

WEIZENBAUM JOURNAL OF THE DIGITAL SOCIETY
Volume 4 \ Issue 1 \ w4.1.6 \ 02-19-2024
ISSN 2748-5625 \ DOI 10.34669/WI.WJDS/4.1.6

Information on this journal and its funding can be found on its website:
<https://wjds.weizenbaum-institut.de>


This work is available open access and is licensed under Creative Commons Attribution 4.0 (CC BY 4.0):
<https://creativecommons.org/licenses/by/4.0/>

KEYWORDS

artificial intelligence
technology promises
myth of intelligent
machine

VOICES FOR THE NETWORKED SOCIETY

Promises and Myths of Artificial Intelligence

Hartmut Hirsch-Kreinsen^{1*} \ Thorben Krokowski¹ 

¹TU Dortmund University

*Corresponding author, hartmut.hirsch-kreinsen@tu-dortmund.de

ABSTRACT

The development dynamics of any new technology are usually associated with promises of its special performance and completely new application possibilities. This is especially true for artificial intelligence (AI), prompting this contribution to inquire into which particular special features the technology promises. However, the imprecise rhetoric of that promise becomes apparent. Although it appears simple, clear, and convincing, it is fundamentally difficult to dispute and introduces multitudes of ambiguity, relying on fuzzy conceptual metaphors, very unspecific assessments, implicit misconceptions, technological determinism, and exaggerations of the future opportunities AI offers for economic and social progress. Ultimately, the promises of AI nourish their lasting persuasive power with notions from the old myth of the intelligent machine.

1 AI Hype

Almost every new technology is linked to promises and expectations regarding its special performance and completely new application possibilities, foreseeing substantial profits and resolutions to often pressing social challenges. This discourse often reaches far into the future with its promises, with the dynamics associated with the promises not infrequently seen as an essential driving force of technological innovation and economic dynamics in general (e.g., Beckert, 2016).

The case of artificial intelligence (AI) offers an exceptional example of a technology whose genesis and diffusion have been repeatedly accompanied by lofty promises. In the past, these types of promises have often proven unrealistic, such that their further development has even threatened to falter in the meantime. For AI, the promises, expectations, and narratives shaping its evolution are apparent in the technology's developmental history (e.g., Nilsson, 2010; Hirsch-Kreinsen, 2023). Meanwhile, AI's rapid and dynamic growth, its foundations, and its possible applications continue to give rise to new promises, which are highly persuasive and raise the expectations and interest of many actors, along with the broader public. Bruce G. Buchanan, an early protagonist of AI in the US and a founding member of the American Association for Artificial Intelligence, sums up this promise-driven dynamic: "The history of AI is a history of fantasies, possibilities, demonstrations, and promises" (Buchanan, 2006, p. 53).

Since at least the beginning of the 2010s, AI development has experienced an unbroken upswing. Following the launch of large language models (LLMs) such as ChatGPT, the discourse on the potential and opportunities of AI has been characterized by exuberant hype. It is no surprise, then, that the general mood is primarily one of fascination with AI's prospects, although the public debate is also characterized by skepticism and anxiety about future social consequences of AI.

2 The False Rhetoric of the Technological Promise

This situation requires asking what special features the technological promise of AI offers and why the rhetoric of this promise is so persuasive and fascinating. It might be that the promise is based on very unspecific assessments, implicit misconceptions, and overlooks fundamental application and functional problems of AI. This rhetoric has shaped the development of AI from its earliest days. The technology promise features a rhetorical architecture characterized by, for example, hypostasizing the new technology, making communicative generalizations and reducing social complexity, and de-contextualizing

the arguments, irrefutable topicality, quantifiability, and relevance to everyday life. Simplifications and generalizations facilitate the pointed communication of a desirable technological future by establishing a link to the experiences and traditional orientations of many actors as well as to already existing discourses on future necessary technological and societal development perspectives. Following the prominent linguist Noam Chomsky, one can also speak of the “false promises” (2023) of AI and, especially, LLMs. This thesis builds upon three main arguments: the ambiguity of the term AI, the myth of technological determinism, and the exaggeration inherent in social and economic promises.

2.1 Ambiguity of the Term

The rhetorical core of AI’s promises ostensibly rests in the ambiguity of the term itself, which is little more than a “very loose umbrella term” (Kalthauer, 2021, p. 23). This is reflected in the fact that although the term has been intensively discussed for decades, it remains a subject area that is difficult to grasp and often controversial, even for experts. At its most coherent, since the 2010s, the term AI has been equated with machine learning and artificial neural networks without specifying individual foundations or modes of operation. Furthermore, the highly specialized applications of realized AI systems are often not assessed as such but are often only associated with the far-reaching promises and visions of artificial general intelligence (AGI) (Ford, 2018).

Thus, AI’s technological promises can be characterized as a sophisticated blend of simplicity and ambiguity. The technological vision appears simple, clear, and convincing, making it difficult to dispute. The prerequisite for this is an easily accessible metaphor linked to AI based on common categories such as “information,” “intelligent,” “learning,” or even “autonomous.” It is an equation of human and machine levels, which critical experts reject as an erroneous ontology of man and machine. These ambiguous metaphors are major reasons for the fascination that emanates from AI, linking it to both everyday notions and wishful thinking about the future of technology.

2.2 Technological Determinism

The discourse on AI, its development, and (above all) its social consequences is highly technology-focused and neglects complex socio-technological contexts of the diffusion and implementation of new technologies, with the apparent inevitability of this diffusion representing an essential component of this fixation on dynamic progress. The discourse generally presents the main drivers of this progress as the continuous development of AI mechanisms such as neural networks, the massive increase in globally available data and

the possibilities for using it, and fast-growing computing power accompanied by dramatically falling costs. An indication of this technology-fixated view is the repeated assertion that AI must be understood as an indispensable and necessary key enabling technology for improving global competitiveness and prosperity levels, as well as for guiding new social transformation processes (e.g., EU, 2020).

These perspectives are based on a fundamentally truncated understanding of the diffusion and implementation conditions of new technologies such as AI. The promise of the technology and the expectations based on it are characterized by a crude technological determinism concerning the processes of diffusion and implementation of AI systems. However, this view neglects the basic findings of innovation and technology research from the social sciences. These findings suggest that there is no clear link between the development of new technologies, their diffusion and implementation processes, and the concrete patterns of their use. Rather, it is a complex and reciprocal relationship shaped by a multitude of non-technical, economic, social, and political factors. Their influence determines which individual diffusion phases a technology will go through, in which way the utilization potential of technologies will be exploited and which consequences for social development will manifest (Rogers, 2003).

2.3 Exaggerated Social and Economic Promises

The technology determinism view implies that the potential of new digital technologies will translate directly into the social and economic reality of their use. Far-reaching prospects for progress and economic growth are expected and continuously publicly emphasized (Ossewaarde & Gülenç, 2020). Meanwhile, the broad range of further possibilities promised includes overcoming the shortage of skilled workers, optimizing road traffic through autonomous driving, improving care in an aging society, enhancing energy efficiency, and contributing to overcoming the climate crisis (e.g., Bareis & Katzenbach, 2021).

These and similar promises admittedly fail to identify which specific AI development steps will be taken to achieve these goals. With promises as far-reaching as they are vague, it is no exaggeration to speak of a hypostatization of AI's capabilities. That is, these promises are largely empty of empirical evidence. Indeed, empirical evidence is frequently not even sought. The economic and sociological forecasts are mainly based on assessments and surveys of scientific and technological experts. This precludes their consideration as hypotheses awaiting verification or realization. Instead, they correspond to drafts of the future that juxtapose imagined alternatives with the present moment. In other words, these predictions about the future development and effects of

digital technologies cannot be validated: “The story is not testable” (Larson, 2021, p. 42).

Additionally, the promises are exaggerated because they neglect factors and conditions that are sociologically classified as side effects. These can stand in the way of intended further developments or even block them completely. They can create new unknown problem areas and challenges that must be addressed anew in terms of innovation policy and may require new types of research and development efforts. Crucial to consider are the uncertain and even contradictory economic effects of a new digital technology, such as the unclear implementation and follow-up costs as well as the ongoing costs of maintaining data. The current AI discourse mostly overlooks these issues. Moreover, technological promises overlook the ways that AI and digital technologies in general always require innovations in the surrounding functional and organizational areas if they are to be functional and successful, that is, indispensable “complementary innovation” (Brynjolfsson & McAfee, 2014).

3 The Myth of the Intelligent Machine

Intertwined with the social focus on digital technologies and the internet that has evolved over decades to produce the prevalent perception of a social megatrend, AI has become a part of the zeitgeist most easily described as the inevitable modernization of society through new technologies.

However, there is much to suggest that the technological promise of AI ultimately draws its lasting persuasive power from the myth of the intelligent machine, which imagines a machine intelligence equal if not superior to human intelligence. The core of this myth is that machines can be created that overcome the physical and intellectual shortcomings of the “deficient being” (Gehlen 1940). The myth of the intelligent machine dates back to antiquity and has always stimulated the visions of philosophers, writers, filmmakers, and artists, among others. It pertains to a long tradition in human history of imaginative stories that are also referred to as AI narratives (Cave et al., 2020). From Ancient Greek myths such as that of the bronze giant Talos who protected Crete to “The Mechanical Turk” imagined by Wolfgang von Kempelen and the computer HAL in Stanley Kubrick’s film “A Space Odyssey” (Marquis et al., 2020), these visions have shaped AI discourse from the very beginning and continue to represent an overarching and generalized development goal. According to Nils Nilsson (2010), human-level AI was already the goal of the founding generation and many other early AI researchers, and it remains the vision of many researchers today, as articulated in the aforementioned formula for AGI (Nilsson, 2010, p. 525).

Accordingly, from the very beginning, AI researchers have aimed to develop models of human cognitive and perceptive abilities and to reproduce these abilities on computers.

Nonetheless, despite the fundamental skepticism towards and criticism of these goals (cf. Kaltheuner, 2021; Larson, 2021), it is indisputable that AI has achieved considerable success in recent decades. It cannot be overlooked that the performance of AI systems is constantly improving, and human abilities have already been significantly surpassed or are about to be surpassed by AI in many fields. However, to conclude from this that a technological breakthrough and a disruptive scenario (cf. Ford, 2018; Struckman & Kupiec, 2023) is imminent and will abruptly initiate far-reaching structural changes in all areas of human life – a steadfast component of the AI myth – is a broad exaggeration of the available facts rather than an objective assessment.

Much more realistic is the position that the future development and spread of AI will probably be incremental. From this perspective, existing AI approaches and methods will evolve, sometimes faster than before, with current limitations and challenges gradually overcome. With AI systems accordingly used on an ever-increasing scale, a trend can be expected that can be called AI routinization, in which AI experiences increasingly become routine and a not inconsiderable part of quotidian systems and, thus, of daily life (Hirsch-Kreinsen, 2023). However, this slow shift will be accompanied by ever-new, far-reaching promises that must be evaluated with great skepticism. For this process to be ecologically effective and socially desirable, the prevailing technology-centric view of AI must be overcome and replaced with a social and organizational contextualization of the diffusion and application of AI. That is, the use of AI – like all technologies before it – requires not only an overarching technological design but also a cohesive social design.

References

- Bareis, J., & Katzenbach, C. (2021). Talking AI into being: The narratives and imaginaries of national AI strategies and their performative politics. *Science, Technology, & Human Values*, 47(5), 855–881.
- Beckert, J. (2016). *Imagined futures: Fictional expectations and capitalist dynamics*. Harvard University Press.
- Brynjolfsson, E., & McAfee, A. (2014). *The second machine age: Work, progress, and prosperity in a time of brilliant technologies*. Norton & Company.
- Buchanan, B. G. (2006). A (very) brief history of artificial intelligence. *AI Magazine*, 26(4), 53–60.

- Cave, S., Dihal, K., & Dillon, S. (2020). *AI narratives: A history of imaginative thinking about intelligent machines*. Oxford University Press.
- Chomsky, N., Roberts, I., & Watumull, J. (2023, March 10). The false promise of ChatGPT. *The New York Times*. <https://www.nytimes.com/2023/03/08/opinion/noam-chomsky-chatgpt-ai.html>
- EU (European Commission) (2020): On artificial intelligence – A European approach to excellence and trust. White Paper, *Communication from the Commission, COM (2020) 65 final*. Brussels.
- Ford, M. (2018). *Architects of intelligence: The truth about AI from the people building it*. Packt Publishing.
- Gehlen, A. (1940). *Der Mensch. Seine Natur und seine Stellung in der Welt*. Junker und Dünnhaupt.
- Hirsch-Kreinsen, H. (2023). Artificial Intelligence: a “promising technology.” *AI Soc.* <https://doi.org/10.1007/s00146-023-01629-w>
- Kaltheuner, F. (2021). *Fake AI*. Meatspace Press.
- Larson, E. J. (2021). *The myth of artificial intelligence: Why computers can't think the way we do*. Harvard University Press.
- Marquis, P., Papini, O., & Prade, H. (2020). Elements for a history of artificial intelligence. In M. Pierre, O. Papini, & H. Prade (Eds.), *A guided tour of artificial intelligence research. knowledge representation, reasoning and learning* (pp. 1–14). Springer.
- Nilsson, N. J. (2010). *The quest for artificial intelligence: A history of ideas and achievements*. Cambridge University Press.
- Ossewaarde, R., & Gülenç, E. (2020). National varieties of artificial intelligence discourses: Myth, utopianism, and solutionism in West European policy expectations. *Computer*, 53(11), 53–61.
- Rogers, E. M. (2003). *Diffusion of innovations* (5th ed.). Free Press.
- Struckman, I. & Kupiec, S. (2023). Why they're worried: Examining experts' motivations for signing the 'pause letter.' *MIT Spring 2023 course 6.S963 Beyond Models – Applying Data Science/AI Effectively*. Massachusetts Institute of Technology.

Acknowledgement

Thorben Krokowski's contribution was made within the framework of the Research Training Group GRK 2193, funded by the German Research Foundation (DFG) – project number: 276879186.