Dec 12th, 12:00 AM

Future imperfect: How AI developers imagine the future

Emmanuelle Vaast
McGill University, emmanuelle.vaast@mcgill.ca

Future imperfect: How AI developers imagine the future

Completed Research Paper

Emmanuelle Vaast
Desautels Faculty of Management, McGill University
Montreal, Canada
emmanuelle.vaast@mcgill.ca

Abstract

This study questions how AI developers consider the potential consequences of their work. It proposes an imagined futures perspective to understand how AI developers imagine the futures associated with AI. It examines qualitatively the case of some AI developers and their work and find that they consider the future consequences of the AI they participate in developing as tangential – i.e., loosely connected to what they do - or integral – i.e., closely associated with what they do - to their work. These imaginations of the future are in tension, prompting some AI developers to work at connecting them as they adjust how they view the future and their work. This study reveals how AI development relies upon distinctive imaginations of the future, illuminates how practitioners engage speculatively with the future, and explains the importance for IT development of developers' answers to what their work may do in the future.

Keywords: Artificial Intelligence, imagined futures, qualitative research, algorithms, expertise, technology development.

Introduction

Recent developments in Artificial Intelligence (AI) associated with machine learning and deep learning hold the promise of profound transformations for people, organizations, and societies (Benbya et al. 2021; Faraj et al. 2018). AI has been defined as "*the frontier of computational advancements that references human intelligence in addressing ever more complex decision-making problems*" (Berente et al. 2021, p. 1435). While the positive potentials associated with AI abound, the development of AI has also led to anxieties and debates (Kane et al. 2021; Riemer and Peter 2020). The potential for ongoing transformations associated with AI is high because AI can learn and act (semi-)autonomously, hence changing the ways in which users and technologies interact (Baird and Maruping 2021). What AI does and what consequences it brings changes over time and may be inscrutable. Information Systems (IS) scholars and practitioners have started to study how people may change their work and identity when AI is introduced (Strich et al. 2021; van den Broek et al. 2021) as well as how organizations may adapt to the introduction of AI (Asatiani et al. 2021; Mayer et al. 2020). Information Systems (IS) scholars have also advocated for engaging critically with AI and for making AI more ethical (Marjanovic et al. Forthcoming; Martin 2019; Stahl and Markus 2021).

IS scholars have called for avoiding an "*oppressive future*" with AI, in particular machine learning (Kane et al., 2021). Importantly, the work of AI developers, i.e., the people involved in designing, creating, and refining AI, affects what AI may do in the future. We thus need to understand how AI developers approach their work and think about the possible future consequences of the technology they participate in developing. So far, evidence of the consequences associated with the use of AI has been conflicting (Lebovitz et al. 2021; Marabelli et al. 2021; Rinta-Kahila et al. Forthcoming). To start unpacking the tension between the promise of future positive potentials and the often disappointing or damaging outcomes, however, we need a better understanding of the ways in which AI developers think about the future associated with the technology they participate in developing. It is especially important to do so because some AI developers have started publicly expressing ambivalence towards their work and toward their companies' policies and

practices. They have for instance demonstrated how AI can discriminate based on ethnicity and gender (Buolamwini and Gebru 2018). Some AI developers have been vocal about the risks that AI can generate (Kellogg et al. 2020). Moreover, employees of technology companies have organized walkouts to protest some AI developments on which they and their organizations have worked. These observations reveal that AI developers contemplate what future their work participates in creating.

This study asks: *How do AI developers consider the future consequences of their work?* To answer this question, this study adopts an imagined future perspective (Frye 2012; Mische 2009) that emphasizes that people imagine what the future may hold and, based on this imagination, act in specific ways in the present. It helps understand how AI developers think about future, ambiguous and uncertain, outcomes that may arise from their work. This study then presents the findings from a qualitative study of some AI developers that reveal how they tended to consider the consequences of their work either as tangential or integral to their work. AI developers found ways to bring these considerations together as they adjusted their work and their imaginations of the future.

This study adds to IS scholarship on AI by bringing in a perspective that illuminates how AI developers' imagination of possible futures shapes what they do. It reveals that AI developers' imaginations of what AI could do in the future affects their current development practices. This study also adds to scholarship that has called for "*speculative engagement*" with the future (Hovorka and Peter 2021) by revealing distinct ways in which practitioners speculate about the future and act according to these speculations. This study further adds to scholarship by revealing that what technology developers consider to be part of or outside of the purview of their work can be contested from within. Finally, it explains how different demographics participate in technology developers imagining distinct futures associated with their work.

Theoretical foundations

The work of AI developers as technology development

AI, especially machine-learning based AI, represents a fascinating new class of technology. The rise of AI holds transformation potential for society and organizations. Yet, this transformative potential of AI has lingered on the cusp of being fulfilled (Canals and Heukamp 2020). Also, AI can lead to risks such as biases in facial recognition systems as well as discrimination in hiring and policing. The rise of AI and the growing needs for AI developers' work have thus led to new uncertainties and ethical dilemmas. There have also been indications of awareness among AI developers and organizations involved in the development of AI of the potential for ambivalent consequences associated with it. For instance, professional associations and companies developing AI have adopted codes of conduct regarding what AI developers may and may not do in their work (Washington and Kuo 2020). Some AI conferences have included discussions panels on ethics and consequences of AI. Since 2014, the ACM Conference on Fairness, Accountability, and Transparency (FAccT) has for instance focused on preventing negative consequences associated with AI. Also, in 2020, the Neural Information Processing Systems (NeurIPS) Conference, one of the largest and most prestigious AI conferences, introduced a mandatory discussion of broader impacts of the work for all submissions. Moreover, massive amounts of data are required to operate AI and AI algorithms have become increasingly sophisticated. Today's AI has often relied upon neural networks and deep learning techniques rather than rule-based or logical AI that dominated earlier generations of AI. Training and refining AI has thus become even more challenging and important to AI development (Zhang et al. 2020).

Furthermore, new generations of AI can learn, predict, and act autonomously or semi-autonomously. Hence, users may not fully understand or anticipate the consequences associated with the implementation of AI-based systems. This black-boxing aspect makes it difficult to retrace how AI reaches the conclusions that it does. Yet, the algorithms that make up AI have properties that AI developers create and refine. AI developers' work is instrumental in shaping AI. What AI developers do, how they work, and which values they privilege has an impact on the AI being developed (Crawford et al. 2019; Orr and Davis 2020).

AI developers are the people who develop AI systems and the algorithms that constitute them. They include computer scientists, data scientists, machine learning researchers and engineers, etc. The demand for AI developers has grown steadily over the last few years. AI developers require in-depth training and constant updates as the AI knowledge domain has expanded rapidly.

AI developers' imagined futures

An imagined futures perspective helps understand how AI developers consider the consequences of their work. AI development impacts the future of the people who will be using the eventual AI or who will be affected by it. The work of AI developers may impact others, organizations, and societies. Scholarship has emphasized the importance of considering the future (Mische 2009) and of studying the relation between the future(s) people imagine and what they currently do (Frye 2012). Such a perspective articulates the connection between an imagined future (i.e., it has not happened yet, may or may not happen in the future, and is imagined) and people's current actions (Beckert 2013). An imagined futures perspective emphasizes that people decide to act in specific ways as they consider future possibilities (Dewey 1922).

An imagined futures perspective is relevant in the case of AI developers. AI developers' work aims at creating AI that others will use in the future (Borup et al. 2006). Developing AI involves an imagination of the future associated with the AI in development. Developers imagine who users will be, what uses they will adopt, and which impacts the future technology will have. An imagined futures perspective enables us to consider how developers' imaginations of the future shape their work. It views development work as a way of enacting AI developers' particular perspectives on what the future associated with AI will entail. There is no certainty about what the future will bring, which is why it is imagined.

Imagined futures bring about practices and create a particular future (Mische 2009). Projections of the future guide practices (Borup et al. 2006). By engaging in practices in the present, AI developers work towards bringing about their visions of the future. Imagined futures thus create and constrain other future possibilities. Moreover, developers may imagine positive and / or negative consequences for the technology in development (Borup et al. 2006). Imagined futures associated with computerized technology, for instance, have long relied upon the expectation that the greater availability of information would necessarily create a better world. Such imagined and often idealized future consequences have in turn affected development practices. The imagined futures associated with AI thus lead AI developers to engage in certain practices rather than in other ones. Furthermore, imagined futures are not fixed once and for all. People develop new ideas about the future on the basis of their actions and their consequences (Mische 2009). Imagined futures thus change. They are dynamic, rather than static, projections. Imagined futures help understand what AI developers do at a certain time but are subject to change. AI developers can re-imagine the future as AI gets created and different AI developers participate in its development.

Imagined futures are also closely associated with people's identity (Frye, 2012, p. 1572). Which futures people imagine for themselves and their work contributes to their sense of self (Vaast and Pinsonneault 2021). Developers cultivate a specific model of expertise that guides what they know, how they work, who they consider themselves to be, and what boundaries they establish for their work. An imagined futures perspective adds a consideration of what AI developers' work may do to this understanding. This consideration is important for AI developers because, on the one hand, the AI they participate in developing will contribute to precipitate the future, and, on the other hand, they can only imagine what the AI they develop will do once it is deployed and used by diverse stakeholders. Answering questions associated with what their work may do can also lead AI developers to demarcate differently the boundaries of their work (Gieryn 1983). Depending upon AI developers' answers to the question of what their work may do, they may pass tasks to others.

Moreover, given the unknowns associated with the future, there may be different imagined futures within a community of AI developers (Jasanoff and Kim 2009). Different AI developers may view AI and what may come from it differently. AI developers can have distinct ways of thinking about the future associated with their work and the new AI in development. This is especially relevant given that AI is developed by a community with AI developers who have different positions, interests, resources, and constraints. AI developers can thus speculate about the future differently. However, scholarship on early AI developers has so far not examined differences in imagined futures. It has instead insisted upon the homogeneity of the epistemic culture and model of expertise.

The distinction between near and distant futures (Augustine et al. (2019) may differentiate AI developers. The distant future refers to a representation of imagined possibilities associated with ambiguity. The near future can be defined in more practical terms and be connected, with some uncertainty, to the present. While much existing scholarship has focused on the near future, emerging research has also started to focus on the distant future, defined in terms of discontinuity with the past and the present and of radical

uncertainty and ambiguity (Augustine et al. 2019). The distinction between near and distant futures clarifies AI developers' perspectives and their work because new AI can be associated with both types of future. AI developers may imagine the future associated with the AI they work on as near and / or distant. A focus on what AI may do in the distant or near future can bring distinct emphases to AI developers' work.

Methods

Research design and data collection

This study relies upon a qualitative design to understand how AI developers considered the consequences of their work. A qualitative design suited the purpose of this study because it enabled to dig deeply onto how some AI developers worked and what perspectives they held regarding AI and the future. This design was also conducive of theoretical elaboration by bringing inspiration for new ideas.

In a typical manner for qualitative research (Denzin and Lincoln 2000; Miles and Huberman 1994), I engaged in the collection of diverse data to examine the research question. For one, I worked with a research assistant to conduct 63 semi-directed interviews. The interviews took place in 2020. Among these interviews, 55 were with AI developers, i.e., with practitioners actively involved in the development of AI, in particular machine learning. These AI developers' job titles included, e.g., computer scientists, professors, research analysts. Vice-presidents of AI development, and PhD students. The other eight interviews were with actors who were part of the AI ecosystem without being actively involved in developing AI (e.g., business development lead in a research institute, lawyer specialized in AI). Semi-directed interviews revealed important ways in which AI developers saw their work and its consequences. We reached AI developers by at first going through [anonymized] University AI resources and Faculty and then expanding contacts via snowballing mechanisms. We engaged in theoretical sampling to get a variety of perspectives among AI developers. We strove to get diversity in terms of stage in their career (from finishing training to very established), gender, country of work, and ethnicity. Table 1 summarizes demographic information about the interviewees.

Gender: • Men: 43 (68.3%) • Women: 19 (30.1%) • Non-binary: 1 (1.6%) **Employing organization**: • Tech company: 20 (31.7%) • University: 10 (15.9%) • Research institute: 10 (15.9%) • Research institute & Univ: 8 (12.7%) • University & tech company: 3 (4.8%) • Finance company: 3 (4.8%) • Non-for-profit org: 3 (4.8%) • Government: 2 (3.2%) • Other: 4 (6.4%)	**Country of work**: • Canada: 25 (39.7%) • US: 25 (39.7%) • France: 3 (4.8%) • Chile: 2 (3.2%) • Switzerland: 2 (3.2%) • China: 1 (1.6%) • Ecuador: 1 (1.6%) • Germany: 1 (1.6%) • Netherlands: 1(1.6%) • South Korea: 1 (1.6%) • UK: 1 (1.6%) **Ethnicity**: • White: 36 (57.1%) • Asian: 16 (25.4%) • Hispanic: 5 (7.9%) • African-American: 3 (4.8%) • Middle-Eastern: 3 (4.8%)
Table 1: Demographic information about interviewees	

During the interviews, respondents discussed their background, career path, and current work. Respondents talked about the challenges and rewards they associated with their work of developing AI.

They discussed the ways in which others (e.g., clients, the general public) perceived AI and whether they considered these perceptions accurate. Interviewees also reflected upon the current situation of their domain as fast-growing and characterized by statistical imbalances in terms of, e.g., gender or ethnicity (Crawford 2016). Interviews happened via phone or teleconference. We recorded, transcribed, and anonymized the interviews.

I also collected archival data from publicly available social media posts and discussions (from Twitter, reddit, Facebook, blogs, and GitHub primarily) among AI developers. The rationale for collecting these archival social media data was that they provided an important, and distinct, perspective on AI developers and on the ways in which they talked about their work and its consequences. Archival data from online sources can be particularly revealing of delicate matters as these discussions occur organically among people. Criteria for the selection of social media data included posts and discussions associated with new AI developments or happenings within the AI community. Also, at times, AI developers discussed issues about the potential associated with their work specifically and / or new AI developments. I collected 125 social media discussions, with more than 2,500 associated posts.

In addition to the interviews and social media archives, I engaged in two years of participant observation in the Observatory on the societal impacts of AI and digital technology, henceforth Observatory on AI (end of 2018-end of 2020). It is an organization that brings together Academics (computer science and social sciences), members of Industry (large technology companies and start-up companies) and non-for-profit organizations dedicated to overseeing developments associated with AI. I had access to email communications, newsletters, reports, and other documentation. I also had access to events organized fully or partly by the Observatory on AI. This participant observation helped gain an understanding of how AI developers thought about the potential of their work and interacted with different actors on the challenges and opportunities of AI development.

I collected new data until theoretical saturation was reached, that is, until the addition of new data did not bring new insights challenging the analyses and theoretical elaboration in progress. Taken together, the different types of data provided complementary perspectives to understand how AI developers considered the future consequences associated with the AI they developed.

Data analysis

The data analysis process followed guidelines for qualitative grounded theorizing and relied upon different stages of coding of the data. The data analysis process also allowed room for changes as new data were added and new insights emerged (Charmaz 2006). Analyses started as the data collection unfolded. They proceeded in stages, with interrelated steps of qualitative coding, memo writing, and gradual theory elaboration (Charmaz 2006). The qualitative data analysis software Atlas.ti helped manage the data, coding, and analysis process.

First, as the interviews and social media data collection were engaged, I went through AI documentation and reviewed the work of the Observatory on AI to become better versed in AI techniques and vocabulary as well as in the AI community. This was helpful to converse with the AI developers in interviews and to make better sense of their social media discussions. Then, I undertook an open coding of the interviews that involved identifying key themes associated with the background, work, and consequences of one's work among AI developers. Based on these early codes, I wrote initial memos that summarized different perspectives that AI developers in the community adopted about their work. At this stage, I was struck by the fact that AI developers struggled with how to deal with the consequences of their work because these consequences were not easy for them to detect but could be positive and negative. The early analyses of the interviews suggested that AI developers considered in two different manners the consequences of their work. Then, I turned to an open coding of the social media data to examine further these initial insights. The open coding of the interviews had brought light to the social media data. Moreover, the social media discussions among AI developers illuminated that the question of the future consequences of the work of developing AI was closely associated with the definition of expertise for some AI developers and

disconnected from it for other AI developers. This insight led to the analytical elaboration of two models of expertise in the AI community, i.e., ideal-typical ways in which AI developers viewed their work and considered its consequences. These two models constituted key categories for this study (Grodal et al. 2021). Having labelled these two models tangential and integral, I went back to the entire data set to see how the two models contrasted and manifested in work in the AI community. This led to specify differences associated with existing concepts in scholarship (near / distant future, identity as expert, and boundary work). I sought to understand what made AI developers lean toward these models.

Then, I searched, in the entire dataset, for ways in which the two models of expertise were connected to each other. I discovered how the AI community resolved some of the tensions between the two models and how this affected AI developers' work. This led me to unpack two major ways in which AI developers connected the two models: they brought some elements of the integral model onto the tangential one and they blended the two models as they adjusted their imaginations of the future associated with their work.

Imagined Futures in AI Development

This section first presents the two models of expertise that were ideal types encompassing contrasting ways in which AI developers imagined the future and thought about the consequences of their work. These two models represented stylized perspectives on the work of AI development and imagined futures towards which AI developers were inclined. This section then details how AI developers connected the two models of expertise by bringing some elements of the integral model onto the tangential one and by blending the models as they adjusted their imaginations of the future.

Tangential and integral models of expertise

The tangential model of expertise considered that the future consequences of AI developers' work were peripheral to the work of AI development. It emphasized that AI developers were scientists creating new possibilities for others to implement and with which to deal. The integral model of expertise highlighted that the future consequences of AI were connected to the work of developing AI. The AI developers who leaned towards the integral model acknowledged the importance of recognizing not only the potential for positive innovations but also for negative consequences associated with the work they accomplished. This section contrasts these models of expertise in terms of their imagined futures, the identity of experts, and boundary work.

The imagined future of the tangential model was "*distant*" (Augustine et al. 2019) in that it was expressed in abstract terms, with a disconnect between the present (i.e., the current situation and the work practices of AI developers) and the future (i.e., what new AI developments would enable in an indeterminate future). For instance, an event organized by the Observatory on AI was dedicated to deep learning. A keynote speaker, a pioneer of deep learning, explained that:

"With all the improvements we have seen in recent years, some people think we are done and just need to scale up what we have already learnt to wider topics and problems, however, I think there are many pieces in the puzzle which are missing."

Such a quote exemplifies how, in a tangential model, AI was making promising advances but that it was difficult to foresee exactly what the distant future would be as well as how it would be reached as seen in the reference to "*many pieces in the puzzle which are missing*." The tangential model of expertise emphasized that AI was progressing in leaps and bounds rather than linearly.

By contrast, the integral model of expertise focused on the "*near*" future (Augustine et al. 2019), that is, on a more concrete future connected to the present. The imagined future of the integral model was practical and associated with existing practices and contexts. In this model, the work of AI developers was to enhance what already existed and to minimize existing limitations, rather than to create unknown possibilities. The model emphasized the concreteness and direct connection between the present work in AI and the near future. For instance, an interviewee was a Faculty at a large University and worked on deep learning and neural networks. He explained the continuity between current AI developments and past technological progress:

"It's important to recognize that at this particular stage of AI, the way that [it] works right now, it's not actually that fundamentally different from any new technology in that what it allows us to do is just

accomplish certain tasks more rapidly at higher volume with, you know, less human effort involved, much as almost any other technology we've invented over time." (Ryan, interview)

This interviewee viewed developing AI as akin to developing prior technologies and as mostly building and improving upon existing foundations.

Identity as AI developers

How AI developers defined who they considered themselves to be differed by model of expertise. In the tangential model, AI developers articulated their identity as that of scientists who contributed to scientific progress and followed rigorous methods whereby they first developed assumptions and then tested and refined them to generate new AI. The scientist identity underlined that, in this model, AI developers considered that they engaged in science to create a better future for many. An interviewee who was a research engineer and project manager in a technology company illustrates this:

"People need to realize we do science, because we need / we want to solve problems and with that, the future will be something that you will be enjoying as well." (Alejandro, interview)

This interviewee considered that AI developers engaged in science to generate a distant future that would benefit society. By contrast, the integral model of expertise articulated an identity for AI developers whereby they were scientists *and* citizens engaged in society. This model presented AI developers as in part activists whose work strove to identify and limit causes of discriminations in the algorithms they and others were building. The mix of expertise and activism was visible in this social media post by the collective "*@Data4BlackLives,*" self-presented as "*a movement of scientists and activists. Data as protest. Data as accountability. Data as collective action*:"

"What is our duty as Black researchers, academics, and community organizers to collect data and tell our stories? Algorithms are not new; they are manifestations of old systems that devalue and dehumanize our people." (@Data4BlackLives, Tweet, 01/2019).

The integral model of expertise also closely connected expertise to the experience, demographics, and social identities of experts. AI developers were scientists engaged in society. In this model, AI developers gained their expertise in part from their social identities as well. For instance, an interviewee worked on developing AI at a large university whose focus was on technology engineering and design. She considered that her gender, and the fact that she was in a social minority within her department, made her more aware than most of her peers of potential negative future consequences of AI.

Boundary work

The two models of expertise were also distinct in their boundary work. They differed in terms of how AI developers demarcated the boundaries between their domain of work and others (Gieryn 1983).

The tangential model emphasized the positive potentials associated with scientific progress in AI. Some AI developers articulated a conviction that they worked towards creating new systems that would yield positive innovations and change. They foresaw benefits for people and societies associated with the development of new working algorithms. Because of this positive potential, AI developers who leaned towards the tangential model cautioned against reining in the development of AI. They considered that limiting progress in AI would ultimately stifle innovations that could bring many positive outcomes to society. In the tangential model, AI developers were responsible for developing new working algorithms. However, AI developers were not responsible for the potentially negative deployment of the algorithms they created. To deal with the negative potential associated with bias in data, for instance, these AI developers often asked other practitioners to engage in activities such as data cleaning and data labelling. These activities were to be done by data engineers rather than by AI developers. This amounted to delegate to others the tasks of dealing with the potential for negative consequences associated with AI development. An interviewee who worked in neural networks and machine learning also illustrates this:

"Some people have a blind trust in these systems… And we are responsible as experts for moderating the distrust and for explaining what the caveats are. But I say: "We as experts," I think in discussion to the general public at that point, maybe the technical experts should have a back seat and the front seat should be more the ethicists and those people with a better understanding of the actual implications of the use of AI." (Tony, interview)

This quote shows how, to limit negative potentials of AI, the tangential model of expertise emphasized that others who were not AI developers but who understood the implications of AI should engage with actors interested in using AI. AI developers from the tangential model appealed to others to explain to the larger public how to use, and not to use, new algorithms.

The integral model of expertise included within the domain of work the importance of recognizing not only the potential for positive innovations but also for negative consequences associated with the work they accomplished and the algorithms they participated in developing. An interviewee who was a senior research scientist in AI at a University noted:

"I do believe that we are held responsible for what we create… We should all be held accountable for the work that we do, and the decisions that we made. And we should have the responsibility to think through those implications. And what damage we could be doing, or what good we could be doing." (Sandra, interview)

This interviewee further explained how, to take in consideration the risks of AI systems early, she relied upon checklists that made complex AI systems less unpredictable. The integral model also emphasized that it was within the purview of AI developers' work to engage via communication with the general public and organizations adopting AI systems.

What made AI developers lean towards these models

Whether AI developers worked on fundamental and / or applied AI contributed in part to which model of expertise they were closer. AI developers who worked mostly on fundamental AI tended to lean towards the tangential model of expertise. AI developers who worked mostly on applied AI tended to lean towards the integral model of expertise. Fundamental AI corresponded to the branch of AI dedicated to making high-level theoretical progress in AI. Fundamental AI was not directly focused towards generating concrete applications that could be tested and used by others. Fundamental AI consisted in generating abstract and high-level possibilities that may one day be applied in specific projects. Fundamental AI was associated with the distant future. For instance, in a 2019 media interview on new directions for deep learning associated with meta-learning and adaptability, a leading authority on deep learning and Turing Award winner, Yoshua Bengio, discussed pioneering advances in AI, but noted that they remained for now more theoretical than applied:

"This is all very basic research using toy problems. That's fine, that's where we're at. We can debug these ideas, move on to new hypotheses. This is not ready for industry tomorrow."

Applied AI corresponds to the branch of AI focused on bringing new applications and services to execute actual tasks such as translation services, image recognition, or recommender systems. AI developers who worked mostly in applied AI tended to lean towards the integral model because their time horizon was near rather than distant. They worked on projects that were already underway and would have delimited outcomes for users. For instance, in a roundtable organized by the Observatory on AI on the opportunities and challenges of AI for agriculture, a participant had a Ph.D. in AI and precision medicine and a postdoc in computer vision and systems biology. In his presentation, he detailed AI developments in the context of the dairy industry and highlighted the potential of AI for predicting milk production, cow resilience, or CO_2 emissions.

Within the AI development community, developers also occupied positions and fulfilled roles that influenced the model toward which they leaned. Some AI developers held positions that made them engage with the distant future by thinking about how to push beyond the current limitations of AI without taking in consideration actual consequences. For instance, a Ph.D. candidate in AI considered it part of his current role to generate breakthroughs but not to worry about what may come:

"My main perspective as a graduate student, I feel like I have practically no accountability for the algorithms that I choose to come up with. And that's kind of the role of a graduate student. I'm supposed to come up with weird, bad, potentially dangerous ideas, because that's how we push the state of the art." (Timothy, interview)

This interviewee further explained how he had created a pilot system that, based on keywords, automatically wrote scientific abstracts. He and his supervisor had been surprised by how well the system had worked but had then moved on to other projects. Other AI developers' roles led them to take in

consideration and act upon potential implications of new AI developments. AI developers who worked in Human Computer Interaction (HCI) and robotics, for instance, had to contend with direct consequences of their work. They were concerned with making sure users could understand new AI applications. For instance, an interviewee was a roboticist whose work focused on interactions between machines and humans. She discussed the challenges of autonomous vehicle projects. She explained that they required experts to be technically proficient, to know the context in which the vehicles would be driving, and to make difficult choices regarding what the vehicle should and should not do when encountering obstacles.

Beyond the distinction between fundamental and applied AI as well as the different roles and positions that they could occupy, AI developers frequently changed positions and organizations in the AI community. This came from the high demand for AI developers and the availability of multiple positions in different industries and types of organizations (e.g., Universities, large companies, start-up companies, non-profit and governmental organizations). Some AI developers went back and forth between more fundamental and more applied AI in different positions. For instance, an interview respondent explained how, after finishing his training, he had worked for an advertising and media company on applied AI projects. After some time, he joined a large technology company, became affiliated with a University, and started working in fundamental AI exclusively. Moreover, some AI developers combined diverse aspects (of fundamental and applied AI and of different types of roles) in their work. Some AI developers whose main work was close to fundamental AI participated in projects that were more applied. Other AI developers who were more on the applied side of AI also worked towards making fundamental progress in their field.

The remainder of the section unpacks how AI developers did so by integrating the tangential model of expertise and by blending the two models.

Integrating the tangential model of expertise

Because of the prestige of the potential for scientific advances associated with AI, some AI developers leaned more towards the tangential model but started integrating some aspects of the integral model onto their expertise and work. They accomplished this by expanding the definition of expertise and by expanding the boundaries of their work.

Some AI developers expanded the definition of expertise in AI development by considering not only scientific and technical skills but also competences arising from social identities, demographics, and experiences. This amounted to incorporating a tenet of the integral model onto the tangential model. A controversy and the ways in which some AI developers responded to it illustrates this expansion of the definition of AI expertise. In June 2020, a team of AI developers released the PULSE algorithm, a computer vision model developed by Duke University researchers that created deepfake pictures from pixelated ones. AI developers soon recognized that the algorithm yielded biased results. The high-resolution pictures it produced were overwhelmingly Caucasian, regardless of the ethnicity of the original picture. AI developers discussed what made the algorithm so obviously biased in its results. Yann LeCun, a prominent figure in AI, Turing Award winner, Professor at NYU, and Chief AI scientist at Facebook, reflected the tangential model of expertise by asserting the prominence of the training data as source of bias:

"ML [Machine Learning] systems are biased when data is biased. This face upsampling system makes everyone look white because the network was pretrained on FlickFaceHQ, which mainly contains white people pics." (Yann LeCun, Twitter post, 06/2020)

This comment led AI developers to debate whether bias could only come from the data. Timnit Gebru, a well-known researcher on algorithmic bias and data mining and co-lead of the Ethical AI team at Google at the time, however criticized LeCun's explanation of the bias:

"Yann, I suggest you watch me and Emily's [Denton] tutorial or a number of scholars who are experts in this are. You can't just reduce harms to dataset bias. For once listen to us people from marginalized communities and what we tell you. If not now during worldwide protests not sure when." (Timnit Gebru, Twitter post, 06/2020).

Gebru explained that her expertise stemmed not only from her formal training, which included a Ph.D. in Electrical Engineering from Stanford University, but also from her being an Ethiopian-American woman working in a domain where these demographics were underrepresented. Her perspective exemplified the integral model. Following the exchange, many AI developers discussed what expertise in AI development entailed and where it came from. Some AI developers who leaned toward the tangential model

acknowledged that they needed to expand their definition of expertise. This was the case for Thomas Dietterich, Emeritus Professor of Intelligent Systems at Oregon State University, who explained:

"Both anticipating and detecting harms requires a diverse team that brings a variety of conceptual models and experience to the table. Engaging such a team from the start is essential to anticipating harms in advance." (Thomas Dietterich tweet, 06/2020).

With this comment, Dietterich expanded the definition of expertise to include "*a variety of conceptual models and experience*." He explained that this was important for AI developers to be able to foresee future negative consequences of new AI developments.

Moreover, some AI developers reflected upon who they were as scientists, as in the tangential model, but also questioned how their gender and / or ethnicity participated in shaping their perspective, as in the integral model. For instance, an interviewee was a researcher in AI focusing on usability and security. He explained that he was proud of the scientific progress that his work in facial recognition was making (as consistent with the tangential model) but that he was uncomfortable with the fact that his work could contribute to discriminatory systems (as consistent with the integral model). He was African-American and aware that facial recognition systems tended to yield adverse results for him and many others. He commented:

"I don't like facial recognition algorithms because [they] don't recognize me well, but you find them in most airports these days… Someone at the end was developing this algorithm so that it is accurate and maybe simplifies the process of check-in. But I don't think they accounted for all the different people that might come through that check-in line. (…) We're developing [these algorithms] for a wide society of people, which includes a wide range of people. That includes both men and women and everybody else. So if we are not inclusive of everybody, then [we] fail at the start, right? [We]'re not achieving the objective." (Kofi, interview)

Such a quote illustrates how some AI developers reflected upon how being part of a social minority within the AI community made them recognize risks with AI developments that others were less likely to be aware of right away. They brought together their social identities and their identity as scientists. They considered that demographics were part of their expertise by making them more or less close to future negative consequences of AI.

Some AI developers who leaned toward the tangential model expanded the boundaries of their workd. They selectively expanded the boundaries of their work either by reflecting upon how to deal with data bias or by engaging with others to discuss potential future consequences of AI development. Doing so, they incorporated some of the characteristics of the integral model onto the tangential one. For instance, an interviewee was Head of a Machine Intelligence research group at a large technology company. In his work, he oversaw AI-based R&D for the company. His outlook was associated with distant future possibilities of AI for new technological developments. He explained that, at first, he and his team had not been involved in the selection of datasets to train new AI models. This was consistent with the tangential model. However, given current developments in AI at his company and in other contexts, he had come to consider that ensuring proper datasets was important to his work. He noted:

"Be bold, but also be responsible. Right? (…) How do we make sure that a responsible decision is made based on data? Because, somehow, we believed that [with] enough data, our algorithms can reflect the real world. And this is obviously not true. There is no unbiased data." (Reinhold, interview)

This interviewee explained how he and his team had recently developed and implemented a system for "*Responsible AI*" that included selecting, cleaning, testing, and refining data for training models. This AI developer had started including in his and his team's work what used to be done by others to limit negative future potential consequences of AI, which was consistent with the integral model.

Blending the two models

AI developers' changing imaginations of the future cultivated a blending of the two ideal-typical models. Blending the two models involved re-imagining the future, bridging near and distant futures, and charting impossible futures.

Re-imagining the future

The imagined futures associated with AI were not fixed. They changed when AI developers engaged in (re-)thinking about the future that could emerge as AI developed and became more sophisticated. Re-imagining the future was still about speculating on what the future technology would do, but it involved doing so while taking into account what past AI developments had accomplished and how the AI community and the broader societal context was changing. AI developers re-imagined the future in an ongoing manner. An interviewee, a research scientist working at a research institute and a university, commented upon how little AI developers who worked on prior generations of AI (here, recommender systems) used to imagine negative consequences of their work and how much things had changed:

"Around maybe 2010, there was a lot of excitement in machine learning around the so-called recommender systems. (…) And now it's pretty clear that actually all of our work has like really strong implications on the way that people think and even in interact with other people." (Jacques, interview)

This interviewee remarked that changes in AI and happenings in society had made it necessary for current AI developers to think about what positive and negative implications new AI developments could bring. To him, re-imagining the future had become essential for AI developers. Another interviewee who was a professor of computer vision and deep learning explained:

"Predicting the future… is really super difficult, right? Some people in the field are [of] the opinion that we're like five years away and some people in the field are saying we're like 100 years away (…). But what we don't see is: what are the problems after that? And this notion of AI is always this receding horizon." (Adam, interview)

This interviewee commented upon the importance and the challenge of re-imagining the future in an ongoing manner because working in AI involved working with a "*receding horizon*." Moreover, another interviewee referred to the NeurIPS conference organizers' decision, in 2020, to require a "*statement on the potential broader impacts of their* [submitters'] *work*" (call for papers, NeurIPS 2020) for all submissions, i.e., a discussion of the potential future consequences of the work. He observed that this decision reflected the growing recognition among AI developers of the need to think and rethink the possible future consequences of AI development:

"The outside world should not underestimate the amount of conversation that happens within the field, whether it [a new AI development]'s bad, which topics to tackle, how to tackle them, what are the implications? (…) This is a moving thing then…" (Tony, interview)

This comment and the new requirement from this conference illustrated how, given ongoing progress in AI, the AI community was re-imagining the future and its own work.

Bridging near and distant futures

Some AI developers worked on bridging the gap between the promises and risks of the distant future related to AI and the near future associated with its current developments. They did so by keeping a perspective of the potential for new progress associated with AI in the distant future while working on realistic improvements for the near future. It was challenging to do so because AI was difficult to comprehend, even for its experts. For instance, in an event from the Observatory on AI, a presentation on recommender systems made the case of AI as a "*fantastic opportunity for good*." It acknowledged that existing recommender systems had current and near future negative consequences but could be changed to make them opportunities for good in the distant future and proposed ways to bridge these horizons.

Bridging near and distant futures was a community effort because it was challenging for AI developers to envision the eventual consequences of the AI they developed. As an example of how AI developers bridged the near and distant futures, during an event organized in part by the Observatory on AI, a computer science professor who was also affiliated with a large technology company presented her perspective on advances in reinforcement learning, an area of machine learning associated with how intelligent agents should make decisions to maximize cumulative rewards. She noted that the (distant future) potential of reinforcement learning was high but far from accomplished and that there was no way to know when it would be achieved. She explained how current efforts from AI developers in industry and Academia worked on incremental improvements (near future) in reinforcement learning as constituting what she called "*Lego bricks*" that would bring the distant potential into realistic realms.

Bridging near and distant futures also entailed that AI developers consider the connections among past, present, and future of AI. AI developers reflected upon their current practices and what they could mean for the near and distant future. A presentation during an event organized in part by the Observatory on AI illustrates this point. The presentation gave insights on few-shot learning, a new technique that aimed at developing machine learning from only a few examples ("few-shot"). This presentation by a researcher at a large technology company focused on the "*Past, present, and future of few-shot learning*." The researcher situated few-shot learning within the history of AI and explained how past and current practices shaped its near and distant future consequences.

Bridging the near and the distant futures was ongoing and oriented the work of AI developers. For instance, an interviewee specialized in Natural Language Processing to develop machine translation algorithms in a start-up company. She considered that the greatest rewards of her work consisted in improving the services that people were already using. In this way, she originally had a perspective of the near future as enhancing existing services in machine translation and of the distant future as enriching human communication. She however noted that she and the other AI developers in her company had had to shift their perspective on the near and distant future services that her company should offer, because users were not using current offerings in the way they had anticipated. She explained that, at first, AI developers at her company were:

"Imagining the user personas [that would be] traveling and studying. Those were the two big personas. But, after we released the product, in reality, actually people use the product to… make friends all over the world, particularly, to make romantic friends. So that was some usage that we didn't expect." (Dasom, interview)

AI developers at her company had adjusted their work according to changes in imagined users and types of uses. They had at first imagined who the users of the service would be and what they would do. Then, based upon actual users, they shifted their expectations of the distant future (from enabling communication to creating new potentials for human connection) and their work as well. They worked towards new future offerings that suited better their current users. Adjustments of what future AI developments should cover were ongoing.

Charting impossible futures

Charting impossible futures involved AI developers reflecting upon what the AI they developed may *not* do. AI developers delimited what they considered to be possible futures associated with AI (i.e., what they wished to continue working on) and what they deemed impossible futures (i.e., potential AI developments that should not be pursued). For instance, many AI developers considered that AI should not be developed or used to predict social outcomes. They noted that AI should not be aimed at estimating the probability of engaging in criminal behavior:

"I received an email for a petition against a paper that had been submitted and accepted to a conference [where] the authors are claiming that they could detect criminals by only looking at their face. (…) How is that cool again? You know, let's say 100 years ago, there were people [who] would take measurements around your head and were speculating that there might be correlation with certain traits, like, is very much related to eugenics. So it's basically the modern version of that." (Jacques, interview)

This comment illustrated how AI developers, as a community, worked to monitor and tamper specific AI developments. In this case, they did so through a petition against a paper originally accepted at a conference. Also, by making a historical connection between current happenings and misguided attempts at eugenics from a century ago, this interviewee reflected on the parallels between current futures (associated with AI for criminality predictions) and past futures (associated with phrenology). Past futures helped AI developers delineate possible and impossible futures in the present state of AI development.

As they charted impossible futures, AI developers questioned the limits of what they should be doing in their work. As an example of such questions, during an event on deep learning organized with the Observatory on AI, a presenter who worked in AI for a technology and consulting company documented how he and his colleagues discussed whether to develop new algorithms:

"What do we build, should we build it how can we build it and is it safe when creating it using AI? Some things don't need AI."

As seen in this quote, charting impossible futures involved questioning the scope and ambitions of future AI developments. Charting impossible futures was distinct from boundary work. It did not imply questions

about which tasks AI developers or others should do. Instead, it consisted in AI developers asking and answering questions regarding what AI in development may not do because future consequences would be harmful. Charting impossible futures delimited what no one should be doing regarding AI development.

Some AI developers thus engaged in thinking and rethinking about what the new technology may not do. Questions regarding what to do and not to do as well as whether to pursue specific projects revealed how some AI developers outlined futures they did not wish to contribute to with their work. They reflected upon how potential negative future consequences of new developments could offset potential positive future ones. Such reflections led to concrete initiatives. For instance, in November 2017, more than 650 AI researchers wrote to the Prime Minister of Canada, ahead of this country's rotating presidency of the G7, to call for an international ban on the development of autonomous weapons. One of the signatories was Geoffrey Hinton, a Turing Award winner who was affiliated with Google and the University of Toronto. Hinton explained his support for the pledge by referring to the need for AI to foster positive rather than negative consequences:

"Artificial Intelligence can improve people's lives in so many ways, but researchers need to push for positive applications of technology by supporting a ban on autonomous weapons systems." (Hinton media interview, publicized with the pledge)

Such a quote illustrates how some AI developers delimited which future they wished to avoid (here, a future where autonomous weapons systems worked and generated harm) and which future they wished to participate in generating (here, a future with beneficial AI applications). Decisions and announcements regarding impossible futures were not necessarily permanent. AI developers charted impossible futures and adjusted what they did in an ongoing manner. Depending upon changing circumstances, developments, and reflections in the AI community, AI developers charted different, impossible and possible, futures over time. For instance, in February 2019, the research institute OpenAI released a partial version of a new text-generating AI program (GPT-2) that it had developed. OpenAI was a research institute whose self-defined mission was to investigate "*the path to safe artificial intelligence.*" Such mission emphasized generating a protected future with AI. AI developers at OpenAI at the time considered that the new program was ground-breaking but dangerous for its potential to develop credible deepfake texts:

"Our model, called GPT-2 (…), was trained simply to predict the next word in 40GB of Internet text. Due to our concerns about malicious applications of the technology, we are not releasing the trained model. As an experiment in responsible disclosure, we are instead releasing a much smaller model for researchers to experiment with, as well as a technical paper." (OpenAI blog post, 02/2019).

The authors of GPT-2 expressed concerns regarding potential misuse of their model and, in response, opted not to make it fully available to others. Yet, they also noted that this decision was "*an experiment in responsible disclosure*" and, hence, that it could be reversed. Several months later, in November 2019, OpenAI fully disclosed its updated model, explaining that, so far, it had not seen much evidence of misuse. The imagined futures regarding potential misuses had changed at OpenAI, prompting a different decision regarding the release of its model. This example involved AI developers thinking about potential future consequences of their work and counteracting such potentiality. By releasing their model at first only partly and then fully, these AI developers worked at ensuring that an unwanted future would remain impossible.

Discussion and Implications

Discussion and theoretical elaboration

Recent developments in AI have prompted reflections regarding the ways in which work, management, and organizations will become affected (Berente et al. 2021; Teodorescu et al. 2021). This study brings in the "*foresight of the future*" (Dewey 1922, p. 313) of AI developers to understand how it shapes their work and AI. It reveals how AI developers' considerations of AI's potential future consequences affects their development work. The imagined futures perspective sheds light on how AI developers reflect upon the consequences of their work when these consequences are not yet formed and cannot be perfectly predicted. The findings revealed two ideal-typical ways in which AI developers imagine the future associated with AI. The tangential model focuses more on the distant future (i.e., associated with a greater gap between current practices and the future potentials of AI) (Augustine et al. 2019) and the integral model on the near future (i.e., associated with a closer connection between current practices and AI in development). Moreover, the

tangential model offers a narrow perspective of the AI developer's identity as that of a scientist detached from the potential consequences of their work. The integral model favors a broader perspective of the AI developer's identity as a scientist involved in society. Boundary work (Gieryn 1983) also reflects the distinction between the two models. The tangential model underlines clear boundaries between AI developers' domain of work and others and the delegation to others of tasks considered outside of the domain of work. The integral model instead proposes a more porous view of the boundaries of the domain of expertise, emphasizing the importance of dealing with tasks that may be considered peripheral to limit potential negative consequences of the AI in development. The tangential and integral models constitute contrasting and stylized ways in which AI developers may consider the potential consequences of their work. However, as they work and participate in the AI development community, AI developers find ways to connect the two models. They do so by bringing some elements of one model (the integral one) onto the other (the tangential one) or by blending the two models.

Integrating the tangential model happens as some AI developers maintain the prestige of their scientific activities and uphold their imagination of the distant future associated with the AI in development. AI developers include some aspects of the integral model onto the tangential model by expanding the definition of expertise to include not only that of a scientist but also to acknowledge expertise associated with experience and social identities. They also expand the boundaries of their work to bring within the purview of their work the consideration and alleviation of some of the consequences of AI.

Changes in imagined futures associated with AI development support the blending of the two models. AI developers get to (re-)think the potential consequences associated with their work. They re-imagine the future as they adjust their expectations of what AI may do and take in consideration prior developments and technologies. AI developers also reflect upon current developments and work at bridging the near and distant imagined futures by rethinking how AI may fit with current and near future users and may unleash a distant future. Blending the two models further involves charting impossible futures, i.e., clarifying what the AI in development may not do. Impossible futures guide AI developers in their work by probing the inevitability of the technology they develop. These three ways of blending the models of expertise (re-imagining the future, bridging near and distant futures, and charting impossible futures) are interconnected. Changing imaginations of the future guide how AI developers approach their work, its priorities, and its restraints.

Scholarly implications

This study holds implications for scholarship on AI and the future, on engaging with the future, and on IT development.

For one, the future has been a key but so far under-examined aspect of AI. Berente et al. (2021, p. 1435) for instance explained that: "*AI is whatever we are doing next in computing*." AI has been viewed as holding the potential to generate the future of work, management, and societies. By proposing an imagined future perspective on AI developers' work, this study brings important light to this future aspect associated with AI. It enables us to adjust our perspective on AI. Instead of considering how AI precipitates the future, this study unpacks AI developers' imaginations of the future with AI. By illuminating how AI developers anticipate the future, then, this study helps gain a better understanding of how AI acquires its potentials.

This study examined how developers involved in the making of AI consider the future. It revealed how AI developers reflect upon AI-in progress, its eventual use, and possible outcomes. Doing so, this study enables scholars to lift the assumption of a clear separation among development, use, and consequences of use, as advocated by Bailey and Barley (2020) for AI and by Williams and Pollock (2012) for IT more generally. The imagined futures perspective is meaningful because AI is sophisticated, takes long to develop, and has consequences that emerge over time. This study revealed how the ambivalent potentials of AI are on the minds of some of its developers. It also unpacked how different perspectives among AI developers regarding what these potentials could be and regarding the roles that their own work may play in unlocking them are associated with diverse imagined futures.

This study further adds to scholarship by revealing how AI developers deal with disparities in imagined futures and how they might change their imaginations of the future associated with their work. People's projections of the future are not static (Mische 2009). This study documented how AI developers' imaginations of the future change. Such changes affect AI development practices and the eventual AI. Past

and current imaginations of the future shape and reshape AI, at times leading AI developers to chart impossible futures that their work may not unleash.

Moreover, this study also holds implications for recent scholarship that has called upon IS researchers to engage with and theorize for the future. This study in particular provides an answer to Hovorka and Peter's (2021) call for "*speculative engagement*" with the future. By bringing emphasis on AI developers' imagined futures, this study reveals how speculative engagement is not only the purview of academics but is also routinely accomplished by practitioners. This speculative engagement affects what they accomplish.

This study connected boundary work and identity work to the engagement with the future. It also revealed how practitioners engage with the near and/or distant future. Doing so, this study adds to seminal scholarship that had so far emphasized how a single model of expertise characterizes work within a technology development community. Unity in the model of expertise had been seen as key to the "*solidification process*" whereby science progresses and technology develops (Knorr Cetina 1981). This study adds to these established insights by revealing how technology developers within the same community may abide by distinct models of expertise based on their imaginations of the future.

Finally, this study adds to scholarship on IT development by revealing how different imagined futures are associated with tensions within IT development. Existing scholarship had shown how people working together from different domains need to engage in boundary work. This study documents how what is and is not part of the work of technology development can be contested from within. Scholarship has examined how people engage in boundary work as they deal with external pressures such as from being from different occupations. It has also explained how external pressures affect boundaries of work (Abbott 2014). This study reveals that boundaries of technology development work can be internally contested. The imagined futures perspective explains how technology developers may anticipate dissimilar future consequences associated with their work and may differ in their considerations of how to address them. Technology developers may have different, broad or narrow, definitions of the community and its boundaries depending upon their perspective on whether and how to include dealing with the consequences of technology development as part of their work.

Furthermore, existing scholarship has long recognized the homogeneity in terms of ethnicity and gender of IT workers (Ahuja 2002). It has shown how this homogeneity can make it difficult for minority IT workers to be recognized for their expertise (McGee 2018). This study revealed how distinct demographics contributed to AI developers imagining different futures associated with their work. As such, this study adds to scholarship by documenting a connection between some demographics and social identities and the closeness of some AI developers with potential negative consequences of their work. Demographics and social identities can thus be anchors for the imagination of future consequences of a technology in development, making them appear more clearly to some than to others.

Implications for practice and policy

Practical and policy implications of this study stem from its highlighting the importance of imagined futures in the work of AI developers. While it is now customary for AI conferences as well as for training programs and codes of conduct to include discussions of the ethical implications of AI, this study reveals that these implications are tied to the ways in which AI imagined the future and to their temporal horizon. AI developers who privilege the distant future may need to also start imagining the near future. Reversely, AI developers who privilege the near future may need to keep an eye as well as on the distant future. Training programs, codes of conducts, and conferences in AI could thus develop guidelines for AI developers to incorporate these two models of expertise in the work.

Future research

The limitations of this study constitute boundary conditions as well as opportunities for further research. For one, this study did not provide a longitudinal perspective on AI developers and their work. It would be useful to adopt a longitudinal perspective to see how AI development and its practitioners have changed over time. Also, this study did not follow the entire process of technology development, implementation, and use, as advocated by Bailey and Barley (2020) and Williams and Pollock (2012). Future scholarship could focus on an AI project and examine its complete process over time, taking in consideration the actions and perspectives of multiple actors throughout the process.

Conclusion

AI holds the potential to change how we work, organize, and function in society. This study illuminated how AI developers imagine the future consequences of their work. This helps us comprehend how the AI development community and its members work and shape AI. This also highlights the importance of AI developers' imaginations of what their work may do. Developing further this understanding and addressing tensions in AI developers' distinct perspectives is significant as increasingly sophisticated AI becomes developed and implemented.

References

Abbott, A. 2014. *The System of Professions: An Essay on the Division of Expert Labor*. University of Chicago press.

Ahuja, M. K. 2002. "Women in the Information Technology Profession: A Literature Review, Synthesis and Research Agenda," *European Journal of Information Systems* (11:1), pp. 20-34.

Asatiani, A., Malo, P., Nagbøl, P. R., Penttinen, E., Rinta-Kahila, T., and Salovaara, A. 2021. "Sociotechnical Envelopment of Artificial Intelligence: An Approach to Organizational Deployment of Inscrutable Artificial Intelligence Systems," *Journal of the Association for Information Systems* (22:2), p. 8.

Augustine, G., Soderstrom, S., Milner, D., and Weber, K. 2019. "Constructing a Distant Future: Imaginaries in Geoengineering," *Academy of Management Journal* (62:6), pp. 1930-1960.

Bailey, D. E., and Barley, S. R. 2020. "Beyond Design and Use: How Scholars Should Study Intelligent Technologies," *Information and Organization* (30:2), p. 100286.

Baird, A., and Maruping, L. M. 2021. "The Next Generation of Research on Is Use: A Theoretical Framework of Delegation to and from Agentic Is Artifacts," *MIS Quarterly* (45:1), pp. 315-341.

Beckert, J. 2013. "Imagined Futures: Fictional Expectations in the Economy," *Theory and Society* (42:3), pp. 219-240.

Benbya, H., Pachidi, S., and Jarvenpaa, S. 2021. "Special Issue Editorial: Artificial Intelligence in Organizations: Implications for Information Systems Research," *Journal of the Association for Information Systems* (22:2), pp. 281-303.

Berente, N., Gu, B., Recker, J., and Santhanam, R. 2021. "Managing Artificial Intelligence," *MIS Quarterly* (45:3), pp. 1433-1450.

Borup, M., Brown, N., Konrad, K., and Van Lente, H. 2006. "The Sociology of Expectations in Science and Technology," *Technology Analysis & Strategic Management* (18:3-4), pp. 285-298.

Buolamwini, J., and Gebru, T. 2018. "Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification," in: *Proceedings of the 1st Conference on Fairness, Accountability and Transparency*. Proceedings of Machine Learning Research: PMLR, pp. 77--91.

Canals, J., and Heukamp, F. (eds.). 2020. *The Future of Management in an Ai World*. Springer.

Charmaz, K. 2006. *Constructing Grounded Theory: A Practical Guide through Qualitative Analysis*. London: Sage publications.

Crawford, K. 2016. "Artificial Intelligence's White Guy Problem," in: *The New York Times*.

Crawford, K., Dobbe, R., Dryer, T., Fried, G., Green, B., Kaziunas, E., Kak, A., Mathur, V., McElroy, E., and Sánchez, A. N. 2019. "Ai Now 2019 Report," *New York, NY: AI Now Institute*).

Denzin, N. K., and Lincoln, Y. (eds.). 2000. *Handbook of Qualitative Research*. London, UK: Sage.

Dewey, J. 1922. "Human Nature and Conduct - an Introduction into Social Psychology." George Allen & Unwin, London.

Faraj, S., Pachidi, S., and Sayegh, K. 2018. "Working and Organizing in the Age of the Learning Algorithm," *Information and Organization* (28:1), pp. 62-70.

Frye, M. 2012. "Bright Futures in Malawi's New Dawn: Educational Aspirations as Assertions of Identity," *American Journal of Sociology* (117:6), pp. 1565-1624.

Gieryn, T. F. 1983. "Boundary-Work and the Demarcation of Science from Non-Science: Strains and Interests in Professional Ideologies of Scientists," *American Sociological Review* (48:6), pp. 781-795.

Grodal, S., Anteby, M., and Holm, A. L. 2021. "Achieving Rigor in Qualitative Analysis: The Role of Active Categorization in Theory Building," *Academy of Management Review* (46:3), pp. 591-612.

Hovorka, D. S., and Peter, S. 2021. "Speculatively Engaging Future(S): Four Theses," *MIS Quarterly* (45:1), pp. 461-466.

Jasanoff, S., and Kim, S.-H. 2009. "Containing the Atom: Sociotechnical Imaginaries and Nuclear Power in the United States and South Korea," *Minerva* (47:2), p. 119.
Kane, G. C., Young, A. G., Majchrzak, A., and Ransbotham, S. 2021. "Avoiding an Oppressive Future of Machine Learning: A Design Theory for Emancipatory Assistants," *MIS Quarterly* (45:1), pp. 371-396.
Kellogg, K., Valentine, M., and Christin, A. 2020. "Algorithms at Work: The New Contested Terrains of Control," *Academy of Management Annals* (14:1), pp. 366-410.
Knorr Cetina, K. 1981. *The Manufacture of Knowledge: An Essay on the Constructivist and Contextual Nature of Science*. Elsevier.
Lebovitz, S., Levina, N., and Lifshitz-Assaf, H. 2021. "Is Ai Ground Truth Really "True"? The Dangers of Training and Evaluating Ai Tools Based on Experts' Know-What," *MIS Quarterly* (45:3), pp. 150-1525.
Marabelli, M., Vaast, E., and Li, J. L. 2021. "Preventing the Digital Scars of Covid-19," *European Journal of Information Systems* (30:2), pp. 176-192.
Marjanovic, O., Cecez-Kecmanovic, D., and Vidgen, R. Forthcoming. "Theorising Algorithmic Justice," *European Journal of Information Systems*), pp. 1-19.
Martin, K. E. 2019. "Designing Ethical Algorithms," *MISQ Executive* (18:2), pp. 129-142.
Mayer, A.-S., Strich, F., and Fiedler, M. 2020. "Unintended Consequences of Introducing Ai Systems for Decision Making," *MIS Quarterly Executive* (19:4), pp. 239-256.
McGee, K. 2018. "The Influence of Gender, and Race/Ethnicity on Advancement in Information Technology (It)," *Information and Organization* (28:1), pp. 1-36.
Miles, M. B., and Huberman, A. M. 1994. *Qualitative Data Analysis - an Expanded Sourcebook*, (2nd ed.). London: Sage.
Mische, A. 2009. "Projects and Possibilities: Researching Futures in Action," *Sociological Forum* (24:3), pp. 694-704.
Orr, W. and Davis, J. L. 2020. "Attributions of ethical responsibility by Artificial Intelligence practitioners." *Information, Communication & Society* (23:5), pp. 719-735.
Riemer, K., and Peter, S. 2020. "The Robo-Apocalypse Plays out in the Quality, Not in the Quantity of Work," *Journal of Information Technology* (35:4), pp. 310-315.
Rinta-Kahila, T., Someh, I., Gillespie, N., Indulska, M., and Gregor, S. Forthcoming. "Algorithmic Decision-Making and System Destructiveness: A Case of Automatic Debt Recovery," *European Journal of Information Systems*), pp. 1-26.
Stahl, B. C., and Markus, M. L. 2021. "Let's Claim the Authority to Speak out on the Ethics of Smart Information Systems," *MIS Quarterly* (45:1), pp. 485-488.
Strich, F., Mayer, A.-S., and Fiedler, M. 2021. "What Do I Do in a World of Artificial Intelligence? Investigating the Impact of Substitutive Decision-Making Ai Systems on Employees' Professional Role Identity," *Journal of the Association for Information Systems* (22:2), p. 9.
Teodorescu, M. H., Morse, L., Awwad, Y., and Kane, G. C. 2021. "Failures of Fairness in Automation Require a Deep Understanding of Human-Ml Augmentation," *MIS Quarterly* (45:3), pp. 1483-1499.
Vaast, E., and Pinsonneault, A. 2021. "When Digital Technologies Enable and Threaten Occupational Identity: The Delicate Balancing Act of Data Scientists," *MIS Quarterly* (45:3), pp. 1087-1112.
van den Broek, E., Sergeeva, A., and Huysman, M. 2021. "When the Machine Meets the Expert: An Ethnography of Developing Ai for Hiring," *MIS Quarterly* (45:3), pp. 1557-1580.
Washington, A. L., and Kuo, R. 2020. "Whose Side Are Ethics Codes On? Power, Responsibility and the Social Good," *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*, pp. 230-240.
Williams, R., and Pollock, N. 2012. "Research Commentary—Moving Beyond the Single Site Implementation Study: How (and Why) We Should Study the Biography of Packaged Enterprise Solutions," *Information Systems Research* (23:1), pp. 1-22.
Zhang, Z., Nandhakumar, J., Hummel, J., and Waardenburg, L. 2020. "Addressing the Key Challenges of Developing Machine Learning Ai Systems for Knowledge-Intensive Work," *MIS Quarterly Executive* (19:4).