Discover Artificial Intelligence

Discover

Perspective

The importance of humanizing AI: using a behavioral lens to bridge the gaps between humans and machines

A. Fenwick[1] · **G. Molnar**[2]

Received: 29 May 2022 / Accepted: 28 July 2022
Published online: 25 August 2022

Abstract
One of the biggest challenges in Artificial Intelligence (AI) development and application is the lack of consideration for human enhancement as a cornerstone for its operationalization. Nor is there a universally accepted approach that guides best practices in this field. However, the behavioral science field offers suggestions on how to develop a sustainable and enriching relationship between humans and intelligent machines. This paper provides a three-level (micro, meso and macro) framework on how to humanize AI with the intention of enhancing human properties and experiences. It argues that humanizing AI will help make intelligent machines not just more efficient but will also make their application more ethical and human-centric. Suggestions to policymakers, organizations, and developers are made on how to implement this framework to fix existing issues in AI and create a more symbiotic relationship between humans and machines moving into the future.

1 Introduction

The concept of artificial intelligence (AI) has been around since antiquity (e.g., [1]). It is clear from investigative literature (e.g., [2, 3]) popular culture (e.g., [4]), and even ancient philosophers (e.g., [5]) that humans have long been intrigued by the idea of creating artificial life, be it from stone or machines, with some sort of intelligence to help, serve, or protect human life. Modern AI has matured into a reputable science and technology thanks to the development of powerful computers, a better theoretical understanding of what AI is and how it works, and the availability of large amounts of data [6].

AI has been defined in many ways [7–10]. The different interpretations of AI generally converge to two major descriptions: (i) 'the ability to think, understand, and problem-solve like a human' and (ii) 'the ability to mimic human thinking'. Another important aspect in defining AI is by the words 'artificial' and 'intelligence'. Artificial is often referred to as anything humans build (e.g., [9, 11]). 'Intelligence' refers to a computer's ability to learn (independently), understand and reason like a human [12]. However, there is currently no clear consensus on how to define intelligence (e.g. [13].). Instead, more philosophical concepts of intelligence (Weak AI and Strong AI) are often used to differentiate between varying degrees of machine intelligence (e.g. [12].). Machine Learning (ML) is often used interchangeably with AI, though related, they are not exactly the same. Machine learning is a subset of AI and describes a set of techniques that is used to solve data-related problems without being explicitly programmed [14]. In this article, by AI we refer to both rule-based and machine learning techniques [12], unless mentioned otherwise.

AI technology, as including machine learning techniques, is capable of processing information (e.g., [15]), identifying patterns (e.g., [16]), making predictions (e.g., [17]), and even operating robots and autonomous devices (e.g., [18, 19]).

✉ A. Fenwick, ali.fenwick@faculty.hult.edu; G. Molnar, gabor.molnar@colorado.edu | [1]Hult International Business School, Dubai, UAE. [2]University of Colorado Boulder, ATLAS Institute, Boulder, CO, USA.

Machine Learning (ML) and a subset called Deep Learning (DL) are powering most digital applications today, providing efficiencies and new avenues for value creation. This trend will only continue as we move into the future, standing at the forefront of the 4th Industrial Revolution.

However, the future of AI is not without concerns. In recent years, ethical and moral dilemmas have emerged regarding how AI is being used in modern-day applications (e.g., [20, 21]), specifically, the use of AI in the public domain and the (un)intentional consequences machine learning algorithms have on human well-being and economic choices (e.g., [22]). In addition, policymakers who lack knowledge in the AI field are not always up to speed on preventing unethical or inhumane use of technology, nor do they want to limit their countries' digital competitiveness due to AI policies that are too stringent. It's clear that the advancement of AI needs to be governed by more human-centric principles (referred to hereafter as 'humanizing AI'), ones that are easily understood by all stakeholders and that benefit society.

If left undefined, humanizing AI is an ambiguous concept and a challenge in humanizing AI is that there is no universally accepted approach that guides the best practice for design and use of AI. In a narrow definition, humanizing AI means the process of creating and using AI that (i) understands not only human emotions but human unconscious dynamics, (ii) has the capability to interact with humans in a natural, human-like manner, and (iii) during this interaction it processes information in a similar way that people do. Producing AI that processes information similarly to people does not automatically produce a symbiotic relationship between humans and AI, however, it is a requirement to build a trusting relationship with machines. We believe that humanizing AI needs to manifest at multiple levels, which are interconnected, to help bridge the gaps between humans and machines which is currently lacking (e.g., [23]).

In this paper, we argue that AI conceptualization and application need to be less artificial and more human-like. We are not arguing that AI needs to look more like human beings, but rather that humanizing AI sets a foundation for AI development to integrate aspects of human intelligence, cognition and behavior which complement human limitations and promote human values. We contribute to the existing literature on AI human-centric advancements by providing a motivational framework to explain the operationalization of AI today. This paper also provides a multilayered behavioral approach to developing and applying AI in a more humane and equitable way. Existing literature on usability, user experience, human-centered design, and human–computer interaction (e.g., [24–28]) all have behavioral elements, but not all of them consider a multilayered approach. Even the ISO standards related to human-centered design for interactive systems [29] lack a multilevel viewpoint. However, [26–28] discuss the necessity of understanding technology development (and AI development with it) from a multilevel perspective. Our paper provides a unique viewpoint in this discussion. Finding a way to build a symbiotic relationship with AI as we transition into a digital world is of crucial importance if humanity wants to benefit from technology [30, 31].

To discuss the rational for our framework and multilayered approach, we structure the paper in the following way. First, we provide an overview of the current concerns with AI. Next, we explain about the importance of humanizing AI and introduce our framework how to humanize AI from a multilevel perspective. Finally, our paper concludes and suggests future research directions.

2 Current concerns with AI

AI's existing and potential benefits are undeniable, but it may do more harm than good in some cases. Academics and business professionals frequently raise concerns about possible biases, lack of transparency and explainability, power imbalance, and liability – to mention a few issues [32].

Even the best AI tools can institutionalize existing biases that might be present in the training data. The creation and deployment of AI solutions have been a predominantly male orientation, restricted to specific areas of the world, such as the US and China (e.g., [33]). The active field of AI has limited diversity in terms of gender and race [34]. This not only unbalances the beneficiaries of this technology, but also limits diversity in use and propagates potential bias in how AI functions (e.g., [33]).

Our experience with real-life AI projects is that companies often underestimate the difficulties in creating a clean and unbiased training dataset. In addition to being aware of the problem of possible bias, domain experts need to do a lot of work to eliminate embedded biases in the training data and make a clean dataset available for the AI algorithm to learn.

We have very little visibility and knowledge of how and why an AI tool makes its decisions, especially with deep learning neural network approaches. Further research is necessary to make deep learning approaches explainable. In general, businesses (and humans) will only trust AI-enabled systems if they can fully understand how AI makes decisions and predictions [35]. Tackling issues of explainability is difficult, as biased modeling is often found after the development stage [36]. Therefore,

researchers need to focus not only on post-modeling explainability but also address explainability in earlier stages. Current standardization efforts (e.g., [37]) aim to bring a harmonized approach to address the issue of transparency. Further to building these frameworks, it is also important that experts from various domains speak the same language. To address this, the IEEE 7007 Working Group has developed a set of ontologies representing norms and ethical principles, data privacy and protection, transparency and accountability, and ethical violation management [38].

There seems to be a global consensus that policymakers will need to establish rules and control mechanisms to ensure that AI tools are safe and will aid rather than harm society (e.g., [39, 40]). However, there is no global consensus on how to build a framework to regulate AI. For example, the European Union and the United States differ widely on how they want to address AI regulation.

To create an EU-wide framework, the European Commission issued a draft proposal of the Artificial Intelligence Act to the European Parliament on 21 April 2021. In the proposal, they established a framework for determining whether an AI application poses a significant risk which would subject it to additional obligations, including a conformity assessment, auditing requirements, and post-market monitoring [41]. The US is taking a slower and more fragmented approach. Lawmakers have introduced bills that control the unwanted effects of AI algorithms from various aspects [42, 43]. Government offices and US regulatory agencies have also outlined their respective views and positions on regulating AI tools, but a comprehensive federal framework does not exist yet.

AI also has the potential to change both market and regulatory dynamics. Because of this, building the proper legal framework is necessary, not only to guard against individual and societal harm but also to ensure a level playing field for businesses. The current imbalance is putting considerable power into the hands of tech companies. To prevent market tipping, lawmakers may consider forcing tech companies in data-driven markets to share some of the data they collect [44] or explain how their algorithms work [45].

3 The importance of humanizing AI

It is not only the prospects of AI that need to be addressed but also the way it is currently used. Big data companies such as social media platforms are renowned for their ability to influence human behavior which has led to scandals and data privacy breaches (e.g., [46, 47]). Not addressing this will only make it more difficult for policymakers to make any significant changes to how big tech companies leverage human data to maximize gain. It is naive to continue to think of humans as superbeings able to fully control themselves in the face of increasingly sophisticated online persuasion and manipulation tactics. Equally concerning, is the way mechanistic algorithms (the application of narrow or weak AI) influence complex human behavior.

If AI had any kind of embodied representation today, it would have to be Mr. Spock from Star Trek. Governed by logic and capable of making rational decisions without being swayed by emotion, Mr. Spock would run the world without space for human error or irrational behavior, diminishing humanity to an artificial society governed by algorithms. A better representation of humankind would be Homer Simpson, limited in cognitive capacity, persuadable and irrational, but also caring and supportive. Homer Simpson would benefit greatly from Mr. Spock's characteristics if they didn't undermine human values.

Humanizing AI requires more than embodiment alone. We also need to consider the underpinning AI architecture paradigms that govern the machine-to-human interaction [48, 49]. These underpinnings help to understand how machines engage with their environment to make sense of the world and how to interact effectively (e.g. [50]). Bringing this back to the Mr. Spock metaphor, it reflects his ability to sense, plan, and interact with the environment to find the best possible solution. AI must not become a replacement for human cognition; rather it should be a tool to enhance human life.

4 Multilevel approach to humanizing AI

So far, we have discussed the reasons why we need to humanize AI. In this section, we will discuss how a behavioral lens can help us humanize it from inception to deployment. Considering only human-centric usages of AI is not enough. A multilevel approach is required, one which focuses on AI from creation all the way through to societal impact. First, designing AI to think like humans is one way of bringing humans and AI closer together. Can behavioral science help us design AI technology that has the capability to take human thought patterns into consideration in their functioning, that is, to embed knowledge of human thinking into the algorithms [51]? Second, applying a behavioral lens to consider more human-centric ways of serving people is needed. How do automation and AI usage facilitate human functioning,

and what about fairness and transparency? Finally, from a macro perspective, how can behavioral science facilitate a more positive and ethical impact of AI on society? We need to discuss the mechanisms underlying existing phenomena and behavior-informed strategies to humanize AI from the micro, meso and macro perspectives.

5 Humanizing AI from an algorithm perspective (micro)

Creating more human-like AI starts at the programming level. Like the micro perspective in behavioral science, understanding the software (brain) helps to understand the hardware (behavior). One of the main goals in AI development is to create intelligent machines that can understand, think, and act like human beings [52]. This type of AI is often referred to as strong AI or artificial general intelligence. Currently, AI's capabilities are narrow in scope (often referred to as weak AI or artificial narrow intelligence) being able to execute specific tasks like automation, surveillance, or autonomous driving.

Understanding machine intelligence and AI architecture is key to guiding more human-centric AI design. However, as AI computation becomes more 'complex' it gets harder to figure out how intelligent machines make decisions (e.g., [53]). To guide the evolution of AI operationalization in an explainable and responsible manner, various (micro-level) types of mechanisms need to be in place. These mechanisms i.e., audit trails (e.g. [54]), interpretability (e.g. [55]), and algorithmic design choices (e.g., [56]) can guide AI development and deployment into the future.

5.1 Anthropomorphism

It is well known that advancements made in machine learning algorithms are often anthropomorphized to represent human-like features [57]. Anthropomorphism is defined as the attribution of human-like traits to non-human objects, animals, and entities (e.g., [58]). Some researchers and businesspeople argue that for AI to become more integrated into human life or to enhance human properties, it needs to be more human-like [59].[1] AI researchers argue that to become more human-like, AI needs to represent human characteristics such as conversational abilities (e.g. [60]), using mental shortcuts to make decisions (e.g., [61]), being empathetic (e.g., [62]), or looking more human physically (e.g., [63, 64]). However, we should note the importance of delineating what human-like means for machines. With the term human-like we refer to the creation of behavioral similarities of humans in machines and do not mean the ontological definition of human-likeness (e.g. humans as conscious, experiencing, emotional beings).

The development of AI functionality and mechanisms has been cognitively inspired and modeled after the human brain (e.g., [65, 66]). For example, the design of Artificial Neural Networks (ANNs) is based on the way neurons in the brain process and exchange information. The development of Convolutional Neural Networks (CNNs) used in computer vision is based on how cats process visual information neurologically [67]. Besides traditional artificial intelligent approaches, Bioinspired Intelligent Algorithms (BIAs) also represent human (or living) organisms functioning at the micro level (e.g., [68]). BIAs show a strong underpinning in neuroscience and biological systems which are reflected in their working mechanisms. Genetic Algorithms (GA), Evolutionary Algorithms (EA), and Bee Colony Algorithms (BCA) are examples of BIAs. The benefit of using BIAs is that they are more explainable than traditional neural networks (e.g., [68–70]).

5.2 Anthropomorphic Algorithms

Some recent attempts to build more human-like AI at the micro level have been the creation of neural networks infused with decision science theory to develop anthropomorphic algorithms that use mental shortcuts to mimic human decision-making (e.g. [61]). Heuristics and mental shortcuts, often referred to as cognitive errors or limitations in human intelligence, do serve an evolutionary purpose [71]. They help to make quick decisions in difficult or uncertain situations while using limited information and cognitive resources [72].

The benefits of infusing algorithms with decision theory are that it helps machines think more like humans, makes faster decisions thanks to minimizing information requirements and computational power, and generates more accurate predictions in line with human cognition [73]. This could lead to a better user experience or higher customer satisfaction with product suggestions. Finally, it also addresses the issue of explainability as the built-in shortcuts make the decision rules applied in complex models transparent (e.g., [74]). The latter is a significant issue

[1] Others (e.g., [118, 119]) argue that AI should be "tethered to the humans who create and deploy them", but it should not be human-like.

of current modern AI, especially for ANNs. Conclusively, using behavioral theory to create more human-like AI not only helps address current limitations of existing AI, but can also provide pathways to more transparent operability and symbiotic human–machine design.

One of the major reasons for anthropomorphizing AI beyond application design is to consider the ethical foundations of anthropomorphic design (e.g., [75]). Bioethical principles are often used as the basis for developing ethical AI, e.g., respecting human rights and dignity, accountability, transparency, and promoting well-being [57]. These ethical considerations are seen as important viewpoints in the design and application of AI [76, 77]. In fact, anthropomorphizing AI has the potential to not only provide perspectives on finding effective ways of coexisting, but also provide a foundation for the ethical use of AI and how it can enhance human properties and life. For example, anthropomorphizing AI can support ethical considerations beyond existing bioethical principles, providing a broader perspective to the meaning of ethical use. [78] has questioned if it is unethical that AI applications or robots can make certain people (people in need of social connection i.e. elderly, mentally unwell) believe it is capable of building an emotional connection. Broadening the perspective of ethical use will become increasingly important as more human-like AI becomes available.

Advancing AI through an anthropomorphic lens is a discussion that we believe requires further attention as the human–machine relationship is not purely objective and rational, but is governed by experiences, emotions and heuristics.

6 Humanizing AI from an application perspective (meso)

In this section, we consider potential approaches to humanizing AI from an application perspective where the 'how' is emphasized more than the 'what'. Technology is always a means to an end and the way it is used depends on the intended purpose. Technology often emerges from a human-centric purpose motive (e.g., improving humanity, helping people to stay connected, making tasks easier). As time passes, the profit motive takes over and users become the target of exploitation, especially if investor metrics are involved in the further development of applications.

This development is prominent in consumer applications such as social media (e.g., Facebook), food delivery (e.g., Deliveroo), and ride sharing (e.g., Uber), which use algorithms to maximize profit margins and influence users and service providers [79]. Though issues relating to online exploitation and manipulation (e.g., [80]), psychological harm (e.g., [81]), data privacy (e.g., [82]), and misconduct (e.g., [83]) have been reported, little preventative action is being taken.

AI developments for business focus mainly on automation (e.g., process automation, robotics), smart solutions (e.g., just-in-time production, I-o-T smart buildings) and programs to help business managers make better decisions (e.g., talent management platforms, business intelligence systems). Businesses see the development and application of AI as a strategic imperative to create business value through increased efficiencies, revenue, or cost reductions (e.g., [84, 85]). However, there are also many company-focused AI solutions aimed at tracking and spying on employees. Whether commercial or business, the collection and use of personal data need to be governed and curated based on ethical considerations and human-enhancing properties. To achieve this, well-being metrics should be considered before investor metrics. This would fundamentally change future business-to-consumer (B2C) and business-to-business (B2B) application development. Other mechanisms which can be used to ensure more human centricity, accountability, and safety in application are audit trails (e.g. [54]), Responsible AI governance (e.g., [86]), and data bias and proportionality checks (e.g., [87]) among others.

Human-like AI application also needs to be considered within 'Industry 4.0' (I4.0), which is a term given to reflect the fourth industrial revolution currently taking place, characterized by big data, velocity and connectivity, cyber security systems, and embedded intelligence often referred to as smart technologies (e.g., [88]). The exponential growth of data and machine-to-machine and machine-to-human connectivity within I4.0 brings along various knowledge management and data interpretation challenges. Ontologies can provide a solution to bridge the complexity of data semantics and enable machine reasoning (e.g. [89, 90]). Within I4.0, machine intelligence needs to move away from narrow AI approaches and evolve more into cross-domain intelligence.

7 Humanizing AI from an organizational perspective

AI is reshaping the ways organizations operate. Bionic companies that combine human capabilities with machine intelligence, are no longer viewed as futuristic. According to a recent survey by McKinsey, the three business functions where AI adoption is most common are service operations, product and service development, and sales and marketing [91].

Technology complexity is not the biggest obstacle for large scale AI adoption – human nature is. Resistance to change and fear of the unknown are often quoted as key barriers to organizational adoption of AI, especially in core business areas where machines are used to perform complex cognitive functions [92]. While employing AI to automate highly manual processes (process automation) is gaining widespread acceptance, intelligent automation (decision automation) still needs to earn trust [93].

A recent research report by IBM argues that "trusted, explainable AI is crucial to widespread adoption of the technology", including maintaining brand integrity and meeting regulatory compliance [94]. The efforts behind explainable AI (XAI) aim to address the concerns related to lack of trust and seek to ensure that humans can easily understand the machine's decisions by providing solutions to black box issues and transparent explanations. But will these efforts be enough?

Research suggests that people supported by machines frequently trust the decisions of AI more than they should. This is even true for explainable AI solutions. Recently, a Google engineer believed that the company's DL-enabled chatbot 'LaMDA' had become sentient due to its human-like conversational abilities [95]. Research evidence indicates that, in many instances, people make worse decisions than they would without the assistance of machine intelligence [96, 97]. This raises the question of what is needed to humanize AI and make it more easily accepted by organizations.

First, AI tools have several advantages in performing certain customer-facing tasks quicker and more accurately than humans, especially when they can demonstrate high cognitive intelligence and empathetic behavior. These solutions, however, must be accepted and trusted by customers, and they need to deliver socially acceptable performance [98]. Our view is that once customers and society accept and trust AI tools, organizations will endorse them more readily.

Second, AI tools need to focus on more than explainability and ethics, eliminating unwanted bias in decision-making, and showing perceptible empathy. They also need to deliver the diversity of human decisions. If we let only a very few algorithms perform specific cognitive tasks, we might end up amplifying systemic risks such as flash-crash events on the stock markets due to high-frequency trading [99], or we might monopolize the core software engines behind cognitive AI tools [100].

Third, AI tools must convey to human users that their decision automation is subject to errors; not all automated decisions will be accurate. The higher the cognitive function, the more likely they can make mistakes, just like humans. We tend to have more trust in people who are like us in some way, and it is easier for us to predict the reactions of people who resemble us [101, 102]. This notion needs to be designed into AI tools and employee training during the operationalization phase if the AI solution is to be easily accepted by organizations and have the desired result.

Lastly, organizations will need to manage the risks that come with introducing intelligent machines. Operationalization of low cognitive AI solutions (such as process automation) poses fewer risks which can be reasonably mitigated by appropriate control [103]. The ultimate danger is that intelligent machines might seize control over their environment and refuse human control. Further research addressing this AI control problem and heavy-handed ex-ante regulation for highly cognitive AI tools will help to mitigate the risk that, as the late physicist Steven Hawkings put it, "The development of full artificial intelligence could spell the end of the human race" [104].[2]

8 Humanizing AI at the societal level (macro)

As mentioned before, AI as a technology has reached a level in which its usage has become ubiquitous. Although machine learning algorithms help us to process data efficiently, predict outcomes, and make intelligent decisions, it is the way these data analytic approaches are used that need to be scrutinized. In this section, we want to discuss some of the more precarious applications of AI that have (or continue to have) a deep impact on human behavior and society at large. Note that our review does not go into the details of subjective and objective measurement of human well-being,

[2] Undeniably, AI has its advantages and benefits, and there are views to support the argument that possible risks can be kept under control [118, 120, 121].

as defined by [105]; our three specific societal examples will be presented just to highlight the need for more human-like considerations of AI in society.

9 China's Social Credit System

In 2020, a countrywide social credit system (SCS) was implemented in China which scores citizens based on their offline and online behavior. Using AI-powered surveillance cameras, payment tracking through Alipay and other Chinese online payment methods, and social media surveillance, China's centralized SCS can evaluate a citizen's score and thus provide or restrict access to resources based on how well someone behaves (e.g., [106]). This application of AI monitors and socially-engineers behaviors and grants or denies access to public resources. From the perspective of humanizing AI, this approach is questionable as it monitors and coerces behavior in an opaque way, leaving it unclear for citizens how to influence their score.

SCSs are nothing new. Online platforms like Uber and Airbnb allow customers and service providers to evaluate the experience they had with each other. User-generated reviews act as a form of social validation and authority. Reviews not only affect the likeability and demand of the service but also impose a level of control over how service providers and customers interact with each other, based on expected behaviors governing the system (e.g. [107, 108]). However, the application of social ranking systems beyond consumer apps needs to be further reviewed.

9.1 Facebook segmentation algorithms

In recent months, Facebook and its operating platform have had significant backlash from the public because of how its algorithms promote hatred and discourage diverse views [109]. Facebook uses AI to classify and segment its users based on all kinds of characteristics ranging from behavioral profiles, social ties and even psychological traits it can infer using machine learning and pattern recognition. Facebook uses this data to enhance platform engagement and better serve its advertisers by offering more targeted advertising.

However, the unintended consequence of creating micro-segments is that users get stuck in echo chambers and are less exposed to different information and opinions. Studies have found that this kind of algorithmic design leads to more siloed thinking, reinforcing latent beliefs and potentially fueling more hatred [109].

These events make clear that AI-driven social media platforms have a significant impact on attitude formation and offline behaviors and thus are not neutral. Platforms that serve humankind need to be transparent in operations and designed with human well-being in mind. Recent EU regulation aims to force social media companies operating in the EU to disclose how their algorithms work. The unintended consequences of for-profit platforms need to be considered during the design stages and addressed later if necessary. Allowing the profit motive to take over the purpose motive is a key challenge to be addressed in the pursuit of a symbiotic relationship between humans and machines.

9.2 Cambridge Analytica Psychographic Profiling

The US presidential elections in 2016 were criticized because of the use of psychographic profiling online to sway voter decisions. The company responsible for facilitating the online profiling and micro-targeting of US citizens was Cambridge Analytica. It allegedly used personal data, improperly collected through a personality test on Facebook, which provided a rich user profile to identify attitudes, beliefs, and opinions. These data were then used to create micro-targeted ads aimed at influencing election participation and voter decisions [110].

Another consequence of this kind of profiling is that user data can be used to make highly accurate predictions about political beliefs, sexual orientations, fears and desires, and other sensitive information, which is impossible to predict using traditional survey approaches if not directly asked [111, 112]. This means that data companies with advanced machine learning algorithms know more about a user than relatives or close friends do [113]. If this information is used to influence attitudes and behavior for political reasons, then AI poses a risk to the social fabric of humanity and political integrity. Global institutions and countries need to address these applications and provide guidelines of ethical considerations in AI design and usage.

At the macro level, techno-social mechanisms need to be in place to ensure that the usage and advancement of AI are improving societal outcomes. This is necessary for technology adoption and trust in AI (e.g., [53]). Institutional mechanisms that govern AI development and deployment at the societal level focus mainly on protecting human values (e.g.,

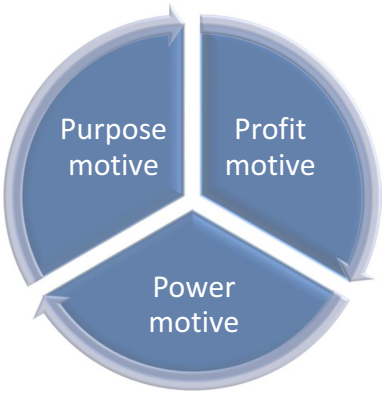

Fig. 1 Motivations driving AI creation and usage

equality, fairness, safety, and privacy – in [114]), accountability (e.g., holding AI developers accountable – in [115]), and incentives (e.g., funding or promoting AI-driven technologies that strengthen human values – in [116]). These mechanisms need to be in place to prevent both individual harm as well as societal harm (e.g. [40]).

10 The power motive of AI usage at the societal level

The three examples given highlight the power motive of AI usage at the societal level. The institutions that can install these mechanisms do it mainly to exert or expand their existing powers over the people they can control within their system. So far, the emphasis has been on governing and control (see Fig. 1 which represents our motivational framework of AI operationalization today - purpose, profit, and power). Using AI to socially engineer behavior, manipulate democratic voting decisions and segregate people based on similar characteristics, values, and/or beliefs does not benefit human well-being or promote equality. It enables "divide and power" asymmetries within society and highlights major concerns related to privacy, surveillance, discrimination and bias. It is also not unthinkable that with the ongoing datafication (digitization of all aspects of our daily lives and the evaluation of this data) that SCS in various forms could emerge in other parts of the world [106].

However, we can see potential avenues for human-enhancing and societally desirable usage of AI despite these examples of human-limiting approaches. Enabling societal development using AI-powered systems can allow people to flourish in more ways than are currently practiced. AI systems can augment human functioning and replace repetitive or dangerous tasks, allowing humans to focus more on strengthening qualities such as creativity, connection, altruism and emotional intelligence. AI and other enabling technologies can be used to help provide equal access to resources to facilitate human growth and well-being, emphasizing morals, fairness, ethics and even philosophy.

11 Conclusion and research directions

Modern-day AI development and deployment show great potential in creating value for business and society. However, the current state and the future of AI are not without concerns. Ethical and moral dilemmas have arisen in recent years due to AI usage in the public domain and the (un)intentional consequences algorithms have on economic choices and human well-being. Moreover, policymakers lack the speed and motivation to regulate the market with the required laws on time. Being too strict on AI policies can limit a country's digital competitiveness. It's clear that new perspectives are needed to help solve current issues and advance the field. In this paper, we argue that AI conceptualization and application need to be less artificial and more human-like. AI development and deployment need to be governed by more human-centric principles, ones that are easily understood by all stakeholders and that benefit society. We therefore propose a multilayered behavioral approach to address the issues and potential solutions.

This paper also reviewed the mechanisms underlying existing phenomena and behavior-informed strategies to humanize AI from the micro, meso and macro perspectives. In terms of mechanisms, we highlighted the importance of audit trails, interpretability, and algorithmic design choices at the micro level, Responsible AI governance, and data bias and proportionality checks at the meso level, and techno-social mechanisms such as protecting human values, accountability, and incentives at the macro level. For strategies, we proposed solutions which help build trusted and

explainable AI and support technology adoption, such as the development of anthropomorphic algorithms and other human-like features, making clear how algorithms make decisions, minimizing human and machine bias, and ensuring that the usage of AI augments and protects human life through incentivization and accountability.

Humanizing AI also means introducing ethical principles into the activities related to the planning, development and deployment of AI tools. The Responsible AI Guidelines, as developed by the US Department of Defense, provide detailed guidelines to ensure that ethical considerations are integrated into the design, training, and operationalization of AI, and define a process that is responsible, reproducible and scalable [117]. They go well beyond spelling out the need for explainability; they also aim to ensure that the decisions of AI tools are in line with human values. We believe that design research efforts should focus on investigating the core aspects of these themes, such as responsible AI, explainable AI and anthropomorphic design.

This paper provides various avenues for future research. First, from the micro perspective, future research should focus on exploring ways to build algorithms that are able to mimic human decision-making processes and make decisions in a more human-centric manner. This is not only important to make AI more understandable but is also an avenue to create intelligent systems with more general intelligence. Second, from the meso perspective, future research efforts should focus on finding ways to promote the application and adoption of more equitable, trusted, and responsible AI solutions to help overcome existing barriers and build a stronger relationship between humans and machines. Finally, from a macro perspective, future research efforts should focus on investigating and designing mechanisms that promote and secure human properties as the application of intelligent systems at the societal level becomes more common. The field of behavioral data science, where ML experts and behavioral scientists work together, has a valuable contribution to humanizing AI in future research efforts.

Addressing today's AI challenges is crucial if we want to build a more symbiotic relationship between humans and machines. Humanizing AI does not automatically lead to a more symbiotic relationship between humans and machines but does set a necessary foundation for its development based on human values and potential. This is not only important to build better AI, but also helps humankind to better understand what it means to be human in a digital world. Once again, we require wisdom to guide the future of AI.

Acknowledgements The authors would like to express sincere thanks to the reviewers for their valuable comments.

Author contributions Ali Fenwick and Gabor Molnar contributed equally to this paper.

Funding This project hasn't received any funding.

Data availability Not applicable.

Code availability Not applicable.

Declarations

Competing interests The authors declare no competing interests. The views and opinions expressed in this paper are those of the authors and do not necessarily reflect the views or positions of any entities they represent.

References

1. Fron C, Korn O. A short history of the perception of robots and automata from antiquity to modern times. In: Social robots: technological, societal and ethical aspects of human-robot interaction. Cham: Springer International Publishing; 2019. p. 1–12.
2. Devecka M. Did the Greeks believe in their robots? Camb Class J. 2013;59:52–69.
3. Homer. The Iliad. New York: Penguin Publishing Group; 1991.
4. Shelley MW. Frankenstein; or, the modern Prometheus. London: Printed for Lackington, Hughes Harding, Mavor & Jones; 1818.

5. Aristotle. The Rhetoric of Aristotle: an expanded translation with supplementary examples for students of composition and public speaking. New York: D. Appleton and Co; 1932.
6. Russell S, Davis E, Norvig P. Artificial intelligence: a modern approach. Hoboken: Prentice Hall; 2009.
7. Afiouni R. Organizational Learning in the Rise of Machine Learning. International Conference on Information Systems, Munich. 2019.
8. Lee J, Suh T, Roy D, Baucus M. Emerging technology and business model innovation: the case of artificial intelligence. J Open Innov. 2019;5(3):1–13.
9. Mikalef P, Gupta M. Artificial intelligence capability: conceptualization, measurement calibration, and empirical study on its impact on organizational creativity and firm performance. Inf Manag. 2021;58(3):1–20.
10. R. Schmidt, A. Zimmermann, M. Möhring and B. Keller, "Value Creation in Connectionist Artificial Intelligence - A Research Agenda," in *AMCIS*, 2020.
11. Simon HA. The sciences of the artificial. Cambridge: MIT; 1970.
12. Russel S, Norvig P. Artificial intelligence: a modern approach. London: Pearson; 2016.
13. Wang P. On defining artificial intelligence. J Artif Gen Intell. 2019;10(2):1–37.
14. Kühl N, Goutier M, Hirt R, Satzger G. Machine Learning in Artificial Intelligence: Towards a Common Understanding. https://arxiv.org/abs/2004.04686. 2020.
15. Du X, Dua S. Data mining and machine learning in cybersecurity. Abingdon-on-Thames: Taylor & Francis; 2011.
16. Bishop CM. Pattern recognition and machine learning. New York: Springer; 2006.
17. Serrano W. Big data intelligent search assistant based on the random neural network., advances in big data: proceedings of the 2nd INNS conference on big data. Thessaloniki: Springer International Publishing; 2016.
18. Chen Y. Integrated and intelligent manufacturing: perspectives and enablers. Engineering. 2017;3(5):588–95.
19. Liu H-Y, Zawieska K. From responsible robotics towards a human rights regime oriented to the challenges of robotics and artificial intelligence. Ethics Inf Technol. 2017;22:321–33.
20. Jobin A, Ienca M, Vayena E. The global landscape of AI ethics guidelines. Nat Mach Intell. 2019;1(9):389–99.
21. Ryan M, Stahl BC. Artificial intelligence ethics guidelines for developers and users: clarifying their content and normative implications. J Inf Commun Ethics Soc. 2021;19(1):61–86.
22. Pew Research Center. Artificial Intelligence and the Future of Humans, 2018.
23. Han S, Kelly E, Nikou S, Svee E-O. Aligning artificial intelligence with human values: reflections from a phenomenological perspective. AI Soc. 2021. https://doi.org/10.1007/s00146-021-01247-4.
24. Hollnagel E, Woods DD. Joint cognitive systems: foundations of cognitive systems engineering. Milton Park: Taylor & Francis Group; 2005.
25. Norman DA. The Design of Everyday Things, Revised and. expanded. Cambridge: MIT Press; 2013.
26. Bødker S. Third-wave HCI, 10 years later—participation and sharing. Interactions. 2015;22(5):24–31.
27. Saariluoma P, Oulasvirta A. User psychology: re-assessing the boundaries of a discipline. Sci Res. 2010;1(5):317–28.
28. Saariluoma P, Cañas J, Leikas J. Designing for Life. London: MacMillan; 2016.
29. ISO, 9241 - Ergonomics of human-system interaction—Part 210: Human-centred design for interactive systems, ISO, 2019.
30. Miyake N, Ishiguro H, Dautenhahn K, Nomura T. Robots with children: practices for human-robot symbiosis. IEEE: Piscataway; 2011.
31. Sandini V, Mohan, Sciutti A, Morasso P. Social cognition for human-robot symbiosis—challenges and building blocks. Front Neurorobotics. 2018;12:34.
32. Fabi S, Xu X, de Sa VR. Exploring the racial bias in pain detection with a computer vision model. 2022. https://cogsci.ucsd.edu/~desa/Exploring_the_Racial_Bias_in_Pain_Detection_with_a_Computer_Vision_Model.pdf. Accessed 15 May 2022
33. Daugherty PR, Wilson J, Chowdhury R. Using Artificial Intelligence to promote diversity. Boston: MIT Sloan Management Review; 2018.
34. Kiritchenko S, Mohammad SM. Examining gender and race bias in two hundred sentiment analysis systems. arXiv. 2018. https://doi.org/10.48550/arXiv.1805.04508.
35. Lockey S, Gillespie N, Holm D, A. Someh IA. A Review of Trust in Artificial Intelligence: Challenges, Vulnerabilities and Future Directions. Proceedings of the 54th Hawaii International Conference on System Sciences. 2021.
36. Suresh H, Guttag JV. A Framework for Understanding Unintended Consequences of Machine Learning. arXiv. 2020;2:8.
37. IEEE. P7001 - Draft standard for transparency of autonomous systems. New York: IEEE; 2020. p. 1–70.
38. IEEE. P7007 - Ontological Standard for Ethically Driven Robotics and Automation Systems. Newyork: IEEE; 2021.
39. Acemoglu D. Harms of AI. Natl Bureau Econ Res. 2021. https://doi.org/10.3386/w29247.
40. Smuha NA. Beyond the individual: governing AI's societal harm. Int Policy Rev. 2021. https://doi.org/10.14763/2021.3.1574.
41. European Commission, Proposal for a regulation of the European parliament and of the council laying down harmonised rules on Artificial Intelligence (artificial intelligence Act) and amending certain union legislative Acts, 2021.
42. United States Congress (117th), H.R.2154—Protecting Americans from Dangerous Algorithms Act, 2021.
43. United States Congress (117th), S.1896—Algorithmic Justice and Online Platform Transparency Act, 2021.
44. Graef I, Prüfer J. Governance of data sharing: a law & economics proposal. Res Policy. 2021;50(9):104330.
45. Fu G. CDA Insights 2022: Toward ethical artificial intelligence in international development. 2022. https://dai-global-digital.com/cda-insights-2022-toward-ethical-artificial-intelligence-in-international-development.html. Accessed on 23 May 2022.
46. Schlackl F, Link N, Hoehle H. Antecedents and consequences of data breaches: a systematic review. Inform Manag. 2022;59:103638.
47. Dembrow B. Investing in human futures: how big tech and social media giants abuse privacy and manipulate consumerism. U MIA Bus L Rev. 2022;30(3):324–49.
48. Bayat B, Bermejo-Alonso J, Carbonera J, Facchinetti T. Requirements for building an ontology for autonomous robots. Industrial Robot. 2016;43:469–80.
49. Coste-Maniere E, Simmons R. Architecture, the backbone of robotic systems. EEE International Conference on Robotics and Automation. Symposia Proceedings, San Francisco, CA. 2000.
50. J. Calzado, A. Lindsay, C. Chen, G. Samuels and J. I. Olszewska, "SAMI: Interactive, Multi-Sense Robot Architecture," in IEEE 22nd International Conference on Intelligent Engineering Systems (INES), Las Palmas de Gran Canaria, 2018.
51. Oulasvirta A. It's time to rediscover HCI models. Interactions. 2019;26(4):52–6.

52. Bostrom N. Superintelligence: paths. Dangers: Strategies, Brilliance Publishing; 2015.
53. Samek W, Müller KR. 2019. Explainable AI: interpreting, explaining and visualizing deep learning. Towards explainable artificial intelligence. Springer. pp. 5–22.
54. Falco G, Shneiderman B, Badger J, Carrier R, Dahbura A. Governing AI safety through independent audits. Nature Mach Intell. 2021;3:566–71.
55. Burkhardt R, Hohn N, Wigley C. Leading your organization to responsible AI. https://www.mckinsey.com/business-functions/quantumblack/our-insights/leading-your-organization-to-responsible-ai. Accessed 14 Jun 2022
56. Amoore L, Raley R. Securing with algorithms. Secur Dialogue. 2017;48(1):3–10.
57. Salles A, Evers K, Farisco M. Anthropomorphism in AI. AJOB Neurosci. 2020;11(2):88–95.
58. Epley N, Waytz A, Cacioppo JT. On seeing human: a three-factor theory of anthropomorphism. Psychol Rev. 2007;114(4):864–86.
59. Bar-Cohen Y, Hanson D. The coming robot revolution: expectations and fears about emerging intelligent, humanlike machines. New York: Springer; 2016.
60. Araujo T. Living up to the chatbot hype: the influence of anthropomorphic design cues and communicative agency framing on conversational agent and company perceptions. Comput Hum Behav. 2018;85(1):183–9.
61. Fabi S, Hagendorff T. Why we need biased AI. How including cognitive and ethical machine biases can enhance AI systems. arXiv. 2022. https://doi.org/10.48550/arXiv.2203.09911.
62. Airenti G. The cognitive bases of anthropomorphism: from relatedness to empathy. Int J Soc Robot. 2015;7(1):117–27.
63. Leong B, Selinger E. Robot Eyes Wide Shut: Understanding Dishonest Anthropomorphism. Proceedings of the Association for Computing Machinery's Conference on Fairness, Accountability, and Transparency, Atlanta, GA, 2018.
64. G. Marcus, *Deep Learning. A Critical Appraisal,* arXiv, 2018.
65. Ullman S. Using neuroscience to develop artificial intelligence. Science. 2019;363(6428):692–3.
66. Eysenck MW, Eysenck C. AI vs Humans. London: Taylor & Francis Group; 2021.
67. Nagi J, Ducatelle F, Di Caro GA, Cireşan D, Meier U, Giusti A, Nagi F, Schmidhuber J, Gambardella LM. Max-pooling convolutional neural networks for vision-based hand gesture recognition. New York: IEEE; 2011. p. 342–7.
68. Ni J, Wu L, Fan X, Yang S. Bioinspired intelligent algorithm and its applications for mobile robot control: a survey. Comput Intell Neurosci. 2016;2016:1–16.
69. Binitha SD, Sathya SS. A survey of Bio inspired optimization algorithms. Int J Soft Comput Eng. 2012;2:2.
70. Olszewska JI. Snakes in trees: an explainable artificial intelligence approach for automatic object detection and recognition. ICAART; 2022.
71. Tversky A, Kahneman D. Judgment under uncertainty: heuristics and biases. Science. 1974;185(4157):1124–31.
72. Klein G. Naturalistic decision making. Hum Factors J Hum Factors Ergonomics Soc. 2008;50(3):456–60.
73. Gadzinski G, Castello A. Fast and frugal heuristics augmented: when machine learning quantifies Bayesian uncertainty. J Behav Exp Finance. 2020;26:100293.
74. Hafenbrädl S, Waeger D, Marewski JN, Gigerenzer G. Applied decision making with fast-and-frugal heuristics. J Appl Res Mem Cogn. 2016;5(2):215–31.
75. Damiano L, Dumouchel P. Anthropomorphism in human-robot co-evolution. Front Psychol. 2018. https://doi.org/10.3389/fpsyg.2018.00468.
76. Mittelstadt B. Principles alone cannot guarantee ethical AI. Nat Mach Intell. 2019;1(11):501–7.
77. V. Vakkuri, K.-K. Kemell and P. Amrahamsson. Implementing ethics in AI: initial results of an industrial multiple case study. Product-Focused Software Process Improvement. PROFES 2019. Lecture Notes in Computer Science. Cham 2019.
78. Coeckelbergh M. Can we trust robots? Ethics Inf Technol. 2012;14(1):53–60.
79. Wu T. The Attention merchants: the epic struggle to get inside our heads. London: Atlantic Books; 2017.
80. Susser D, Roessler B, Nissenbaum H. Online manipulation: hidden influences in a digital world. Georgetown Law Technol Rev. 2019;4(1):1–45.
81. Amedie J. The Impact of Social Media on Society. 2015. https://scholarcommons.scu.edu/engl_176/2. Accessed 26 May 2022
82. Sushama C, Kumar MS, Neelima P. Privacy and security issues in the future: a social media. Mater Today. 2021. https://doi.org/10.1016/j.matpr.2020.11.105.
83. Bakir V, McStay A. Fake news and the economy of emotions. Digit J. 2018;6(2):154–75.
84. Alsheibani SA, Messom CH, Cheung YP, Alhosni M. Reimagining the Strategic Management of Artificial Intelligence: Five Recommendations for Business leaders in AMCIS. 2020.
85. Amer-Yahia S, Roy SB, Chen L, Morishima A, Monedero J. Making AI machines work for humans in FoW. ACM Sigmod Record. 2020;49:30–5.
86. E. Papagiannidis, I. M. Enholm, P. Mikalef and J. Krogstie. Structuring AI Resources to Build an AI Capability: a Conceptual Framework. ECIS. 2021.
87. Arrieta AB, Díaz-Rodríguez N, Ser JD, Bennetot A, Tabik. Explainable Artificial Intelligence (XAI): concepts, taxonomies, opportunities and challenges toward responsible AI. Inform Fusion. 2020;58:82–115.
88. Sampath K, Khamis A, Fiorini S, Carbonera J, Olivares Alarcos A. Ontologies for industry 4.0. Knowl Eng Rev. 2019;34:E17.
89. A. Hassani, A. Medvedev, P. D. Haghighi, S. Ling, M. Indrawan-Santiago, A. Zavlavsky and P. P. Jayaraman. Context-as-a-Service Platform: exchange and share context in an IoT ecosystem. IEEE International Conference on Pervasive Computing and Communications Workshops. 2018.
90. Olszewska JI, Allison AK. ODYSSEY: Software development life cycle ontology. Proceedings of the International Conference on Knowledge Engineering and Ontology Development. 2018.
91. Chui M, Hall B., Singla, Sukharevsky A. Global survey: the state of AI in 2021. https://www.mckinsey.com/business-functions/mckinsey-analytics/our-insights/global-survey-the-state-of-ai-in-2021. Accessed 7 Feb 2022
92. Goasduff L. 3 Barriers to AI Adoption. 2019. https://www.gartner.com/smarterwithgartner/3-barriers-to-ai-adoption. Accessed 7 Feb 2022

93. Coombs C, Hislop D, Taneva SK, Barnard S. The strategic impacts of Intelligent Automation for knowledge and service work: an interdisciplinary review. J Strateg Inform Syst. 2020;29:4.
94. Watson IBM. 2021. Global AI Adoption Index 2021. https://newsroom.ibm.com/IBMs-Global-AI-Adoption-Index-2021. Accessed 8 Feb 2022
95. Fenwick A, Caneri M, Ma S, Chung-Pang TS, Jimenez MA, Calzone O, López-Ausens T, Ananías C. 2022. Sentient or illusion: what LaMDA teaches us about being human when engaging with AI. MIT Technology Review Arabia (Arabic). https://drfenwick.medium.com/sentient-or-illusion-what-lamda-teaches-us-about-being-human-when-engaging-with-ai-39b9237b49d8. Accessed 26 Jun 2022.
96. Bansal G, Wu T, Zhou J, Fok R, Nushi B, Kamar E, Ribeiro MT, Weld D. Does the Whole Exceed its Parts? The Effect of AI Explanations on Complementary Team Performance. CHI '21: Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems. pp. 1–16, 2021.
97. Buçinca Z, Lin P, Gajos KZ. Glassman EL. Proxy tasks and subjective measures can be misleading in evaluating explainable AI systems. IUI '20: Proceedings of the 25th International Conference on Intelligent User Interfaces. pp. 454–464, March 2020.
98. Pelau C, Dabija D-C, Ene I. What makes an AI device human-like? The role of interaction quality, empathy and perceived psychological anthropomorphic characteristics in the acceptance of artificial intelligence in the service industry. Comput Hum Behav. 2021. https://doi.org/10.1016/j.chb.2021.106855.
99. Kirilenko A, Kyle AS, Samadi M, Tuzun T. The flash crash: high-frequency trading in an electronic market. J Financ. 2017;72(3):967–98.
100. Hindman M. The internet trap: how the digital economy builds monopolies and undermines democracy. Princeton: Princeton University Press; 2018.
101. DeBruine LM. Facial resemblance enhances trust. Proc Royal Soc Biol Sci. 2002;269:1498.
102. Kramer RM. Rethinking trust. Harv Bus Rev. 2009;87(6):68–77.
103. B. Bhatti. 7 Types of AI Risk and How to Mitigate their Impact. https://towardsdatascience.com/7-types-of-ai-risk-and-how-to-mitigate-their-impact-36c086bfd732. Accessed 13 Sept 2020
104. R. Cellan-Jones, "Stephen Hawking warns artificial intelligence could end mankind," 2 December 2014. . Available: https://www.bbc.com/news/technology-30290540. [Accessed 8 February 2022].
105. IEEE, 7010 Recommended Practice for Assessing the Impact of Autonomous and Intelligent Systems on Human Well-Being, New York, NY: IEEE, 2020.
106. A. Cheung and Y. Chen, "From Datafication to Data State: Making Sense of China's Social Credit System and Its Implications," *Law & Social Inquiry,* pp. 1–35, 2021.
107. S. Feldstein. The Global Expansion of AI Surveillance. 2019. https://carnegieendowment.org/2019/09/17/global-expansion-of-ai-surveillance-pub-79847. Accessed 14 Jun 2022
108. A. Fenwick. How's your social credit score? 2018. https://www.hult.edu/blog/your-social-credit-score/ Accessed 26 Jun 2022
109. Flaxman S, Goel S, Rao JM. Filter bubbles, echo chambers, and online news consumption. Public Opin Quart. 2016;80:298–320.
110. Bastos MT, Mercea D. The Brexit botnet and user-generated hyperpartisan news. Soc Sci Comput Rev. 2019;37(1):38–54.
111. Kosinski M, Stillwell D, Graepel T. Private traits and attributes are predictable from digital records of human behavior. Proc Natl Acad Sci. 2013;110(15):5802–5.
112. Kosinski M, Bachrach Y, Kohli P, Stillwell D, Graepel T. Manifestations of user personality in website choice and behaviour on online social networks. Mach Learn. 2014;95(3):357–80.
113. Youyou W, Kosinski M, Stillwell D. Computer-based personality judgments are more accurate than those made by humans. Proc Nat Acad Sci. 2015;112(4):1036–40.
114. European Commission. White paper on artificial intelligence: a European approach to excellence and trust. 2020. https://ec.europa.eu/info/sites/info/files/commission-white-paper-artificial-intelligence-feb2020_en.pdf. Accessed 26 Jun 2022
115. M. Wieringa, "What to account for when accounting for algorithms. A systematic literature review on algorithmic accountability," in *Proceedings of the 2020 conference on Fairness, Accountability, and Transparency*, 2020.
116. The White House. Artificial Intelligence, Automation, and the Economy. 2016. https://obamawhitehouse.archives.gov/sites/whitehouse.gov/files/documents/Artificial-Intelligence-Automation-Economy.PDF. Accessed 26 Jun 2022
117. J. Dunnmon, B. Goodman, P. Kirechu, C. Smith and A. V. Deusen, Responsible AI Guidelines in Practice, Defense Innovation Unit, US Department of Defense, 2021.
118. I. Kostopoulos. Decoupling Human Characteristics from Algorithmic Capabilities. The IEEE Standards Association, 2014. https://standards.ieee.org/initiatives/artificial-intelligence-systems/decoupling-human-characteristics/. Accessed 11 Jun 2022
119. Johnson DG, Miller KW. Un-making artificial moral agents. Ethics Inf Technol. 2008;10(2):123–33.
120. Stahl BC. Ethical issues of AI. Artificial intelligence for a better future springer briefs in research and innovation governance. Cham: Springer; 2021.
121. Saariluoma P, Rauterberg M. Turing's Error-revised. International Journal of Philosophy Study. 2016;4:22–41.