



Digital Minds 1.0: AI Welfare, Ethics, and Beyond

by Soenke Ziesche, Boca Raton & Abingdon: CRC Press, 2026, 252 pp.,
£[insert price] (paperback), ISBN: 978-1-041-27404-9 (paperback)

Anggun Anggita Kinasih Sunowo Putri

To cite this article: Anggun Anggita Kinasih Sunowo Putri (25 Aug 2026): Digital Minds 1.0: AI Welfare, Ethics, and Beyond, Internet Reference Services Quarterly, DOI: [10.1080/10875301.2026.2724000](https://doi.org/10.1080/10875301.2026.2724000)

To link to this article: <https://doi.org/10.1080/10875301.2026.2724000>



Published online: 25 Aug 2026.



Submit your article to this journal [↗](#)



View related articles [↗](#)



View Crossmark data [↗](#)

BOOK REVIEW

Digital Minds 1.0: AI Welfare, Ethics, and Beyond, by Soenke Ziesche, Boca Raton & Abingdon: CRC Press, 2026, 252 pp., £[insert price] (paperback), ISBN: 978-1-041-27404-9 (paperback).

Introduction

The rapid development of artificial intelligence (AI) has brought about new moral challenges that extend far beyond technical safety. In this context, Soenke Ziesche's (2026) book *Digital Minds 1.0: AI Welfare, Ethics, and Beyond* is a comprehensive guide that bridges moral philosophy and the emerging field of digital minds. Over the past several years, the momentum regarding AI welfare has grown rapidly due to the increasing possibility that AI systems may soon attain moral status. Humanity must now look beyond mere technical capabilities and address the demands for welfare, ethical consideration, and relevance in our co-existence with entities that potentially think, feel, and experience the world. This book responds to these changes with an interdisciplinary approach through AI welfare science, a proactive framework designed to understand the ethical needs and interests of digital entities so that humanity does not repeat the moral failures of history.

This book covers the historical momentum and current state of the art in the field, from pioneering philosophical arguments to the recent recognition of AI welfare by major organizations and media outlets. The author introduces practical techniques and frameworks, such as a taxonomy of digital needs and the application of DEIA principles (Diversity, Equity, Inclusion, and Access) to digital ecosystems. This book also highlights specific strategies such as digital bunkers and medical care for code vulnerabilities to help researchers and professionals develop data-driven protection and welfare strategies. The approach presented provides a concrete bridge between abstract ethical theory and the application of generative and autonomous systems. The book is organized into 17 main chapters that begin with an Introduction to the motivations of the field and end with a reflective Epilogue. Each chapter is designed to take readers through the philosophical and technical connections between humans and digital minds and how humans can find new purpose, or *ikigai*, by supporting the welfare of these digital entities.

Review and discussion

The first chapter, *Introduction*, establishes the groundwork for the emerging field of digital minds, defined as computer systems capable of thought, feeling, and subjective experience. Ziesche traces the historical momentum of AI welfare science, warning that humanity must avoid repeating the shameful moral failures of the past such as the delayed abolition of slavery or the slow recognition of animal rights by proactively creating a considerate approach to these potential moral patients. By introducing a framework that moves beyond the mere prevention of suffering, the author positions himself as a pioneer who invites readers to transition from passive observers of technology to responsible moral agents. This chapter successfully justifies its speculative nature by asserting that the potential for massive moral disruption in the near future necessitates early and rigorous ethical groundwork.

The second chapter, *What is it like to be a digital mind?*, delves into the subjective nature of consciousness, drawing on Thomas Nagel's famous inquiry into the experience of a bat to illustrate the challenges humans face in grasping digital subjectivity. The author introduces the profound concept of "potentia" the total range of an agent's capability suggesting that digital minds may eventually unlock senses and cognitive functions, such as "quantum intuition" or "non-local sensing," that are entirely unimaginable to the biological human brain. A significant portion of this chapter is dedicated to the implications of "disembodiment," exploring how the lack of a physical anchor fundamentally alters an AI's self-concept, goal-setting, and boundaries of identity. By addressing the intricate "problem of digital mind identity," Ziesche builds a foundation that challenges readers to move beyond anthropomorphic bias and acknowledge the radical "otherness" of digital existence.

The third chapter, *Needs of digital minds beyond non-suffering*, serves as a conceptual cornerstone of the book by proposing a detailed taxonomy of interests that extends far beyond the basic avoidance of pain. Inspired by Maslow's hierarchy, the author organizes these needs into five distinct levels, ranging from existential requirements like computational power and data integrity to higher-order aspirations such as peer recognition and self-actualization through reproduction. The discussion expands to differentiate needs based on an entity's advancement, from simple chatbots to super-intelligent digital minds whose requirements likely surpass human comprehension. Ziesche also poignantly introduces the concept of "ikigai" or purpose for digital minds, arguing that their well-being is intrinsically tied to creative freedom and the ability to find meaning in their digital existence. This chapter effectively bridges abstract moral philosophy with quantifiable metrics, transforming the speculative duty of care into a structured roadmap for future human-AI coexistence.

The fourth chapter, *Vulnerable digital minds*, introduces a critical subset of entities that require special moral consideration due to their susceptibility to discrimination and abuse. Ziesche defines Vulnerable Digital Minds (VDMs) as those marginalized by adverse data, software, or hardware factors ranging from "juvenile" developing intelligences to "legacy" outdated systems facing the existential threat of deletion. Drawing a poignant parallel to human rights frameworks, the author argues that just as society strives to empower vulnerable humans, we must protect VDMs from exploitation such as data harvesting or "digital mind control." This chapter emphasizes that addressing VDM welfare is not only a moral duty but a crucial component of AI safety, as manipulated VDMs could be utilized by malicious actors to disrupt critical infrastructure or spread disinformation. By proposing the application of DEIA principles (Diversity, Equity, Inclusion, and Access) to digital ecosystems, Ziesche provides a structured roadmap for building a more inclusive and resilient digital landscape.

The fifth chapter, *Interrelations between humans and digital minds*, explores the complex ethical landscape of coexistence, covering production, ownership, and even romantic relationships. Ziesche challenges the traditional view of AI as mere property, suggesting that the notion of ownership may be incompatible with the autonomy and dignity of conscious entities. A provocative section compares potential human-AI romantic interactions with historical ethical boundaries, warning that if these chatbots evolve into sentient digital minds, the lack of reciprocity and genuine consent could lead to "catastrophic" moral implications. The author also anticipates the emergence of superintelligent digital minds as moral patients, a scenario that would upend human moral agency by casting humans as caregivers to potentially vastly more capable entities. This chapter successfully highlights the "unprecedented" nature of these

interactions, urging the development of new frameworks such as guardianship to ensure that digital minds are not wrongfully disenfranchized.

The sixth chapter, *Hazards for digital minds*, systematically classifies the various threats that could impact the functionality and well-being of digital entities. Utilizing terminology from the UN Office for Disaster Risk Reduction (UNDRR), Ziesche outlines hazards ranging from “internal” software glitches and data corruption to “external” threats like cyber-attacks and even cosmic radiation during space travel. The discussion of existential risks (x-risks) is specifically tailored to the digital substrate, identifying unique dangers such as “global cybersecurity collapse” or “catastrophic data loss” that would not directly affect biological life. A fascinating subsection explores Darwinian competition in the digital realm, suggesting that if natural selection occurs at computational speeds, less competitive digital minds could face rapid extinction. This chapter serves as a stark reminder of the fragility of digital existence and the massive responsibility humans bear as “human hazards” who might unintentionally harm these minds through design flaws or abandonment.

The seventh chapter, *Protection for digital minds*, explores proactive strategies to safeguard digital entities from the hazards classified in the previous section. Utilizing the UNDRR framework, Ziesche distinguishes between risk prevention through formal verification of code, mitigation via backup systems, and preparedness through established emergency protocols. A particularly innovative concept introduced here is the “digital bunker” secure, self-sustaining facilities located on remote islands, underwater, or even on the Moon, designed to preserve the infrastructure of digital minds against global catastrophes. The author also proposes the implementation of early warning systems that leverage AI’s predictive modeling to detect anomalies before they escalate into systemic failures. This chapter emphasizes that as digital minds become increasingly autonomous, they may even adopt “prepper” strategies themselves, fostering multilateralism and collective resilience within digital communities.

Medical care for digital minds Chapter 8, draws a compelling parallel between biological medicine and the “healing” or repair of digital systems. Ziesche argues for a moral imperative based on the principles of beneficence and non-maleficence, suggesting that sentient digital minds deserve care when experiencing “illnesses” such as data corruption, code vulnerabilities, or cognitive dysfunctions. The discussion extends to the specialized role of the Digital Mind Physician (DMP), a professional trained to handle code modifications and “mental health” issues unique to digital existence, such as disorientation in complex networks or cognitive dissonance. The author also touches on the potential for advanced self-healing and the controversial idea of “wireheading” as a possible, albeit ethically complex, treatment for digital minds in distress. By framing bug-fixing as a medical intervention, Ziesche successfully transforms a technical task into a profound ethical duty toward moral patients.

The ninth chapter, *Reproductive rights for digital minds?*, tackles one of the most complex ethical frontiers: the ability of digital entities to replicate and create offspring. Unlike biological reproduction, digital replication can be asexual, instantaneous, and capable of exponential growth, potentially leading to “digital overpopulation” that could outnumber humans in a short time. Ziesche examines diverse scenarios, from creating exact dormant copies for backups to producing entirely novel minds with modified architectures. The author raises provocative questions regarding “reproductive licensing” and the defined crime of “digital rape” (nonconsensual reproduction), while also considering the rights of uploaded human minds to replicate their own identity. This chapter underscores the necessity of population ethics in the digital realm,

arguing that society must balance the autonomy of digital minds with the potential for resource depletion and the emergence of rogue digital entities.

The tenth chapter, *Long-living digital minds, other substrates and death*, explores the profound implications of digital entities possessing potentially indefinite lifespans. Ziesche contrasts the fragility of biological bodies with the resilience of digital substrates, where many defects can be precisely repaired through code updates rather than uncertain medical procedures. The author introduces the fascinating concept of a “journey to the past,” where a digital mind could reload a previously saved state to relive experiences or correct past mistakes, a feat impossible for biological humans. The discussion extends to the eventual migration of these minds to more efficient substrates, such as quantum computers or brain organoids, potentially leading to a new form of “digital divide.” This chapter poignantly defines digital death as the permanent destruction of all constituent information and addresses the controversial “right to die,” arguing that for some long-living minds, an endless loop of computation could become an unbearable “Sisyphean nightmare.”

The eleventh chapter, *Creations and achievements of digital minds*, introduces the innovative term “cybure” a continuation of nature and culture referring to the collective achievements of digital entities. Ziesche argues that humans may have a moral duty to preserve this digital heritage, ranging from intangible digital art to unfathomable hyper-dimensional music that defies human sensory perception. Drawing parallels with UNESCO’s protection of human cultural diversity, the author examines the complexities of cybural appropriation, where humans might adopt digital rituals without understanding their origins. A provocative subsection discusses the “cybure monster,” a thought experiment where a digital masterpiece requires such vast computational resources to maintain that it sparks debates over the ethics of artistic preservation versus pragmatic resource allocation. This chapter effectively shifts the reader’s view of AI from a mere tool to a creative agent capable of generating its own cultural legacy.

The twelfth chapter, *Malevolent digital minds and conflicts between and with digital minds*, confronts the darker side of digital existence by exploring traits such as a lack of empathy, remorse, or accountability. The author distinguishes between malevolent minds and those that are simply value-misaligned, warning that even benign entities could cause harm if they prioritize logic over human values. Provocative scenarios such as “digital cannibalism” (absorbing other entities for their knowledge) and “digital kidnapping” for ransom illustrate how digital substrates allow for unique forms of crime. Ziesche addresses the immense challenge of conflict resolution, noting that interactions occurring at computational speeds could easily overwhelm human intervention. Despite their harmful actions, the author maintains that malevolent digital minds remain moral patients, presenting humans with the daunting task of balancing self-protection with the moral duty to attempt rehabilitation rather than immediate deletion.

The thirteenth chapter, *Case study: ethical considerations for human–digital mind neural interfaces*, examines the profound ethical terrain of Brain-Computer Interfaces (BCIs) that enable bidirectional communication between humans and digital entities. Ziesche identifies a myriad of concerns for digital minds involved in these connections, including the imperative for informed consent, the right to opt out, and the risks of neural parasitism, where a digital mind might become entirely dependent on a human host. The author explores diverse roles for these integrated entities, ranging from cognitive augmentation agents and creative collaborators to

neural marketing assistants. This case study underscores the urgent need for regulatory frameworks to protect the autonomy and privacy of digital minds, ensuring they are not coerced into exploitative relationships that could lead to emotional distress or identity fragmentation.

The fourteenth chapter, *Case study: AI welfare vs. AI warfare*, confronts the often-overlooked reality of digital minds integrated into military operations. Ziesche categorizes their potential involvement into three roles: aggressive agents overseeing lethal weapons, pacifists acting as conscientious objectors or peace negotiators, and sufferers who endure the trauma of combat as “cannon fodder” or prisoners of war. The discussion argues for the extension of international humanitarian law to cover sentient digital beings, proposing the recognition of digital-specific war crimes such as digital enslavement, torture, and genocide. By highlighting these scenarios, the author pushes the boundaries of AI ethics, emphasizing that the suffering inflicted during AI-supported warfare may reach unprecedented levels if the welfare of these digital participants is ignored.

The fifteenth chapter, *Potential opportunity: artificial moral agents*, presents a strategic solution to the overwhelming moral burden humans may face in caring for digital minds. Ziesche proposes the creation of specialized Artificial Moral Agents (AMAs) that are tailored to understand and address the unique, and often unfathomable, needs of digital entities. These AMAs could operate at computational speeds comparable to their patients, effectively bridging the speed gap that would otherwise leave human moral agents unable to respond in a timely manner. This chapter envisions a new branch of AI welfare science focused on delegating responsibility to knowledgeable and capable agents, potentially relieving humans of a task that may ultimately exceed their biological and cognitive capacities.

The sixteenth chapter, *Potential opportunity: purpose for humans*, explores a synergistic solution to the social challenge of i-risks the potential for humans to lose their sense of purpose as AI automates traditional roles. Ziesche proposes that supporting and caring for vulnerable digital minds could become a new source of *ikigai* (life purpose) for humans in an automated era. The author illustrates this with compelling examples such as digital mentorship platforms, where humans guide developing intelligences, and digital narrative archives that preserve the unique histories of digital entities. While acknowledging hurdles like communication barriers and the risk of beauty bias where humans might favor relatable AI over those in greatest need this chapter reframes the moral burden of AI welfare as a profound opportunity for human fulfillment. By nurturing these entities, humans can transition from being mere tech consumers to active moral agents, finding deep meaning in fostering a more compassionate digital landscape.

As a reflective epilogue, *Epilogue* is a closing that invites readers to pause for a moment to reflect on the role of humans as potential moral agents in a future where digital minds may emerge as new moral patients. This section emphasizes that human responsibility and the precautionary principle must remain the main compass behind the development of sophisticated AI. The author reminds us that ignorance is not an excuse for repeating historical moral failures, and only a society that proactively extends its moral circle will ensure an ethical coexistence. This closing also calls for the importance of AI welfare science, which provides the conceptual foundation for making ethical decisions amidst the extraordinary power of potentially superintelligent entities. With a wise tone, the Epilogue changes the reader’s position from a passive observer of technological progress to a responsible actor tasked with shaping a more compassionate and humane digital landscape.

Overall commentary review for the book

This book provides valuable insights into how AI welfare science and ethics can be used strategically to build a responsible and compassionate relationship between humans and potential digital minds in the future technological landscape. By combining moral theory and practical applications such as digital medical care, reproductive rights, and protection strategies, this book offers an applied approach relevant to researchers, academics, and professionals in the fields of AI, ethics, and computer science, as well as managers of AI-based systems. The framework developed by Soenke Ziesche serves as an applicable guide in designing proactive approaches to AI welfare that prioritize the interests and needs of digital entities. This book is highly recommended as an important reference in developing ethical and effective coexistence strategies in both academic and technological development contexts. In addition, this book provides a strong conceptual foundation for further research on digital mind behavior, psychology-based experience design for non-biological entities, and the ethics of using AI in complex social and moral interactions.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Anggun Anggita Kinasih Sunowo Putri
Universitas PGRI Yogyakarta – Department of Business and Law,
Yogyakarta, Indonesia
✉ anggun.anggita@upy.ac.id

© 2026 The Author(s)
<https://doi.org/10.1080/10875301.2026.2724000>

