

Association for Information Systems

AIS Electronic Library (AISeL)

ICIS 2022 Proceedings

AI in Business and Society

Dec 12th, 12:00 AM

Managing Organizational Identity Challenges Caused By AI Implementation: The Role Of AI Principles

Alexandra Haimerl

University of Passau, alexandra.haimerl@icloud.com

Anne-Sophie Mayer

Vrije Universiteit Amsterdam, a.s.mayer@vu.nl

Marina Fiedler

University of Passau, marina.fiedler@uni-passau.de

Follow this and additional works at: <https://aisel.aisnet.org/icis2022>

Recommended Citation

Haimerl, Alexandra; Mayer, Anne-Sophie; and Fiedler, Marina, "Managing Organizational Identity Challenges Caused By AI Implementation: The Role Of AI Principles" (2022). *ICIS 2022 Proceedings*. 10. https://aisel.aisnet.org/icis2022/ai_business/ai_business/10

This material is brought to you by the International Conference on Information Systems (ICIS) at AIS Electronic Library (AISeL). It has been accepted for inclusion in ICIS 2022 Proceedings by an authorized administrator of AIS Electronic Library (AISeL). For more information, please contact elibrary@aisnet.org.

Managing Organizational Identity Challenges Caused By AI Implementation: The Role Of AI Principles

Short Paper

Alexandra Haimerl
University of Passau
Innstraße 27
94032 Passau
alexandra.haimerl@gmx.de

Anne-Sophie Mayer
Vrije Universiteit Amsterdam
De Boelaan 1105
1081 HV Amsterdam
a.s.mayer@vu.nl

Marina Fiedler
University of Passau
Innstraße 27
94032 Passau
marina.fiedler@uni-passau.de

Abstract

Alongside growing external pressure for implementation of Artificial Intelligence (AI) technologies, multi-stakeholder demands for responsible conduct have led to an increasing number of organizational AI principles. As previous research on AI principles has mainly focused on their content, restrictions, and external functions, little is known about their relevance for organization-internal stakeholders. Concurrently, while organizational identity was shown to play a central role in technology implementation's success or failure, with respect to AI implementation, the concept has remained unexplored. Building on 25 expert interviews as part of an ongoing research that involves a qualitative, cross-industry multiple-case study with 13 organizations, we reveal AI principles' capacities for managing organizational identity towards AI implementation by: (1) redefining organizational identity, and (2) aligning organizational identity's facets. Our findings accentuate the relevance of organizational identity for AI implementation, and indicate AI principles' role as a tool to manage this transformative change in an identity-conforming way.

Keywords: AI principles, organizational identity, AI implementation

Introduction

Sometimes, [employees' feedback] is 'Software Corp. should not be doing AI!' (...) I think, you know, when we're a large multinational tech firm, whether or not we like it, AI is part of the future of our company, and to just say we can't do it is not the right answer. It's: How do we do it appropriately? So, we have to balance all of these different interests. (Sarah, Software Corp.)

The above statement demonstrates both the increasing external pressure on organizations to adopt Artificial intelligence (AI) technologies, and organizational stakeholders' conflicting interests resulting from the technology's disruptive potential (Faraj et al. 2018; Hagendorff 2020). AI technologies are developed to generate predictive models that generate new knowledge that exceeds experts' knowledge by inferring patterns from large amounts of data (Faraj et al. 2018). Thereby, these technologies can be used

for augmenting or automating work and decision-making (Raisch and Krakowski 2021). AI technologies are increasingly introduced in organizations to enhance productivity, promote innovation, enhance creativity, and create new competitive advantages (Brynjolfsson and Mitchell 2017; Faraj et al. 2018). Despite the multiple potentials AI systems hold for organizations, AI technologies are also characterized by their unforeseeable future trajectories and an exceptional variety of related potential risks such as non-transparency, discrimination or uncertainty (Faraj et al. 2018; Harari 2017). Thus, the implementation of AI technologies amplifies both the pressure for organizational change (Harari 2017) and multi-stakeholder demands to control AI's potential negative impacts (Hagendorff 2020).

Arguably in response to such demands, recently, there has been an increasing trend among organizations to issue documents containing non-legislative policy, most often referred to as *AI (ethics) principles* or *AI (ethics) guidelines* (Jobin et al. 2019; Kelley 2022; Morley et al. 2020). AI principles are formal documents that are developed or selected by organizations to state normative declarations about how AI systems are intended to be used by organizations' managers and employees (Kelley 2021). They act as a mutation of business codes (Kelley 2021) and reflect the general need for ethical guidance of AI in absence of external regulation (Jobin et al. 2019). So far, literature has mainly focused on AI principles' organization-external functions. In this regard, corporate motives have been questioned, assuming, for example, that companies intended to impede the development of a regulatory policy infrastructure (Calo 2017), or shape, based on their own preferences and interests, what is perceived to be ethical AI (Greene et al. 2019).

From an organization-internal perspective, the implementation of transformational innovation, such as AI, significantly challenges organizations and their identity, as perceived by organizational members (Altman and Tripsas 2015; Hinings et al. 2018; Kammerlander et al. 2018; Tyworth 2014). As the key institutional change agents (Hinings et al. 2018), employees play a decisive role in organizational change initiatives' success or failure in general (Neves et al. 2018) and with respect to technology implementation in particular (Altman and Tripsas 2015; Kammerlander et al. 2018). At the same time, employees' central role often remains underestimated, even though social factors are among the major barriers to AI implementation (Cubric 2020). Especially "attempting to adopt technologies seen as incompatible with the organization's identity are [*sic*] likely to meet significant institutional resistance" (Tyworth 2014, p. 81), and, therefore, to eventually fail (Anthony and Tripsas 2016).

Organizational identity refers to employees' shared concepts of the characteristics that define their organization, and how it, therefore, relates to and differs from other organizations (Altman and Tripsas 2015). The concept consists of substantive identity aspects, which are shaped by an organizations' products and processes, and intangible identity aspects, which comprise organizational members' perceptions of core organizational values and beliefs (Ashforth et al. 2011; Kammerlander et al. 2018). Prior research emphasized that the introduction of new information systems (IS) can challenge employees' organizational identity (Anthony and Tripsas 2016). Especially the introduction of technologies that do not align with the organization's identity is often associated with high institutional resistance (Tyworth 2014), and, thus, high failure rates (Anthony and Tripsas 2016). However, in contrast to conventional IS, AI systems are able to learn (Faraj et al. 2008), take over entire core business activities (Strich et al. 2021), and cause new ethical and moral questions (Faraj et al. 2018; Rai et al. 2019). Therefore, these systems are assumed to challenge employees' organizational identity to a new extent. Given the repeatedly shown relevance of *organizational identity* for technology implementation and the variety of AI-related stakeholder concerns, it is surprising that, despite sharply rising numbers in publications on AI implementation (Cubric 2020), so far, the role of organizational identity has remained unexplored.

In an attempt to reduce this research gap, we set out to shed light on the effect AI principles have from an internal stakeholder perspective, which acknowledges the complexity of the new pressures and demands organizations face with AI implementation. Simultaneously, addressing AI principles from an internal stakeholder perspective, we re-examine and ultimately challenge previous academic literature's predominantly negative perception of AI principles. As organizational identity is a central parameter for the understanding and managing of social dynamics that come along with organizational technology and which are essential for successful technology outcomes (Tyworth 2014), we address the following research question: "*How do AI principles contribute to managing organizational identity challenges caused by AI implementation?*"

To answer this research question, we build on findings from an ongoing qualitative research study. This study involves to date 25 interviews with experts from 13 organizations operating in or across different

industries in Germany. Following an iterative grounded theory approach (Gioia et al. 2013; Strauss and Corbin 1990), we identify several categories that show how AI principles can contribute to managing organizational identity conflicts caused by the implementation of AI. Thereby, we illustrate the interactions between AI principles, organizational identity, and AI implementation. Our findings accentuate the relevance of organizational identity for AI implementation, and suggest AI principles' potential to (1) mitigate the identity threats of AI technologies and (2) manage AI implementation in an identity-conforming way. Therewith, our study provides theoretical contributions and practical implications for organizations facing AI implementation alike.

Methods

Our research is a cross-industry multiple-case study. Given that (1) organizational identities are organization-specific, cognitive schemas and (2) organizational domain identity is fundamentally shaped by products and markets, we considered including diverse cases conducive to answering the research question. Ideally, the respective organization had already published (about) their AI principles. Additionally, we contacted organizations, which we considered likely to (have) issue(d) AI principles, as they applied AI in fields prone to ethical challenges. AI principles were designed and structured in a similar way across organizations. In all organizations, AI principles were issued as formal documents that stated how the respective AI technology or how AI technologies in general are intended to be used (e.g. emphasis on work augmentation and support of employees), which AI technologies will and will not be used (e.g. focus on supervised learning and no use of neural networks) and which values are important for the organization when working with AI technologies (e.g. fair and explainable systems, highest possible degree of transparency, focus on data and privacy protection etc.). In this context, we particularly selected organizations that use AI technologies for work augmentation rather than for work automation because we wanted to understand how AI principles affect employees' perception of their organization's organizational identity in the context where employees work with the AI technology instead of where they are replaced. In most cases, we emailed several potential informants within the same organization, hoping to gain broader and more balanced insights into the views of experts in different positions. Determined by positive responses and confirmed suitability for our research, the final case selection comprises 13 organizations operating in six industries in Germany. Besides eleven corporations, our study also includes two associations (hereinafter referred to as unions) that represent either companies or their employees. Integrating the latter provided perspectives beyond the boundaries of an individual corporation or industry.

We conducted 25 semi-structured interviews with at least one expert from each of the 13 organizations between May 2020 and June 2021. As developments in this study's underlying topics are highly dynamic and fast-moving, three of the respondents were interviewed twice. While we originally focused on top-level and mid-level managers as informants, for some organizations, it turned out beneficial to also interview data scientists. The informants' individual roles, as reflected in their job titles, show a wide range of their respective area of expertise. Contacting experts in different roles, we chose this diversity in the cases deliberately in order to cover both the technical and management perspective and, thus, contribute a balanced view of AI technologies' and AI principles' influence on employees' perceptions.

An interview guide informed by relevant literature led the conduct of all interviews. Its questions were based on concepts established in previous studies in the fields of organizational identity or AI. They either surveyed focal concepts in the applicational context of the respective organization or synthesized such concepts for the context of organizational identity in relation to AI (principles). Specifically, questions centered around informants' perception of (1) their organization's identity, (2) how AI technologies' application influences their organizational identity, and (3) how AI principles affect the relationship between AI technologies and organizational identity.

As we had conducted interviews in the setting of a previously unexplored topic, our data analysis followed an iterative grounded theory approach (Strauss and Corbin 1990). Thereby, we analyzed the interview data inductively, following Gioia et al.'s (2013) recursive three-step coding process. In the first step, we perused each transcript individually, adhering faithfully to informants' statements about AI principles' organization-internal functions, potentials, and capabilities. Starting with an open coding, we also encoded text segments and passages, which conveyed challenges and potentials of organizational AI implementation in general. Considering the interrelated complexity of the organizational context, we

anticipated that these might become relevant at a later, more abstract stage of the coding process, where connections might emerge. In the second step, we clustered the identified first order concepts and categorized them to more abstract second order themes. Thereby, we moved back and forth between emergent themes and paying particular attention to different aspects of organizational identity as derived from theory (Gustafson and Reger 1995; Kammerlander et al. 2018). In the data analysis' third and final step, we will combine the second order themes to aggregate dimensions. As this paper is part of an ongoing research project, we will focus on two aggregate dimensions in the following.

Findings

To gain an in-depth understanding of how AI principles contribute to managing organizational identity, we first explored what challenges AI technologies can cause in terms of organizations' organizational identity. In a second step, we examined AI principles' perceived relevancy and function for organization-internal stakeholders.

AI technologies as a challenge to organizational identity

AI technologies often evoke ethical challenges that affect organizational identity. AI systems embody values through technological development and design choices; yet these can conflict with the individual organization's implicit and explicit values, that is, organizations' intangible identity. In view of the diverse ethical issues related to AI technologies, with respect to organizational identity, management faces the challenge to ensure that employees can continue identifying with the ways the individual organization – including its artificial actors – operates.

Our interviews disclosed that the implementation of AI technologies can challenge organizational identity because of their (1) ability to learn and act autonomously, and (2) promise to outperform human experts. First, in contrast to former disruptive technologies, which primarily led to an automation and information of existing organizational processes, AI technologies are distinguished by their human-like ability to think and act intelligent and to learn and operate autonomously. While AI technologies rely on human-controlled training data as starting instructions only, they autonomously refine or even evolve their logic and patterns, connections, and weighting, as they learn from additional data points. As a result, employees are confronted with increased opacity and inscrutability of AI technologies' decisions and outputs which shapes the way employees perceive their organization. As one expert states:

Many organizations always paid high attention to being transparent towards their employees, and of course towards customers and other stakeholders, but especially towards their employees. And this underlying organizational guideline or standard is now being challenged or even endangered by the increasing use of AI tools that have contradicting attributes and values. So the way employees perceive the values, visions and standards of their organization can be drastically affected. (Peter, Fintec Corp.)

Second, AI technologies promise to generate knowledge and decisions superior to human experts. Moreover, AI technologies perform increasingly complex functions, which, by far, exceed the narrowly defined and less sophisticated tasks executed by previous technologies. As a result, many AI technologies are able to take over not only isolated knowledge work tasks, but knowledge workers' entire core activities. This potential to devalue expertise and displace knowledge workers can impact employees' perception of their organization's organizational identity. As one interviewee explains:

I worked for one organization that always promoted itself as a secure, loyal and fair employer. But this organization also struggled a lot with competitive pressure in several domains and the use of algorithms offered in many fields a more effective and cost-reducing way of working. They didn't quit employees but they also didn't hire new employees for several jobs that could mainly be managed by algorithms. And of course this had an impact on employees and their image of their employer. How can you promise to be a stable employer while increasingly buying these AI tools? That was a very common question I heard a lot. (Milo, Consulting Corp.)

AI principles' potential to manage organizational identity

To manage the described challenges in organizations' organizational identity due to the implementation of AI technologies, many organizations issue AI principles. In the following, we will elaborate on two aggregate dimensions that emphasize how AI principles can contribute to managing organizational identity, namely (1) redefinition of organizational identity, and (2) alignment of organizational identity's facets.

Redefinition of organizational identity

The first group of AI principles' functions contributes to redefining organizational identity by fostering a shared understanding, a common vision, and employer branding. First, our interviews showed that, as AI introduces a range of new concepts, questions, and risks, AI principles are a useful tool to create a shared understanding of AI-related concepts within an organization. Accordingly, a shared understanding establishes an organizational frame for the management of AI-related issues. This frame provides a basis for inter-divisional and inter-project congruence, and prevents employees and organizations from unwanted uncertainties and risks. In this regard, AI principles emerged as a beneficial tool as "there is just nothing worse than when one department answers them one way and the other department answers them another way, and you are confused afterwards" (Carl, Communication Corp.). In several cases, the need to create a shared understanding that represents the organization's definition of AI-related concepts is the major motivation for formulating AI principles:

There was also a discussion between me and our boss that, presumably, we all have the same understanding somehow, but as long as it's not written down anywhere, we don't know. And that's why we formulated [our AI principles] that way, both in terms of certain principles and certain measures. And then we discussed them in various team meetings - we had a few iterations to make sure that it was a team view and that the different perspectives were included. (Isaac, Insurance Corp.)

Furthermore, our interviews indicated that due to the fact that AI and AI ethics contain controversial issues, it is not uncommon for diverse organization-internal stakeholders to hold differential views on related topics. The process of defining AI principles opens an inner-organizational discourse that brings together different perspectives, and ultimately demands a decision for a certain option, which all organizational members will commit to. Our interviews disclosed that a team-based approach to defining AI principles facilitates a decision-making process that yields a clear organizational stance in consideration of the respective organizational identity. Consequently, organizational members share and commit to the organizational approach, even if, on some topics, they may have diverging personal preferences.

We're just not a tech booth like Google that just does things because they're feasible. Even though, personally, I find that totally fascinating and would love to work that way, but that's not Communication Corp., and sometimes that's also our pain. But the question [at Communication Corp.] must always be: What is the added value, especially for people? (Carl, Communication Corp.)

Second, our interviews demonstrated that discrepancies in views on AI do not only restrict to specific concepts, but also appear in employees' general attitudes towards the technology. As the development of AI principles is often the first step to an organization-wide consideration of AI-related topics, AI principles "open up a culture of conversation and enable communication about things that could not really be talked about otherwise – especially not so across different levels of hierarchy" (Danna, Digital Union). Thereby, identity-shaping questions are subjects of discussion of notable relevance. Especially in cases of differing ideas on AI, AI principles appeared to be a useful tool to consolidate organizational members in the pursuit of a common goal.

And it is only through this stakeholder discourse that you actually get the acceptance that both employees who are fearful of this and employees who are (...) very open to it, that they sit down together and say: We have AI principles here to work jointly on one goal. (Gregory, Consultancy Corp.)

Our interviews indicated that including organizational values in such goals is momentous. Accordingly, a common vision "to achieve more, not in the sense of more sales and more efficiency, but actually achieving a positive impact" (Thomas, Technology Corp.) contributes to organizational AI implementation.

I see a strong connection to our values, and that's why (...) we have to let [technology] become part of our organization. That's why a cultural change is also the case there. I also observe this partially in more innovative companies: I notice that where artificial intelligence is only perceived as a means of improving productivity and increasing efficiency, and this common vision is not being drawn, that it is difficult at that point. It simply is not imaginable that people would be open to a disruptive technology there. (Grant, Consultancy Corp.)

The third potential of redefining organizational identity through AI principles emerging from our interviews, is a conscious amelioration of the organization as an employer. Accordingly, while “artificial intelligence – much like digitalization – does not necessarily have a positive radiant effect on people” (Grant, Consultancy Corp.), “one wants to identify in that AI is also synonymous with progress and with brand new technology” (David, Digital Union). Given this ambivalence, tying the technology’s use to certain principles, AI principles hold the potential to add to an organization’s attractiveness as an.

Furthermore, the values codified in principles were also linked to employees’ “motivation to work in this company and identify with the company. And if that's not the case, then the employees won't realize their potential and the company's success will suffer in the long term” (Conor, Communication Corp.). Accordingly, AI principles also foster employees’ self-identification with their organization, which was also linked to both current and potential employees’ employment decisions.

For me, it is important that I have that good feeling that we are working on a value structure, which I can support and where I don't have to bend (...). As an employee of the company, that gives me an extremely good feeling. And, yes, as a human being in the sense of working here. I wouldn't want to work here if that wasn't the case. (Thomas, Technology Corp.)

Alignment of organizational identity’s facets

The second group of AI principles’ functions contributes to the *alignment of organizational identity’s facets*. It demonstrates the significant interrelation between substantive and intangible aspects of organizational identity, and the ability of AI principles to harmonize both facets of organizational identity.

First, AI principles contribute to an alignment of organizational identity’s facets by embedding AI in the organization’s established identity. Our interviews demonstrated that AI implementation inevitably causes inner-organizational changes, many of which diverse groups of employees (e.g., regarding age, openness to change, particular jobs) hold diverging views on. In this regard, AI principles can be a tool to express that an organization’s “values are still the superordinate frame that AI should move in” (Isaac, Insurance Corp.). Accordingly, establishing AI principles can be motivated by the ambition to communicate AI technologies’ congruence with organizational values, thereby embedding AI in the established value system.

These central values and characteristics, which also include topics such as diversity (...), are supposed to be expressed again in the principles. Only then, in relation to algorithms, so that we say: we generally don't want to have any prejudices in our company and, of course, there shouldn't be any unconscious biases in our AI levels. So, we generally don't want to have them, and to that extent I would say it reflects the values again in relation to AI. (Zoe, Transformation Corp.)

Thus, “when [companies] develop ethical principles, then of course they want to strongly express, emphasize, underline these values. And, of course, companies intend to reinforce the values in the sense of (...) how they think about themselves.” (David, Digital Union). Analogously, the interviews showed that AI principles are also a means to express the continuity of and commitment to central organizational beliefs and organizational culture. As such, they emphasize that the organizational use of AI integrates into existent organizational culture, often with regard to corporate responsibility and business ethics.

It has always been the case at Transformation Corp. that this principle, that we want to bring good technology into the world, we always talk about 'good tech', has been around.

(...) If we bring such technologies into the world, which are so powerful and which can somehow influence and change society, then we also have a responsibility to society. And

that was probably the starting point that triggered [our AI principles]. (Zoe, Transformation Corp.)

Thereby, AI principles show the potential to establish a link between existent company culture and AI technologies, which was indicated to play a substantial role in AI technologies' implementation.

Changing a culture is extremely difficult and lengthy and, honestly, I think we've rather gone down the path of thinking: 'How do I marry our culture with the technical possibilities and how do we find that?' The principles should actually be the link, because we don't want to demonize technology. But we also don't just want to implement technology per se and because it's new technology, but our culture is to think: (...) How can I make people happier, better, enable them in some way through technology? (Carl, Communication Corp.)

Second, the interviews showed that AI principles are a basis for aligning substantive aspects of organizational identity – that is, products, processes, markets, and core competencies – with organizational values. In this regard, informants reported that organizations select products, projects, and customers dependent on their perceived fit with the organizational values as codified in their AI principles.

[AI principles] really are the official commitment to certain principles and also an expression of the corresponding responsibility that one [company] imposes on oneself or to which one submits. In some cases, this goes so far that certain principles stipulate, for example, that certain technologies are not to be sold to certain countries, certain stakeholders or other industrialists, because the company may not want to endorse their use or may expect the further development of its own technology in a direction that does not correspond to its own principles. The decisions that are made are definitely also financially drastic, simply in order to live up to the basic principles that one has imposed on oneself. (Danna, Digital Union)

Similarly, the codification of organizational values in AI principles also appeared to encourage the development of AI solutions that integrate and reflect core aspects of the respective organization's intangible identity.

The question arose very, very quickly about the ethical responsibility or of how to program this AI, so that we can also [incorporate] this orientation to the common good and how we can redeem this very, very high trust that citizens place in municipal companies (...), and how we can emphasize it as the determining element in the use of AI. (Paula, Public Union)

Moreover, AI principles were frequently mentioned as a basis for the establishment of new processes and tools that assure the implementation of organizational values. Corresponding to the respective organization's context and requirements, supporting processes and tools may range from workshops and trainings to checklists to technical toolkits, some of which are compulsory in AI-related work.

We have formulated very clearly, which basic requirements and specific requirements must be applied in projects. It is described very precisely, which steps have to be taken, which documents have to be created, which documentation has to be made, which workshop coordination has to take place. This is a task that is written in the books of every manager and every team manager, to actually implement the compliance [with our AI principles] in the applications. (Thomas, Technology Corp.)

Finally, the interviews also indicated that AI principles help organizations anticipate and manage changes regarding their core competencies. On the one hand, "a great deal of new know-how and technical understanding is required in order to feed AI with the appropriate data" (Paula, Public Union). Accordingly, especially in less technological industries, the implementation of AI necessitates new core competencies. AI principles were shown to bring awareness to such requirements, and to induce organizations to acquire capable talent. On the other hand, AI is often associated with employee replacement, even if this is not the intention by the management. AI principles emerged as an apt tool for exhibiting organizational value-coherence, for example, commitment to employees.

And the fact is, there are people behind it, and something has to happen to these people,

and we also support this process of how to deal with the client's employees – who had certain tasks that can now perhaps be taken over by an AI, or at least be simplified so that they take up less time – and how to train these employees, (...) so that they can, for example, take on other tasks that an AI simply cannot do yet. (Zoe, Transformation Corp.)

Conclusion

Our research set out to explore how AI principles can contribute to managing to organizational identity. We show that AI principles offer an important management tool that allows to address, guide, and form organizational identity in two major ways: AI principles enable the (1) redefinition of organizational identity, and the (2) alignment of organizational identity's facets. These aspects are crucial to manage possible organizational challenges through the implementation of AI technologies.

Our findings contribute to the literature on AI principles and organizational identity. First, contrary to widespread propositions that AI principles are primarily directed at external stakeholders in pursuit of corresponding goals, our study demonstrates AI principles' relevance for internal stakeholder management. Specifically, literature on AI ethics has often criticized corporations issuing AI principles for their presumed externally directed motives, such as impeding the development of regulatory infrastructure (Calo 2017; Jobin et al. 2019); deliberately shaping the AI ethics discussion in their own interests (Benkler 2019; Jobin et al. 2019); and calming public concerns, while not mitigating the underlying issues (Boddington 2017). Facing these evaluations, our study sheds light on AI principles' organization-internal significance. It indicates that AI principles are a multifunctional instrument, capable of demonstrating and enacting organizational (value) commitment, which – as binding in its nature – might also entail the forgoing of short-term financial profits. Additionally, our findings suggest that AI principles trigger the alignment of processes with organizational values, therewith influencing both organizational and individual action. Thus, as opposed to propositions that AI principles would be ineffective in influencing employee behavior (Hagendorff 2020), our findings indicate that AI principles are a practical tool for guiding employee action.

Second, our study demonstrates that while an organization's diverse groups of employees (as regards, e.g., age, openness to change, profession) frequently have diverging views on the change that AI introduction inevitably causes to aspects of substantive identity, aspects of intangible identity potentially have a uniting effect. Consequently, organizations employ AI principles to consciously frame AI introduction within the wider context of the organization's established intangible identity. Specifically, informants described AI principles' function to express the endurance of their organization's intangible identity and AI's congruence with the respective values, beliefs, or culture perceived to be core. Thus, by *integrating AI into the organization's established identity*, AI principles reveal the capacity to “managing identity threats from innovation (...) in a way that mitigates the identity threat and leaves identity intact” (Anthony and Tripsas 2016, p. 428). From a technology introduction perspective, this makes AI principles an important tool, especially, as other change strategies appear unviable in the AI context. For example, Tyworth (2014) recommends overcoming employees' resistance to identity-challenging technologies by aligning innovation and identity through an incremental change strategy. However, given the rapid pace of innovation and competitive forces in the AI context (Faraj et al., 2018), organizations are unlikely to have the time to implement incremental changes. In light of this dilemma, our study suggests AI principles as a tool to reduce AI-induced identity threats by aligning the technology with core aspects of organizational identity – thereby turning organizational identity from a constraint to an enabler of AI introduction (compare Anthony and Tripsas 2016).

Overall, our research shows the important role of AI principles in AI implementations processes. In our ongoing research project, we will dive deeper into AI principles' potentials for managing organizational identity and develop a framework that helps organizations to establish AI principles for organizational-internal processes and functions.

References

Altman, E., and Tripsas, M. 2015. Product-to-Platform Transitions: Organizational Identity Implications,

- in *The Oxford Handbook of Creativity, Innovation, and Entrepreneurship*, C. Shalley, M. A. Hitt, and J. Zhou (eds.), New York, NY, United States: Oxford University Press, pp. 379–394.
- Anthony, C., and Tripsas, M. 2016. “Organizational Identity and Innovation,” in *The Oxford Handbook of Organizational Identity*, M. G. Pratt, M. Schultz, B. E. Ashforth, and D. Ravasi (eds.), New York, NY, United States: Oxford University Press, pp. 417–435.
- Ashforth, B. E., Rogers, K. M., and Corley, K. G. 2011. “Identity in Organizations: Exploring Cross-Level Dynamics,” *Organization Science* (22:5), pp. 1144–1156.
- Benkler, Y. 2019. “Don’t Let Industry Write the Rules for AI,” *Nature* (569), pp. 161–161.
- Boddington, P. 2017. *Towards a Code of Ethics for Artificial Intelligence, Artificial Intelligence: Foundations, Theory, and Algorithms*, Cham, Switzerland: Springer Nature.
- Brynjolfsson, E., and Mitchell, T. 2017. “What Can Machine Learning Do? Workforce Implications,” *Science* (358:6370), pp. 1530–1534.
- Calo, R. 2017. “Artificial Intelligence Policy: A Primer and Roadmap,” *University of Bologna Law Review* (3:2), pp. 180–218.
- Cubic, M. 2020. “Drivers, Barriers and Social Considerations for AI Adoption in Business and Management: A Tertiary Study,” *Technology in Society* (62), Elsevier Ltd, Article 101257.
- 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.
- Gioia, D. A., Corley, K. G., and Hamilton, A. L. 2013. “Seeking Qualitative Rigor in Inductive Research,” *Organizational Research Methods* (16:1), pp. 15–31.
- Greene, D., Hoffmann, A. L., and Stark, L. 2019. “Better, Nicer, Clearer, Fairer: A Critical Assessment of the Movement for Ethical Artificial Intelligence and Machine Learning,” in *Proceedings of the 52nd Hawaii International Conference on System Sciences*, pp. 2122–2131.
- Gustafson, L. T., and Reger, R. K. 1995. “Using Organizational Identity To Achieve Stability and Change in High Velocity Environments,” in *Academy of Management Proceedings* (1995:1), pp. 464–468.
- Hagendorff, T. 2020. “The Ethics of AI Ethics: An Evaluation of Guidelines,” *Minds and Machines* (30:1), Springer Netherlands, pp. 99–120.
- Harari, Y. N. 2017. “Reboot for the AI Revolution,” *Nature* (550), pp. 324–327.
- Hinings, B., Gegenhuber, T., and Greenwood, R. 2018. “Digital Innovation and Transformation: An Institutional Perspective,” *Information and Organization* (28:1), pp. 52–61.
- Jobin, A., Ienca, M., and Vayena, E. 2019. “The Global Landscape of AI Ethics Guidelines,” *Nature Machine Intelligence* (1:9), pp. 389–399.
- Kammerlander, N., König, A., and Richards, M. 2018. “Why Do Incumbents Respond Heterogeneously to Disruptive Innovations? The Interplay of Domain Identity and Role Identity,” *Journal of Management Studies* (55:7), pp. 1122–1165.
- Kelley, S. 2022. “Employee Perceptions of the Effective Adoption of AI Principles,” *Journal of Business Ethics*.
- Morley, J., Floridi, L., Kinsey, L., and Elhalal, A. 2020. “From What to How: An Initial Review of Publicly Available AI Ethics Tools, Methods and Research to Translate Principles into Practices,” *Science and Engineering Ethics* (26:4), pp. 2141–2168.
- Neves, P., Almeida, P., and Velez, M. J. 2018. “Reducing Intentions to Resist Future Change: Combined Effects of Commitment-Based HR Practices and Ethical Leadership,” *Human Resource Management* (57:1), pp. 249–261.
- Rai, A., Constantinides, P., and Sarker, S. 2019. “Next-Generation Digital Platforms: Toward Human–AI Hybrids,” *MIS Quarterly* (43:1), iii–x.
- Raisch, S., and Krakowski, S. 2021. “Artificial intelligence and management: The automation–augmentation paradox,” *Academy of Management Review* (46:1), pp. 192–210.
- Strauss, A., and Corbin, J. M. 1990. *Basics of Qualitative Research: Grounded Theory Procedures and Techniques*, Thousand Oaks, CA, US: Sage Publications, Inc.
- 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), pp. 304–324.
- Tyworth, M. 2014. “Organizational Identity and Information Systems: How Organizational ICT Reflect Who an Organization Is,” *European Journal of Information Systems* (23:1), pp. 69–83.