



# The mechanisms of AI hype and its planetary and social costs

Alva Markelius<sup>1</sup> · Connor Wright<sup>1</sup> · Joahna Kuiper<sup>1</sup> · Natalie Delille<sup>1</sup> · Yu-Ting Kuo<sup>1</sup>

Received: 20 November 2023 / Accepted: 5 March 2024 / Published online: 2 April 2024  
© The Author(s) 2024

## Abstract

Our global landscape of emerging technologies is increasingly affected by artificial intelligence (AI) hype, a phenomenon with significant large-scale consequences for the global AI narratives being created today. This paper aims to dissect the phenomenon of AI hype in light of its core mechanisms, drawing comparisons between the current wave and historical episodes of AI hype, concluding that the current hype is historically unmatched in terms of magnitude, scale and planetary and social costs. We identify and discuss socio-technical mechanisms fueling AI hype, including anthropomorphism, the proliferation of self-proclaimed AI “experts”, the geopolitical and private sector “fear of missing out” trends and the overuse and misappropriation of the term “AI” in emerging technologies. The second part of the paper seeks to highlight the often-overlooked costs of the current AI hype. We examine its planetary costs as the AI hype exerts tremendous pressure on finite resources and energy consumption. Additionally, we focus on the connection between AI hype and socio-economic injustices, including perpetuation of social inequalities by the huge associated redistribution of wealth and costs to human intelligence. In the conclusion, we offer insights into the implications for how to mitigate AI hype moving forward. We give recommendations of how developers, regulators, deployers and the public can navigate the relationship between AI hype, innovation, investment and scientific exploration, while addressing critical societal and environmental challenges.

**Keywords** AI hype · Anthropomorphism · AI narratives · Climate change · AI ethics · Socio-technical systems

## 1 Introduction

The notion of artificial intelligence (AI) as a transformative technology has emerged as a dominating narrative, influencing the collective understanding of societies worldwide. This global AI enthusiasm, which is typically referred to as AI hype, spans academia, geopolitics, major technology firms, startups, investors, and even early adopters. The growing prominence of AI and associated technologies in press and media coverage, further intensifies this perception [1]. There are a range of mechanisms, both social and technical, which are fueling and driving AI hype. This paper is set out to identify and discuss those mechanisms, in particular how they are made present from a socio-technical perspective. We define AI hype as a trending global fixation and prioritisation of AI-related technologies, ideas and investments. This is a stark contrast to the so-called AI winters,

an historical period of time marked by a lack of interest and investment in AI technologies (such as in the 1970s) [2]. It can be argued that AI hype is disproportionate against the potential of AI technologies, however, that debate is beyond the scope of this paper, as we focus on the mechanisms and consequences of said phenomenon.

The magnitude of the current AI hype surpasses previous periods in the history of AI, and this unprecedented wave of excitement and anticipation is undeniably on a significantly larger scale than previous eras of AI hype we have seen historically [3]. The IBM Global AI Adoption Index 2022 reports that 35% of businesses are using AI, a four percentage point increase over 2021 [4]. While this does differ by region (60% of Chinese and Indian companies are said to be using AI, compared to 24% in Australia), the global engagement by non-tech companies has also been significant thanks to the accessibility of large language models (LLMs) like ChatGPT (which now has over 100 million users globally [5]). For example, 27/100 charities polled in the Charity Digital Skills Report are now using AI in their day-to-day activities [6], while AI-related businesses in the UK has increased by 688% over the last 10 years [7].

✉ Alva Markelius  
ajkm4@cam.ac.uk

<sup>1</sup> Leverhulme Centre for the Future of Intelligence, University of Cambridge, Cambridge, UK

Yet, amidst this hype around the possibilities and benefits of AI development, there exists a disquieting apprehension of experts, leaders as well as the public, encouraging a slowing down of the development and urging calls for regulation and governance to keep up with the rapid development of AI technologies [8]. These calls, while ostensibly motivated by positive intentions, play a role in the broader global trend of AI hype. The current level of AI hype prompts significant concerns that extend beyond the immediate necessity for regulation and governance in the deployment and development of AI. In this paper, we seek to explore the phenomenon of AI hype through the lens of these concerns. Our aim is to draw meaningful comparisons between the current AI hype and similar historical instances, shedding light on the unique mechanisms that drive the hype in the present era. Furthermore, we aim to draw attention to the often neglected and overseen consequences of the current AI hype, focusing primarily on the planetary costs and increasing inequalities. In doing so, we recognise AI hype as a socio-technical narrative which is spun through a multitude of mechanisms globally and affect societies, and our planet, worldwide.

The first part of this article accounts for historical instances of AI hype, and how they differ from the current wave of AI hype, particularly in terms of magnitude and scale. Subsequently, our analysis will identify and discuss the most prominent sociotechnical mechanisms that drive AI hype. These include for instance the phenomenon of anthropomorphism [9], geopolitical and private sector “fear of missing out” (FOMO)-trends [10, 11], the overuse and misuse of the term “AI” in emerging technologies [12] and the influential narratives and notions by the different stakeholders [13], and exaggerated AI-literacy in the field [14]. The second part of the article accounts for the consequences of the current wave of AI hype, that are often overlooked. We scrutinise its implications in particular for the environment and the planet, and the intense pressure it creates on limited resources and energy consumption [15, 16]. Additionally, we focus on the tendency of this phenomenon to reinforce and reproduce socio-economic injustices and inequalities through job loss and polarisation [17], and how it affects human intelligence through knowledge decay and what we call post-truth. In the concluding section, we synthesise our findings to offer insights into the implications for developers, regulators and the public moving forward.

## 1.1 Historical perspectives of AI hype

Over recent decades, a multitude of emerging technologies have been championed by techno-optimists, entrepreneurs, corporations, media outlets, and investors, notably venture capitalists. These stakeholders have enthusiastically claimed and foresaw the forthcoming ubiquity of such technological advancements. Indeed, innovations such as personal

computers, GPS, smartphones, and the Internet have materialised and profoundly and broadly changed our society in many ways. On the other hand, certain technologies, such as nuclear fusion, quantum computing, the metaverse, and cryptocurrency, have either experienced limited adoption or remain in the nascent stages of development [18]. Supersonic air travel was once touted as the future of transcontinental travel. Despite its marvellous engineering achievement, it struggled with insurmountable challenges, ranging from economic viability and fuel efficiency to noise pollution, environmental ramifications, and operational complexity. The famed supersonic Concorde flights terminated operations in 2003 [19]. Similarly, magnetic propulsion trains, the hyperloop concept, and the segway were once heralded as transformative transport solutions. Yet, they either persist as developing concepts or fail to achieve widespread adoption.

The modern history of AI has been characterised by periods of intense optimism followed by disappointment and scepticism since the term was coined in the 1950s. These periods of optimism and subsequent retreats are commonly called “AI hype” and “AI winter.” The initial surge of enthusiasm for AI began in the 1950s and 1960s. Herbert Simon, who was awarded the Nobel Prize in Economic Sciences and the ACM Turing Award, once declared that “machines will be capable, within twenty years, of doing any work a man can do.” [20]. Similarly, Marvin Minsky, a leading pioneer in AI, predicted in 1967 that the challenge of creating AI would be substantially solved within a generation [21]. Although there were some noteworthy achievements in the early days, such as General Problem Solver, checkers, and the invention of the AI programming language LISP [2], by the late 1960s and early 1970s, researchers faced considerable hurdles in advancing the technology. It became apparent that the AI systems of the time struggled to address real-world problems effectively. As a result, securing funding for AI projects became increasingly difficult, leading to the first AI winter. A brief resurgence of AI enthusiasm emerged in the late 1970s and early 1980s, spurred by the proliferation of rule-based expert systems [22]. Many well-known expert systems were developed during that time, e.g., SRI International’s PROSPECTOR (mineral exploration), the University of Pittsburgh’s CADUCEUS (medical diagnosis), and DEC’s XCON (R1) (VAX computer system configuration). However, this momentum was short-lived as it became evident that it is challenging to develop and maintain expert systems for complex domains [2], and by the end of the decade, a second AI winter set in, persisting until the late 1990s to early 2000s.

Unlike previous AI hypes, which were marked by technologically over-ambitious promises followed by so-called AI winters due to the failure to fulfill these expectations, the last two decades have witnessed a significant acceleration in the development of certain AI technologies, notably,

machine learning algorithms, and recently generative AI, and such advancement has facilitated substantial integration of these technologies into various practical applications, including security screening, surveillance, manufacturing, drug discovery, social networking, office productivity, and e-commerce. This progress has occurred despite technical constraints and ethical dilemmas, such as those associated with facial recognition technology, as Buolamwini and Gebru [23] highlighted in their paper on intersectional accuracy disparities in commercial gender classification.

These advancements are primarily due to the significant increase in computational power, such as cloud computing and GPUs, the proliferation of data collection and availability from the digital economy, smartphones, and IoT sensors, and breakthroughs in algorithms, especially machine learning algorithms such as deep learning, reinforcement learning, and transformer foundation models. The tech industry had claimed that certain machine learning algorithms were already on par with or even outperformed humans in tasks such as image classification, captioning, speech recognition, video games, and chess games. For example, Microsoft's claims of achieving human parity in speech recognition and image captioning [24, 25], along with Google's AlphaGo defeating human grandmasters. Geoffrey Hinton, commonly known as one of the godfathers of modern AI, once even suggested that "People should stop training radiologists now. It's just completely obvious within five years, deep learning is going to do better than radiologists... It might be 10 years, but we've got plenty of radiologists already." [26]. More recently, with the rapid advancement of large language models (LLMs) since the release of GPT-3 in 2020, some techno-optimists and technology companies have declared that current LLMs, such as GPT-4, have shown the "sparks" of intelligence and would be considered as the first true examples of artificial general intelligence (AGI) [27, 28].

The level of investment and techno-optimism in generative AI technologies has been particularly noteworthy since the release of ChatGPT by OpenAI in 2022. CB Insights [29] estimated that funding for generative AI grew from \$2.5 billion in 2022 to \$14.2 billion in the first half of 2023 and that the global generative AI market is expected to reach \$42.6 billion in 2023. As of July 2023, there were over 330 generative AI startups [30]. Generative AI technologies, such as LLMs, have been rushed into integration into various product lines by large technology companies such as Microsoft, Google, and Meta. For example, Microsoft has incorporated ChatGPT and GPT into a wide range of products and services, including its Bing search engine, the Windows operating system, the Office suite, the Edge web browser, security products, and developer tools. Similarly, Google has integrated its LLMs into services such as Google Search, Google Gemini, Vertex AI, Google Workspace, and many more. Fearing they will miss out, many businesses are

scrambling to form their generative AI strategy. The speed of end-user adoption has also been rapid. For example, ChatGPT was considered the fastest-growing "app" of all time [31], with 1 million users within a week of its release and crossing 100 million active users within a month.

These observations suggest that the current AI hype differs from the previous ones in terms of the level of enthusiasm and investment, speed of deployment and adoption, and the breadth of the applicable domains. However, despite this momentum, there has also been a growing skepticism and criticism of generative AI technology limitations, its pitfalls, the scope and scale of applications, and the overhype it (and AI in general) may have received [16, 32, 33]. We argue that the mechanisms of AI hype stretches far beyond the actual capabilities and presumed transformative power of AI technologies and that it is rather a result of complex socio-technical forces. Furthermore, we argue that the detrimental consequences of AI hype indeed should make us reconsider the validity of the narratives spurring the hype around AI. The following sections will account for these two parts.

## 2 Mechanisms of AI hype

Situating emerging AI technologies in a socio-technical context is necessary to understand the underlying mechanisms contributing to its hype. Although this analysis is not complete, we here attempt to point out and unpack some of the socio-technical mechanisms which we identify as significantly contributing to the current discourse, namely anthropomorphism, exaggerated AI literacy, directing narratives and FOMO and overuse of the term AI. Ultimately, AI hype can be viewed as a global, sociotechnical imaginary, a narrative, created collectively by these mechanisms, whether deliberately or not. We will address the often overlooked costs and consequences of this narrative in Sect. 3.

### 2.1 Anthropomorphism

In this section we explore the phenomenon of anthropomorphism as one of the mechanisms which is driving and enabling the current AI hype as outlined in section one. Anthropomorphism has been widely studied in relation to AI and many attempts have been made to conceptualise, measure and theorise the phenomenon [34]. One of the most commonly used definitions is offered by Epley et al. [35] and states that anthropomorphism involves attributing human characteristics (e.g. intentions, motivations, and emotions) to the behaviour of nonhuman entities, such as animals, natural forces, deities, and machines, whether they are real or imagined (pp. 864–865). Many AI technologies are deliberately designed intentionally to be anthropomorphised, as a means to facilitate social interaction, improve user experience or

for marketing and monetary purposes [36, 37]. Anthropomorphism as a phenomenon makes itself present as a shared narrative among users, technologies, designers, innovators and regulators alike, with significant epistemological and ethical consequences [38]. Humans tend to be quick to anthropomorphise, whether it is a chatbot interface, a digital avatar or solely a statistical “AI solution” tool, even after just brief exposures [39]. This includes the mind perception of agency and emotion [39, 40], attribution of gender [41–43] as well as judgement and competence [44]. However, despite these common user interaction reactions, what is particularly worth noting is that anthropomorphization of AI technologies often is a deliberate design choice made by innovators, developers and deployers alike [45]. Anthropomorphism has previously been discussed in relation to AI hype [9] through two dimensions of ethical considerations, exaggeration and misrepresentation of AI capabilities and distortion of moral judgments about AI [9]. We here build on this work and identify another dimension; neglect of AI infrastructures. Personification of AI, including ascribing characteristics like gender, emotion and physical attributes do all have significant ethical and social implications. However, in terms of AI hype, it is the attribution of capabilities, such as intentions, competence and motivations of the AI systems that has implications in particular [9]. Researchers have previously cautioned against using rich psychological terms, such as understanding, motivation and creativity in contexts of AI, because of the possibility of over-attributing capabilities which in turn might have ethical ramifications for regulation, scientific understanding and expectations among broader social and societal contexts [46]. This over-attribution, as reproduced and reinforced by the AI hype, may cause overestimations or misunderstandings of AI systems’ actual capabilities. The consequences of this are multiple, for instance since it might give rise to either disproportionate fear of AI technologies or on the contrary of uncritical optimism, but ultimately it may also blur moral and ontological boundaries between humans and technologies [38]. The question of anthropomorphism, and the resultant conception of AI as increasingly agentic entities is also tied to questions of accountability and responsibility [9, 47]. As AI systems exhibit a human-like semblance of autonomy in decision-making and execution of actions, they gradually become perceived as moral agents [9, 48]. This has sparked a debate of whether it is right or not to create artificial moral agents (AMAs) [49] and how to endow those with ethical mechanisms to be capable of making sophisticated moral decisions as humans do [50]. Finally, worth pointing out is the risk that actors profiting from the development of AI might deliberately use anthropomorphism to obscure and complicate issues around accountability, leaving them unaccountable for risks and harms associated with technologies they develop and deploy [9].

Another consequence of anthropomorphism as a mechanism of AI hype is a harmful oversight of the complex and multifaceted AI infrastructure that underpins the development, operation, and maintenance of AI technologies [51]. This infrastructure constitutes a vast network of interconnected components, encompassing human labour [52–54], substantial planetary resources [15, 51], and a complex framework of societal institutions. The current discourse surrounding AI frequently obscures this underlying structure, resulting in what can be aptly termed the “mystification” of AI technologies [55]. This mystification, highly related to the anthropomorphisation of AI, perpetuates the faulty perception of AI technologies as autonomous entities, seemingly existing in isolation from the infrastructure of human, environmental, and societal dependencies that are vital for their functioning [47, 55]. Such misconceptions not only obfuscate the ethical implications of AI, but also undermine the imperative to scrutinise the ramifications of AI within the broader context of their interdependence with humanity, ecosystems, and established social norms (*ibid.*). As AI hype is reinforcing the notion of AI as autonomous, anthropomorphised entities, the more the AI infrastructure risks being further neglected and detached from public awareness, regulation and oversight. AI technologies need to be seen in the light of the human labour, planetary resources, and the dynamics of political and social discourses it is deeply dependent upon.

## 2.2 AI “Experts” and exaggerated literacy

While the concept of AI can stretch back to writings about automotons in 12th century Türkiye [56], the relatively recent release of large language models like ChatGPT has led to an expansion in the accessibility of the field of AI [57]. Consequently, the perceived recency of the industry provided by the AI hype has led to a proliferation of self-proclaimed AI experts. To illustrate, such is the importance of AI-related knowledge when looking for jobs; a LinkedIn report found that 40% of Gen Z staff (aged 18–26) exaggerated their knowledge of AI to seem more informed [14]. It is this drive to stand out from the rest that motivates workers to exaggerate their AI literacy level. The above reality allows us to explore the techno-deterministic nature of AI hype that has been spurred on by and feeds into the techno-optimism attitude surrounding the technology [58]. While technodeterminism has several formulations varying in severity, we primarily use the concept to evoke the perspective-shaping and persuasive power of the AI hype. In this case, technodeterminism involves two pillars [59]: one being the unstoppable nature of progress being marked by technology (whereby resisting using technology is a losing game) and the other being that, given the inevitable evolution of technology, society then adjusts itself around this narrative.

In relation to the first pillar, AI technology becomes treated like a hammer, making every business and social problem look like a nail in need of a technological solution. This can help explain how Google insiders are still pondering the usefulness of the Bard AI chatbot in their business practice [60]. In reference to the second pillar, AI knowledge turns into a highly valued and sought-after skill in a wide variety of disciplines, from policymakers to marketing professionals. Both of these pillars show the high lock-in potential of AI technologies, whereby the further integrated and used they become in corporate and civilian lives, the more any corporate or social progression seems to require AI technologies. Consequently, presenting oneself as competent with AI knowledge becomes a strategy for job security as well, with 41% of employees saying how they fear being left behind if they do not know how to leverage AI in a study conducted by Canva [61]. The resulting panorama thus includes a determined attitude towards figuring out how best to use AI and those wanting to work in the space over-selling their abilities to join in [14]. Comparatively, seeking advice from recent experts in a field can be analogous to the self-proclaimed ‘COVID-19 experts’ during the pandemic. For instance, Stanford Professor John Ioannidis initially mocked the US government for its fear over a mounting death toll with Covid, and claiming that the number of deaths would not pass 10,000 [62]. Yet, the total number of deaths (as of April 6 2023) stands at 1,132,662, a farcry away from Ioannidis’ claim. Hence, when it comes to the AI hype, those without complete credibility are able to present themselves as AI experts given the demand for AI skills [14] via the technologically deterministic narrative that is presented. The latter topics of narratives and the generated fear of missing out (FOMO) we will now explore further.

### 2.3 Directing narratives and FOMO

The AI hype is furthermore made present not only in industry, academia and among media and users, but also takes shape in a very political manner [13]. The fear of missing out on the increasing development does not solely belong to startups and businesses deploying AI-based technologies to keep up, but can be seen in a larger geopolitical scale [63]. Indeed, recent rethorics have largely elucidated the so-called US–China AI Arms Race [64], which refers to the polarised notion of urgent competitive development and deployment of AI technologies between the two nations. This notion is strongly institutionalised and driven by a combined narrative from governance and legislation, military bodies, as well as actors in the tech-industry [11]. The sense of urgency, and the notion of AI as not only commercial products deployed to effectivise our daily lives, but rather as necessary “strategic national assets” in a geopolitical sense indeed largely contributes to the current AI hype [65]. Large

tech companies, and the AI technologies they develop, are increasingly being seen as indispensable for the safety and sovereignty of nations, and subsequently, initiatives for regulation and legislation of AI technologies are being advocated against [11]. Indeed, this polarisation and antagonising narratives between nations that are being woven regarding the development and deployment of AI technologies highly impact the general narrative and hype by creating a sense of fear and urgency among a range of social contexts. To illustrate, Bareis and Katzenbach [13] conducted a mapping of different nations’ so-called sociotechnical imaginaries, that is narratives about AI found in the official AI strategy documents of the nations. While focusing predominantly on US, China, Germany and France they found striking similarities and consistencies of how these nations create narratives around AI in their official strategy documents, in particular regarding the notion of AI as inevitable, the aforementioned technodeterminist rhetorics as well as policy suggestions which are both “bold and vague” (p. 857). Despite these similarities, they also find noteworthy differences between the four nations in terms of how they formulate and focus their sociotechnical imaginaries. Shortly summarised, Germany focuses on AI applications in the manufacturing industry with notions of efficiency, innovation and climate. The French strategy is rooted in a human-centred ethos, prioritising the use of AI in sectors that is said to enhance the quality of human life. In the US, AI is portrayed as an expression of national patriotism, striving to equate the nation’s technological progress with overall societal advancement. In contrast, the Chinese Communist Party portrays AI as a tool for maintaining social order and enforcing regulation [13]. In other words, on a geopolitical scale, the AI hype is manifested into technodeterminist rethorics, where nations build imaginaries of AI as both inevitable and necessary for the nations flourishing and survival. Although some differences can be seen in the ideological underpinnings of these imaginaries and rethorics, the resulting AI hype is present in them all.

The geopolitical forces are not the only significant forces in directing and creating global narratives and imaginaries around emerging AI technologies. Two other significantly powerful groups of contributors are multinational corporations [66] and academic institutions [67]. The last decade the total number of AI publications more than doubled, and of those 75.23% of the AI documents were in the education sector, making academia a significant voice in directing the global narratives around AI.

65% of journal publications originated from China, the UK, the US, or Europe [67], resulting in a notably uneven geographic and cultural influence on the development of epistemologies and the direction of knowledge and narratives. Furthermore, despite an historically even distribution of AI research between academia and industry, industry has

now taken the lead in access to computing power, data and talent and is currently becoming increasingly more influential [68]. A noteworthy aspect here is the manner in which narratives and perceptions within academia and industry are guided concerning the debate on the long-term versus short-term risks associated with current AI technologies [69]. This debate has been increasingly present and subject to change the developmental discourse of AI in different directions, and researchers are cautioning that this debate might create unnecessary division between two camps who are in fact closely related [70]. Although the distinction between long and short term risk is not a new debate, it takes shape in a different way in light of the current AI hype. For instance, the Future of Life Institute's call for a pause for at least 6 months of training of AI systems more powerful than GPT-4 [8] gained large attention worldwide and was significantly impacting the global narratives around emerging AI technologies, and not least contributing to the current AI hype. Similarly, industry leaders such as Sam Altman, CEO of OpenAI, have warned that advanced AI can cause serious risk and that potential harm needs to be regulated accordingly [71]. It is notable that individuals responsible for developing and deploying technologies with significant societal impact are advocating for enhanced regulation (*ibid.*). Within the context of the current AI hype, it is common for influential entities to shape definitions of risk and harm, presenting themselves and their solutions as primary mitigators of these challenges.

This is particularly relevant to understanding the mechanisms behind the AI hype, where directing the narratives is one of the most powerful of those mechanisms [72, 73]. Whether it is industry, academia or on the geopolitical level, who and how the narratives around AI are created is a significant mechanism of hype, especially considering the current asymmetries in who is able to have a voice heard in directing the hype [74]. In fact, one could even go so far as to stating that the hype itself is a creation of narratives and imaginaries about emerging AI technologies [72], and thus, it becomes increasingly relevant who is having the privilege to have their voice heard and the power to direct and shape these narratives and who does not. It is also relevant if the fear of missing out is one of the main driving forces in directing the narratives, and consequently if that fear is justified or not [10]. In other words, there is a need to establish whether there is a cost of “missing out” and if so, if it is in any way corresponding to the fear of missing out narratives that are being spun.

The consequences of this are multiple. For instance, the rhetoric within nations of AI as necessary and inevitable tools for the flourishing and even survival of the nations, in particular for military or national security reasons, threatens to be a way to justify infringements on individual data privacy and human rights [11, 13, 75]. When individual data

is used to further surveillance and control parts of societies, the risk for this data to be used in harmful ways is also increasing, e.g. to control or surveil already marginalised or historically controlled communities and hence reinforce and reproduce violent structures of power [76–78]. Furthermore, when the private sector and industries increasingly further ideas about their products as necessary parts of nations' survival and flourishing, this also increases their perceived credibility as governing bodies, which is directly related to democracy. An increasing government dependence on AI companies for e.g. policing and security comes with a shift in power for AI companies to hold central positions in democracies [79]. As these sectors increase their influence over narratives, we believe there is a corresponding shift away from democratic principles in society. Additionally, this dynamic results in marginalized communities having reduced opportunities to contribute to global discussions and narratives about AI, including indigenous communities and nations, which are often overlooked as significant contributors to AI development [80].

## 2.4 Overuse of the term AI

Over the years, an overwhelming influx of products and software solutions have emerged asserting the incorporation of AI within their offerings and its ability to transform the world [81]. This surge in AI-related claims has given rise to a concern that parallels the historical notion of “snake oil,” which refers to extravagant and often unsubstantiated marketing of products [82]. The current landscape is marked by numerous companies vying to make bold pronouncements about their utilisation of AI, accompanied by substantial investment in AI product development and discussions surrounding the potential societal impact of AI [83]. While Sam Altman, former CEO of the parent company of ChatGPT, has expressed to the media a worstcase scenario for AI that could entail “lights out for all of us” [84], there are many contrasting views on the AI's current state and its potential. Meta's Chief AI scientist, Yann LeCun, characterises ChatGPT as “nothing revolutionary” and in a similar vein, University of Washington professor Emily Bender has issued a cautionary note, emphasising that the concept of an all-knowing computer program belongs firmly in the realm of science fiction and should remain there [84]. We must consider that the field of AI is incredibly diverse, encompassing various subfields and applications, from natural language processing and computer vision to machine learning and expert systems. Notably, some areas of AI have progressed further and are more well-developed than others [85], which underscores the importance of distinguishing between various AI domains and their practical limitations.

Numerous headlines and substantial investments have emphasised the notion of AI as an existential threat to

humanity [86]. Nevertheless, Narayanan [12] argues that this perspective is significantly flawed, constituting a “tower of fallacies”. One prominent misconception involves the expectation of the arrival of AGI, primarily rooted in predictions based on the scaling trends of AI models. In reality, he points out that there are technical limits that are slowing down this progress, making it less likely that AGI will arrive as fast as some suggest. Another misleading belief concerns the idea that AI might gain independence and turn rogue, where Narayanan [12] highlights the lack of real-world evidence supporting these claims, which frequently hinge on theoretical scenarios. He argues that the real risks linked to highly capable AI systems are more likely to stem from human misuse or manipulation, rather than AI autonomously straying from its intended programming and developing its own agency. Furthermore, we believe that there’s a troubling issue where current security vulnerabilities seem to be ignored in favour of speculating about a possible rogue AI in the distant future. This disregard for immediate risks is especially alarming, notably in the corporate sector, where security concerns and risks of mass surveillance, misinformation and manipulation, and the inadequacy of our current economic paradigm in a world where AI plays an increasingly prominent role [87] appear to take a backseat to speculative scenarios. While discussions regarding rogue superintelligent AI could hold value in some respects, the exaggerations and erroneous arguments within these discussions can prove detrimental.

Within organisations, “predictive AI-driven systems” are being employed by teams to make assessments of an individual’s personality and suitability for a job role based on as little as a 30-s video [88]. Notably, these systems claim to perform this evaluation by focusing on non-verbal cues such as body language and speech patterns, effectively bypassing the content of the candidate’s speech. Evidence that algorithms may help reduce human biases advocates for the adoption of algorithmic techniques in hiring with a variety of computational metrics proposed to identify and prevent unfair behaviour [89]. But to date, little is known about how these methods are used in practice and how successful they actually are [90]. In fact recent research have pushed back on the notion that evaluations of intricate human attributes and job suitability can be accurately rendered from such a brief and content-agnostic video and that it instead serves to uphold powers of structural sexist and racist ideals [91]. In essence, these AI systems exhibit characteristics more aligned with a complex random number generator than a dependable tool for genuine assessment and decision-making. As [92] suggests, technology owners must be held accountable, prompting essential questions regarding the data on which language models are trained and the model’s capacity to provide explanations or cite references for the answers they generate. But the questions should be extended

to whether or not these technologies are desirable in the first place, and fundamentally seek to interrogate notions of consent, marginalisation and monetisation of algorithmic and datafied high-stakes decisions.

### 3 Consequences of AI hype

Having accounted for the main mechanisms of AI hype, we conclude that they are mainly driven by socio-technical narratives, and a multitude of forces, often overlooked in the current discourse. The following section will cover the planetary and social costs of the hype. Consequences of AI hype have previously been discussed in terms of risks for public safety, legal practices and worker displacement [93]. Here, we identify some of the often overlooked consequences of AI hype, such as use of planetary resources, disruption on socio-economic structures and threats to human intelligence. While we recognise that these consequences does not constitute a full account for all potential costs of AI hype, we hope to contribute with a perspective to spark further debate and research into the consequences of AI hype.

#### 3.1 Planetary costs

Having established what the current AI hype consists of and mechanisms driving said phenomenon, it’s necessary to consider the very tangible consequences it presents. Research has showed how training the language model BERT (based on transformer architecture) is almost equivalent to a trans-American flight in terms of carbon emissions (p. 4) [94]. Furthermore, LLMs require hardware (such as chips) to run. Such is the demand for chips that are required to develop and deploy LLMs that the world’s largest chip manufacturer [95], TSMC, expects revenues from AI chip manufacturing to grow by 20% this coming year [96]. An example of additional hardware that sustains the subjects of AI hype is data centres, which require substantial amounts of water to sustain. We have broken down the planetary costs of AI hype into two main sections; data centres and water and electricity. We further discuss the opportunity costs of the AI hype planetary costs.

##### 3.1.1 Data centres and water

Data centres can be hyperscale (warehouses containing servers) or much smaller (a backroom cupboard in an office). For our purposes, the planetary costs of AI hype are clearest when focusing on hyperscale data centres. To illustrate, the National Security Agency in the US at Fort Meade, Maryland, uses 5 million gallons of water a day to maintain ambient temperatures for the data servers [97] (p. 3). Similarly, the Utah Data Centre in Bluffdale uses 1.7 million gallons of

water/day, while Microsoft's San Antonio data centre uses 8 million gallons/day [98].

Now, it can be levelled that such water consumption is miniscule in comparison to other areas where water is used. For example, in 2014, 626 billion litres of water was used to supply data centres for the whole year [99]. Yet, in the US during 2015, sectors such as thermoelectric power used 503 billion litres per day, whereas irrigation used 446 billion litres per day [100]. Hence, while data centres do impact the environment through their crucial part in sustaining LLMs in the AI hype, it could be argued to not be a significant impact when compared to other sectors' use of water. Furthermore, it could also be levelled that AI can aid in the battle against climate change [101]. AI's ability to process multidimensional and unstructured data allows for our understanding of past climate trends (such as global mean temperatures [102]) to deepen and for more accurate forecastings to be created (ibid.). In addition, companies such as Microsoft and Google have sought to combat the water usage levels in their data centres, with Google's Hamina data centre in Finland using seawater for cooling since 2011 [103] and Microsoft undergoing Project Natick to explore the possibility of using the Pacific Ocean to cool some of its servers [104].

However, Monserrate's work [15] shows how devastating an effect on water supplies these data centres can have. He notes how residents in Bluffdale, Utah, are suffering from water shortages due to the nearby Utah Data Center used by the US National Security Agency (ibid.). Microsoft's data centre in Northlake, Illinois, required an intergovernmental agreement between Northlake and Franklin Park to organise the necessary extra water capacity for Northlake to sustain the data centre's water demands [105]. Furthermore, the Taiwanese drought in 2021 meant that TSMC had to respond by ordering truckloads of water [106] to continue its operations. Above all, while applaudable, initiatives like Google's and Microsoft's above are not enforceable (meaning the companies are not bound to commit resources for long periods of time to such green practices). Hence, regarding the AI hype, this stress on water resources will only become more pronounced [15]. It is a tangible, and worrying planetary cost, that is already causing significant consequences for numerous humans living in the affected areas.

### 3.1.2 Electricity

Alongside water, it is important to consider other forms of resource scarcity and the expansion of the AI industry when considering the planetary impacts of the AI hype. As a result of the sustained hype, ChatGPT's electricity in January 2023 alone was the equivalent of 175,000 Danish families' usage in a whole year [107]. Furthermore, almost 20% of the Republic of Ireland's electricity usage is attributable to data centres, more than all of the country's urban

households [108]. Moreover, this energy consumption is 31% higher than 2021 and almost 400% higher than in 2015 (ibid.). Recent research has in depth investigated the growing energy footprint of emerging AI, and conclude it is immensely energy-intensive. Both the training and the implementation phase of many AI technologies are particularly energy-costly, and increasing as models gets bigger [109]. With increasing demands for efficiency, so is demand for technology adoption, so increased energy efficiency is not really save net energy in the end. The author is suggesting AI regulation should demand from developers to disclose of energy usage come from renewable sources or not [110]. Generative AI is particularly thirsty for energy, every prompt sent to LLM-based chatbots such as ChatGPT does require energy, up to 100 times more energy than an email. This is much faster than renewable energy sources can supply [111]. Considering this, we continue to hold that the planetary effects of the AI hype will only continue to increase as companies (and the wider population in general) continue to utilise the technology. We echo other researchers in the field who call for investigation and disclosure of energy sources used for training and deploying AI technologies, and to what extent they rely of renewable versus fossil based sources. Furthermore, we hope that more research can determine to what extent the increased global energy demand has on net carbon emissions and climate impact and be considered as part of the conversation on where the actual existential risk from AI lies.

### 3.1.3 Opportunity cost

With this in mind, we posit that there is an opportunity cost in buying into the AI hype. While companies pour billions into integrating generative AI into their business (such as accounting firms Ernst and Young and KPMG committing to \$1.4 billion and \$2 billion investments respectively [112]), there are various business and social aspects that are neglected as a result. For example, it is claimed that Taiwanese rice farmers are being subsidised to not grow rice so that more water can be diverted to TSMC facilities, meaning less rice is grown, increasing the demand on food imports for the population [113]. Building on this, while generative AI can use different forms of learning to machine learning, a poll of data science professionals by KDnuggets [32] revealed how only 0–20% of all machine learning projects get deployed. This is further reinforced by how, according to a McKinsey report [114], data science professionals reported that only 15% of machine learning projects had been realised. As a result, while billions are invested into generative AI, there is a relatively small guarantee that this investment will lead to a tangible outcome.

Instead, these funds could rather be diverted to other projects or initiatives. For example, we believe this could be

investigating how to optimise current models to be more efficient, or even investing in workforce training to better equip data scientists and engineers in understanding current company models. This could involve introducing engineers deploying the models to stakeholder analysis, which [115] observe is an unlisted part of some data scientist positions. Furthermore, we hold that a more measured approach to AI investment allows for more careful consideration about why AI technologies are being used, and how to best use them. Such considerations can then involve prioritising tackling current important issues, such as data centres' reliance on water for data cooling [103]. Confronting this problem will then focus efforts on dealing with water distribution prioritisation to data centres during droughts (like in Taiwan [106]). Failing to explore such alternatives to AI investment may, thus, lead to the aforementioned substantial planetary costs, as well as the social costs we shall now explore.

### 3.2 Social costs

A significant early contributor to the hype regarding AI (inclusive of the entire array of technologies algorithms, networks, models, data, and interfaces that are referred to commonly as Artificial Intelligence) was economist Klaus Schwab's proclamation that the 4th Industrial Revolution was upon us in 2016 [116]. His reasoning included AI as one of the primary innovations bringing about this revolution, and declared that the proliferation and integration of AI would bring about "nothing less than a transformation of humankind" (ibid.). Since this time, as we have seen earlier on in this paper, the investment in AI technologies has grown at an incredible pace. In 2022 the amount of private investment in AI was 18 times greater than it was in 2013 [117] and mentioning the new flavor of AI, presently generative AI, in a pitch or prospectus continues to be a key element to getting high valuation or the attention of venture capitalists [118]. Given the amount of investment, and, therefore, the requirement for that investment to generate exponential return, expectations for AI to deliver commercial value are high [119]. The opportunities to meet those expectations also increase as people embrace digital experiences across many areas of their lives. As Schwab continued in his description of the 4th industrial revolution critical elements, he noted that digital capabilities in "the societal front" also were contributing to how "a paradigm shift is underway in how we work and communicate, as well as how we express, inform and entertain ourselves." [116]. We believe that whether the hype of AI is warranted or not is almost immaterial to discussing the potential and real costs of industries acting on the premise that the hype is founded. The velocity at which industry and consumers alike continue expansion of AI into all digital realms and the already significant infiltration of these digital realms into so much of our personal and social

lives place so much of society in AI's path [120, 121]. We hold that this breadth and momentum make it a certainty that the societal impacts associated with AI, for better or for worse, will be just as significant.

One of the most commonly discussed social issues arising from AI is that of biased or inequitable outcomes that result from AI decision-making [122]. Indeed, a quick search of "bias in AI systems" on Google Scholar returned no less than three million results at the time of this paper's writing. The breadth of remediation needed to address bias when it comes to AI spans not only the data which powers these systems, but the human biases which create that data, and the systemic biases in which the AI participates [123]. It can be extrapolated from this article (ibid.) that the hype around AI will only accelerate the impact these layered sources of bias will have, both from a scale and speed perspective. This is especially true if the sense of urgency to reap the benefits of AI exceeds the pace at which deliberation and consequence understanding regarding AI and inequality can be exercised. Indeed, the social costs of AI are manifold, and stretches from emotional dependence of social AI [124] to algorithmic neocolonialism [76] and surveillance [78]. However, we focus on three themes we find directly caused by the AI hype, namely job loss and job polarization [116], knowledge decay [125] and finally knowledge corruption and post-truth [126]. These issues are deeply intertwined with other social costs, and we hope to contribute to the conversation by raising issues that might often be overlooked related to AI hype.

#### 3.2.1 Job loss and job polarization

When Krause announced the 4th Industrial Revolution, he expressed concern over the fundamental changes across all aspects of life, "We are at the beginning of a revolution that is fundamentally changing the way we live, work, and relate to one another. In its scale, scope and complexity (it) is unlike anything humankind has experienced before" [116]. Statements like these bolster the belief in the power of these technologies, leading executives and technology decision-makers to act on the sense of urgency echoed by headlines and AI vendors [127]. With these leaders seeking to capitalize on the promise of employee efficiency and time savings, a wide variety of jobs are seeing not only a renewed process and decision automation interest but an intense focus on employee augmentation in the administrative and repetitive elements of their work [128]. The current hype about generative AI's capabilities touts productivity by reassigning routine mental tasks of workers to machines [129]. McKinsey predicts that, if AI delivers on this promise, there is the potential to automate work activities that absorb 60–70 percent of employees' time today (ibid.). Such statements create a sense of a new technology "silver bullet" in the competitive scramble to continuously cut costs in business

leaders' planning priorities and resulting technology investments. Businesses acting on this promise of productivity via generative AI technologies would result in a dramatic reduction in person-hours, and, if existing employees are not deliberately redeployed to other productive work, a drastic loss of jobs [17, 130]. This category of job loss will include those that have the most administrative or repetitive elements, like clerical or customer service work on the lower end of the wage scale and therefore widens the economic gap brought about by the shrinking of the middle class (*ibid.*). This being the case, one of the likely outcomes this particular focus of AI hype will have is a broad socio-economic impact [131]. Even more costly to workers would be the combination of AI capabilities being overstated (hyped) and companies moving ahead with deployment because they are afraid to not keep up with the productivity gains they believe their competitors are realizing with these same hyped AI capabilities (especially seen in the national AI strategies we have explored [11]. We believe that such a condition would still result in job loss as mentioned while also adding significant burden on both individual employees that are left with suboptimal augmentation and on company productivity overall as processes and teams struggle to fill the gaps and pay for the ever increasing costs of their AI implementations. Even if little job loss results directly from staff reductions made by companies, AI augmentation may contribute to declines across multiple dimensions of work that contribute to worker well-being including, “worker freedom, sense of meaning, cognitive load, external monitoring, and insecurity” [132]. These declines in job satisfaction have the potential to counteract some of the productivity gains the tools were intended to achieve or lead to both productivity decline due to increased sick time or decreased employee performance (such as the burnout problem in responsible AI [133]. Of course, some efficiency and productivity benefits will be realized, but many consulting firms recommend taking a slow and measured approach to Generative AI implementation as the technology works out some of its ethical challenges [120, 128]. If this advice is ignored, we believe that the negative impacts will outweigh the potential gains, and society, individuals, and businesses will all suffer the consequences.

### 3.2.2 Knowledge decay

As AI capabilities are touted as being the equivalent of or superior to human knowledge [134], there is real danger in accelerated intelligence and skill loss due to AI implementations made with inflated confidence in the machine's intelligence [135]. The definition of skill loss can be summarized as “loss or decay of trained or acquired skills (or knowledge) after periods of nonuse” [125]. This is not a distant, imagined potentiality; examples of skill decay exist

across professions and industries today due to a variety of automated or technology-delegated tasks and digital augmentation of skilled jobs [135]. The most well-researched area of this condition may be aviation. As more and more technologies were introduced to the cockpits of airplanes, for both flight and navigation tasks, the human skills to do those tasks degraded [136] leaving pilots in increased danger if and when those technologies failed or became unavailable. When jobs are automated, workers lose skills they possess specific to those jobs because they no longer perform, or perform only in exceptional events, the tasks that require those skills (*ibid.*). Similar impacts are seen when exercising intellectual skills or retaining knowledge [137]. Another example, exists in the military [138]. With the proliferation of map and GPS technologies the knowledge of how to navigate a landscape (in those populations where those technologies are commonly available) has been largely forgotten, “of the 914 soldiers who have been through the (navigation) training, half have failed that portion... the high failure rate is a troubling sign as the service gears up for conventional warfare” (*ibid.*). A key factor of the present hype about generative AI cycle is that this technology claims broader application than prior technologies by producing human-like output [139]. We believe that this factor, which is not specific to any one industry or to a certain type of job, extends its impact significantly. With this broad of a task, automation impact estimates have been as high as 25% of current “work” across the global job market [140] also means that the breadth and depth of skill loss has the potential to be at a scale never before experienced. This knowledge decay will intensify the more we assign intellectual tasks to the AI technologies that are implemented in work environments, or are available to substitute for thinking in day to day life (as seen in the military [138]). Drawing from the principles of skill decay [125], the more we delegate “thinking” work—from calculation or formula application in chemistry to product and policy knowledge in customer service or sales—to machines, the more our brains will eventually retire our relevant knowledge. Whether general knowledge (e.g. the ability to calculate) or specific industry knowledge (e.g. a company's product details), we hold that lacking the ability to retrieve that knowledge will leave workers unable to validate generative outputs and reduce their capacity to compete with machines in the knowledge and skill spaces that humans are removed from or minimized in.

### 3.2.3 Knowledge corruption and post-truth

Generative AI both produces content and is trained (in part) by consuming large amounts of content. The training of the largest current LLMs has been done primarily with content that exists on the internet [141]. In fact, it is believed that in the case of the most popular models, the hundreds of

terabytes of data that were the basis of the training included ALL of the content on the internet as of approximately 2018, courtesy of the “Common Crawl... a non-curated corpus consisting of multilingual snapshots of the web” [126]. One problem with this breadth of training data is the wide range of quality to be found in such a public and unqualified mass of information. Indisputably, the entirety of the internet contains content that is often inaccurate or blatantly wrong, with a wide span of opinion and conjecture in the mix. This problem is exacerbated by the fact that LLM processes do not distinguish between high quality and low quality sources [141]. This makes the AI tools unreliable in the quality and validity of their outputs, and worse, very difficult for users of these tools to know whether or not an output is accurate [142]. Indeed, today’s generative content tools are prone to hallucinations wherein generated text “is nonsensical, or unfaithful to the provided source input” [143] but also presented as confidently as accurate text generation, “Similar to psychological hallucination, which is hard to tell apart from other ‘real’ perceptions, hallucinated text is also hard to capture” (ibid.). AI generating outputs containing false information that is indistinguishable from accurate outputs is a significant problem in a number of ways. If users of this technology develop a level of trust that encourages them to review the generated output less carefully, wrong information can be perpetuated. To illustrate, lawyers using Generative AI were so convinced by OpenAI’s hallucination of case law, they submitted a legal brief full of fictitious cases [144]. Another situation resulted in a customer of Air Canada being told inaccurate policy information during a chat with a bot on the airline’s website, ultimately resulting in a lawsuit where the customer won his claim that the bot’s representation of the policy was indistinguishable as false [145]. Furthermore, there is a circular pattern here: generative AI tools generates content, that content can be retained in the tool for later training, that content also may be shared on the internet via interest sites, blogs, social media, and perhaps more than on one channel depending on the purpose. This leads to an ongoing cycle of questionable information being generated, used, disseminated, and ultimately being scooped up in the next round of model training [146]. If one considers the unavoidable addition of the actual deep fake and deliberate misinformation participating in the same cycle as generated hallucinogenic content, the corruption of knowledge is even more severe [147]. Our ability to know what is actual and what is invented will be increasingly difficult to maintain.

## 4 Moving forward

Having explored the mechanisms of AI hype, as well as its consequences, we now offer some practical points for managing the phenomenon going forward. These will serve as

useful guidelines towards effectively navigating and managing the deluge and fervour that the AI hype brings.

- Why?—asking why AI technologies are being implemented in your organisation, and not just how, will help to assess the core reasons behind why AI is being considered for implementation. Do we need these technologies at all or are we just following along with the current hype?
- Design—the above can also apply to the design process. In our context, asking why an AI application is being designed in an anthropomorphic way helps to fully assess whether it is a necessary strategy or not. Robustly assessing (cognitive) capabilities of AI technologies can also help inform and reduce risk of misinterpretation.
- AI expert awareness—before engaging with AI experts, make sure that they possess the depth of knowledge that will be required to appropriately implement the AI technology at hand. This could be proven over several conversations, as well as evaluating their experience in the field.
- Opportunity costs—as shown in our sections on the planetary and social costs of the AI Hype, investing in AI brings its own opportunity costs which are worth weighing up. Do not neglect the entire AI infrastructure and supply chain.
- The AI infrastructure—be aware of the sustainability of the entire AI infrastructure, including natural resources, human labour and energy usage. Are you aware of the entire AI supply chain and the ethical, economical, social and ecological considerations in each component?
- Context—being aware of the narratives at play (such as the difference in national AI strategies we have observed) can prove a valuable tool in evaluating the need for AI in your organisation. Acknowledging that there are actors with vested interests in presenting AI as the only way forward can help produce a measured judgement on its necessity.

The above can also be reinforced through investigating how AI hype narratives have been crafted in the past. While this AI hype cycle is far greater than its previous iterations, contextualising the current AI obsession helps to create a grounded understanding of how the AI hype cycle works.

## 5 Conclusion

In this article, we have addressed the mechanisms of AI hype and accounted for the planetary and social consequences of that hype. The mechanisms include anthropomorphism, exaggerated AI literacy or so called “AI-experts,” the geopolitical narratives and FOMO the overuse and

misappropriation of the term “AI”. While we recognise that there are other mechanisms also contributing to the hype, we conclude that these mechanisms are significant in shaping the global socio-technical narratives and imaginaries woven around emerging AI technologies in our contemporary society. The consequences of this hype, that are often overlooked, include planetary costs, such as the material aspects of data centers, energy usage and vast amounts of limited natural resources required realise the development and deployment of AI technologies. Furthermore we have identified the social costs, which include socio-economic costs such as job loss and job polarisation as well as costs to human intelligence including knowledge decay and knowledge corruption and post-truth. To round off we provided suggestions for moving forward, which can be adopted by developers, designers, regulators and the public alike. We hope this paper serves to cut through the AI hype, bringing a socio-technically grounded perspective on the material infrastructure of AI.

## Declarations

**Conflict of interest** All authors certify that they have no affiliations with or involvement in any organisation or entity with any financial interest or non-financial interest in the subject matter or materials discussed in this manuscript. The authors did not receive support or funding from any organisation for the submitted work.

**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

- Leaver, T., Srdarov, S.: Chatgpt isn’t magic: The hype and hypocrisy of generative artificial intelligence (ai) rhetoric. *M/C J.* **26**(5) (2023)
- Russell, S.J., Norvig, P.: *Artificial Intelligence: a modern approach*, 3rd edn. Pearson, London (2009)
- Smith, P., Smith, L.: This season’s artificial intelligence (ai): is today’s ai really that different from the ai of the past? Some reflections and thoughts. *AI Ethics* 1–4 (2023)
- IBM: IBM Global AI Adoption Index 2022. <https://www.ibm.com/watson/resources/ai-adoption> (2022)
- Porter, J.: ChatGPT continues to be one of the fastest-growing services ever. *Verge*. <https://www.theverge.com/2023/11/6/23948386/chatgpt-active-user-count-openai-developer-conference> (2023)
- Amar, Z., Ramsay, N.: Charity digital skills report 2023. <https://charitydigitalskills.co.uk/the-charity-digital-skills-report-introduction/> (2023)
- AI activity in UK businesses: executive summary. <https://www.gov.uk/government/publications/ai-activity-in-uk-businesses/ai-activity-in-uk-businesses-executive-summary> (2022)
- Pause Giant AI Experiments: An open letter. <https://futureoflife.org/open-letter/pause-giant-ai-experiments/> (2023)
- Placani, A.: Anthropomorphism in ai: hype and fallacy. *AI Ethics* 1–8 (2024)
- Pereira, L.M., Santos, F.C., Lenaerts, T., et al.: To regulate or not: a social dynamics analysis of an idealised ai race. *J. Artif. Intell. Res. Artif. Intell. Res.* **69**, 881–921 (2020)
- AINowInstitute: US-China AI race: AI policy as industrial policy. AINow Institute. <https://ainowinstitute.org/publication/us-china-ai-race> (2023)
- Matsakis, L.: The Princeton researchers calling out “ai snake oil”. *Semafor*. <https://www.semafor.com/article/09/15/2023/the-princeton-researchers-calling-out-ai-snake-oil> (2023)
- Bareis, J., Katzenbach, C.: Talking ai into being: the narratives and imaginaries of national ai strategies and their performative politics. *Sci. Technol. Hum. Values* **47**(5), 855–881 (2022)
- Walton, D.: Gen Z staff admit to pretending they have used AI tool. <https://www.talintinternational.com/gen-z-staff-admit-to-pretending-they-have-used-ai-tool/> (2023)
- Monserrate, S.G.: The cloud is material: On the environmental impacts of computation and data storage. MIT Schwarzman College of Computing (2022)
- Bender, E.M., Gebru, T., McMillan-Major, A., Shmitchell, S.: On the dangers of stochastic parrots: can language models be too big? In: *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, pp. 610–623 (2021)
- Jaimovich, N., Siu, H.E.: Job polarization and jobless recoveries. *Rev. Econ. Stat.* **102**(1), 129–147 (2020). <https://doi.org/10.1162/resta00875>. (Accessed 2023-11-09)
- Howard, C.: *Innovating in uncertain times: Lessons from 2022*. Harvard Business Review. <https://hbr.org/2022/12/innovating-in-uncertain-times-lessons-from-2022> (2022)
- Glancey, J.: *Concorde: the Rise and Fall of the Supersonic Airliner*. Atlantic Books Ltd, London (2015)
- Simon, H.A.: *The Shape of Automation for Men and Management*, vol. 13. Harper & Row, New York (1965)
- Minsky, M.L.: *Computation*. Prentice-Hall, Englewood Cliffs (1967)
- Hendler, J.: Avoiding another AI winter. *IEEE Intell. Syst. Intell. Syst.* **23**(02), 2–4 (2008)
- Buolamwini, J., Gebru, T.: Gender shades: intersectional accuracy disparities in commercial gender classification. In: *Conference on Fairness, Accountability and Transparency*, pp. 77–91 (2018). PMLR
- Microsoft Research: Human parity speech recognition. <https://www.microsoft.com/en-us/research/project/human-parity-speech-recognition/#:~:text=This%20ongoing%20project%20aims%20to,focus%20on%20unconstrained%20conversational%20speech>. Accessed 19 Feb 2024
- Microsoft Azure: Azure cognitive services has achieved human parity in image captioning. <https://azure.microsoft.com/en-us/updates/azure-cognitive-services-has-achieved-human-parity-in-image-captioning/>. Accessed 19 Feb 2024
- Windsor, M.: This radiologist is helping doctors see through the hype to an AI future. <https://www.uab.edu/reporter/people/achievements/item/9925-this-radiologist-is-helping-doctors-see-through-the-hype-to-an-ai-future> (2022)
- Arcas, B.A.Y.: Artificial general intelligence is already here. *NOEMA*. <https://www.noemamag.com/artificial-general-intelligence-is-already-here/> (2023)

28. Bubeck, S., Chandrasekaran, V., Eldan, R., Gehrke, J., Horvitz, E., Kamar, E., Lee, P., Lee, Y.T., Li, Y., Lundberg, S., et al.: Sparks of artificial general intelligence: early experiments with gpt-4. *arXiv:2303.12712* (2023)
29. The generative AI landscape: top startups, venture capital firms, and more. CB Insights. <https://www.cbinsights.com/research/generative-ai-funding-top-startups-investors/> (2023)
30. The generative AI market map: 335 vendors automating content, code, design, and more. CB Insights. <https://www.cbinsights.com/research/generative-ai-startups-market-map/> (2023)
31. Tung, L.: Chatgpt just became the fastest-growing “app” of all time. ZDNET (2023). <https://www.zdnet.com/article/chatgpt-just-became-the-fastest-growing-app-of-all-time/>
32. Siegel, E.: Models are rarely deployed: an industry-wide failure in machine learning leadership. KDnuggets. <https://www.kdnuggets.com/models-are-rarely-deployed-an-industry-wide-failure-in-machine-learning-leadership.html> (2022)
33. Strasser, T.: Narrow AI-powered visualization facilitation tools in foreign language learning: a visual approach promoting equal opportunities in foreign language grammar teaching. In: Facilitating Visual Socialities: Processes, Complications and Ethical Practices, pp. 85–98. Springer, Berlin (2023)
34. Li, M., Suh, A.: Anthropomorphism in ai-enabled technology: a literature review. *Electron. Mark.* **32**(4), 2245–2275 (2022)
35. Epley, N., Waytz, A., Cacioppo, J.T.: On seeing human: a three-factor theory of anthropomorphism. *Psychol. Rev.* **114**(4), 864 (2007)
36. Xie, Y., Zhu, K., Zhou, P., Liang, C.: How does anthropomorphism improve human-ai interaction satisfaction: a dual-path model. *Comput. Hum. Behav.* **148**, 107878 (2023)
37. Blut, M., Wang, C., Wunderlich, N.V., Brock, C.: Understanding anthropomorphism in service provision: a meta-analysis of physical robots, chatbots, and other AI. *J. Acad. Mark. Sci.* **49**, 632–658 (2021)
38. Salles, A., Evers, K., Farisco, M.: Anthropomorphism in AI. *AJOB Neurosci. Neurosci* **11**(2), 88–95 (2020)
39. Jacobs, O., Pazhoohi, F., Kingstone, A.: Brief exposure increases mind perception to chatgpt and is moderated by the individual propensity to anthropomorphize. *PsyArXiv* (2023)
40. Spatola, N., Wudarczyk, O.A.: Ascribing emotions to robots: explicit and implicit attribution of emotions and perceived robot anthropomorphism. *Comput. Hum. Behav. Hum. Behav.* **124**, 106934 (2021)
41. Wong, J., Kim, J.: Chatgpt is more likely to be perceived as male than female. *arXiv:2305.12564* (2023)
42. Shaw-Garlock, G.: Gendered by Design: Gender Codes in Social Robotics. *Social Robots: Boundaries, Potential, Challenges*, pp. 199–218. Routledge, New York (2017)
43. West, M., Kraut, R., Ei Chew, H.: I’d blush if i could: closing gender divides in digital skills through education. UNESCO (2019)
44. Farisco, M., Salles, A.: Ch(e)atgpt. dissimulation of reality and the need for an ethical leap. SSRN (2023)
45. Renzullo, D.: Anthropomorphized AI as capitalist agents: the price we pay for familiarity. Montreal AI Ethics Institute (2019)
46. Shevlin, H., Halina, M.: Apply rich psychological terms in ai with care. *Nat. Mach. Intell.* **1**(4), 165–167 (2019)
47. Campolo, A., Crawford, K.: Enchanted determinism: power without responsibility in artificial intelligence. *Engaging Science, Technology, and Society* (2020)
48. Waytz, A., Cacioppo, J., Epley, N.: Who sees human? the stability and importance of individual differences in anthropomorphism. *Perspect. Psychol. Sci.* **5**(3), 219–232 (2010)
49. Formosa, P., Ryan, M.: Making moral machines: why we need artificial moral agents. *AI Soc.* **36**, 839–851 (2021)
50. Cervantes, J.-A., Lopez, S., Rodriguez, L.-F., Cervantes, S., Cervantes, F., Ramos, F.: Artificial moral agents: a survey of the current status. *Sci. Eng. Ethics* **26**, 501–532 (2020)
51. Crawford, K.: *The Atlas of AI: Power, Politics, and the Planetary Costs of Artificial Intelligence*. Yale University Press, London (2021)
52. Altenried, M.: The platform as factory: crowdwork and the hidden labour behind artificial intelligence. *Cap. Class* **44**(2), 145–158 (2020)
53. Schmidt, F.A.: Crowdsourced production of ai training data: how human workers teach self-driving cars how to see. Technical report, Working Paper Forschungsforderung (2019)
54. Ekbia, H.R., Nardi, B.A.: *Heteromation, and other stories of computing and capitalism*. MIT Press, Cambridge (2017)
55. Pasquinelli, M., Joler, V.: The noosphere manifested: AI as instrument of knowledge extractivism. *AI Soc.* **36**, 1263–1280 (2021)
56. Hurt, A.: Robots and artificial intelligence have ancient mythology origins. *Discover Magazine*. <https://www.discovermagazine.com/technology/robots-and-artificial-intelligence-have-ancient-mythology-origins> (2022)
57. Adeshola, I., Adepoju, A.P.: The opportunities and challenges of chatgpt in education. *Interact. Learn. Environ.* (2023). <https://doi.org/10.1080/10494820.2023.2253858>
58. Poel, I.V.D.: Three philosophical perspectives on the relation between technology and society, and how they affect the current debate about artificial intelligence. *Hum. Aff. Aff.* **30**(4), 499–511 (2020). <https://doi.org/10.1515/humaff-2020-0042>
59. Heder, M.: AI and the resurrection of technological determinism. *Informacios Tarsadalom XXI* **21**(2), 119–130 (2021). <https://doi.org/10.22503/infstars.XXI.2021.2.8>
60. Alba, D.: Critics say Google’s AI tool generates madeup facts, potentially dangerous advice. The Seattle Times Company. <https://www.seattletimes.com/business/even-google-insiders-are-questioning-bard-ai-chatbots-usefulness/> (2023)
61. Peralta, P.: Employees are lying about their AI skills to keep their jobs. *Employee Benefit News*. <https://www.benefitnews.com/news/why-employees-feel-the-need-to-exaggerate-their-ai-skills> (2023)
62. Sridhar, D.: Why can’t some scientists just admit they were wrong about Covid? *Guardian News and Media*. <https://www.theguardian.com/commentisfree/2022/mar/24/scientists-wrong-covid-virus-experts> (2022)
63. Schmidt, E.: Innovation power: why technology will define the future of geopolitics. *Foreign Aff.* **102**, 38 (2023)
64. Cai, C., Zhang, R.: In: Pashentsev, E. (ed.) *Malicious Use of Artificial Intelligence, Uncertainty, and U.S.–China Strategic Mutual Trust*, pp. 377–396. Springer, Cham (2023). <https://doi.org/10.1007/978-3-031-22552-914>
65. Kim, J.: Traveling ai-essentialism and national ai strategies: a comparison between south korea and france. *Rev. Policy Res.* (2023)
66. Bharadiya, J.P.: Machine learning and ai in business intelligence: trends and opportunities. *Int. J. Comput.* **48**(1), 123–134 (2023)
67. The AI Index report measuring trends in artificial intelligence. Stanford University—AI Index (2023). <https://aiindex.stanford.edu/report/>
68. Ahmed, N., Wahed, M., Thompson, N.C.: The growing influence of industry in ai research. *Science* **379**(6635), 884–886 (2023)
69. Baum, S.D.: Reconciliation between factions focused on near-term and long-term artificial intelligence. *AI Soc.* **33**(4), 565–572 (2018)
70. Cave, S., Ohigeartaigh, S.S.: Bridging near-and long-term concerns about ai. *Nat. Mach. Intell.* **1**(1), 5–6 (2019)
71. Roose, K.: A.I. Poses “risk of extinction,” industry leaders warn. *The New York Times*. <https://www.nytimes.com/2023/05/30/>

- [technology/ai-threat-warning.html?bgrp=c&smid=url-share](#) (2023)
72. Hermann, I.: Artificial intelligence in fiction: between narratives and metaphors. *AI Soc.* **38**(1), 319–329 (2023)
  73. Bode, I., Qiao-Franco, G.: AI geopolitics and international relations: a divided world behind contested conceptions of human control. In: *Handbook on Public Policy and Artificial Intelligence*. Edward Elgar Publishing, Cheltenham (2022)
  74. Coeckelbergh, M.: Narrative responsibility and artificial intelligence: how AI challenges human responsibility and sense-making. *AI Soc.* **38**(6), 2437–2450 (2023)
  75. Sloane, M.: Threading innovation, regulation, and the mitigation of ai harm: examining ethics in national ai strategies. Sloane, Mona.” In: *Threading Innovation, Regulation, and the Mitigation of AI Harm: Examining Ethics in National AI Strategies*.” *The Global Politics of Artificial Intelligence*. Chapman and Hall, pp. 1–28 (2022)
  76. Birhane, A.: Algorithmic colonization of Africa. *SCRIPTed* **17**, 389 (2020)
  77. Benjamin, R.: Race after technology. In: *Social Theory Rewired*, pp. 405–415. Routledge, UK (2023)
  78. Browne, S.: *Dark Matters: On the Surveillance of Blackness*. Duke University Press, Durham (2015)
  79. Jungherr, A.: Artificial intelligence and democracy: a conceptual framework. *Soc. Media Soc.* **9**(3), 20563051231186350 (2023)
  80. Png, M.-T.: At the tensions of south and north: critical roles of global south stakeholders in ai governance. In: *Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency*, pp. 1434–1445 (2022)
  81. Marr, B.: The 10 best examples of how companies use artificial intelligence in practice. Bernard Marr. <https://bernardmarr.com/the-10-best-examples-of-how-companies-use-artificial-intelligence-in-practice/> (2021)
  82. Gandhi, L.: A history of “snake oil salesmen”. NPR. <https://www.npr.org/sections/codeswitch/2013/08/26/215761377/a-history-of-snake-oil-salesmen> (2013)
  83. Babina, T., Fedyk, A., He, A., Hodson, J.: Artificial intelligence, firm growth, and product innovation. *J. Financ. Econ.* **151**, 103745 (2024). <https://doi.org/10.1016/j.jfineco.2023.103745>
  84. Jackson, S.: ‘The CEO of the company behind AI chatbot ChatGPT says the worst-case scenario for artificial intelligence is ‘lights out for all of us’’. *Business Insider*. [https://www.businessinsider.com/chatgpt-openai-ceo-worst-case-ai-lights-out-for-all-2023-1?r=US&IR=T&utm\\_source=copy-link&utm\\_medium=referral&utm\\_content=topbar](https://www.businessinsider.com/chatgpt-openai-ceo-worst-case-ai-lights-out-for-all-2023-1?r=US&IR=T&utm_source=copy-link&utm_medium=referral&utm_content=topbar) (2023)
  85. PricewaterhouseCoopers: PwC’s global artificial intelligence study: sizing the prize. <https://www.pwc.com/gx/en/issues/data-and-analytics/publications/artificial-intelligence-study.html>
  86. Federspiel, F., Mitchell, R., Asokan, A., Umana, C., McCoy, D.: Threats by artificial intelligence to human health and human existence. *BMJ Glob. Health* **8**(5), 010435 (2023)
  87. Richards, B.: The illusion of ai’s existential risk. *NOEMA*. <https://www.noemamag.com/the-illusion-of-ais-existential-risk/> (2023)
  88. Raghavan, M., Barocas, S., Kleinberg, J., Levy, K.: Mitigating bias in algorithmic hiring: evaluating claims and practices. In: *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*, pp. 469–481 (2020)
  89. Sunstein, C.R.: Governing by algorithm? no noise and (potentially) less bias. *Duke LJ* **71**, 1175 (2021)
  90. Chamorro-Premuzic, T., Akhtar, R.: Should companies use ai to assess job candidates. *Harvard Bus. Rev.* **17** (2019)
  91. Drage, E., Mackereth, K.: Does ai debias recruitment? race, gender, and ai’s “eradication of difference.” *Philos. Technol.* **35**(4), 89 (2022)
  92. Reed, J.: Can enterprise tech redeem itself with Generative AI? Vishal Sikka on doing AI right, and avoiding generative AI Snake Oil Vendors. *Diginomica*. <https://diginomica.com/can-enterprise-tech-redeem-itself-generative-ai-vishal-sikka-doing-ai-right-and-avoiding-generative> (2023)
  93. LaGrandeur, K.: The consequences of ai hype. *AI Ethics* **1–4** (2023)
  94. Strubell, E., Ganesh, A., McCallum, A.: Energy and policy considerations for deep learning in nlp. *arXiv:1906.02243* (2019)
  95. Ting-Fang, C., Li, L.: TSMC predicts revenues could grow 26. <https://asia.nikkei.com/Business/Tech/Semiconductors/TSMC-predicts-revenues-could-grow-26-this-year-on-AI-boom> (2024)
  96. Wong, J.: TSMC thinks 2024 will repower chip sales. *Wall Street J.* <https://www.wsj.com/business/earnings/tsmc-thinks-2024-will-repower-chip-sales-aed4ac6a> (2024)
  97. Hogan, M.: Data flows and water woes: the utah data center. *Big Data Soc.* **2**(2), 2053951715592429 (2015). <https://doi.org/10.1177/2053951715592429>
  98. Zahodiakin, P.: Report: “Critical Action” needed to fight enormous energy waste at data centers. *MintPress News*. <https://www.mintpressnews.com/report-critical-action-needed-fight-enormous-energy-waste-data-centers/196535/> (2014)
  99. FitzGerald, D.: Data centers and hidden water use. *Wall Street J.* <https://www.wsj.com/articles/data-centers-1435168386> (2015)
  100. Dieter, C.A., Maupin, M.A., Caldwell, R.R., Harris, M.A., Ivahnenko, T.I., Lovelace, J.K., Barber, N.L., Linsey, K.S.: *Estimated Use of Water in the United States in 2015*, vol. 1441. <https://doi.org/10.3133/cir1441>. <https://pubs.usgs.gov/publication/cir1441> (2018)
  101. Cowsls, J., Tsamados, A., Taddeo, M., Floridi, L.: The AI gambit: leveraging artificial intelligence to combat climate change—opportunities, challenges, and recommendations. *AI Soc.* **38**(1), 283–307 (2023). <https://doi.org/10.1007/s00146-021-01294-x>
  102. Ise, T., Oba, Y.: Forecasting climatic trends using neural networks: an experimental study using global historical data. *Front. Robot. AI* (2019). <https://doi.org/10.3389/frobt.2019.00032>
  103. Mytton, D.: Data centre water consumption. *NPJ Clean Water* **4**(1), 11 (2021)
  104. Sutherland, T., Bopp, G.: The pacific futures of subsea data centers. *New Media Soc.* **25**(2), 345–360 (2023). <https://doi.org/10.1177/14614448221149944>
  105. Miller, R.: Northlake buys extra water for microsoft site. *Data Center Knowledge*. <https://www.datacenterknowledge.com/archives/2008/08/20/northlake-buys-extra-water-for-microsoft-site> (2008)
  106. Moss, S.: TSMC orders water by the truckload to keep chip production going during growing drought. *All Content RSS*. <https://www.datacenterdynamics.com/en/news/tsmc-orders-water-truck-load-keep-chip-production-going-during-growing-drought/> (2021)
  107. Pascual, M.G.: The dirty secret of artificial intelligence. *EL PAIS English*. <https://english.elpais.com/science-tech/2023-03-23/the-dirty-secret-of-artificial-intelligence.html> (2023)
  108. Smith, I.: Ireland’s data centres are using as much electricity as all city homes. *Euronews*. <https://www.euronews.com/green/2023/06/13/data-centres-gobble-up-18-of-irelands-electricity-as-country-struggles-with-climate-target> (2023)
  109. Vries, A.: The growing energy footprint of artificial intelligence. *Joule* **7**(10), 2191–2194 (2023)
  110. Leffer, L.: The ai boom could use a shocking amount of electricity. *Sci. Am.* (2023)
  111. Cohan, P.: As chatgpt and other ai tools increase energy demand, here’s what investors need to know. *Forbes* (2023)

112. Mauer, M.: KPMG plans 2 billion investment in AI and cloud services. *Wall Street J.* <https://www.wsj.com/articles/kpmg-plans-2-billion-investment-in-ai-and-cloud-services-e4fd0dd5>
113. Feng, E.: Epic drought in Taiwan pits farmers against high-tech factories for water. *NPR.* <https://www.npr.org/sections/goatsandsoda/2023/04/19/1170425349/epic-drought-in-taiwan-pits-farmers-against-high-tech-factories-for-water> (2023)
114. Panikkar, R., Saleh, T., Szybowski, M., Whiteman, R.: The potential of machine learning in services operations—McKinsey. <https://www.mckinsey.com/capabilities/operations/our-insights/operationalizing-machine-learning-in-processes> (2021)
115. Davenport, T., Malone, K.: Deployment as a critical business data science discipline. *Harvard Data Sci. Rev.* (2021). <https://doi.org/10.1162/99608f92.90814c32>
116. Schwab, K.: *The Fourth Industrial Revolution*, First U.S. Edition edn. Crown Business, New York (2016)
117. AI Index Report. Technical report, HAI, Stanford University. <https://aiindex.stanford.edu/report/#individual-chapters> (2023). Accessed 8 Nov 2023
118. Temkin, M.: In the world of startup valuations, there’s generative AI and everything else. *PitchBook*. Section: News & Analysis. <https://www.pitchbook.com/news/articles/early-stage-valuations-generative-ai-compare-vc> (2023)
119. Perifanis, N.-A., Kitsios, F.: Investigating the influence of artificial intelligence on business value in the digital era of strategy: a literature review. *Information* **14**(2), 85 (2023)
120. Chui, M., Hazan, E., Roberts, R., Singla, A., Smaje, K., Sukharevsky, A., Yee, L., Zimmel, R.: Economic potential of generative AI. <https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/the-economic-potential-of-generative-ai-the-next-productivity-frontier#/>
121. Elliott, A.: *The Culture of AI*, p. 268. Routledge, 2 Park Square, Milton Park, Abingdon, Oxon (2018)
122. Islam, M.M., Shuford, J.: A survey of ethical considerations in AI: navigating the landscape of bias and fairness. *J. Artif. Intell. Gen. Sci. (JAIGS)* **1**(1) (2024) (ISSN: 3006-4023)
123. Schwartz, R., Vassilev, A., Greene, K., Perine, L., Burt, A., Hall, P.: Towards a standard for identifying and managing bias in artificial intelligence. <https://doi.org/10.6028/NIST.SP.1270>. <https://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.1270.pdf>
124. Laestadius, L., Bishop, A., Gonzalez, M., Illencic, D., Campos-Castillo, C.: Too human and not human enough: a grounded theory analysis of mental health harms from emotional dependence on the social chatbot replika. *New Media Soc.* 14614448221142007 (2022)
125. Arthur, W., Jr., Bennett, W., Jr., Stanush, P.L., McNelly, T.L.: Factors that influence skill decay and retention: a quantitative review and analysis. *Hum. Perform.* **11**(1), 57–101 (1998). <https://doi.org/10.1207/s15327043hup11013>. (Accessed 2023-11-09)
126. Luccioni, A.S., Viviano, J.D.: What’s in the box? A preliminary analysis of undesirable content in the common crawl corpus. <http://arxiv.org/abs/2105.02732> (2021). Accessed 2 Dec 2022
127. Werbos, P.J.: The new ai: Basic concepts, and urgent risks and opportunities in the internet of things. In: *Artificial Intelligence in the Age of Neural Networks and Brain Computing*, pp. 93–127. Elsevier (2024)
128. Dykes, B.: Generative AI: why an AI-enabled workforce is a productivity game changer. *Forbes*. section: entrepreneurs. <https://www.forbes.com/sites/brentdykes/2023/04/12/generative-ai-why-an-ai-enabled-workforce-is-a-productivity-game-changer/> (2023)
129. Economic potential of generative AI. Technical report, McKinsey & Company. <https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/the-economic-potential-of-generative-ai-the-next-productivity-frontier#/> (2023). Accessed 9 Nov 2023
130. Shine, I., Whiting, K.: These are the jobs most likely to be lost—and created—because of AI. *World Economic Forum* (2023). <https://www.weforum.org/agenda/2023/05/jobs-lost-created-ai-gpt/>
131. Bonab, A.B., Rudko, I., Bellini, F.: A review and a proposal about socio-economic impacts of artificial intelligence. In: *Business Revolution in a Digital Era: 14th International Conference on Business Excellence, ICBE 2020, Bucharest, Romania*, pp. 251–270. Springer (2021)
132. Nazareno, L., Schiff, D.S.: The impact of automation and artificial intelligence on worker well-being. *Technol. Soc.* **67**, 101679 (2021). <https://doi.org/10.1016/j.techsoc.2021.101679>
133. Heikkilä, M.: Responsible AI has a burnout problem—MIT technology review. <https://www.technologyreview.com/2022/10/28/1062332/responsible-ai-has-a-burnout-problem/> (2022)
134. Korteling, J.H., Boer-Visschedijk, G.C., Blankendaal, R.A., Boonekamp, R.C., Eikelboom, A.R.: Human-versus artificial intelligence. *Front. Artif. Intell.* **4**, 622364 (2021)
135. Cramarencu, R.E., Burca-Voicu, M.I., Dabija, D.C.: The impact of artificial intelligence (ai) on employees’ skills and well-being in global labor markets: a systematic review. *Oeconomia Copernicana* **14**(3), 731–767 (2023)
136. Volz, K., Yang, E., Dudley, R., Lynch, E., Dropps, M., Dorneich, M.C.: An evaluation of cognitive skill degradation in information automation. In: *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, vol. 60 (2016). <https://doi.org/10.1177/1541931213601043>
137. Arthur, W., Day, E.A.: Skill decay: the science and practice of mitigating skill loss and enhancing retention. In: *The Oxford Handbook of Expertise*, pp. 1085–1108. Oxford University Press, Oxford (2018). <https://doi.org/10.1093/oxfordhb/9780198795872.013.47>
138. Beynon, S.: Have soldiers forgotten how to use a map and compass? *Military*. <https://www.military.com/daily-news/2022/12/16/have-soldiers-forgotten-how-use-map-and-compass.html> (2022)
139. Fui-Hoon Nah, F., Zheng, R., Cai, J., Siau, K., Chen, L.: *Generative AI and ChatGPT: Applications, Challenges, and AI-Human Collaboration*. Taylor & Francis, New York (2023)
140. Briggs, J., Kodnani, D.: Global economics analyst the potentially large effects of artificial intelligence on economic growth. Technical report, Goldman Sachs & Co LLC. <https://www.gspublishing.com/content/research/en/reports/2023/03/27/d64e052b-0f6e-45d7-967b-d7be35fabd16.html> (2023)
141. Kasneci, E., Seßler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günemann, S., Hüllermeier, E., et al.: Chatgpt for good? On opportunities and challenges of large language models for education. *Learn. Individ. Differ.* **103**, 102274 (2023)
142. Narayanan, A.: [@random walker]: people are excited about using ChatGPT for learning. It’s often very good. But the danger is that you can’t tell when it’s wrong unless you already know the answer. I tried some basic information security questions. In most cases the answers sounded plausible but were in fact BS. Publication Title: Twitter Type: Tweet. <https://twitter.com/randomwalker/status/1598383507214020608> (2022). Accessed 9 Nov 2023
143. Ji, Z., Lee, N., Frieske, R., Yu, T., Su, D., Xu, Y., Ishii, E., Bang, Y., Chen, D., Chan, H.S., Dai, W., Madotto, A., Fung, P.: Survey of hallucination in natural language generation. *ACM Comput. Surv.Comput. Surv.* (2023). <https://doi.org/10.1145/3571730>. (arXiv:2202.03629 [cs])
144. Merken, S., Merken, S.: New York lawyers sanctioned for using fake Chat- GPT cases in legal brief. *Reuters*. <https://www.reuters.com/legal/new-york-lawyers-sanctioned-using-fake-chatgpt-cases-legal-brief-2023-06-22/> (2023)
145. Cecco, L.: Air Canada ordered to pay customer who was misled by airline’s chatbot. *Guardian*. <https://www.theguardian.com/world/2024/feb/16/air-canada-chatbot-lawsuit> (2024)

146. Heikkilä, M.: How AI-generated text is poisoning the internet. Algorithm. <https://www.technologyreview.com/2022/12/20/1065667/how-ai-generated-text-is-poisoning-the-internet/> (2022)
  147. Leib, M., Kobis, N., Rilke, R.M., Hagens, M., Irlenbusch, B.: Corrupted by algorithms? How AI-generated and human-written advice shape (dis) honesty. *Econ. J.* **134**(658), 766–784 (2024)
- Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.