Perspectives Beyond P(doom): Power, Uncertainty and Mundane Engineering in AI Ethics

Per Erik William Strandberg[1,*], Eduard Paul Enoiu[2] and Mirgita Frasheri[3]

[1]Westermo Network Technologies AB, Västerås, Sweden
[2]Mälardalen University, Västerås, Sweden
[3]Aarhus University, Aarhus, Denmark

Abstract

The informal concept P(doom), the probability of an AI triggered existential catastrophe, has become a visible part of AI safety discourse in scientific, engineering and public spheres. In this paper, we examine P(doom) from cultural, epistemic, and practical perspectives through which AI risk is made meaningful. Using collaborative autoethnography, we develop three situated perspectives: AI and power imbalances, deep uncertainty and decision making, and engineering culture and responsibility. Our synthesis suggests that conditions under which knowledge about AI risk is produced, controlled, trusted, and acted upon are important. Across the three accounts we identified cross-cutting themes concerning narratives, epistemic limits, socio technical responsibility, human judgment, and appropriate AI use. We argue that AI ethics should move toward understanding how to responsibly address uncertainty through democratic and institutional control, risk management, ethical formation, AI literacy, and reviewable engineering practices. The contribution is a situated reframing of P(doom) into a set of perspectives about power, knowledge, use and responsibility.

Keywords

philosophy of risk, ethics of artificial intelligence, decision theory, collaborative autoethnography

1. Introduction

The informal concept P(doom), the probability of an AI triggered existential catastrophe, has become a rather visible part of AI safety discourse [1, 2, 3]. It appears in scientific debates, public interviews, policy discussions, social media conversations, but also in engineering communities. The concept is quite attractive because it gives a seemingly simple form to a difficult question: how worried should we be about AI? Yet this simplicity can also be problematic, making a deeply uncertain future appear more knowable than it is.

In this paper we examine what becomes visible when P(doom) is viewed as a socio-technical problem involving people, processes, technologies, institutions, markets, and forms of knowledge. This is important as AI becomes increasingly embedded in everyday life and work. This socio-technical framing connects our work to ongoing debates in AI ethics [4, 5]. Much of AI ethics has been organized around principles such as fairness, transparency, privacy, accountability, human oversight, robustness, and societal well-being.

Previous work has also shown that principles are often difficult to translate into practice, especially in concrete contexts [6, 7, 8]. This paper builds on that concern by exploring how the problem of P(doom) can be reframed. The paper also connects to recent discussions in the TETHICS community, like big tech's capitalization of AI ethics [9], the fallacy of autonomous AI [10] and deskilling in human AI augmentation [11].

We use collaborative autoethnography from within overlapping research and professional communities concerned with AI ethics, risk, software engineering and socio-technical responsibility. Our aim is to develop an interpretive and analytical account of how P(doom) is made meaningful from

TETHICS'26: 9th Conference on Technology Ethics, November 11–12, 2026, Lahti, Finland
*Corresponding author.
per.strandberg@westermo.com (P. E. W. Strandberg); eduard.paul.enoiu@mdu.se (E. P. Enoiu); mirgita.frasheri@ece.au.dk (M. Frasheri)
0000-0003-1688-6937 (P. E. W. Strandberg); 0000-0003-2416-4205 (E. P. Enoiu); 0000-0001-7852-4582 (M. Frasheri)

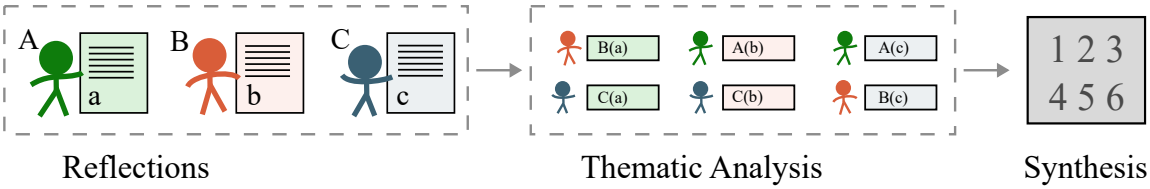

Figure 1: Overview of the method. First, each author (A, B, and C) wrote an individual reflection on P(doom) (a, b, and c). Second, each author made an analysis of the other reflections (B(a), C(a), ...). Third, the analyses were merged in a synthesis workshop where six topics emerged.

different positions. Through this approach, we reach the main contribution of this paper: six topics in the form of five cross-cutting themes of narratives under uncertainty, epistemic limits, socio-technical responsibility, human judgment and oversight, appropriate AI use, as well as the question of perspective – from democracy to mundane engineering. The paper explores the framing of P(doom) as a question about the conditions under which AI risk is known, communicated, governed, and acted upon. In addition, we connect such AI risk discourse with more immediate concerns in AI ethics, including power, regulation, risk management, ethical formation, monitoring, and engineering practice.

The rest of this paper is organized as follows. In Section 2 we cover the collaborative autoethnographic method and Section 3 contains our practical and professional perspectives. Results are synthesized in Section 4. Sections 5 and 6 discuss and conclude the paper.

2. Research method: Collaborative autoethnography

This paper uses a collaborative autoethnographic approach to examine how P(doom) is experienced and interpreted from three situated practice and research positions, see Figure 1. Autoethnography [12, 13] is a qualitative research method that uses personal experience to describe and interpret texts, experiences, beliefs, and practices. This method is appropriate for our study because we examine P(doom) here from both practical and professional perspectives, as well as a cultural narrative in technology. The authors are not external observers of AI development, use and discourse. We write from within overlapping research and professional communities concerned with AI ethics, risk, technology, software engineering and socio-technical responsibility.

Our own experiences can be situated accounts through which broader ethical and cultural patterns can be examined. They consist of three first-person reflective accounts written by the authors. Each account was developed from the author's own research background, professional experience, and engagement with AI-related debates and selected literature on AI ethics, catastrophic AI risk, risk management, and software engineering responsibility. The methodology included four steps:

- Each author wrote an individual reflective account responding to the same guiding question: *How does P(doom) appear from my research and professional position, and what does this perspective reveal about AI ethics?* The authors wrote this from their own standpoint rather than producing a neutral review. This allowed each perspective to show different assumptions, professional experiences and ethical concerns.
- After the initial accounts were written, each author conducted a separate thematic analysis of the other accounts. In this step, we annotated recurring ideas, assumptions and ethical concerns, while also noting differences between the perspectives.
- We then organized a synthesis workshop in which the authors compared and discussed their themes. The workshop focused on identifying shared patterns, disagreements and any cross-cutting themes across the three accounts. The session was recorded and transcribed using Microsoft Teams and Copilot, and the transcript was used as additional material to refine the analysis. Based on the individual thematic analyses, the workshop discussion and the transcript, we identified several cross-cutting themes across the three accounts.

- The individual accounts were developed through three iterations. The first versions were written before the synthesis workshop. They were then revised based on the reciprocal analyses and workshop discussion. A final revision was conducted after the conference peer review to clarify the accounts and their relationship to the shared themes. The rewriting served a reflexive purpose; it clarified how each position related to the themes without removing differences between the positions.

These three accounts are not intended to represent a general view of AI ethics or software engineering. Our contribution is instead interpretive and analytical with situated insight into how P(doom) is experienced and made meaningful across three distinct research positions. The approach also has limitations. Because we analyze our own experiences, the accounts are partial and shaped by our own positions. We address this through collaborative reading, explicit comparison across these perspectives, and engagement with existing literature shared among all authors. In addition, because all three authors work within technical, research and engineering communities, the accounts do not include the perspectives of policymakers, non-technical workers and user communities. Also, collaborative revision may smooth over differences between these accounts. We therefore try to retain explicit differences where possible and do not claim representativeness or thematic saturation.

3. Three situated accounts of P(doom)

This section presents three situated perspectives on P(doom). These perspectives provide an overview of how AI ethics can be interpreted from different research positions.

3.1. Mirgita's Perspective: AI and Power (Im)Balances

I write from the perspective of a computer engineer, as well as a researcher working on autonomous and multi-agent systems, robotics, and digital twins, wondering how the technologies we building are impacting the world, and contemplating on whether we are actually helping people, or just making fodder for Big Tech.

3.1.1. Ethics, Regulations, and Practical Realities

One of the main questions seems to be whether there is any point to AI ethics, from a real world, practical standpoint. Munn [14] argues precisely on the fact that while indeed we academics are busy with philosophical questions on the ethics of AI, how we can build such ethical agents etc., most of what is argued remains meaningless. For example, when considering AI principles such as beneficence, non-maleficence, and justice, among others, upon closer inspection, it becomes quite quickly clear that the definitions of what these terms mean are vague and ambiguous. In addition, Munn argues that, the development of AI happens in well documented toxic tech environments (with questionable practices and goals at best), and not in vacuum. And crucially, there are no teeth in these AI manifestos [15], as they are not meant to enforce consequences, rather provide a philosophical framework on what is ethical given cultural and historical settings. Even more so, Big Tech appears to capitalize on AI Ethics, exploiting the discourse for its own gains, rather than for purely ethical motivations [9].

Several regulations have been put forward, such as the GDPR [16] (2016), and the AI Act[1] (2024), by the European Union, and CCPA[2] (2018), CPRA[3] (2020), in California, that target aspects such as the privacy of individuals, and protection thereof. While these regulations represent important milestones in the societies where they apply, it is not necessarily clear what their reach is. E.g., in the GDPR it is stated that consent for data processing should be "freely given, specific, informed and unambiguous", as well as in the CPRA, where businesses have to "clearly inform consumers about how they collect and

[1] https://artificialintelligenceact.eu/
[2] https://theccpa.org
[3] https://thecpra.org

use personal information and how they can exercise their rights and choice". It is practically impossible to give consent when users navigate seas of online legal click-through documents everyday even for checking the weather, let alone dealing with AI bots [17]. In fact, Vardi argues that, in order to regulate Tech, we need to nullify such contracts [18]. The AI Act promotes a human-centric use of AI, as well as cross border, free movement of AI services. Are people being educated on these matters, and even given a say in how AI is to impact their lives?

I think it all comes down to the following question: Is it possible to regulate Big Tech in a meaningful way? We are at a point in time where these companies provide the digital infrastructures that underpin both private and public services, potentially acting as gatekeepers to such technology, with non-negligible social and market power [19]. In addition, some of these companies have a market worth more than GDPs of entire nations [20], e.g., as of November 2025, Apple's market capitalization hit 4.01 trillion dollars, outranked only by the GDPs of the US, China, Germany, and Japan at the time[4]. The comparison does not, of course, provide the full story, but it is quite telling that a handful of companies are worth more than most of the countries in the world. These companies are primarily driven by commercial interests and shareholder value, that are not necessarily aligned with broader societal interests, as well as the benefit of nations and peoples. They can influence elected officials and elections to bring forward candidates who will protect these interests. In addition, they have the financial upper hand for lobbying, and in certain countries, officials do not have to pledge not to receive gifts from lobbyists[5]. How is the nurse, teacher, blue collar worker, etc., to compete for representation? And if we are able to rein in Big Tech in developed rich countries, what happens to underdeveloped and developing countries?

Pitt [21], argues bravely on the abolition of Big Tech organizations, and compares them to the robber barons of the 19th century. What could have been tools for democracy, have denigrated to platforms for bots, trolls, and ragebait, for the benefit of some unnamed shareholders somewhere. Indeed he writes that

> "... digital enterprises have become increasingly monopolistic, individualistic beyond narcissistic to the point of lamentably solipsistic, devastatingly careless of human potential, and environmentally wasteful. For all that their revolutionary "move fast and break things" zeal is branded with the ostensible motives of freedom and liberty, the reality is private profit and unrestricted individual sovereignty for themselves, but techno-feudalism and modern indentured servitude for those deemed, damned and doomed as other or surplus."

Pitt is one of a few who actually points directly at the problem that stems from what Big Tech is and represents today. Others have delved into the weighing of the pros and cons of AI adoption [22], taking into consideration developing countries, providing hopeful insights on the new job opportunities that will be available, prompt engineer as a prominent example here (is this the new call centre job?), fundamentally neglecting the many realities of developing countries, where deep-seated corruption is a fact, with the number of emigrants simply increasing[6], many who are economically motivated. What chance do third-world countries with shaky pseudo democracies have to regulate these technologies, where it is evident that money will speak, no matter at what societal or environmental cost, where people already have a tough time regulating their governments, crucially failing to do so? Are we just instead witnessing a new era of techno-colonialism?

3.1.2. AI Everything, Everywhere all at Once

It would seem that, for better or worse, AI is to be integrated in everything we do. Nvidia's CEO has expressed that "Every young person should become an AI expert. It is an incredible technology that

[4] https://finance.yahoo.com/news/apple-richer-4-countries-164523140.html

[5] In the US, the executive order requiring the executive branch to sign ethics pledges regarding gifts from lobbyists has been rescinded https://www.yahoo.com/news/trump-just-made-bribing-politicians-214623958.html

[6] It is possible to see emigration trends for some countries in South-eastern, Eastern Europe and Central Asia from the United Nations migration platform https://seeecadata.iom.int/msite/seeecadata/

significantly lowers the knowledge barrier of any professional field. It is so easy to use—there's a reason why it's the most widely adopted technology in history. AI fluency will empower you and elevate your chosen craft."[7] OpenAI's CEO shares a similar point of view "The obvious tactical thing is just get really good at using AI tools."[8] Notably, it does not seem to be the case of understanding the underlying mechanics of the technology, but rather just using it. Perhaps the latter can lead to the former, the question remains, if we use AI to lower the knowledge barrier, skipping some of the steps of classical training, which required one to sit down and think first, can we get competent professionals able to critically appraise what the AI suggests? Are universities going to pump out prompt engineers or critical thinking individuals?

And we do not have to guess that much any longer, as studies are already coming out showing the impact of using AI on cognitive activities. Kosmyna et al. report on EEG results of people writing on given SAT topics, considering three groups: large language model (LLM) users, search engine users, and brain only users [23]. What they find is that the latter group has the highest connectivity, and is able to increase performance when switching to LLMs at a second stage. The opposite is the case for LLM users. Oakley et al. show empirical findings of how premature integration of AI during learning inhibits the process itself, with negative impacts on declarative and procedural memory [24]. So it seems that, similarly to the invention of the pocket calculator, even if we have and use it, we should still probably teach children how to do the math with pen and paper. Countries like Sweden, initially at the forefront of digital learning, are now going back to physical books instead[9]. Fundamentally, I think that we should not eat up any new technology without some reasonable resistance or scepticism, and look at big tech bros with rose coloured glasses, with profound but shallow admiration, and do whatever they tell us. Note here also, that when it was about their own children, the likes of Bill Gates, and Steve Jobs have raised them with limited access to tech (is it the classical do as I say not as I do?).[10] Big tech bros are thinking of their shareholders, not of societal consequences, they can potentially buy an island and retreat there should society collapse. The rest of us cannot afford such extravagance. This does not mean becoming a Luddite, but rather approaching the topic from the following question: What is AI's place, as a tool to help people, and not make them obsolete by taking over thinking and creative processes? Paradoxically, this is why we invest in technology, so we are able to have more time to think and be creative[11].

Finally, I would like to note on a current research trend that aims to build mechanisms for providing guarantees for AI (specifically LLMs), and challenge this effort. The nature of the tool is stochastic; it is able to learn from huge amounts of data, make rather good predictions some of the time, and hallucinate profoundly other times. Should we then, instead of throwing it at every problem, adopt it in those places where it would be most useful and where it is strongest, such as synthesizing large amounts of text, as an additional tool for brainstorming, using the stochastic nature and even possible hallucinations, as a trigger for our own creative processes?

To use AI to replace for example therapists, youth counsellors, or even safety mechanisms in a power plant, or even as a digital friend for children etc., one would need to provide strict guarantees on its behaviour. It is not clear that such effort is warranted. Instead, let us find appreciation for existing methods that have been proven to do the job well, without the extra cost of training and processing power. Just because we have a hammer, it does not mean that we should see everything as a nail.

[7] https://fortune.com/2026/04/29/nvidia-ceo-jensen-huang-engineering-path-to-success-ai-era-gen-z-advice-ieee-medal-of-honor-winner/

[8] https://www.businessinsider.com/sam-altman-ai-coding-job-market-software-engineer-students-openai-2025-3

[9] https://www.bbc.com/news/articles/cly0vk77vdko

[10] https://fortune.com/2026/02/21/peter-thiel-bill-gates-steve-jobs-steve-chen-tech-billionaires-publicly-shielding-their-children-from-tech-products-social-media/

[11] The reader is also pointed to the commentary of Diaconescu, for a discussion on efficiency vs creativity when it comes to socio-technical systems [25].

3.2. Per's Perspective: P(doom), deep uncertainty and decisions

I work as a company-employed project manager in research projects with industry-academia-collaborations. The company makes cyber-physical systems and has about 600 employees. During my time as an industrial doctoral student I grew a personal off-work interest in research ethics leading to a more general interest in applied ethics and philosophy. After spending three months in Silicon Valley and meeting AI developers at tech giants, and also leading AI champions at the company, I write from the perspective of corporate AI adoption with a focus on philosophy of risk and ethics.

3.2.1. Statements on P(doom) are prophecies, not facts

It seems knowledgeable people, in their speculation of the value of P(doom) give an average of 5-10%. Each such statement, e.g. "I think P(doom) is 5%" is not a prediction, rather a prophecy - and often from people with agendas. Deutsch [26] argues that trustworthy statements of the future come from an explanatory power in knowledge. Genuine understanding can be collected from operational data of nuclear power plants under normal conditions, combined with various durability calculations and risk analyses, as well as from commissions after minor incidents or catastrophic meltdowns. For classic nuclear power plants, the technology has not changed radically in many decades - we have good explanatory theories and models for how they work. We can describe very well the probabilities of operational disruptions. Similarly, when Covid-19 came, we had a good understanding of what an airborne virus is and different biological models for the spread of infection and protection. Even though the virus was new, we had genuine knowledge that explained systems and mechanisms to us. However, on the AI front, knowledge is growing rapidly and it is partly a new domain for us. We lack the operational history and the explanations that could come from it.

Hendrycks et al. [27] mention four types of catastrophic AI risks: antagonistic use of evil AI agents, AI races such as autonomous arms races, organizational risks where companies with weak security cultures cause a disaster, or that AIs get their own evil goals or goals that are not in line with human goals ("misalignment"). An example of the fourth scenario is a destruction of all humans as they stand in the way of making as many paper-clips as possible (the "paperclip maximizer" scenario is from an interview with Bostrom [28]).

A similar observation to that mentioned by Deutsch is the tuxedo fallacy described by Hansson [29]. This fallacy involves believing that one is in a situation with known outcomes, all of which have known probabilities, i.e. like in a tuxedo at a casino. Instead, one finds oneself in a situation with unknown outcomes and unknown probabilities, a bit more like in an alien jungle on an unknown planet with unknown creatures that have an unknown probability of wanting to eat people. When we do not know probabilities or even all outcomes, it is appropriate, Hansson believes, to think about classic engineering principles: like when we design bridges for higher loads than anticipated, and to think about human error.

My point with bringing up Deutsch and Hansson is that people making prophecies about P(doom) don't know what they are talking about, their statements are worthless as they are not anchored in explanatory power. For P(doom), the tuxedo is not applicable, the possible outcomes and their probabilities are not known.

3.2.2. We need (boring) risk management

Marchau et al. [30], describe a levelled model for uncertainty. At level 1, we are basically in a casino with known probabilities for known possible outcomes. Level 2 is about systems that can be described with probabilities or alternative futures that can be described with probability intervals. At level 3 we can name outcomes and also bound or rank their probabilities (*this* is more likely than *that*). At level 4 we reach deep uncertainty: we can name multiple plausible futures, but as we don't understand system mechanisms, the probabilities cannot be predicted. At level 5, we have unknown unknowns, we cannot name or rank possible futures. From the perspective of P(doom), Levels 3 and 4 are particularly relevant. As mentioned, experts have opinions on what the scenarios could be, their mechanisms, and

also propose likelihoods. For these types of uncertainties, Marchau et al. suggest using robust strategies. These do not maximize expected utility based on uncertain probabilities, instead the goal ought to be to minimize maximum regret: a good decision is one that yields a favorable outcome over most scenarios. In retrospect we aim then to minimize the feeling of "we should have done it differently."

Furthermore, for uncertainties at level 3 the authors advocate traditional scenario analysis. Unfortunately, it seems that scenario analysis is often poor at predicting the future in such a way that the analyses do not capture disruptive scenarios, since we often extrapolate from known knowledge instead of thinking about disruptive technology such as large language models are affecting society. However, in the case of P(doom) I would argue that we have largely already identified many potential disasters. With these scenarios we could apply classic (boring) risk management practices of supervision, mitigation, prevention: e.g., a guard-rail around a toxic LLM could prevent encouragement to suicide in a similar way as a reinforced bridge prevents collapse. And if an LLM was found to try to replicate itself, then its plug should be pulled, and its servers disconnected from the internet. This is similar to setting up a quarantine for a ship with an infected crew. In other words: we minimize the worst consequences of the harmful event.

3.2.3. We need to build ethical developers, and it can be done

Three dominating schools of thought related to ethics are duty ethics, consequentialism and virtue ethics. Duty ethics looks at the reasoning *behind* an act; I must not lie, because if I am allowed to lie, then everyone could lie, and when everyone always lies, there cannot be truth or even a meaningful society. Consequentialism looks at the consequences *following* an act: Ideally converting the harms and benefits for every living person (and probably also every inanimate object) now and in the future from today until the end of the universe. A more relaxed consequentialism allows for following rules of thumb: usually, kicking dogs brings more harm and pain to the universe than it gives pleasure, so we ought to have a rule against kicking dogs (as opposed to having to do a separate utility calculation every time someone contemplates kicking a dog). It seems that both duty ethics and consequentialism strive to reduce ethics to following rules, to "checklistification." Checklists are great, sometimes, like when you're a pilot and you observe a rare condition in the plane that some safety engineer has already specified a solution to (in the form of a checklist). For AI ethics, Hayes et al. [31] argue that these principles are not meaningful, the technology changes faster than we can write principles, by the time the system user observes a problem, the checklist might be obsolete. Therefore, they argue (and I agree) that we ought to focus on virtue ethics: by exploring not the reasoning behind the act, or its consequences, but by focusing on the actor or agent. In an ambitious project on virtues and technology, Vallor [32] builds a technomoral system of global virtue ethics. She argues that virtue ethics has been adopted promisingly in environmental ethics, bioethics, media ethics, and business ethics, and she moves on to technosocial ethics. She is clearly not against technology in general, instead she argues that "to be antitechnology is … to be antihuman." Of particular interest is her suggestion on moral habituation and her six approaches to it: relational understanding, reflective self-examination, and intentional self-direction, moral attention, prudential judgment, and appropriate extension of moral concern.

Someone might object to the idea that AI developers ought to get virtue ethics training, it might seem like I am fighting windmills with my opinion. However, a controversial organization like Volkswagen could reinvent itself following the 2015 "Dieselgate" emissions fraud scandal (that cost them about 30 billion Euro [33]). They first implement business ethics training and now AI ethics training. In fact, Volkswagen's ethical principles for AI, under the slogan of "We value heartbeats over bytes" came into place in their AI training program (covering not only ethics) and has now reached 130 000 people [34]. If Volkswagen could do it, then why couldn't the tech giants in Silicon Valley? (In my anecdotal experience, the developers at these companies are very aware and care deeply about the techno-social impact of AI, e.g. the water consumption [35]. However, they might be repressed by toxic company culture, or toxic business goals.)

Like many before me, I argue that there is a problem with engineers and ethicists belonging to two

cultures that never meet or even speak the same language [36]. An ethics training grounded in a company context that considers not only the principles behind decisions or consequences of actions, but also the inner states, the virtues, of the decision-maker is needed.

3.3. Eduard's Perspective: Software engineering culture and AI responsibility

I write from the perspective of a researcher working in software testing and AI-assisted software engineering practices. From this position, AI risk does not first appear to me as an abstract future catastrophe. It appears in the everyday culture of software engineering: in what engineers trust, what they question, what they review, and what they gradually stop noticing. This does not mean that catastrophic AI risk is irrelevant to AI engineering or to the use of AI in engineering; rather, such risk becomes meaningful also when it is connected to the practices through which systems are built, justified, maintained, and governed. For me, P(doom) is also useful for asking what kinds of engineering communities and cultures can support responsible action under uncertainty.

3.3.1. Narratives and AI engineering myths

Under uncertainty, narratives, heuristics and simplified accounts are unavoidable; therefore, I do not see the existence of narratives about AI catastrophe as the main problem with P(doom). The problem begins when these stories become unquestioned. A P(doom) number, an AI hype story or an AI myth can create false confidence when it is repeated often enough without being connected to evidence, practice, and responsibility. This is similar to what we already see in software engineering communities. We often hear claims such as "*AI will solve everything*", "*AI will replace testers*", or "*software engineering is over*" [37, 38, 39]. These statements are not only predictions. They are cultural signals. Such narratives seem to simplify a confusing transition. They also risk making engineers less attentive to the concrete conditions under which AI-assisted work is reliable, understandable, and accountable.

In the same way, P(doom) can be useful or harmful depending on whether it opens a discussion about uncertainty and responsibility or it turns into a symbol rather than an argument. If it is treated as a sign of inevitable catastrophe, it may create resignation. But if it is dismissed as pure science fiction, it may prevent serious engagement with current risks. In both cases, the narrative can displace more practical questions: What evidence supports an AI-generated software engineering artifact? Which assumptions are embedded in it? What happens when the speed and volume of artifact generation exceed the capacity for meaningful review? From my perspective, the value of P(doom) discourse lies in whether it prompts such questions in software engineering practice, since narratives can shape how AI risks, responsibilities and possible responses are understood. For example, conceptions of AI as autonomous can obscure the human agency and responsibility involved in how AI systems are built and deployed [10].

3.3.2. Human responsibility under uncertainty

In software engineering research, many activities can now be partially automated, including coding, testing, debugging, mining software repositories, generating benchmarks, running experiments, and drafting results [40]. However, theory building remains difficult to automate. This involves deciding what matters, why a phenomenon occurs, what a result means and which questions are worth asking in the first place [41]. These activities require argumentation, judgment, contextual understanding and the ability to connect empirical observations to broader explanations [42].

This leads me to view P(doom) discourse as a socio-technical challenge for responsible software engineering. It raises questions about how AI is developed and the institutions, practices, and engineering communities that support responsible action under uncertainty. From my perspective, the relevant question is what kinds of engineering cultures can support responsible action when the future cannot be fully known. A useful starting point is to make AI-assisted software engineering work inspectable, traceable, contestable, monitored and reversible.

3.3.3. Mundane mechanisms of risk in AI-Assisted software engineering

My experience with AI-powered software test generation suggests that many of the important research questions arise in quite mundane places. They appear when a generated test is accepted without much review, when an AI suggestion is trusted because it looks plausible, when evidence is insufficient, when no one knows who is responsible for an AI generating such artifacts or even when automation creates a false sense of quality [43, 44]. These are not spectacular apocalyptic scenarios, but they are the everyday mechanisms through which responsibility can become unclear and a practical solution is needed. They are mundane because they occur in routine software engineering workflows. When repeated across teams and communities, weakly understood artifacts and weak oversight can accumulate into organizational dependencies and governance gaps. This is one way in which engineering decisions become relevant to broader questions of AI risk.

3.3.4. What can software engineering contribute?

Software engineering can make a concrete contribution by translating broad ethical aspirations into reviewable practices. It offers practical mechanisms for working under uncertainty, including explicit rationale, argumentation and review, traceability, monitoring, human oversight, and stakeholder involvement [45, 46, 47]. From my perspective, a useful way to approach P(doom) is to examine how AI is already entering everyday practice. In my own work, this is most visible in software testing, test automation, and AI-assisted engineering decision-making. AI is no longer only discussed as part of future scenarios or speculation. It is already becoming part of how software engineering research is conducted, and how software is designed, tested, justified, monitored, and governed. This makes software engineering a practical site where broader questions about AI risk, responsibility, and human judgment become concrete.

3.3.5. Agentic workflows and engineering culture

I experience AI in software engineering as a transformation from writing and reviewing software artifacts to AI agents that generate and maintain artifacts in orchestration with humans. The software engineer is a coordinator of agents, tools, prompts, tests, pipelines and reviews rather than someone who crafts executable software artifacts [48]. An AI assistant may soon generate more code, tests, reports and metrics faster than humans can meaningfully interpret. The danger is bad code, as it always was, but also the loss of understanding and de-skilling [11]. This creates a new kind of socio-technical debt, ranging from messy, complex code to unclear rationales, invisible assumptions and unreviewed AI-assistant decisions. This makes engineering judgment and oversight more important, not less. Software engineering practices can make AI-generated claims contestable through testing, review, traceability, monitoring and documented evidence.

Responsible software engineering culture should make such socio-technical debt visible and negotiable. Drawing on Sharp et al. [49], I use *engineering culture* to refer to the shared and often implicit beliefs, values and practices within the software engineering communities that shape how evidence is assessed, which tools and authorities are trusted, how decisions are challenged, and how responsibility is assigned. Agentic software engineering broadens this problem by introducing AI that can plan activities, use tools and generate or modify software artifacts with varying degrees of autonomy. Consequently, software engineering is increasingly concerned with how AI-assisted work is coordinated, how knowledge is transferred, and how semi-executable systems composed of natural language, code, tools, policies, and workflows are created and maintained [48]. From this perspective, the engineer is not only a developer of software artifacts, but also a reviewer, interpreter, orchestrator and controller of AI-generated artifacts. However, meaningful oversight depends on whether engineering teams have sufficient time, evidence, authority and transparency to challenge AI outputs despite organizational incentives and tight deadlines.

4. Cross-perspective analytical synthesis

In this section, we describe the findings from the synthesis workshop (described in Section 2). The analysis reveals six topics: P(doom) as a narrative under uncertainty, knowledge and epistemology, responsibility is socio-technical in nature, human judgment and oversight, appropriate use of AI, and a reflection on the scale at which P(doom) is relevant (from a macro level where it is a global democracy issue, to micro in mundane engineering tasks).

A first cross-cutting theme is that *P(doom) functions as a narrative under uncertainty*. The problem is perceived as beginning when such narratives become detached from evidence, practice, and responsibility. An AI hype narrative or the claim that AI should be used everywhere can create false certainty. Also, an apocalyptic narrative can create resignation. If AI is imagined as an unstoppable future catastrophe, then humans could become bystanders rather than responsible actors. In this sense, P(doom) is both a statement about the future and also a cultural signal that affects what kinds of action appear reasonable in the present. It can also shift attention toward apocalyptic narratives, while immediate harms such as the erosion of democratic processes, the concentration of power, and the reduced ability of humans and institutions to contest AI-made decisions could become less visible.

A second theme concerns *knowledge and epistemology*. All three perspectives ask, but in different ways, what kind of knowledge about AI risk is possible: who controls that knowledge, and what should be done when knowledge is incomplete. From the perspective of deep uncertainty, P(doom) is problematic because it presents unknown futures. From a power perspective, the issue can also be seen as political. If AI systems become dominant interfaces to finding and processing information, then the actors who control AI infrastructures may also gain power over what counts as knowledge. From an engineering perspective, the epistemic problem also seems to be relevant in everyday practice, since plausibility is not the same as correctness, accountability, or understanding of what information is generated and reviewed using AI.

A third theme is that *responsibility is socio-technical in nature* and none of the accounts treat AI responsibility as something that can be reduced to individuals, technical safeguards or abstract ethical principles alone. Responsibility is distributed across developers, users, organizations, educational institutions, regulators, states, and technology companies. This distribution creates a practical difficulty, since many actors are involved and responsibility can easily become diluted.

A fourth theme concerns *human judgment and oversight*. At the macro level, this appears as a problem of public literacy and democratic agency. At the meso level, it appears as the need for robust decision making, ethical formation, and risk management under uncertainty. At the micro level, the need is shown for engineers and teams to remain capable reviewers and orchestrators of AI agents. The central issue is if humans can retain the competence, time, authority and support required to challenge AI outputs.

A final fifth theme relates to the *appropriate use of AI* and all accounts share a resistance to the assumption that AI should be applied everywhere simply because it is available. This is supported by tasks that require smaller models, conventional tools, human deliberation, or no AI system at all. This point connects the technical suitability of AI with ethics, sustainability, and democratic control. One could argue that asking where AI should not be used is maybe as important as asking how AI can be made safer where it is used.

Sixth, during the synthesis workshop, authors recognized that *differences between the perspectives are equally important*. They show that P(doom) is interpreted through different forms of knowledge (e.g., political knowledge, risk knowledge, engineering knowledge, and even moral knowledge). These differences are analytically productive because they reveal that AI risk cannot be understood on a single scale (e.g., societal, planetary scale). The *three perspectives differ mainly in the scale of the mentioned concerns*. One stresses democratic power and institutional means, another focuses on decision-making under deep uncertainty, and the third examines everyday practice. Together, they show that the problem with P(doom) is about what this narrative hides or reveals about power, uncertainty, responsibility, and appropriate AI use. A response could require democratic institutions, more robust risk strategies, ethical education, and practices that keep AI systems reviewable, traceable, and open to human judgment.

5. Discussion

This framing also contributes to debates about the principle practice gap in AI ethics. Some of the AI ethics work has been criticized for remaining at the level of broad principles that are difficult to operationalize in real contexts [14, 15]. Prem's review of approaches from ethical AI frameworks to tools is useful here because it shows that bridging the gap requires more than declaring principles; it requires methods, instruments, organizational practices [6]. Our paper supports this view but adds that ethical practice also depends on power relations, institutional and professional cultures. The challenge is therefore to combine ethical principles, practical tools, risk management, and professional formation.

The risk perspective shows that uncertainty does not justify inaction. On the contrary, uncertainty can justify robust action. This connects to approaches in risk management, including scenario analysis, monitoring, redundancy, human oversight, and mitigation [29, 30, 50]. These practices require attention to mechanisms, vulnerabilities, and consequences, but not only from a technical perspective.

This paper is based on three situated perspectives. The collaborative autoethnographic method provides some interpretive depth and the accounts are also shaped by the three professional backgrounds of the authors. In addition, each perspective tackles P(doom) at a different level, from the macro to the meso, and finally to the micro view. It is apparent that all three angles are required to understand and conceptualize AI risk, prompting the need for a collection of mechanisms, from regulations and their enforcement to risk analysis to the education of the public and engineers. Three complementary dimensions are the ones often associated with a socio-technical perspective (people, processes, and technology), and we argue that the challenges coming from P(doom) cannot be reduced to only a question of people, or of processes, or of technology. As many have argued before us, a holistic perspective is needed.

We argue that education and moral habituation of developers and AI users are critical. This is a debate at least as old as Snow's classic paper on the two cultures [36]. More recently, work on teaching AI Ethics to engineers [51, 52] reveals that doing so ought to be done from a holistic perspective with case studies or group projects – again, the topic of perspective and level comes back.

At the macro level, the debate on P(doom) raises questions on governance of AI, as well as AI for governance [53, 54, 55]. In both cases, classic ethical principles such as transparency, accountability, and the role of the human-in-the-loop matter.

The idea from Eduard's perspective on socio-technical debt (STeD) is a relatively recent construct, but it has been partially addressed in previous work [56, 57, 58]. It helps connect the rather mundane level of AI-assisted engineering to broader questions of AI responsibility. In a recent doctoral thesis, Persson found that organizations struggle more from social debt than from technical challenges (e.g. from governance fragmentation, procurement processes or organizational inertia) [56]. Herath and Ahmad, [57] found that topics such as personal essence and growth, collaborative perseverance, and organizational dynamics could have an impact on STeD, e.g. by not allowing time for reflection and knowledge sharing. An empirical evaluation suggests that STeD can be addressed using debt stories (similar to user stories) [58]. In our context, STeD represents a specific risk of AI, with systems becoming easier to generate than to understand, review, justify, or govern.

Future work should extend this through interviews with practitioners involved in this problem. It could also examine how P(doom) circulates in specific communities, how AI risk narratives influence engineering decisions, and how organizations translate AI ethics principles into concrete practices.

6. Conclusion

This paper has examined P(doom), the probability of an AI triggered existential catastrophe, as a cultural, epistemic, and practical framing of AI risk. Through collaborative autoethnography, we developed three situated perspectives on AI and power imbalances, deep uncertainty and decision making, and engineering culture and responsibility. The synthesis identified six topics in the form of five shared themes and an important difference: P(doom) functions as a narrative under uncertainty, it relates

knowledge and epistemology, responsibility for P(doom) is socio-technical in nature, we need human judgment and oversight, we also need an appropriate use of AI, and finally we need to reflect on the scale at which P(doom) is relevant: at a macro level it can be a global democracy issue, and on a micro level we can observe it in mundane engineering tasks. We therefore suggest that the value of P(doom) lies in leading us to ask about these six topics. While less visible than the apocalypse, these topics give guidance here and now. In this sense, moving beyond P(doom) means asking about how power, uncertainty, and mundane engineering shape AI ethics.

Acknowledgments

Eduard Paul Enoiu received support from Software Center and the AI and Society Fellowship.

Declaration on Generative AI

During the preparation of this work, the author(s) used: Google Gemini and Microsoft 365 Copilot (GPT-5) for checking grammar, spelling, explanations and reviewing; Copilot was also used for transcribing speech. These were not used for writing, to make final analytical or interpretive decisions. After using these tools/services, the authors reviewed and edited the content as needed and took full responsibility for the publication's content.

References

[1] S. Field, Why do Experts Disagree on Existential Risk and P(doom)? A Survey of AI Experts, arXiv preprint arXiv:2502.14870 (2025). doi:10.48550/arXiv.2502.14870.

[2] R. E. Guingrich, M. S. Graziano, P(doom) Versus AI Optimism: Attitudes Toward Artificial Intelligence and the Factors That Shape Them, Journal of Technology in Behavioral Science (2025) 1–19. doi:10.1007/s41347-025-00512-3.

[3] J. Danaher, Why AI Doomsayers are Like Sceptical Theists and Why it Matters, Minds and Machines 25 (2015) 231–246. doi:10.1007/s11023-015-9365-y.

[4] A. Jobin, M. Ienca, E. Vayena, The global landscape of AI ethics guidelines, Nature machine intelligence 1 (2019) 389–399. doi:10.1038/s42256-019-0088-2.

[5] L. Floridi, J. Cowls, A Unified Framework of Five Principles for AI in Society, Machine Learning and the City: Applications in Architecture and Urban Design (2022) 535–545. doi:10.1002/9781119815075.ch45.

[6] E. Prem, From ethical AI frameworks to tools: a review of approaches, AI and Ethics 3 (2023) 699–716. doi:10.1007/s43681-023-00258-9.

[7] K.-K. Kemell, V. Vakkuri, F. Sohrab, How Do AI Ethics Principles Work? From Process to Product Point of View, Conference on Technology Ethics (TETHICS'23), CEUR Workshops 3582 (2023).

[8] K.-K. Kemell, J. K. Nurminen, V. Vakkuri, Monitoring Machine Learning Systems from the Point of View of AI Ethics, Conference on Technology Ethics (TETHICS'24), CEUR Workshops 3901 (2024).

[9] J. Parviainen, Big Tech's Capitalization of AI Ethics: Reflections on the Politicization of AI Ethics, Conference on Technology Ethics (TETHICS'24), CEUR Workshops 3901 (2024).

[10] D. Schlienger, The Fallacy of Autonomous AI, Conference on Technology Ethics (TETHICS'24), CEUR Workshops 3901 (2024).

[11] F. Huseynova, Addressing Deskilling as a Result of Human-AI Augmentation in the Workplace., Conference on Technology Ethics (TETHICS'24), CEUR Workshops 3901 (2024).

[12] S. J. Cunningham, M. Jones, Autoethnography: a tool for practice and education, in: Proceedings of the 6th ACM SIGCHI New Zealand chapter's international conference on Computer-human interaction: making CHI natural, 2005, pp. 1–8. doi:10.1145/1073943.1073944.

[13] C. Ellis, T. E. Adams, A. P. Bochner, Autoethnography: an overview, Historical Social Research 36 (2011) 273–290. doi:10.12759/hsr.36.2011.4.273-290.

[14] L. Munn, The uselessness of ai ethics, AI and Ethics 3 (2023) 869–877. doi:10.1007/s43681-022-00209-w.
[15] A. Rességuier, R. Rodrigues, AI ethics should not remain toothless! A call to bring back the teeth of ethics, Big Data & Society 7 (2020). doi:10.1177/2053951720942541.
[16] EU, Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation) (Text with EEA relevance), 2016.
[17] M. Frasheri, A. Bakhtiarnia, L. Esterle, A. Iosifidis, Are LLM-based methods good enough for detecting unfair terms of service?, arXiv preprint arXiv:2409.00077 (2024).
[18] M. Y. Vardi, To Regulate Tech, Nullify Click-Through Contracts, Communications of the ACM 66 (2023) 5. doi:10.1145/3609981.
[19] K. Birch, K. Bronson, Big tech, Science as Culture 31 (2022) 1–14. doi:10.1080/09505431.2022.2036118.
[20] Y. Tokat, The Big Tech versus the Nation-States: Clash of Economic Interests and Struggle to Compete on a Global Scale, Eurasian Conferences on Language & Social sciences, 2022. doi:10.2139/ssrn.4539475.
[21] J. Pitt, On the abolition of all big tech organizations, IEEE Technology and Society Magazine 44 (2025) 17–22.
[22] N. R. Mannuru, S. Shahriar, Z. A. Teel, T. Wang, B. D. Lund, S. Tijani, C. O. Pohboon, D. Agbaji, J. Alhassan, J. Galley, et al., Artificial intelligence in developing countries: The impact of generative artificial intelligence (ai) technologies for development, Information development 41 (2025) 1036–1054.
[23] N. Kosmyna, E. Hauptmann, Y. T. Yuan, J. Situ, X.-H. Liao, A. V. Beresnitzky, I. Braunstein, P. Maes, Your brain on chatgpt: Accumulation of cognitive debt when using an ai assistant for essay writing task, arXiv preprint arXiv:2506.08872 4 (2025).
[24] B. Oakley, M. Johnston, K.-Z. Chen, E. Jung, T. J. Sejnowski, The memory paradox: Why our brains need knowledge in an age of ai, arXiv preprint arXiv:2506.11015 (2025).
[25] A. Diaconescu, Efficiency versus creativity as organizing principles of socio-technical systems: Why do we build (intelligent) systems?[commentary], IEEE Technology and Society Magazine 38 (2019) 13–22.
[26] D. Deutsch, The Unknowable: how to prepare for it, TEDxBrussels, YouTube video, 2011. URL: https://youtu.be/SVgGYQ_5ID8, accessed: 2026-03-22.
[27] D. Hendrycks, M. Mazeika, T. Woodside, An overview of catastrophic ai risks, arXiv preprint arXiv:2306.12001 (2023).
[28] K. Miles, Artificial Intelligence May Doom The Human Race Within A Century, Oxford Professor Says, Huffington Post, 2014. URL: https://www.huffpost.com/entry/artificial-intelligence-oxford_n_5689858, accessed: 2025-12-30.
[29] S. O. Hansson, From the casino to the jungle: Dealing with uncertainty in technological risk management, Synthese 168 (2009) 423–432. doi:10.1007/s11229-008-9444-1.
[30] V. A. W. J. Marchau, W. E. Walker, P. J. T. M. Bloemen, S. W. Popper, Decision Making under Deep Uncertainty: From Theory to Practice, Springer Nature, Cham, Switzerland, 2019. doi:10.1007/978-3-030-05252-2, open Access.
[31] P. Hayes, N. Fitzpatrick, J. M. Ferrández, From applied ethics and ethical principles to virtue and narrative in AI practices, AI and Ethics 5 (2025) 1217–1239. doi:10.1007/s43681-024-00472-z.
[32] S. Vallor, Technology and the Virtues: A Philosophical Guide to a Future Worth Wanting, Oxford University Press, New York, 2016. doi:10.1093/acprof:oso/9780190498511.001.0001.
[33] G. B. Turna, How "Dieselgate" Changed Volkswagen: Rushing to Erase the Traces of Greenwashing, in: Socially responsible consumption and marketing in practice: Collection of case studies, Springer, 2022, pp. 255–273. doi:10.1007/978-981-16-6433-5_16.
[34] RCR Wireless News, Volkswagen's AI drive: Powering the future of automotive innovation, https://www.rcrwireless.com/20250910/industry-4-0/volkswagen-ai-drive, 2025. Accessed:

2026-05-03.

[35] N. Jegham, M. Abdelatti, C. Y. Koh, L. Elmoubarki, A. Hendawi, How Hungry is AI? Benchmarking Energy, Water, and Carbon Footprint of LLM Inference, arXiv preprint (2025). doi:10.48550/arXiv.2505.09598.

[36] C. P. Snow, Two cultures, Science 130 (1959) 419–419. doi:10.1126/science.130.3373.419.

[37] S. J. Bronner, Folklore and Ethnology of the Modern World: A Theory of Cultural Engagement, Taylor & Francis, 2025. ISBN: 9781040763612.

[38] V. Antinyan, M. Grah, Fashion-driven software engineering: Industry experiences of form over substance, IEEE Software (2025). doi:10.1109/MS.2025.3575696.

[39] E. Enoiu, J. Malm, G. Gay, Folklore in software engineering: A definition and conceptual foundations, arXiv preprint arXiv:2601.21814 (2026).

[40] I. Ahmed, A. Aleti, H. Cai, A. Chatzigeorgiou, P. He, X. Hu, M. Pezzè, D. Poshyvanyk, X. Xia, Artificial intelligence for software engineering: The journey so far and the road ahead, ACM Transactions on Software Engineering and Methodology 34 (2025) 1–27. doi:10.1145/3719006.

[41] A. Zeller, IEEE Computer Society Harlan D. Mills Award and Talk by Andreas Zeller: Should Computer Scientists Experiment Less? On the Past, Present, and Future of Software Engineering Research, ICSE 2026 Main Plenaries, 2026. URL: https://conf.researchr.org/details/icse-2026/icse-2026-main-plenaries/8/IEEE-Computer-Society-Harlan-D-Mills-Award-and-Talk-by-Andreas-Zeller-Should-Comput, conference talk, ICSE 2026, Rio de Janeiro, Brazil, 15 April 2026. Accessed: 2026-05-08.

[42] W. C. Booth, G. G. Colomb, J. M. Williams, The craft of research, University of Chicago press, 2009. ISBN: 9780226062648.

[43] A. Aleti, Software Testing of Generative AI Systems: Challenges and Opportunities, in: 2023 IEEE/ACM International Conference on Software Engineering: Future of Software Engineering (ICSE-FoSE), IEEE, 2023, pp. 4–14. doi:10.1109/ICSE-FoSE59343.2023.00009.

[44] N. Alshahwan, M. Harman, A. Marginean, Software testing research challenges: An industrial perspective, in: 2023 IEEE Conference on Software Testing, Verification and Validation (ICST), IEEE, 2023, pp. 1–10. doi:10.1109/ICST57152.2023.00008.

[45] A. Nguyen-Duc, B. Cabrero-Daniel, A. Przybylek, C. Arora, D. Khanna, T. Herda, U. Rafiq, J. Melegati, E. Guerra, K.-K. Kemell, et al., Generative Artificial Intelligence for Software Engineering – A Research Agenda, Software: Practice and Experience 55 (2025) 1806–1843. doi:10.1002/spe.70005.

[46] P. E. Strandberg, E. P. Enoiu, M. Frasheri, Ethical challenges and software test automation, AI and Ethics 5 (2025) 6185–6206. doi:10.1007/s43681-025-00804-7.

[47] D. Lo, Trustworthy and Synergistic Artificial Intelligence for Software Engineering: Vision and Roadmaps, in: 2023 IEEE/ACM International Conference on Software Engineering: Future of Software Engineering (ICSE-FoSE), IEEE, 2023, pp. 69–85. doi:10.1109/ICSE-FoSE59343.2023.00010.

[48] R. Feldt, Keynote: Agentic Software Engineering Will Eat the World: AI-Based Systems as the New Operating System of Society, International Workshop on Agentic Engineering, AGENT 2026, co-located with ICSE 2026, 2026. URL: https://conf.researchr.org/details/icse-2026/agent-2026-papers/31/Keynote-Agentic-Software-Engineering-Will-Eat-the-World-AI-Based-Systems-as-the-New, keynote talk, Rio de Janeiro, Brazil, 14 April 2026. Accessed: 2026-05-08.

[49] H. Sharp, H. Robinson, M. Woodman, Software engineering: community and culture, IEEE Software 17 (2000) 40–47.

[50] International Organization for Standardization, Risk management — Risk assessment techniques, iso standard no. 31010:2019 ed., International Organization for Standardization, Geneva, Switzerland, 2019. URL: https://www.iso.org/standard/72140.html.

[51] L. Tuovinen, A. Rohunen, Teaching AI ethics to engineering students: Reflections on syllabus design and teaching methods, Conference on Technology Ethics (TETHICS'21), CEUR Workshops 3069 (2021).

[52] L. J. Wiese, I. Patil, D. S. Schiff, A. J. Magana, AI ethics education: A systematic literature review,

Computers and Education: Artificial Intelligence 8 (2025) 100405. doi:10.1016/j.caeai.2025.100405.

[53] A. Ghosh, A. Saini, H. Barad, Artificial intelligence in governance: recent trends, risks, challenges, innovative frameworks and future directions, AI & SOCIETY 40 (2025) 5685–5707. doi:10.1007/s00146-025-02312-y.

[54] M. Nieminen, N. Gotcheva, J. Leikas, R. Koivisto, Ethical AI for the governance of the society: Challenges and opportunities, Conference on Technology Ethics (TETHICS'19), CEUR Workshops 2505 (2019) 20–26.

[55] S. Pasas-Farmer, R. Jain, From discovery to delivery: Governance of AI in the pharmaceutical industry, Green Analytical Chemistry 13 (2025) 100268. doi:10.1016/j.greeac.2025.100268.

[56] P. Persson, Managing Socio-Technical Debt: Causes and Design-Science Solutions for Citizen-Centred Digital Public Services, Ph.D. thesis, University of Gothenburg, 2025. ISBN: 978-91-8115-551-8.

[57] P. Herath, M. O. Ahmad, A case study on decoding human factors and socio-technical debt in large scale agile project, in: Proceedings of the 2025 29th International Conference on Evaluation and Assessment in Software Engineering Companion, 2025, pp. 1–9. doi:10.1145/3727967.3756846.

[58] N. Riquet, X. Devroey, B. Vanderose, Debt Stories: Capturing Social and Technical Debt in the Industry, in: Proceedings of the 7th ACM/IEEE International Conference on Technical Debt, 2024, pp. 40–44. doi:10.1145/3644384.3644473.