P&D PHILOSOPHY & DIGITALITY
AN OPEN ACCESS, PEER-REVIEWED JOURNAL

# Beyond Artificial Intelligence: How Algorithms Communicate without Understanding

Elena Esposito

**Abstract:** The remarkable performance of recent algorithms grounded in machine learning and Big Data is often portrayed as a milestone in Artificial Intelligence (AI), suggesting a replication of human cognitive processes by machines. This article challenges such interpretations, arguing that systems like ChatGPT excel not by achieving intelligence akin to humans, but by generating outputs that can be used by humans to create their own relevant information. Without understanding content, algorithms have learned to participate in communication. The evolving interaction between humans and such technologies is likely to significantly influence the future of intelligence. However, understanding these effects requires shifting focus away from direct comparisons and competitions between human and machine cognitive abilities, toward exploring their complementary roles in communication.

**Keywords:** Generative AI; artificial intelligence; artificial communication; machine learning; singularity

## The intelligence of algorithms

1 What does intelligence have to do with recent developments in Artificial Intelligence (AI)? Or put another way: when we talk about AI, is it intelligence we are talking about?

2 Apparently, that seems to be the point, as the term itself indicates. The debate on Artificial Intelligence is recently focused on generative AI, and the lively discussion on these new technologies is primarily about the comparison/competition between human intelligence and alleged machine intelligence. Generative AI may be feared or welcomed, but many are convinced that it constitutes "profound change in the history of life on Earth"—as stated, for example, in



Contact: Elena Esposito <elena.esposito@uni-bielefeld.de>
https://doi.org/10.18716/pd.v2i1.11657

the alarming open letter signed by thousands of researchers worldwide (Future of Life Institute 2023). This change is traced to the fact that "we have summoned an alien intelligence" (Harari, Harris, and Raskin 2023), a non-human mind that could achieve a non-human intelligence, with outcomes yet to be discovered.

3 It may be, as many fear, the arrival of the "singularity" (Kurzweil 2005) or a "superintelligence" (Bostrom 2014) that may surpass human capabilities and take over. There is concern that an increasingly autonomous alien intelligence may empower itself and follow its own priorities and criteria, different from those of human programmers, to the point of possibly threatening humanity's existence. Others speak rather of augmented intelligence (Rheingold 1985) as collaboration between natural intelligence and artificial intelligence, where technologies support and assist human capabilities in processing data and performing various tasks. This would increase efficiency and lead to better outcomes; a much-cited example are developments in precision medicine where advanced AI systems make it possible to analyze genomic data, identify variations, find patterns, and propose therapies that could not be achieved without their input.

4 In all these discourses, comparisons are made, alarmed or confident, between more or less different and more or less competent types of intelligence. This article argues that it is precisely this comparison that is misleading. It implies the idea, which has accompanied research on automatic information processing since its inception, of reproducing with machines (artificially) the forms of human intelligence. Today, however, it is becoming evident that imitating or simulating human thought processes with digital tools is no longer the goal of recent programming techniques. Algorithms explicitly and intentionally work differently from the human mind, and I maintain that this is precisely why in the last 10 to 15 years they are achieving the results we all observe with wonder—and often awe. What are these successes due to?

5 In coincidence with the new "spring" of AI projects, there has been a profound shift in project design related to two interconnected innovations. The first is the enormous progress in *machine learning*. In itself, machine learning is nothing new; it has been talked about since long (Nilsson 2010) and with mixed success, but today algorithms seem to be capable not only of learning, but of learning on their own: they decide for themselves what to learn and how. This is called unsupervised machine learning and deep learning: Techniques in which machines learn to perform tasks in ways that were not intended by their programmers and that in some cases are incomprehensible to humans, including those who designed the algorithms (Goodfellow, Bengio, and Courville 2016). In many cases, even the programmers do not understand how the machine proceeds and how it achieves its outcome (Burrell 2016).

6 The second novelty, without which the first could not be accomplished, are the so-called *Big Data*: the enormous amount of diverse data that are available today for machine operations (Mayer-Schönberger and Cukier 2013). The definition of Big Data is complicated and controversial (see, e.g., (Kitchin 2014)), but there is general agreement that this new type of data is related to the spread of Web 2.0—the so-called participatory web that has emerged over the past 15 to 20 years. It involves a series of innovations in programming technologies, which have

made it possible to create more dynamic web pages that are open to the contributions and interventions of their users—who participate directly in the creation of content. This has led to an enormous multiplication of available data. Each of us produces a great deal of it; we do it with our behavior, with the browsing and choices we make on the web, with our participation in social media, but also with all the data collected by GPS services or the so-called Internet of Things. It is heterogeneous, unsupervised, unselected data, often biased in many different ways (O'Neil 2016)—but it is the data that algorithms feed on to perform the tasks that appear intelligent to us, as ChatGPT and other forms of generative AI show in striking ways.

7 The result of these innovations is that now the algorithms that use them seem to become intelligent. They are capable of doing more and more things and doing them better and better, and in particular they know how to perform tasks that were previously the prerogative of humans, who are indeed intelligent. Algorithms are able to have conversations, provide complex information, write texts that are usually competent and appropriate (apart from so-called hallucinations), compose music, and create completely realistic images and videos from natural language descriptions. But if one looks at how programs work, one sees that these amazing performances have become possible not because machines have finally become intelligent, but rather because they have given up trying to become intelligent. Back to our initial question: It is not intelligence that is involved.

8 Indeed, as programmers have long recognized, recent algorithms based on machine learning and Big Data do not try to imitate the forms of human intelligence—that would be too heavy a burden, and they do not need to (Borgo 2020). They do something completely different, looking for patterns and regularities in the huge amount of text and materials they have at their disposal, and using them to produce text or images that make sense to their users—but not to the algorithms, which understand nothing of the content they process. These machines succeed because they have learned to use human intelligence autonomously, even without understanding it. Recent algorithms are programmed to use human input at various stages of their processes to structure and direct their own behavior. Algorithms reprocess the results of users' intelligence and present them back to them in a way that is surprising, appropriate, and informative (Esposito 2017).

9 How does it happen? ML algorithms are able to compute, combine, and process with surprising efficiency the information they find in data, but they are not able to produce it on their own. Algorithms "feed" on the information and cues generated (consciously or unconsciously) by individuals and their behavior to produce new, surprising, and potentially instructive information. Human contributions intervene in at least three phases of the operation of algorithms: in producing data for training, in tuning their behavior, and in generating the feedback (reinforcements) that direct it (Esposito 2024 ch. 8). Thus, they are able to respond intelligently to user requests, without having to be intelligent themselves.

10 None of us process information in this way. An often-cited example is machine translation programs, which work very well today, but this has been happening since programmers stopped

trying to teach algorithms the different languages, with their grammars and rules. Today algorithms translate texts from Chinese without knowing Chinese, and neither do their programmers. They just look for regularities in the vast amounts of multilingual materials they find in their training data, and use them to produce texts that the machines do not need to understand, but that are informative to the human users who requested them.

11 Even the most vocal critics of Generative AI recognize this. Noam Chomsky, for example, who wrote a vibrant indictment of the false promises of ChatGPT, notes that these programs "differ profoundly from how humans reason and use language" (Chomsky, Roberts, and Watamull 2023)—but for him, as for many others, this diversity is evidence of the stupidity of machines. In my view, however, the difference from human forms of processing is not a weakness of these technologies, but the very root of their success. Machines are neither smart nor stupid; they are different. The results of ChatGPT and the like are not the realization of the dream (or nightmare) of artificial intelligence, but a signal that we need a different approach. This is why I propose to move from the idea of artificial intelligence to the exploration of new forms of artificial communication (Esposito 2022).

## The mysteries of intelligence

12 This change of approach can be helpful first of all because it is still not at all clear what intelligence is. After millennia of analysis and reflection, we do not have a shared definition of intelligence. The more we study it, the more we discover different forms of intelligence, from the logical to the practical (Sternberg 1985), to the emotional, to the musical or linguistic, to (Gardner 1983)'s influential proposal of a multiplicity of intelligences distinct from each other. What is striking about all these intelligences is how diverse they are, making the unifying notion (the idea that they are all intelligences) increasingly elusive.

13 It does not get simpler if we decide to refer to the organic level, overcoming the much-discussed mind/brain dualism (Dennett 1992). Neurophysiological research recognizes that the very relevant advances that have been made in recent years show that many aspects of brain functioning are still unknown. We still know little about the connectivity of the brain, the flows of information within it, about neuronal coding, about the precise mechanisms of memory formation, storage, and retrieval—and especially about the organic nature of consciousness. Also, the reference to neural networks in recent machine learning algorithms is more a suggestion than the precise reproduction of the processes of the brain.

14 Whether one refers intelligence to the mind or to the brain, it is in any case still a mysterious object. The idea of artificially reproducing something we know so little about is indeed strange. About communication, on the other hand, we already know a lot, and this is certainly an advantage. We know, for example, how communication has changed over the centuries and as society has evolved, from only interaction between people who are co-present to increasingly flexible

and inclusive forms that also allow us to communicate with partners distant in space and time, gradually becoming more inaccessible, anonymous and impersonal. Communication by voice and gesture has been joined by handwriting and then print, transmission of sound and of still, moving and even three-dimensional images. That the forms of communication change is not new, nor is it an enigma. Rather, it is a matter of identifying and understanding the differences and continuities between familiar forms of communication and the still mysterious ones that are made possible by technological innovation.

15 The reference to communication makes it possible to observe the ongoing evolution in a more realistic and informative way than the reference to intelligence—even and especially since we do not know exactly what we mean by intelligence. Whatever intelligence is, machines have not learned to become intelligent, but to do something else: they have learned to participate in communication. The reason we tend to attribute to LLMs an intelligence of their own is that users who interact with them get in response a communication that is tailored precisely to their request, the circumstances, and the context of the query. In many cases no one had previously thought of it and formulated it in that precise form—it is in a sense an autonomous creation of the algorithm. In this setting, then, the communicative partner is not the one or more human beings who author of the materials being accessed. The partner is directly the algorithm that provides the response. One communicates *with* the algorithm. And since until now communicative responses had always been provided by humans on the basis of intelligence, one tends to attribute also to the algorithms that generate the communications a specific form of intelligence, albeit artificial—called in fact generative AI.

## The communication of algorithms

16 Of course, we must now address what is meant by communication. As it is usually understood, the notion of communication requires that the mental processes of the participants converge on a common content. Referring to the Latin root of the term communication (*communicatio*), communication is assumed to take place when the partners arrive at sharing a thought, or at least a part of it. A communication is then said to have taken place if at the end of the process the receiver gets at least some of the information that the sender had in mind and put into the communication channel. Even allowing for noise and differences in encoding/decoding, interpretation, and expertise, the idea remains in this understanding of communication that in a successful communication some element of the identity of the information must be preserved[1]. Dealing with interaction with machines, however, the problem is that we face a situation where one of the communication partners is an algorithm that does not understand content, meaning

1. Shannon and Weaver's model of communication transmission is still the basis (revised and supplemented) for the notion of communication in most sociological and semiotic approaches ((Shannon and Weaver 1949); (Fiske 1990); (Eco 1975, 65-69)).

or interpretation. It only deals with data. The user, therefore, does not share any information (even in part) with the partner, because the partner does not know any information. Can we still say that this is communication? Are we dealing with an "aberrant" condition[2] or with an unprecedented form of communication?

17 I refer here instead to a different concept of communication, proposed by Niklas Luhmann, according to which participation in communication does not presuppose the sharing of thoughts among participants (Luhmann 1997, 92). For example: If the communication conveyed by this article succeeds, it is not because the readers think the same things as I who wrote it. Fortunately, readers do not need to have access to my thoughts, which are and remain mine alone. The communication succeeds if each of them uses what is written in the article to produce their own thoughts and information, which depend on their own history, interests, and specific perspective, and are inevitably different from mine and anyone else's. The thoughts of each of the participants in the communication are his or her own alone, triggered by what is said or written in the communication, and not a reproduction (maybe partial) of the thoughts of the speaker/writer. In this view, even algorithms, which do not think, can serve as communication partners if they have the ability to produce contributions that enable their users (us intelligent human beings) to produce their own relevant and interesting information by participating in the communication. Algorithms do not need to think and understand information to enable us to produce meaningful information and circulate it in communication.

## Social effects of generative AI

18 What can we say about ChatGPT and similar algorithms if we shift the focus from intelligence to communication? Obviously, communicating with algorithms is not the same as communicating with intelligent human beings. The artificial partner is not an alter-ego[3], empathy, when it exists, is only simulated, the algorithm has no direct access to the world and is not able to perceive independently. But observing the use and effects of generative AI from a communicative perspective can lead to insights that we do not get if we deal with them as autonomous forms of intelligence. Artificial communication is different from the communication we are familiar with, which now becomes "natural" and can be observed as such. Comparing the two forms can enable us not only to understand more adequately the use of recent technologies, but also to observe our society and its communications from a new perspective.[4]

19 First of all, some concerns become void. Concerns about generative AI as an alien mind that could make us obsolete and replace our intelligence, in this view appear groundless—and with it

2. In the sense of semiotics' aberrant coding: cf. (Eco and Fabbri 1965).

3. Despite the different variants of the Eliza effect: (Weizenbaum 1976).

4. Cf. for example, the symposium (Pilati, Munk, and Venturini 2024).

all the imaginations about machines rebelling, following their desires and their will to be autonomous and powerful. But this does not mean for sure that there is nothing to worry about. Generative AI is not directly intelligent but intervenes in communication, and we have to be very careful about its effects. If algorithms work differently from the human mind and often incomprehensibly, how can we control their results, knowing also that they are inevitably biased in many different ways (boyd and Crawford 2012)?

20 Even if machines—which do not think—have no desires or preferences of their own, in fact, that does not make their results objective. The functioning of the algorithms that work with Big Data inevitably depends on the perspective and inclinations of the programmers who designed them, and also—in an even more elusive and uncontrollable way—on all the biases implicit in the behavior of users on the Web, which are inherited from the data with which the machines work. Bias is inevitable, not least and precisely because the machines themselves have no perspective of their own. And of course, the processing of algorithms can be used by users in the most diverse, even malicious and harmful ways—and machines per se cannot resist it.

21 First of all, the spread of generative AI produces much-discussed (and very difficult) *alignment* problems: How do we get learning algorithms to do what we want, in the way we want them to do it? (Khamassi, Nahon, and Chatila 2024). If systems obey us without understanding what we have in mind, there is always the danger that they will be misaligned, i.e., accomplish their assigned tasks in a way that is different from what the user intended or approves. Asked to design a way to drastically reduce workplace injuries, for example, the system might suggest shutting down all factory operations permanently—or it may respond to requests for suggestions to increase the stars rating of a service with the advice of coercing or bribing customers, or creating fake positive reviews.

22 Then there are serious concerns about the effects of algorithms on public debate and democracy —the main concern that led the "godfather of AI," Geoffrey Hinton, to leave Google, warning that generative AI poses a threat to humanity. Tools such as Dall-E, Stable Diffusion and Midjourney can create, from natural language descriptions, realistic images of objects or people that do not exist, or even completely believable images and videos of people who exist but never did or said the things being shown. There is a widespread and justified fear that generative AI could be used to easily, quickly, and cheaply produce and disseminate news that appears plausible but is actually baseless (deepfakes).

23 These algorithms compel us to confront novel ways of communicating, producing new information from available information, and they require specific controls and criteria—which are not yet available. The European Community's AI Act intends to give guidelines in this regard, but it is a very general draft with many uncertain points. Private organizations are also taking action. The *Content Authenticity Initiative* project[5], supported by major companies such as Adobe, Photoshop and the New York Times (and many others), has the stated goal of producing (and

5. https://contentauthenticity.org/

potentially mandating) "content credentials" to be included in images to reveal that they were generated by AI. Groups of researchers at MIT and Columbia University, moreover, are working on AI systems that do not just give answers, but help people develop specific critical thinking skills (Heikkilä 2023).

24 If, as seems certain, the intelligence of each of us depends on the type of communication we are exposed to and the stimuli it offers us, interaction with nonintelligent machines will also have profound effects on intelligence development. In order to grasp them, however, we should abandon the comparison and competition between the cognitive abilities of humans and those of machines.

## References

Borgo, Stefano. 2020. "Ontological Challenges to Cohabitation with Self-Taught Robots." *Semantic Web* 11 (3): 161–67.

Bostrom, Nick. 2014. *Superintelligence: Paths, Dangers, Strategies*. Oxford: Oxford University Press.

boyd, danah, and Kate Crawford. 2012. "Critical Questions for Big Data." *Information, Communication and Society* 15 (5): 662–79.

Burrell, Jenna. 2016. "How the Machine 'Thinks': Understanding Opacity in Machine Learning Algorithms." *Big Data & Society* 1:1–12.

Chomsky, Noam, Ian Roberts, and Jeffrey Watamull. 2023. "The False Promise of ChatGPT." *The New York Times*, March 8, 2023.

Dennett, Daniel. 1992. *Consciousness Explained*. New York: Back Bay Books.

Eco, Umberto. 1975. *Trattato Di Semiotica Generale*. Milano: Bompiani.

Eco, Umberto, and Paolo Fabbri. 1965. *Prima Proposta per Un Modello Di Ricerca Interdisciplinare Sul Rapporto Televisione/Pubblico*. Perugia: Mimeo.

Esposito, Elena. 2017. "Artificial Communication? The Production of Contingency by Algorithms." *Zeitschrift Für Soziologie* 46 (4): 249–65.

———. 2022. *Artificial Communication: How Algorithms Produce Social Intelligence*. Cambridge: MIT Press.

———. 2024. *Kommunikation Mit Unverständlichen Maschinen*. Wien: Residenz Verlag.

Fiske, John. 1990. *Introduction to Communication Studies*. London: Routledge.

Future of Life Institute. 2023. "Pause Giant AI Experiments," March 22, 2023.

Gardner, Howard. 1983. *Frames of Mind: The Theory of Multiple Intelligences*. New York: Basic Books.

Goodfellow, Ian, Yoshua Bengio, and Aaron Courville. 2016. *Deep Learning*. Cambridge: MIT Press.

Harari, Yuval, Tristan Harris, and Aza Raskin. 2023. "You Can Have the Blue Pill or the Red Pill, and We're Out of Blue Pills." *The New York Times*, March.

Heikkilä, Melissa. 2023. "A Chatbot That Asks Questions Could Help You Spot When It Makes No Sense." *MIT Technology Review*, April.

Khamassi, Mehdi, Marceau Nahon, and Raja Chatila. 2024. "Strong and Weak Alignment of Large Language Models with Human Values." *Scientific Reports* 14:19399.

Kitchin, Rob. 2014. "Big Data, New Epistemologies and Paradigm Shifts." *Big Data & Society* 1 (1).

Kurzweil, Ray. 2005. *The Singularity Is Near*. New York: Viking Books.

Luhmann, Niklas. 1997. *Die Gesellschaft Der Gesellschaft*. Frankfurt/M.: Suhrkamp.

Mayer-Schönberger, Viktor, and Kenneth Cukier. 2013. *Big Data. A Revolution That Will Transform How We Live, Work, and Think*. London: Murray.

Nilsson, Nils John. 2010. *The Quest for Artificial Intelligence. A History of Ideas and Achievements*. New York: Cambridge University Press.

O'Neil, Cathy. 2016. *Weapons of Math Destruction*. New York: Crown.

Pilati, Federico, Anders Kristian Munk, and Tommaso Venturini. 2024. "Repurposing Generative AI for Social Research." *Sociologica* 18 (2).

Rheingold, Howard. 1985. *Tools for Thought: The History and Future of Mind-Expanding Technology*. Cambridge: MIT Press.

Shannon, Claude Elwood, and Warren Weaver. 1949. *The Mathematical Theory of Communication*. Urbana: University of Illinois Press.

Sternberg, Robert. 1985. *Beyond IQ: A Triarchic Theory of Intelligence*. Cambridge: Cambridge University Press.

Weizenbaum, Joseph. 1976. *Computer Power and Human Reason*. New York: W. H. Freeman and Company.