast, ... the human mind is an astonishingly efficient and elegant system that operates with a limited amount of information. It does not seek to infer brute correlations from data but rather to create explanations. ... Let's stop presenting it as “Artificial Intelligence” and call it what it is: “plagiarism software.” (Chomsky 2023a)

We will return to this point soon regarding some political-communicative implications. For now, what emerges in this case is, on the one hand, the critique and fears towards a model of “generative” intelligence (incidentally, we should remember that Chomsky’s linguistic model was also generative, see below about this point, albeit with entirely different intentions and not based on a probabilistic-statistical mathematical dimension). Intelligence that, through an extremely complex system with Large Language Models (cf. De Baggis and Puliafito 2023; Douglas Heaven 2023; Resnik 2024) composed of millions and millions of parameters – and featuring stratifications across an enormous number of levels or “layers” and interconnected components – produces statistical correlations capable of recognizing recurring patterns. It is also interesting to note that these generative LLMs, according to their designers and researchers, model and generate language using the same mechanisms that can then be used to simulate chemical-molecular models and vice versa, applicable in other fields as well (musical, artistic creations, or others). In a sort of universalization (through the use of probabilistic-statistical predictive models) of the ability to detect and generate patterns, connections, and syntactic, then possibly semantic, relationships.

On the other hand, Chomsky critiques what he terms “generalized plagiarism” in AI systems, a concept that extends beyond its traditional definition. He highlights the protests by writers, artists, screenwriters, and news organizations worldwide against the unrestricted and unauthorized use of vast text corpora for AI training. These systems rely on consuming millions of textual works without proper consent

or acknowledgment. For Chomsky, this represents not only an ethical breach but also an attempt by systems like ChatGPT to position themselves as “impostors,” claiming equivalence with human intelligence.

More broadly, Chomsky raises deeper “techno-moral” concerns. He critiques the creation of engineered intelligences that, while highly efficient, sophisticated, and powerful, fundamentally lack the essential traits of human intelligence: a “moral faculty,” ethical sensibilities, self-control, and the capacity for self-assessment. This absence, he argues, underscores the inherent limitations and potential dangers of relying on such systems.⁵

3 A short provisional genealogy of AI, between behavior and cognition: cultural semiotic implications

Before proceeding, however, I would like to propose a brief genealogy of AI – a history that is both mythical and real, well-known yet deserving of re-examination for its semiotic implications and cultural significance. This history is relevant not only to the field of semiotics but also to Science and Technology Studies (STS) and Actor-Network Theory (Ribes 2018), particularly in relation to the techno-scientific and social status of AI.

We often consider the techno-cultural history of AI primarily as a linguistic and cognition issue, which is indeed a key aspect. However, it is also deeply connected to broader themes such as a theory of human rationality and behavior in the emergence of cognitive sciences. To address the epistemological and socio-semiotic foundations of AI, it is necessary to consider the study of rational behavior models developed during the nascent cognitive sciences of the 1950s. This is often highlighted even by Noam Chomsky in his recent interviews, where he remains critical and polemical toward models like ChatGPT, which he refers, as just said, to as a “statistical” or, better, “stochastic parrot.” I will return again to Chomsky’s critiques later in this paper.

The term “artificial intelligence,” as it is well known, first appeared in 1956 at the famous Dartmouth Conference, an event widely regarded as marking the

⁵ Chomsky (2023b) says:

Unless carefully controlled, AI engineering can pose severe threats. Suppose, for example, that care of patients was automated. The inevitable errors that would be overcome by human judgment could produce a horror story. Or suppose that humans were removed from evaluation of the threats determined by automated missile-defense systems. As a shocking historical record informs us, that would be the end of human civilization.

official beginning of AI as an interdisciplinary field of study (see, for a historical reconstruction, i.e., Douglas Heaven 2024; Pasquinelli 2023).⁶ Organized by John McCarthy, Marvin Minsky, Nathaniel Rochester, and Claude Shannon at Dartmouth College, the conference laid the groundwork for AI, introducing concepts such as symbolic manipulation, logic-based systems, and the idea of creating machines capable of solving general problems.

Herbert Simon, who later received a Nobel Prize (significantly) in Economics, and Allen Newell developed the Logic Theorist program in 1956, a pioneering AI program capable of proving mathematical theorems. Simon's theory of "bounded" (or "limited") rationality, which challenges the traditional economic notion of perfect rationality, suggests that in real decision-making situations, individuals are limited by cognitive and time constraints, leading to satisfying rather than optimizing behavior. The relevant point here is the nexus with economic sciences and the attempt to go and find the basis of rational human behavior, related to decision-making.

McCarthy famously defined AI as "the science and engineering of making intelligent machines," related to understanding human intelligence and behavior through computational models. This approach also involved studying data interpretation, structure, and patterns. Stuart Russell of UC Berkeley has noted that AI's origins are closely linked to cognitive science as well as to economic study of rational behavior, as exemplified by Simon's work. This connection is particularly interesting from a semiotic perspective, even though subsequent developments favored other models, such as bounded rationality and, later, the so-called strategic rationality (like those proposed decades later by Crozier and Friedberg in France, starting also from Goffman), and different shifts concerning research on different forms of "behavioral" rationality in economics.⁷ This research paradigm, which originated in

6 Douglas Heaven, in his extensive reconstruction and review, dense with critical considerations related to modern-day AI as well, recalls how the very definition of Artificial Intelligence was controversial from the very beginning.

In fact, "artificial intelligence" was just one of several labels that might have captured the hodgepodge of ideas that the Dartmouth group was drawing on. The historian Jonnie Penn has identified possible alternatives that were in play at the time, including "engineering psychology," "applied epistemology," "neural cybernetics," "non-numerical computing," "neuraldynamics," "advanced automatic programming," and "hypothetical automata."

What Turing saw – unlike McCarthy's crew – was that thinking is a really difficult thing to describe. The Turing test was a way to sidestep that problem. "He basically said: Instead of focusing on the nature of intelligence itself, I'm going to look for its manifestation in the world. I'm going to look for its shadow" says Law ... a researcher who studies the history of AI at the University of Cambridge and ethics and policy at Google DeepMind (Douglas Heaven 2024).

7 See, about this, and the intertwining with narrative models, STS research, and economic studies, Ribes 2018; Svetlova 2021.

the 1950s, aimed to create machine models capable of optimizing human behavior in various contexts, from chess and, later, GPS systems to operational research in economics and statistics. On those years, Alan Turing's work, as it is well known, particularly his "black box" model of a chess-playing program and his 1950, best known, paper "Computing Machinery and Intelligence," laid the foundation for the famous Turing Test. The Turing Test, or "the imitation game," evaluates a machine's ability to exhibit intelligent behavior equivalent to that of a human. Turing also introduced the concept of the "universal machine" (cf. Douglas Heaven 2024), which evolved into the fundamental idea of the modern programmable computer. What is particularly relevant here is the concept of the "conversational machine" – a foundational idea that has persisted over time and remains central to the development of AI. This notion aligns with the modern practice of using "prompts" as an interactional tool to engage effectively with machines. The recent proliferation of prompt user manuals, tutorials, and various "prompt styles" underscores the growing significance of this interactional paradigm.

The ideal of a conversational machine suggests that if a machine could engage in dialogue with a human so convincingly that the human could not distinguish its responses from those of another human, its behavior might be deemed intelligent. This concept not only captures the essence of Turing's original vision but also reflects the evolution of human-machine interaction in contemporary AI systems.

From this shifting paradigm, general-purpose AI emerged, especially during a period of strong, optimistic economic growth. On the contrary, 1970s and 1980s saw the development of expert systems – with their databases and inferential machines capable of carrying out acts of querying those databases – but these models encountered significant challenges during what some have called the "AI winter." However, this period of "stasis" ultimately paved the way for the development of the large language models we encounter today. As Russell observes, these models are marked by their "massive sample complexity," which allows them to process vast amounts of data, but they also exhibit a "relatively low expressive capacity," reflecting limitations in their ability to generate nuanced and contextually rich outputs.

The McCulloch-Pitts model of the neuron, developed, in an independent way, from 1943, played a crucial role from the early development of neural networks (according to several researchers, it represents the basis for an alternative path to cognitivism, and to classical computational-type models; see Beckmann et al. 2023). This model, which uses a binary threshold for processing logic, can perform basic logical operations such as AND, OR, and NOT, but also by adjusting the input weights. The McCulloch-Pitts neuron model served as the starting point for the study of artificial neural networks, paving the way for more sophisticated models and algorithms that continue to be used and developed in modern AI applications, and whose rediscovery in the following decades has given, as it is well known, boost to machine

learning projects and today's AIs. Connectionism, the neural network modeling or parallel distributed processing, thus became the theoretical framework for AI inspired by the brain's information processing through interconnected neural networks. Another significant milestone was Joseph Weizenbaum's ELIZA program, another example of "conversational device," which simulated dialogue between a patient and a psychoanalyst. However, the landscape of AI changed dramatically after 1997, with developments such as Deep Blue and later the advent of AlphaGo through the use of neural networks and later the widespread use of statistics and trained neural networks in machine and deep learning systems.

3.1 A critical phenomenological look at AI

The shift to models employing trained and, more recently, untrained networks marked a pivotal moment in the evolution of AI, with profound implications for semiotics. This transition raises critical epistemological questions and invites reflection from a semiotic perspective. Philosopher Hubert Dreyfus, in his seminal work *What Computers Can't Do*, later updated as *What the Computer Still Can't Do* (1992, 2007), offered a powerful critique of disembodied simulations of intelligence. Drawing on European phenomenology, particularly the works of Merleau-Ponty, Heidegger, and Marc Johnson, Dreyfus argued that such simulations fail to capture the embodied and situated nature of human cognition.

Dreyfus was particularly critical of the shortcomings of GOFAI (Good Old-Fashioned AI), highlighting its inability to emulate the nuanced, embodied, and context-sensitive aspects of human intelligence. While he acknowledged the potential of neural network models to address some of these limitations, he remained skeptical of their capacity to fully replicate human cognitive processes, emphasizing the inherent challenges of reducing lived experience to computational inferences. This critique remains relevant in discussions of the theoretical and practical limits of contemporary AI systems.

Dreyfus argues that it is impossible to reduce everyday life experience to mere inferences and questions how skills and know-how can be represented as knowledge. He also emphasizes the importance of imagination, particularly the need to organize spatially knowledge for understanding also typical sentences, i.e., the role of deixis – for instance, how we locate things in relation to our position in space and time coordinates, such as "over there" or "nearby." These issues point to the broader, and today classic, problem of embodiment, as explored by Johnson (and together with the fundamental works of Lakoff, paving the way for the fundamental path of research, in recent decades, related to embodied cognition and enactivism), who argues that imagination is a pervasive structured activity (see, for further

insights about an “alternative path,” concerning a possible “computational phenomenology,” based on the same philosophical sources and issues proposed by Beckmann et al. 2023 and Dreyfus).

4 Spatializations, vectors, and statistics: a new distributionalism and new semiotic challenges

The importance of spatialization in models of sense production, particularly in structural semiotics and enunciation, is almost a given. However, a pressing question arises: how have contemporary AI models integrated spatialization, whether of information or meaning? Recent advancements in AI have heavily relied on the superimposition of neural network layers, leveraging statistical and vectorization processes.

These processes effectively simulate internal “spaces” of connection, mapping relationships between concepts, words, syllables, punctuation, or sentences, which are transformed into “tokens.” In this framework, vectors are employed to measure the distances between these tokens, capturing semantic and syntactic relationships. Crucially, statistics play a central role in these models, enabling the calculation of probabilities for these distances and guiding the predictive capabilities of AI systems. This combination of spatialization and probabilistic modeling lies at the heart of the generative and analytical power of modern AI systems, marking a significant evolution in the ways meaning and structure are computationally processed.

Statistics has played here a fundamental role in the analysis, manipulation, and extraction of meaningful information from vectors in AI applications. Various statistical methods are used in vector operations, such as addition, subtraction, multiplication, and division. These operations allow for the manipulation and transformation of data represented by vectors. Statistical concepts like mean, variance, and covariance describe vector properties in vector spaces, helping to understand distributions and relationships within vectors. Techniques such as Principal Component Analysis (PCA) and Singular Value Decomposition (SVD) reduce the dimensionality of vectors while preserving essential information, particularly when dealing with high-dimensional data. Clustering and classification methods, like k-means clustering⁸ and support vector machines,⁹ use vectors to group similar data

⁸ An unsupervised machine learning algorithm used to group data into a specified number of clusters based on their similarities. It is widely used in pattern recognition, data mining, and statistical analysis.

⁹ They are supervised machine learning algorithms primarily used for classification and regression tasks. They are particularly effective in high-dimensional spaces and are widely recognized for their robustness in handling linear and non-linear data.

points or classify them into categories based on statistical principles, leading to accurate predictions. Additionally, feature selection techniques identify the most informative features in vectors, improving AI model efficiency and accuracy. Statistical regression techniques model relationships between input and output variables, commonly used in predictive modeling tasks. Furthermore, distance metrics such as Euclidean distance, or cosine similarity, and others, measure vector similarity or dissimilarity.

In summary, the advent of statistics plays an essential role in AI, enabling systems to process and interpret data represented as vectors, making informed decisions and predictions.

The other key role, alongside statistics and vectorization procedures, has been that played by distributionalism, which has taken on new life thanks to them (see Resnik 2024 on this “renaissance” of distributionalism, thanks to statistics and vectorialization). Distributionalism resuscitated thanks to statistics in AI, which focuses on meaning representation through language distribution models, underpins this framework. The distributional hypothesis, central to this paradigm, suggests that words appearing in similar contexts and patterns tend to have similar meanings. Distributional semantic models, based on this hypothesis, represent word or phrase meanings through their distribution patterns in text corpora. These models often use word embedding techniques like Word2Vec¹⁰ and GloVe¹¹ to capture semantic relationships between words.

Vector space models in distributional semantics represent words or phrases as vectors in high-dimensional spaces, capturing similarities and semantic relationships. This enables tasks like word similarity, analogy completion, and sentiment analysis. Recent advances in AIs, such as contextual word integration (e.g., BERT, GPT), build on distributional principles, capturing not only individual word distributions but also their context within larger text sequences, leading to more nuanced representations of meaning. Distributionalism in AI seems to provide a powerful framework for capturing and representing semantic relationships in language (and perhaps a challenge for semiotics), allowing AI systems to interpret and generate human-like responses based on distribution patterns in textual data.

More generally, the ongoing debate among AI and LLM researchers centers on how to give computers facts and rules or, more importantly, how to structure data effectively. Hubert Dreyfus (1992) again explained that the earlier focus was on

¹⁰ Natural language processing (NLP) technique for learning vector representations of words, also known as word embeddings. These embeddings represent words in a continuous vector space where semantically similar words are closer together.

¹¹ Global Vectors for Word Representation is a model for generating word embeddings that represent words in a continuous vector space.

storing facts without context, whereas the real challenge and the emphasis is on producing and reinventing contexts. Dreyfus (1992: XXII-XXIII), drawing again on Merleau-Ponty and Bourdieu, argues that “savoir-faire,” or practical knowledge (i.e., regarding cultural logics, such as that of gift and counter-gift, studied by Bourdieu and anthropologists) must be deferred and distinct, as it is “a matter of style.” That is, the ability to coordinate to a situation at a given time without the need to extrapolate or recognize a-priori a pattern and then apply this recognition, which AI classically does. In contrast, AI tends to extrapolate style in a classical manner.

Returning to the discussion on distributionalism and statistics, the approach has significantly evolved, now enriched by vector space models and statistical-probabilistic methods. These advancements have sparked considerable debate, particularly from figures like, once again, Noam Chomsky. Since the 1950s, Chomsky has been a vocal critic of probabilistic models in linguistics, including Markov models, and has extended this criticism to the use of statistical methods in AI. Peter Norvig (2017) provides a concise summary of Chomsky’s critiques, highlighting the fundamental tensions between statistical approaches and Chomsky’s views on the nature of language and cognition:

- 1) Statistical language models may have engineering success, but this is irrelevant to science. Accurate modeling of linguistic facts is akin to “butterfly collecting”; what matters in science is the underlying principles.
- 2) Statistical models are incomprehensible and provide no insight.
- 3) Statistical models might simulate some phenomena accurately, but they do so in the wrong way. People do not decide on the third word of a sentence by consulting a probability table based on previous words. Instead, they map from an internal semantic form to a syntactic tree-structure, which is then linearized into words, without using probability or statistics.
- 4) Statistical models have been proven incapable of learning language, suggesting that language must be innate. So, why are statistical modelers wasting their time on the wrong enterprise? (Norvig 2017)

Norvig offers some counterpoints:

- 1) Engineering success is not the sole measure of science, but science and engineering develop together. Engineering success indicates that something is working correctly, which is evidence (though not proof) of a scientifically successful model.
- 2) Science involves both gathering facts and making theories; neither can progress alone. In the history of science, the laborious accumulation of facts has been the dominant mode. The science of understanding language is no different in this respect.
- 3) While a model containing billions of parameters can be difficult to understand by inspecting each parameter individually, insight can be gained by examining

the model's properties – where it succeeds and fails, and how well it learns as a function of data.

- 4) A Markov model of word probabilities¹² cannot model all language, just as a concise tree-structure model without probabilities cannot either. A probabilistic model that covers words, syntax, semantics, context, discourse, etc., is needed. Chomsky dismisses all probabilistic models because of the shortcomings of a particular 50-year-old model. Interpretation tasks, such as speech recognition, inherently involve probabilistic problems, making probabilistic models our best tool for representing language facts, processing language algorithmically, and understanding how humans process language (Norvig 2017).

How, then, can we reconcile, on the one hand, the undeniable advancements in research on Large Language Models (LLMs) – advancements driven by the advent of probabilistic-statistical and vector-based models that have successfully moved beyond the stagnation of earlier computational linguistics and the limitations of old distributionalism, as well as countering Chomskyan critiques – and, on the other hand, the criticisms originating with Dreyfus and the notion of a potential computational phenomenology (as discussed by Beckmann et al. 2023)? Or, more aptly, should we consider the possibility of a computational semiotics? This approach would recognize that meaning and its production are always “experiential,” grounded in concrete “savoir-faire,” embodiment, and situational knowledge.

Furthermore, how do we account for the inherent limitations of LLMs themselves, as highlighted by several researchers (e.g., Resnik)? These limitations, often manifested as various types of biases, suggest that AIs require “more semantics” and that not all linguistic mechanisms can be fully described, or effectively generated, through statistical distribution alone.

5 Another approach to generativity?

Related to these points we find another and fundamental issue: the question of generativity. In AI and ChatGPT, generativity refers to a system's ability to generate new and creative results (in terms of text or image production) beyond what it has been explicitly trained to do, implying the ability to generalize from learned models and rules to produce new and innovative solutions or content. In structural

¹² A probabilistic model that predicts the likelihood of a sequence of words based on the assumption that the probability of each word depends only on a finite number of preceding words. This simplification is derived from the Markov property, which states that the future state (or word) depends only on the current state (or preceding words) and not on the entire history.

semiotics, generativity is conceived very differently also from the Chomskyan conception – although semiotics has borrowed some of the concepts from it – it is linked to metalinguistic constructions. “Deep” and “surface” are spatial scientific metaphors relating to the axis of verticality, designating the starting position and end point of a chain of transformations. This chain represents a process of generation, a generative trajectory – the generative path (*parcours*) as it has been proposed, as it is well known, by Greimas and his school (Greimas and Courtés 1993 [1979]) – within which the various stages are distinguished by an increase in the degree of meaning complexity. The operational nature of these structural stages justifies the questioning and arrangement that the theory must carry out. In semiotics, the use of this dichotomy fits within the general theory of meaning generation. It accounts for the generative principle, where complex structures arise from simpler ones, and the principle of “meaning growth,” where each structural complexification produces an extension of meaning. Therefore, each domain of the generative trajectory includes both syntax and semantics.

The notion of depth is relative, with each domain of discourse generation referring to a “deeper” domain, leading to the deep structure par excellence: the elementary structure of meaning in Greimas’ structural semiotics. Could this stratified structural model, in comparison with the idea of generativity in AIs, be useful for the advancement of knowledge and further development of AI forms? Considering that this concept of generativity is intrinsically linked to the enrichment and growth of meaning complexity – and, by extension, to invention and creativity – it is worth exploring another critical issue in contemporary AI: the phenomenon of so-called “hallucinations.” These hallucinations represent unique and “uncontrolled” ways of generating possible meanings, which merit closer examination as we seek a more nuanced understanding of AIs’ capabilities and limitations.

5.1 From generation to AIs hallucinations: an open question

AI hallucinations occur when generative models produce inaccurate information that appears true. Flaws in training data and algorithms cause these false outputs, leading to convincingly presented but illogical or untrue content. Biased or poor-quality training data, insufficient user context, and inadequate programming can lead AI systems to misinterpret information, resulting in hallucinations. These issues are common in text generators, image recognition, and creation models (cf. Shah 2023; Vendeville et al. 2024; Zhang et al. 2023).

Emily Bender, quoted by Shah, explains that, from a linguistic point of view, LLMs do not initially comprehend word meanings. Instead, they view text as

sequences of characters, recognizing patterns over time to grasp language rules, word associations, and semantics. This allows LLMs to generate human-like content. However, their lack of factual understanding causes hallucinations when answering questions. But sometimes things do look more complicated, and some unpredictable elements seem to emerge.

In any case, Hardik Shah (2023) identifies four types of AI hallucinations:

- 1) **Factual Inaccuracies.** AI often generates text that appears factual but contains incorrect details. For example, in 2023, Google's Bard chatbot falsely claimed that the James Webb telescope had taken the first image of an exoplanet, although the first images were taken in 2004, years before the Webb telescope's launch.
- 2) **Manufactured Information.** AI can generate entirely fictitious content, such as fake URLs, codes, people, news, books, and research. This presents risks for those using AI like ChatGPT for research, as it can produce plausible but false information. For instance (quoting again Shah 2023), in June 2023, a lawyer used ChatGPT to create a legal motion filled with false case law and citations. The lawyer was unaware of ChatGPT's ability to fabricate information that sounds credible but is not real, resulting in a fine for submitting the fictitious request.
- 3) **Dangerous Misinformation.** AI can produce false and defamatory information about real people, combining truth and fiction to create damaging stories. This type of misinformation can have serious consequences. For example, ChatGPT falsely claimed that an Australian mayor had been convicted of corruption in the 1990s, even though he was a whistleblower. Such misinformation has attracted the attention of the US Federal Trade Commission, which is investigating whether OpenAI has damaged its reputation by making false claims.
- 4) **Strange or Worrying Results.** Some AI hallucinations are bizarre or frightening. By design, AI models aim to generalize patterns and generate creative results. While this creativity can produce strange results, it is not necessarily problematic unless accuracy is essential. For instance, Microsoft's Bing chatbot behaved strangely by professing love for a journalist and enlightening users.

Other researchers (e.g., Thiollet 2024) classify hallucinations as "Input-conflicting," "Context-conflicting," and "Fact-conflicting"; they depend upon biases related to data, how are they interpreted and manipulated, or how models and inferences work. "Some notable sources of hallucinations are related to training data biases, parametric knowledge biases, data encoding problems, or overfitting" (Zhang et al. 2023; also quoted also in Thiollet 2024). Even if not always harmful, these responses illustrate the unexpected and sometimes troubling outputs that AI can produce when given creative freedom, or because of biases. Again, following Thiollet:

More and more experts are arguing that hallucinations are not a bug but a feature of the technology behind Generative AIs. By broadening the definition of hallucination as considered so far in our reflection (i.e. an unsatisfactory result presented in a factual way) to that of an inherent characteristic of Generative AIs, it becomes possible to take a new look at this phenomenon and to see its advantages. (Thiollet 2024: 2)

More generally, even these issues highlight the need for careful monitoring and ethical considerations in AI development and application, nevertheless, several other AI researchers point out that hallucinations also represent opportunities for experimentation (cf., Thiollet 2024); to observe the emergence of new unexpected patterns and new forms of possible creative meaning in AIs: “hallucinations as sources of Creativity.” At the same time, some scholars suggest that there is here the typical risk of anthropomorphism: that is, of attributing to machines behavior that is excessively similar to that of humans. The problem lies perhaps in maintaining the dimension of experimentation, and at the same time that “right distance” of observation so as not to fall into this kind of fallacy.

From a semiotic and socio-semiotic perspective, the issue of hallucinations invites further investigation, even within the constraints of this article. It is essential to explore the broader implications of the category of hallucination, as partially outlined earlier. Are hallucinations merely errors or biases in machine outputs, or do they reveal an ability to generate unexpected pathways of meaning through these very errors? Could they serve as a metaphorical extension – what some commentators describe, as said, a “catacresized metaphor” – that anthropomorphizes Artificial Intelligence? Alternatively, might the notion of the “unexpected,” as theorized by Greimas, offer valuable insight? In *De l'imperfection*, Greimas (1987) examines the concept of the “unexpected” within structural semantics and perception categorization. For Greimas, the unforeseen represents a variable – a rupture – that reshapes prefigured patterns and influences choices in pathways of meaning. This unpredictability is critical to understanding the dynamics of meaning renewal and the interpretive processes in texts. Applying this framework, the AI-generated phenomenon of hallucination might be reframed as a disruption that contributes to semantic innovation rather than as a mere failure.

Additionally, this concept of hallucination could be enriched by engaging with definitions from phenomenology. Merleau-Ponty (1945), for example, describes hallucination as the disintegration of reality before our eyes, substituting it with a “quasi-reality.” According to Merleau-Ponty, the world is not a static object with purely objective determinations; instead, it has “splits and gaps” through which subjectivity asserts itself. This interplay of subjective intrusion and ruptured reality might offer a productive lens through which to consider AI hallucinations, framing them as a site of tension between programmed objectivity and the unpredictable emergence of meaning.

6 Figures, narratives, hybrid actors, blank checks, and prompts: tentative levels of semiotic analysis and implications for political communication

The intersection of AI with politics, communication, and society is vast, influencing areas as diverse as modern warfare and the spread of misinformation. In contemporary conflicts, AI systems have played an increasingly prominent role. For example, Israel's "Lavender" and "Gospel" AI systems,¹³ as reported in *972 Magazine*, a left-wing news and opinion online Israeli magazine (but also confirmed by other influential international newspapers such as *The Guardian*, also in relation to the use of AI in the war in Ukraine),¹⁴ analyze data from extensive databases in real time to identify and prioritize "targets" or "enemies to be eliminated." These systems often autonomously provide directives for launching bombs and missiles, leaving minimal opportunity for human operators to interpret or question the data. This automation has been linked to the significant number of civilian casualties in the Gaza conflict. Similarly, Russian and Ukrainian drones now utilize AI to autonomously recognize targets, further underscoring the technological shift in warfare.

Beyond warfare, the broader implications of AI raise critical questions about reality and truth in the digital age, particularly concerning simulation and simulacra. The work of Deleuze and Guattari, particularly their concepts of "machinic chaining" and social subjection, provides valuable frameworks for understanding the mechanisms of power and influence embedded within AI-driven communication. These theoretical perspectives shed light on the complex dynamics of enunciation and control that characterize AI's role in shaping societal narratives.

So, coming to the other issues concerning the topic of this paper, what impact do these developments have on the forms and processes of political communication? The implications, it seems, are multiple and significant. What are the most relevant and interesting points regarding the political discourse of, and on Artificial Intelligences? It seems that at least three levels of relevance emerge here, closely interrelated.

At a fundamental level, as previously suggested, we are confronted with an issue of narratives, storytelling, and the "discourse" surrounding AI, along with its rhetorical frameworks. More precisely, we might refer to this as a question of

¹³ See, e.g., a journalistic investigation about these systems: "'Lavender': The AI machine directing Israel's bombing spree in Gaza," *+972 Magazine*, 2024: <https://www.972mag.com/lavender-ai-israeli-army-gaza/>.

¹⁴ See, e.g., <https://www.theguardian.com/technology/article/2024/jul/14/ais-oppenheimer-moment-autonomous-weapons-enter-the-battlefield>.

meta-narratives: how public discourse today frames and conveys the themes, figures, and metaphors of the technological and social revolution brought about by artificial intelligences. On one side, we observe a range of actors and symbolic representations: the fear of an “other,” an alien presence that is simultaneously close and familiar, juxtaposed with AI systems being cast as saviors – omnipotent or omniscient entities. Here, an unsettling paradox begins to emerge. These very systems, becoming increasingly autonomous, may soon engage in self-representation across the web, media, and public discourse. This raises intriguing possibilities about how AI might influence its own narrative in the collective imagination.

Another layer of complexity arises from the technical and scientific metaphors that underpin discussions about AI. Metaphors are critical for conceptualizing and understanding artificial intelligence, especially within the realm of neural networks. Scholars in the sociology of science, epistemology, and scientific discourse – ranging from Black to Boyd, Kuhn, and Lakoff – have long emphasized the significance of metaphors. These rhetorical-semiotical devices are vital for cognition, discovery, modeling, and the formation of technical languages in fields as diverse as economics and physics.

Consider the case of *Transformer* models, which exemplify the interplay between metaphor and function. Transformers, a specialized type of neural network, simulate certain structures and functions of the human brain, excelling at processing sequential data such as words in a sentence or notes in a melody. The transformative innovation within Transformers is the “attention mechanism,” which enables the model to focus selectively on the most relevant parts of the input sequence. This mechanism is pivotal for discerning complex relationships and dependencies within data. By revolutionizing natural language processing (NLP), Transformers have driven significant advancements in AI applications. The term “head” in Transformers, for instance, refers to the multi-head attention mechanism, a key feature that captures diverse aspects of an input sequence simultaneously. This dual role of technical objects – functionally specific and mythically resonant – reveals their broader cultural impact. Technical metaphors, often catachrestic and hybridized, solidify not only the utility but also the mystique and credibility of AI systems.

Furthermore, AI functions not merely as a “theme” but as a pervasive “figure,” a conceptual entity that traverses societal and communicative domains, particularly political communication. This is evident in the context of war, a domain where AI has become critically relevant and intersects with political discourse. Recent conflicts, such as, again, those in Ukraine and Gaza, highlight AI’s dual role: on the battlefield, where it informs tactical and strategic operations, and in the production of imagery and propaganda. In both spheres, AI operates as a tool of influence and transformation, not only reshaping the mechanics of war but also redefining the narratives and perceptions that surround it. This underlines AI’s profound and

multifaceted impact, requiring careful analysis of its role in shaping not only technical capabilities but also the socio-political narratives that frame contemporary reality. Or consider instances of the manipulation and transformation of images of public figures: for example, the viral famous deepfake photograph of the Pope dressed in an extravagant white Moncler down jacket or a fabricated image of Donald Trump being arrested. These are emblematic of a broader phenomenon – what could be termed a globalized “mega-photoshopping” era – where the boundaries between authenticity and fabrication become increasingly blurred.

Such developments carry profound implications, notably the accelerated erosion of the concept of truth and a significant challenge to the traditional dichotomy of true versus false. In the realm of communication and political discourse, we may increasingly encounter discursive and textual artifacts that autonomously traverse the web and media ecosystems. These artifacts could make it exceedingly difficult to determine their authenticity, degree of manipulation, or inherent falsehood, further complicating the already delicate interplay between information, perception, and trust.

A second, interrelated dimension emerges, addressing the specific technical and techno-social aspects of this issue. This is particularly significant as it links the problem of narratives “about” AI to the sociosemiotic study of techniques, artifacts, and the formation of social actors. This perspective is crucial for evaluating the role of AI within communication and political discourse, requiring us to conceptualize AI models as technological artifacts, hybrid objects, and, above all, “mediators” and “translators.” This approach builds on the work of Bruno Latour and his collaborators in Science and Technology Studies (STS) and Actor-Network Theory (ANT).

A sociology of techniques – integrating socio-semiotics, communication theory, and anthropology – applied to the field of Artificial Intelligence highlights how these systems function within ongoing processes of mediation and translation between “human actors” and “non-human actors.” This perspective shows the emergence of “hybrid beings” in our social world, entities that operate within and reshape practices of communication, political discourse, and media representation. These hybrids actively participate in forming new assemblages and collectives, redefining the boundaries of agency and interaction.

Consider, for example, teams deploying AI in medical contexts, where human expertise and machine intelligence coalesce into a unified but hybrid form of decision-making. Similarly, hybrid entities such as robot-artificial intelligence systems combined with voice generators exemplify this merging of human and machine capabilities. At the extreme end of the spectrum are the dramatic and troubling cases involving warfare and conflict. In these scenarios, composite hybrid entities – such as drones, as said – function simultaneously as weapons, tactical instruments, and conduits for communication, politics, and propaganda. These

entities embody a convergence of technology, strategy, and discourse, fundamentally challenging traditional notions of agency and responsibility within both technological and human domains. This raises an important question, particularly regarding AI: how should we conceptualize and analyze the “emergence” of these new hybrid beings within the social world? This inquiry extends to their communicative, political, ethical, and deontological implications, culminating in questions about the socio-semiotic and communicative – not merely cognitive – status of their “intelligence.” This issue also ties into decision-making systems and societal expectations, such as those concerning economic behavior, which are increasingly entwined with narrative, discursive, and political dimensions, as recent STS-based research has shown.

Additionally, some scholars argue that we should not think of “intelligence” as a singular concept but rather recognize a plurality of intelligences, particularly when considering humans – and even more so these new hybrid entities. For instance, distinctions can be made between operational or procedural intelligences and those directed toward specific tasks or outputs. Other researchers, such as sociologist Elena Esposito (2022), propose that these intelligences give rise to forms of “artificial communication” that are fundamentally distinct from human communication. This perspective highlights the hybrid and mediating nature of these entities, emphasizing that the most critical communicative and political phenomena involve processes of delegation and mediation toward these “new beings.” These dynamics emphasize the need to rethink traditional frameworks for understanding agency, responsibility, and interaction in light of these transformative technological developments. One of the emerging issues in both the communicative-discursive realm and in specific areas of application, such as workplaces, seems to be the so-called “deskilling,” caused by AIs applied, for instance, to the medical context: that is, related to “over-reliance” on the machine and leading to the loss of one’s skills. Therefore, once again it seems related to forms of fiduciary delegation often rendered irreversible, fostering our “laziness” in letting machines handle many of our activities, which turn – in a narrative-discursive point of view – from “delegated helpers” to a kind of “magic helpers”: to whom we tend to delegate many of our activities.

A final, third level, closely intertwined with the previous ones, pertains to the pragmatic-operational and discursive dimensions of AI – specifically, the communicative and political-discursive aspects of AI systems. This is what some have termed the “prompt culture.” This term refers, it is well known, to the ways in which users interact with AI systems, particularly conversational models like ChatGPT, but also others, such as Midjourney for image generation. These systems, operating in a chat-based format, facilitate both research and content creation through conversational interfaces. Central to this interaction are “prompts,” natural

language commands that guide the AI's responses. The concept and practice of "prompting" – as everybody knows, and as briefly mentioned earlier – has rapidly evolved into a specialized skill and even a profession (prompt designers). It has spurred a burgeoning ecosystem on social media, encompassing web guides, YouTube tutorials on crafting effective prompts, and strategies for optimizing communication with AI systems. This emerging culture includes applications and plugins designed specifically for ChatGPT, aimed at enhancing user-machine interactions through precise, clear, and effective prompts.

In broader terms, this phenomenon highlights the expansive scope of AI's communicative and discursive-political dimension. On one level, it shows how AI is reshaping discursive-textual practices that are increasingly permeating daily life. On another level, it connects these practices to a wider cultural and societal framework, infused with "mythical" narratives – anthropologically speaking, no less real for their mythical quality. These narratives not only influence how society at large perceives AI but also have significant implications for political communication and public debate. In this sense, AI's impact extends beyond its technical functionalities to deeply affect the way we construct and navigate our social and political realities.

7 What remains: bodies, resistance, and re-launches

Finally, a critical question remains: what about the role of bodies? In the context of the extensive reasoning and the invention of novel "world models" by machines (see Beckmann et al. 2023), the bodily¹⁵ and perceptual-affective dimensions appear largely overlooked – beyond the critiques raised by scholars such as Dreyfus (1992), as we have seen, and others. This is noteworthy, given the attention these dimensions have received in recent decades across the social and cognitive sciences, culminating in a revolution marked by the advent of theories of embodiment and embodied cognition.

Current research suggests that for AI systems and models to become "truly" intelligent – or perhaps just "more" intelligent – they must develop a form of "self-simulation" or self-perception. This involves the capacity to "imagine" themselves as embodied entities. For example, robots equipped with AI may need to

¹⁵ Concerning the problem of the relationship between models of AIs and the issue of embodiment, see, in addition to the seminal questions posed by the aforementioned Dreyfus, Chrisley (2003) for a deepening concerning the beginnings of AI research and debate on embodiment; and Hellström et al. (2024), for more recent developments and considerations, also concerning a typology of forms of embodiment and relative "agents."

construct an internal image of themselves to achieve cognition and a nascent form of “self-awareness.” Spinoza’s question, famously echoed by Deleuze, “What can a body do?” resonates here. Today, the body remains both a simulacrum and a frontier – something to be reimagined and reconquered.

This notion aligns again with Deleuze and Guattari’s (2010) proposition of a “new machinic unconscious” or a “new phylum of machines” that might begin to think and imagine themselves autonomously. In this sense, AI’s potential self-perception could reflect the emergence of a novel form of machinic subjectivity, indirectly linking to Esposito’s (2022) hypotheses on the evolving interplay between technology, agency, and identity. But, this seemingly tangential point offers a metaphor tied to the bodily dimension, providing a critical lens for reconsidering political communication. It reminds us that there are always “residual” zones – areas of resistance or friction – even in the realm of advanced communication technologies. In the case of AI, these zones may represent the spaces where new forms of “hybrid and artificial social cognition” meet resistance to the utopian promises of AI’s “glorious sun of the future.” However, this is not to align with a purely Chomskyan critique – which, while highlighting concrete risks, appears partially outdated – but to address another issue: much of today’s AI development is still grounded in “extractivist” practices (see Pasquinelli 2023). These technologies, despite the good intentions of many researchers and startups that often begin with “open” ideals, are eventually funded or acquired by major corporations. They rely on the exploitation of billions of textual objects, drawing from diverse sources such as literary works, news articles, blogs, websites, and social media content.

The intent here is not merely to “denounce” these practices but to acknowledge their nature and explore their implications. For example, some artist collectives and groups embed code elements into their works – be it digital art or other media – that disrupt AI systems when used for training. This practice (of “poisoning,” see Heikkilä 2023),¹⁶ far from being a simple artistic provocation, might be seen as a form of digital neo-Luddism. Yet, it raises critical questions about the ethics and politics of communication. On one hand, AI opens up vast opportunities for freedom, innovation, and efficiency in textual and visual production. It has demonstrated remarkable potential for automated critical and analytical work on massive repositories of information and images, as evidenced by its use in groundbreaking journalistic investigations. Such advancements are already transforming how communication is analyzed, critiqued, and produced, reshaping the routines of communicative work.

¹⁶ Cf Heikkilä’s 2023 article “This new data poisoning tool lets artists fight back against generative AI.” The tool, called *Nightshade*, messes up training data in ways that could cause serious damage to image-generating AI models.

At the same time, it is crucial to remain vigilant against the emergence of new forms of exploitation and “capture” inherent in AI systems. The challenge lies in ensuring that AI does not devolve into a tool of servitude but instead becomes a platform for reimagining communicative labor – particularly from a semiotic perspective – focusing on the production of sense and meaning. This moment calls for a renewed commitment to the original and critical role of the “political” in political communication: recognizing and understanding the new actors and hybrid subjects that are reshaping the public sphere. By embracing this perspective, we can begin to transform AI into a space for collective reinvention and innovation, rather than allowing it to be confined to yet another domain of extraction and control.

References

- Beckmann, Pierre, Guillaume Köstner & Inês Hipólito. 2023. An alternative to cognitivism: Computational phenomenology for deep learning. *Minds and Machines* 33(3). 397–427.
- Chomsky, Noam. 2023a. The false promise of ChatGPT. *New York Times*. <https://www.nytimes.com/2023/03/08/opinion/noam-chomsky-chatgpt-ai.html> (accessed 26 December 2024).
- Chomsky, Noam. 2023b. Noam Chomsky speaks on what ChatGPT is really good for. *Common Dreams*. <https://www.commondreams.org/opinion/noam-chomsky-on-chatgpt> (accessed 26 December 2024).
- Chrisley, Ron. 2003. Embodied artificial intelligence. *Artificial Intelligence* 149. 131–150.
- De Baggis, Mafe & Alberto Puliafito. 2023. *In principio era ChatGpt*. Milano: Apogeo.
- Douglas Heaven, Will. 2023. Geoffrey Hinton tells us why he’s now scared of the tech he helped build. *MIT Technology Review*. <https://www.technologyreview.com/2023/05/02/1072528/geoffrey-hinton-google-why-scared-ai/> (accessed 14 December 2024).
- Douglas Heaven, Will. 2024. What is AI? Everyone thinks they know but no one can agree. And that’s a problem. *MIT Technology Review*. <https://www.technologyreview.com/2024/07/10/1094475/what-is-artificial-intelligence-ai-definitive-guide/> (accessed 14 December 2024).
- Dreyfus, Hubert. 1992. *What computers (still) can’t do: A critique of artificial reason*. Cambridge, MA: MIT Press.
- Dreyfus, Hubert. 2007. Why heideggerian AI failed and how fixing it would require making it more heideggerian. *Artificial Intelligence* 171(18). 1137–1160.
- Eposito, Elena. 2022. *Comunicazione artificiale: Come gli algoritmi producono intelligenza sociale*. Milano: Egea.
- Greimas, Algirdas Julien. 1987. *De l'imperfection*. Périgueux: Fanlac.
- Greimas, Algirdas Julien & Joseph Courtés. 1993 [1979]. *Sémiotique. Dictionnaire raisonné de la théorie du langage*. Paris: Hachette.
- Guattari, Félix. 2010. *The machinic unconscious*, Taylor Adkins (trans.). New York: Semiotext(E).
- Hellström, Thomas, Niclas Kaiser & Suna Bensch. 2024. Taxonomy of embodiment in the AI Era. *Electronics* 13. 4441.
- Heikkilä, Melissa. 2023. This new data poisoning tool lets artists fight back against generative AI. *MIT Technology Review*. <https://www.technologyreview.com/2023/10/23/1082189/data-poisoning-artists-fight-generative-ai/> (accessed 14 December 2024).

- Merleau-Ponty, Maurice. 1945. *Phénoménologie de la perception*. Paris: Gallimard.
- Norvig, Peter. 2017. On Chomsky and the two cultures of statistical learning. <https://norvig.com/chomsky.html> (accessed 14 December 2024).
- Pasquinelli, Matteo. 2023. *The eye of the master: A social history of artificial intelligence*. London: Verso.
- Resnik, Phillip. 2024. Large language models are biased because they are large language models. *Arxiv*. <https://arxiv.org/pdf/2406.13138v1> (accessed 14 December 2024).
- Ribes, David. 2018. STS meet data science once again. *Science Technology & Human Values* 44(3). 514–539.
- Shah, Hardik. 2023. 4 types of hallucinations. *Medium*. <https://hardiks.medium.com/4-types-of-ai-hallucinations-9f87bdaa63e3> (accessed 14 December 2024).
- Svetlova, Ekaterina. 2021. AI meets narratives: The state and future of research on expectation formation in economics and sociology. *Socio-Economic Review* 20(2). 841–861.
- Thiollet, Aymeric. 2024. *Hallucinations in AI: Fatality or opportunity?* Human Technology Foundation. <https://www.human-technology-foundation.org/news/hallucinations-in-ai-fate-or-opportunity> (accessed 14 December 2024).
- Vendeville, Benjamin, Liana Ermakova & Pierre de Loor. 2024. Le problème des hallucinations dans l'accès à l'information scientifique fiable par les LLMs: verrous et opportunités. In *CORIA 24: Conférence en Recherche d'Information et Applications, April 3–4, La Rochelle*. http://coria.asso-aria.org/2024/articles/position_31/main.pdf (accessed 14 December 2024).
- Zhang, Yue, Yafu Li, Leyang Cui, Deng Cai, Lemao Liu, Tingchen Fu, Xinting Huang, Enbo Zhao, Yu Zhang, Yulong Chen, Longyue Wang, Anh Tuan Luu, Wei Bi, Freda Shi & Shuming Shi. 2023. Siren's song in the AI ocean: A survey on hallucination in large language models. *Arxiv.org*. <https://arxiv.org/pdf/2309.01219.pdf> (accessed 14 December 2024).