AI and Ethics (2026) 6:27
https://doi.org/10.1007/s43681-025-00884-5

RESEARCH

The ethics of analog AI

Maximilian Kiener[1] · Jonas Bozenhard[1] · Gitta Kutyniok[2,3,4,5] · Sven Nyholm[2,3]

Received: 8 May 2025 / Accepted: 25 September 2025

Abstract
AI ethics has matured as a field, yet its concepts and tools have been developed almost entirely with digital AI in mind. Emerging "analog" AI approaches, that is, systems that compute over continuous, rather than binary, signals, challenge this digital default. Analog AI promises substantial energy savings and alternative modes of information processing. However, it also departs in ethically significant ways from the widely discussed digital AI systems such that its differences from digital AI could mean that current ethical frameworks may not address analog AI's unique challenges. This paper highlights how the distinctive characteristics of analog AI raise new ethical questions that require careful consideration. We examine four key areas of AI ethics, namely fairness, privacy, explainability, and safety, and propose a roadmap for the ethical examination of analog AI. Our aim is to lay the foundation for an important but hitherto unrecognized subfield within AI ethics: the ethics of analog AI.

Keywords Artificial Intelligence · Analog · Fairness · Privacy · Explainability · Safety

1 Introduction

Over the past decade, AI ethics has evolved into a distinct field at the intersection of philosophy, technology, and the social sciences. However, discussions have largely concentrated on digital AI, that is, systems that rely on binary computation (0s and 1s). This is mainly because the vast majority of AI systems shaping our world in the past and the present are digital. In fact, AI and digital computing are often so closely linked that many are surprised to learn that there is an alternative: so called "analog" AI. Unlike digital AI, analog AI processes continuous (and not binary) signals and holds the promise of two key advantages over the digital paradigm. First, it demonstrates higher energy efficiency, which can help significantly reduce energy consumption and deliver considerable sustainability benefits. Second, analog AI resembles the brain's information processing more closely, thus presenting the prospect of developing machines with more human-like intelligence. Thus, although analog AI's potential is only barely realized today, it offers the promise of more sustainable and "intelligent" AI. For this reason, it has recently been suggested that "[c]omputing today is digital, but analog has a future" [26] p. 354), with some even going so far as to claim that "[t]he future of AI is analogue" [37].

However, despite this compelling vision for the future of AI, the ethical implications of analog AI are practically unexplored. Since current AI ethics frameworks were designed for digital AI, they are insufficient for addressing the distinct ethical challenges emerging from analog AI's unique information processing features that vary substantially from those of digital AI. Thus, the aim of this paper is to lay the foundation for an important but hitherto unrecognized subfield within AI ethics: the ethics of analog AI.

✉ Maximilian Kiener
maximilian.kiener@tuhh.de

Jonas Bozenhard
jonas.bozenhard@tuhh.de

Gitta Kutyniok
kutyniok@math.lmu.de

Sven Nyholm
S.Nyholm@lrz.uni-muenchen.de

1 Hamburg University of Technology, Hamburg, Germany
2 Ludwig-Maximilians-Universität München, Munich, Germany
3 Munich Center for Machine Learning (MCML), Munich , Germany
4 University of Tromsø, Tromsø, Germany
5 DLR-German Aerospace Center, Cologne, Germany

This new ethics deals with the novel ethical questions that arise from the distinctive characteristics of analog AI. The ethics of analog AI identifies the ethical risks posed by this emerging form of AI and guides its responsible design and adoption.

This paper proposes a roadmap for the ethical examination of analog AI. To that end, we attend to four key areas of AI ethics and explore how analog AI might impact each of them: fairness, privacy, explainability, and safety. Our analysis focuses on these four topics, as they span a wide range of concerns in AI ethics and represent the areas where the potential benefits and risks of analog AI are arguably most pronounced and socially impactful. We assess the extent to which analog AI presents advantages and disadvantages relative to digital systems in each of these areas. These investigations aim to provide the basis for further exploration of the ethical dimensions of analog AI. But before we do so, we shall lay out the key features of analog AI and the current state of research.

2 What is analog AI?

The success of artificial intelligence (AI) and its broad societal impact have led many to call it the Fourth Industrial Revolution. However, concerns about reliability such as missing security, robustness issues, biased decisions, and AI's black box behavior as well as the enormous energy consumption of AI-computing have lately raised the question to which extent those problems are of a more fundamental nature than previously thought [7]. Digital systems, which rely on 0 s and 1 s, sometimes struggle to accurately represent real-world problems that are continuous by nature. This mismatch could lead to inaccuracies and stability issues, particularly for AI tasks like classification problems, inverse problems, and solving partial differential equations. These challenges highlight the need for next-generation AI to be developed alongside advanced computing technologies. Recent research suggests that analog computing, when analyzed through the theoretical framework of a Blum-Shub-Smale (BSS) machine, offers insights into addressing various challenges. However, since BSS machines are an abstract model that cannot be physically realized, the extent to which practical analog computing can approximate this model remains an open question [6, 9].

Advances in analog computing have started to drive the development of neuromorphic computing, a technology designed to mimic the brain's processing methods. Neuromorphic chips can use analog components to process real-world data through electrical signals like currents and voltages [23]. This brain-inspired design enables more efficient and natural information processing compared to traditional digital methods. By harnessing analog computation, neuromorphic systems offer significantly higher energy efficiency and better performance for AI tasks [3]. To unlock their full potential, however, specialized software models are essential. For example, spiking neural networks, which mimic the way neurons transmit signals in the brain, are well-suited for implementation on neuromorphic hardware. One particularly promising application is neuromorphic robotics, where extremely low energy consumption and minimal latency are critical, for example, in autonomous robots searching for survivors in collapsed tunnels or operating in deep-sea environments for extended periods. Another key area is neuromorphic vision for automotive, consumer, medical, and security systems. Here, capabilities such as on-chip learning and online adaptation offer enormous advantages, enabling devices to become more adaptable, responsive, and self-optimizing over time [36].

Neuromorphic chips are not the only innovation in the realm of analog computing. Certain quantum systems can also meet similar demands [28]. Another active line of research is biocomputing, which uses biological systems or components, such as single neurons, either in place of, or in combination with, traditional silicon-based electronics [25, 42].

Despite its potential and advantages, however, analog AI also comes with limitations that warrant careful consideration. Both its strengths and weaknesses are crucial for an ethical analysis of analog AI. Interestingly, what initially seems like technical limitations could be, as we shall discuss below, ethical advantages under the right circumstances. For the purposes of this paper, the following four features of analog computing are particularly relevant (cf. [31,33]):

(a) **Processing Method**: Analog AI processes continuous data through physical properties like voltage, current, and resistance, whereas digital AI processes discrete data using binary digits. Because of this, analog AI can process (analog) data directly without conversion, reducing bottlenecks and increasing speed, whereas digital AI, particularly in traditional von Neumann architectures, faces memory bottlenecks due to the need to convert incoming data into binary digits.

(b) **Power Consumption**: Analog AI typically consumes less power, which is often discussed in the promising field of neuromorphic systems that aim to mimic the brain's efficiency. Digital AI, however, often requires more power, though optimizations can improve efficiency in specific applications. This difference is partly due to the computational efficiency that results from analog AI's method of processing information.

(c) **Precision**: In theory, analog AI is very precise because it works with real numbers directly. Digital AI, on the

other hand, can only approximate real numbers using fractions, even though it is more accurate when working with exact values. However, in practice, analog AI is limited by noise, environmental factors, and physical component imperfections. Physical variations impact the operation of analog computers in ways that do not affect digital computers.
(d) **Flexibility**: Analog AI is less flexible as hardware must often be reconfigured physically for different tasks, while digital AI is highly flexible as software-based reprogramming allows dynamic adaptability. This also affects the scalability: while analog circuits can scale well for specialized tasks, they are harder to reconfigure dynamically. By contrast, digital computing is highly scalable for general purpose systems.

2.1 Analog AI ethics

Currently, there is no dedicated research on the ethics of analog AI that asks, head-on, if the ethical questions pertaining to analog AI differ in relevant ways from those pertaining to digital AI.

What we do have, however, are the following contributions relevant in the context of our paper:

- Technical analog AI papers on aspects that ethicists can build on. Although the do not themselves provide any normative assessment of analog versus digital computing, they highlight important technical aspects for an ethical analysis [14, 30, 38].
- Research on how the choice of hardware can affect outcomes and hence, indirectly, questions of fairness, though this is not framed as an "ethics of analog AI" [34].
- Shorter opinion pieces, blogs, and research articles [1] that gesture at analog AI's energy benefits [20]; IBM-Research [21, 22] and thus also motivate further ethical sustainability-related research.
- Genuine ethics work appears under adjacent labels, neuromorphic or brain-inspired AI, and neurotechnology, and that's where the substantive ethical analysis arises, but more on topics like privacy and neurorights, identity and agency, or equity [15, 41, 46]. But here it is important to distinguish neuromorphic AI (a brain-inspired computational paradigm, often spiking/event-driven) from analog AI (implementations that compute with continuous physical signals). Neuromorphic systems can be implemented digitally, analog, or mixed-signal, whereas many analog-AI devices accelerate conventional non-spiking networks and are not "neuromorphic" per se.
- Within this general lack of systematic exploration of analog AI ethics, [8] are an exception and explicitly connect legal-ethical requirements, especially algorithmic transparency, to the computing model itself and compare digital (Turing) vs. analog (Blum–Shub–Smale, BSS) computation. They formalize algorithmic transparency as a precondition for trustworthiness and show that whether it can be achieved may depend on computability in the chosen model [8].

These strands of work indicate the need of a deeper analysis of the ethics of analog AI. To provide such an analysis, we selected four core ethical aspects for our discussion of of analog AI, namely fairness, privacy, explainability, and safety, and we decided to focus on these for two reasons. First, these four aspects are central to AI ethics more generally. So, if there is anything distinctive about analog AI from an ethical perspective, it should surface here. Second, as we shall show in the following sections, each of these aspects has specific links to what is characteristic of analog AI and provides therefore a needed component in motivating this new field of analog AI ethics. Yet, the field of analog AI ethics, which we want to introduce in this paper, will eventually have to go beyond this scope and include further considerations too.

1. Fairness

In digital AI ethics, fairness is a key topic, particularly regarding algorithmic bias and discrimination. Researchers have examined how biases in data or model design can produce unequal outcomes for groups defined by race, gender, or socioeconomic status. To address this, various fairness metrics assess whether AI systems perform equitably across groups [18, 19]. For example, fairness can be evaluated by comparing accuracy rates or ensuring consistent false positive and false negative ratios across populations. Challenges arise when multiple fairness metrics conflict due to inherent trade-offs. Optimizing for equal accuracy across groups may clash with minimizing disparities in false positives or false negatives. These conflicts have sparked debates over which fairness criteria to prioritize and how to make principled decisions when trade-offs are unavoidable [19].

With analog AI, two key differences complicate fairness evaluation. Its sometimes imprecise outputs as well as its continuous data processing challenges traditional fairness metrics, which rely on the discrete computations of digital AI. Analog AI's variability makes it difficult to measure outcomes like accuracy, false positives, or false negatives, necessitating new assessment approaches. This raises several important questions: How does imprecision in practice affect the reliability of fairness audits, and what strategies can address this? How might output variability skew audit results, leading to over- or underestimation of bias? In real-time applications like surveillance or hiring, how can we ethically mitigate the risk of parallel processing amplifying biases? Finally, what new fairness metrics or evaluation methods are needed to accommodate analog AI's practical imprecision and noise-related variability?

Second, analog AI raises fairness questions that go way beyond fairness metrics and have to do with analog AI's potential to be significantly more environmentally sustainable than digital AI [44]. This matters because the unsustainable nature of current digital AI exacerbates global inequities: while some people develop and benefit from these technologies, others bear the environmental harms they create [43]. Specifically, individuals and organisations in the Global North - particularly wealthy tech companies and consumers - reap the advantages of AI systems, while these technologies often rely on unsustainable practices that exploit natural resources and damage the environment. In contrast, communities in the Global South typically have less access to the benefits of AI, yet they disproportionately suffer the environmental consequences and negative externalities [13]. Analog AI, with its potential for greater sustainability, offers an opportunity to mitigate this form of injustice. By reducing the environmental footprint of AI technologies, analog AI could help distribute both the benefits and burdens of AI development more equitably.

Analog AI's potential for greater sustainability introduces key ethical considerations. Do its environmental benefits justify its use over digital AI, even when digital systems sometimes offer greater precision? How should sustainability and precision be balanced in different contexts? Can deploying more sustainable analog AI reduce the unfair distribution of environmental harms, where communities in the Global South disproportionately suffer the costs of digital AI systems? Addressing these questions requires attention to the specifics of analog AI beyond the current framework in digital AI ethics.

2. Privacy

Analog AI introduces several technical features that can support privacy principles such as data minimization, purpose limitation, and storage limitation. Since analog AI improves speed and reduces latency in certain situations, it can eliminate the need to transfer large amounts of data to centralized servers (cf. [31, 33]). This supports data minimization by processing only the data necessary for immediate decision-making. Additionally, analog in-memory computing (AiMC) can reduce the volume of data required for computations. As analog AI is designed for specific tasks, it naturally supports purpose limitation by minimizing the potential for data to be repurposed for unintended uses. Moreover, the ability to process data locally and in real-time means, at least for certain applications, that analog AI does not rely on storing large datasets over long periods. Instead, it operates with transient data, reducing the need for extensive data retention [23]. This approach aligns with storage limitation principles, addressing privacy concerns associated with the long-term storage of sensitive information.

The striking point here is that what may seem like technical limitations at first, i.e. less data, lack of transferability to another purpose, and storage limitation, can be recast as ethical strengths in the protection of data. To support this hypothesis further, consider the theory of privacy as Contextual Integrity proposed by Helen Nissenbaum [35]. Contextual integrity posits that privacy is maintained when personal information flows appropriately according to the norms of the specific context. It emphasizes the importance of context-specific norms and the roles of data subjects and data recipients. Contextual integrity could therefore favor using analog AI rather than digital AI in many contexts because of analog AI's ability to process data locally and only for specific, well-defined purposes.

However, analog AI also poses unique challenges that may not apply, or apply differently, to digital AI. Rights such as the right to erasure under the EU's GDPR [4], the right to delete personal information under California's Consumer Privacy Act (CCPA) [2], or similar data protection laws in other jurisdictions, such as Brazil's LGPD [16], become harder to exercise due to analog AI's rigidity. While digital systems in large language models (LLMs) also face difficulties in removing individual data points - sometimes requiring costly retraining of the entire system - analog AI adds a distinct layer of complexity with its hardware components. Analog hardware is often physically rigid and challenging to reprogram once programmed, making data extraction or erasure even more impractical. This rigidity reinforces purpose limitation but significantly limits individuals' control over their data post-processing. This raises several questions: To what extent can individuals assert the right to erasure with analog AI? Would this require complete hardware reprogramming, and what costs are reasonable to facilitate this? How should these costs be balanced against data subjects' rights? If withdrawal of consent is impractical, should consent protocols be stricter upfront? How do these constraints affect data controllers' obligations and liabilities, and what safeguards or policies should be in place when using analog AI?

Moreover, another challenge involves security and privacy in analog AI systems. Since analog AI systems rely heavily on physical hardware (cf. [31, 33]), this makes them vulnerable to tampering, theft, and damage, leading to risks of unauthorized access, data manipulation, or system compromise. Additionally, analog AI's susceptibility to noise and imprecision can corrupt data, causing privacy breaches through unintended data leakage or misinterpretation. Such inaccuracies can result in misidentifications, producing false positives or negatives that improperly disclose personal data. These challenges necessitate tailored security protocols and raise questions about the distribution of responsibility between data controllers and subjects. What safeguards should controllers implement to mitigate these risks, and how should liability be shared when hardware or environmental vulnerabilities cause privacy breaches?

3. Explainability

Explainability can be defined as the "level of understanding [of] how the AI-based system […] came up with a given result" (ISO/IEC TR 29119-11:2020(en), 3.1.31). Analog computing introduces both distinctive opportunities and challenges in this context. One theoretical advantage lies in its ability to process real numbers directly. While digital systems approximate real numbers through binary states, analog computation operates on continuous values and thus avoids the need for discretization. Analog computing is like measuring a curved line with a flexible tape, whereas digital computing is like using a rigid ruler for the task. As shown by theoretical models like Blum-Shub-Smale (BSS) Machines, digital computing can, in certain cases, produce false computations that the analog BSS model can avoid [8]. So, in theory, analog AI holds promise for explainability. It has the potential to align more closely with real-world phenomena and allows for a more intuitive understanding of how inputs are transformed into outputs.

However, there are challenges in realizing these theoretical advantages in practice. As mentioned before, real-world analog systems are susceptible to noise, hardware imperfections, and environmental interference, which compromise the precision that theoretical models promise. These physical limitations introduce variability, making analog systems less reliable and repeatable than their digital counterparts, which rely on error correction and standardized binary processing to maintain precision [14, 30]. Thus, while analog systems can be very accurate and explainable in theory, practical constraints undermine their ability to deliver easily foreseeable results in noisy environments.

Relatedly, a significant challenge for analog AI is what might be termed explainability at the information processing level [5]. The fact that analog systems process information continuously makes it difficult to pinpoint exact states or transitions that lead to specific outcomes. This problem is exacerbated by the non-repeatability of analog systems, as their sensitivity to noise and environmental factors makes it harder to trace and verify their decision-making processes [11]. By contrast, digital AI systems, with their reliance on discrete and well-defined states, offer a higher degree of explainability at the information processing level.

However, when paired with specific architectures, there is potential to enhance the practical explainability of analog AI systems in the future. Analog computing shows distinctive advantages for "truly" neuromorphic AI, as it aligns more closely with the way biological neurons process signals [24]. The human brain is a highly modular system, consisting of specialized regions that handle different functions, such as vision, memory, and decision-making, while still working as a whole [27]. Therefore, to mimic the human brain as accurately as possible, an advanced neuromorphic AI demands not only analog computing, but also a high degree of modularity. In fact, increased modularity might be necessary to mitigate the inherent limitations of analog computing, such as signal variability and noise sensitivity, making practical implementation more viable. Greater modularity can improve the explainability of an AI system by allowing for easier localization of specific functions. Such modularity offers what we might call *systemic* explainability, i.e., a better understanding of how the system, with its various modules, works as a whole and generates specific outcomes.

In summary, analog AI's direct processing of real numbers offers theoretical transparency and an intuitive connection to real-world phenomena. Yet, practical limitations such as noise sensitivity and challenges in information processing-level explainability temper these advantages. Modular neuromorphic architectures may enhance the systemic explainability of analog AI systems. However, this raises the question of whether systemic explainability can compensate for analog AI's lack of explainability at the information processing level. In contrast, digital deep neural networks, characterized by their "black box" nature and highly interconnected structures, lack systemic explainability, but offer superior explainability at the information processing level compared to analog systems. In conclusion, unlike digital AI, which excels in precision, repeatability, and robust error correction, analog AI's performance is highly context-dependent: Its strengths may shine in controlled settings and specific architectures, while its weaknesses become more pronounced in unpredictable, dynamic environments and most practically feasible architectural implementations.

This has implications for questions about responsibility too, in particular with respect to how easily human agents can satisfy the so-called epistemic condition of responsibility when deploying analog AI. The epistemic condition refers to the idea that for someone to be responsible for an outcome, they must have had the ability to (reasonably) foresee the consequences or outcomes of their action [39]. The reflections above add new layers to the responsibility debate by introducing new distinctions - such as theoretical versus practical explainability and systemic versus information processing-level explainability - that enable more nuanced discussions of the epistemic condition as it relates to AI and responsibility.

4. Safety

Over the past decades, AI has grown increasingly powerful, autonomous, and capable. The recent success of large language models highlights the challenges of predicting the development, performance, and emerging capacities of AI. This raises the question of how to control increasingly powerful and autonomous AI systems and align them with ethical values to ensure AI safety [12, 17]. Analog AI holds significant potential to enhance AI safety in several ways.

First, one approach in trying to retain control over advanced AI systems is known as *capacity control*, which involves limiting a machine's capabilities to ensure human oversight, even as the AI becomes more powerful. An intuitive strategy is equipping the system with an off-switch, but if the AI is highly advanced, it may perceive the switch as a threat to fulfilling its objectives and preemptively replicate itself across multiple systems [10, 40]. This could make it effectively indestructible and uncontrollable. Analog AI presents a much lower risk in this regard, as its unique hardware characteristics make copying model weights between systems far more difficult [29].

Second, another capacity control strategy, known as *boxing* [45], proposes to isolate AI by limiting its input and output channels, restricting it to specific tasks and environments. Notably, analog AI naturally shares this characteristic due to its lack of digital computing's flexibility. Analog systems are confined to specialized applications and their "programs" are fixed to the hardware. This inherent limitation reduces the risk of unintended interactions with the wider world, making analog AI easier to contain and control.

Third, potentially dangerous AI capabilities may arise not only from training a single model but from integrating various AI systems. However, due to the distinctiveness of analog hardware, merging multiple analog AI systems is considerably harder than merging digital ones. Moreover, unlike digital neural networks, analog neural networks cannot use data parallelism, meaning they cannot process multiple data streams simultaneously across identical units, because the hardware of each analog neural network is different, making it difficult to share information directly [32].

Fourth, when it comes to preventing advanced AI systems from developing unexpected, misaligned capacities, explainability at the systemic level (which highly modular neuromorphic analog AI may offer) is more relevant than explainability at the information processing level (which digital AI offers). This is because the emergence of new capacities is a systemic phenomenon, not one that occurs at the level of information processing. Moreover, misbehavior is to be attributed to the system (or at least individual modules), rather than to single computations within the system. Hence, highly modular neuromorphic analog AI might not only enhance systemic explainability but also improve safety.

However, analog AI also exhibits limitations related to safety. Due to its non-repeatability, context-dependent imprecision, and the limitations in signal transmission and storage, analog computing seems less controllable when attempting to ensure AI reliably adheres to specific values or points of compliance. Considering these points, we can identify two concerns related to AI safety: one focuses on preventing or mitigating emergent and potentially dangerous capabilities, while the other concentrates on ensuring that AI systems reliably align with specific values we want them to comply with. With its limitations in scalability, information sharing, and model integration and its advantages in capacity control, boxing, and potentially in systemic explainability, analog AI reduces the risk of developing AI with dangerous emergent capacities. In contrast, digital AI is better suited for aligning systems with specific values due to its higher reliability, efficient learning capabilities, and greater explainability at the information processing level. Therefore, in contexts, in which high precision and reliable system behavior are essential, digital AI seems favorable.

At this point, it is worth noting the EU AI Act, which embeds a safety-centred, risk-based scheme: uses that can significantly affect health, safety, or fundamental rights are classed as high-risk (Annex III) and must satisfy, among other things, risk management across the lifecycle (Art. 9), human oversight (Art. 14), and appropriate accuracy, robustness, and cybersecurity (Art. 15). Our analysis of the ethics of analog AI in terms of safety can inform these assessments. To highlight a few relevant aspects:

Art. 9 (Risk management): High-risk systems must identify hazards, set operating limits, and verify controls over time. In these critical domains, digital AI can be prima facie preferable for its greater reliability and precision, which typically makes it easier to keep the system behaving as expected and to manage updates safely. Analog AI, by contrast, can introduce too much variability in specific environments, so safety measures have to be stricter: more frequent checks and recalibration, and clear limits on where and how the system may be used. That said, analog architectures can reduce hazards tied to capability escalation and easy replication, as mentioned earlier, which matters where containment is the overriding safety goal.

Art. 14 (Human oversight): Oversight must be "effective" during use. At present, digital AI systems offer more explainable decision-making and easier human oversight due to their higher reliability and precision. Insofar as effective oversight presupposes that operators can understand what the system is doing well enough to predict when to trust, slow, or stop it, this speaks in favour of digital AI: reproducible behaviour, consistent logs, and stable interfaces make it easier to detect out-of-scope operation and to intervene in time. Analog AI can be harder to oversee today if behaviour varies across devices or degrades with conditions; operators then need simple tolerances, on-the-spot checks, and clear escalation paths. However, when combined with neuromorphic architectures, future analog AI systems might - thanks to greater systemic explainability - provide clearer, more intuitive pictures of module roles and interactions, thereby supporting more meaningful oversight than current digital pipelines.

Art. 15 (Accuracy, robustness, cybersecurity): Digital systems typically make it easier to demonstrate stable accuracy under stress tests. Analog systems must evidence stability under environmental conditions and aging, and resilience to physical-layer interference. Cybersecurity measures for analog AI must therefore address both digital attack surfaces and vulnerabilities inherent in physical signal behavior.

Overall, in areas in which the emergence of dangerous AI capacities is likely or may have catastrophic effects, analog AI seems to be the safer option - for example, in highly advanced robotics, biotechnology, simulation-heavy research, and general-AI/AGI-oriented work - because limiting integration and replication is itself a safety control. At the same time, digital AI still leads in terms of current performance, and this raises further trade-off questions in the context of safety when analog AI is an alternative. For example, would we be willing to compromise not only compliance reliability but also performance in these areas to prevent the potential emergence of harmful AI capabilities?

3 Summary of comparing digital and analog AI

Digital AI is the current default. But with analog AI becoming a more viable option in the not-too-distant future, we need to ask when to use analog rather than digital AI. Building on our discussion, several ethically significant contrasts ought to be considered:

- Sustainability: Analog AI is energy-efficient for low-power applications, while digital AI can achieve sustainability only if there are data centers powered by renewable energy.
- Fairness: Analog AI can help advance group-level fairness for collective or even global questions, while digital AI is more appropriate in the context of individual fairness for personalized outcomes or metrics.
- Privacy: Analog AI suits fixed tasks where purpose limitation is needed to prevent illegitimate repurposing, while digital AI excels in adaptable tasks needing frequent updates.
- Explainability: Analog AI shows significant potential in low-noise, controlled environments and future modular neuromorphic architectures, while digital AI remains the better choice for noisy, unpredictable settings or applications requiring greater explainability at the information processing level.
- Safety: Analog AI is favorable in areas in which preventing or mitigating emergent and potentially dangerous capabilities is paramount, while digital AI is better suited for ensuring conformity with specific values or points of compliance.

4 Conclusion

Doing justice to the issues mentioned throughout the paper requires the development of a new subfield: analog AI ethics. We hope to have laid the foundation for this new area of inquiry.

By calling for “analog AI ethics” we name a domain where ethical appraisal is sensitive to how computation is realized, so that the central normative question becomes: when is it ethically better to implement a given capability analog rather than digital (and vice versa), as well as which further normative questions follow from our decision in this regard? This is not reducible to engineering performance. It concerns distinct risk profiles and duties that attach to different compute models. Which duties hinge on the computation we choose (analog vs. digital)? For which tasks does a given computation make key ethical aims, like transparency, stable behavior, or robust privacy, realistically attainable? And what tests must a system pass before and during deployment? These are all questions that the future of this field will include.

To be clear, however, we do not claim that analog AI ethics has to reinvent the wheel, that it needs novel moral theories from the ground up, or that it rejects or disconnects from prior AI ethics. Our claim is more about agenda-setting: as with autonomous vehicles or generative AI, the novelty lies in a targeted problem list that can adapt already familiar methods or approaches from existing debates while adding analog-specific questions and challenges. Hence, the next step is to treat analog computing architectures not as curiosities but as live options whose shape matters for ethics: when does that computational paradigm help us keep the promises we owe about fairness, privacy, explainability, and safety and when doesn't it? In the near term, progress could look practical and human-facing. Before we build, we should ask: What is this system *for*? What will it *not* be used for? What kind of stability or adaptability does this task actually need? How will we explain what the system can and cannot do? Finally, longer term, the field may create the expectation that we justify, not just *what* we compute, but also *how* we compute. If we can normalize that stance, then "analog AI ethics" will mark a way of designing systems that keeps people's interests in view, and leaves room for revision as our understanding deepens.

Author contributions Maximilian Kiener and Jonas Bozenhard are the lead authors. Gitta Kutyniok and Sven Nyholm are contributory authors. All authors reviewed the entire manuscript.

Funding Open Access funding enabled and organized by Projekt DEAL. No funding was received for this paper.

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

Declarations

Competing interests The authors declare no competing interests.

Ethics declaration No conflict of interest. No ethics approval required for this research.

References

1. Ambrogio, S., Narayanan, P., Okazaki, A., Fasoli, A., Mackin, C., Hosokawa, K., Friz, A.: An analog-AI chip for energy-efficient speech recognition and transcription. Nature. **620**(7975), 768–775 (2023)
2. Ambrose, M.L., Ausloos, J.: The right to be forgotten across the pond. Journal of Information Policy **3**, 1–23 (2013)
3. Andrei, V.C., Drăgutoiu, A.P., Béna, G., Akl, M., Li, Y., Lohrmann, M., Boche, H.: *Deep-Unrolling Multidimensional Harmonic Retrieval Algorithms on Neuromorphic Hardware.* Paper presented at the 2024 58th Asilomar Conference on Signals, Systems, and Computers. (2024)
4. Ausloos, J.: The Right to Erasure in EU Data Protection Law. Oxford University Press, Oxford (2020)
5. Bitar, A., Rosales, R., Paulitsch, M.: Gradient-based feature-attribution explainability methods for spiking neural networks. Front. NeuroSci. **17**, 1153999 (2023)
6. Blum, L.: *A theory of computation and complexity over the real numbers.* Paper presented at the Proceedings of the International Congress of Mathematicians. (1990)
7. Boche, H., Fono, A., Kutyniok, G.: Limitations of deep learning for inverse problems on digital hardware. IEEE Transactions on Information Theory (2023). https://doi.org/10.1109/TIT.2023.3326879
8. Boche, H., Fono, A., Kutyniok, G.: Mathematical algorithm design for deep learning under societal and judicial constraints: the algorithmic transparency requirement. arXiv preprint arXiv:2401.10310. (2024)
9. Boche, H., Fono, A., Kutyniok, G.: Inverse problems are solvable on real number signal processing hardware. Applied and Computational Harmonic Analysis **74**, 101719 (2025)
10. Bostrom, N.: Superintelligence: paths, dangers, strategies. Oxford University Press, Oxford (2014)
11. Cartier, E.A., Kim, W., Gong, N., Gokmen, T., Frank, M.M., Bishop, D.M., Wu, E.Y.: *Reliability challenges with materials for analog computing.* Paper presented at the 2019 IEEE International Reliability Physics Symposium (IRPS). (2019)
12. Clymer, J., Gabrieli, N., Krueger, D., Larsen, T.: Safety cases: How to justify the safety of advanced AI systems. arXiv preprint arXiv:2403.10462. (2024)
13. Crawford, K.: The Atlas of AI: power, politics, and the planetary costs of artificial intelligence. Yale University Press (2021)
14. Demirkiran, C., Nair, L., Bunandar, D., Joshi, A.: A blueprint for precise and fault-tolerant analog neural networks. Nat. Commun. **15**(1), 5098 (2024)
15. Farisco, M., Baldassarre, G., Cartoni, E., Leach, A., Petrovici, M.A., Rosemann, A., van Albada, S.J.: A method for the ethical analysis of brain-inspired AI. Artif. Intell. Rev. **57**(6), 133 (2024)
16. Gadoni Canaan, R.: The effects on local innovation arising from replicating the GDPR into the Brazilian general data protection law. Internet Policy Rev. **12**(1). (2023)
17. Gyevnar, B., Kasirzadeh, A.: AI safety for everyone. Nat. Mach. Intell., 1–12. (2025)
18. Hedden, B.: On statistical criteria of algorithmic fairness. Philos. Public Affairs. **49**, 209–231 (2021)
19. Hellman, D.: Measuring algorithmic fairness. Va. Law Rev. **106**(4), 811–866 (2020)
20. Huang, J.: The Next Frontier of AI: lower emission processing using analog computers: (2022). https://montrealethics.ai/the-next-frontier-of-ai-lower-emission-processing-using-analog-computers/?utm_source=chatgpt.com
22. IBM-Research: How analog in-memory computing could power the AI models of tomorrow: (2025). https://research.ibm.com/blog/how-can-analog-in-memory-computing-power-transformer-models
21. IBM-Research: IBM Research's latest analog AI chip for deep learning inference: (2023). https://research.ibm.com/blog/analog-ai-chip-inference
23. Indiveri, G., Liu, S.-C.: Memory and information processing in neuromorphic systems. Proc. IEEE. **103**(8), 1379–1397 (2015)
24. Ivanov, D., Chezhegov, A., Larionov, D.: Neuromorphic artificial intelligence systems. Front. NeuroSci. **16**, 959626 (2022)
25. Kagan, B.J., Kitchen, A.C., Tran, N.T., Habibollahi, F., Khajehnejad, M., Parker, B.J., Friston, K.J.: In vitro neurons learn and exhibit sentience when embodied in a simulated game-world. Neuron. **110**(23), 3952–3969 (2022). e3958

26. Kalinin, K.P., Gladrow, J., Chu, J., Clegg, J.H., Cletheroe, D., Kelly, D.J., Falck, F.: Analog optical computer for AI inference and combinatorial optimization. Nature, 1–8. (2025)
27. Kanwisher, N.: Functional specificity in the human brain: a window into the functional architecture of the Mind. Proc. Natl. Acad. Sci. **107**(25), 11163–11170 (2010)
28. Killoran, N., Bromley, T.R., Arrazola, J.M., Schuld, M., Quesada, N., Lloyd, S.: Continuous-variable quantum neural networks. Physical Review Research **1**(3), 033063 (2019)
29. Levy, S.: The 'Godfather of AI' Has a Hopeful Plan for Keeping Future AI Friendly. WIRED Magazine (2023)
30. Mackin, C., Rasch, M.J., Chen, A., Timcheck, J., Bruce, R.L., Li, N., Nandakumar, S.: Optimised weight programming for analogue memory-based deep neural networks. Nat. Commun. **13**(1), 3765 (2022)
31. MacLennan, B.J.: A review of analog computing. Department of Electrical Engineering & Computer Science, University of Tennessee, Technical Report UT-CS-07-601 (September), 19798–19807. (2007)
32. McAleese, S.: Geoffrey hinton on the past, present, and future of AI. *Lesswrong*. (2024)
33. Mehonic, A., Ielmini, D., Roy, K., Mutlu, O., Kvatinsky, S., Serrano-Gotarredona, T., Balanov, A.G.: Roadmap to neuromorphic computing with emerging technologies. APL Materials (2024). https://doi.org/10.1063/5.0179424
34. Nelaturu, S.H., Ravichandran, N.K., Tran, C., Hooker, S., Fioretto, F.: On the fairness impacts of hardware selection in machine learning. arXiv preprint arXiv:2312.03886. (2023)
35. Nissenbaum, H.: Privacy as contextual integrity. Wash. L Rev. **79**, 119 (2004)
36. Novo, A., Lobon, F., Garcia de Marina, H., Romero, S., Barranco, F.: Neuromorphic perception and navigation for mobile robots: a review. ACM Comput. Surveys. **56**(10), 1–37 (2024)
37. Peters, J.: The Future of AI is Analogue. IAI News: (2024). https://iai.tv/articles/the-future-of-ai-is-analogue-auid-2917
38. Rasch, M.J., Mackin, C., Le Gallo, M., Chen, A., Fasoli, A., Odermatt, F., Tsai, H.: Hardware-aware training for large-scale and diverse deep learning inference workloads using in-memory computing-based accelerators. Nat. Commun **14**(1), 5282 (2023)
39. Robichaud, P., Wieland, J.W.: Responsibility: the epistemic condition, First edn. Oxford University Press, Oxford (2017)
40. Russell, S.: Human-Compatible: AI and the problem of control. Penguin Random House UK (2019)
41. Salles, A., Farisco, M.: Neuroethics and AI ethics: a proposal for collaboration. BMC Neurosci. **25**(1), 41 (2024)
42. Smirnova, L., Caffo, B.S., Gracias, D.H., Huang, Q., Morales Pantoja, I.E., Tang, B., Harris, T.D.: Organoid intelligence (OI): the new frontier in biocomputing and intelligence-in-a-dish. Front. Sci. **1**, 1017235 (2023)
43. Van Wynsberghe, A.: Sustainable AI: AI for sustainability and the sustainability of AI. AI Ethics **1**, 213–218 (2021)
44. Wang, H.: Analog chip paves the way for sustainable AI. In: Nature Publishing Group UK London (2023)
45. Yampolskiy, R.V.: Leakproofing singularity-artificial intelligence confinement problem. J. Conscious. Stud. JCS (2012)
46. Yuste, R., Goering, S., Arcas, B.A.Y., Bi, G., Carmena, J.M., Carter, A., Huggins, J.E.: Four ethical priorities for neurotechnologies and AI. Nature **551**(7679), 159–163 (2017)

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