Roberto LAGHI
Avignon Université

From Human to Binary and Back: On the Need to Explain and Understand Digital Machines in the Humanities

Human-Centred AI in the Translation Industry. Questions on Ethics, Creativity and Sustainability

Katharina Walter,
Marco Agnetta
[eds.]

5.1/2024
Yearbook of Translational Hermeneutics
Jahrbuch für Übersetzungshermeneutik

Journal of the Research Center
Zeitschrift des Forschungszentrums

Hermeneutics and Creativity, University of Leipzig
Hermeneutik und Kreativität, Universität Leipzig

DOI: 10.52116/yth.vi1.103

Cite this article:
Laghi, Roberto (2025): "From Human to Binary and Back: On the Need to Explain and Understand Digital Machines in the Humanities." In: *Yearbook of Translational Hermeneutics* 5.1: *Human-Centred AI in the Translation Industry. Questions on Ethics, Creativity and Sustainability* (ed. by Katharina Walter, Marco Agnetta), pp. 117–135. DOI: <10.52116/yth.vi1.103>.

Yearbook of Translational Hermeneutics 5.1/2025
ISSN: 2748-8160 | DOI: 10.52116/yth.vi1.103

From Human to Binary and Back: On the Need to Explain and Understand Digital Machines in the Humanities

Roberto LAGHI
Avignon Université

Abstract: This article aims to bring attention to some usually overlooked aspects of the relationship between humans and complex digital technologies. Before engaging with artificial intelligence (AI), it is indeed pivotal to address some key questions about it. Specifically, I will try to focus on our ability to understand how AI technologies work and determine creative and critical uses we can make of them. To do so, I will first discuss problems associated with using the current definitions of AI and suggest that we should make a creative effort to re-translate these terms in order to find better-suited expressions. I will call attention to the need for a different kind of translation, which negotiates between what machines do and what we can understand about them, because one of the biggest challenges of machine learning is to make the internal processes explainable and understandable for us humans. I will close with elaborations on some creative forms of interaction with language models and image models which support artists, writers and creators (who do not want to see their work stolen by AI crawlers and used to train datasets), with the overall goal of building an ethical, critical and sustainable relationship between humans and digital machines.

Keywords: Artificial intelligence, AI ethics, Explainability, Data poisoning, Anti-computing.

1 Introduction: On Artificial Intelligence and its Possible Definitions

In this article I will not address specific issues on translation and artificial intelligence (AI) but rather bring the focus to some aspects of AI and of our relationship with it, aspects that I consider foundational for developing a digital hermeneutics able to engage the humanities (and thus translation studies as well) in a deeper understanding of digital tools that are already in common use. This is why, throughout this article, I will use the word 'translation' in a broader sense, which includes forms of encoding and decoding, and this is because of the cultural and social impact of these forms (cf. Laghi 2023: 52).

Digital technologies are ubiquitous, and they always involve writing (at least as computer code) and translation (between computer code and human language, or between different languages). If we consider every interaction with digital technologies as an act of writing and translation, we can understand how our use of devices based on AI is always an act of writing and translation as well. This is why I would like to start my argument with an effort to challenge the current use of the expression "artificial intelligence," a "magic" marketing catchphrase that creates confusion about what this technology is capable of and about the actual dangers it entails, and I intend to do so as an experiment in translation, to help us better define some of the key concepts related to AI. Although the expression 'AI' is already widely accepted and has entered common usage (even in academia), I still believe that, as scholars, we need to be very careful when we define new technologies, because how we define them can change the way we perceive them, thus influencing the way we critically think about them.

To make a provocative point, what if, instead of AI, we talked about "SALAMI"? This acronym was created by the

Italian tech journalist Stefano Quintarelli in 2019, and it stands for "Systematic Approaches to Learning Algorithms and Machine Inferences" (Quintarelli 2019). If we use the acronym SALAMI instead of AI, the questions that we often hear in the AI debate seem to lose meaning completely, as Quintarelli provocatively asks:

> Will we still support the idea that SALAMI will develop some form of consciouness [sic]? Will SALAMI have emotions? Can SALAMI acquire a "personality" similar to humans'? Will SALAMI ultimately overcome human limitations and develop a self superior to humans? Can you possibly fall in love with a SALAMI? Can we suddenly perceive a sense of how all these far flung (unrealistic) predictions look somewhat ridiculous? (Quintarelli 2019)

Of course, Quintarelli's questions are just a provocative invitation to think about our relationship with technology, starting from how we define it, and about the supernatural/magical approach we often have (cf. Mohamed 2022; Dalmasso 2020: 177).

The main point still remains, however, that we do not have an accepted definition of AI and, even more importantly, we do not have a shared and accepted definition of intelligence or, better, "we have a lot of partial definitions [of intelligence], all of which are bound to specific contexts" (Loukides 2022). The philosopher Luciano Floridi argues that "AI is better understood as a new form of agency, not intelligence," which prompts him to say that "the digital has changed the nature of agency, but we are still interpreting the outcome of such changes in terms of modern mentality, and this is generating some deep misunderstanding" (Floridi 2023: XIII, 10). He then argues that the success (and usefulness) of AI lies in the "decoupling [of] the ability to solve a problem or complete a task successfully from any need to be intelligent to do so" (Floridi 2023: 12). It is true that AI relies on statistics and probability and not on understanding: As we know, large language

models (LLMs) and other automatic translation tools do not understand the meaning of the texts they are prompted with and their outputs, but they still manage to produce results that are often satisfactory or, at least, good enough. If this happens, however, Floridi argues that it is because the world is becoming an infosphere better and better adapted to what AI can do:

> If drones, driverless vehicles, robotic lawnmowers, bots, and algorithms of all kinds can move "around" and interact with our environments with decreasing trouble, this is not because productive, cognitive AI (the Hollywood kind) has finally arrived. It is because the "around," the environment our engineered artefacts need to negotiate, has become increasingly suitable to reproductive engineered AI and its limited capacities. (Floridi 2023: 26)

However, if we consider definitions of intelligence such as the one proposed by Cristianini, that is, "the ability to behave effectively in novel situations" (Cristianini 2023: 116),[1] then the perspective changes completely because this definition also includes AI, as these systems are increasingly able to relate to novel situations by finding effective solutions.[2]

---

1 Cristianini proposes a definition that is in line with the one given by Max Tegmark: "ability to accomplish complex goals" (Tegmark 2017: 869; Cristianini 2023: 116).

2 It does not make sense, for the purpose of this article, to enter the debate about the concept of intelligence, because it would take us too far. We can, however, underline the need for a change in our anthropocentric view, in order to take into account different forms of intelligence, cognition and agency. Moreover, Cristianini stresses how important it is "to give up the illusion that we, human beings, are the paragon of all intelligent things, an illusion that is hindering our understanding of the world" (Cristianini 2023: 293). We are faced with the need for a paradigm shift, and the need for this shift comes from the advent of digital technologies as a fundamental part of our lives, our relationship to knowledge, to reality and to the way we experience it. Likewise, it is also appropriate to develop new critical capacities and new logics of resistance to explore "the configurations of meaning

We can see that the issue is not resolved, and better-defining expressions could be helpful to understand the technologies we are dealing with. Floridi (2023: 14) says that "the absence of a definition for AI is evidence that the expression is not a scientific term. Instead, it is a helpful shortcut for referring to a family of sciences, methods, paradigms, technologies, products, and services." AI is first and foremost a marketing expression or, as Lanier/Weyl (2020: n.p.) say, it "is an ideology, not a technology."

The expression was created in 1956 by John McCarthy for the organization of a conference, the Dartmouth Summer Research Project on Artificial Intelligence. Before that date, the same field was referred to as 'automata studies' and the reason for this change was due to the fact that McCarthy wanted "to escape association with 'cybernetics'" (McCarthy 1996: 73) but also, apparently, because using 'AI' instead of 'automata studies' made it easier to access research funding (cf. Hunger 2023). Moreover, we should not forget that "[a]ll artificial intelligence is built on the same foundation of code, data, binary, and electrical impulse. Understanding what is real and what is imaginary in AI is crucial" (Broussard 2018: 186). The distinction between what is real and what is imaginary passes through the use of appropriate definitions for the kinds of technologies we have, because the terminology that has been imposed on the public debate is deceiving and confusing. Whilst using more specific expressions and proposing concepts that better define what AI is and does, my use of the term 'AI' should be taken as a "helpful shortcut," as intended by Floridi (2023: 14).

made possible today by the unprecedented alliance between biology, philosophy, and cybernetics" (Malabou 2017: 30).

2 Translating AI Concepts to Better Define them

Instead of using AI, we could try to be more specific, also in order to avoid biological metaphors that foster a pernicious tendency toward anthropomorphizing these technologies. To highlight the difference between what is real and what is imaginary, we can start with the expression "machine learning," because

> [w]hen a machine 'learns', it doesn't mean that the machine has a brain made out of metal. It means that the machine has become more accurate at performing a single specific task according to a specific metric that a person has defined. This kind of learning does not imply intelligence. (Broussard 2018: 1754)

The ambiguity of the term 'AI' that I explained in these first paragraphs makes the debate about it more difficult, as Bender and Hanna point out:

> In one sense, it is the name of a subfield of computer science. In another, it can refer to the computing techniques developed in that subfield, most of which are now focused on pattern matching based on large data sets and the generation of new media based on those patterns. Finally, in marketing copy and start-up pitch decks, the term "AI" serves as magic fairy dust that will supercharge your business. (Bender/Hanna 2023: n.p.)

This is why Francis Hunger proposes to use a different terminology: Instead of AI, for example, "automated pattern recognition" seems more appropriate and correct to Hunger (2023: n.p.). Instead of "machine learning" he suggests the use of "machine conditioning"; instead of "deep learning," "deep conditioning"; instead of "neural network", "weighted network"; and "node" or "weight" instead of "neuron" (ibid.). We would probably not imagine an existential threat posed by "automated pattern recognition," nor would we ask questions about whether a "deep conditioning" system could gain con-

sciousness (as we have already seen with the provocative acronym SALAMI).

A better definition can lead us to a better understanding. But to better understand what these technologies do, it is not enough to define them properly. We also need to take a look at what happens inside them. Technologies are a human product. Therefore, we should be in a position to understand what they do, based on what assumptions and for what purposes. This means adopting a sociotechnical approach capable of investigating digital services and devices (such as LLMs and neural machine translation systems such as DeepL) not just in terms of their mere technical functions, but also for the sociocultural meaning they have in the context of our relationship with them and within society at large. The questions we ask about these technologies can also mirror the questions we ask about human beings as scientific progress transforms what we know. But, if understanding how the human brain works is indeed a challenge (and will still be for a long time), understanding what complex technologies do is pivotal for them to be implemented and used safely and ethically.

3 Inside the Black Box: Understanding 'the Machine'

As digital systems become more and more complex, the problem of explaining and understanding what they do and how they do it is increasingly crucial. Can we understand "the machine"? How can we translate its functioning in ways that humans can understand and trust? Cynthia Rudin (2019) has pointed out the importance of building interpretable machine learning models that are transparent and accurate instead of building new models which can try to explain what "black boxes" models do. Her hypothesis is convincing, especially in a

time when machine learning is applied to fields like criminal justice, medicine and finance. But we could add some insights from the humanities to her approach as a computer scientist in order to get a wider perspective on the issue. Floridi, whose work focuses on AI ethics, proposes five core principles for an ethical development of these systems: beneficence, nonmaleficence, autonomy, justice (these first four are commonly used in bioethics) and explicability (cf. Floridi 2023: 57). This last principle is "understood as incorporating both the *epistemological* sense of *intelligibility*—as an answer to the question 'how does it work?'—and in the *ethical* sense of *accountability*—as an answer to the question 'who is responsible for the way it works?'" (Floridi 2023: 57–58).

"How does it work" and "who is responsible for the way it works" are fundamental questions. When it comes to creative uses of LLMs or text-to-image models (TTIs), researcher and poet Allison Parrish argues that coders, artists and engineers are responsible for the outputs of models based on machine learning since digital machines cannot be considered responsible for their outputs (cf. Parrish 2021).[3] Understanding how AI systems work and establishing forms of accountability should be a strong imperative in computer science and in the humanities, and a combined approach to the problem could be greatly beneficial. This sociotechnical approach is made quite clear in the concept of "understandability," coined by David Berry, who argues that

> rather than providing descriptions purely from the domains of a formal, technical and causal model of *explanation* (dominant in the sci-

3 This is extremely important on another level: on one side, systems like ChatGPT would have been impossible to develop without (often not legally granted) access to copyrighted material to be trained on (cf. Milmo 2024), on the other, because copyright might not be applicable to AI-generated art (cf. Brittain 2023).

> ences), these technologies would benefit from critical approaches that take account of *understanding*, more common in the humanities and social sciences […]. (Berry 2023: §37)

The impact of AI is social and cultural in a broad sense before it is technical, and the same goes for explainability, but it is difficult to add explainability after these complex technologies have already been developed. The best option is probably to embed explainability/interpretability directly at the beginning of the development process, but this is in harsh contrast with what the companies developing AI are doing—making profits from "the intellectual property afforded to a black box" (Rudin 2019: 209)—and with the technologies that are already available to the public.

Many researchers work on the explainability issue, though, and a few tools have been developed which can help us take a look inside these tools and understand how decisions are taken and outputs are created. This includes the What-If tool developed by *Google* (Wexler et al. 2020) or *AI Explainability 360* (by IBM and now a Linux Foundation project), even if many of the LLMs built by companies like *Google*, *Meta* and *Microsoft* are not really open, in the sense commonly used for open source but also meaning open to independent investigations (cf. Gibney 2024). It is worth mentioning the research conducted by institutes such as *DAIR* (*Distributed AI Research*), whose work goes in the direction of building "frameworks for non-exploitative community rooted research practice" including, for example, datasets accountability (Khan/Hanna 2022) and Data&Society, especially on AI and algorithmic accountability and justice (cf. Moss et al. 2021). There are also more experimental efforts, such as the one by Thomas Parr and Giovanni Pezzulo, who draw on the theory of "active inference" to create a theoretical model of machine understanding "such

that, when queried, a machine is able to explain its behaviour" (Parr/Pezzulo 2021: 1).

While the research on explainability and interpretability continues, we need tools to engage critically and creatively with the AI models already available right now. We cannot limit ourselves to a passive use of these black boxes, unaware of the bias embedded in them and the real harm they are already causing (cf. Bender/Hanna 2023).

4 Critical Creative Engagement with AI

As an introduction to this part, I would first like to start with an observation that I found interesting: "Surprising errors are AI imagery's best approximation of genuine creativity, or at least its most joyful" (Herrman 2022: n.p.). Of course, AI has no imagery: the only imagery we are talking about is the one we, as humans, project on AI outputs. But looking for errors and glitches helps us to look at what happens behind the screen, inside these models. These errors can reveal some aspects of the functioning of LLMs or TTIs in unexpected ways because they show us these machines going wrong, as we can see in the research by Giuseppe Sofo about translating the city in the digital era (cf. Sofo 2025). I then would like to refer to a literary author to highlight an idea that should be taken into consideration whenever we interact with a language model or image model with a creative intent. Science fiction writer Ted Chiang, in a piece that appeared in the New Yorker, defines ChatGPT as "a blurry jpeg of the web" and he underlines the fact that using AI is not a good way to create original work because the process of writing itself is what enables the eventual creation of something original.

> Some might say that the output of large language models doesn't look all that different from a human writer's first draft, but, again, I think this is a superficial

> resemblance. Your first draft isn't an unoriginal idea expressed clearly; it's an original idea expressed poorly, and it is accompanied by your amorphous dissatisfaction, your awareness of the distance between what it says and what you want it to say. That's what directs you during rewriting, and that's one of the things lacking when you start with text generated by an A.I. (Chiang 2023: n.p.)

Even if Chiang does not speak for all writers, authors and artists, we may think that short-circuiting a large part of creative work to machines, whose functioning we essentially know nothing about, could not only impact human creativity capabilities, but also make us the weakest part of the relationship between humans and digital machines. Moreover, we need to add another element to these remarks, and this is the fact that the scraping of texts and images produced by writers and artists to train AI is the

> largest and most consequential theft in human history. Because what we are witnessing is the wealthiest companies in history (Microsoft, Apple, Google, Meta, Amazon…) unilaterally seizing the sum total of human knowledge that exists in digital, scrapable form and walling it off inside proprietary products, many of which will take direct aim at the humans whose lifetime of labor trained the machines without giving permission or consent. (Klein 2023: n.p.)

In this light, we can see that the issue with AI and its use is not just a matter of defining the concept of AI properly and understanding and explaining what it does, but also of regaining control over human cultural production and creativity. This is also to reclaim more control over the digital tools that are pervasively taking an ever-increasing space in our lives, limiting our agency to what the affordances of the devices and black boxes allow us. Critical analysis of the co-existence of human and machine cognition are necessary—especially in the fields of learning and education (cf. Markauskaite et al. 2022; Tafani/Pievatolo 2024)—, also when they take the form of creative resistance to the unilateral imposition of AI technologies.

5 From Algorithmic Sabotage to Data Poisoning

Forms of critical engagement and creative resistance to AI are being developed by artists, researchers and hacktivists. The Algorithmic Sabotage Research Group (ASRG) published a "Manifesto on 'Algorithmic Sabotage'" (2024) that claims to be "a figure of techno-disobedience for the militancy that's absent from technology critique" and "a form of counter-power" which aims at "dismantling contemporary forms of algorithmic domination and reclaiming spaces for ethical action from generalized thoughtlessness and automaticity". The intent of ASRG goes in the direction of a "dissonant reading of computational history" and of "an alternative understanding of the technological," thus working on an "alternative account of the actual and potential impacts of the computational" (Bassett 2021: 206), inscribing itself in the line of anti-computing positions.

It is of extreme interest that some of these forms and tools are developed inside universities, thus combining academic research with a creative/subverting intent and an overall focus on the common good of society. *Glaze*, for example, is a data poisoning tool developed by a team led by Ben Zhao and Shawn Shan at the University of Chicago. As the research team explains on their website, they wanted to develop "technical tools with the explicit goal of protecting human creatives against invasive uses of generative artificial intelligence or GenAI" that "artists can use to disrupt unauthorized AI training on their work product. Ultimately our goal is to ensure the continued vitality of human artists, and to restore balance and ensure a healthy coexistence between AI and human creatives, where the human creatives retain agency and control over their work products and their use" (*The Glaze Project*: "Our Mission and Vision" n.d.).

Another data poisoning tool is *PhotoGuard*, and it was developed by a research team at MIT: it uses an encoder attack which makes the AI model interpret the image it tries to scrape as something else, and a diffusion attack which disrupts the way the AI models generate images (cf. Salman et al. 2023). *Kudurru*, another tool that helps artists to protect their work from AI crawlers, is also worth mentioning because it "gives artists two options to disrupt scraping. First, they can simply block the blacklisted IP addresses. Second, […] they can also choose to sabotage or 'poison' the scrapers' efforts by sending back a different image than the one requested" (Knibbs 2023: n.p.). Data poisoning ranks among the actions of data leverage (cf. Vincent et al. 2021), one of the few tools in the hands of users to counter the indiscriminate use of their data by large companies for their sole economic profit.

The examples mentioned in this section should show that a critical engagement with AI is pivotal in order to fully grasp the social and political implications of its development and diffusion. If we consider that AI models are trained, as we have seen, on stolen human writing, visual arts, music and even voices, we can understand that disrupting the digital forms of enclosure AI companies are imposing could be considered a practical contribution to the critical reflection about technological development. These disruptions are also a practical way to reappropriate digital tools and to deepen our understanding of how the digital affects artistic creation and human cognitive processes. It seems to me that the knowledge and awareness I tried to highlight in this article should be at the foundation of every interaction with AI, be it creative, professional or academic.

6 Conclusions

This article attempted to illustrate a three-stage process to deal with the challenges posed by artificial intelligence in a creative and informed way: at first, by translating the concepts that are hidden in AI companies' marketing jargon to better-defining expressions; secondly, by taking a look inside digital machines to try to translate their inner functioning for human understanding; thirdly, by engaging with AI in a critical and creative way which aims to rebalance the inequality of forces at play when it comes to complex technologies such as machine learning. And this is because "[u]nderstanding AI means understanding its specific computational operations and everything that is being carried along by them; the history that AI has absorbed, the world in which it is emerging, and the futures that it calls forth" (McQuillan 2022: 2).

To properly understand the challenges we are facing with AI applications (automatic translation, general-purpose LLMs, TTIs, but also the use of AI in fields such as the criminal justice system and healthcare) we should therefore take into account the social, economic and political underpinnings—that is to say, the ideology—that inform the development of AI technologies from their very beginning.

The critical analysis proposed in this article aims to demonstrate that an up-to-date digital hermeneutics can only be developed through a profound and critical engagement with the technologies with which we are constantly interconnected. This means opening up the devices and services we use to the critical scrutiny of research, with a multidisciplinary approach that calls upon different fields, such as philosophy of technology, computer science, cognitive science, critical code studies, just to name a few. This socio-technical analysis will help us not only to be informed and knowledgeable users but also to

translate demands for social justice into concrete actions that counter technological determinism, helping society as a whole to shape digital technologies for the common good.

7 References

ASRG – Algorithmic Sabotage Research Group (2024): "Manifesto on 'Algorithmic Sabotage.'" 13.06.2024. URL: <https://algorithmic-sabotage-research-group.github.io/asrg/manifesto-on-algorithmic_sabotage/> (24.08.2025).

BASSETT, Caroline (2021): *Anti-Computing: Dissent and the Machine*. Manchester: Manchester University Press.

BENDER, Emily M. / HANNA, Alex (2023): "AI Causes Real Harm. Let's Focus on That over the End-of-Humanity Hype." In: *Scientific American*, 12.08.2023. URL: <https://www.scientificamerican.com/article/we-need-to-focus-on-ais-real-harms-not-imaginary-existential-risks/> (24.08.2025).

BERRY, David M. (2023): "The Explainability Turn." In: *Digital Humanities Quarterly* 17/2, n.p. URL: <https://www.digitalhumanities.org/dhq/vol/17/2/000685/000685.html> (24.08.2025).

BRITTAIN, Blake (2023): "AI-Generated Art cannot Receive Copyrights, US Court Says." In: *Reuters*, 21.08.2023. URL: <https://www.reuters.com/legal/ai-generated-art-cannot-receive-copyrights-us-court-says-2023-08-21/> (25.08.2025).

BROUSSARD, Meredith (2018): *Artificial Unintelligence. How Computers Misunderstand the World*. Cambridge / London: The MIT Press.

CHIANG, Ted (2023): "ChatGPT Is a Blurry JPEG of the Web." In: *The New Yorker*, 09.02.2023. URL: <https://www.newyorker.com/tech/annals-of-technology/chatgpt-is-a-blurry-jpeg-of-the-web> (24.08.2025).

CRISTIANINI, Nello (2023): *La scorciatoia. Come le macchine sono diventate intelligenti senza pensare in modo umano*. Bologna: il Mulino.

DALMASSO, Anna C. (2020): "I poteri divinatori degli schermi. Previsione, premediazione, feed-forward." In: CARBONE, Mauro / DALMASSO, Anna C. / BODINI, Jacopo [eds.]: *I Poteri Degli Schermi. Contributi Italiani a Un Dibattito Internazionale*. Milano: Mimesis, pp. 173–188.

FLORIDI, Luciano (2023): *The Ethics of Artificial Intelligence: Principles, Challenges, and Opportunities*. New York: Oxford University Press.

GIBNEY, Elizabeth (2024): "Not All 'Open Source' AI Models are Actually Open: Here's a Ranking." In: *Nature*, 19.06.2024. DOI: <10.1038/d41586-024-02012-5>.

THE GLAZE PROJECT (n.d.): "Our Mission and Vision." URL: <https://glaze.cs.uchicago.edu/aboutus.html> (24.08.2025).

HERRMAN, John (2022): "AI Art Is Here and the World Is Already Different." In: *Intelligencer*, 19.09.2022. URL: <https://nymag.com/intelligencer/2022/09/ai-art-is-here-and-the-world-is-already-different.html> (24.08.2025).

HUNGER, Francis (2023): "Unhype Artificial 'Intelligence'! A Proposal to Replace the Deceiving Terminology of AI." In: *Zenodo*, 12.04.2023. DOI: <10.5281/zenodo.7524493>.

KHAN, Mehtab / HANNA, Alex (2022): "The Subjects and Stages of AI Dataset Development: A Framework for Dataset Accountability." In: *SSRN Scholarly Paper*, 11.10.2022. DOI: <10.2139/ssrn.4217148>.

KLEIN, Naomi (2023): "AI Machines Aren't 'Hallucinating.' But Their Makers Are." In: *The Guardian*, 08.05.2023. URL: <https://www.theguardian.com/commentisfree/2023/may/08/ai-machines-hallucinating-naomi-klein> (24.08.2025).

KNIBBS, Kate (2023): "A New Tool Helps Artists Thwart AI—With a Middle Finger." In: *Wired*, 12.10.2023. URL: <https://www.wired.com/story/kudurru-ai-scraping-block-poisoning-spawning/> (24.08.2025).

LAGHI, Roberto (2023): "Humain-Machine : Une relation des traductions (entre numérique et cognition)." In: FROELIGER, Nicolas / LARSONNEUR, Claire / SOFO, Giuseppe [eds.]: *Human Translation and Natural Language Processing Towards a New Consensus?* Venezia: Edizioni Ca' Foscari, pp. 51–64. DOI: <10.30687/978-88-6969-762-3/004>.

LANIER, Jaron / WEYL, E. Glen (2020): "AI is an Ideology, not a Technology." In: *Wired*, 15.03.2020. URL: <https://www.wired.com/story/opinion-ai-is-an-ideology-not-a-technology/> (24.08.2025).

LOUKIDES, Mike (2022): "The Problem with Intelligence." In: *O'Reilly Media*, 13.09.2022. URL: <https://www.oreilly.com/radar/the-problem-with-intelligence/> (24.08.2025).

MALABOU, Catherine (2017): *Métamorphoses de l'intelligence: Que faire de leur cerveau bleu?* Paris: PUF.

MARKAUSKAITE, Lina / MARRONE, Rebecca / POQUET, Oleksandra / KNIGHT, Simon / MARTINEZ-MALDONADO, Roberto / HOWARD, Sarah / TONDEUR, Jo / DE LAAT, Maarten / SHUM, Simon B. / GA-

ŠEVIĆ, Dragan / SIEMENS, George (2022): "Rethinking the Entwinement between Artificial Intelligence and Human Learning: What Capabilities Do Learners Need for a World with AI?" In : *Computers and Education: Artificial Intelligence* 3, N° 100056, pp. 1–16. DOI: <10.1016/j.caeai.2022.100056>.

MCCARTHY, John (1996): *Defending AI Research: A Collection of Essays and Reviews*. Stanford: CSLI Publications.

MCQUILLAN, Dan (2022): *Resisting AI: An Anti-Fascist Approach to Artificial Intelligence*. Bristol: Bristol University Press.

MILMO, Dan (2024): "'Impossible' to Create AI Tools like ChatGPT without Copyrighted Material, OpenAI Says." In: *The Guardian*, 08.01.2024. URL: <https://www.theguardian.com/technology/2024/jan/08/ai-tools-chatgpt-copyrighted-material-openai> (24.08.2025).

MOHAMED, Alana (2022): "Magic Numbers." In: *Real Life*, 05.12.2022. URL: <https://reallifemag.com/magic-numbers/> (24.08.2025).

MOSS, Emanuel / WATKINS, Elizabeth A. / SINGH, Ranjit / ELISH, Madeleine C. / METCALF, Jacob (2021): "Assembling Accountability: Algorithmic Impact Assessment for the Public Interest." In: *Data & Society*. URL: <https://datasociety.net/library/assembling-accountability-algorithmic-impact-assessment-for-the-public-interest/> (24.08.2025).

PARR, Thomas / PEZZULO, Giovanni (2021): "Understanding, Explanation, and Active Inference." In: *Frontiers in Systems Neuroscience* 15, N° 772641, pp. 1–13. DOI: <10.3389/fnsys.2021.772641>.

PARRISH, Allison (2021): "Language Models can only Write Poetry." In: *Allison Posts*, 13.08.2021. URL: <https://posts.decontextualize.com/language-models-poetry> (24.08.2025).

QUINTARELLI, Stefano (2019): "Let's Forget the Term AI. Let's Call Them Systematic Approaches to Learning Algorithms and Machine Inferences (SALAMI)." In: *Quinta's Weblog*. URL: <https://blog.quintarelli.it/2019/11/lets-forget-the-term-ai-lets-call-them-systematic-approaches-to-learning-algorithms-and-machine-inferences-salami/> (24.08.2025).

RUDIN, Cynthia (2019): "Stop Explaining Black Box Machine Learning Models for High Stakes Decisions and Use Interpretable Models Instead." In: *Nature Machine Intelligence* 1/5, pp. 206–215. DOI: <10.1038/s42256-019-0048-x>.

SALMAN, Hadi / KHADDAJ, Alaa / LECLERC, Guillaume / ILYAS, Andrew / MADRY, Aleksander (2023): "Raising the Cost of Malicious

AI-Powered Image Editing." In: *arXiv*. URL: <http://arxiv.org/abs/2302.06588> (24.08.2025).

SOFO, Giuseppe (2025): "Traduire en archipel(s). Translating the City and Performing Translation in the Digital Era." In: VIDAL, Ricarda / CAMPBELL, Madeleine [eds.]: *The Translation of Experience*. London: Routledge, pp. 128–152. DOI: <10.4324/9781003462569-10>.

TAFANI, Daniela / PIEVATOLO, Maria C. (2024): "Omini di burro. Scuole e università al Paese dei Balocchi dell'IA generativa." In: *Bollettino telematico di filosofia politica*, 12.10.2024. URL: <https://btfp.sp.unipi.it/it/2024/10/omini-di-burro-scuole-e-universita-al-paese-dei-balocchi-dellia-generativa/> (24.08.2025).

TEGMARK, Max (2017): *Life 3.0: Being Human in the Age of Artificial Intelligence*. New York: Alfred A. Knopf.

VINCENT, Nicholas / LI, Hanlin / TILLY, Nicole / CHANCELLOR, Stevie / HECHT, Brent (2021): "Data Leverage: A Framework for Empowering the Public in Its Relationship with Technology Companies." In: *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency* (*FAccT '21*). New York: Association for Computing Machinery, pp. 215–227. DOI: <10.1145/3442188.3445885>.

WEXLER, James / PUSHKARNA, Mahima / BOLUKBASI, Tolga / WATTENBERG, Martin / VIÉGAS, Fernanda / WILSON, Jimbo (2020): "The What-If Tool: Interactive Probing of Machine Learning Models." In: *IEEE Transactions on Visualization and Computer Graphics* 26/1, pp. 56–65. DOI: <10.1109/TVCG.2019.2934619>.

About the Author: After a master's degree in History and another in Public and Political Communication, Roberto Laghi (https://orcid.org/0000-0003-4084-0437) worked for about ten years as a journalist, editor and consultant. In 2023, he obtained his PhD from the University of Avignon ("Langues et littératures romanes") in joint supervision with the University of Parma ("Scienze filologico-letterarie, storico-philosofiche e artistiche"). His dissertation, focused on contemporary Italian digital writing, opens up to the analysis of human writing through the technologies used by authors to create it, considering not only the possibilities offered by specific devices and services, but also the economic, social and political implications of software and hardware. After completing his doctorate, his research now focuses on digital writing and its central role in the transformative shifts that are taking place in contemporary societies. Through an interdisciplinary ap-

proach derived from the Digital Humanities, he has turned to disciplines such as cybernetics, neuroscience and biotechnology as additional tools for critical analysis to build a theoretical framework adapted to the complexity of our digital world.

Contact: roberto.laghi@alumni.univ-avignon.fr