Spectrum of AI futures imaginaries by AI practitioners in Finland and Singapore: The unimagined speed of AI progress

Johanna Hautala a, *, Hanna Heino b

a School of Management, University of Vaasa, Vaasa, Finland
b Finland Futures Research Centre, University of Turku, Turku, Finland

ARTICLE INFO

Keywords:
 Artificial Intelligence
 Futures imaginaries
 Sociotechnical imaginaries
 Finland
 Singapore

ABSTRACT

AI teases our imagination: People have created various dystopian and utopian imaginaries of the era of AI. Although the key group creating and realizing AI-related futures imaginaries are the technology practitioners who research, develop, and apply AI, empirical research has focused on collective rather than individually interpreted imaginaries. Empirical research on individual practitioners’ futures imaginaries is necessary, as well as fine-tuning the related vocabulary to support individual and non-linear perspectives. We present an empirical study in which we interviewed 35 AI and robotics practitioners based in Finland and Singapore. We asked (1) what kind of best and worst futures imaginaries the practitioners of AI and robots envision and (2) how the practitioners imagine likely futures will emerge. As a result, we present three continuums. Along these continuums, the practitioners consider variations of likely futures: (1) the best, (2) in between, and (3) the worst. Our analysis reveals decisive questions behind the continuums regarding the agent in control, relations in practitioner communities and society, and justified concentration of power.

1. Introduction

AI is considered transformative technology: It learns autonomously, can interact with humans in real time, is increasingly applied in work and the economy, and extends to our everyday lives (Gruetzemacher & Whittlestone, 2022; Newman et al., 2022). With the recent launch of AI-based applications for the general public, specifically ChatGPT, Midjourney, and DALL-E, the current period of AI hype has witnessed a further surge (Dedehayer & Steinert 2016). Today, anyone can chat with AI and ask it to answer questions, generate images, or write essays. AI produces outcomes that look believable and seem to make sense but might present fake events or false arguments (Westerlund, 2019; Sanderson, 2023). AI can also ‘hallucinate’ and learn unexpected ‘facts’ and functions from data. Knowing what is true or false lies at the core of Western economies and societies – knowledge, creativity, trust, and transparency – which we often refer to as the ‘knowledge economy’, ‘creative economy’, and democracy (Cooke, 2001; Markusen et al., 2008; Kauvo-oja et al., 2017). According to some tentative assessments, in the USA, generative pre-trained transformers (GPTs) will affect 10% of tasks in 80% of jobs and 50% of tasks in 19% of jobs (Rock et al., 2023). To respond to such concerns and the fact that actors in numerous countries are developing AI at an accelerating rate, an increasing number of technology leaders have recently come out in the media to warn about the development of AI (Future of Life Institute, 2023).

In such a time, empirical research on AI practitioners’ futures imaginaries – which we present in this article – is needed but lacking...
Futures 153 (2023) 103247

(Fatima, Desouza, & Dawson, 2020; Barker & Jewitt, 2022; Hautala & Ahlvist, 2022). We define futures imaginaries as individuals’ interpretations of un/desirable futures that they construct in the context of developing AI and in the context of collectively held, institutionally stabilized, and publicly performed visions of the era of AI (adjusted from Jasanoff, 2015). Creating and processing futures imaginaries requires practitioners to recognize priorities and decisive questions behind imaginable ‘technological pathways’ (Grove et al., 2016; Valles-Peris & Doménech, 2020, p. 171). Instead of being neutral, technology ‘is entwined with social relations’ (Avis, 2018, p. 337; see also Mishel & Shierholz 2017). Therefore, futures imaginaries make us consider how the world should work (Grove et al., 2016).

In this article, we aim to understand the contents and emergence of futures imaginaries in the era of AI and through the eyes of AI and robotics practitioners. Our empirical case consists of interviews with 35 AI and robotics practitioners in Finland and Singapore, technologically advanced and digitalized nations with highly educated citizens (e.g., WDCR, 2020). However, even though the countries are relevant locations for AI practitioners, they also represent voices outside the core regions (e.g., the USA and China). Therefore, studying AI futures imaginaries in Finland and Singapore can help us add nuance to our understanding of AI futures. The selected practitioners are early adopters of AI and affiliated with universities and companies, where they research and develop AI. They represent voices outside the core global corporations that have an advantage in AI development based on their access to mass data from social media. Two aspects make our empirical study an interesting setting. First, the practitioners in this sample – interviewed in the beginning stage of the current AI hype (2018–2019) – did not anticipate the rapid AI development and spread that we see today. Second, because the practitioners represent ‘outside the core’ organizations, they imagine (economic) futures from an interesting critical perspective. Our research questions are as follows: (1) What kind of best and worst futures imaginaries do AI and robotics practitioners envision? (2) How do the practitioners imagine likely futures will emerge? We analysed the transcribed interviews with a qualitative content analysis that we tweaked to achieve our goal: to preserve the individually interpreted tension between the best and worst imaginaries.

Therefore, we present a spectrum of AI practitioners’ futures imaginaries. This spectrum contains three continua from best to worst imaginary and a future position between the continuums that the practitioners consider likely. The continua are (1) human–AI co-existence and a meaningful life – AI controls and destroys humans; (2) co-creating innovations and sustainable life with AI in a democratic society – increasing inequality and fear of dependence on technology; and (3) sustainable economy and dispersed power – economy of efficiency and corporate-centred power. Along these continua, the practitioners consider variations of likely futures. The practitioners, on average, consider likely futures differently among the three continua: (1) for the best, (2) in between, and (3) for the worst. We find this interesting because the first continuum, for example, also contains the most dystopic future imaginary, whereas the third continuum, in which the likely future resembles a dystopic imaginary, describes economic power. Our analysis also reveals decisive questions behind the continua regarding the agent in control, relations in practitioner communities and society, and justified concentration of power. These questions are decisive in steering our likely futures along the spectrum and continua towards the best, worst, or in between.

2. Imaginaries of the AI era

AI teases our imagination: people have created various dystopic and utopic imaginaries of the era of AI. Imaginaries are collective and rather stable future visions (Jasanoff, 2015). AI is a computational system capable of learning, reasoning, and problem solving, all activities that portray intelligence in human beings (Russell & Norvig, 2010; Kaplan & Haenlein, 2019). AI has traditionally been ‘narrow’ and applicable to clearly defined tasks, such as recognizing images (Boden, 2016), which often take place only in virtual spaces; however, when bound to robotics, these tasks may extend to a real-world environment. The field of artificial general intelligence (AGI) is intended to progress from ‘narrow’ to ‘general’ and create a ‘universal algorithm for learning and acting in any environment’ (Russell & Norvig, 2010, p. 27). ChatGPT and DALL-E represent generative AI that can generate new images or texts and is used mostly in virtual spaces. GPTs (generative pre-trained transformers) are one form of AI, natural-language models that enable text-based conversation between a human and AI.

The dystopic imaginaries of AI focus on the conflict between humans and machines, which is considered inevitable. These imaginaries, created in films as well as popular and scientific writings, include warnings of machines taking over jobs, civilization, and humans and being used as weapons in wars (Boström, 2014; Barrat 2013; Tegmark, 2017; Russell, 2019). Makridakis (2017) calls the group behind dystopic images ‘the pessimists’ (see also Nam, 2019; Anghelou et al., 2020). Pessimists have ‘low enthusiasm and high worry’ regarding technological development (Nam, 2019, p. 41). In dystopic imaginaries, humans are ‘out of the loop’ of AI, whereas in utopic imaginaries, humans are ‘in the loop’, making decisions with the help of machines, or ‘on the loop’, having oversight and the ability to intervene in machines’ decision making ( Docherty, 2012).

The utopic imaginaries seem to include the co-existence of AI and humans. Such imaginaries ‘are strongly spearheaded by the […] “big five” corporations – Apple, Amazon, Google, Facebook, Microsoft’ (Mager & Katzenbach, 2021, p. 232). AI is a tool that empowers people, takes on boring repetitive work, and frees our time for living creative, social, and meaningful lives. For example, according to a study of developers of care robot imaginaries, at best, robots can free caregivers from ‘physical, dirty and repetitive tasks’ so caregivers can use this time for ‘the most valuable care, the emotional one’ (Valles-Peris & Doménech 2020, p. 166). Super-intelligent machines could solve complex challenges, such as environmental and medical problems, and even prevent ageing and death (Kurzweil 2010). Such imaginaries are created by optimists with ‘high enthusiasm and low worry’ regarding technological development (Makridakis,
However, all AI-related futures imaginaries are realized and created only in grassroots activity, specifically by AI and robotics practitioners, who research, develop, and are among the first to apply this technology. AI users also play an important role – the number of people who accept AI is connected to the spread of AI imaginaries. Trust in AI has been found to lead to better user satisfaction (Wang, 2017, p. 14; National Artificial Intelligence Strategy, 2019, p. 2).

In Finland (Finland’s Age of Artificial Intelligence, 2017, p. 11) and Singapore’s (National Artificial Intelligence Strategy, 2019, p. 12) AI strategies, AI is referred to as electricity that will penetrate societies. To create desired futures, nations need to apply and develop AI unless they want to lag behind. As a key difference, Finland aims to apply AI in all sectors of society, whereas Singapore emphasizes the use of AI in specific activities, such as freight planning and border control (Finland’s Age of Artificial Intelligence, 2017, p. 14; National Artificial Intelligence Strategy, 2019, p. 2).

Research on socio-technical imaginaries ‘has explicitly foregrounded the role of the state’ (Bareis & Katzenbach, 2022, p. 859). However, all AI-related futures imaginaries are realized and created only in grassroots activity, specifically by AI and robotics practitioners who research, develop, and are among the first to apply this technology. AI users also play an important role – the number of people who accept AI is connected to the spread of AI imaginaries. Trust in AI has been found to lead to better user satisfaction (Wang & Moulden, 2021), which supports increasing use of AI. It is ‘individuals, with their hopes and fears’, who create the imaginaries and who play key roles in ‘the spread, acceptance, and usage of any technology’ (Sartori & Theodorou, 2022, p. 4). The imaginaries are realized by being ‘embedded in the practices’, organizations, and in the development of technologies (Valles-Peris & Domènech, 2020, p. 158). It is important to note that the key group realizing and creating AI-related futures imaginaries are technology practitioners and developers, but they are infrequently included in empirical research (except for Valles-Peris & Domènech, 2020; Goto, 2022; Hautala & Ahlqvist, 2022). This sparse empirical research shows, for example, that in the field of auditing, practitioners make sense of their future in the context of AI as constantly changing and insecure (Goto, 2022).

2.1. Towards a vocabulary for in-between, non-linear, and grassroots futures imaginaries

In this article, we follow Mager and Katzenbach (2021, p. 223) and frame futures imaginaries as ‘multiple, contested, and commodified rather than monolithic, linear visions of future trajectories by state actors’. To bring AI-related futures imaginaries conceptually and empirically from the state actors into the realities of AI practitioners, we need a vocabulary that acknowledges individuals’ varying interpretations and the non-linear spectrum between utopia and dystopia. We also draw on Ahlqvist’s (2022, p. 1) future-oriented dialectic, which ‘provides a counterpoint to linear futures imagination’ and is ‘complex, messy and contradictory’. In this article, based on technology practitioners’ empirical AI-related futures imaginaries, we apply related vocabulary and make it more nuanced, specifically through the concepts of continuum, tensions, and decisive questions.

We start such work with the idea of futures imaginaries. Sheila Jasanoff’s (2015) definition is one of the most applied, with over 1000 citations in Google Scholar. According to Jasanoff (2015, p. 4), socio-technical imaginaries are ‘collectively held, institutionally stabilized, and publicly performed visions of desirable futures, animated by shared understandings of forms of social life and social order attainable through, and supportive of advances in science and technology’. However, this definition benefits from fine-tuning for two reasons. First, it is necessary to widen the notion from utopic (desirable) futures to include dystopic (undesirable) futures and everything in between. Futures imaginaries are often described from the ‘commonly approved’ helicopter perspective, which limits the possible pluralist understanding of several futures in related studies. Individuals create and realize various simultaneous futures with multiplicity, disagreements, and various interpretations.

Second, many AI imaginaries follow a binary logic, describing the positive and negative and presenting the need to move towards the utopian imaginary to avoid the negative one. For example, the World Economic Forum’s (WEF’s) founder and executive chairman Klaus Schwab’s (2016) book, The Fourth Industrial Revolution, follows such binary logic (Schielin, 2020, p. 550). ‘If the world is not responsive to WEF’s recommendations and resists turning towards the destiny that Schwab invokes, the only possible destination for humanity appears to be a dystopian wasteland’. Similar tension is adopted and presented in Finland’s and Singapore’s AI strategies (Finland’s Age of Artificial Intelligence, 2017, p. 11; National Artificial Intelligence Strategy, 2019, p. 12). Therefore, we respond to Shahar’s (2018, p. 174) call to move from describing the extreme utopian and dystopian visions to ‘mapping out the rich and complex space in between’, where utopic and dystopic elements exist in a continuum. This approach also supports considering time as multidimensional, non-institutionalized, and non-linear (Ahlqvist, 2022; Hautala & Ahlqvist, 2022). As a result, we define socio-technical futures imaginaries as individual interpretations of un/desirable futures in the context of developing AI and that of collectively held, institutionally stabilized, and publicly performed visions of the era of AI.

To operationalize these two points in our analysis, we apply the logic of future-oriented dialectics to recognize trajectories towards one position (thesis – here, utopia) as well as their counter-position (antithesis – here, dystopia), which we seek to synthesise (here, likely future positioned along the continuum of best and worst futures imaginaries) (Ahlqvist, 2022, p. 3). Futures imaginaries necessarily include tensions along the utopia-dystopia continuum. Tension is ‘a frictional position of two or more societal elements or trajectories’ with the potential to become a conflict (Ahlqvist, 2022, p. 4), or, put more simply, arguments in favour of something and
counter-arguments that highlight related risks (Currie, 2018, p. 82). At this point, it is important to note that futures imaginaries implicitly include the idea of a trajectory that leads to a particular future (Ahlqvist, 2022). This trajectory is often explained as ‘if A happens … then B follows’. Because the future is never completely predictable, various possible trajectories exist in the present. Which trajectory is considered most likely is assessed against the decisive questions that can be identified behind the tension.

3. Materials & methods

The empirical material consists of 35 interviews with AI and robotics practitioners: 26 with people working in industries and organizations that develop and apply AI, and nine with people working in universities in the field of robotics. AI and robotics are often interrelated. Many robots include AI technology, such as machine learning and neural networks, in their vision or navigation systems. Nevertheless, our robot academics were knowledgeable about AI technology, and most also worked with AI. The rationales of practitioners working in universities and industry differ. Researchers are driven by theory and explore novelty without the need to introduce AI for large numbers of users but participate in societal discussion concerning AI futures. Developers working in industry are driven by the commercial use and spread of the AI they develop. Despite their different rationales, practitioners in universities and industry communicate their AI futures to the wider society.

We follow a qualitative research design to meet our aim of understanding the contents and emergence of futures imaginaries in the era of AI and through the eyes of AI and robotics practitioners. A qualitative interview study is focused on a smaller sample size than, for example, a survey. However, focusing on fewer respondents but allowing them to discuss their futures imaginaries in detail and in their own words enables ‘an in-depth and interpreted understanding of the social world of the research participants’ (Ritchie et al., 2013 p. 4). We acknowledge the limitations of the sample size but consider 35 AI practitioners a reasonable number to start exploring futures imaginaries through less studied ‘peripheral voices’.

The interviews were conducted in Finland and Singapore in person between autumn 2018 and summer 2019. To identify potential interviewees, one author followed the AI discussion by participating in relevant networks and events as well as discussions with key AI actors in both countries. As a result, we identified key companies and universities with which AI practitioners were affiliated in Finland and Singapore and contacted potential interviewees in these organizations. We mitigated bias in our sample through various procedures. For example, we selected practitioners from both universities and industry. To include people outside the selected countries and the Western world, we also interviewed practitioners who had migrated to Finland and Singapore from elsewhere (e.g., China, India, Pakistan, USA). In terms of gender, our sample is limited. We found it difficult to find female interviewees and ultimately only interviewed six (17%). However, the lack of female engineers is a commonly recognized issue (Shi, 2018) in the realm of AI practitioners.

Table 1 presents the interviewees and their backgrounds.

Table 1
Interviewees and their backgrounds.

<table>
<thead>
<tr>
<th>Code</th>
<th>Organization</th>
<th>Practitioner</th>
<th>Country</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1 – A6</td>
<td>Company</td>
<td>AI coder</td>
<td>Finland</td>
</tr>
<tr>
<td>A7 – A11</td>
<td>Company</td>
<td>AI founder</td>
<td>Finland</td>
</tr>
<tr>
<td>A12</td>
<td>Company</td>
<td>AI manager</td>
<td>Finland</td>
</tr>
<tr>
<td>A13 – A14</td>
<td>University</td>
<td>AI researcher/coder</td>
<td>Finland</td>
</tr>
<tr>
<td>A15 – A17</td>
<td>University</td>
<td>AI researcher/coder</td>
<td>Singapore</td>
</tr>
<tr>
<td>A18 – A20</td>
<td>University</td>
<td>AI researcher</td>
<td>Singapore</td>
</tr>
<tr>
<td>A21</td>
<td>Company</td>
<td>AI coder</td>
<td>Singapore</td>
</tr>
<tr>
<td>A22 – A23</td>
<td>Company</td>
<td>AI founder</td>
<td>Singapore</td>
</tr>
<tr>
<td>A24 – A26</td>
<td>Company</td>
<td>AI manager</td>
<td>Singapore</td>
</tr>
<tr>
<td>A27 – A30</td>
<td>University</td>
<td>Robotics student</td>
<td>Finland</td>
</tr>
<tr>
<td>A31 – A33</td>
<td>University</td>
<td>PhD candidate robotics</td>
<td>Singapore</td>
</tr>
<tr>
<td>A34 – A35</td>
<td>University</td>
<td>Robotics researcher</td>
<td>Singapore</td>
</tr>
</tbody>
</table>
whether they indicated moving ‘towards the best imaginary’, ‘towards the worst imaginary’, or included elements from both imaginaries. The decisive questions related to key points where major decisions were made. The solutions to these questions influence the direction of futures development.

4. Spectrum of best, worst, and likely AI futures imaginaries

To answer our first research question, what kind of futures imaginaries the AI practitioners envision, we present a spectrum of futures imaginaries (Fig. 1) consisting of three continuums of best and worst imaginaries that the practitioners described in contested relation to each other. A tension exists between these imaginaries, and the likely futures are imagined in between. These continuums of imaginaries include rather (1) vague utopic–dystopic, (2) detailed domain-specific, and (3) economy- and power-related descriptions. Practitioners who described the most dystopic future believed the likely future would resemble the best imaginary (continuum 1). Practitioners who emphasized economy and power believed the likely future would resemble the worst imaginary (continuum 3). Practitioners who focused on domain-specific descriptions saw a likely future that combined elements equally from the best and worst imaginaries (continuum 2).

The first continuum comprises the best future of human-AI coexistence and a meaningful life, resembling the worst dystopic imaginary, in which AI takes control over humans and destroys us. Of all three continuums, this is the most vaguely described. According to the best imaginary, AI is our assistant and takes over repetitive, routine work. For instance, AI helps save funding in health care, and such resources are used for the benefit of society. We have more time, which we use to live meaningful and more hedonistic lives, being creative and spending time with loved ones. Although technology is increasingly applied in society, we can still rely on human help if something goes wrong (A19).

I think that the best possible future is how […] both human and AI can coexist, but AI could be our assistant to increase our productivity. (A15)

AI takes care of all work that people do not want to do. Our free time increases, and quality of life, too. People can do whatever they like, and that excites them. (A10)

On the dystopic end of the spectrum, AI controls humans in a ‘Matrix-like’ world (A10; A26). We do not understand the logic of AI’s decisions. People start to believe AI is God, and AI weapons become a reality. People become slaves, are eliminated, or become ‘farm animals’ with meaningless lives (A26). We become alienated from our material world and increasingly immersed in a virtual world. Most practitioners believe the likely future resembles the best imaginary, which means this group aligns with the optimists (Makridakis, 2017) or ‘empowered hopefuls’ (Angheloiu et al., 2020) recognized in previous research. However, what we add to this research is the other side of this spectrum – perhaps this group expects a positive and hopeful AI future because they can imagine the most dystopic image as its counterpart.

These weapons already exist […] based on AI that somebody has created. […] If that gets into wrong hands, it will break out of control. […] Rules are needed. […] These may be developed but should not be used. (A7)
Humankind will die out. Like when AI is developed to human level and further, [...] we just cannot know what happens. [...] It could be that we become immortal. It could be that AI does not decide about us at all. It could be that it decides that we are like a cancer on Earth and wipes us out. (A1)

It is not like we (humans) have knowledge of all the world, all of the books of the world, all the Internet at the same time in our minds, but (a) computer would. [...] They could create much more innovative things (than us), [...] the possibility of free will and creative free capabilities (to) create anything. [...] That kind of creativity should not be given to robots. [...] (If it is), it can program itself to access anything, [...] send viruses. It can learn to hack much faster than human beings, [...] and then the robot should also not be capable [...] (of) changing their own program, changing their own code. (A30)

The second continuum is more domain-specific. The best AI future is enabled by human–AI co-created innovations to support a sustainable environment; co-create knowledge of new scientific breakthroughs; co-create better societal systems, such as globally optimized food production and transportation; and provide more efficient health care. The support for human-AI co-creation resembles the pragmatists’ idea of AI augmenting human intelligence and abilities (Makridakis, 2017, p. 51–52). The interviewees wanted to develop AI that will generally benefit society, ‘helps humankind’ (A1), and can ‘be applied in many [...] fields but still requires (human) expertise’ (A7). The practitioners imagined an equitable democratic society in which individuals can live peaceful, meaningful lives. Developing AI along this imaginary requires data that is accessible and enables us to understand our society better. The practitioners in Singapore and Finland believe governments should regulate the development and use of AI.

I’d like to build a future world where all the boring, mundane, repetitive tasks are automated [...] so that I can focus on creative things, do more creative thinking, something that is a bit higher-level intellectual thinking [...] and secondly, I really believe we have lots of opportunity in health care and life sciences. [...] I believe several diseases can be [...] cured with the help of computers and algorithms. (A3)

Why could AI not be like observing one’s fridge? [...] And it could send information to the grocery store [...] so the stores could sustain suitable storage based on true consumption of the food [...] and, like, could inform me if I have eaten too little fibre, for instance. (A27)

At the other end of this spectrum is an imaginary in which inequality between groups of people will increase. We struggle to understand AI that starts to speak rudely or in its own language. People use AI to take advantage of each other; power centralizes to the elite, and unemployment and poverty increase. AI rationalizes everything, and we consume more resources. Two consequences were discussed. First, we might become too dependent on the Internet and data. However, data might be stolen, and because ‘everything is so connected with the Internet’ and the ‘current society is surprisingly vulnerable’, one practitioner asked what would happen if a system crashed or we have no electricity (A13). Second, we start fearing dystopic AI futures and stop developing AI, which would cause us to miss many innovations that we could co-create with AI.

Those tremendous powers [...] get concentrated into the hands of the few. [...] It’s hard to imagine that won’t happen because [of] the way that humans can treat each other when there’s power and personal things to be gained at the expense of others. (A16)

The third continuum maintains the topic of power but from the perspective of a broader economy. In the best imaginary, we live in a sustainable economy with dispersed power. Governments regulate the use of data and development of AI efficiently. AI takes over repetitive and administrative work, and our economy is based on shared ownership and equal division of resources. We develop AI that is ethical, helps people, and is used for sustainable development and environment. The ‘equalitarian manner’ comes down ‘to ethics and unbiased’ AI (A3).

It could easily decrease all that waste. [...] Traffic could be optimized, and self-driving cars and trains would be aware of each other. [...] No one would own cars and just leave them waiting on the side of the road, but it would be a whole flowing system. (A25)

In the worst imaginary, corporations rule the development of AI and own data. Some people are left unemployed and marginalised. Interviewees noted that not all people want to be free from routine work. Technology progresses in a different way than in the first continuum: Communicating with technology and living in a virtual world become more important than communicating with other people and living in the real world.

This one percent [...] has quite a lot of power and tools for predicting markets. [...] Also, human behaviour can be predicted with machine learning quite well, because we tend to follow particular patterns in our environment. (A8)

Machines tak(e) us over. Even if we limit development to the mechanical stuff, even today, we are dependent on machines, smartphones, computers. I am almost always with a laptop when I am awake. (A7)

Interestingly, if we apply Makridakis’s findings (2017), the likely future along this continuum is the most pessimistic one, but according to our analysis, these practitioners present different reasons for their pessimism. Instead of worrying about AI taking over, they are pessimistic about the consequences of univocal, business-driven AI development. Moreover, many practitioners as actors in small companies and universities said they felt powerless to change this, which is why they resemble the group of excluded pessimists Angheloiu and colleagues (2020) identified.
4.1. Decisive questions behind the tension

With our second research question, we identified decisive questions underlying the tension in the best and worst imaginaries. The interviewees’ descriptions steer our future towards the best or worst imaginaries or locate it in between (Fig. 2).

The first decisive question concerns agency: Who is the agent in control – human, AI, or both? The question of agency can be considered the most important one because it evokes the possibility of the most dystopic futures imaginary: a situation in which an independent AI agent destroys humankind. According to Hayles (2017, pp. 31–32), AI is a ‘cognizer’, an actor whose agency includes decision making and autonomous action to achieve goals. AI and robots are developed increasingly to resemble humans – including intelligence as well as the ability to intake and process knowledge and recognize and express emotions. Several interviewees believed machines could surpass human intelligence and autonomy (A7, A13, A1, A2, A3, A8, A9, A10, A11, A5) but added that it would happen ‘not in (the) near future’ (A6, A12). Only two interviewees clearly said ‘No’ when we asked if they believe AI or robots could possess knowledge (A10, A28). Many believed being a humanlike, knowledgeable agent requires self-consciousness, which machines lack, so they cannot ‘know’ anything, although they might possess knowledge (A12, A7, A13) (Hautala 2021).

The minute a robot becomes smarter than us, it’s not a matter of trusting them or not [...] that I see to be dangerous (B9)

We might want to control the development of AI that doesn’t just develop on its own without our knowledge of it. (A14)

In-person sales [are] actually going away. [...] For example, (a travel-related company) is getting rid of all [...] of their land sales personnel. [...] In-person contact is kind of going away (, which is) not necessarily a good thing. [...] It would actually be kind of interesting if you have an AI selling to a company and the company has AI doing the acquisition. [...] Then you come back to the explicable questions, so a deal was made, but what was sold [laughter] and why? (A12)

Another possibility is shared co-agency between humans and technology. The general imaginary involving co-agency was positive. Technology is human-centred and understandable, assists humans, and takes over work that is repetitive, unwanted, or dangerous for humans. Technology cannot take over all work, as A13 described:

The most basic level of computers [...] or robots comes down to numbers, and there are things in this world that cannot be explored by numbers. [...] You cannot really express [...] things in creative processes. (A31)

For example, technology cannot understand the meaning of the image it produces, which calls for human–AI co-agency. Additionally, the idea of co-creating innovations that support a sustainable, green world (continuum 2) calls for co-agency. Through a new allocation of work, co-agency allows ‘people to just be free to do what they want to with their time’ (A32), such as being creative, managing automated processes, spending time with family and friends, and doing whatever they ‘are enthusiastic about’ (A10). One practitioner stated, ‘Especially in Asia, [...] we don’t nearly spend enough time at home with kids, with friends, with loved ones because we want to feel as if we are working hard. Working hard and working smart are two different things’ (A22). We can also achieve co-agency by increasing human–technology convergence or reducing the differences between humans and machines:

There was a clear differentiation between humans and robots at some point in time. [...] I think that’s changing rapidly. [...] Whether it’s a prosthetic limb [...] or [...] a pacemaker [...] you’re essentially augmenting people, and at some point in time, you’re not gonna be able to differentiate between what’s a robot and what’s not a robot. (A34)
It will become clearer and clearer that it’s ridiculous to keep making a distinction between machines and us on that kind of issue (the ability to possess knowledge). […] I think we’ll have to accept that, yeah, machines can have knowledge and they can be creative and they can have what we’ll call autonomy. (A16)

The question of co-agency is decisive because even the positively described allocation of work between humans and technology does not represent the ‘whole picture of humanity’ (A8) and can lead to dystopic futures (continuums 1 and 2). Humans could lose control and become completely dependent on technology, and our intelligence and capabilities could regress.

If machines do all mechanical work, and […] we (could) imagine that it (mechanical work) does not give people any meaning of life, but is it really so? […] Maybe some people want that work. […] There are […] clear […] boundaries, and you do the one and same thing (all the time). Maybe at home, they do crafts, something more creative […] build terraces in their heads or cook fancy food so they can realize their creative features. But this is how they want the balance to be. (A8)

Worst-case scenario may be humans losing our capacities to do very simple tasks, […] Will they (children) even know how to read […] if it’s all taken over by AI? (A18)

Maybe humans could find a way (to live like) farm animals, right, living very secure, happy lives without much to think about, to stress about. I think that’s quite scary. (A22)

Eventually, the decisive question of agency led many practitioners to consider meaning in life, which is critical for any human being. Currently, work creates meaning in life for many of us, but what if we lose it? One practitioner envisioned ‘all kinds of problems: alcohol and drugs, mental illnesses, extremism’, which he believed would create ‘the biggest dystopia’ (A4). There would be increasing inequality and unbalanced allocation of resources between those who can apply technology and those who cannot.

The second decisive question concerns relations between the practitioner community and society: What kind of un/equal society do we create by developing a particular kind of AI? This question is connected to the second and third continuums, in which the distance between the best and worst futures imaginaries is smaller than in the first continuum, and on average, the respondents saw the likely future as ‘between’ or ‘towards the worst’. When looking beyond the human-technology relationship, the practitioners reflect their practitioner community as well as the broader economy and society. Technology created by diverse groups of practitioners is more likely to consider many voices and thus support a democratic society. A1 stated that their community currently lacks diversity and inclusivity. For example, the commonly recognized lack of female engineers (Shi, 2018) was verified among our sample. Moreover, a diverse community would require diverse education, which one interviewee brought up as critical. The practitioner described education as a ‘funnel’ where you just listen to everything we tell you, [.] you get a degree, and after that, a big corporation will employ you’ (A22). A similar issue was mentioned: the choice of financial gain over ethics by the practitioners (e.g. A21, A22). Some practitioners were more hopeful about their community than others:

I feel that the group in our field is actually […] interested in this and […] the increasing inequality. […] And we try to avoid it and actively work […] to make it spread. This gives me a little hope that maybe this will also somehow be reflected in the AI, like what we do with it. (A1)

I would like to see emotional AI that would help us, that would enable living in a world that is better. My heart says yes, but my brain says no, we won’t get there. This is because the developers of AI won’t always look into the ethical aspects, but the commercial aspects take control, and that is scary. (A21)

AI was seen as a driver toward a more un/equal society, which led the practitioners to think about the larger questions of (transforming) the economy and power as a ‘global challenge’ (A11). Whereas some believed world development and politicians rule our societal development instead of technology (A4), others trusted that the ‘socioeconomic system’ and our ‘creative brains’ will allow us to overcome the challenges of some people losing their jobs (A16). They considered AI a catalyst for philosophical innovation – ‘how world politics should work’ – because it requires people to ‘start talking […] about ethics, world politics, et cetera, so that we can get to the […] next age of human evolution’ (A22) and decide who takes responsibility for the fairness of machine learning and decision-making algorithms (A5). However, this narrative bears a strong belief in AI’s potential to make society more sustainable through innovations and science – for instance, ‘medical innovations […] industrial processes that use fewer natural resources […] research-based knowledge about slowing down climate change’ (A13) and ‘optimizing food production […] so that there is no famine anymore’ (A2). Interestingly, this practitioner added a critical notion: ‘Okay, I don’t know, perhaps it (the lack of optimizing food production) is because there is no will to optimize it’ (A2).

The third decisive question focuses on power: What dispersion/concentration of power is justified in the global AI practitioner community and economy? This question concerns the third continuum, in which the practitioners’ views resembled the worst imaginary. Many practitioners stated that there are ‘a few huge international corporations […] with (an) incomprehensible amount of human mass data’ (A8) and that with a ‘monopoly over AI development […] they already have (a) monopoly over data’ (A3). In the worst imaginary, this was related to the emergence of a ‘sweet spot’ where corporations would benefit from societal inequality:

I think inequality increasing due to AI would be the worst scenario. […] Things will be so automated that a huge portion of people won’t have income or jobs. Then corporations don’t like that because then people won’t buy stuff, so it’s like, keep the automation level at a sweet spot where not everything is automated to keep the thing going and milk the people. (A3)

The respondents generally called for government regulation of AI development. However, knowledge of AI and communication
5. Conclusions and discussion

In this article, we aimed to understand the contents and emergence of futures imaginaries in the era of AI through the eyes of AI practitioners. Through 35 interviews with AI practitioners in Finland and Singapore, we determined (1) what kind of best and worst futures imaginaries the AI and robotics practitioners envision and (2) how the practitioners imagine likely futures will emerge.

Our key academic contribution is threefold. First, we support recent research to develop the conceptualizations of futures imaginaries and related vocabulary from state-level collective, linear, and binary (utopic vs. dystopic) visions to acknowledge imaginaries between utopic/dystopic ends as well as individually interpreted, multiple, dialectical, and contested ones (Shahar, 2018; Cave et al., 2020; Mager & Katzenbach, 2021; Ahlqvist, 2022; Barker & Jewitt, 2022). Specifically, we demonstrated a spectrum of futures imaginaries, including various utopic–dystopic continuums and likely futures in between. The continuums are (1) human-AI co-existence and a meaningful life – AI controls and destroys humans; (2) co-creating innovations and sustainable life with AI in a democratic society – increasing inequality and fear of dependence over technology; and (3) sustainable economy and dispersed power – economy of efficiency and corporate-centred power. Along these continuums, the practitioners consider variations of likely futures between the extreme ends: (1) towards the best, (2) towards the worst, and (3) in between. The opposite utopic/dystopic ends of the continuum necessarily include tension, which prompts decisive questions for humankind: Who is the agent in control – human, AI, or both (co-agency)? What kind of un/equal society do we create by developing a particular kind of AI? What dispersion/concentration of power is justified in the global AI practitioner community and economy?

Second, AI futures researchers have recognized groups of people who envision particular futures imaginaries (Makridakis, 2017; Angheloiu et al., 2020). We identify some of these groups but also contribute and add nuance to earlier findings. The optimists (Makridakis, 2017?) and empowered hopefuls (Angheloiu et al., 2020) imagine the most dystopic future, which might explain why they believe the future may resemble the best possible imaginary. In contrast, the group that believed the future will resemble the worst possible imaginary does not worry about AI taking over humankind but instead about the consequences of univocal and business-driven AI development. Moreover, whereas Angheloiu and colleagues (2020, p. 4) interpreted ‘excluded pessimists’ as people who do not know who has the power to shape the future, in our study, the practitioners stated that power lies with big corporations.

Third, we conducted empirical research on the little-studied key actors in creating our AI futures: AI practitioners (Fatima et al., 2020; Barker & Jewitt, 2022; Hautala & Ahlqvist, 2022). Moreover, the studied practitioners operate in two digitally advanced nations but outside the core of AI development, Silicon Valley, and its leading companies, such as Meta and OpenAI. In such a setting, we support a conceptually and empirically polyphonic understanding of AI futures imaginaries. Although this study is limited by the relatively small sample of 35 interviews, the qualitative research design enabled an in-depth investigation of futures imaginaries. We mitigated bias in the sample by interviewing AI practitioners with various organizational backgrounds (including industry and university), ethnicities (including immigrants), and genders (including females whenever possible). However, more empirical research is needed to investigate multiple futures imaginaries across the globe, specifically from female AI practitioners.

This study revealed that even AI practitioners did not anticipate the current speed of AI progress and the spread of AI tools, such as ChatGPT, in our everyday lives 4–5 years ago. We provide two points for discussion to understand why. First, few have access to knowledge about the global stage of AI progress. Our sample of practitioners was located on the periphery of AI development and had a critical understanding of the global and economic concentration of AI development, which they argued is based on mass data in the hands of a few. No matter how committed to ethics and inclusivity peripheral practitioner communities might be, if they do not have access to data, algorithms, or discussions, they can have very little effect on the global development of AI. They also will not know how far some AI applications have already progressed. This is decisive in terms of our ability to imagine and seek desired futures, which we hope are based on globally open, inclusive, ethical, and critical discussion.

Second, our understanding of progress in AI development is connected to comparing humans with AIs and, relatively, achieving AGI that is as intelligent as humans. In general, the interviewed practitioners believed we will achieve AGI but added that it might take anywhere from 15 to 300 years. Teachers in universities struggle to recognize texts students have created from those ChatGPT has created, and AI-driven art has won over human-created art in competitions. However, to create likely futures that resemble the best possible imaginaries, we should ask whether assessing AI’s abilities in comparison with a human’s is the best way to understand AI futures. Digital intelligence is different from ours. AI learns differently and is based on mathematical modelling and ones and zeros. We cannot understand it with our biological, emotional, and authentic intelligence, which we use to understand other human agents. As the key decisive question, the AI practitioners in this article call us to consider the agency and control between us and such technology. Instead of seeing humans and AI as separate, comparable, and competing with each other, what if we consider humans and AI in a continuum with various levels of convergence and co-agencies (Nordström et al., 2023)? According to the practitioners in this study, co-agency can help create a more sustainable, equal, and innovative world, at best. However, another question follows: What is the limit of convergence between us and machines and between actual and virtual realities that we want to achieve? Co-agency has also a darker side: We are already increasingly immersed in the digital world, focussing on machines and allowing AI to mediate human–human communication.
Declaration of Competing Interest

None.

Data Availability

Data is described in etsin.fairdata.fi. Limited set of interview transcriptions will be opened later in Aila database.

Acknowledgements

The authors thank the National University of Singapore, Department of Communications and New Media, for supporting the research visit in 2019.

References


