Journal of Responsible Technology 17 (2024) 100075

Contents lists available at ScienceDirect

Journal of Responsible Technology

journal homepage: www.sciencedirect.com/journal/journal-of-responsible-technology

Research Article

Principles of digital humanism: A critical post-humanist view

Erich Prem

The University of Vienna, Institute of Philosophy, Philosophy of Media and Technology, Universitätsstraße 7, Vienna A-1010, Austria

ARTICLE INFO

Keywords:
Digital humanism
Principles
Computer ethics
AI ethics
Digital sovereignty
Humanism

ABSTRACT

Digital humanism emerges from serious concerns about the way in which digitisation develops, its impact on society and on humans. While its motivation is clear and broadly accepted, it is still an emerging field that does not yet have a universally accepted definition. Also, it is not always clear how to differentiate digital humanism from other similar endeavours. In this article, we critically investigate the notion of digital humanism and present its main principles as shared by its key proponents. These principles include the quest for human dignity and the ideal of a better society based on core values of the Enlightenment.

The paper concludes that digital humanism is to be treated as a technical endeavour to shape digital technologies and use them for digital innovation, a political endeavour investigating power shifts triggered by digital technology, and, at the same time, as a philosophical endeavour including the quest to delineate its scope and to draw boundaries for the digital.

Methodologically, digital humanism is an interdisciplinary effort to debate a broad range of digitisation shortfalls in their totality, from privacy infringements to power shifts, from human alienation to disownment. While it overlaps with a range of established fields and other movements, digital humanism reflects a new academic, engineering, and societal awareness of the challenges of digital technologies.

1. Introduction

1.1. Aims and methodology

Since its inception, digital humanism[1] has received substantial praise for its objectives and for bringing together various fields in their efforts of working on a better digital future. It has been included in political programmes, was taken up as a topic by funding agencies,[1] and led to multinational political agreements.[2] For some, digital humanism seems a contradiction in terms, i.e. an oxymoron that puts the two opposites of the human and the non-human in one phrase to reveal a paradox. For others, it is an aspiration joining forces for a vision of a better digital technology. Others again, regard it a political movement and intellectual initiative that is still very much evolving today. The objective of this article therefore is to shed light on digital humanism as it originated following the *Vienna Manifesto on Digital Humanism* (Werthner, 2020; Werthner et al., 2022a p. xi-xiii), to clarify its status and reach, and to delineate it with some precision from other fields of scientific, engineering, and scholarly inquiry. I shall pursue this goal as objectively as possible, but also as a participant in the digital humanism discussions and meetings, and with the caveat of a preliminary assessment because digital humanism is a young movement and still developing.

There is considerable demand for a clarification of the name, the concepts, and directions of digital humanism. Humanism is already a contested concept (Gallie, 1955), and digital humanism even more so. A now traditional criticism is directed against humanism as a predominantly male, white, and colonial effort that is overly anthropocentric and lacking inclusivity. From Adorno and Horkheimer's criticism of the Enlightenment (Adorno & Horkheimer, 1944; Nida-Rümelin & Winter, 2024) to Foucault, humanism has been criticized for being overly focused on humans and their rights (Foucault, 1974) disregarding both animals and the environment.[3] Furthermore, despite a shared interest in human rights and a humanistic education, the notion of digital humanism may not always necessarily resonate well with contemporary

E-mail addresses: erich.prem@univie.ac.at, prem@eutema.com.

[1] The European Union called for a project on digital humanism in its 2023-24 work programme 'Digital, Industry, and Space'. Digital humanism is listed as a topic in the Austrian coalition agreement of 2020. The Vienna Science and Technology Fund launched its first call on Digital Humanism in 2021.

[2] The 'Poysdorf Declaration on Digital Humanism' was signed by the Foreign Ministers of Austria, the Czech Republic, and Slovakia on June 30, 2021. https://www.bmeia.gv.at/en/the-ministry/press/news/2021/06/austrias-chairmanship-of-the-slavkovausterlitz-format-ends-with-the-signing-of-the-poysdorf-declaration-on-digital-humanism (last accessed 23-11-23).

[3] Cf. Johnson (2009) for a differentiated position.

https://doi.org/10.1016/j.jrt.2024.100075

Available online 20 January 2024

humanists who usually focus on what it means to be human and, hence, on the non-digital.

This paper starts with analysing the origins and motivation of digital humanism (Section 1.2) such as the Vienna manifesto and a broad range of critical issues connected to digital technologies. It then describes the main interests of its proponents (Section 1.3) and lists current topics in digital humanism from automation of work to privacy, algorithmic decision making, surveillance, AI ethics, technology monopolies, threats to democracy, state sovereignty to geopolitics. As will be explained, this breadth of topics is a challenge, but also an opportunity for digital humanism.

Section 2 explores the relation of digital humanism with traditional humanism (Section 2.1) and explains its connection to historical humanism, the human-centricity or *homo mensura* principle, and a possible focus on care or *humanitas*. It proceeds to briefly position digital humanism with respect to the Age of Enlightenment (Section 2.2). Section summarises the ideas, activities, and discussions of digital humanism with the help of five principles that address the impact of the digital on people, the aim to protect people and the environment, the call to use digital technology for strengthening democracy, the affirmation that digital technology can be shaped and, finally, that humans and machines are fundamentally different. Section 4 discusses the epistemology of digital humanism as a technical field, a political idea, and a philosophy. It shows how digital humanism aims to bring a broad range of disciplines together including computer science, law, social science, and ethics. Here, the paper also differentiates it from other disciplines that debate the relation of humans in the digital realm (Section 4.3). The paper concludes with a discussion section that assesses digital humanism as a type of *critical post-humanism* that addresses important criticisms of traditional humanism. The paper lists open issues that will need more attention and pointers for future research in Section 5.

The sources for this analysis include the *Vienna Manifesto on Digital Humanism*, the anthologies by Werthner et al. (2022b) on the subject, publications referencing the Vienna version of digital humanism and the Vienna manifesto, the online lecture series, workshops, debates, conferences, the digital humanism roadmap (Prem et al., 2022), and selected authors who have referenced Digital Humanism in their work.

1.2. Origin, manifesto, and initiative

Digital humanism emerges from serious concerns about the way in which the digitisation of societies progresses and how digital technologies develop as well as from the impact of this development on humans and on society. While its motivation is clear and broadly accepted, it is still an emerging field that does not yet have a precise definition. Werthner et al. (2024) call for a digital humanism that "describes, analyses, and influences the complex interplay between technology and mankind with the aim of a better society and a better life while fully respecting universal human rights." As will be explained in more detail below, this should be interpreted as an interest in the understanding of the development of digitization and as the aim to influence it both technically and societally, i.e. politically.

As a name for the initiative and movement, *digital humanism* was coined in May 2019 following an international conference at the Vienna University of Technology that resulted in the drafting of the Vienna Manifesto as a "call to deliberate and act on technological development."[4] The manifesto lists 31 academic authors from universities in Austria, Italy, the Netherlands, Switzerland, Germany, and USA. After its publication, the Vienna Manifesto on Digital Humanism became a position statement signed by more than thousand experts and organisations worldwide. It lays out the motivation and goals for the Digital Humanism Initiative and was translated into eight languages by the end of 2023. It is easy to see how the manifesto aims at a balanced view of the opportunities of digital progress while clearly identifying important problems that motivated it in the first place. Its central starting point is the diagnosis of serious concerns about digitalization from monopolization to digital surveillance. Key observations of the current state-of-affairs of digital technology are:

- A diagnosed co-evolution of technology and humankind,
- a new "quest for enlightenment and humanism" (manifesto),
- and a call to influence technological development.

In addition, the manifesto lists eleven core principles for the development of a better digital future. These principles start by emphasizing democracy and democratic values (e.g. inclusion, freedom of speech or expression, privacy etc.) arguing that there is a need for stricter regulation in the digital sphere and of technology monopolies and platforms based on a broad academic and public discourse. The principles include a call for computer-supported decisions rather than entirely machine-made decisions whenever people are affected. The principles argue for an interdisciplinary approach and a strong role of universities to create the necessary knowledge and the engagement of researchers with wider society. The manifesto calls upon practitioners to take responsibility for the impact of technologies they develop. Its principles also include a call for new curricula and improved education in computer science that starts as early as possible to combine technical skills with ethical and social awareness. The manifesto has an academic and political undertone with its focus on debate, knowledge, on the roles of universities and with its call to actively shape technology development, i.e. technology policy.

1.3. Drivers and interests

Four years into its existence, the Vienna Manifesto still offers a valid description of the motivation and interests of digital humanism and summarizes many of the interests of its proponents. The following table (Table 1) provides a list of current topics of interest to proponents of the field grouped along four focal areas of interest: the human, our society, the economy, and the state.

Naturally, some topics are relevant in more than one area. Digital humanists debate automation at the individual human level and also discuss its economic and societal impact. Note that some topics are in fact core concepts of several disciplines, e.g. privacy is a concept that is often considered central to legal debates, but it is also a research field in computer science (e.g. in privacy-preserving machine learning) and debated in philosophy. Similarly, automation is both a technical endeavour, a concern in work science, in economics etc.

These topics are broad, and the issues and concerns are many which poses the question as to whether it is reasonable to discuss all these various problems within a single large initiative. As I will argue in more

Table 1
Overview of current topics in digital humanism grouped into four areas of interest.

Human	Society	Economy	State
• Automation of work	• Surveillance	• Platform economics	• Democracy
• Human identity	• Technology ethics (esp. AI)	• Technology monopolies	• Digital politics and regulation
• Human dignity	• Online discourse, fake news	• Regulation of technology and tech companies	• Sovereignty
• Privacy	• Freedom of speech	• Consumer rights	• Geopolitics
• Algorithmic control and decisions	• Human rights		
• Education	• Resilience		

[4] Earlier mentions of the term date back to Douhei (2011) and an exhibition at the Vienna Museum for Applied Arts (Vienna Biennale Guide, 2017).

detail below, digital humanism aims at considering many of these aspects in conjunction if not in their totality. This is a central point of difference to other initiatives that are critical of the state and development of digital technology such as those focused on privacy and surveillance, online censorship and freedom of speech, or democracy, consumer rights and monopolist practices, digital sovereignty and democracy, and many others. Digital humanists argue that these issues are interlinked and therefore require joint attention. Before returning to this point in Section 4 below, let us investigate just precisely what type of humanism its digital version is and how its objectives relate to those that are usually associated with the Age of Enlightenment.

2. Digital humanism and enlightenment

2.1. Humanism

Digital humanism is clearly embedded in a tradition of Enlightenment and humanism both of which the manifesto also describes as part of its objectives. In fact, the manifesto claims that "[*t*]he quest is for enlightenment and humanism." However, linking humanistic ideals with critical thoughts about technological progress, as demanded in the manifesto, is not a completely straightforward endeavour given that these ideals are not fully detailed. There are at least four important human-centric concepts that can be linked to digital humanism, namely (i) historical and contemporary humanism, (ii) *homo mensura* (the human as the measure), (iii) *humanitas* (kindness) and (iv) human rights (discussed in the next Section).

(i) Humanism is a term used for a rather diverse range of historical developments that some trace back to antiquity, others root it predominantly in Italian renaissance, while some consider it as resulting from a mostly German pedagogical tradition. This is unsurprising as humanism has developed over the centuries into many different forms and with varying intentions. Milad Douhei first used the concept of a digital humanism (2011), in the French original "humanisme numérique". Douhei builds on the Lévi-Strauss (1973) classification of the humanist movement into three phases of (i) an aristocratic Renaissance humanism, (ii) a bourgeois and exotic humanism of the 19th century, and (iii) a 20th century democratic humanism which he extends with the contemporary 'digital humanism'.

A central interest of humanism (and a key concern of anthropology) throughout the ages has been the nature of the human being. It poses the question of the *specificum humanum*, of what is special and unique about humans and what therefore results as the *conditio humana*, the human condition, cf. Schmölz (2020). The specification of what makes us human has changed over time and with different authors and was predominantly undertaken as an essentialist endeavour, i.e. looking for the essence or a list of features that characterize the humankind. It includes *rational thinking* as a characteristic of key importance in antiquity, e.g. in the work of Aristotle, the capacity of *language* as an important element in religious and philosophical thinking, for example of Thomas Aquinas, Johann Gottfried Herder or Ernst Cassirer. More modern conceptions have emphasized human *creativity* or *consciousness* and the resulting responsibility of humans for their actions. There is a clear opposition to viewing humans as animals (or as machines) and an appreciation of human autonomy. It is no coincidence that key characteristics such as autonomy, creativity, language, rationality, and consciousness often appear in contemporary debates of digital humanism, e.g. in questions about artificial intelligence, cf. Nida-Rümelin (2022).

(i) The way in which digital humanism makes humankind the measure of all things digital is a reminder of the homo mensura principle of Protagoras which for Hegel marks the point in time from when everything in philosophy revolves about the human. From here, the human becomes the way to measure and compare everything. In the digital realm the human is often the measure of digital technology, but not in ways that digital humanists consider appropriate. Many digital systems are designed with a sole interest in the user as a means of commercial success. This in many cases also limits the perspective and design focus to the single user. It is quite the opposite of putting humankind at the centre of attention as it reduces the relation of a digital system to a relationship with the user's interaction for the sake of profit maximization. This leaves out all societal aspects in the user interaction – even when the interface is a social network the interaction focuses on the relation of the system to the individual. In a sense, society disappears at the interface, and it is therefore difficult to regain a social or societal dimension at later stages of the technology development. While there is a debate about the homo mensura principles of Protagoras in whether it concerns the individual human, humankind, or a group of humans, digital humanism is rather clear about the many problems of a type of computer science that is overly focused on individualism (Werthner et al., 2022a, p. vii-viii). This also describes the contrast between a notion of freedom that includes a collective dimension and a purely neo-liberal and economic notion reduced to just individual freedom and limited state power (Akkermans et al., 2022, p. 58).

(ii) A third concept that is perhaps better suited for digital humanism than classical humanism with its anthropological connotations and metaphysical search for the essence of humans is humanity or humanness understood in the sense of the Latin word humanitas (see also Nida-Rümelin and Winter 2024). In this interpretation, digital humanism is more concerned with acts of humanity and kindness and of course with digital technology as an enabler of inhuman acts. This implies a more relational view of human and machine and the concept of care emerging from this relation rather than an essentialist perspective of focussing on what it means to be human in terms of ontological features for the specificum humanum.

Such care may concern the individual and also the society that becomes affected by system-human interactions – if only as an emergent property of many such interactions. The task for system development and for constraining the power of digital systems and their owners then becomes one of caring for individuals and society that is affected by digitization. Very much in the tradition of an ancient conception of '*humanitas*', this leads us to a virtue ethics of digital development and a consideration of both wisdom and praxis. It also connotes *dignity,* a notion that plays a central role not just in humanism but also in the Age of Enlightenment and the ensuing human rights.

2.2. Enlightening our dark digital ages

The year 2018 witnessed both the publication of Zuboff's *Age of Surveillance Capitalism* and a special issue of the magazine *Wired* announcing that the internet was broken on its titlepage. They are examples of a wave of sobering criticism about the status quo of our digital environment, the internet, and of social networks. The criticism is particularly drastic because it declares the ideals of a previous period as having disappeared. These ideals included the use of electronic networks for a productive exchange of ideas between individuals, academics, and nations; the joint shaping and use of a creative commons facilitated by the internet; and the free and open access to knowledge, software, and

tools. These are all ideals inherited from the Age of Enlightenment and a *Republic of Letters* as the first wave of international academic exchange of ideas, sharing of knowledge, and a solid belief in progress including through science and engineering. Famously, Immanuel Kant explains Enlightenment as "man's emergence from his self-incurred immaturity" hoping that this happens in a society of free thinking individuals (Kant, 1784).[5]

Instead, the articles in said *Wired* magazine lament that the free sharing of creative output started to threaten the creative industries. The early ideal of anonymity became a factor in abusive trolling and hacking innocent people. Big data lost its innocence by enabling a surveillance tool of unprecedented power and the rise of online platforms created an oligopoly of attention (Prat & Valletti, 2022). These descriptions are all the stronger for being put in contrast with the old ideals with which Berners-Lee and others sought to develop the internet. Since 2018, a range of diagnoses were published that fall nothing short of describing our digital environment as a return to the Dark Ages – a time of cultural and intellectual decline and of economic exploitation. Digital humanism hence seeks to enlighten the Dark Ages including through education, critical thinking, and systematic critique. Much of the work in digital humanism is directed at enlightening broader audiences about technology monopolies, big tech powers, digital network effects, technologies for privacy protection etc. It is also directed at liberation – in this case from the powers of oligopolies and state surveillance while also standing up for tolerance, equity, gender equality and anti-discrimination. And it underlines the importance of the commons (e.g. of language and knowledge technologies or access to data), of equal opportunities, and fairness in harvesting the benefits of the digital world.

The development of universal human rights is often put in the context of the Age of Enlightenment. Rooted in antiquity the movement towards setting limits to despotism gained huge momentum with the French revolution. Pollmann (2022) characterises modern human rights as an expression of a revolutionary progress of the law of nations that is by no means irrevocable. The fact that such progress indeed seems revocable today is one of the drivers of the emphasis on human rights in digital humanism. Human rights rest on human dignity and often both refer to each other. Despite their roots in many philosophical and political traditions (e.g. Plato, Cicero, the Magna Charta, and Confuzius) the idea started to develop throughout the 17th century and post-revolution France. It is only consequential that human rights and how to guarantee them in an increasingly digital and interconnected world are a focus of digital humanism. There are proposals to expand the current set of human rights, for example with a right to human decision-making because of increasing AI-based decisions in all sorts of applications (cf. Paola (2022), Blume and Rauchbauer (2022)).

The programme of digital humanism builds on the historic achievements of both humanism and Enlightenment. The next section summarizes this programme and the positions of digital humanism in five principles that include the quest for human dignity and the ideal of a better society based on core values of the Enlightenment.

3. Five principles of digital humanism

Here, I propose five principles to summarize and to structure the activities and academic debates that exist within digital humanism. The principles are necessarily abstract and omit many of the details of current debates but help to understand the core claims and lines of thinking. We begin by diagnosing the current situation as disturbing followed by making demands about the future of digital technology. We also clarify that a better future and change are feasible and that there are important human-machine differences to be preserved. Hence, the principles are concerned with

1. the impact of technology on people and their co-evolution,
2. a mandate for technology to protect people and environment,
3. a demand for technology to strengthen democracy and society,
4. the assertion that technologies are malleable,
5. the confirmation of the differences between people and machines.

3.1. Digital technologies are changing us and our lives

Digital technology has revolutionized our daily lives. It can help save human lives, facilitate the sharing of knowledge and development of culture, bring people together, and enable resource-saving economic activity. Already the manifesto emphasizes how digital technologies have led to a disruption of "the very fabric of society" from communication to institutions and political structures. This also includes science and the humanities.

Digital humanism critically examines narratives that are associated with the development of digital technologies. As much as the positive aspects of our digital world are acknowledged, we are reminded that not every digital innovation also means progress for people and society. This is an important reservation against the omnipresent claim that digitization was per se beneficial for individuals and society. Such claims are now characteristic of the marketing and public relation messages that technology firms use to advertise new products and more broadly improve the public perception of their brands. They are also an important component in political messages ranging from the advertising of modernization policies to public technology development programmes, the renewal of education curricula, or the digital transformation of public services. Digital humanism reminds us that digital technologies and their applications enable unprecedented intrusions into our privacy. They promote private and government surveillance and can threaten our fundamental rights. It thus questions narratives that suggest digitization as an inherently beneficial form of progress for people and society. And, perhaps more importantly, it tends to also question narratives of *invisible hands* that guide technologists towards the design of better futures for everybody simply by the power of markets. Quite to the contrary, digital technology has led to undesirable changes and has unfavourably shifted important private and state power relations without prior discussion and without a vote on it. Network effects of digital systems can support the formation of monopolies and intensify economic imbalances.

The fact that digital technologies are changing people can also be taken quite literally. There are unwanted changes in how young people develop through extended interactions with digital systems. They range from increasing short-sightedness (e.g. Liu et al. 2021, Foreman et al. 2021) to debates about decreased attention spans[6] and the problematic self-image of young people emerging from their interaction with social media (Timmers, 2022).

Such developments may in turn trigger and guide the development of new digital technologies, a phenomenon that has been called co-evolution of technology on the one hand, and people and society on the other. Lee (2020) describes several such co-evolutions and explains how technological development is neither fully under the control of individual engineers, organisations, nor society. This however should not discourage us from developing improved visions for digitization and technology development. It rather implies a strong call for improved ways of shaping our digital future. There are two important caveats, however. This design must be feasible (see below) and secondly, we should have clear objectives and agree on where we want to go. This is

[5] In the interpretation of Foucault (1994), Immanuel Kant describes Enlightenment as the moment when humanity starts putting its own reason to use, without subjecting itself to religious or state authority and to limit illegitimate uses of reason that give rise to dogmatism, heteronomy, and illusion.

[6] Cf. Prensky and Berry (2001) for arguments for the thesis and Vedechkina et al. (2021), for an overview and more cautious conclusions.

the subject of the next two principles.

3.2. Digital technology must protect people and the environment

Unsurprisingly, the central message of digital humanism is its focus on humans and their dignity. Their freedom and self-determination should be measures of the development and use of digital technologies. Digital technologies should promote human autonomy and empower people to make their own decisions. Where it supports human decisions, it must support fairness and counteract the possibility of biased decisions.

This is humanism at work, i.e. in practice. A key identification in this philosophy concerns the identification of humans with the autonomy to act (agency) and the central role of human dignity. *Agency* is central because it relates to action, both human and computer-controlled action and potential conflicts arising from algorithmic decision-making on the one hand or limitations and interferences of computer-controlled actions with people's autonomy (Floridi, 2023).

The concept of *dignity* is key where humans should be protected even when they are not directly affected by the actions of computer-controlled machines or automated decision-making. For example, protecting a person's privacy is not only a right or a cautionary principle in case some hackers succeed in stealing personal data from a computing system. It should also be regarded a basic human right or indeed a trait of all humans. As such, questions of financial compensation for giving up privacy may often miss the point in that privacy cannot be upended contractually in principle–to list just one example where such a conception of privacy as dignity will make a difference in designing our digital future.

Throughout history, the idea of putting humans and human society at the centre of a philosophy has not been without criticism. One historical critique concerns the appropriate definition of who has been considered 'human' as in traditional humanism such definitions have often at least implicitly focused on the European white academic man. This view can be linked with colonialism and the imposition of values and views without any consideration of other people, minorities, and different ways of life around the globe. Clearly, this also concerns the export of digital technologies designed in only a few parts of the world and from a background of specific values and little to no interest in different cultures. It is therefore systematically important that digital humanism aims for digital technologies that recognize and promote people in their differences, no matter where and how they live. This requires not just critical inquiry, but an openness to and inclusion of other cultures in the formulation of design objectives and in the design of digital technologies. It requires an onboarding of diversity including feminist positions which are still far from mainstream in IT design. It may at times require to set limits to technologies for cultural reasons – a fact that has not yet found universal acceptance for a technology that is fundamentally open and networked and therefore usually impacts internationally.

The other historically relevant and hugely contemporary criticism of humanism concerns the focus on *just* the human disregarding nature or the environment. It is therefore of great systematic importance that digital humanism moves beyond the human individual and human society by debating and critically analysing the impact of digital systems on the environment. Increasingly, digital humanism calls for sustainability of digital technologies and the utilization of digital tools for reducing harmful human impacts on the environment (Kaack et al., 2022; Rolnick et al., 2022). Digital technology should keep our environment liveable and sustainable. The current debate in digital humanism clearly shows how complicated these questions are. Not only are there huge gaps in our understanding of the environmental impacts of digital technologies, but the field is also moving very quickly and the separation of potential positive effects of the technology from its detrimental impact is proving notoriously difficult.[7]

3.3. Digital technology should strengthen democracy and society

Digital humanists have called for the development and implementation of digital technology that is put at the service of democracy, of basic and human rights. Digitisation should help to promote justice and freedom and to meet the needs of all parts of society. This demand is easy to agree with, but at odds with how digital technology appears today. In fact, digital technology is widely blamed for reinforcing or even enabling recent threats to liberal democracies worldwide. These threats operate at different levels. At the level of applications, there is a lively debate about the distribution of illegal content, online bullying, and the distribution of fake news disguised as alternate 'facts' in online networks (Prem & Krenn, 2024). At the level of companies, the concerns for democracy emerge from the tendencies towards winner-takes-all phenomena and network effects strengthening a few large tech companies rather than producing a rich and competitive market of many companies. In addition, Zuboff (2019) has pointed out how large technology companies use their power to influence political decision making.

Furthermore, digital humanists demand that digitization should contribute to social welfare and solidarity, preserve and expand social and cultural achievements and should make them accessible for all. It should enable a broad education and contribute to rational debate. It must enable meaningful work that is socially recognised.

Digital technology is closely linked with network effects that are often associated with the formation of monopolies. Where the value of a service increases with the number of users, there is also a tendency of users moving towards that service effectively limiting the number of service providers and, hence, choice. Such service providers can harvest scale effects and create huge market barriers for new entrants given the enormous costs of infrastructure. An important discussion in digital humanism concerns the precise role and nature of possible economic and power effects of large technology firms. A key question in these debates is whether the concept of monopoly applies to the internet giants and whether our current legal and economic frameworks are sufficient to effectively limit the power of the large players to avoid the abuse of consumers. Although the question of monopolies remains a key driver of many legal discussions and an important driver of legislation, there are also signs that different conceptions are required to better control problematic developments. As an example consider the proposal of Staab (2019) to regard large internet giants as market owners rather than monopolists. Companies such as Amazon or Apple can practically design the rules of their markets and web stores as they wish. Different from producer monopolies, the owner of a market has various means of

[7] David Rolnick's lecture provides an example of the complicated ways in which AI may accelerate or impede climate change. On the one hand, AI can facilitate climate change mitigation and adaptation strategies within a variety of sectors, on the other AI can also contribute to rising greenhouse gas emissions through applications that benefit high-emitting sectors or drive increases in consumer demand, as well as via energy use associated with AI itself. https://caiml.dbai.tuwien.ac.at/dighum/dighum-lectures/david-rolnick-is-ai-good-or-bad-for-the-climate-its-complicated-2022-11-15/ Cf. Rolnick et al. (2022), Kaack et al. (2022).

control: information, access, price of participation, service levels and provision. Such markets are not neutral but conform to strict access controls exerted by the large players. This results in a combination of access control and commission. This goes beyond mere 'neoliberalism' – a term often used to describe the underlying economic philosophy, especially of large digital actors. Staab suggests the term *neo--mercantilism* for the idea of owned markets. In addition, there is a growing debate also in digital humanism that a focus on traditional economic monopolies, i.e. on distorted prices and a focus on consumer markets, is insufficient for overcoming the current challenges and more attention is needed, for example on marketing power (e.g. Baeza-Yates and Fayyad 2022).

3.4. Use and application of digital technology can be shaped

Realizing a more positive digitally enabled future requires the malleability of our digital future. This is more than a logical truism. It demands that our digital future can be realistically influenced following a great vision and that society takes a clear stance and corresponding actions. In this aspect digital humanism antagonises the narrative of a hardly controllable evolution of technological progress; it is a call for self-determination and activism. The core digital humanist claim is that *technology is not a destiny:* It can be shaped to support democracy and must comply with the rules of democratic societies.

Note that this is not necessarily in contradiction with Lee's aforementioned analysis of technology development as an evolutionary phenomenon. While Lee is probably right that we are not in full control of every aspect, this does not imply that we have no control at all. Lee compares the situation with the control of our natural environment. It evolves without our full control, but humans still actively shape their environments from agriculture to breeding animal species. The description of the course of technological development as necessary and following a path of technological evolution driven by competitive market forces often arises in debates about technology regulation. The *Silicon Valley* narrative includes an aspect where technology development is described as a sequence of necessary steps to stay competitive and to give the people what the people want or need. Zuboff has called this a rhetoric of inevitabilsm that distracts public debate from the intentions of large tech companies that drive technology development (Zuboff, 2019 p. 342). Beyond this merely evolutionary narrative, Daub (2020) describes how a discourse about technological and business disruption led by tech giants paints an image in which continuous development poses risks of losing out. Daub explains the believe in a certain kind of danger when progress is only slow as a feature of modernity (Daub 2019, p. 120). Radicality of progress has become a condition of modernity, hence rendering a more considerate, perhaps more careful consideration of progress unmodern. Digital humanists, however, point out that technological innovation is not equivalent to societal progress: "[Digital humanism…] maintains a positivist goal for technology to create societal progress rather than just innovation for the sake of economic growth" (Werthner et al., 2022a, preface).

Significant parts of the debate therefore address the self-interest of market participants such as the large technology platform that pushes certain technologies into the market while hindering others. Where applications of digital technologies have a significant impact on our lives, the regulation of technology should come from people in a democratic manner. Neither companies nor the market nor technologies make the law, but the democratically legitimized institutions. The enforcement of these rules is to be ensured everywhere by democratically legitimized state authority. Digital technology is a driver of innovation and should be used creatively. Using it for progress requires the cooperation of administration, businesses, and research as well as transnational cooperation. Promoting the desirable effects of digital technology is the responsibility of everyone involved in its development.

3.5. Machines are not people

Given the central motivation and ideas of humanist thinking it is unsurprising that most if not all digital humanists firmly believe that there are fundamental differences between humans and machines, that these differences matter, and that digital technology should not blur these differences. This implies that digital humanists are neither post- nor transhumanists.[8] *Trans-humanists* debate a future in which technology plays a key role in expanding human bodies and cognitive functions by means of digital technology or bioengineering. In transhumanism, there is an underlying assumption that humans benefit from digital implants or other improvements and that such additions to the human are desirable or perhaps even necessary. The typical examples include illness and weakness of the human body, but also mortality. The transhumanist therefore holds a *deficit perspective* of human beings as weak and inherently requiring technical improvement, eventually perhaps reaching immortality by means of technology. From both the humanist and digital humanist perspectives this is a highly questionable assumption. Instead, digital humanists will often argue that humans are simply the best humans that we can and should consider. Rather than trying to improve people with the help of digital technology, the aim of digital humanism is an improved technological development that accepts and supports humans in what and how they are or what they can become as human beings.

It should therefore be clear that digital humanism has no sympathy for the ideas of *post-humanists* who believe that an inescapable technological evolution will eventually lead to overcoming humans by means of a super-AI (e.g. Moravec 1998, Kurzweil 2005). The post-humanist idea suggesting that the development of AI may eventually render the human body useless is not just found improbable in digital humanism, but completely alien in its disregard for the human and of human society as goals-in-themselves (Blume & Rauchbauer, 2022). For a digital humanist, neither post- nor trans-humanism are very clear about the sense in which they represent progress and improvement, i.e. against which measure (if not the human) and for what precise purpose such improvements should evolve. Very much in line with the rejection of equating machines and people, digital humanism is also opposed to claims that people must put themselves at the service of technology. People are not obliged to support digital technology nor its development – a line of thinking that is sometimes implicit in Silicon Valley narratives about technological progress.

4. Disciplines of digital humanism

From its onset, digital humanism has been perceived as an endeavour that requires contributions from diverse disciplines such as philosophy, social science, law, or economics. Interdisciplinarity emerges naturally from the motivation to analyse and influence the design of digital systems with a focus on the human and on society. Digital humanism actively encourages contributions from and interactions of various disciplines that have contributed to related debates. This ranges from technical disciplines, sciences of the human and of society to scholarly fields of philosophy and legal studies.

4.1. Designing a desirable digital future

Digital humanism is a multi-disciplinary and multi-purpose endeavour. It is multipurpose in its aims because it seeks

[8] As used here, post-humanism refers to the idea of a ,'super-AI' that exceeds humans rendering human bodies and eventually humans useless. The latter must be distinguished from *critical post-humanism* that develops in reaction to criticisms of humanism, e.g. against colonialism or individualism. However, the notions of trans- and post-humanism are not consistently used in the literature.

(i) to support the development of digital technologies that are better aligned with the digital humanists' vision of progress (e.g. liberal democracy),
(ii) to understand the implications of digital technologies for individuals, societies, and our environment,
(iii) and to intervene politically to limit and better balance the power of large enterprises with those of states and individuals.

4.1.1. Digital humanism as a technical discipline

Improving digital technologies is an engineering endeavour that focuses on improving existing technologies, on using already developed technologies (e.g. new methods for privacy preservation, technologies with less negative impact on the environment) more widely, and on creating the right types of systems and applications. While supportive of these directions for technical development, this is not to imply technical solutionism such as putting an app where a change in society may be required.

Digital humanism has made the case that these developments cannot be achieved by computer science alone. Not only does it require the involvement of IT system users and other stake holders in the design, there often is also a need for further critical examination and investigation, for example by technology impact studies, sociological and psychological assessments. Social science plays a strong role in digital humanism with its many branches from the study of political systems to group behaviours or economics. Much in the tradition of German humanism, pedagogy plays an important role in enlightening all generations about digital technology, its impact, its malleability, and ways to improve and use it.

With the intensification of geopolitical debates and calls for stricter regulation, for example of artificial intelligence, links with legal studies and political science have intensified (e.g. Bradford 2023). There are strong indications that digital technology and large digital enterprises call for innovation in these fields. Examples include the debate about state sovereignty, new concepts for effectively limiting the power of large tech enterprises beyond monopoly law or new ways of thinking about privacy. Given the trans-national character of many digital systems, these debates require an international collaboration which also calls for new legal concepts, bodies, and instruments.

4.1.2. Digital humanism as a political movement

Like many other '-isms', digital humanism offers a position that can be seen as a political endeavour or a societal movement, similar to other examples such as socialism, liberalism, Marxism, or conservativism etc. As such it is to be distinguished from '-ics' and '-logy' as in the cases of logic, physics, or biology.[9] An '-ism' includes a set of foundational beliefs of its proponents and usually comes with the intention to disseminate it and help it succeed, i.e. reach its objectives. Hence, analysing digital humanism exclusively as a study of the human in relation to digital systems or just as an engineering effort falls short of its political aspects, opinions, and societal aspirations.

An important part of this political endeavour is to investigate power shifts triggered by digital technology. This includes shifts at the individual level, e.g. from persons to companies or the state. It also includes shifts at the societal level where large technology companies have entered a competition for power with nation state. Certainly, digital humanism may often study the relation of digital systems to society, but it is not shy to offer its normative views and beliefs in doing so. As a movement, the positions of digital humanism are often liberal, e.g. when emphasizing personal liberty and autonomy, occasionally mildly Marxist, e.g. in investigations about alienation of work through algorithmic systems or automated production, and many times social democratic, e.g. when emphasizing social fairness and equality in relation to economic and social order. At times, it is also conservative, especially regarding traditional values and virtues so that it results as an initiative above party lines.

Legal scholars are of key importance given their long tradition in debating, specifying, and practically implementing concepts and regulations. They also play a key role in proposing new regulation fit for an increasingly digital world. This may require adapting existing legal frameworks and questioning millennia-old legal concepts, e.g. the fundamental separation of things from persons in many jurisdictions. And it may require to go beyond the conventional borders of legal systems (e.g. continental Roman law or case-law) to effectively work at the international level. New technologies with high degrees of autonomous decision-making are threatening some old dichotomies (e.g. between things and persons) and scholars have started to investigate new concepts. Another important line of thinking in legal studies focuses on basic and human rights and their application in highly internationalized digital contexts. This includes, for example, debates concerning freedom of speech versus censorship, conceptions of privacy, and questions of responsibility and accountability. They also contribute to the analysis of power relations – traditionally also a field of political scientists. The investigation of digitally enabled power relations and how to tame them from a digital humanist perspective have played an increasing role in the last years.

4.1.3. Digital humanism as a philosophical endeavour

Digital humanism is also a philosophical endeavour, especially with regard to debating established terms and concepts or proposing new terms and relations. As already mentioned, it touches upon many key concepts that have been debated in the history of philosophy. From the foundational issues of human/machine differences, the whole discipline of philosophical ethics is experiencing renewed dynamics with issues such as ethics of AI or algorithmic decision making about humans. Philosophers are also contributing to the question of how to deal with and relate to digital systems as humans, for example regarding the question of human-robot relations. A particular philosophical aim is the quest to delineate the scope of the digital realm and to draw boundaries for the digital, e.g., in making life-and-death decisions about humans.

There are also important epistemological questions about the 'theory-free' nature of knowledge accumulated in statistical AI systems such as neural networks (Anderson, 2008) and the ensuing debate about understanding and explanation of AI-based decisions (Speith, 2022). Unsurprisingly, the advent of generative AI and advanced chatbots have triggered substantial debates about notions of meaning and understanding as well as questions of proper speech, fairness, and bias of AI systems. In addition, questions of responsibility for the impacts of technical systems such as robots or AI systems are debated within digital humanism as well as more fundamental questions about the methodological paradigm shifts presented by digital technologies and disciplines such as computer science, cf. Schiaffonati (2022).

4.2. A holistic and interdisciplinary approach

Methodologically, digital humanism is an interdisciplinary effort to debate a broad range of digitisation shortfalls in their totality, from privacy infringements to power shifts, from human alienation to disownment. Digital humanism is also transdisciplinary and has received enormous interest and participation from artists. The Arts can play an important role in the debate. This is evidenced in many art-based initiatives, festivals, and events that focus on many of the issues that digital humanism debates as well. Institutionalised festivals or museums such as *Ars Electronica*[10] or the *Museum of Applied Arts* in Vienna have

[9] This distinction is even clearer in other languages such as German where we can distinguish '-ismus' as in Humanismus from '-istik' as in Humanistik.

[10] Ars Electronica is a prime annual electronic arts festival and initiative, https://ars.electronica.art

debated topics from privacy to AI and from questions about "Who owns the truth?" to "What do we want?"[11] The latter is a particularly difficult question that requires answers from visionaries who are unbounded by predicted trajectories of a technology progress that is considered to be predestined. It will have to remain at the core of digital humanism for years to come because it has remained unsatisfactorily answered ever since the old ideals of the internet have disappeared.

The interdisciplinary nature of digital humanism roots in a holistic approach in trying to understand and making actionable unwanted consequences of our increasingly digital world. The following example may explain the rationale for a necessarily holistic approach. Consider how an abundant and ever more efficient digital hardware facilitated the creation of digital service platforms. People's online interaction with these platforms have led to harvesting personal data at an unprecedented scale which in turn facilitates the creation of prediction models of people. These models are used for digital marketing and control that have led to establishing large tech firms as new digital powerhouses (Zuboff et al., 2019). This is but one line of causation from digital hardware to digital power that demonstrates how specific characteristics of digital technology interact with each other and affect individuals and society (Fig. 1).

Several elements in this development are now recognized drivers of problems debated in digital humanism, for example political and democratic questions arising from digital control over communication and people or issues of resilience that directly relate to digital hardware and to the power of large tech companies. It is a central claim in digital humanism that the challenges as much as the technology development require *joint* consideration. It is insufficient to only focus on isolated problems such as AI ethics or privacy. Such debates may often be hugely valuable, for example to design new privacy-preserving technologies, but they are at the same time insufficient in addressing the problems. Privacy or AI ethics require a consideration in the context of the underlying technological development, questions of human autonomy and democratic societies, and the related power structures. This holistic perspective is necessary to avoid the type of technological solutionism that critics have often diagnosed of narrow engineering-based approaches.

4.3. Related fields

While digital humanism overlaps with a range of established fields and other movements, it reflects a new academic, engineering, and societal awareness of the challenges of digital technologies. This is not to say that the issues had not been subject to scholarly debate before; quite the opposite: From *science and technology studies* to *anthropology*, from the *philosophy of technology* to *human-centred computing*, there is a broad range of academic fields that were similarly motivated by current societal problems of digital technologies. However, there are important differences and distinctions:

Anthropology as the science of humans focuses on biological aspects and it also includes philosophical anthropology that is closer to humanism in its investigation of what is it that makes us human, i.e. a *specificum humanum*. This concerns human characteristics such as the autonomy to make decision, personality, self-reflection (including as a human about being human). Social and cultural anthropology include societal and cultural aspects. As explained above, the focus in digital humanism is much more about technology aspects and its implications as well as about a normative forward-looking perspective about a better future for humanity in living with digital systems.

Science and technology studies (STS) have addressed many important problems of digital technologies and their relationship with humans and society. STS investigates technologies in their historic, cultural, and social contexts. It is often described as a confluence of various fields with a focus on social constructions of technology. It has in the past often critically analysed digital technologies and their problematic impacts on society. However, it was rarely strongly perceived in engineering, nor has it triggered massive directional change in developing digital technologies.[12] Traditionally, STS is methodologically closer to social science than to formal or technical sciences and therefore often not interested in devising better technologies. This may be a reason for its relatively small impact on computing disciplines in the past. There is clearly a role for STS to play in digital humanism to contribute its analyses and theories much more to computer science and the design of digital technology than in the past.

Human-centred design and computing is a specialised discipline of human-computer interaction with a strong focus on human factors and ergonomic design. It addresses several concerns that are associated with digital humanism, in particular effects of digital systems on individuals. However, it is usually less concerned with societal aspects or power relations that are central to digital humanism.

Finally, the *philosophy of technology and media* is an area that has significant overlaps with digital humanism. It is, however, another non-technical discipline that focuses on the analysis of concepts, provides theories, and principles aiming to clarify a broad range of aspects of technology. In particular, it is concerned with the meaning of technology for humans and for society. Today, a significant portion of the debates in the philosophy of technology and media concern ethical aspects of digital technologies such as AI and online social networks. This links with the interests of digital humanism although it is not concerned with building digital systems.

Digital humanism therefore appears as a unique domain of scholarly investigation in its critical yet constructive approach to digital technology. It has so far proven to maintain a productive dialogue with many other disciplines although the differences in objectives, methods, and vocabulary between these fields clearly provide significant challenges as well.

5. Discussion and conclusions

5.1. Assessing digital humanism

Digital humanism as it emerged from the Vienna Manifesto has become a considerable academic movement with societal and political aspirations to shape the post-digital. It cannot yet be considered a scientific, engineering, or scholarly discipline on its own for its lack of typical components of a research or engineering discipline such as its own methodology, body of knowledge, curricula, or institutes. It has earned importance in describing and influencing power shifts and societally detrimental directions of technology development. It has influenced research programmes and has created research roadmaps with pointers for future research (see below).

Despite its unqualified reference to humanism, digital humanism should be characterised as a type of *critical post-humanism* as it is hardly an unreflecting, backwards oriented, or romantic movement that aims to revive a Renaissance ideal. Digital humanism has convincingly and credibly criticised parts of computer science and its application for being overly focused on the individual. This puts it in the following of approaches aiming to overcome an excess of individualism characterising historical humanism. It thematised colonialism and the mis- and underrepresentation of women online (Canca, 2022). In addition, much of today's efforts in the field concern a societal and political dimension and

[11] The Austrian Museum of Applied Arts (MAK) held an exhibition on "What do we want? Dimensions of a new digital humanism" in 2017 as a part of the Vienna Biennale. (Vienna Biennale Guide, 2017) https://www.mak.at/en/program/exhibitions/what_do_we_want_

[12] Cum grano salis. To be fair, STS has often provided important input to political decision making, for example in early stages of debates about privacy and data protection.

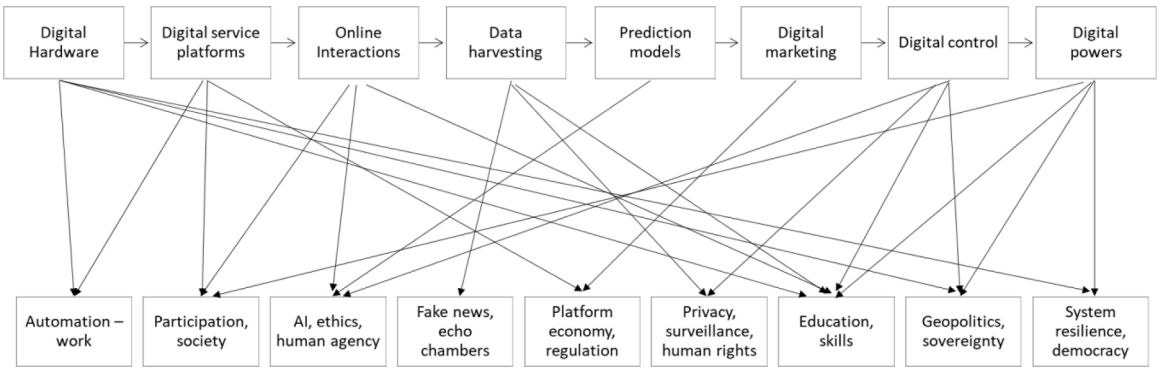

Fig. 1. An exemplary linkage of digital technology development with key issues in digital humanism for the case of large online social networks.

not just a human factor aspect in engineering.

It has recently also included an environmental dimension in its debate although this is an extremely challenging area. While the enormous computational effort for calculating AI models has been rightly criticized, precise figures about the environmental impact are changing with technical advances. More importantly, new digital applications are also helping a transition towards renewable energy systems and could play an important role in mitigating the effects of climate change.

The response of digital humanism is not yet fully convincing, especially given its lack of proposals for technological alternatives to current developments. It has a holistic perspective that debates the connections of various problems but poses the question of practicability. It cannot yet be directly used as an implementation guideline in the development of IT systems. It may be more successful in societal delineation (i.e. setting limits) and enlightening academics, a broader public, and decision makers. However, the involvement of companies has only just started and will be key for ensuring the impact of digital humanism beyond regulation. The industry is interested and started to embrace digital humanism, but at the same time is also reinterpreting it (cf. Krause 2023). This reinterpretation focuses mostly on the idea of human factors, human-centric interfaces, and inclusivity. It also occasionally includes environmental effects. But it is, unsurprisingly, more difficult for industry to embrace the societal aspects and the calls for limits on the power of digital industry.

There remains an inherent paradox in the idea of a *digital* humanism. The digital realm as a general model of computation (i.e. the algorithm) and a model of practically everything including living beings practically proposes an equivalence of humans and machines. It is based on the idea of a fully mechanised explanation of nature using numbers and rules. Humanism and digital humanism, however, aim to reserve a special place for people. They emphasize the differences of humans and machines but the challenge of clearly delineating the two worlds remains. Digital humanism therefore asks questions about the human and human society in a digital world from a normative perspective of traditional humanism. It is pressed to provide answers based on those derived from a pre-digital human condition. But it should also progress beyond a traditionally normative perspective where it acknowledges the post-digital, for example human-machine co-evolution. The quest then becomes to investigate, define, and expand the human relation to itself, to others, and to the environment in the post-digital. This task goes far beyond a principled investigation into morality (i.e. ethics) in improving our understanding of human dignity in a post-digital world.

Such a post-digital humanism is a call for empowerment of users, of those affected by digital technology including society. It proposes a normative preference for digital technologies that increase access to knowledge, put the human in control, lend a personal voice to the user. It aims to develop AI that works in partnership with humans. It urges the use and development of IT tools to ensure personal privacy and to mitigate the risks of categorization and prediction, e.g. through transparency of and individual power over recommendation and personalisation.

It calls to further democratisation, participation, and inclusion in society, support diversity, and guarantee fundamental rights. It aims to develop and use digital technology for strengthening the social contract (e.g. between the rich and the poor, between the powerful and the disempowered, between the generations, between the workers and the unemployed, between the healthy and the sick).

5.2. *Outlook and challenges for research*

Beyond its political intentions, digital humanism points to important future research areas or fields that require more attention from researchers, engineers, and scholars of all involved disciplines. The current strong focus on AI is unavoidable, but many more issues require and deserve attention. If the diagnosis of a *failing system* (as in Section 2.2) is correct, we will need a new vision or at least visionary elements that future and better digital technologies should help realize. We require a much stronger development of technologies for the greater good, for democracy, and for human rights than what is developed today. This will perhaps mean investments from the public rather than from private actors alone who necessarily will have to consider commercial success. Digital humanism shows the way to debating sovereignty and humanity in times of undirected development where innovation is mistaken for progress. It is unlikely that a new digital world can embrace precisely the same values and established principles that have governed our non-digital history. At the same time, this cannot mean to throw all established principles and learnings overboard. The difficult question then is which values to keep core and where to embrace new or adapted principles.

Digital humanism requires research at the fundamental and applied levels (cf. Prem et al. 2022). The former includes research into the nature of computation and its design (e.g. in the light of human/machine co-evolution), the nature of privacy, and digital power, of dematerialisation of products into services, etc. Putting the human in control requires new architectures and tools for governing personalization and recommender systems. Improved privacy and security architectures and tools, and fundamental questions about how we would like to live in a post-digital world. This includes, for example, research at the intersection of philosophy, ethics, linguistics, psychology, and AI to better understand which types of language models should be developed and how they should speak with us.

Other key topics at the societal level include sovereignty and resilience, new regulatory schemes, and concepts beyond just monopoly law – topics that will require the collaboration of political scientists with engineers, regulators, and legal scholars. Similarly, we may have to work towards overcoming the ontological, feature-based perspective of what it means to be human towards a more relational view of our environment and the objects (including AI and robots) to which we are developing specialised relations. Finally, digital humanism should also help provide answers regarding the limits of digitalisation including, for

example, a right to a human decision and investigate proper levels of non-digital solutions for those unable or perhaps unwilling to use digital tools as opposed to a future where life is prescribed as necessarily digital.

Declaration of competing interest

Research reported here was supported by the University of Vienna. There are no conflicts of interest.

Acknowledgments

The author would like to thank the anonymous reviewers for their comments and the Vienna Science and Technology Fund (WWTF) for stimulating the debate about five principles for digital humanism.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.jrt.2024.100075.

References

Adorno, T.W., & Horkheimer, M. (1944). Dialektik der aufklärung (Dialectic of enlightenment), Frankfurt: Suhrkamp, 1997.
Akkermanns, H., Gordijn, J., & Bon, A. (2022). Return to freedom: Governance of fair innovation ecosystems, 2022. In H. Werthner, E. Prem, E. A. Lee, & C. Ghezzi (Eds.), *Perspectives on digital humanism*. Cham, CH: Springer Nature, 53 ff.
Anderson, C. (2008). The end of theory: The data deluge makes the scientific method obsolete. In Wired. Jun 23, 2008.
Baeza-Yates, R., & Fayyad, M. (2022). The attention economy and the impact of artificial intelligence, 2022. In H. Werthner, E. Prem, E. A. Lee, & C. Ghezzi (Eds.), *Perspectives on digital humanism*. Cham, CH: Springer Nature, 123 ff.
Blume, C., & Rauchbauer, M. (2022). How to be a digital humanist in international relations: Cultural tech diplomacy challenges Silicon Valley, 2022. In H. Werthner, E. Prem, E. A. Lee, & C. Ghezzi (Eds.), *Perspectives on digital humanism*. Cham, CH: Springer Nature, 101 ff.
Bradford, A. (2023). *Digital empires: The global battle to regulate technology*. Oxford: Oxford University Press.
Canca, C. (2022). Did you find it on the internet? Ethical complexities of search engine rankings, 2022. In H. Werthner, E. Prem, E. A. Lee, & C. Ghezzi (Eds.), *Perspectives on digital humanism*. Cham, CH: Springer Nature, 135 ff.
Daub, A. (2020). *What tech calls thinking: An inquiry into the intellectual bedrock of Silicon Valley*. FSG Originals.
Douhei, M. (2011). *Pour un humanisme numérique*. Paris: Editions du Seuil.
Floridi, L. (2023). *The ethics of artificial intelligence: Principles, challenges, and opportunities*. Oxford: Oxford University Press.
Foreman, J., Salim, A. T., Praveen, A., Fonseka, D., Ting, D. S. W., He, M. G., & Dirani, M. (2021). Association between digital smart device use and myopia: A systematic review and meta-analysis. *The Lancet Digital Health, 3*(12), e806–e818.
Foucault, M. (1974). *Von der subversion des wissens* (pp. 141–170). München: Hanser. In Seitter W. (ed.).
Foucault, M. (1994) Qu'est-ce que les Lumières? (What is Enlightenment?) Dits et Ecrits, Vol. 4.
Gallie, W. B. (1955). Essentially contested concepts. In , *56*. *Proceedings of the aristotelian society* (pp. 167–198). Aristotelian Society, Wiley.
Guide, V.B. (2017) Roboter. Arbeit. Unsere Zukunft (Robots. Work. Our future) MAK wien, Verlag für moderne Kunst.
Johnson, P. (2018). *Feminism as radical humanism*. London: Routledge.
Kaack, L. H., Donti, P. L., Strubell, E., Kamiya, G., Creutzig, F., & Rolnick, D. (2022). Aligning artificial intelligence with climate change mitigation. *Nature Climate Change, 12*(6), 518–527.
Kant, I. (1784). *Was ist aufklärung? Ausgewählte kleine schriften, 512* p. 1999). Felix Meiner Verlag.
Krause, G. (2023). *Die praxis des digitalen humanismus. (The practice of digital humanism)*. Wiesbaden: Springer Vieweg.
Kurzweil, R. (2005). *The singularity is near: When humans transcend biology*. New York: Penguin.
Lee, E. A. (2020). *The coevolution: The entwined futures of humans and machines*. Cambridge: MIT Press.
Lévi-Strauss, C. (1973). *Anthropologie structurale deux*. Plon-Julliard.
Liu, J., Li, B., Sun, Y., Chen, Q., & Dang, J. (2021). Adolescent vision health during the outbreak of COVID-19: Association between digital screen use and myopia progression. *Frontiers in Paediatrics, 9*, Article 662984.
Moravec, H. (1998). *Robot: Mere machine to transcendent mind*. Oxford: Oxford University Press.
Nida-Rümelin, J. (2022). Digital humanism and the limits of artificial intelligence, 2022. In H. Werthner, E. Prem, E. A. Lee, & C. Ghezzi (Eds.), *Perspectives on digital humanism* (p. 72 ff). Cham, CH: Springer Nature.
Nida-Rümelin, J., & Winter, D. (2024). Humanism and enlightenment, 2022. In H. Werthner, E. Prem, E. A. Lee, & C. Ghezzi (Eds.), *Perspectives on digital humanism* (p. 3 ff). Cham, CH: Springer Nature.
Paola, I. (2022) The challenge of human dignity in the era of autonomous systems. In.
Pollmann, A. (2022). *Menschenrechte und menschenwürde. (Human rights and human dignity)*. Frankfurt: Suhrkamp.
Prat, A., & Valletti, T. (2022). Attention oligopoly. *American Economic Journal: Microeconomics, 14*(3), 530–557.
Prem, E., Hardman, L., Werthner, H., & Timmers, P. (2022). Research, innovation, and education roadmap for digital humanism. *The digital humanism initiative*. Vienna: TU Wien. https://dighum.ec.tuwien.ac.at/.
Prem, E., & Krenn, B. (2024). On algorithmic content moderation, 2024. In H. Werthner, C. Ghezzi, J. Kramer, J. Nida-Rümelin, B. Nuseibeh, E. Prem, & A. Stanger (Eds.), *Introduction to digital humanism* (pp. 481–494). Cham: Springer Nature.
Prensky, M., & Berry, B. D. (2001). Do they really think differently. *On the Horizon, 9*(6), 1–9.
Rolnick, D., Donti, P. L., Kaack, L. H., Kochanski, K., Lacoste, A., Sankaran, K., & Bengio, Y. (2022). Tackling climate change with machine learning. *ACM Computing Surveys (CSUR), 55*(2), 1–96.
Schiaffonati, V. (2022). Explorative experiments and digital humanism: Adding an epistemic dimension to the ethical debate, 2022. In H. Werthner, E. Prem, E. A. Lee, & C. Ghezzi (Eds.), *Perspectives on digital humanism*. Cham, CH: Springer Nature, 77 ff.
Schmölz, A. (2020). Die conditio humana im digitalen zeitalter: Zur grundlegung des digitalen humanismus und des wiener manifests (The human condition in the digital age). *MedienPädagogik: Zeitschrift für Theorie Und Praxis Der Medienbildung, 2020*, 208–234. https://doi.org/10.21240/mpaed/00/2020.11.13.X
Speith, T. (2022). A review of taxonomies of explainable artificial intelligence (XAI) methods. In *Proceedings of the ACM conference on fairness, accountability, and transparency* (pp. 2239–2250).
Staab, P. (2019). *Digitaler kapitalismus: Markt und herrschaft in der ökonomie der unknappheit (Digital capitalism: Power and governance in the economy of abundance)*.
Timmers, P. (2022). The technological construction of sovereignty, 2022. In H. Werthner, E. Prem, E. A. Lee, & C. Ghezzi (Eds.), *Perspectives on digital humanism* (pp. 213–218). Cham, CH: Springer Nature.
Vedechkina, M., & Borgonovi, F. (2021). A review of evidence on the role of digital technology in shaping attention and cognitive control in children. *Frontiers in Psychology, 12*, Article 611155.
Werthner, H. (2020). The Vienna manifesto on digital humanism. In M. Hengstschläger (Ed.), *Digital transformation and ethics* (pp. 338–357). Elsbethen: Ecowin. http://hdl.handle.net/20.500.12708/30434.
Werthner, H., Ghezzi, C., Kramer, J., Nida-Rümelin, J., Nuseibeh, B., Prem, E., & Stanger, A. (2024). *Introduction to digital humanism*. Cham: Springer Nature.
Werthner, H., Prem, E., Lee, E. A., & Ghezzi, C. (2022b). *Perspectives on digital humanism*. Cham, CH: Springer Nature.
Werthner, H., Prem, E., Lee, E. A., & Ghezzi, C. (2022a). *Perspectives on digital humanism* (p. 25). Cham, CH: Springer Nature. ff.
Zuboff, S. (2019). *The age of surveillance capitalism: The fight for a human future at the new frontier of power*. New York: Public Affairs.
Zuboff, S., Möllers, N., Wood, D. M., & Lyon, D. (2019). Surveillance capitalism: An interview with Shoshana Zuboff. *Surveillance & Society, 17*(1/2), 257–266.