

**Karsten Weber**

ORCID: 0000-0001-8875-2386

Ostbayerische Technische Hochschule Regensburg

# Public Images of Artificial Intelligence: An Overview

**Keywords:** artificial intelligence, media representation, public perception, cross-country comparison, mass media, social media

**Słowa kluczowe:** sztuczna inteligencja, przedstawianie w mediach, odbiór społeczny, porównanie między krajami, media masowe, media społecznościowe

## Introduction

Artificial intelligence (AI) is not only a ubiquitous topic in scholarly debates but in recent years has also received increased media and political attention. Debates about AI and its (potential) social impacts have long since ceased to be purely academic. Discussions, initially strongly influenced by autonomous vehicles but currently dominated by expectations concerning Large Language Models such as ChatGPT, have contributed to public awareness; yet, AI and robotics often are conflated, as are AI and autonomous vehicles as well as AI and big data – in short, everything that seems to somehow fit into the even broader and even more vaguely defined narrative of digitalization is associated with AI. This mixture increases media presence and alleged relevance for society but of course does not guarantee that questions and challenges raised by technologies like AI will be discussed properly.

Scientific and technological issues do go through cycles of debate, often characterized by a significant decline in public attention after a certain period of time, and the issue to a large extent or completely disappears from mass media's agenda – until the cycle may start all over again. In this regard, Naudé (2019, p. 3) writes about AI: “Between 1956 and 2007, however, research and business interest in AI waxed and waned. AI winters, when interest and funding in AI declined, occurred in the 1970s and late 1980s until about 2006.” To some extent, these ups and downs are like *hype cycles* (e.g., Dedehayir and Steinert, 2016; Steinert and Leifer, 2010). At the same time, debate cycles are often specific to certain countries; it is very likely that public debates on AI are somewhat different in, for instance, Germany, France, Poland, or the Netherlands. A better

understanding of the existence and course of such debate cycles and their causes could probably contribute to a better understanding of the topic itself.

There are certainly different ways to approach AI and the associated debates on social and ethical concerns, and thus to contribute to a better understanding of the above-mentioned debate cycles. One would be to examine the history of AI and the historical relationship between design, scholarly discussions of its social and normative implications, and the public perception thereof. Even a cursory reconstruction of AI's history would have the advantage of separating public from scholarly debates, at least analytically. This would allow the many nuances and complex issues that in science are at the forefront, but which tend to be pushed into the background in the public debate, to be made visible again and subjected to critical reflection. In this sense, the statement of Alison E. Adam (1990, p. 237) is still valid more than thirty years later:

As several authors have criticized AI and speculated how it may affect our lives, it becomes doubly important to address the history of AI technology from a non-deterministic stance in order to offer something more than existing histories of computer technology. Rather than asking how AI has affected society and changed the nature of work, it is vital to ask instead what choices have been made in the type of AI technology that has been developed and who has made these choices. This view accords more control to the individual, thus making the technology appear less sinister and less mystifying and ultimately more open to criticism.

In this paper, the emphasis is on media representations of AI. The use of AI in healthcare will be particularly highlighted, as there are currently high expectations regarding the benefits of AI which are communicated both in publications addressing healthcare professionals as well as the public. In addition, the deployment of AI in healthcare often raises particular concerns on the part of potential users and stakeholders since digital technology in general and AI in particular are considered to have the potential to replace humans in providing healthcare services; this is considered to be a major challenge for the professionals in this field of application as well as to the patients and other stakeholders. Also, in healthcare privacy seems to be of particular importance to all stakeholders – albeit for different reasons – and data-driven AI might deeply invade this privacy. Finally, the use of AI in healthcare seems to particularly challenge the notion of autonomy of patients as well of healthcare professionals. In general, deploying AI in healthcare creates relationships between human beings and technology that is much more intimate than, for instance, using AI in the production of goods and services in industry.

## Media representation and societal impact

A comprehensive empirical study of how AI is represented in the media was beyond the scope of this paper; the following is based on a rather cursory evaluation of existing studies and is meant to be a first exploratory approach to the issue. It is also important to note that a distinction must be drawn between how AI affects the public and how AI is perceived or portrayed in the public. For example, it is quite likely that the subjective perceptions of journalists reporting on AI (may) differ from objective facts.

One problem with focusing on media representations of AI is that studies exist only for some countries. Therefore, the following statements should always be taken with a grain of salt – it is important not to generalize uncritically, as cultural and social differences are likely to have a strong influence on attitudes towards AI, and this in turn has a considerable impact on their representation in the relevant media.

When talking about AI, the term is used rather vaguely – even a quick glance at the science sections or feature pages of daily and weekly newspapers as well as other news outlets shows that there is usually no distinction between weak and strong AI, nor between symbolic and connectionist approaches; big data, machine learning, and much more are lumped together without much differentiation; often, it is not taken into account that much of what was developed in the first heyday of AI research in the 1960s to 1980s is not even perceived as AI today. Zhai et al. (2020) illustrate this situation when, based on an extensive analysis of print media, they write:

First of all, robots, as a concrete entity from imagination to realization, have been a hotspot in news since 1980s. Speech recognition and autonomy, as the keywords related to robotics, have also attracted the attention of the mass media in the past 40 years. Second, LISP and Prolog, as two major computer languages, attracted much attention in the early stage. [...] Finally, driverless cars, big data and machine learning have gradually become hot news in recent years.

One more limitation must also be emphasized: When speaking of media representation, reference is primarily made to mass media; this is sometimes deviated from when reference is made to social media. For a more comprehensive and less tentative overview it would also be necessary to investigate the impacts of fictional representations of AI on public images of this technology. It can be assumed that representations of AI (e.g., HAL 9000 in “2001: A Space Odyssey”), robots (e.g., Gort from the film “The Day the Earth Stood Still”, Robby from the film “Forbidden Planet” or Wall-E from the film of the same name), or androids (e.g., Data from the TV series “Star Trek: The Next Generation” or the T-800 in the “Terminator” franchise) has seeped into the collective memory of many societies and has influenced attitudes towards technology. At least, studies on representations of other social groups in movies suggest that they inform the public, as Petra Pansegrau (2009, p. 374) points out regarding science and scientists:

The transformation of science and technology has always evoked ambivalent ideas, and so science and everyday scientific life regularly find their way into feature films. This happens much more regularly than is usually assumed: even a cursory search for science as a film motif yields well over 400 films. In these films, scientists or real scientific issues or warnings are usually at the center of a completely fictional plot, but they still convey a supposedly realistic idea of everyday scientific life (my translation).

A similar point could be made about AI, which have appeared regularly on screen since the 1950s in the form of intelligent computers or robots. John Belton (1994, p. xxi) is certainly right in saying that “films [...] cannot be viewed as simple mirrors of cultural reality”, but that “cinema reveals [...] something about [national] experience, identity and culture.” If one follows this view, then many stories and films about power-hungry computers or killer robots tell us something about the mental condition of a society in which such films are produced and/or consumed. For robots, which in the public perception often are associated or even conflated with AI, there are empirical findings:

While expectations about a robot’s capabilities seem to be based on interactional situations and the physical form a robot takes, there is no doubt that cultural artifacts such as science fiction novels and films play a role in people’s understanding of robots. Science fiction films, television, and literature commonly portray robots, and these fictional depictions have been found to contribute to people’s expectations about real robots. For instance, researchers have found that people often refer to science fiction films and books when they are asked to discuss robots (Kriz et al., 2010, p. 458).

On the next page of their paper, the authors become even more explicit:

Because most Americans have not ever interacted with a real robot, it is likely that their understanding of robots and their expectations about robots’ capabilities are at least in part based on what they know about fictional robots. Thus, it is probable that robots such as *The Terminator*, *R2D2*, and *Wall-E* contribute in some way to Americans’ understanding of robots and robotic technologies (Kriz et al., 2010, p. 459).

Another issue that might be investigated images of and beliefs about AI is how media representations of AI affect different stakeholders beyond the public (cf., Ouchchy et al., 2020). This is certainly another question that would have to be answered primarily with the help of empirical data, although it would be methodologically difficult to prove causal relationships. Nevertheless, one can think about basic forms of causal links (Cave and Dihal, 2019, p. 74):

There are at least three ways in which these narratives could shape the technology and its impacts. First, they could influence the goals of AI developers. [...] Second, narratives could influence public acceptance and uptake of AI systems [...]. Third, narratives could influence how AI systems are regulated, as they shape the views of both policymakers and their constituents.

Accepting this systematization, it must be added that such chains of effects on the one hand are certainly not a one-time affair and, on the other hand, have

repercussions: Assuming, for example, the first mode of impact from the above quote, it is very likely that the design of AI influenced by media representations will in turn inform their future media representations, and the process will begin anew (cf. Schmitz et al., 2008). In the end, one will be confronted with a network of cause-and-effect relationships rather than a simple causal chain. This alone will not facilitate the empirical demonstration of concrete influences. Yet, in a most recent study, Brewer et al. (2024) were able to show that viewing science fiction is a predictor regarding beliefs and opinions about AI.

Even if fictional representations of AI should not play a major role here, nevertheless one can draw on a systematization developed by Cave and Dihal (2019). Based on their analysis of a large number of (fictional) representations, they identified four pairs or dichotomies of hopes and fears regarding AI. These are:

1. “immortality versus inhumanity”,
2. “ease versus obsolescence”,
3. “gratification versus alienation”, and
4. “dominance versus uprising”.

The first dichotomy addresses the hope that AI could contribute to the fulfillment of the old human dream of immortality by, for example, loading a person’s mind into a computer, where it would continue to exist virtually forever. Further hopes are that such a digitized mind could be provided with a synthetic body (which would not necessarily have to have a human shape), so that not only the mental but also the physical existence of a person could in principle be guaranteed forever. However, this also raises the fear of inhumanity, since immortality would come at the price of an inhuman (or perhaps better: non-human) existence.

By contrasting “ease versus obsolescence” as second dichotomy, Cave and Dihal (2019) emphasize that AI systems could make people’s lives easier to an unprecedented extent by taking over tasks previously performed by humans; conversely, however, this also raises fears of becoming completely obsolete in the labor process. Indeed, concerns about the possibility of job losses are being mentioned in various disciplines (e.g., Ernst et al., 2018; Ford, 2013). Before the COVID-19 pandemic made headlines, the threat of massive job losses due to AI was repeatedly reported in newspapers, magazines, and on television. It should be noted, however, that in western countries similar debates have been taking place since the 1970s, in Germany, for example, under headings such as “workmate computer” or “workmate robot”.

According to Cave and Dihal (2019), AI also raises the hope of instant gratification of all desires, not least with regard to intimacy and sexuality, because androids could read all the wishes of their human masters (and here, for once, gender is really important) from their eyes and fulfill them compliantly – since they are machines, no one’s rights would be violated. However, this conclusion is quite controversial (e.g., Danaher and McArthur, 2017; Richardson, 2016; Scheutz and Arnold, 2016; Sullins, 2012) since some scholars argue that robots and androids might have some rights. The underlying fear is that people will become obsolete not only in the labor process, but also in the context of interpersonal

relationships, and thus become alienated from one another. If one searches for content on this topic on the Internet, one finds recent television and newspaper reports (cf. Javaheri et al., 2020) but the number is still limited – maybe editors are cautious because of the subject’s slippery nature. However, Sophie Wengerscheid emphasizes that this topic is actually not new (2018, p. 37):

Ever since Pygmalion succeeded in creating the perfect lover, the idea of intimate relationships between humans and artificially created beings has become more and more popular, especially in the 21st century. While engineers and computer scientists are still hard at work on the technological development of robots with humanlike capacities, contemporary science fiction film and literature has already been showing us a variety of humans and posthumans interacting with each other intensely and entering into posthuman love affairs.

Finally, according to Cave and Dihal (2019), AI could reshape or reinforce power relations, for example in the context of international conflicts or even wars. However, this is also linked to the fear that machines will eventually turn against humans, and possibly subjugate them.

At first glance, these four dichotomies appear quite compelling, but it remains to be seen whether media representations actually follow similar patterns. Debates of “immortality versus inhumanity” (e.g., Cutas and Harris, 2006) already appeared in connection with so-called cyborgs and genetic engineering; technology has been discussed in terms of “ease versus obsolescence” since the first industrial revolution at the latest; “gratification versus alienation” are part of considerations in media theory, particularly regarding audiovisual media; and even “dominance versus uprising” already occurred when Taylorism and Fordism were criticized. This could be an indication that media representations of AI are not specific to this kind of technology, but that they reflect both the hopes and fears that have been with people since they began to use the first simple machines to improve their lives.

## Knowledge and attitudes about AI

As mentioned above, one problem with media representations of AI and their effects is that little data is available for many countries, that existing studies used rather different methods, and that the transferability of corresponding studies from one country to other countries must be questioned. However, despite these methodological problems, the studies mentioned here can at least provide first insights.

Cave et al. (2019) report on an online survey as part of an online marketing panel with over 20,000 participants who are repeatedly asked about various topics. Compared to standard online surveys with a convenience sample, this approach has the advantage that there is no self-selection, and that representativeness can be guaranteed; 1,078 people completed the questionnaire. Questions were asked about knowledge of AI and about the four dichotomies mentioned above.

According to the results, knowledge of AI is very widespread, at least among respondents of that survey, as 85% said they had heard of AI – age differences were (surprisingly) small. Comparing this level of knowledge with, for instance, Germany, the results of a survey conducted by Fischer and Petersen (2018) are rather sobering, as many people there know little to nothing about algorithms that ultimately implement AI (Fischer and Petersen, 2018, p. 3):

For Germany, the survey shows that a broad societal discussion beyond the professional discourse and the relevant interested parties has not yet begun. Rather, ignorance, indecision, and unease dominate the debate on algorithms in Germany. Very few people have a clear opinion on the subject, but they are very much against decisions made by or with the help of algorithms (my translation).

Although not representative for a whole society, a survey among Russian students (Merenkov et al., 2021) reports rather similar results. Even well-educated young people like students seem to be uninformed as almost a quarter of the respondents in this survey did not know AI.

According to Fischer and Petersen, lack of knowledge seems to be reflected in discomfort and rejection. Indeed, this finding is supported by a study commissioned by the Association of German Banks (Bundesverband deutscher Banken) (Knorre et al., 2020, p. 13):

A similarly representative survey conducted by GfK for the Association of German Banks in June 2018 [...] comes to considerably more skeptical conclusions. According to this survey, 75% of Germans are familiar with the term, but one in four has never heard of it. In contrast to the results of the Bitkom survey, the majority of Germans associate it with fears (63%), while only 37% see opportunities. [...] The conclusion of this survey is that mistrust of digitally controlled processes is generally still high (my translation).

Knorre et al. also refer to other studies; their findings could be summarized by saying that many people in Germany have basic knowledge of the subject, but largely lack detailed knowledge. In addition, there is an undecided to rather skeptical attitude towards AI (Knorre et al., 2020, p. 14):

Despite some differences in the design of the surveys and their results, there is – not surprisingly – no clear formation of opinion on the use of big data, artificial intelligence, and algorithms. The surveys show that a clear majority of respondents are uneasy about big data and artificial intelligence, having a vague fear of losing control, at least a wait-and-see attitude, but in some cases an almost “fatalistic” attitude (my translation).

This does not seem to be the case to the same extent in the UK, as Cave et al. (2019) write that the fears of their respondents fall far short of their positive expectations; Gherheș and Obrad (2018) show something similar for Romania. Even if this can only be seen as an indication, it is already clear that one cannot and should not simply draw conclusions about people in another country from respondents in one country. Cave et al. (2019, p. 336) also emphasize this in the



four recommendations they derive from their study for further research, calling for an investigation of the public perception of AI across cultural boundaries.

Cave et al. (2019) explicitly emphasize that at least the people they interviewed had predominantly positive attitudes towards AI and that only a small fraction assumed that the scenarios implied by the aforementioned dichotomies could occur in a negative way. However, this contrasts with other statements, as Colin Garvey (2018, p. 1) stresses that

[i]n my anthropological fieldwork among AI tribes, one trait common to most groups was the belief that negative media coverage of AI (and specifically, the use of Terminator imagery) is to blame for public concerns about AI. I often heard the lament that if laypeople could only see what AI people do in the lab, they would have no reason to be concerned. In other words, concerns about AI arise only because an ignorant public is being propagandized to fear AI by an unscrupulous media.

Referring to the android played by Arnold Schwarzenegger in several movies, Garvey calls this the “Terminator syndrome” and says that it is based on three assumptions:

1. media coverage of AI is overwhelmingly negative,
2. negative coverage directly determines public opinion on AI, and
3. public opinion on AI is driven by irrational fears.

Garvey tested these assumptions in an exploratory study. Despite its limitations, which the author himself explicitly points out, he was unable to find any confirmation of the Terminator syndrome; the majority of the well over 3,000 media articles included in the study show neutral or positive attitudes toward AI, and only about a quarter contain negative attitudes. This suggests that the public image of AI – at least in the U.S. – does not seem to be characterized by the Terminator syndrome; the results of this study are also corroborated by an investigation of newspaper articles of a 30-year period by Fast and Horvitz (2017). However, these results contrast with a prediction made by Zhai et al. (2020), who suggest that an increase in critical to negative media reports about the effects of AI is to be expected, and that this will require scientists involved in AI development to pay more attention to normative issues: “It can be expected that, with the increasing number of negative reports on AI in the media, researchers will have more and more discussions on AI ethics, and more people will join the discussion on AI design.” In this respect, the comparison with Germany is quite revealing, as what Zhai et al. predict seems to be already a reality there. Even if the corpus of 169 media articles examined, collected over the course of a year from major German daily and weekly newspapers and news magazines, does not yet allow for too far-reaching generalizations, the conclusions of Knorre et al. (2020, p. 16) seem rather plausible: “It becomes clear that the public discourse on the handling of personal data is dominated by the ‘big brother’ narrative, which is firmly rooted at least in Western industrialized countries” (my translation).



Although the authors limit their analysis to the issue of the handling of personal data and, as noted, the dataset was somewhat narrow, their historical reconstruction of the genesis of this narrative is convincing; specific events that took place in Germany rather recently and have cultural, political, and social references to even more distant periods determine the perception of data protection in Germany in general and in the context of AI. Conversely, in countries where certain historical experiences do not exist (or, better to say, does not exist in the same way) – particularly the Nazi dictatorship from 1933 to 1945 and the experience of a socialist dictatorship from 1949 to 1989 in East Germany – certain associations with AI, in this case that of “big brother,” may not have the same impact as in Germany.

### **Strategic communication or deliberate misdirection?**

The statement by Knorre et al. (2020, p. 29) that “[...] it may seem opportune to focus on the term artificial intelligence instead of big data, which is presumably less loaded with anxiety and is often used synonymously with big data anyway” (my translation) must come as a surprise. This passage might indicate that the authors suspect that, given the considerable ambiguity in public debates on AI, some stakeholders involved might be tempted to exploit this ambiguity for their own purposes. Since Knorre et al. actually provide evidence for this assumption, it could be assumed that the public image of AI is also influenced by the strategic considerations of at least some stakeholders.

However, such an interpretation also runs the risk of assuming conspiratorial tendencies, although it would perhaps be more obvious and reasonable to assume that this is a manifestation of stakeholders’ frustration and, above all, interests. For instance, while Herbert Weber (2017) writes about digitalization, what he writes could also be applied to AI – and he mentions AI as part of digitalization. Weber (2017, p. 7) states under the headline “Digitalization as camouflage”, without citing any evidence: “‘Digitalization’ is probably the deliberately chosen wrong term for what is really meant by its use in parts of the public discourse” (my translation). Thus, a deliberate misuse is assumed, the reasons for which Weber sees in the aim of breaking up social, economic, and bureaucratic dependencies; however, as already mentioned, no evidence for this assumption is provided. There is therefore at least a risk of misunderstanding, because if one assumes a deliberate abuse, the question of the beneficiaries and their goals immediately arises.

Weber continues and writes that “[i]f one continuously follows the public discourse, one can only come to the conclusion that there is not only reluctance, but perhaps even obstruction in society towards digitization, and that this is a significant problem in society” (ibid., my translation); this is biased at best, but seems more like a conspiracy theory – which, however, was printed in an edited book published by a renowned publisher, not on some obscure website.

There is no doubt that such isolated views should not be overestimated, but neither should they be completely disregarded, since the path from reasonable doubt to conspiracy theory can be quite short; for that, the COVID-19 pandemic and the violent protests against containment measures that repeatedly occurred in numerous countries are one example; the events surrounding the 2020 U.S. presidential election and the unfounded allegations of widespread electoral fraud is another. If some stakeholders continuously insinuate that AI is a tool for sinister purposes, it is at least conceivable that a militant social movement will develop that sees dark forces at work behind the use of AI.

A preliminary conclusion that can be drawn from such considerations is that transparency is paramount if initial public skepticism about AI is not to turn into (militant) opposition. For example, efforts regarding “explainable AI” can be interpreted in this way: Only when all stakeholders can know, at least in principle, how AI in general and a specific AI system in particular works and produces results – i.e., is not a black box – will these stakeholders no longer be influenced in their decisions and attitudes solely by irrational assumptions, but, at least to some extent, by well-founded knowledge (cf., Araujo et al., 2020; de Fine Licht and de Fine Licht, 2020; Felzmann et al., 2019).

## Artificial intelligence in healthcare

Some scholarly publications on AI in healthcare – understood here as the combination of medicine, nursing, and other health-related professions – include rather positive assessments of the (potential) benefits of AI:

Artificial intelligence (AI) has been used for image analysis in dermatology, pathology and radiology with good accuracy and fast diagnostic speed. AI helps decrease medical errors, recommend precision therapies for complex diseases, optimise the care procedures of chronic illnesses and increase patient enrolment into clinical trials (Wang and Alexander, 2020, p. 19).

Numerous authors even assume that, at least for some areas of medicine, physician activities will be largely or completely taken over by AI systems (e.g., Jannes et al., 2018; Hardy and Harvy, 2020; Obermeyer and Emanuel, 2016). According to a survey by Doraiswamy et al. (2020), 83% of respondents (N=791) believe it is unlikely that AI will have the skills to take over key tasks in this area. But at the same time, more than half of the respondents believe it is likely that AI systems will take over diagnostic and prognostic tasks, and three-quarters believe that AI will take over documentation tasks.

These statements are based on professional judgment. In the context of this paper, however, it is important to ask how patients (will) react to AI in healthcare, since it is not least their acceptance that will determine whether and how quickly such systems will disseminate (cf. Young et al., 2021). Here, a look at the U.S. provides interesting insights. Stai et al. (2020) write about their study based on a convenience sample:

Participants had nearly equal trust in AI vs. physician diagnoses, however, they were significantly more likely to trust an AI diagnosis of cancer over a doctor's diagnosis when responding to the version of the survey that suggested an AI could make medical diagnosis [...]. [...] Almost all (94%) stated they would be willing to pay for an AI to review their medical imaging, if available.

In contrast to these results, Rojahn et al. (2023) report that U.S. citizens would prefer human doctors to make medical decisions. Comparing these results with the above-mentioned statements by Knorre et al. (2020) on attitudes of the German population towards AI, it should be noted that in contrast to the preliminary findings, attitudes in Germany are significantly more skeptical and negative than in other countries. Gao et al. (2020), based on their own social media analysis of a corresponding platform in China, point out that there, attitudes towards AI systems in medicine have improved compared to previous studies. They also find that the majority of social media posts they examined, which were thematically aligned accordingly, assume that AI systems will either fully or partially replace doctors. Thus, at least in their study, the expectations of scientists and those persons potentially affected by AI seem to be in line.

Since Wang and Alexander (2020) focus specifically on the potential replacement of radiologists, it is interesting that Gupta et al. (2020) take up and analyze this topic. They conclude that press reports, which then often influence the content of social media posts, are usually written by people who have only limited knowledge about the functioning, capabilities, and limitations of AI in medicine. This could lead to the conclusion that public opinion, which is formed or expressed both in mass media and social media, is in some cases not influenced by factual information but rather by incomplete knowledge and possibly also prejudices. Gupta et al. therefore recommend that members of the relevant professions should engage in such media debates.

## Conclusion

When new technologies are introduced, it is often said that stakeholders need to build trust in order to accept its use (e.g., Zhang and Allan, 2020). This hypothesis can also be found with regard to AI in medicine (e.g., Araujo et al., 2020; Glikson and Woolley, 2020; Hengstler et al., 2016). However, it is not always clear by whom and how this trust should be established, as it involves technological, regulatory, moral, cultural and, of course, medical considerations. Yu and Kohane (2019, p. 240) go so far as to say that building AI systems for medical purposes is the rather easy part:

Building an intelligent automated entity to evaluate, diagnose and treat patients in research settings is arguably the easiest part of designing an end-to-end medical AI system. In the context of the hype and hopes surrounding emerging AI applications in medicine, we need to acknowledge the brittleness of these systems, the importance of defining the correct frameworks for their

application, and ensure rigorous quality control, including human supervision, to avoid driving our patients on autopilot towards unexpected, unwanted and unhealthful outcomes.

This quote sums up what can be said about AI systems in medicine – and, by extension, in other fields of application. However, if experts are already warning that the road to safe and beneficial AI systems will be a long one, it should come as no surprise that in many countries the public sometimes is making some very exaggerated statements about AI, both positive and negative. These attitudes are not necessarily based on reliable knowledge, but on expectations, fears and hopes, and perhaps even prejudices – on the part of potential patients, journalists as well as decision-makers in politics, government, and business.

It must be emphasized that being skeptical regarding AI in itself should not be judged as being irrational even if this skepticism is not based on reliable knowledge (cf. Luca Liehner et al., 2023) – “better safe than sorry”. Yet, from the point of view of public understanding of science and science communication, it must give cause for concern when images of technology in general and of AI in particular are often characterized by Manichean dichotomies, as reported by Cave and Dihal (2019). While it would be an important task to induce more differentiated images in this regard, much more comprehensive and systematic research than the cursory overview presented in this paper can provide is needed to identify the mechanisms that influence the public image of AI across countries more precisely.

## References

- Adam, Alison E. (1990). What can the history of AI learn from the history of science? *AI & Society*, 4(3), 232–241.
- Araujo, Theo; Helberger, Natali; Kruikemeier, Sanne and de Vreese, Claes H. (2020). AI we trust? Perceptions about automated decision-making by artificial intelligence. *AI & Society*, 35(3), 611–623.
- Belton, John (1994). *American Cinema/American Society*. New York: McGraw-Hill.
- Brewer, Paul R.; Cuddy, Liam; Dawson, Wyatt and Stise, Robert (2024). Artists or art thieves? Media use, media messages, and public opinion about artificial intelligence image generators. *AI & Society*, <https://doi.org/10.1007/s00146-023-01854-3>.
- Cave, Stephen and Dihal, Kanta (2019). Hopes and fears for intelligent machines in fiction and reality. *Nature Machine Intelligence*, 1(2), 74–78.
- Cave, Stephen; Coughlan, Kate and Dihal, Kanta (2019). ‘Scary robots’: Examining public responses to AI. *Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society*. Honolulu/U.S.: ACM, 331–337.
- Cutas, Daniella and Harris, John (2006). The ethics of ageing, immortality and genetics. In: Ashcroft, Richard E.; Dawson, Angus; Draper, Heather and McMillan John R. (eds.), *Principles of health care ethics*. Chichester: Wiley, 797–801.
- Danaher, John and McArthur, Neil (eds.) (2017). *Robot sex: Social and ethical implications*. Cambridge/Massachusetts: The MIT Press.
- de Fine Licht, Karl and de Fine Licht, Jenny (2020). Artificial intelligence, transparency, and public decision-making: Why explanations are key when trying to produce perceived legitimacy. *AI & Society*, 35(4), 917–926.

- Dedehayir, Ozgur and Steinert, Martin (2016). The hype cycle model: A review and future directions. *Technological Forecasting and Social Change*, 108, 28–41.
- Doraiswamy, P. Murali; Blease, Charlotte and Bodner, Kaylee (2020). Artificial intelligence and the future of psychiatry: Insights from a global physician survey. *Artificial Intelligence in Medicine*, 102, 101753.
- Ernst, Ekkehardt; Merola, Rossana and Samaan, Daniel (2018). *Economics of artificial intelligence: Implications for the future of work*. ILO Future of Work Research Paper Series. Geneva: International Labour Office.
- Fast, Ethan and Horvitz, Eric (2017). Long-term Trends in the public perception of artificial intelligence. *Proceedings of the AAAI Conference on Artificial Intelligence*, 31(1), 963–969.
- Felzmann, Heike; Villarronga, Eduard F.; Lutz, Christoph and Tamò-Larrieux, Aurelia (2019). Transparency you can trust: Transparency requirements for artificial intelligence between legal norms and contextual concerns. *Big Data & Society*, 6(1), 205395171986054.
- Fischer, Sarah and Petersen, Thomas (2018). *Was Deutschland über Algorithmen weiß und denkt: Ergebnisse einer repräsentativen Bevölkerungsumfrage*. Gütersloh: Bertelsmann-Stiftung.
- Ford, Martin (2013). Could artificial intelligence create an unemployment crisis? *Communications of the ACM*, 56(7), 37–39.
- Gao, Shuquin; He, Lingnan; Chen, Yue; Li, Dan and Lai, Kaisheng (2020). Public perception of artificial intelligence in medical care: Content analysis of social media. *Journal of Medical Internet Research*, 22(7), e16649.
- Garvey, Colin (2018). *Testing the ‘Terminator Syndrome’: Sentiment of AI news coverage and perceptions of AI risk*. SSRN Electronic Journal.
- Generali Deutschland AG (ed.) (2017). *Generali Altersstudie 2017*. Berlin, Heidelberg: Springer.
- Gherheș, Vasile and Obrad, Ciprian (2018). Technical and humanities students’ perspectives on the development and sustainability of artificial intelligence (AI). *Sustainability*, 10(9), 3066.
- Glikson, Ella and Woolley, Anita W. (2020). Human trust in artificial intelligence: Review of empirical research. *Academy of Management Annals*, 14(2), 627–660.
- Gupta, Sonia; Kattapuram, Taj M. and Patel, Tirath Y. (2020). Social media’s role in the perception of radiologists and artificial intelligence. *Clinical Imaging*, 68, 158–160.
- Hardy, Maryann and Harvy, Hugh (2020). Artificial intelligence in diagnostic imaging: Impact on the radiography profession. *British Journal of Radiology*, 93(1108), 20190840.
- Hengstler, Monika; Enkel, Ellen and Duelli, Selina (2016). Applied artificial intelligence and trust – The case of autonomous vehicles and medical assistance devices. *Technological Forecasting and Social Change*, 105, 105–120.
- Jannes, Mark; Friele, Minou; Jannes, Christiane and Wopen, Christiane (2018). *Algorithmen in der digitalen Gesundheitsversorgung. Eine interdisziplinäre Analyse* (im Auftrag der Bertelsmann Stiftung). Köln: ceres.
- Javaheri, Alireza; Moghadamnejad, Navid; Keshavarz, Hamidreza; Javaheri, Ehsan; Dobbins, Chelsea; Momeni-Ortner, Elaheh and Rawassizadeh, Reza (2020). Public vs media opinion on robots and their evolution over recent years. *CCF Transactions on Pervasive Computing and Interaction*, 2(3), 189–205.
- Knorre, Susanne; Müller-Peters, Horst and Wagner, Fred (2020). Big Data im öffentlichen Diskurs: Hindernisse und Lösungsangebote für eine Verständigung über den Umgang mit Massendaten. In: Knorre, Susanne, Müller-Peters, Horst and Wagner, Fred (eds.), *Die Big-Data-Debatte*. Wiesbaden: Springer Fachmedien, 1–62.
- Kriz, Sara; Ferro, Toni D.; Damara, Pallavi and Porter, John R. (2010). Fictional robots as a data source in HRI research: Exploring the link between science fiction and interactional expectations. *19<sup>th</sup> International Symposium in Robot and Human Interactive Communication*. Viareggio: IEEE, 458–463.
- Luca Liehner, Gian; Biermann, Hannah; Hick, Alexander; Brauner, Philipp and Ziefle, Martina (2023). Perceptions, attitudes and trust toward artificial intelligence—An assessment of the public opinion. *Artificial Intelligence and Social Computing*, 72, 32–41.
- Merenkov, A. V.; Campa, R. and Dronishinets, N. P. (2021). Public opinion on artificial intelligence development. *KnE Social Sciences, Culture, Personality, Society in the Conditions of Digitalization: Methodology and Experience of Empirical Research Conference*, 565–574.

- Naudé, Wim (2019). *The race against the robots and the fallacy of the giant cheesecake: Immediate and imagined impacts of artificial intelligence*. IZA Discussion Papers. Bonn: Institute of Labor Economics (IZA).
- Obermeyer, Ziad and Emanuel, Ezekiel J. (2016). Predicting the future—big data, machine learning, and clinical medicine. *New England Journal of Medicine*, 375(13), 1216–1219.
- Ouchchy, Leila; Coin, Allen and Dubljević, Veljko (2020). AI in the headlines: The portrayal of the ethical issues of artificial intelligence in the media. *AI & Society*, 35(4), 927–936.
- Pansegrau, Petra (2009). Zwischen Fakt und Fiktion – Stereotypen von Wissenschaftlern in Spielfilmen. In: Hüppauf, Bernd and Weingart, Peter (eds.), *Frosch und Frankenstein*. Bielefeld: transcript, 373–386.
- Richardson, Kathleen (2016). Sex robot matters: Slavery, the prostituted, and the rights of machines. *IEEE Technology and Society Magazine*, 35(2), 46–53.
- Rojahn, Jessica; Palu, Andrea; Skiena, Steven and Jones, Jason J. (2023). American public opinion on artificial intelligence in healthcare. *PLOS ONE*, 18(11), e0294028.
- Scheutz, Matthias and Arnold, Thomas (2016). Are we ready for sex robots? *11th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. Christchurch/New Zealand: IEEE, 351–358.
- Schmitz, Michael; Endres, Christoph and Butz, Andreas (2008). A survey of human-computer interaction design in science fiction movies. *Proceedings of the 2<sup>nd</sup> international conference on INtelligent TEchnologies for interactive entertainment (ICST INTETAIN '08)*. Brussels: ICST, 1–10.
- Stai, Bethany; Heller, Nick; McSweeney, Sean; Rickman, Jack; Blake, Paul; Vasdev, Ranveer; Edgerton, Zach; Tejpaul, Resha; Peterson, Matt; Rosenberg, Joel; Kalapara, Arvin; Regmi, Subodh; Papanikolopoulos, Nikolaos and Weight, Christoph J. (2020). Public perceptions of artificial intelligence and robotics in medicine. *Journal of Endourology*, 43(10), 1041–1048.
- Steinert, Martin and Leifer, Larry (2010). Scrutinizing Gartner's hype cycle approach. *Proceedings of PICMET 2010 Technology Management for Global Economic Growth*. Phuket/Thailand: IEEE, 1–13.
- Sullins, John P. (2012). Robots, love, and sex: The ethics of building a love machine. *IEEE Transactions on Affective Computing*, 3(4), 398–409.
- Wang, Lidong and Alexander, Cheryl A. (2020). Big data analytics in medical engineering and healthcare: Methods, advances and challenges. *Journal of Medical Engineering & Technology*, 44(6), 267–283.
- Weber, Herbert (2017). Digitalisierung im öffentlichen Diskurs. In: Weber, Herbert and Viehmann, Johannes (eds.), *Unternehmens-IT für die Digitalisierung 4.0*. Wiesbaden: Springer Fachmedien, 3–13.
- Wennerscheid, Sophie (2018). Posthuman desire in robotics and science fiction. In: Cheok, Adrian D. and Levy, David (eds.), *Love and sex with robots*. Cham/Switzerland: Springer International Publishing, 37–50.
- Young, Albert T.; Amara, Dominic; Bhattacharya, Abhishek and Wei, Maria L. (2021). Patient and general public attitudes towards clinical artificial intelligence: A mixed methods systematic review. *The Lancet Digital Health*, 3(9), e599–e611.
- Yu, Kun-Hsing and Kohane, Isaac S. (2019). Framing the challenges of artificial intelligence in medicine. *BMJ Quality & Safety*, 28(3), 238–241.
- Zhai, Yujia; Yan, Jiaqi; Zhang, Hezhao and Lu, Wei (2020). Tracing the evolution of AI: Conceptualization of artificial intelligence in mass media discourse. *Information Discovery and Delivery*, 48(3), 137–149.
- Zhang, Baobao and Dafoe, Allan (2020). U.S. public opinion on the governance of artificial intelligence. *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society*. New York: ACM, 187–193.



## Summary

Artificial intelligence (AI) is not only a ubiquitous topic in scholarly debates but, in recent years, has also received increased media and political attention. Discussions, initially strongly influenced by autonomous vehicles but currently dominated by expectations concerning Large Language Models such as ChatGPT, have contributed to public awareness. There are debate cycles, often characterised by a significant decline in public attention after a certain period of time, and the issue, to a large extent or completely disappears from mass media's agenda, and they are often specific to certain countries. In this paper, a cursory and tentative cross-country overview concerning studies on media representations and the public perception of AI is provided. The use of AI in healthcare is particularly highlighted, as there are currently particularly high expectations regarding the benefits of AI, but other areas of application are also considered. It was tentatively concluded that in many countries, public perceptions and public attitudes toward AI are often based on superficial knowledge and even prejudices.

## Postrzeżanie sztucznej inteligencji – przegląd literatury

### Streszczenie

W ostatnich latach sztuczna inteligencja (ang. Artificial Intelligence, AI) jest częstym tematem debat naukowych, przyciąga również uwagę mediów, polityków. Dyskusje, początkowo koncentrujące się na autonomicznych pojazdach, a obecnie zdominowane przez oczekiwania dotyczące dużych modeli językowych, takich jak ChatGPT, przyczyniły się do wzrostu świadomości społecznej na temat AI. W niniejszym artykule przedstawiono pobieżny przegląd badań dotyczących prezentowania sztucznej inteligencji w mediach i postrzegania jej przez społeczeństwo. W pracy, obok innych obszarów zastosowań, szczególnie podkreślono wykorzystanie sztucznej inteligencji w opiece zdrowotnej ze względu na wysokie oczekiwania względem korzyści płynących z jej zastosowania w medycynie. Wstępnie stwierdzono, że w wielu krajach postawy społeczne wobec sztucznej inteligencji często opierają się na powierzchownej wiedzy na jej temat, a nawet na uprzedzeniach płynących z możliwych zagrożeń jej zastosowania.

