



# Model of ethical analysis of digital technologies: towards true digital humanism

Cristina Díaz de la Cruz<sup>1</sup> · José Luis Fernández Fernández<sup>2</sup> · Carolina Villegas-Galaviz<sup>3</sup>

Received: 10 January 2025 / Accepted: 22 March 2025  
© The Author(s) 2025

## Abstract

This study introduces a comprehensive model for the ethical analysis of digital technologies to foster a full implementation of digital humanism. Drawing on key ethical traditions, the model acts as a framework for evaluating emerging technologies to ensure that they contribute to human development and societal well-being rather than solely driven by economic objectives. The theoretical framework integrates principles from ethical theories such as virtue ethics, deontological ethics, and consequentialism and applies them to digital technology development. The analysis emphasizes the need for transparency, accountability, and the prioritization of human dignity. The model is applied to augmented reality (AR) devices to illustrate its applicability, highlighting their potential benefits—such as enhanced education and healthcare capabilities—and the risks, including privacy concerns, digital distraction, and the exacerbation of social inequalities. Through this framework, developers and policymakers can better understand the ethical implications of digital technologies before their widespread implementation to mitigate risks and enhance positive outcomes. Ultimately, this model advocates for a balanced approach to digital innovation, where technological progress is aligned with universal human values and ethical standards, ensuring that the digital world remains a servant to human flourishing.

**Keywords** Digital humanism · Ethics · Technological development · Digital technology

## 1 Introduction

Recently, a renowned historian, Yuval Harari, and two colleagues wrote about the risks to humanity of opening some of the currently available artificial intelligence (AI) technologies to the public (2023). According to them, a thorough analysis of the dangers of using each of these technologies should have been conducted, as is done with pharmaceuticals, mass transportation, and other sectors. The technological race undertaken by large digital technology

companies entails omitting the critical analysis processes of these tools before they are presented for public use.

An example of the above-mentioned authors' concern about using certain digital technologies is reflected in the following sentence: "Social media was the first contact between A.I. and humanity, and humanity lost" (Yuval Harari et al. 2023, p. 3). Regardless of whether this analysis is accurate, we could agree that virtually any indiscriminately available technology presents certain risks (Dhirani et al. 2023).

Some of the most well-documented risks associated with new digital technologies include: privacy risks (Cavoukian et al. 2010; Solove 2011), cybersecurity risks (European Union Agency for Cybersecurity 2023), social risks (Turtle 2012), and psychological risks (Nisafani et al. 2020; Ramos-Diaz et al. 2020; Cavalcante Siebert et al. 2023; McLoughlin and Brady 2024), worsened by the risk of misinformation and fake news due to the increasing amount of information available and the difficulty in differentiating true from false information (Aghajari et al. 2023; Efstratiou and De Cristofaro 2022; Rubin 2022; Shin 2024; Zhou et al. 2023).

---

✉ Cristina Díaz de la Cruz  
cdiaz@utpl.edu.ec

José Luis Fernández Fernández  
jlfernandez@comillas.edu

Carolina Villegas-Galaviz  
cvillegasgalaviz@fordham.edu

<sup>1</sup> Universidad Técnica Particular de Loja, Loja, Ecuador

<sup>2</sup> Comillas Pontifical University, Madrid, Spain

<sup>3</sup> Fordham University, New York, USA

The increasing ubiquity of AI-based systems raises significant ethical and governance challenges, particularly concerning meaningful human control. Cavalcante Siebert et al. (2023) highlight that AI systems can generate unpredictable consequences, often misaligning human moral values and societal norms. The difficulty in attributing moral responsibility for autonomous AI decisions underscores the need for frameworks that ensure human oversight remains central. The concept of meaningful human control has emerged as a crucial principle to address these responsibility gaps, advocating for explicit moral accountability in AI design and implementation. However, operationalizing this principle remains a challenge, as it requires technical mechanisms and systemic socio-technical considerations that integrate policy, law, and human–AI interactions.

In addition to these more common risks, some authors also mention threats such as manipulation of human consciousness and behavior, distortion of thinking and culture, and the growth of new forms of criminalization of economic activity (Berg et al. 2023). There are also those who draw attention to the risk of dehumanization, spiritual lumpenization, loss of cognitive competencies, crisis of intellectual culture, and threat to creative development (Kotlyarova et al. 2021).

Given all the risks now pointed out and returning to the concern presented by Yuval et al. (2023) about the importance of testing digital technologies before being made available for public use, we should ask ourselves what the most effective and realistic way would be to perform this type of pre-testing, considering the unstoppable dynamism of the current technological race.

Some authors and institutions advocate that the precautions to be taken in the face of different technological development proposals should promote those that entail the least risk and the greatest benefits (Reijers 2016; UNESCO 2023; Wright 2011).

In this line of thought arises the concept of digital humanism, which aims to provide a secure framework for the technological development of the digital age, ensuring that technology is at the service of human beings and not the other way around (Fuchs 2022; Hofkirchner and Kreowski 2022; Hofkirchner and Kreowski 2022; Nida-Rümelin and Staudacher 2024; Lucci and Osti 2024).

The present study aims to provide a tool capable of reinforcing the evaluation of developing digital technologies based on universal ethical principles to reduce their risks and increase awareness of their need of improvement, all of this oriented to promote an effective implementation of digital humanism. To this end, this theoretical-practical study is proposed, guided by the following research questions:

- Is it possible to develop a model of ethical analysis capable of assessing the degree of digital humanism of digital technologies?
- If possible, what ethical issues should be included in the analysis to ensure the most comprehensive assessment possible?
- Finally, can this analysis be applied to existing technologies as an example?

To answer these questions the study will be presented in four parts: The first, consists of a theoretical exposition of digital humanism, based on a review of the literature on the subject, seeking to make some of its main principles explicit. The second, also based on the literature review, contains an exposition of the main ethical theories and their fundamental questions, relating them to how to guide technological development. Based on the results of these first two parts, some reflections on the complementarity between the principles of digital humanism and fundamental ethical issues will be presented. Thus, the first two research questions should have been answered. Finally, the third part presents an explanation on how to apply the ethical analysis model, to exemplify its usefulness in real cases. Specifically, the model is applied to critically examine augmented reality (AR) devices from the perspective of digital humanism.

This study bolsters on two key premises: First, on the consideration of digital humanism as a valuable proposal to establish a line of work to protect the human being from the risks of the unbridled development of digital technologies; second, on the consideration that the best way to deepen the analysis of the impact of these technologies on human nature and the construction of a better world is to rethink the critical ethical questions about each of them. In short, this model aims to provide a form of ethical analysis of digital technologies to ensure a technological development at the service of the integral well-being of people and the authentic development of societies.

## 2 Digital humanism

The concept of digital humanism emerged in response to the rapid rise of technology and its impact on human life, emphasizing human rights, social responsibility, and sustainability. This notion was further developed in the 2018 Vienna Manifesto on Digital Humanism (Mayer and Strassnig 2020; Various authors 2018), which sought to integrate technological development within a contemporary humanistic vision.

Rooted in critical theory, the manifesto promotes the Common Good as an ethical–political foundation for a democratic and inclusive society that safeguards human dignity and rights. It resulted from interdisciplinary

collaboration among scholars and practitioners in philosophy, computer science, law, and political science and has been endorsed by organizations such as the European Commission and the UN Secretary-General's High-level Panel on Digital Cooperation. Acknowledging the dual nature of technological progress—its opportunities and risks—the manifesto calls for a shift beyond efficiency and profit toward human-centered, ethical, and sustainable approaches to technology.

The Vienna Manifesto was conceived to ensure that technological progress serves humanity and well-being rather than being dictated by economic and technocratic imperatives. Without ethical and moral considerations, unchecked automation could pose risks to democratic societies by fostering self-referential and uncontrolled technological systems. The initiative aimed to define strategies that align digitalization and AI with ethical and democratic values, placing human dignity at the core of technological development.

Its core principles include placing technology at the service of humanity, respecting democratic values and human rights, ensuring transparent and accountable digital governance, regulating AI democratically, and preventing unchecked automation from undermining human autonomy. Since its publication, digital humanism has gained traction as a counterbalance to profit-driven technological development, advocating for responsible digitalization that enhances human life rather than replacing or diminishing it (Fuchs 2022; Werthner et al. 2023).

A systematic review of AI-related ethical considerations in research from 1986 to 2021 (Giarmoleo et al. 2024) analyzed 309 articles and identified two primary categories of ethical concerns: those arising from AI design and those emerging from human–AI interactions. The study underscores the need to ensure that technological development aligns with human well-being, addressing risks while upholding privacy, democracy, human dignity, and ethical values.

Within the framework of digital humanism, six key principles have been selected for the ethical analysis in this study: human-centered design, value-sensitive design, responsible innovation, technological determinism, ethics, and democratic governance (Broadbent et al. 2015; Burr and Floridi 2020). These principles serve as a foundation for evaluating the ethical implications of technological development, as they encompass broader concerns frequently discussed in this field. A more detailed explanation of each principle follows:

*Human-centered design* focuses on designing technologies centered on the needs and desires of human beings. This involves engaging users in the design process, understanding their needs and preferences, and prioritizing their well-being over technological advancement or profit

(Fernandez-Carames and Fraga-Lamas 2018; Shneiderman 2020).

*Value sensitive design* is a theoretically grounded approach to the design of technology that accounts for axiology and human values in a comprehensive and principled manner throughout the whole design process (Friedman et al. 2006).

*Responsible innovation* considers the social and ethical implications of new technologies before they are developed and implemented. This includes assessing potential risks and benefits, engaging stakeholders in decision-making, and mitigating negative consequences (Owen et al. 2013; Van den Hoven et al. 2015). Responsible innovation seeks to align technological development with societal needs by integrating ethical considerations into design and implementation. However, one of its main criticisms is defining what constitutes “responsibility” in a rapidly evolving digital landscape. The interests of developers, corporations, policymakers, and users often diverge, making it difficult to establish a universally accepted framework. In addition, critics argue that responsible innovation can be selectively applied as a rhetorical device rather than a concrete regulatory mechanism. A multi-stakeholder approach incorporating iterative ethical assessments and transparent accountability mechanisms is necessary to prevent the misuse or dilution of responsible innovation principles.

*Technological determinism* acknowledges that technology can shape society and culture, while *social determinism* defends that society shapes technology. Considering both positions, it is critical to assess social and cultural contexts in which technology is developed and used to design technologies that align with societal values and promote human flourishing (Borgmann 1984; Winner 2010). Moreover, the notion of technological determinism—the idea that technology drives societal change inevitably and autonomously—warrants further critical scrutiny. While technological advancements undeniably shape social, economic, and political structures, attributing change solely to technology risks oversimplifying the complex interplay between human agency and innovation. A more balanced perspective, often referred to as soft determinism, recognizes that while technology influences societal development, human decision-making and policy interventions shape its trajectory. This perspective underscores the importance of proactive ethical frameworks that do not passively react to technological change but actively guide its course in alignment with human values.

*Ethics* encompasses a broad tradition of reflections on the good and evil, also on the rationale and consequences of human actions. In the context of digital humanism, it is generally present in the form of codes of ethics that typically include principles such as transparency, respect for privacy,

and consideration of the social and environmental impacts of technology (Google 2022; Pugh 2009). Many professional organizations and technology companies have developed ethical codes of conduct to guide their work (Becker et al. 2023).

*Democratic governance* of technology emphasizes the importance of democratic participation and decision-making in the development and governance of technology. This involves ensuring that diverse voices and perspectives are heard in technology development and implementation and that technology's benefits and risks are distributed fairly across society (van Dijk 2020). Nevertheless, democratic governance in digital technologies presents both opportunities and limitations. While it aims to foster inclusivity and participatory decision-making, the feasibility of ensuring broad representation in digital policy remains a significant concern. Technological governance is often dominated by corporations and governments with asymmetric power over data and infrastructure, raising questions about how democratic ideals can be upheld in practice. Some scholars argue that democratizing technology requires more than just inclusive dialogue; it demands structural reforms that redistribute power among stakeholders. One potential approach is the establishment of independent oversight bodies that include ethicists, civil society representatives, and technologists to mediate ethical concerns and ensure that governance structures genuinely reflect the interests of diverse communities.

All these principles of digital humanism respond, in a certain way, to fundamental ethical questions that have always been present in the theoretical and practical discussion about what is good and convenient for human beings and what rather degrades or potentially destroys our societies. For this reason, the discussion based on principles has been moved to a broader ethical approach, which will be explained below. Several ethical considerations will be presented, followed by a discussion of the relationship between universal ethical issues and the fundamental principles of digital humanism.

### 3 Universal ethical questions

#### 3.1 Ethical foundations

As has just been pointed out, the need to address the ethical dimension of technological development in general, and of digital technologies in particular, is a generally accepted requirement. In this sense, many initiatives expressly appeal to ethics as a kind of intellectual antidote and condition of possibility (not only to avert a potential dystopia presented by proposed scenarios as transhumanism, singularity, posthumanism, *mort de la mort*, from the development

of the NBICs), but, above all, to deploy a technological development in favor of humanity and the human. That means, a technology that places people at the center and that can contribute to the emergence of digital humanism (Jobin et al. 2019; Larsson 2020).

In this respect, it is important not to forget the teaching of that double ethical axiom, which states that although it is not always technically possible to achieve everything that at a given historical moment can be considered desirable from a moral point of view, it is also true, from the other point of view, that not everything technically possible at a given moment is always and at the same time ethically desirable and morally good (Fernández Fernández 2022).

In any case, the key to discernment about what constitutes, at the same time, a technological advance and a parallel growth in humanity should probably be sought in a triple instance of philosophical reflection:

- (1) From a coherent and universalizable Anthropology accompanied by a sociological paradigm, able to shed light on what kind of human development is intended to be promoted.
- (2) From an Axiology capable of proposing a hierarchical table of values with the greatest possible logic and objectivity, establishing a parameter for prioritizing some values among other a priori.
- (3) From a well-founded Ethics, which takes the form of guidelines for action, including different perspectives and making possible the reduction of undesired risks and fostering positive consequences. Furthermore, this philosophical-moral approach to the implications of digitalization on humanity already has a long history (Baker-Brunnbauer 2021; Floridi and Taddeo 2016; Martin et al. 2019; Mittelstadt et al. 2016).

As a result of the convergence between, on the one hand, the guidelines for action and reports produced by various professional bodies and regulatory entities (European Commission 2019) and, on the other, academic reflection, the emergence of a kind of common ethical factor in relation to digital technologies is becoming increasingly evident. The ethical requirements for the development of a responsible digital technology are taking shape in a kind of adaptation of the usual series of principles of the different Applied Ethics, which have their origin, to a large extent, in the matrix of Bioethical reflection itself (Jonsen 1998; Toulmin 1982). Such are the appeals to the traditional principles of non-maleficence, beneficence, autonomy, and justice, to which others have been added, complementary to those specific to the digital context and the new moral demands of cyber society. In addition to the principles already mentioned above, it is worth highlighting some more specific aspects, such as explainability, human control, and supervision of

algorithms and AI systems, as well as the principles of transparency, responsibility, solidarity, and sustainability (Fernández Fernández 2021).

Although the list of ethical principles tends to be presented as recurrent and homogeneous, the perspectives from which we seek to provide a foundation for those principlist proposals point to a philosophical-moral pluralism in which various methodological approaches and different ethical theories find a place. This plurality is justified by the fact that some approaches are more appropriate than others for shedding light on the different aspects of the human act (consciousness, freedom, will) or of moral action (objectives and aims; intentions of the agent; means employed; circumstances in which the action is carried out; results obtained or foreseeable, as well as any other issues like duties or values, that may be relevant in a given context, and the like).

It is at this point that the plurality of ethical theories—for example, Virtue Ethics (Walker and Ivanhoe 2007), Duty Ethics (Kant 2008), Utilitarian Ethics (Bentham and Mill 1973; Sen 1979), Value Ethics (Gracia 2013; Hartman 1932; Méndez 2023; Scheler 1973), together with other complementary approaches and perspectives such as the Weberian appeal to the difference between the ethics of conviction and the ethics of responsibility (Jonas 1984; Weber 2021)—serve, on the one hand, as guides for deliberation in discerning between different axiological proposals or different possibilities of action and, on the other hand, as elements to reasonably justify the decisions that might have been already adopted in a given context (Broad 1930).

In general terms, it could be said that attempts to justify moral norms can be classified by reference to two main aspects. One would respond to a teleological conception, and the other is oriented toward deontological positions. Teleological ethics (*telos*: end) identify as good the aspiration to happiness or an equivalent category, and from this assumption, they try to orient humanly correct conduct. The supporters of deontologism (*deon*: duty), starting from Kant (1785), warn that what is genuinely human is not the desire for happiness—shared, moreover, with the rest of the animals—, but rather autonomy, the fact of being capable of imposing duties on oneself. From this perspective, it insists on the obligation to respect and promote the autonomous human person as an end in itself.

For its part, consequentialist ethical theories determine the rightness of an action based on its outcomes, advocating for the maximization of overall welfare from an agent-neutral perspective. Utilitarianism, a key form of consequentialism, asserts that actions are right if they maximize well-being. Some debates extend consequentialist evaluation beyond actions to rules, motives, character traits, and institutions (Andersen 2022; McElwee 2020).

A common critique is that consequentialism neglects reasons for action, though some argue it can account for them, adding depth to its framework (Woodard 2020). Subjective consequentialism struggles with assigning expected values to unforeseen outcomes, leading to deontic indeterminacy (Jay 2020). Efforts to generalize it, such as rule consequentialism, have conceptual limitations, while functional decision theory offers a more robust approach (Greene and Levinstein 2020). In addition, consequentialism acknowledges that actions often produce unintended consequences, a phenomenon known as counterfinality, where uncoordinated individual actions result in collective irrationality (Latypov 2021; Moroni 2012).

Thus, consequentialist analysis urges careful consideration of both intended and unintended effects of large-scale digital technology use.

In contrast, axiological ethics, within the philosophy of values, posits that humans intuit values across four levels: economic and vital, ethical, esthetic, and transcendent. Ethical values, in particular, can be structured by height and strength, as proposed by Scheler (1973) and Hartman (1932), and later integrated coherently by Méndez (2023).

Rather than mutually exclusive paradigms, ethical theories offer complementary perspectives for evaluating digital technologies. Ethical deliberation should integrate multiple frameworks to guide praxis effectively. Three core axioms emerge as guiding principles: (1) ethical ends should be pursued without violating deontological principles (e.g., human dignity, universalizability) or consequentialist considerations regarding expected outcomes. (2) Deontological ethics should prevail unless it contradicts eudaimonic goals or leads to greater harm than good. (3) Actions should maximize collective well-being—especially for the most vulnerable—while respecting dignity, universalizability, and human potential.

### 3.2 The underlying vision of the ideal human being (anthropological model)

When we study human behavior and develop some kind of ethical theory, there is always present, implicitly, or explicitly, a way of understanding the human being. Thus, we could say that every ethical theory is underlain by an anthropological theory, since ethical behavior is always oriented towards a vision of the good and of that which most signifies and develops human potential (Bainton 2012; Fontrodona and Sison 2006; Melé and Cantón 2015).

With this premise in mind, we can ensure the relevance of not only answering questions about the means, values, and ends of each technology but also about how this proposal intends to promote integral human welfare and development. To this end, it is necessary to make explicit what model of the person underlies a given technological proposal.



Although it may seem difficult for the author of a given technology to state what his or her anthropological vision is, in reality, this can often be partially deduced from the answers given to the above questions about purpose, principles, values, and objectives.

To identify the visions of the human being behind different technologies, the following steps can be taken: First, examining the purposes and goals of current technological development proposals. For example, asking about what these proposals are aiming to achieve in terms of enhancing human capabilities, addressing societal needs, or transforming human experiences. Second, analyzing the values and beliefs embedded in these proposals. For example, examining our beliefs about what the core assumptions about human nature, behavior, and potential are, that underlie these technological trends. Third, looking for recurring themes in technological discourse. For example, questioning if there are common visions of the ideal human being that emerge across different technological domains, such as AI, biotechnology, or sustainability.

This analysis can provide valuable insights into the complex interplay between technology and society, and how different technological trends may shape our understanding of what it means to be human in a rapidly evolving technological landscape.

To make this philosophical exercise even more explicit, some of the most common visions of human beings in the analysis of technological development at the historical level will be presented.

Humanism is a philosophical perspective that strongly emphasizes on human values, ethics, and dignity. Technologies developed with a humanistic vision aim to enhance the well-being and quality of life for individuals. For example, medical technologies are often designed with the goal of alleviating human suffering and improving health (Davydov 2020; Doukas et al. 2022; Kozlerek 2021). Most humanist proposals are based on a dualistic vision of the human being, recognizing the importance of a non-material dimension (soul, spirit) as a fundamental element for promoting true human development (Villegas-Galaviz et al. 2021).

From a philosophical anthropology that values social justice and equity, technologies may be developed to address societal inequalities and promote equal access to resources and opportunities. This can include technologies that aim to bridge the digital divide or promote economic and educational equity (Holeman and Kane 2020). This approach is a complement to the humanistic proposition.

Some technological developments are shaped by an environmentalist or ecologically minded philosophical anthropology. From this perspective, an ideal human being is one who coexists harmoniously with nature. Sustainable technologies and practices are designed to minimize the

environmental impact of human activities (Abram 2012). This approach reduces the differences between human beings and other species of living beings present in the cosmos, thus not considering some specific characteristics and potentials of humankind.

On the other hand, recently, the transhumanist proposal has gained greater strength in relation to technological development. Transhumanism is a philosophical movement that advocates for using technology to enhance human capabilities and transcend the limitations of the human condition. Technologies such as genetic engineering, AI, and human–computer interfaces are often developed with the goal of improving or augmenting human abilities (Besnier 2013; Hopkins 2012). Generally, transhumanist currents are based on a monistic vision of the human being, fundamentally based on materialism, and propose the improvement of the human being only from a physical and cognitive perspective, ignoring the spiritual dimension (Cole-Turner 2011).

In contrast to transhumanism, posthumanism challenges traditional notions of the human and often emphasizes the blurring of boundaries between humans and technology. Technologies informed by posthumanist thought may explore concepts like cyborgs, AI, and the integration of technology into the human experience (Ahn 2023; Buchanan-Oliver and Cruz 2015). The anthropological conception of posthumanists is dynamic since they consider humanity a species in continuous evolution.

When examining the relationship between philosophical anthropology and technology, one can consider how different philosophical frameworks influence the goals, values, and ethical considerations behind technological development. It is important to explore how these philosophical perspectives impact the choices made in the creation and implementation of technology and how they shape our understanding of what it means to be an ideal human being in a technological society.

### 3.3 The underlying vision of the ideal society (sociological model)

Just as every technological proposal is based on a specific anthropological model, it is also grounded in a particular model of society. This section examines various sociological perspectives that shape the understanding of technological development, as well as the main sociological paradigms that have historically influenced this field.

Utopianism envisions an ideal and often perfect society characterized by harmony, justice, and a utopian community. Utopian thinkers propose alternative social structures to eliminate social ills and create a better world (Moro 1516). Utopians may see technology as a means to achieve their ideal society. They might envision technological

advancements that promote social equality, improve living conditions, and enhance human well-being. The use of technology in a utopian society is likely guided by a desire to eliminate poverty, suffering, and social injustice.

Marxism envisions a classless society where the means of production are collectively owned, emphasizing the role of class struggle in historical development. It seeks to overcome capitalist exploitation (Marx 2004; Marx and Engels 2019). Marxists often view technology as a tool for social transformation. They may see technological development as a means to liberate the working class, reduce inequality, and increase economic efficiency. In a Marxist society, technology is expected to serve the common good, eliminating profit motives in production.

Liberalism promotes individual freedom, equality, and limited government intervention in society. It envisions a society where individual rights and freedoms are protected (Hayek 1969; Mill and Mill 1966). Liberals typically support technological development that enhances individual liberty and economic prosperity. They may advocate for technological innovation in areas such as communication, personal autonomy, and economic efficiency. Technological progress is seen as a means to empower individuals and expand personal freedoms.

Conservatism emphasizes tradition, authority, and the preservation of existing social structure (Burke 2016; Kirk 1987). It values stability and order in society. Conservatives may approach technological development cautiously, prioritizing the preservation of cultural and social traditions. They are likely to support technologies that maintain social stability and reinforce existing institutions. Technological progress should align with conservative values and not disrupt the social order.

Anarchism envisions a society without government and hierarchy, emphasizing individual autonomy and voluntary cooperation. Anarchists may see technology as a tool for decentralization and empowerment of individuals and communities (Kropotkin 2015; Malatesta 1909). They are likely to support technologies that enable direct democratic participation and self-governance. Technological development in an anarchist society is guided by the principles of autonomy and cooperation.

The vision of authors regarding technological development within these proposals can vary, but it is generally influenced by the core principles and goals of each philosophical current. Technology is often seen as a means to achieve the broader societal ideals advocated by each school of thought, whether it be social justice, individual freedom, or community cooperation.

For all these reasons, it is important to know the vision of society that underlies a technological proposal to gain a deeper understanding of the social consequences that a given technology may have.

### 3.4 Relationship between principles of digital humanism and main ethical approaches

Returning to the previously presented principles of digital humanism, these can be connected to the outlined ethical approaches as follows:

To carry out the human-centered design, it is necessary to establish, first, the anthropological model that is seen as ideal. The human needs and desires to be defended and promoted are a function of the anthropological vision being considered.

To implement a value-sensitive design, it is necessary to define the axiological framework on which the design is carried out, establishing which values have priority over others that can be neglected.

Responsible innovation must include the analysis derived from ethical questions about the purposes, consequences and principles assumed in advance by a technological proposal, only then can the personal and social implications of its implementation be determined.

If a technological determinism's position on the significant influence of technology in the construction of anthropological and societal models is considered, it becomes more apparent the importance of previously establishing, at a theoretical level, what type of person and society should be promoted. Therefore, it is necessary to determine in advance what anthropological and sociological models underlie each of the proposals of digital technologies.

In relation to the principle of democratic governance, its requirements are directly related to the importance of taking into account different ethical voices or perspectives. As mentioned above, our model advocates the need to respond to different ethical approaches to allow for a more complete analysis.

Similarly, as shown with some principles of digital humanism, other principles can be related to the main ethical issues discussed. It is important to note that principles such as autonomy, non-maleficence, equity, and explainability have not been addressed, as these have already been briefly introduced within the principlist ethical tradition.

### 3.5 Proposed ethical evaluation table

Based on the considerations presented above, Table 1 summarizes the ethical aspects that must be analyzed for the development of a specific digital technology.

It is intended, therefore, that the people and organizations responsible for developing a specific digital technology can respond to the questions posed below by filling out the table in question. Once the answers about a specific digital technology are known, a group of ethics experts will be able to make the pertinent observations before the technology is given to public use.

**Table 1** Questions for ethical analysis of digital technologies

Name of digital technology	Parameter	Answer
	1. Provide a description of the technology	
	2. What is the ultimate purpose or intention (finality) of this digital technology?	
	3. What are the means used (how) for its development?	
	4. What are the most important principles and values (a priori criteria/axiology) taken into account for its creation?	
	5. What are the desired or positive consequences (goods) expected with the use of this technology?	
	6. What are the risks and undesired negative consequences (evils) that could not be avoided?	
	7. What is the vision of the ideal human being (anthropological model) underlying this technological proposal?	
	8. What is the vision of the ideal society (sociological model) underlying this technological proposal?	

Similar to what currently happens with ethics committees in research with human beings, which are mandatory for most projects that entail any risk to people, this proposed tool could serve as a basis for the creation of ethical committees for the development of digital technologies. This is a form of practical implementation of digital humanism.

In the following section, the tool will be applied to a real case as a practical exercise, followed by a brief ethical analysis.

#### 4 Application example: augmented reality devices

This section examines the application of the ethical analysis model to a specific technology: augmented reality (AR) devices, focusing on the current release of spatial computing smart glasses. These devices integrate digital content into the physical environment, enabling users to interact with online content through their eyes, hands, and voices, without the need for direct physical engagement. This allows users to experience augmented content while maintaining their physical surroundings, visible only to them.

To apply the model to augmented reality devices, an analysis will be conducted from the perspective of external examiners evaluating the ethical and societal implications. The analysis is relevant to those designing and deploying digital technologies, though for illustrative purposes, this example will focus on AR devices.

Regarding the technology's description, Turner, (2022) defines AR technologies as tools that “augment” perception by superimposing virtual objects onto the user's visual field. Examples include the Apple Vision Pro, Meta Quest Pro, and XREAL Air, which blend virtual and physical spaces, transforming interactions (Turner 2022). For example, an individual wearing Apple Vision Pro in a coffee shop could read emails while drinking coffee, without the need for a laptop, tablet, or smartphone.

On the technological development front, companies are aiming to offer users more immersive experiences. Moving from desktop computers to smartphones, now smart glasses are proposed as the next step, merging physical and digital worlds. These devices represent a natural progression, extending beyond wearables like smartwatches to a fully integrated digital experience.

The possibility of spatial computing is rooted in the vast data generated by users and corporations, along with permanent internet connectivity, which makes constant digital engagement possible. Though the principles and values of technology companies cannot always be fully known, their pursuit of technological innovation, user engagement, and profit are evident. These motives are not inherently harmful, but ethical concerns arise when such developments lead to user harm, excessive engagement, or profit-driven practices that conflict with broader ethical frameworks like Freeman's Stakeholder Theory (Freeman 2010), which calls for creating value for all stakeholders.

Regarding the positive consequences of AR technologies, it is crucial to recognize their potential to address societal challenges. For example, AI and AR technologies are already being explored under initiatives like “AI for Good” to enhance sectors such as healthcare and education. AR devices, for instance, could revolutionize education by offering immersive historical experiences, facilitating a deeper understanding of events. Turner (2022) also highlights the potential of AR to enhance cognitive capacities, particularly in therapies for cognitive disorders like Alzheimer's, or through medical applications such as surgical planning and behavioral health improvement (Apple Newsroom 2024).

However, the risks of AR devices must also be considered. Turner (2022) identifies three key epistemic issues: digital distraction, digital deception, and digital divergence. These concerns are amplified by surveillance capitalism (Zuboff 2019), a data-driven model that exploits user attention. AR devices exacerbate digital distraction by overlaying ads and reminders on the user's visual field, while their ability to



track eye movement raises privacy issues. Digital deception occurs when users are misled by false or manipulated digital content, and digital divergence refers to the creation of filter bubbles that isolate users from reality.

A comprehensive ethical analysis of AR devices must include frameworks such as deontology, virtue ethics, and utilitarianism to evaluate duties, consequences, and the broader social impact. It is also necessary to address the anthropological and sociological models embedded in the technology. The humanist vision underlying AR devices should be questioned: do these devices prioritize human values and dignity? Do they enhance individual well-being, or do they exacerbate mental health issues such as anxiety and depression?

Furthermore, the affordability of these devices poses an issue, as their high cost limits access to wealthier social classes, thus increasing social inequality (Roose 2024). Environmental concerns are also relevant, as the AI systems underlying AR devices contribute significantly to energy consumption (Calvert 2024).

The transhumanist and posthumanist implications of AR technologies must also be considered. By enhancing human cognitive abilities, these devices align with ideologies that view human nature as evolving and mutable. Finally, the sociological model behind AR technologies should be assessed for its societal impact. These devices may contribute to both utopian visions of self-constructed realities and capitalist exploitation, raising concerns from Marxist, liberal, and conservative perspectives. They also challenge traditional social structures, potentially fostering anarchist ideals or threatening individual freedoms in the face of digital manipulation.

A key challenge in implementing digital humanism lies in its scalability across diverse contexts. While the proposed model provides a structured approach, it may not always align with local cultural, legal, and economic realities. Future research should explore adaptive ethical models that respect local values while upholding universal humanistic principles. Comparative case studies could inform a more context-sensitive approach to digital humanism.

Furthermore, practical challenges exist in operationalizing ethical analysis. Policymakers may struggle to enforce standards in a rapidly evolving technological landscape, while developers may lack clear guidelines for integrating ethics into product design. Multidisciplinary governance bodies, including ethicists, technologists, and policymakers, could help bridge this gap. In addition, embedding ethics early in the development process, such as through value-sensitive design, may facilitate the alignment of digital technologies with humanistic values.

Finally, the limitations of ethical frameworks in anticipating unintended consequences must be recognized. While AR devices offer numerous benefits, such as

enhancing education and healthcare, they may also exacerbate social inequalities and privilege those with access to advanced digital infrastructure. Continuous evaluation of ethical guidelines will be necessary to ensure that digital humanism evolves in response to the challenges and opportunities presented by emerging technologies.

## 5 Conclusions

The ethical analysis model proposed in this study aims to ensure that emerging digital technologies align with the principles of digital humanism, emphasizing human dignity and social well-being over economic interests. As digital technologies continue to advance, particularly with innovations such as AI and augmented reality (AR), this approach is increasingly critical. The model provides a structured framework for evaluating these technologies, anticipating both their benefits and risks, and offering a path forward for responsible development.

Digital humanism represents a transformative shift in how technology is conceptualized. Rather than focusing solely on efficiency or profit maximization, it advocates for creating technologies that improve human life and promote principles of equity, justice, and privacy. As technologies become more embedded in all aspects of life, the goal is to enhance human capabilities and contribute to the creation of more just societies.

The ethical model facilitates the evaluation of digital technologies, identifying potential risks and offering a proactive approach to mitigate harm while maximizing benefits. Its application before widespread implementation allows for a more responsible trajectory for technological advancement, helping to safeguard societal well-being in the face of rapid technological growth. As we look toward the future, the model will be increasingly relevant in guiding innovations across industries, particularly in fields where new opportunities for positive societal change coexist with significant ethical challenges.

The model's emphasis on transparency is particularly important in the context of evolving digital technologies. In the wake of privacy scandals and the erosion of trust in digital platforms, transparency will be key to rebuilding public confidence in new technologies. By advocating for clearer communication about technological purposes, principles, and consequences, the model paves the way for greater accountability in future tech development, ensuring that societal benefits are prioritized over short-term corporate gains.

However, the ethical model is not designed to stifle innovation. On the contrary, it seeks to guide the development of transformative technologies by ensuring they align with ethical standards. As technologies continue

to mature, this model can play a crucial role in ensuring that innovation remains rooted in human values. Looking forward, its application could extend into new areas such as virtual and augmented workplaces, further enhancing human productivity and connectivity but also requiring careful consideration of new ethical dilemmas.

While the study reveals significant potential in AR for sectors such as education and healthcare, it also identifies risks that must be addressed proactively. As these technologies evolve, it will be necessary to focus on reducing digital distraction, safeguarding privacy, and addressing issues of social exclusion. Future research and ethical deliberations must continue to focus on mitigating these risks, ensuring that digital advancements do not exacerbate existing social divides but rather contribute to inclusivity.

The importance of interdisciplinary collaboration in the ethical analysis of emerging technologies cannot be overstated. As digital technologies become more pervasive, their impact will span a wide range of sectors, including psychology, sociology, and economics. The continued involvement of technologists, ethicists, sociologists, and policymakers will be essential in fostering an ongoing dialogue that ensures digital humanism remains relevant and adaptable to the fast-evolving technological landscape.

While the proposed ethical analysis model aims to align digital innovation with human-centered values, it is essential to consider the potential challenges and counterarguments. One critique concerns the balance between ethical oversight and technological progress. Some argue that imposing strict ethical frameworks might slow down innovation, as it could add bureaucratic hurdles that delay development. Technology companies, especially those in competitive markets, may view ethical compliance as an obstacle to progress. In addition, the fast pace of digital transformation means that ethical guidelines could become outdated before they are fully implemented. Addressing this concern will likely require a flexible, iterative approach to ethical assessment—one that balances accountability with adaptability, ensuring that ethical scrutiny does not hinder innovation.

Another challenge involves the variability of ethical principles across cultures and industries. While digital humanism advocates for universal values like human dignity and well-being, ethical priorities differ among regions and sectors. For example, what is considered an essential safeguard in one society might be viewed as a hindrance to technological or economic advancement in another. Future iterations of the framework could incorporate modular or sector-specific guidelines to allow for contextual adaptations while preserving core ethical principles. Engaging a broader range of stakeholders—ethicists from diverse cultural backgrounds and representatives from various technological

sectors—would enrich the discourse and help ensure the model remains both globally relevant and practically viable.

In conclusion, the ethical analysis model presented in this study offers a practical tool to help ensure that digital technologies are developed in alignment with the principles of digital humanism. By promoting transparency, accountability, and human well-being, the model can contribute to responsible technological advancement, ensuring that digital development serves humanity rather than diminishing it.

**Author contributions** CDC conceptualized the main ideas of the article and organized the distribution of its sections. CDC and JLF collaborated on writing the philosophical framework, ensuring a coherent and robust theoretical foundation. CVL contributed to the development of the practical example, enriching the article with relevant and illustrative applications. All three authors participated in reviewing and editing the manuscript, ensuring the clarity, accuracy, and consistency of its final version.

**Data availability** No datasets were generated or analyzed during the current study.

## Declarations

**Conflict of interest** The authors declare no competing interests.

**Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

## References

- Abram D (2012) *The spell of the sensuous: perception and language in a more-than-human world*. Vintage
- Aghajari Z, Baumer EPS, DiFranzo D (2023) Reviewing interventions to address misinformation: the need to expand our vision beyond an individualistic focus. *Proc ACM Hum-Comput Interact* 7:1–34
- Ahn S (2023) Humanism contra post-humanism. *Hist Mater* 31(1):63–92. <https://doi.org/10.1163/1569206x-20232009>
- Andersen MM (2022) Consequentialism. In: *Oxford research encyclopedia of politics*
- Apple Newsroom (2024) Apple Vision Pro unlocks new opportunities for health app developers. <https://www.apple.com/newsroom/2024/03/apple-vision-pro-unlocks-new-opportunities-for-health-app-developers/>
- Bainton BR (2012) Applied ethics: anthropology and business. *Int J Bus Anthropol* 3(1):114–133

- Baker-Brunnbauer J (2021) Management perspective of ethics in artificial intelligence. *AI Ethics* 1:173–181
- Becker SJ, Nemat AT, Lucas S, Heinritz RM, Klevesath M, Charton JE (2023) A Code of Digital Ethics: laying the foundations for digital ethics in a science and technology company. *AI Soc* 38:2629–2639
- Bentham J, Mill JS (1973) *The utilitarians*. Anchor Press/Doubleday
- Berg J, Green F, Nurski L, Spencer DA (2023) Risks to job quality from digital technologies: are industrial relations in Europe ready for the challenge? *Eur J Ind Relat* 29(4):347–365. <https://doi.org/10.1177/09596801231178904>
- Besnier J-M (2013) Humanity: a failed experiment? Versions of trans-humanism. *Futuribles Anal Prospect* 397:5–20
- Borgmann A (1984) *Technology and the character of contemporary life: a philosophical inquiry*. University of Chicago Press
- Broad CD (1930) *Five types of ethical theory*. Harcourt Brace & Co.
- Broadbent S, Dewandre N, Ess CM, Floridi L, Ganascia J-G, Hildebrandt M, Laouris Y, Lobet-Maris C, Oates S, Pagallo U, Simon J, Thorseth M, Verbeek P-P (2015) *The onlife manifesto*. In: Floridi L (ed) *The onlife manifesto: being human in a hyperconnected era*. Springer, pp 7–13. [https://doi.org/10.1007/978-3-319-04093-6\\_2](https://doi.org/10.1007/978-3-319-04093-6_2)
- Buchanan-Oliver M, Cruz AGB (2015) The spectre of posthumanism in technology consumption: the death of the human? In: *Death in a consumer culture*. <https://doi.org/10.4324/9781315726175-31>
- Burke E (2016) *Reflections on the revolution in France*. In: *Democracy: A Reader*. Columbia University Press, pp 234–238
- Burr C, Floridi L (2020) The ethics of digital well-being: a multidisciplinary perspective. In: Burr C, Floridi L (eds) *Ethics of digital well-being: a multidisciplinary approach*. Springer, pp 1–29. [https://doi.org/10.1007/978-3-030-50585-1\\_1](https://doi.org/10.1007/978-3-030-50585-1_1)
- Calvert B (2024) AI already uses as much energy as a small country. It's only the beginning. <https://www.vox.com/climate/2024/3/28/24111721/climate-ai-tech-energy-demand-rising>
- Cavalcante Siebert L, Lupetti ML, Aizenberg E, Beckers N, Zgonnikov A, Veluwenkamp H, Abbink D, Giaccardi E, Houben G-J, Jonker CM, van den Hoven J, Forster D, Lagendijk RL (2023) Meaningful human control: actionable properties for AI system development. *AI Ethics* 3(1):241–255. <https://doi.org/10.1007/s43681-022-00167-3>
- Cavoukian A, Taylor S, Abrams ME (2010) Privacy by design: essential for organizational accountability and strong business practices. *Identity Inf Soc* 3:405–413
- Cole-Turner R (2011) Transhumanism and transcendence: Christian hope in an age of technological enhancement, pp 1–219. <https://www.scopus.com/inward/record.uri?eid=2-s2.0-84903255303&partnerID=40&md5=429a1ff7866f7f860657b59c939f1f2d>
- Davydov DA (2020) The crisis of humanism and emerging post-anthropocentric epoch: a personalistic view. *Changing Soc Personal* 4(2):220–247. <https://doi.org/10.15826/csp.2020.4.2.099>
- Dhirani LL, Mukhtiar N, Chowdhry BS, Neue T (2023) Ethical dilemmas and privacy issues in emerging technologies: a review. *Sensors* 23:1151
- Doukas DJ, Ozar DT, Darragh M, de Groot JM, Carter BS, Stout N (2022) Virtue and care ethics & humanism in medical education: a scoping review. *BMC Med Educ*. <https://doi.org/10.1186/s12909-021-03051-6>
- Efstratiou A, De Cristofaro E (2022) Adherence to misinformation on social media through socio-cognitive and group-based processes
- European Commission (2019) *Ethics guidelines for trustworthy AI*. <https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai>
- European Union Agency for Cybersecurity (2023) Identifying emerging cyber security threats and challenges for 2030. <https://doi.org/10.2824/117542>
- Fernández Fernández JL (2021) Hacia el Humanismo Digital desde un denominador común para la Ciber Ética and la Ética de la Inteligencia Artificial. *Disputatio Philos Res Bull* 10:107–130
- Fernández Fernández JL (2022) Ethical considerations regarding biases in algorithms. In: E. A. I. Ethics, H. Education. G. Practice, & G. for Educators (eds) Chia. *Globethics.net*, pp 71–90
- Fernandez-Carames TM, Fraga-Lamas P (2018) A review on human-centered IoT-connected smart labels for the Industry 4.0. *IEEE Access* 6:25939–25957. <https://doi.org/10.1109/ACCESS.2018.2833501>
- Floridi L, Taddeo M (2016) What is data ethics? *Philos Trans R Soc* 374:1–5
- Fontrodona J, Sison AJG (2006) The nature of the firm, agency theory and shareholder theory: a critique from philosophical anthropology. *J Bus Ethics* 66(1):33–42. <https://doi.org/10.1007/s10551-006-9052-2>
- Freeman RE (2010) *Strategic management: a stakeholder approach*. Cambridge University Press. <https://doi.org/10.1017/CBO9781139192675>
- Friedman B, Kahn P, Borning A (2006) Value sensitive design and information systems. In: Zhang P (ed) *Human–computer interaction in management information systems: foundations*. M. E. Sharpe, Inc., pp 1–27
- Fuchs C (2022) What is digital humanism? In: *Digital humanism*. <https://doi.org/10.1108/978-1-80382-419-220221003>
- Giarmoleo FV, Ferrero I, Rocchi M, Pellegrini MM (2024) What ethics can say on artificial intelligence: insights from a systematic literature review. *Bus Soc Rev* 129(2):258–292. <https://doi.org/10.1111/basr.12336>
- Google (2022) 2022 AI principles progress update. <https://ai.google/static/documents/ai-principles-2022-progress-update.pdf>
- Gracia D (2013) *Construyendo valores*. Triacastela
- Greene P, Levinstein BA (2020) Act Consequentialism without free rides. *Philos Perspect* 34(1):88–116. <https://doi.org/10.1111/phpe.12138>
- Hartman N (1932) *Ethics*. George Allen & Unwin
- Hayek FA (1969) *The road to Serfdom*, by Friedrich A. Hayek. University of Chicago Press
- Hofkirchner W, Kreowski H-J (2022) Digital humanism: how to shape digitalisation in the age of global challenges? <https://doi.org/10.3390/proceedings2022081004>
- Holeman I, Kane D (2020) Human-centered design for global health equity. *Inf Technol Dev* 26(3):477–505. <https://doi.org/10.1080/02681102.2019.1667289>
- Hopkins PD (2012) Why uploading will not work, or, the ghosts haunting transhumanism. *Int J Mach Conscious* 4(1):229–243. <https://doi.org/10.1142/S1793843012400136>
- Jay C (2020) Subjective consequentialism and the unforeseeable. *Utilitas* 32(1):33–49. <https://doi.org/10.1017/S095382081900030X>
- Jobin A, Ienca M, Vayena E (2019) The global landscape of AI ethics guidelines. *Nat Mach Intell* 1(9):389–399
- Jonas H (1984) *The imperative of responsibility*. In: *Search of an ethics for the technological age*. University of Chicago Press
- Jonsen A (1998) *The birth of bioethics*. Oxford University Press
- Kant I (1785) *Fundamentación de la metafísica de las costumbres*
- Kant I (2008) *Groundwork of the metaphysics of morals*. Yale University Press
- Kirk R (1987) *Ten conservative principles*. Heritage Foundation, Washington, DC
- Kotlyarova V, Rudenko A, Yaksa N, Shubina M (2021) Digital technologies in modern higher educational space: analysis of risks and threats. *E3S Web Conf* 273:12050
- Kozlarski O (2021) From the humanism of critical theory to critical humanism. *Eur J Soc Theory* 24(2):246–263. <https://doi.org/10.1177/1368431020960958>

- Kropotkin P (2015) *The conquest of bread*. Penguin UK
- Larsson S (2020) On the governance of artificial intelligence through ethics guidelines. *Asian J Law Soc* 7:437–451
- Latypov IA (2021) ‘Counterfinality’ in sociological theory: reconceptualization of the concept. *RUDN J Sociol* 21(4):697–710
- Lucci A, Osti A (2024) Exit (digital) humanity: critical notes on the anthropological foundations of “digital humanism.” *J Responsible Technol* 17:100077. <https://doi.org/10.1016/j.jrt.2024.100077>
- Malatesta E (1909) *Anarchy* (Issue 5). Freedom Office
- Martin K, Shilton K, Smith J (2019) Business and the ethical implications of technology: introduction to the symposium. *J Bus Ethics* 160:307–317
- Marx K (2004) *Capital*: volume I, vol 1. Penguin UK
- Marx K, Engels F (2019) *The communist manifesto*. In: *Ideals and ideologies*. Routledge, pp 243–255
- Mayer K, Strassnig M (2020) The digital humanism initiative in Vienna. In: *Digitaler Humanismus*. Waxmann Verlag, Wien, pp 27–39
- McElwee B (2020) The ambitions of consequentialism. *J Ethics Soc Philos* 17(2):198
- McLoughlin KL, Brady WJ (2024) Human-algorithm interactions help explain the spread of misinformation. *Curr Opin Psychol* 56:101770. <https://doi.org/10.1016/j.copsyc.2023.101770>
- Melé D, Cantón CG (2015) *Fundamentos antropológicos de la dirección de empresas*. EUNSA
- Méndez JM (2023) *Curso Completo sobre Valores Humanos*. Innovaética Editorial
- Mill JS, Mill JS (1966) *On liberty*. Springer
- Mittelstadt BD, Allo P, Taddeo M, Wachter S, Floridi L (2016) The ethics of algorithms: mapping the debate. *Big Data Soc* 3(2):1–21
- Moro T (1516) *Utopia*. Ediciones Colihue SRL
- Moroni S (2012) Land-use planning and the question of unintended consequences. In: Emanuel Andersson D (ed) *The spatial market process*, vol 16. Emerald Group Publishing Limited, pp 265–288. [https://doi.org/10.1108/S1529-2134\(2012\)0000016013](https://doi.org/10.1108/S1529-2134(2012)0000016013)
- Nida-Rümelin J, Staudacher K (2024) Philosophical foundations of digital humanism. In: *Introduction to digital humanism*. Springer Nature Switzerland, pp 17–30. [https://doi.org/10.1007/978-3-031-45304-5\\_2](https://doi.org/10.1007/978-3-031-45304-5_2)
- Nisafani AS, Kiely G, Mahony C (2020) Workers’ technostress: a review of its causes, strains, inhibitors, and impacts. *J Decis Syst* 29(sup1):243–258. <https://doi.org/10.1080/12460125.2020.1796286>
- Owen R, Bessant JR, Heintz M (2013) *Responsible innovation: managing the responsible emergence of science and innovation in society*. Wiley
- Pugh EW (2009) Creating the IEEE code of ethics. In: 2009 IEEE conference on the history of technical societies, pp 1–13. <https://doi.org/10.1109/HTS.2009.5337855>
- Ramos-Díaz J, Jiang Q, Ramos-Sandoval R, Griffiths MD (2020) Problematic mobile phone use and risky behaviors among college students in Peru. In: *Proceedings—2nd International Conference on Informatics, Multimedia, Cyber, and Information System, ICIMCIS 2020*, pp 253–258. <https://doi.org/10.1109/ICIMCIS51567.2020.9354318>
- Reijers W (2016) *SATORI*. European Commission
- Roose K (2024) The apple vision pro is a marvel. But who will buy it? <https://www.nytimes.com/2024/01/31/technology/roose-apple-vision-pro.html?searchResultPosition=1>
- Rubin VL (2022) *Misinformation and disinformation*. Springer. <https://doi.org/10.1007/978-3-030-95656-1>
- Scheler M (1973) *Formalism in ethics and non-formal ethics of values: a new attempt toward the foundation of an ethical personalism*. Northwestern University Press
- Sen A (1979) *Utilitarianism and wellfarism*. J Philos New York NY 76(9):463–489
- Shin D (2024) *Artificial misinformation*. Springer Nature Switzerland. <https://doi.org/10.1007/978-3-031-52569-8>
- Shneiderman B (2020) Human-centered artificial intelligence: reliable, safe & trustworthy. *Int J Hum-Comput Interact* 36(6):495–504. <https://doi.org/10.1080/10447318.2020.1741118>
- Solove DJ (2011) Nothing to Hide: the false tradeoff between privacy and security
- Toulmin S (1982) How medecine saved the life of ethics. *Perspect Biol Med* 25(4):736–749
- Turkle S (2012) *Alone together: why we expect more from technology and less from each other*. Basic Books
- Turner C (2022) Augmented reality, augmented epistemology, and the real-world web. *Philos Technol* 35(1):19. <https://doi.org/10.1007/s13347-022-00496-5>
- UNESCO (2023) *Ethical impact assessment*. UNESCO
- Van den Hoven J, Vermaas PE, Van de Poel I (2015) *Handbook of ethics, values, and technological design: sources, theory, values and application domains*. Springer
- van Dijk J (2020) *The network society*, 4th edn. SAGE Publications Ltd
- Various authors (2018) *Vienna manifesto on digital humanism*. [https://caiml.org/dighum/dighum-manifesto/Vienna\\_Manifesto\\_on\\_Digital\\_Humanism\\_EN.pdf](https://caiml.org/dighum/dighum-manifesto/Vienna_Manifesto_on_Digital_Humanism_EN.pdf)
- Villegas-Galaviz C, Fernández-Fernández JL, Díaz de la Cruz C (2021) La obligación moral del mejoramiento humano: límites y posibilidades. *Revista Fe y Libertad* 4(1):25–38. <https://doi.org/10.55614/27093824.v4i1.87>
- Walker RL, Ivanhoe PJ (2007) *Working virtue: virtue ethics and contemporary moral problems*. Clarendon Press
- Weber M (2021) *Politics as a vocation*. Hassell Street Press
- Werthner H, Stanger A, Schiaffonati V, Knees P, Hardman L, Ghezzi C (2023) Digital humanism: the time is now. *Computer* 56(1):138–142. <https://doi.org/10.1109/MC.2022.3219528>
- Winner L (2010) *The whale and the reactor: a search for limits in an age of high technology*. University of Chicago Press
- Woodard C (2020) Consequentialism and Reasons for Action. In: Portmore DW (ed) *The Oxford handbook of consequentialism*. Oxford University Press, pp 178–196
- Wright D (2011) A framework for the ethical impact assessment of information technology. *Ethics Inf Technol* 13:199–226
- Yuval H, Harris T, Raskin A (2023) you can have the blue pill or the red pill, and we’re out of blue pills. *The New York Times*, pp 1–4. <https://www.nytimes.com/2023/03/24/opinion/yuval-harari-aichatgpt.html>
- Zhou J, Zhang Y, Luo Q, Parker AG, de Choudhury M (2023) Synthetic lies: understanding ai-generated misinformation and evaluating algorithmic and human solutions. In: *Proceedings of the 2023 CHI conference on human factors in computing systems*. <https://api.semanticscholar.org/CorpusID:257633591>
- Zuboff S (2019) *The age of surveillance capitalism: the fight for a human future at the new frontier of power*. Public Affairs

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.