



Images of Artificial Intelligence: a Blind Spot in AI Ethics

Alberto Romele¹

Received: 21 June 2021 / Accepted: 11 January 2022 / Published online: 29 January 2022
© The Author(s) 2022

Abstract

This paper argues that the AI ethics has generally neglected the issues related to the science communication of AI. In particular, the article focuses on visual communication about AI and, more specifically, on the use of certain stock images in science communication about AI — in particular, those characterized by an excessive use of blue color and recurrent subjects, such as androgyne faces, half-flesh and half-circuit brains, and variations on Michelangelo's *The Creation of Adam*. In the first section, the author refers to a “referentialist” ethics of science communication for an ethical assessment of these images. From this perspective, these images are unethical. While the ethics of science communication generally promotes virtues like modesty and humility, similar images are arrogant and overconfident. In the second section, the author uses French philosopher Jacques Rancière's concepts of “distribution of the sensible,” “disagreement,” and “pensive image.” Rancière's thought paves the way to a deeper critique of these images of AI. The problem with similar images is not their lack of reference to the “things themselves.” It rather lies in the way they stifle any possible forms of disagreement about AI. However, the author argues that stock images and other popular images of AI are not a problem per se, and they can also be a resource. This depends on the real possibility for these images to support forms of pensiveness. In the conclusion, the question is asked whether the kind of ethics or politics of AI images proposed in this article can be applied to AI ethics tout court.

Keywords AI images · Stock images · AI ethics · Ethics of scientific communication · Jacques Rancière

This article is part of the Topical Collection on *Philosophy of Technology and the French Thought*

✉ Alberto Romele
alberto.romele@uni-tuebingen.de

¹ IZEW, Universität Tübingen, Tübingen, Germany

1 Introduction

The main hypothesis of this article is that there is a blind spot in the current debate in AI ethics. Let us consider the recently edited *Oxford Handbook of Ethics of AI* (Dubber et al., 2020). In this eight-hundred-and-eighty-one-page volume, not a single line is devoted to an issue that should be at the center of the inquietude of scholars that are concerned with AI ethics, especially of those who come from humanities and social sciences.

We are referring to communication, in particular science communication, *about* AI. The literature on AI ethics seems to ignore the problem. A search for “ethics AND communication AND Artificial Intelligence” via Google Scholar does not return any useful results.

Before continuing the discussion, an important clarification is needed. One might argue that the ethics of communication about AI is not AI ethics, but ethics of science communication applied to the specific topic of AI. This is disputable, however. Indeed, it would be misleading to think of AI as a collection of techniques and technologies independent from the way innovation in AI is mediated and communicated. Communication about innovation in AI essentially contributes to frame expectations or, as we prefer to name them, “imaginaries”¹ about AI that we believe play a fundamental role in the concrete development of AI and its implementation in our societies. The philosophy of AI, and the philosophy of technology in general, have paid little or no attention to the issue of technological imaginaries. On the contrary, disciplines like Science and Technology Studies (STS) and media studies have been particularly attentive to the ways the representations of technology, be they visual or written, institutional or not, etc., become conditions of possibility for the existence and development of specific technologies.

For instance, Flichy (2007, pp. 8–12) refers to Ricoeur’s connection between ideology and utopia to understand the role of the *imaginaire* in what he calls the “technological action.” At the center of Ricoeur’s use of these concepts is the idea that ideology and utopia should not be defined in opposition to reality because reality is always symbolically mediated. Ideology and utopia are the two constitutive poles of the social *imaginaire*, one trying to maintain the social order, the other trying to disrupt it. The same holds true for the technological action, which is always embedded and supported by the *imaginaire* through a process that goes from utopia to ideology.

The ontological premise of this article is that one cannot develop any comprehensive understanding of AI without taking into account the imaginaries about AI. Since these imaginaries are crystallized in visual or written representations, we also contend that a comprehensive AI ethics should include considerations on the representations of and the communication about AI.

¹ Expectations are generally understood as present statements that say something about the future. The concept of imaginary is broader because it includes in the same dynamic *past* prejudices and *future* expectations about the *present*.

This ontological premise has effects on both the definition of AI and the definition of AI ethics. McCarthy defines AI as the “science and engineering of making intelligent machines, especially intelligent computer programs.”² This definition has the merit of suggesting that intelligence is not a human prerogative. Today, the term “Artificial Intelligence” is mostly used to indicate all sorts of machine learning algorithms. This, one might say, further frees AI from the obligations of resembling human intelligence or being embedded in a human-like body. However, we think that such a definition is still reductive insofar as it focuses exclusively on AI as a technical and scientific phenomenon. We argue that AI has also become a social and cultural fact.

Regarding AI ethics, it is interesting to note that in his entry for the *Stanford Encyclopedia of Philosophy*, Vincent C. Müller (2020) distinguishes between two main areas of focus that of the ethical issues that arise with AI systems as *objects*, that is, tools made and used by humans, and that concerning AI systems as *subjects*, that is, ethics for the AI systems themselves. Our idea is that there is room for a third area, the one that considers AI as a technological imaginary and its consequences. There is already emerging research in this area. Think, for example, of Cave and Dihal (2020), who focus on the fact that AI is predominantly portrayed as white — both in color and ethnicity.

In this paper, we do not pretend to deal with the communication about AI in general. Rather, we want to focus on a specific aspect of it, namely the visual communication about AI. Even more specifically, we intend to deal with the use of AI images that are produced by non-experts who, presumably, did not consult any expert or scientific source during the production of these images. We are referring to the many popular (popular in the sense of something intended to suit the general public) visual representations of AI that one can find on the homepages of university departments and laboratories (some of which are considered to be leading in the field of AI), on the posters of academic events about AI, official research communications from public institutions, in specialized courses, on the cover of books, etc. Many of these images are stock images, that is, pre-produced images made available for license by paying a fee to both the creators and the stock agencies managing the images.

The domain of popular visual representations of AI is broader than just AI stock images. There are popular AI images that are in fact produced in other contexts and for other purposes. One thinks of DeviantArt, a kind of online social network and art gallery where one can find many images labeled as “artificial intelligence” that can be liked, commented on, and in some cases downloaded for free.³ But stock images represent a quantitatively impressive phenomenon: consider that the search engine of Getty Images gives at this moment (December 2021), 27,901 images for “Artificial Intelligence” research. In addition, there is an economic and algorithmic logic behind stock images, whose ultimate purpose is to be sold, so they are always among the first results of our searches via search engines. Finally, we should not forget that many public institutions and private companies have specific economic

² <http://www-formal.stanford.edu/jmc/whatisai/node1.html>. Accessed December 1st, 2021.

³ <https://www.deviantart.com/>. Accessed December 1st, 2021.

agreements with stock imagery agencies, with the result that these institution's and company's communication services routinely use stock imagery to represent emerging technologies such as AI.⁴

The images we want to deal with have been for long time dismissed, both by “hard” and “soft” sciences, as mere fantasies. However, they are invasive in the current imagery of AI, and for this reason, they deserve to be questioned.

The article is structured in two sections. In the first section, we present two cases of stock images of AI in science communication, and we apply to them a standard ethics of science communication. The output is a foregone conclusion: from the perspective of this ethics, which is oriented by a form of “referentialism,” these images of AI are unethical. Similar stock images of AI do not “humbly” represent the “things themselves”; they let more than what they are supposed to show be seen (certainly more than what is concretely done in technological innovation in AI). At this point, two choices seem possible. The first one consists of criticizing the use of similar images and inviting scholars, and, more generally, all stakeholders involved in science communication about technological innovation in AI, to be more cautious with how they visually represent AI. The second one, which we believe to be more interesting, consists in accepting that AI is difficult to visually represent, but is represented nevertheless. Similar images cannot simply be dismissed, because they are produced continuously and so occupy an important part of AI's present imagery. In the second section, we engage with the reflections on aesthetics, politics, and images developed by the French philosopher Jacques Rancière, in particular the notions of “distribution of the sensible” (Rancière, 2004), “disagreement” (Rancière, 1999), and “pensive image” (Rancière, 2009a). We contend that Rancière's perspective offers the possibility of a different critique of stock images and other popular images of AI. From his perspective, similar images of AI are problematic not because they are “unethical” but rather because they are “unpolitical.” The problem with them does not lie in their lack of reference to the “things themselves.” It rather lies in the way they mark a gap between experts and non-experts, insiders and outsiders; it also lies in their incapacity to promote forms of disagreement among concerned groups beyond a simplistic logic of oppositions—goodness/badness, risk/opportunity, humans/nonhumans,⁵ etc. We contend as well that Rancière's perspective offers the possibility to think of these images beyond their criticisms, that is, not only as a danger but also as potential resources. This depends on the concrete possibility for these images to support, rather than stifle, forms of pensiveness. In the conclusion,

⁴ See *Research*eu*, the monthly magazine of CORDIS, European Commission's primary source of results from the projects funded by the EU's framework programs for research and innovation. The magazine, as well as CORDIS' website, makes abundant use of images of science and technology retrieved from Shutterstock. <https://cordis.europa.eu/research-eu/en>. Accessed December 1st, 2021.

⁵ In an article devoted to the visual representations of data centers, Taylor (2019) has theorized the notion of “technological wilderness.” According to him, what characterizes these images is the absence of human beings. This corresponds to a representational strategy related to “to emic and etic fantasies and futures of human-free security, automation and data objectivity” (Taylor 2018, 3). Interestingly enough, popular imagery of AI is usually characterized by the presence of both humans and machines, most often represented in terms of transition from one to the other.

we briefly ask ourselves if the kind of ethics (or politics) of AI images we propose in this article can be applied to AI ethics tout court.

2 The Unethics of AI Images

Type “Artificial Intelligence” in a web browser and look for images: Among the results, you will see unreal holographic interfaces, half-flesh half-circuit brains, lines of code waving in space, robots tapping on smart touchscreens, and at least one of the hundred variations of Michelangelo’s *The Creation of Adam* in a human–robot version. Most of these images are stock images. What usually characterizes stock images is their clichéd way of representing aspects of reality. Stock images have been mocked for this, for instance where women are pictured laughing alone eating salad⁶ or seem unable to drink water from a bottle or glass.⁷ In the case of AI and other emerging technologies, stock images have the tendency to be overly “unrealistic” and “hyperbolic.” A limited group of scholars in media studies has undertaken analyses of stock images and their social consequences, without, however, focusing on the images of science and technology or specifically on AI—see, in particular, Frosh (2003, 2020) and Turlow, Aiello, and Portmann (2019). It has been observed that while stock images are generally dismissed as the “wallpaper” of consumer culture, they are also “central to the ambient image environment that defines our visual world” (Aiello, 2016, np).

Stock images of AI have not only invaded the popular Web. They are widely used, both online and offline, to communicate about events, publications, courses, etc., on AI proposed and organized by scientific institutions that are often considered to be leading in the field of AI research (be it in engineering or in social sciences and humanities). In this regard, without any claim of exhaustiveness, we started to collect stock images of AI used in science communication and marketing through an Instagram profile called “ugly.ai.”⁸ We collected from the profile over eight months (May–December 2021), collecting more than a hundred images. From these images, we choose in this article to focus on two images, both relating to the field of AI ethics. It is indeed interesting to note that AI ethics itself sometimes shows little attention to the ethical implications of visual representations of AI. It is important to stress that it is not our intention to offer a detailed analysis of these images. In fact, this article has a theoretical intent. Also, it is important to highlight that we do not want to criticize the use of images when communicating about AI or AI ethics in general, nor the use of stock images as such. Rather, we want to problematize the use, however abundant, of a certain type of stock image, which is characterized by some common traits, for example: colors (mainly blue), subjects (robots,

⁶ <https://www.thehairpin.com/2011/01/women-laughing-alone-with-salad/>. Accessed December 1st, 2021. The example is retrieved from Aiello and Woodhouse (2016).

⁷ https://www.boredpanda.com/women-dont-know-how-to-drink-water-stock-photos/?utm_source=google&utm_medium=organic&utm_campaign=organic. Accessed December 1st, 2021.

⁸ <https://www.instagram.com/ugly.ai/>. Accessed December 1st, 2021.

half-artificial brains, human hands meeting artificial hands, female and androgynous faces, zero and ones, etc.), and certain dynamics related to time, space, and subjectivity/intersubjectivity.⁹

Image 1 is a screenshot of the cover of *The Oxford Handbook of Ethics of AI*.¹⁰ As one can read on the bottom-left of the back cover of the book, the cover image is retrieved from iStock, a company owned by Getty Images, the most important stock image supplier worldwide, and the author is the professional Moldovan illustrator Fiodora Chiosea. At the color level, a predominance of blue and white can be seen. Regarding white, we refer again to Cave and Dihal (2020). Regarding blue, Pastoureau (2018) concludes his historical research by stating that if blue is the most appreciated color in the world today, it is because it is not a strong color. It is a color that does not cause shock, does not hurt. Instead, it is a calming, peaceful, distant, anesthetizing color. It is no coincidence that many international organizations use blue to represent themselves visually: the UN, UNESCO, the Council of Europe, and the European Commission. These observations on the color blue give strength to the thesis about the “anaesthetics” of similar AI images that will be advanced later. The subject of image 1 is a classic androgynous face that, in this case, is made of “digital particles” that become a printed circuit board. On the website of iStock, the image is presented as follows: “Vector of a face made of digital particles as symbol of artificial intelligence and machine learning. Abstract human head outline with a printed circuit board. Technology and engineering concept.”¹¹ Finally, something interesting can be said about the internal dynamics of the image, particularly with regard to time. In the original image, the one available on iStock’s website, the image goes in the opposite direction, going from the printed circuit board on the left to the face made of digital particles on the right. According to classical Western logic, as manifested, for example, in the practice of reading, time flows from left to right. This means that the sense of the original image is that of a digital object made of circuits, which now becomes a *quasi*-human (an artificial intelligence). Once inverted, as in the case of the *Oxford Handbook*’s cover, the image might suggest something very different: a human being who transforms and becomes *non*-human, a digital object — in fact, in this case, the circuit board represents a principle of dematerialization.

Image 2 comes from a webpage of the website of Futurium, a European Commission’s platform “dedicated to European citizens for discussing EU policies.” From

⁹ For a semiotic analysis of images around these three elements (subjectivity and intersubjectivity, time, and space) see Dondero (2020). Semiotics is concerned with the enunciative marks of subjectivity and intersubjectivity, time, and space in the discourse. In verbal semiotics, these marks are the pronouns (I, you, etc.), the verbal tenses (from present to past), and the adverbs (here and there). In visual semiotics, these marks are different. Subjectivity and intersubjectivity correspond to the system of gazes that circulate within the image, or between its subject and the potential viewer. In particular, the profile corresponds to the third person, while the front view to the first and second person. Time is expressed through the disposition of the figures on different levels of depth. Finally, space is mainly manifested through perspective.

¹⁰ https://www.instagram.com/p/CPH_Iwmr216/. Accessed December 1st, 2021.

¹¹ For image 1, as well as for image 2, we have used the reverse image search engine TinEye (<https://tineye.com/>). Accessed December 1st, 2021), which enables one to find the original source of the image by looking for results among only stock and collection.

this webpage, one could download (and read about the piloting process concerning) the “European Guidelines for a Trustworthy AI” in multiple languages, first published in April 2019.¹² This is an image from iStock as well, by Thai illustrator Kit-tipong Jirasukhanont (Phonlamaipphoto). European institutions have been engaged for years in the development of a “European way” to AI, which should be characterized not only by technological excellence but also, and above all, by ethical values. It is then interesting to observe how there is here a lack of attention to the ethical implications of the images through which Europe’s ethical commitment to AI is represented. Blue is the dominant color in this image. Moreover, there is a movement from left to right that suggests a shift from the past (the human being) to the present and future (the robotic hand). But in this case, the most interesting aspect is perhaps the subject of the image itself, in which there is a clear reference to Michelangelo’s *The Creation of Adam*. In this way, a general aura of transcendence is attributed to AI, as if AI were the result of a divine emanation rather than a human creation subject to possible imperfections. Incidentally, it is interesting to note that in *The Creation of Adam*, the right side of the image is occupied by God and not by Adam, so one might wonder whether in Image 2 it is the AI itself, represented as a robot hand, that is divinized. In truth, there are two elements in this image that moderate such an “extreme” interpretation. The first one is the illuminated finger of the robotic hand, which can only recall the finger of E.T., the character in Spielberg’s movie.¹³ The second is the presence of a touch screen between the two. The fact that the touch screen is transparent suggests that there is no longer a “behind” and “in front” of the screen. These two elements make visible in the image not only the idea of divine creation, but also that of an encounter between two conscious entities.

Our goal in the rest of this section is to apply to these two examples what we believe to be a very common ethical perspective in science communication. We refer to Dahlstrom and Ho (2012), who investigate the ethical implications of using

¹² <https://www.instagram.com/p/CPH8xoCLTm7/>. Accessed December 1st, 2021. The image is also retrievable at https://ec.europa.eu/futurium/sites/futurium/files/capture_1_0.jpg. Accessed December 1st, 2021. Since mid-May 2021 the website is archived, and a new Futurium platform has been launched. On this specific image and the “AI creation meme,” see Singler (2020). The author collected 79 images of this kind, and analyzed them from different perspectives: colors, background imagery, online locations of the images, relative positions of the human and the artificial arm (right or left), etc. She observed the emergence, beyond aesthetics, of “post-humanist narratives that express the apocalyptic,” where “apocalypse” means “the transformation of the world, either through a transformation of humanity or of a new creation, and the relationship between the human and the created machine that that suggests” (Singler 2020, p. 14).

¹³ <https://www.instagram.com/p/CPH8xoCLTm7/>. Accessed December 1st, 2021. The image is also retrievable at https://ec.europa.eu/futurium/sites/futurium/files/capture_1_0.jpg. Accessed December 1st, 2021. Since mid-May 2021 the website is archived, and a new Futurium platform has been launched. On this specific image and the “AI creation meme,” see Singler (2020). The author collected 79 images of this kind, and analyzed them from different perspectives: colors, background imagery, online locations of the images, relative positions of the human and the artificial arm (right or left), etc. She observed the emergence, beyond aesthetics, of “post-humanist narratives that express the apocalyptic,” where “apocalypse” means “the transformation of the world, either through a transformation of humanity or of a new creation, and the relationship between the human and the created machine that that suggests” (Singler 2020, p. 14).

narrative to communicate science to a non-expert audience — and, of course, not only texts, but also photographs and images in general can have narrative properties.

Based on the existing literature on narrative and its cognitive and social effects, the authors state that narrative can have a multitude of consequences, such as improving comprehension, generating more interest and engagement, increasing self-efficacy through modeling, influencing real-world beliefs, and persuading an otherwise resistant audience.

The authors introduce three ethical considerations concerning the use of narrative in science communication: (1) What is the underlying purpose of using narrative: comprehension or persuasion? This includes two sub-considerations: (a) Do I want to facilitate potential controversy through greater understanding or reduce potential controversy through greater acceptance?; (b) Can I justify manipulating my audience?; (2) What are the appropriate levels of accuracy to maintain within the narrative? This includes the following sub-considerations: (a) Which elements of my topic must remain rigidly accurate and which can be relaxed to construct a more effective narrative?; (b) Is it necessary that my narrative portrays a generalizable example or can it justifiably portray an extreme example? (3) Should narrative be used at all? This also includes two sub-considerations: (a) Will my audience accept a narrative from my position?, and (b) Will others within my issue be using narrative?

We hypothesize that behind Dahlstrom and Ho's ethical considerations about using narrative in science communication, there is the issue of reference or adherence of the narrative to the scientific or technological object/fact in question, and to the evidence-based kind of reasoning that presumably characterizes scientific discovery and technological innovation. Our thesis is that if considered from this point of view, visual representations of AI like those we have selected, representative of the dominant imagery of AI in our view, are simply unethical.¹⁴

In all three ethical considerations, Dahlstrom and Ho propose, the ethical value of narrative is directly proportionate to its capacity to leave room, in the end, for science and technology and their way of reasoning. It is not by chance that virtues like humility, sincerity, transparency, openness, honesty, *kairos* (meaning in Greek “the opportune moment”), and generosity have been put at the center of virtue ethics of science communication—see in particular Medvecky and Leach (2019, Chapter 9), and, for a critical perspective, John (2018). The first consideration concerns the possibility of resorting to narrative in science communication either for persuasion or for comprehension. The authors have two frameworks in mind: PUS (Public Understanding of Science) and PEST (Public Engagement in Science and Technology), respectively. It is important to highlight that in this context, we are not discussing what scientific practice is or what kind of attitude or reasoning best fits science and technology. As such, the difference between PUS and PEST can be disregarded here.

¹⁴ Of course, naive referentialism has been abandoned for a long time in the reflections on the visual representations of science and technology in disciplines like the STS and the philosophy of technology. Think, for instance, of Latour's (2014) motto about scientific images: “the more manipulations, the better.” However, one should also consider the fact that, no matter if naively embraced or heavily problematized, reference to the things themselves remains the major issue for the great majority of these studies.

The fact is that in both cases, the use of narrative has the sole function of paving the way to a dynamic that is entirely internal to science and technology. The same holds for the second consideration, which is about the appropriate level of accuracy (strict or relaxed) to maintain in the use of narrative in science communication. Once again, there is no recognition of narrative per se: its ethical value is always measured on the basis of its “accuracy,” that is, its capacity to properly refer to the things themselves. Finally, the last consideration is about the possibility of not using narrative at all, which means that narrative in science communication is somehow reduced to a sometimes necessary, but always unpleasant, stratagem to realize the scopes of science. With theological terminology, we could say that the logic of the use of narrative in science communication is a logic of *kenosis*, which in Ancient Greek means “self-emptying.”

Images such as images 1 and 2 follow an opposite logic. They are not humble, honest, sincere, or transparent. Rather, they are arrogant, and overconfident. In sum, they are not “accurate.” They indicate more scientific progress than they should, certainly more than actually exists in current science and technology. No human head/brain/mind has ever been turned into, and probably will never be, “digital particles”; the robotic hand depicted in Image 2 is a fantasy: whoever has visited a prosthetic center, or even a scientific laboratory working on upper-limb prostheses, knows that the status of research and innovation in the field is very different. Not to mention the transparent touch screen, which is very different from the screens we deal with in our everyday lives. According to the “referentialist” framework proposed by Dahlstrom and Ho, these images are simply unethical.

It must be acknowledged that representing AI to a non-expert audience is a real challenge. Especially when it comes to showing not only the technology itself, but also its social and cultural implications. In this regard, one could distinguish three levels of visual representation: (1) The first is the one that wants to be closer to the “thing itself,” that is, the algorithm. Think of the representation of a decision tree learning, of a network of artificial neurons, or the way the algorithm is encoded in a computer program. Yet, not only might one wonder if such representations really show the “thing itself.” One could say that such a representation does not take into account AI as a social and cultural phenomenon; (2) The second is the one that represents AI as being embedded in different technologies (drones, smartphones, mechanical arms, etc.) and specific contexts (agriculture, medicine, military actions, etc.). In this case, however, AI is clearly black-boxed into another technology. Moreover, such images are often already third-level images, for example, when they choose specific technical objects (in particular, humanoid robots), or when they “augment” existing technologies (e.g., by adding elements that come out three-dimensionally from the screen of a smartphone), or even when they place simple objects against backgrounds (e.g., sunsets or particularly clear skies) that instill feelings of hope or fear; (3) Finally, the third level is that of the images we consider in this article. From a referentialist point of view, these are definitely the worst ones. The fact is that they do not so much refer to the “thing themselves” as they do to expectations and imaginaries, whether those of engineers, organizations, and companies, or potential spectators. Each of these levels, we believe, is legitimate in its own way, but under certain conditions. Therefore, it is important to emphasize that

we do not want to advocate for a return from the third to the first level, according to what would be a classical referentialist approach. It would be wrong to think that the abundant use of third-level images would be merely transient and that; afterwards, a process of “normalization” would follow. The images of this level respond to a specific need that the other levels are unable to address. Finally, we also want to say that the problem of representability is not unique to AI. It is common to many digital, often emerging, technologies such as cloud and quantum computing. And it is a problem that also concerns non-digital technologies, such as nanotechnologies, on whose visual representation there is already a large literature (except, interestingly enough, on this third-level kind of visual representations) — see, for instance, Slaattelid and Wickson (2011). For this reason, the specific case of AI that we consider here could then be extended to other representations of science and technology.

3 The Unpolitics¹⁵ of the Images of AI

At the end of the previous section, we have reached a possible result. Such a result would consist of formulating recommendations for stakeholders about not using images of AI like Images 1 and 2 in their science communication, and, more generally, about moderate use of visual representations of AI—moderation is commonly considered to be a virtue. In its most radical, iconoclastic version, such recommendations might consist of inviting stakeholders to not use images at all when it comes to science communication of innovation in AI. While a similar conclusion is at the opposite of our intention in this article, we believe that it is already a good result, at least insofar as it problematizes a topic—the ethics of AI images—that the literature has completely ignored. In this section, however, we want to go beyond this perspective. The thesis of this section is that images of AI like Images 1 and 2 are not unethical; or, at least, the fact of being unethical according to a referential perspective is so evident that it does not represent a true problem. Similar images are rather “unpolitical.” To demonstrate this, we use in this section the thought of the French philosopher Jacques Rancière, in particular the concepts of “distribution of the sensible” (Rancière, 2004), “disagreement” (Rancière, 1999), and “pensive image” (Rancière, 2009a).

The recourse to Rancière’s thought is not accidental here. This article is included in a topical collection devoted to “Philosophy of Technology and the French Thought.” Esposito (2018) distinguishes German Philosophy, French Theory, and Italian Thought. For Esposito, the salient feature of the first, particularly the Frankfurt School, is Negativity. In Hegel, negativity was still only a passing moment; with Adorno and Horkheimer, it becomes insuperable instead. As for French Theory, its core category is the Neutral. For example, “Deconstruction is neutral, suspended

¹⁵ The terms “unpolitics” and “unpolitical” must be understood here in light of the distinction between politics and police that we will introduce later. Similar images of AI are certainly “political” insofar as they contribute to a certain distribution of the sensible. They are, however, “unpolitical” because the manner in which they do so is that which serves to maintain the *status quo* rather than undermine it.

between yes and no, positioned at their point of intersection. It marks its distance both from the paradigm of crisis and that of critique” (Esposito 2018, p. 16). In either case, he says, one ends up in an impasse of thought. The Italian Thought — Esposito thinks of a tradition in political philosophy going from Machiavelli to Agamben — would instead be able to avoid this impasse, being an “affirmative thought”: “it can be argued that, by and large, the main effort of Italian philosophers has been to think not in a reactive but in an active, productive, affirmative way” (Esposito 2018 p. 17). Following this distinction (admittedly simplistic in many ways, but nonetheless useful), we assert that Rancière’s thought is properly a representative of the French *Thought* because, while embracing a certain critical, and even neutral, attitude of French theorists (evident especially in the notion of “distribution of the sensible”), he also embraces the affirmative attitude of Italian thinkers, as emerges especially from the concepts of “disagreement” and “pensive image.” For instance, as will be shown in the conclusion, disagreement does not coincide with mere chaos, but rather with the concrete possibility of thinking and building a new form of (technological) democracy.

Our first hypothesis is that AI images like Images 1 and 2 are “unpolitical” because they contribute to the framing of a specific “distribution of the sensible” in the technological innovation in AI. For Rancière (2004, p. 12), the expression indicates “the system of self-evident facts of sense-perception that simultaneously disclose the existence of something in common and the delimitations that define the respective parts and positions within it.” In other words, the distribution of the sensible regards the constitution of a shared time, space, and horizon of understanding, and the distribution of access and roles (that is, recognition, legitimacy, and ultimately power) within such a delimited space, time, and horizon of understanding. The distribution of the sensible, and the consequent distribution of access and roles, imply exclusions, sometimes from specific access and roles, sometimes from the whole space, time, and horizon of understanding. The distribution of the sensible is for Rancière a political practice, because “politics revolves around what is seen and what can be said about it, around who can see and the talent to speak, around the properties of space and the possibilities of time” (Rancière, 2004, p. 13). Politics and aesthetics are strongly connected, where “aesthetics” is to be understood both in the sense of the Greek *aisthesis*, which means “perception,” and in the sense of art and cultural productions in general. On the one hand, politics is a matter of distribution (or exclusion from) roles and access to perception—seeing/being seen, listening/being listened to, etc. On the other hand, art and cultural productions can either contribute to the reproductions of the dominant regimes of perception or contribute to their suspension and eventual transformation.¹⁶

¹⁶ In what concerns the link between art and politics, see Rancière (2009b). For him, there is a continuity between authentic art and politics insofar as authentic art, as well as politics, represent the possibility to suspend the ordinary forms of the sensible experience, that is, the ordinary distribution of the sensible. In Rancière’s words (2009b, p. 25–26. Translation modified. Italics is our), “art and politics do not constitute two permanent separate realities [...]. They are two *suspended* forms of distribution of the sensible [...]” The English translation does not include the adjective “suspended,” hence radically distorting the meaning of the sentence here. Indeed, for Rancière, what characterizes both authentic art and politics is the possibility to suspend, that is, to offer alternatives to, the dominant regime of the sensible exercised by the police and of which what is called and legitimized as art is mostly an expression.

We contend that the dominant imagery of AI implies a specific distribution of the sensible whose ultimate effect is to mark a gap between experts and non-experts, insiders and outsiders. It has been argued that the use of images in science popularization has an introductory function. For instance, Gigante (2018) coined the term “portal images.” However, we contend that stock images of AI in science communication are “screen images,” where “screen” refers to its etymology, meaning “to cut, divide, cover, shelter, and separate.” The fact is that one can watch thousands of similar images of AI without having to develop any critical reasoning about AI. These images instead have an “anesthetic” effect, which means that the reiterated contact with them makes non-experts and outsiders less and less sensitive to the most urgent issues related to AI and increases their feelings of resignation about AI.

We propose to apply these considerations to our object of study. In particular, we introduce the notion of “*anaesthetics*,” a word referring to the fact that the distribution of the sensible related to similar images (aesthetics) has anesthetic effects on those who are “outside.” The concept of *anaesthetics* is also important for another reason. One might think that the loss in terms of both ethics and politics at the level of the single image of AI is somehow retrieved at the level of the context in which the image is used, and to which it finally belongs. Hence, a possible criticism of our discourse might consist in affirming that there is no ethics or politics of similar images per se, because similar images are always used in context, and the ethical or political assessment should be made not on the single image, but with regard to the whole context. To put it plainly, science communication on AI is full of ugly and bad images, yet these images can still be used ethically or politically whenever they are integrated into a rigorous discourse. However, such criticism not only forgets that in the media environment in which we live, images are most often detached from, and perceived outside from, their context. Think of how often we content ourselves with scrolling the home screen of our news feeds without actually reading the article or even the titles. This criticism also forgets that similar images can, through their “force,” anesthetize the communicational context in which they are supposed to be embedded and on which they are supposed to depend.¹⁷

Our second hypothesis is that stock images of AI are also unpolitical because they impede or anesthetize any form of “disagreement.” Above, we have argued that politics has to do with the distribution of the sensible. However, on other occasions, Rancière proposes distinguishing more carefully between politics and police. We might say that the distribution of the sensible as a form of domination is related to

¹⁷ One of the anonymous reviewers stated that “according to certain theories of interpretation (see the intentionalism of Quentin Skinner), authorial intention is what matters.” For this reason, space should be given to the author’s reasons — in this case, either the author of the images or the one who chose them. To this objection, one can reply that there are many other theories of interpretation that argue that authorial intentions do not matter or that it would not be fair to take them into account. Once produced, a text, an image, or the articulation of the two is instead to be considered as autonomous both in its contents and in its effects. We are referring, in particular, to authors such as Gadamer and Ricoeur. The latter writes, for example: “Dialogue is an exchange of questions and answers; there is no exchange of this sort between the writer and the reader. Rather, the book divides the act of writing and reading into two sides, between which there is no communication. The reader is absent from the act of writing; the writer is absent from the act of reading” (Ricoeur 1991, p. 107).

the police, while politics in a proper sense is rather related to the practice of disagreement, which can also be understood as a suspension of the dominant distribution of the sensible. Rancière (1999, p. 29) defines the police as “an order of bodies that defines the allocation of ways of doing, ways of being, and ways of saying [...]; it is an order of the visible and sayable that sees that a particular activity is visible and another is not, that this speech is understood as discourse and another is noise.” He defines politics as “an extremely determined activity antagonistic to policing: whatever breaks with the tangible configuration whereby parties and parts or lack of them are defined by a presupposition that, by definition, has no place in that configuration” (Ibid.). He also says that “political activity is whatever shifts a body from the place assigned to it or changes a place’s destination. It makes visible what had no business being seen and makes heard a discourse where once there was only place for noise” (Rancière, 1999, p. 30).

Politics in a proper sense implies the possibility of disagreement, which is neither ignorance nor misunderstanding. Disagreement is neither a matter of teaching to others what they do not know yet, nor is it a question of explaining more, to allow better understanding. Disagreement is somehow more radical: it is “a specific type of speaking situation (*situation de parole*): one where one of the interlocutors does not hear what the other is saying. Disagreement is not the conflict between the one who says white and the one who says black. It is the conflict between the one who says white and the one who says white but does not hear the same thing” (Rancière, 1995, p. 12. Translation is ours).¹⁸ Police anesthetize disagreement and promote consensus, but the consensus is nothing but the disappearance of politics.

Let us now apply these ideas to the use of stock images and the like in science communication about AI. We already said that stock images are usually characterized by their generalized and stereotyped way of representing reality. These images regard the imaginaries, that is, the fears and hopes, enthusiasms, and hostilities about AI that the concerned group of non-experts (but also experts, insofar as experts are not constantly reasoning and acting as experts) has about AI. Stock images and all sorts of popular representations of AI might be considered public arenas that attract different audiences trying to cope with AI despite its inaccessibility and “black-box-ness.” However, this is still a desideratum, because many stock images of AI currently have little to do with disagreement. On the contrary, one can say that they anesthetize disagreement by promoting forms of consensus about the general hopes and fears about AI.

A general optimism permeates the visual representations of AI one can find on the online catalogs of stock image providers like Getty Images and Shutterstock. For instance, Image 3 is the first result for the quest “facial recognition AND Artificial Intelligence” on the search engine of Getty Images (December 2020, options “most popular” and “creative”). This image, titled “Businessman using face recognition outdoor” and authored by Wonry, has no alarming element, although facial

¹⁸ Surprisingly, the English translation of this text lacks two important paragraphs of the preface in the original French version (Rancière 1995, p. 12–13) in which Rancière specifies the meaning for him of “disagreement” (*mésentente*).

recognition is a much-debated topic in AI ethics precisely because of its potential risks.¹⁹ Rather the image recalls progress, future, and, we could even say, quiet and security. Even when pessimism emerges, for instance via more explicit searches, stock images tend to represent it as a caricature, as if the quest for differentiation and opposition were more important than a real engagement with the issue. Image 4 is the first result for “war AND Artificial Intelligence” on the search engine of Getty Images (December 2020, options “most popular” and “creative”).²⁰ Two robots (one of them recalls the ED-209 villain robot of the movie RoboCop) and a drone are represented in what looks like to be a post-apocalyptic environment. Incidentally, the alarming red color of the image opposes the reassuring blue color that dominates the current popular imagery of AI.

In the past years, some efforts have been made by stock image providers to promote different visual representations of social reality. For instance, in 2016, Getty Images, in collaboration with Women’s Sport Trust, launched a new collection on its online catalog called “Sporting Women,” whose aim is “increasing the visibility of female athletes” and “challenging the way female athletes are portrayed in imagery.”²¹ Citizen Stock is an attempt to produce generic images from “real people.” The initiative is presented as follows: “Citizen Stock was launched in May 2010 as a source of new images [...] depicting real people. Models are not role models at all, but children, moms, dads, grandparents, skateboarders, lawyers, teachers, musicians [...] and small business owners [...].”²²

Our third hypothesis is that a similar initiative might be undertaken in what concerns stock images of AI, and more generally stock images of science and technology.²³ In particular, we believe that more engaged imagery of AI could be created not so much following the classic urge for reference, but rather pursuing what Rancière has called the “pensiveness” of the image. According to the French philosopher, “a pensive image is [...] an image that conceals (*recèle*) unthought thought, a thought that cannot be attributed to the intention of the person who produces it and has an effect on the person who views it without her linking it to a determinate object” (Rancière, 2009a, p. 107. Translation modified).²⁴ Among the several

¹⁹ <https://media.gettyimages.com/photos/businessman-using-using-face-recognition-outdoors-picture-id866481488?s=2048x2048>. Accessed December 1st, 2021.

²⁰ <https://media.gettyimages.com/illustrations/futuristic-robotic-war-illustration-id975923624?s=2048x2048>. Accessed December 1st, 2021.

²¹ <http://press.gettyimages.com/getty-images-partners-with-womens-sport-trust-to-redefine-imagery-of-female-athletes-in-commercial-and-editorial-storytelling/>. Accessed December 1st, 2021. Another case is the Getty Images “Genderblend” collection, launched in 2015, which is supposed to portray gender identities and relations in ways that are more inclusive and diverse.

²² <https://www.citizenstock.com/>. Accessed December 1st, 2021. The example is retrieved from Frosh (2020).

²³ On the critique of the fact that a corporation like Getty Images defines visual politics, see Aiello and Woodhouse (2016).

²⁴ The English translator misses an important passage of the French text. In the French text it is written “An image is not supposed to think. It is supposed to be just an object of thinking. A pensive image is then [...]”, the English translation is “An Image is not supposed to think. It contains unthought thought [...]”.

examples, he considers the famous 1865 photo by Alexander Gardner of the sentenced-to-death Lewis Payne.²⁵ The pensiveness of this photography depends on the tangle between several forms of indeterminacy: (1) The one concerning the visual composition: we cannot know if the position—Lewis Payne is seated according to a highly pictorial arrangement—has been chosen by the photographer or not. We do not even know whether the photographer has simply recorded the wedges and marks on the wall, or whether he has deliberately highlighted them; (2) The one concerning the work of time: the body, the clothes, the posture of Lewis Payne are at home in our present, yet the texture of the photograph bears the stamp of times past; (3) The one concerning the attitude of the character: we know that Lewis Payne is going to die, but we cannot read his feelings in his gaze.

It might be thought that the pensiveness of the images depends exclusively on our ignorance and the resistance of the image to be interpreted—for instance, when its provenance or the thought of its author is unknown. However, Rancière insists on the fact that pensiveness rather depends on the capacity of the image to bring together different regimes of expression without homogenizing them. He talks, for example, of “dis-appropriate similarity” (Rancière, 2009a, p. 129), which is more than mere juxtaposition and yet less than identification. In other words, images are pensive insofar as they form always-open and never-exhausted metaphors on different spatial and temporal levels.

The concept of the pensive image is particularly interesting because it detaches the possibility for an image to be pensive from the need for it to be adherent to the reality it represents. Whether adherent to reality or not, an image can be thoughtful to the extent that it can provoke thought in the spectator. The presence of multiple planes, spatial and temporal, of interpretation, in short a metaphoricity intrinsic to the image itself, is what allows it to be pensive. Now, why are the AI stock images we have considered not pensive? Precisely insofar as they direct thought in a unique direction, for example, hope, fear, or trust. Without going into the details of the analysis, we can consider again the abundant use of a calming, anesthetizing color like blue as an example.

The paradigm through which Rancière builds his notion of a pensive image is art. We believe as well that artistic productions today offer several possibilities for visually representing AI beyond the usual clichés, and without much concern for the reference to the technical artifact. Let us consider the robotic sculpture *Black Box* by the French artist Fabien Zocco.²⁶ Robotic black cubes move slowly on the ground. Their movements let a sort of enigmatic behavior emerge, lending a semblance of life to these minimalist artifacts. *Black Box* thus aims to give substance to the often used, but less often thought of, metaphor of the “black box,” which in the ethical discourses on AI indicates the inaccessibility to the internal functions of a system such as a machine learning algorithm. This work does not refer to AI as a collection

²⁵ [https://en.wikipedia.org/wiki/File:Lewis_Payne_cwpb.04208_\(cropped\).jpg](https://en.wikipedia.org/wiki/File:Lewis_Payne_cwpb.04208_(cropped).jpg). Accessed December 1st, 2021. the same picture was chosen by Roland Barthes. Indeed, Rancière's notion of pensive image is a critical reply to Barthes distinction between *stadium* and *punctum*.

²⁶ <https://www.fabienzocco.net/blackbox.html>. Accessed December 1st, 2021.

of techniques and technologies—we do not know how the black boxes move. It rather refers to AI as an imaginary, which is, however, not anesthetized according to the easy opposition between fear and hope. *Black Box* inspires both fascination and uncanniness, attraction, and repulsion. The black boxes move, they behave and seem alive, and yet they cannot be understood. A second example is the *Anatomy of an AI System* by Crawford and Joler,²⁷ whose goal is to present Amazon Echo as “an anatomical map of human labor, data and planetary resources.” We believe that this map can be approached from two different levels. The first one is the level of representativeness. For instance, one can download and *read* the map in its details to have a better understanding of AI not in isolation, but rather in its multiple human and environmental implications. The second other one consists of *perceiving* the map as a whole. In this second case, the spectator is taken by a kind of vertigo, given the complexity and the many dimensions that are suggested by the opening of the AI black box — like the opening of a human body and the arrangement of all its internal organs. The effect, after all, is not unlike that of the *Black Box*. Certainly, this latter work extremizes opacity, while the other one extremizes “monstration.” Yet, in both cases, it is a matter of problematizing AI and our daily relationship with it.

We believe that the main challenge for the ethics of AI images would consist of going beyond the limits of the artistic (and hence most often elitist) production to import the pensiveness of works like *Black Box* and the *Anatomy of an AI System* in more popular contexts, in particular in the context of the production of stock images about AI, and science and technology in general.

4 Conclusion

In this paper, we have argued that there is a blind spot in the current debate about the ethics of AI. This blind spot consists of ignoring the ethical issues related to science communication about AI. In particular, we have focused on visual communication, and even more specifically on the use of certain stock images of AI. In the first section, we have referred to Dahlstrom and Ho (2012), who investigated the ethical implications of using narrative to communicate science, with a view to making an ethical assessment of the dominant imagery in science communication about AI. The result has been a foregone conclusion: similar images are unethical. While the ethics of science communication generally promotes the practice of virtues like modesty, humility, sincerity, transparency, openness, honesty, and generosity, stock images and other popular visual representations of AI are arrogant, pompous, and overconfident. In this section, we have also sketched the outlines of a general theory of visual representability of AI — which is today mostly identified with machine learning algorithms. We have distinguished between (1) the possibility of representing the algorithm itself; (2) the depiction of those technologies (drones, autonomous vehicles, etc.) in which AI is embedded; (3) the images, like those considered in this article, that focus on the expectations, fears, and hopes about AI. Our idea is that (1)

²⁷ <https://anatomyof.ai/>. Accessed December 1st, 2021.

and (2) are not intrinsically better than (3). Each of these levels is unsatisfying in its own terms; at the same time, each of them responds to needs that the other levels are unable to address.

In the second section, we have referred to the theories about aesthetics, politics, and images of the French philosopher Jacques Rancière. We hypothesized that Rancière's philosophy paves the way to a deeper critique of similar images of AI. The problem with these images is not their lack of reference. Rather, it lies in the way they anesthetize any debate and disagreement about AI. In particular, we have mobilized the three notions of "distribution of the sensible," "disagreement," and "pensive images." While the first two had the function of criticizing the current status of the dominant imagery of AI, the third one had the role of seeing not only a danger in the broad use of stock images in the science communication about AI, but also a potential resource.

Our final question concerns the possibility of applying similar ethics or politics of AI images to AI ethics tout court. The debate in AI ethics has been for a long time oriented by the ideal of consensus. Think of the several reports and guidelines about ethical AI that have proposed "universal" principles such as transparency, trustworthiness, and beneficence. Jobin, Jenca, and Vayena (2019) have stated that while a convergence around some of these principles is observable today, disagreement arises when it comes to putting them into practice. A similar disagreement depends, for instance, on the social and cultural contexts in which the principles must be applied. Scholars are increasingly attentive to the contextualization of AI ethics and the kind of misunderstandings and disagreements that the implementation of a globalized product such as AI technologies can cause in a specific social and cultural context or whenever different "spheres of justice" enter into conflict. We believe that Rancière's aesthetics and political philosophy might represent a good theoretical framework to think about AI ethics on a different basis. On such a basis, disagreement would be less an obstacle to be sooner or later overcome than a resource. This is what Rancière (Rancière, 1999, p. 102) says about consensus democracy:

According to the reigning idyll, consensus democracy is a reasonable agreement between individuals and social groups who have understood that knowing what is possible and negotiating between partners is a way for each party to obtain the optimal share that the objective givens of the situation allow them to hope for and which is preferable to conflict. But for parties to opt for discussion rather than a fight, they must first exist as parties who then have to choose between two ways of obtaining their share. Before becoming a preference for peace over war, consensus is a certain regime of the perceptible: the regime in which the parties are presupposed as already given, their community established, and the count of their speech identical to their linguistic performance. What consensus thus presupposes is the disappearance of any gap between a party to a dispute and a part of society. It is the disappearance of the mechanisms of appearance, of the miscount and the dispute opened up by the name "people" and the vacuum of their freedom. It is, in a word, the disappearance of politics.

In other words, consensus is already based on a certain distribution of the sensible that legitimates some actors, discourses, and ways of argumentation, while excluding in principle some others. The consensus is the "disappearance of politics"

because it is always-already legitimized by the police. Consensus excludes any form of disagreement, so one can suppose that several efforts currently undertaken to include marginalized individuals or groups in what concerns technological innovation in AI are rather forms of anesthetization of the disagreement that these marginalized individuals or groups may manifest. So, the question arises if a radically different AI ethics is possible; one in which the search for inclusion and consensus (on universal principles and virtues, for instance) leaves room for the creativity of disagreement and agonism²⁸ among the multiple concerned groups.

Acknowledgements The ideas presented in this paper have been previously discussed on three occasions: at the seminário permanente de fenomenologia of the center PRAXIS (Centro de Filosofia, Política e Cultura) at the University of Evora, Portugal, on December 11, 2020; at the colloquium of the University of Delft Philosophy Department on December 14, 2020; at the colloquium of the University of Twente Philosophy Department on January 12, 2021. We are grateful to the organizers of those events (Irene Borges Duarte, Pieter Vermaas, and Bas de Boer) and to all the participants for their precious insights that contributed to improving the article.

Funding Open Access funding enabled and organized by Projekt DEAL. The research related to this article is partially funded by the FCT (Fundação para a Ciência e a Tecnologia) as part of the project *From Data to Wisdom. Philosophizing Data Visualizations in the Middle Ages and Early Modernity*, POCI-01-0145-FEDER-029717.

Availability of data and material The data (images) that support part of the findings of this study are available at <https://www.instagram.com/ugly.ai/>.

Declarations

Ethics approval and consent to participate Not applicable.

Consent for publication Not applicable.

Competing interests The author declares no competing interests.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Aiello, G., & Woodhouse, A. (2016). When corporations come to define the visual politics of gender: The case of Getty Images. *Journal of Language and Politics*, 15(3), 352–368.

²⁸ The reference is to Chantal Mouffe (2013), whose thought has been partially inspired by that of Rancière. On the application of Mouffe's agonistic approach to technological conflict, see Popa, Block, and Wessenlik (2020).

- Aiello, G. (2016). Taking stock. *Ethnography Matters*, <https://ethnographymatters.net/blog/2016/04/28/taking-stock/>. Accessed December 1st, 2021.
- Cave, S., & Dihal, K. (2020). The whiteness of AI. *Philosophy & Technology*, 33, 685–703.
- Dahlstrom, M. F., & Ho, S. S. (2012). Ethical considerations of using narrative to communicate science. *Science Communication*, 35(5), 592–617.
- Dondero, M. (2020). *The language of images: The forms and the forces*. Springer.
- Dubber, M., Pasquale, F., & Das, S. (2020). *The Oxford handbook of ethics of AI*. Oxford University Press.
- Esposito, R. (2018). German philosophy, French theory, Italian thought. In D. Gentili, E. Stimilli, & G. Garelli (Eds.), *Italian critical thought. Genealogies of the present*. Rowman & Littlefield, pp. 11–22.
- Flichy, P. (2007). *The internet imaginaire*. The MIT Press.
- Frosh, P. (2003). *The image factory: Consumer culture, photography, and the visual content industry*. Berg.
- Frosh, P. (2020). Is commercial photography a public evil? Beyond the critique of stock photography. In M. Miles, & E. Welch (Eds.), *Photography and its publics*. New York: Bloomsbury. Pre-print version available at https://www.researchgate.net/publication/338880857_Is_Commercial_Photography_a_Public_Evil_Beyond_the_Critique_of_Stock_Photography. Accessed December 1st, 2021.
- Gigante, M. E. (2018). *Introducing science through images: Cases of visual popularization*. University of South Carolina Press.
- Jobin, A., Ienca, M., & Vayena, E. (2019). Artificial intelligence: The global landscape of ethics guidelines. *Nature Machine Intelligence*, 9(2), 389–399.
- John, S. (2018). Epistemic trust and the ethics of science communication: Against transparency, openness sincerity and honesty. *Social Epistemology*, 32(2), 75–87.
- Keohane, R. O., Lane, M., & Oppenheimer, M. (2014). The ethics of science communication under uncertainty. *Politics, Philosophy & Economics*, 13(4), 343–368.
- Kerr, A., Barry, M., & Kelleher J.D. (2020). Expectations of artificial intelligence and the performativity of ethics: Implications for communication governance. *Big Data & Society*, <https://journals.sagepub.com/doi/full/https://doi.org/10.1177/2053951720915939>. Accessed December 1st, 2021.
- Latour, B. (2014). The more manipulations, the better. In C. Coopmans, J. Vertesi, M. Lynch, & S. Woolgar (Eds.), *Representation in Scientific Practice Revisited*. Cambridge, MA: The MIT Press.
- Medvecky, F., & Leach, J. (2019). *An ethics of science communication*. Palgrave.
- Meyer, G., & Sandoe, P. (2012). Going public: Good scientific conduct. *Science and Engineering Ethics*, 18, 173–197.
- Mouffe, C. (2013). *Agonistics: Thining the world politically*. Verso.
- Müller, V.C. (2020). Ethics of artificial intelligence and robotics. *Stanford Encyclopedia of Philosophy*, <https://plato.stanford.edu/entries/ethics-ai/>. Accessed December 1st, 2021.
- Pastoureau, M. (2018). *Blue: The history of a color*. Princeton University Press.
- Pimple, K. (2002). Six domains of research ethics. *Science and Engineering Ethics*, 8, 191–205.
- Popa, E.O., Blok, V. & Wessenlik, R. (2020). An agonistic approach to technological conflict. *Philosophy & Technology*.
- Priest, S., Goodwin, J., & Dahlstrom, M. F. (Eds.). (2018). *Ethics and practice in science communication*. The University of Chicago Press.
- Rancière, J. (1995). *La mésentente: Politique et philosophie*. Galilée.
- Rancière, J. (1999). *Disagreement: Politics and philosophy*. Minnesota University Press.
- Rancière, J. (2004). *The politics of aesthetics: The distribution of the sensible*. Continuum.
- Rancière, J. (2009a). *The emancipated spectator*. Verso.
- Rancière, J. (2009b). *Aesthetics and its discontents*. Polity.
- Ricoeur, P. (1991). *From text to action*. Northwestern University Press.
- Singler, B. (2020). The AI creation meme: A case study of the new visibility of religion in artificial intelligence discourse. *Religions*, 11(5), 1–17.
- Slaattelid, R. T., & Wickson, F. (2011). Imag(in)ing the nano-scale: Introduction. *NanoEthics*, 5, 159–163.
- Thurlow, C., Aiello, G., & Portmann, L. (2019). Visualizing teens and technology: A social semiotic analysis of stock photography and news media imagery. *New Media & Society*, 22(3), 528–549.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Terms and Conditions

Springer Nature journal content, brought to you courtesy of Springer Nature Customer Service Center GmbH ("Springer Nature").

Springer Nature supports a reasonable amount of sharing of research papers by authors, subscribers and authorised users ("Users"), for small-scale personal, non-commercial use provided that all copyright, trade and service marks and other proprietary notices are maintained. By accessing, sharing, receiving or otherwise using the Springer Nature journal content you agree to these terms of use ("Terms"). For these purposes, Springer Nature considers academic use (by researchers and students) to be non-commercial.

These Terms are supplementary and will apply in addition to any applicable website terms and conditions, a relevant site licence or a personal subscription. These Terms will prevail over any conflict or ambiguity with regards to the relevant terms, a site licence or a personal subscription (to the extent of the conflict or ambiguity only). For Creative Commons-licensed articles, the terms of the Creative Commons license used will apply.

We collect and use personal data to provide access to the Springer Nature journal content. We may also use these personal data internally within ResearchGate and Springer Nature and as agreed share it, in an anonymised way, for purposes of tracking, analysis and reporting. We will not otherwise disclose your personal data outside the ResearchGate or the Springer Nature group of companies unless we have your permission as detailed in the Privacy Policy.

While Users may use the Springer Nature journal content for small scale, personal non-commercial use, it is important to note that Users may not:

1. use such content for the purpose of providing other users with access on a regular or large scale basis or as a means to circumvent access control;
2. use such content where to do so would be considered a criminal or statutory offence in any jurisdiction, or gives rise to civil liability, or is otherwise unlawful;
3. falsely or misleadingly imply or suggest endorsement, approval, sponsorship, or association unless explicitly agreed to by Springer Nature in writing;
4. use bots or other automated methods to access the content or redirect messages
5. override any security feature or exclusionary protocol; or
6. share the content in order to create substitute for Springer Nature products or services or a systematic database of Springer Nature journal content.

In line with the restriction against commercial use, Springer Nature does not permit the creation of a product or service that creates revenue, royalties, rent or income from our content or its inclusion as part of a paid for service or for other commercial gain. Springer Nature journal content cannot be used for inter-library loans and librarians may not upload Springer Nature journal content on a large scale into their, or any other, institutional repository.

These terms of use are reviewed regularly and may be amended at any time. Springer Nature is not obligated to publish any information or content on this website and may remove it or features or functionality at our sole discretion, at any time with or without notice. Springer Nature may revoke this licence to you at any time and remove access to any copies of the Springer Nature journal content which have been saved.

To the fullest extent permitted by law, Springer Nature makes no warranties, representations or guarantees to Users, either express or implied with respect to the Springer nature journal content and all parties disclaim and waive any implied warranties or warranties imposed by law, including merchantability or fitness for any particular purpose.

Please note that these rights do not automatically extend to content, data or other material published by Springer Nature that may be licensed from third parties.

If you would like to use or distribute our Springer Nature journal content to a wider audience or on a regular basis or in any other manner not expressly permitted by these Terms, please contact Springer Nature at

onlineservice@springernature.com