

June 2024

## Translating AI Ethics Principles into Practice to Support Robotic Process Automation Implementation

Dörte Schulte-Derne

Ulrich Gnewuch

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

---

### Recommended Citation

Schulte-Derne, Dörte and Gnewuch, Ulrich (2024) "Translating AI Ethics Principles into Practice to Support Robotic Process Automation Implementation," *MIS Quarterly Executive*: Vol. 23: Iss. 2, Article 6. Available at: <https://aisel.aisnet.org/misqe/vol23/iss2/6>

This material is brought to you by the AIS Journals at AIS Electronic Library (AISeL). It has been accepted for inclusion in MIS Quarterly Executive by an authorized administrator of AIS Electronic Library (AISeL). For more information, please contact [elibrary@aisnet.org](mailto:elibrary@aisnet.org).

## Translating AI Ethics Principles into Practice to Support Robotic Process Automation Implementation

*When organizations leverage artificial intelligence (AI) to automate processes previously performed by people, it frequently causes uncertainty and fear among those affected. An often suggested way for organizations to navigate such challenges is to seek guidance from AI ethics principles. Leaders, however, find it difficult to make practical use of these abstract, high-level principles. Based on a case study of the large-scale implementation of robotic process automation at an energy service provider in Germany, we provide recommendations for translating AI ethics principles into practice.<sup>1,2</sup>*

**Dörte Schulte-Derne**  
HHL Leipzig Graduate School of  
Management (Germany)

**Ulrich Gnewuch**  
University of Passau (Germany)

### Ethical Challenges of Artificial Intelligence

With growing pressure to reduce costs, improve efficiency and maximize productivity, organizations increasingly rely on artificial intelligence (AI). However, using AI to automate processes and tasks that were previously performed by humans often causes uncertainty, raises concerns and instills fear among employees.<sup>3</sup> Such reactions are typically triggered by questions about the possibility of AI making employees' jobs obsolete, eroding their professional identity, disrupting their work routines and forcing them to reskill or upskill. Regardless of whether AI systems use rule-based or machine-learning approaches, these questions appear and must be addressed. As a consequence, organizations often face ethical challenges when using or planning to use AI. On the one hand, they are seeking to deploy AI to help them remain competitive in the long term by reducing costs. Hence the concerns about job losses; employee costs form a substantial proportion of the overall cost structure, particularly in the service sector. On the other hand, most organizations care deeply about their employees and want to provide them with a safe and secure working environment.



<sup>1</sup> Mary Lacity is the accepting senior editor for this article.

<sup>2</sup> This research was conducted as part of the "MeKIDI" project (project no. EXP.01.00019.20) funded by the German Federal Ministry of Labour and Social Affairs. The authors of this article are responsible for its content. We thank the senior editor, Mary Lacity, and the reviewers for their valuable feedback and guidance throughout the review process. We are also grateful to Claudia Lehmann, Alexander Maedche and Daniel Schloß for their helpful comments on earlier versions of this article, and to Anja Seiffer and Davinny Sou for their help in data collection and analysis. Finally, the authors thank all members of the project team at ESP (an anonymous German energy service provider).

<sup>3</sup> For an overview of employee reactions, see Seiffer, A., Gnewuch, U. and Maedche, A. "Understanding Employee Responses to Software Robots: A Systematic Literature Review," *Proceedings of the 42nd International Conference on Information Systems* (ICIS), 2021, December 2021.

To help organizations address the ethical challenges surrounding AI, many international organizations, private companies and research institutions have published AI ethics principles and guidelines over the past years. For example, the European Commission appointed a High-Level Expert Group on Artificial Intelligence that released its *Ethics Guidelines for Trustworthy AI* in 2019.<sup>4</sup> And the Berkman Klein Center for Internet and Society at Harvard University published a metastudy of 36 prominent sets of AI ethics principles in 2020.<sup>5</sup>

Research has shown that while there are differences in wording, a global consensus is emerging on the importance of five common AI ethics principles: *responsibility, justice and fairness, transparency, non-maleficence and privacy*.<sup>6</sup> Though these principles were identified through a considered and comprehensive approach, they tend to be formulated at a high level of abstraction, which makes them difficult to put into action.<sup>7</sup> As a result, a gap exists between high-level AI ethics principles and practical support for leaders to help them address the AI-related ethical challenges that their organizations face. In other words, even if organizations are willing to apply AI ethics principles, their leaders often lack the necessary knowledge to translate these abstract principles into practical guidance.

To address this challenge, we conducted a longitudinal case study at a German energy service provider, referred to as ESP for reasons

of anonymity. The case study's focus was the large-scale introduction and implementation of robotic process automation (RPA), starting in 2020. Over the next two and a half years, ESP implemented 45 RPA bots to automate a large number of back-office processes, primarily in the customer service and billing departments. These RPA bots enabled the automated processing of more than 200,000 back-office transactions per year, with 70% of them completed without any human involvement. Given that each transaction would have taken a human employee between three and five minutes to complete, the total working time saved was more than 9,300 hours, corroborating earlier findings on the business value of RPA.<sup>8</sup> However, rather than cutting jobs, ESP decided to use the freed-up capacity to redeploy employees to other, more value-adding, tasks (e.g., personalized customer consultations, cross-selling activities).

When faced with ethical challenges during its RPA implementation journey, ESP made significant efforts to address them by respecting the rights and interests of employees impacted by the introduction of the technology. These efforts were much in line with the underlying idea of AI ethics. For example, to increase transparency and reduce uncertainty among employees early on, ESP invited employees and works council members to watch, at any time, a live RPA bot perform its work on a dedicated workstation.

To better understand ESP's efforts and their impact on attitudes toward the RPA implementation, we interviewed, surveyed and observed employees in different roles at different levels of the organization, including back-office employees, process owners, leaders, works council members and consultants over the 30-month period starting in 2020 (see the Appendix for details of our research method).

Based on the lessons learned from the ESP case, we provide actionable guidelines for translating AI ethics principles into practice. Before presenting the case study in detail, we first briefly describe RPA and provide an overview of AI ethics principles.

4 *Ethics Guidelines for Trustworthy AI*, Independent High-Level Expert Group on Artificial Intelligence, European Commission, April 8, 2019, available at <https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai>.

5 Fjeld, J., Achten, N., Hilligoss, H., Nagy, A. and Srikumar, M. *Principled Artificial Intelligence: Mapping Consensus in Ethical and Rights-based Approaches to Principles for AI*, The Berkman Klein Center for Internet and Society, January 15, 2020, available at <https://cyber.harvard.edu/publication/2020/principled-ai>.

6 Our focus on these five principles is grounded in research by Jobin et al. who conducted a systematic meta-analysis of 84 AI ethics documents published by government agencies, private companies and research institutions from different geographic regions, including the U.S. (21), the European Union (19), the U.K. (13) and Japan (4). Though their results showed that no single principle appeared in all documents, five were referenced in more than half of them: *responsibility, justice and fairness, transparency, non-maleficence and privacy*, indicating an emerging global convergence across stakeholders on the importance of these five principles. For details, see Jobin, A., Ienca, M. and Vayena, E. "The Global Landscape of AI Ethics Guidelines," *Nature Machine Intelligence* (1), September 2019, pp. 389-399.

7 Munn, L. "The Uselessness of AI Ethics," *AI and Ethics* (3:3), August 2022, pp. 869-877.

8 For in-depth coverage of the business value of RPA, see: 1) Lacity, M. C. and Willcocks, L. P. "Robotic Process Automation at Telefónica O2," *MIS Quarterly Executive* (15:1), March 2016, pp. 21-35; and 2) Lacity, M., and Willcocks, L. "Becoming Strategic with Intelligent Automation," *MIS Quarterly Executive* (20:2), June 2021, pp. 169-182.

## Robotic Process Automation

Robotic process automation (RPA), which belongs to the class of “Automation AI,”<sup>9</sup> is a technology that enables digital processes to be automated through software robots (“bots”) that operate on the user interface in the same way as humans do, by logging in with an account and password, entering data and clicking buttons.<sup>10</sup> Typically, RPA bots are designed to interact with existing IT systems (e.g., Microsoft Outlook, SAP ERP) and perform routine tasks in a rule-based manner, such as copying and pasting data from one system to another. Though RPA is less “intelligent” than other AI technologies, it is usually considered part of AI, especially as RPA vendors are adding more intelligence to their software (e.g., integrating computer vision and machine learning capabilities).<sup>11</sup> Moreover, our case study revealed that, regardless of the intelligence level of RPA, almost everyone at ESP referred to it as “AI.”

RPA has been implemented successfully across a wide range of industries, such as finance, telecoms and healthcare, and its business benefits include increased operational efficiency, error reduction and overall cost savings.<sup>12</sup> Employees can also benefit from RPA because it frees them to focus on more interesting and value-adding tasks.<sup>13</sup>

However, the implementation of RPA, similar to other AI or IT implementation projects, can be challenging. According to HFS Research, only six in 10 RPA initiatives are meeting expectations.<sup>14</sup>

9 Reis, L., Maier, C., Mattke, J., Creutzenberg, M. and Weitzel, T. “Addressing User Resistance Would Have Prevented a Healthcare AI Project Failure,” *MIS Quarterly Executive* (19:4), December 2020, pp. 279-296.

10 Lacity, M. C. and Willcocks, L. P., op. cit., March 2016.

11 For a discussion of why RPA can be considered as part of AI, see Davenport, T. H. and Ronanki, R. “Artificial Intelligence for the Real World,” *Harvard Business Review* (96:1), January 30, 2018, pp. 108-116.

12 Techatassanasoontorn, A. A., Waizenegger, L. and Doolin, B. “When Harry, the Human, met Sally, the Software Robot: Metaphorical Sensemaking and Sensegiving around an Emergent Digital Technology,” *Journal of Information Technology* (38:4), February 2023, pp. 416-441.

13 Ibid.

14 HFS Research surveyed 511 Global 2000 companies to understand the state of automation. The results show that more than a third of RPA initiatives fail to deliver the expected improvements in business outcomes and more than 40% are unable to realize the projected cost savings. See *Market Impact Report: Automation is Back with a Bang!*, HFS Research, September 6, 2022, available at <https://www.hfsresearch.com/research/automation-is-back-with-a-bang/>.

An often-overlooked challenge is that the implementation of RPA can lead to disruptions in work routines, changes in organizational structures and redesigned roles for employees, which could result in fears of job losses and competition between RPA bots and employees.<sup>15</sup> Thus, when implementing RPA, organizations must proactively address employees’ fears, obtain their buy-in, and prevent panic and sabotage.<sup>16</sup>

## Ethics Principles for Artificial Intelligence

The growing sophistication and adoption of AI has raised several ethical challenges, including issues of responsibility, fairness, and transparency. As a result, there is intense public discussion about AI ethics, and many politicians have highlighted the need for new AI policies. To help address ethical challenges surrounding the design, development, implementation and use of AI, many international organizations (including the OECD, European Union and Future of Life Institute), private companies (e.g., Microsoft, SAP), and research institutions (e.g., the Association for Computing Machinery, the Institute of Electrical and Electronics Engineers) have published AI ethics principles and guidelines over the past years.<sup>17</sup> These principles offer valuable insights for typical change management activities, such as communication, training and obtaining stakeholder buy-in.

In a meta-analysis of 84 AI ethics documents published by different stakeholders from different geographic regions,<sup>18</sup> researchers found that no single principle appeared in all documents, but that five principles were referenced in more than half of them: responsibility, justice and fairness, transparency, non-maleficence and privacy. Table

15 Techatassanasoontorn, A. A., Waizenegger, L. and Doolin, B., op. cit., February 2023.

16 Lacity, M. C. and Willcocks, L. P., op. cit., June 2021.

17 For an overview of AI ethics principles, see: 1) Jobin, A., Ienca, M. and Vayena, E., op. cit., September 2019; and 2) Fjeld, J., Achten, N., Hilligoss, H., Nagy, A. and Srikumar, M., op. cit., January 15, 2020.

18 Jobin, A., Ienca, M. and Vayena, E., op. cit., September 2019.

**Table 1: AI Ethics Principles and Example Definitions**

AI Ethics Principle	Example Definitions	Related Terms and Concepts
<b>Responsibility</b>	According to UNESCO, “The ethical responsibility and liability for the decisions and actions based in any way on an AI system should always ultimately be attributable to AI actors corresponding to their role in the lifecycle of the AI system. Appropriate oversight, impact assessment, audit and due diligence mechanisms ... should be developed to ensure accountability for AI systems and their impact throughout their lifecycle.” <sup>19</sup>	<ul style="list-style-type: none"> <li>• Accountability</li> <li>• Liability</li> <li>• Integrity</li> </ul>
<b>Justice and Fairness</b>	According to the European Commission: “The development, deployment and use of AI systems must be fair ... fairness has both a substantive and a procedural dimension. The substantive dimension implies a commitment to ensuring equal and just distribution of both benefits and costs and ensuring that individuals and groups are free from unfair bias, discrimination and stigmatization. ... The procedural dimension of fairness entails the ability to contest and seek effective redress against decisions made by AI systems and by the humans operating them.” <sup>20</sup>	<ul style="list-style-type: none"> <li>• Inclusion</li> <li>• Prevention of bias</li> <li>• Non-discrimination</li> <li>• Equality</li> </ul>
<b>Transparency</b>	The OECD recommends that AI actors “should provide meaningful information, appropriate to the context ... to foster a general understanding of AI systems, to make stakeholders aware of their interactions with AI systems, including in the workplace, to enable those affected by an AI system to understand the outcome, and to enable those adversely affected by an AI system to challenge its outcome.” <sup>21</sup>	<ul style="list-style-type: none"> <li>• Explainability</li> <li>• Interpretability</li> <li>• Disclosure</li> <li>• Right to information</li> </ul>
<b>Non-Maleficence</b>	According to the United Nations, “AI systems should not be used in ways that cause or exacerbate harm, whether individual or collective, and including harm to social, cultural, economic, natural and political environments.” <sup>22</sup>	<ul style="list-style-type: none"> <li>• Harm</li> <li>• Security/safety</li> <li>• Protection</li> </ul>
<b>Privacy</b>	The Future of Life Institute recommends that “people should have the right to access, manage and control the data they generate, given AI systems’ power to analyze and utilize that data.” <sup>23</sup>	<ul style="list-style-type: none"> <li>• Data usage control</li> <li>• Consent</li> <li>• Personal information</li> </ul>

*Note:* Because there is no universally agreed-upon set of definitions for AI ethics principles, this table contains examples of definitions from prominent sources. In the definitions, AI actors include all those who play an active role in the lifecycle of an AI system, including organizations and individuals that develop, deploy or operate AI.



1 provides example definitions for each of the five principles.

Despite the emerging global convergence across stakeholders on the importance of these five principles, little agreement exists on how they are defined or how they should be implemented. For example, to increase transparency, a common recommendation is that those developing or implementing AI should disclose information, but the definitions of what should be communicated, to whom and how can vary greatly.<sup>24</sup> As a consequence, there is a large gap between the high-level principles (“what”) and organizational practices that can readily be put into action (“how”). Unfortunately, this gap is widened by the complexity, variability, subjectivity and lack of standardization of these principles.<sup>25</sup> Organizations therefore struggle to translate the AI ethics principles into practices that would help them address the AI-related ethical challenges they face.

## Overview of the Implementation of Robotic Process Automation at ESP

Our case study is of a German energy service provider, referred to anonymously as ESP. It acts as the shared service organization of one of the largest energy companies in the German stock market. The majority of ESP’s 280 employees handle back-office processes related to metering, billing and customer service for more than

700,000 customers at approximately 900,000 metering points.

ESP and its parent company’s strategy is to transform the company into a zero-carbon business. Core elements of this strategy for all business areas are a “safe working environment,” “accountability,” “diversity” and “learning and developing.” The works council has a strong position in the company and consistently strives to address employees’ needs and concerns. In addition, ESP fosters a culture of compliance and offers a whistleblower hotline to report wrongdoings or unethical behavior. Moreover, ESP operates in a heavily regulated energy sector that legally requires a high level of safety and security to ensure a stable energy supply and the protection of critical infrastructure.

Against this backdrop, any implementation of AI at ESP must not only consider technical challenges (e.g., technical feasibility) and business outcomes (e.g., implementation costs), but must also meet the legal requirements of the energy sector and address possible concerns of those affected by AI (e.g., back-office employees) and their representatives in the works council.

Like many other companies in the energy sector, ESP and its parent company are under significant pressure to cut costs, because they operate in a commodity business with low margins and intense competition while simultaneously having to make massive investments in the transition to climate neutrality. In addition, with many employees approaching retirement, ESP faces significant challenges in filling job vacancies.

For these reasons, ESP decided in 2019 to evaluate the feasibility of using RPA to automate routine back-office processes in the areas of billing and customer service. The technology choice was made based on reports of other companies that had successfully implemented RPA and ESP’s need for a technology that could manage its high volume of identical or similar back-office transactions within tight timeframes to fulfill service-level agreements. From the start of its RPA implementation journey, ESP made substantial efforts to listen to the needs of its back-office employees whose processes or tasks were going to be automated. Rather than cutting jobs, the main goal was to enable these employees

19 From *Recommendation on the Ethics of Artificial Intelligence*, UNESCO, 2021, available at <https://unesdoc.unesco.org/ark:/48223/pf0000380455>.

20 From *Ethics Guidelines for Trustworthy AI*, op. cit., April 8, 2019.

21 From *Recommendation of the Council on Artificial Intelligence*, OECD, May 2019, available at <https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449>.

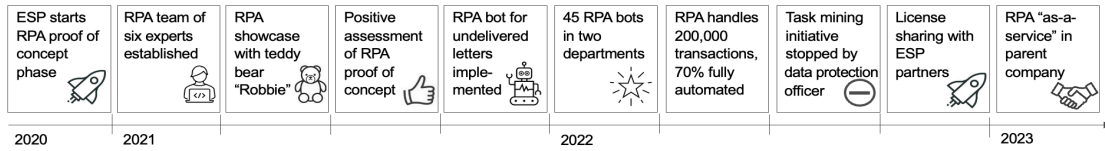
22 From *Principles for the Ethical Use of Artificial Intelligence in the United Nations System*, Chief Executives Board for Coordination, United Nations System, October 27, 2022, available at [https://unseeb.org/sites/default/files/2023-03/CEB\\_2022\\_2\\_Add.1%20%28AI%20ethics%20principles%29.pdf](https://unseeb.org/sites/default/files/2023-03/CEB_2022_2_Add.1%20%28AI%20ethics%20principles%29.pdf).

23 From *Asilomar AI Principles*, Future of Life Institute, August 11, 2017, available at <https://futureoflife.org/open-letter/ai-principles/>.

24 Ibid.

25 Morley, J., Floridi, L., Kinsey, L. and Elhalal, A. “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), August 2020, pp. 2141-2168.

**Figure 1: Timeline of Robotic Process Automation Implementation at ESP**



to undertake more complex tasks in their current department through upskilling or reskilling.

The RPA implementation at ESP started with a small agile team consisting of a process manager, a “key user” with a customer service background and external consultants responsible for programming the RPA bots using the software provided by UiPath.<sup>26</sup> The implementation followed a collaborative approach in the initial exploration of possible RPA use cases by involving all stakeholders, including back-office employees, leaders, and members of the IT department. After analyzing the potential time savings of each use case, a first proof of concept was developed. Insights gleaned from the proof of concept formed the basis for the design, development, introduction and evaluation of each RPA bot. The timeline of ESP’s RPA implementation journey is depicted in Figure 1.

Overall, more than 45 RPA bots were implemented at ESP. In total, these bots successfully processed more than 200,000 back-office transactions per year. Of these, 70% were handled end to end by RPA without any human involvement, while the remaining 30% were preprocessed before being forwarded to human employees, thus further reducing the manual workload. Given that each transaction would have taken a human employee between three and five minutes to complete, the total working time saved was more than 9,300 hours.

ESP used the freed-up capacity to redeploy employees to other more value-adding tasks. For example, a back-office employee who had previously focused on meter reading feasibility

checks was retrained to handle customer calls involving personalized consultations (e.g., addressing invoice issues or discussing energy-saving measures) and cross-selling activities (e.g., promoting new gas or electricity tariffs). In another example, a key user in the customer service team who viewed RPA as an opportunity to advance her career participated in training on the RPA software and implementation process. Subsequently, she was promoted to a process manager role and took on the responsibility of supervising the bots’ daily operations and managing the exceptions generated by them. From a business perspective, the RPA implementation was considered a success because it enabled ESP to handle processes for external customers much faster. In addition, it served as a blueprint for RPA initiatives in other parts of the parent company (“RPA as a service”), thereby helping to extend ESP’s portfolio for internal customers.

One of the most successful RPA bots implemented at ESP was designed to automatically process undelivered letters (i.e., scanned physical letters that could not be delivered via mail to customers and were returned to ESP). This bot operated according to the following four steps: 1) log into the ERP system, go to the list of undelivered letters, read the corresponding information of a letter and retrieve the customer ID of the letter’s addressee using optical character recognition; 2) log into the customer relationship management system and check if the customer reported a change of address; 3) if yes, put the new address on the letter and resend it; if no, send a request to the local resident registration office and leave the case open with a deadline to check if a new address was received; and 4) if no new address is received within a defined number of days, send an

<sup>26</sup> UiPath (<https://www.uipath.com>) is one of the top vendors of RPA software and is consistently named as a leader in the RPA market. See, for example, *Gartner Magic Quadrant for Robotic Process Automation*, Gartner, Inc., available at <https://www.gartner.com/en/documents/4595599>.

**Table 2: Summary of Issues Faced by ESP for the Five AI Ethics Principles**

AI Ethics Principle	Issues that Arose at ESP
<b>Responsibility</b>	<ul style="list-style-type: none"> <li>• Who takes (personal) responsibility for the actions of a particular RPA bot? Who is in charge of the overall RPA implementation?</li> <li>• How should the workload be distributed to employees when a bot makes mistakes due to programming errors or a bot refers a case for further processing?</li> <li>• Are leaders higher up in the hierarchy aware of the real impact of RPA on employees and do they assume responsibility for negative consequences themselves?</li> </ul>
<b>Justice and Fairness</b>	<ul style="list-style-type: none"> <li>• Who can suggest ideas for potential processes to be automated through RPA? Do all suggestions receive equal consideration?</li> <li>• Do only certain employees enjoy the benefits of RPA (e.g., being freed from “boring” tasks) without being expected to take on more complex tasks and having to learn something new?</li> <li>• How can the company treat all employees with the same respect, regardless of their differing levels of motivation and enthusiasm toward the opportunities of RPA?</li> </ul>
<b>Transparency</b>	<ul style="list-style-type: none"> <li>• How should the company address the uncertainty among employees and works council members about what RPA is and how it might affect work?</li> <li>• Who should communicate changes (e.g., new RPA bots)? What should be communicated and to whom?</li> <li>• How should employees be involved in the process of implementing new bots?</li> </ul>
<b>Non-Maleficence</b>	<ul style="list-style-type: none"> <li>• How do employees cope with the fear of possible job losses? Do they believe the promise that there will be no layoffs?</li> <li>• Does RPA relieve more stress (e.g., by reducing the workload during peak times) than it creates (e.g., by disrupting established work routines)?</li> <li>• How does the implementation of RPA affect other stakeholders, such as customers who value the personal interaction with a human employee?</li> </ul>
<b>Privacy</b>	<ul style="list-style-type: none"> <li>• Are there any data privacy issues that may have been overlooked, particularly in the context of using task mining to generate ideas for new RPA bots?</li> <li>• What does the company need to do to comply with local data privacy laws—e.g., the European Union’s General Data Protection Regulation (GDPR)?</li> </ul>

email to a human employee to manually look up the address and process the letter.

When the idea of processing undelivered letters using RPA was discussed in early 2021, employees were skeptical. For example, one employee said: “It would be nice if it worked. But I don’t think in reality something like that works.” However, a year later, the automation of this process turned out to be the most successful RPA bot at ESP. More than 65% of undelivered letters were handled without human involvement, freeing up substantial time for employees on the customer service team. As one team member commented: “Before the use of the robot, we always had ... 7,000 to 8,000 undelivered letters to take care of and now we are at 2,500 to 3,000 letters.”

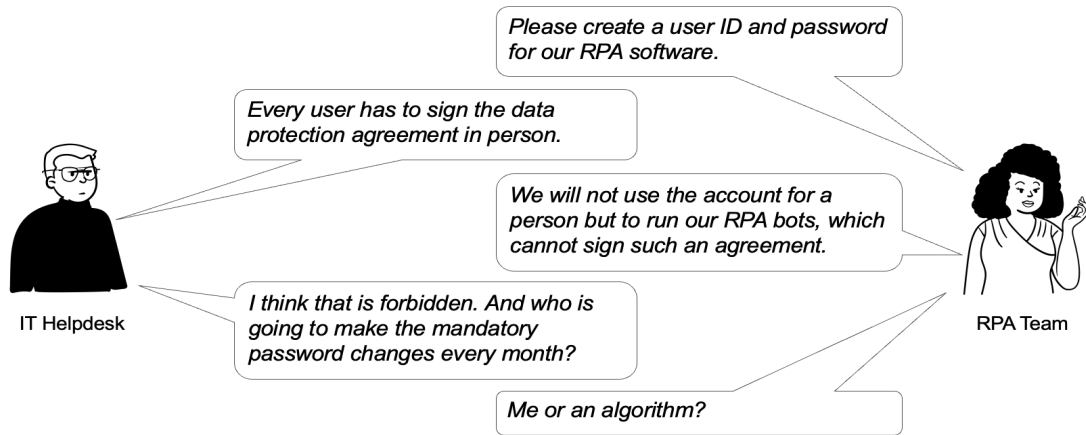
This example illustrates that, despite employees’ initial skepticism, an RPA bot was able to handle a sizable portion of the transaction volume of a process that back-office employees believed to be too difficult and complex for a bot. Though ESP committed itself to no layoffs, the success of this bot did cause uncertainty among employees about the possible long-term consequences of implementing RPA at ESP.

### How ESP Addressed Ethical Issues in its RPA Journey

Throughout its RPA implementation journey, ESP encountered ethical challenges on multiple fronts as it made significant efforts to respect the rights and interests of employees impacted by the introduction of RPA technology. Our



**Figure 2: Discussion about Responsibility Between the IT Helpdesk and RPA Team**



primary aim in studying ESP was to better understand these efforts and their impact on employees' attitudes toward RPA. Following our 30-month period of interviewing, surveying and observing ESP employees, we analyzed the data collected through the lens of five common AI ethics principles: *responsibility, justice and fairness, transparency, non-maleficence and privacy*. We used this theoretical framing because employees' comments in a company-wide survey, conducted in 2022, revealed that they were often concerned about issues relating to responsibility (e.g., for the negative consequences of the RPA implementation), justice and fairness (e.g., fair distribution of RPA benefits) and transparency (e.g., uncertainty about what RPA is and how it might affect their work). However, they appeared to be less concerned about non-maleficence (e.g., the fear of job losses) and privacy (e.g., the protection of private data). The issues that arose at ESP during its RPA implementation journey for each of the five AI ethics principles are summarized in Table 2 and described in detail below.

### Issues Concerned with Responsibility

Discussions about *responsibility* started at the very beginning of ESP's RPA implementation journey, even before the first prototype was developed. Initially, the IT helpdesk refused to create a user account for the RPA bots, arguing that a machine cannot possibly fulfill the same

responsibilities as a human employee. This lively discussion between the IT helpdesk and members of the RPA team, depicted in Figure 2, raised broader questions about who would take (personal) responsibility for the actions of a particular RPA bot or the RPA implementation in general and about how these responsibilities could manifest in the organization.<sup>27</sup>

Similar discussions took place after an RPA bot was tested in a live environment and made several serious mistakes,<sup>28</sup> resulting in complaints from back-office employees: "... there was a test run. And there were just too many errors that came up, which we then had to deal with afterwards. We ... had more work than before!" Though back-office employees did not program the bot, they ended up paying for the errors in its design by manually having to go through and correcting all the transactions it had touched. Similarly, leaders were concerned not only about the immediate consequences but also about potential damage to the company. For example, a leader commented: "We have a few hundred

27 Concerns about responsibility for the outcomes generated by "black box" AI are common. For an example in the healthcare context, see van Giffen, B. and Ludwig, H. "How Siemens Democratized Artificial Intelligence," *MIS Quarterly Executive* (22:1), March 2023, pp. 1-21.

28 It is important to note that RPA bots do not make "mistakes" in the traditional sense. However, they can produce incorrect outputs due to errors or flaws in their programming or design. At ESP, such errors predominantly stemmed from misunderstandings about the details of a specific process.

thousand customers, [and] it can end badly [if the bot doesn't work properly]. So, the fear is that [wrong] payments go out. Or that our own CEO receives an incorrect invoice.”

Against this backdrop, a senior process manager was asked to be responsible for all RPA bots, including their performance, actions and logins, and to report directly to the head of the customer service department to ensure rapid decision-making if needed. This process manager had the power and responsibility to start an RPA bot or stop it if it produced too many errors. For example, he had to stop a bot and adjust its instructions when the German tax rate on gas changed. Otherwise, the bot would have used incorrect templates, resulting in excessive manual work for back-office employees who would have had to fix the mistakes.

As the number of RPA bots continued to increase, a key user from the customer service team assumed this “bot supervisor” role<sup>29</sup> and was later given the nickname “Robbie’s mum,”<sup>30</sup> highlighting her responsibility for taking care of the bots. Being a member of the customer service team, it was easier for her to distribute the workload to colleagues when a bot made mistakes or referred cases for further processing.

In addition, back-office employees also demanded that leaders higher up in the hierarchy take responsibility for the RPA implementation: “Our leaders ... must be aware of the responsibility of what automation means for employees. They should lead the initiatives and encourage employees to get in touch with the RPA team.” This quote from a company-wide survey showed that responsibility should not be limited to the RPA team. In a similar vein, one ESP leader wanted his peers to engage more actively with employees impacted by the RPA implementation: “I can only say, take on your responsibility as a leader. ... You have to make time for this topic, you have to listen to many conversations. Above all, you must show interest as a leader yourself.”

29 For in-depth coverage of the different roles in an RPA implementation project, see Kokina, J., Gilleran, R., Blanchette, S. and Stoddard, D. “Accountant as Digital Innovator: Roles and Competencies in the Age of Automation,” *Accounting Horizons* (35:1), August 2021, pp. 153-184.

30 “Robbie” was the name given to the giant teddy bear placed in front of the initial RPA workstation. Research shows that anthropomorphizing RPA bots is a common phenomenon—e.g., “Robbie missed his mum.” For more information, see Techatassanasoontorn, A. A., Waizenegger, L. and Doolin, B. op, cit., February 2023.

In general, back-office employees understood the need for automation, but they wanted senior management not just to follow a trend but also to listen to the concerns of affected employees. The company-wide survey revealed that about a fifth of ESP’s employees believed that their leaders both underestimated the risks associated with RPA and lacked the competencies necessary for its successful implementation.<sup>31</sup> Employees also challenged their leaders’ assumption that everyone wants to be freed from routine tasks so that they can focus on more complex and meaningful work. Some employees really liked their routine tasks, while others simply did not want to learn something new (especially those nearing retirement). For leaders, responsibility therefore meant being aware of the real impact of RPA on their lower-level employees and assuming responsibility for negative consequences themselves, rather than delegating them to the RPA team.

## Issues Concerned with Justice and Fairness

The AI ethics principle of *justice and fairness* typically focuses on addressing issues of bias and discrimination, though these issues did not arise during the RPA implementation at ESP. However, this principle also emphasizes the importance of fair access to AI (procedural justice) and the fair distribution of its benefits and costs (distributive justice). During early discussions at ESP, we recognized the importance of granting an equal opportunity for all employees to make suggestions about possible processes to be automated using RPA. Several back-office employees believed that their suggestions were not considered because they were not members of the RPA team tasked with making decisions about possible RPA use cases. As a result, they experienced a sense of exclusion and felt that they were being treated unfairly because their colleagues benefitted from RPA while they still had to do all the “boring” work.

Other employees complained when a new RPA bot was introduced in their area of responsibility because they realized that taking on other,

31 For a discussion of why AI implementation projects demand new kinds of expertise and skills, see Someh, I., Wixom, B. H., Beath, C. M. and Zutavern, A. “Building an Artificial Intelligence Explanation Capability,” *MIS Quarterly Executive* (21:2), June 2022, pp. 143-163.

more complex tasks would require upskilling and reskilling. For example, one employee explained that he would have liked to keep some of the simple tasks that were previously his responsibility: “In the past, when I still worked in telephony, tasks like changes in advance payments and meter readings were exactly the ones that were quite enjoyable because you could do them ‘effortlessly,’ without having to think too much.” In his view, fairness meant that everyone should give up some of their current tasks and learn something new. Overall, these findings show that perceptions of fairness differed considerably among employees and were strongly influenced by their firsthand experiences with RPA’s benefits and costs.

Another fairness issue that arose during the RPA implementation was the difficulty of bringing together a diverse workforce with different sociodemographic backgrounds and levels of technical expertise. One leader described how several employees were confident that their own tasks would not be taken over by a bot before they retired: “Some of them are older, more established and of course, they thought at the beginning that only we [humans] can do that, no robot can do that.” As a result, when their tasks were taken over by an RPA bot, they ended up losing face.

In contrast, other employees with little technical expertise had unrealistically high expectations that led to a fear of job losses and resistance to any involvement in RPA implementation. On the other hand, many younger, more tech-savvy employees quickly understood what RPA would be capable of and how they could use its implementation as an opportunity to accelerate their careers. Though both groups generally respected each other’s attitudes toward RPA, negative age stereotypes about older employees occasionally led to tensions between them.

### Issues Concerned with Transparency

*Transparency* emerged as one of the most important AI ethics principles during ESP’s RPA implementation journey. Initially, there was a great deal of uncertainty among back-office employees and leaders about what RPA is and how it might affect their work. They also questioned whether RPA bots could perform

tasks correctly and faster than humans. The RPA team tried to answer all of these questions but realized that this not only required a lot of effort but sometimes also created more confusion than clarity.

To provide greater clarity, the RPA team decided that everyone should be able to see for themselves how a live RPA bot performed its work. They set up the RPA software on a normal desktop PC placed in one of the offices and invited employees and works council members to observe its functioning at any time. However, on several occasions, someone switched off the PC and thereby accidentally stopped the RPA bot, thinking that a colleague had forgotten to turn off the computer before leaving. The RPA team addressed this issue by placing a giant teddy bear on the chair in front of the desktop PC to remind everyone that it should not be switched off (see Figure 3). The teddy bear, subsequently named “Robbie,” often attracted the attention of passing employees and helped to convey the impression that RPA implementation was not happening behind their backs. One leader explained that the teddy bear changed the perception of RPA from an army of invisible workers to a “support [tool] for trivial things that are no longer a burden.” In addition to increased transparency, the teddy bear not only reduced employees’ uncertainty and works council members’ concerns but also turned fear into curiosity.

As with any IT implementation project, open and transparent communication is key to creating a trusting atmosphere when implementing RPA.<sup>32</sup> Insights from the company-wide survey showed that 81% of ESP’s employees described themselves as knowing little about RPA but interested in receiving more information in the future. Indeed, 59% registered an interest in becoming actively involved in the RPA initiative. However, questions remained about who should communicate, what should be communicated and to whom. One leader also pointed out that it was often challenging to reach everyone in a timely manner: “In the hierarchy, [the communication] goes downwards, and of course, something might get lost along the way, but we have always tried

32 Lacity, M., Willcocks, L. and Gozman, D. “Influencing Information Systems Practice: The Action Principles Approach Applied to Robotic Process and Cognitive Automation,” *Journal of Information Technology* (36:3), September 2021, pp. 216-240.

**Figure 3: Teddy Bear “Robbie” Sitting in Front of the Initial RPA Workstation**



to inform everyone. Unfortunately, sometimes it turns out that we haven't completely succeeded, ... even if it's just that someone who finds the topic [of RPA] really exciting was on vacation when it was discussed for the first time."

Though employees appreciated the town hall meetings and top-down communication of senior management (e.g., on the topic of job security), they felt that face-to-face meetings with their colleagues and team leaders, who explained both positive and negative consequences to them, were more helpful. For example, one back-office employee recommended: "Involve all employees in the process and [don't] just present the end product. This will make it easier for many to understand the function of RPA." The RPA team therefore encouraged lower-level employees who participated in one of the pilot projects to openly share their own RPA experiences with their colleagues. These informal face-to-face meetings helped other employees ask questions without feeling judged and become involved in the RPA implementation—for example, by identifying new RPA use cases in their own processes.

Overall, the RPA team and senior management were satisfied with their communication efforts.

Nevertheless, the company-wide survey also showed that not everyone shared this view. For example, one employee requested that ESP leaders make sure that "everyone is introduced to the new [RPA] system and not [just] the people ... sitting at the PC." This suggestion prompted ESP's leaders to reconsider how employees who would not be affected by RPA, such as mechanics or field representatives, should be informed about the implementation of RPA.

### Issues Concerned with Non-Maleficence

The AI ethics principle of *non-maleficence* emphasizes the avoidance of harm.<sup>33</sup> In contrast to physical robots, software robots, such as RPA bots, cannot directly cause or exacerbate physical harm.<sup>34</sup> However, issues related to psychological

<sup>33</sup> "Do no harm" is a fundamental principle of many ethical and philosophical systems. For a more comprehensive coverage of this principle, based on business ethics and institutional economics, see Suchanek, A. and Entschew, E. M. "Ethical Focal Points as a Complement to Accelerated Social Change," *Humanistic Management Journal* (3), November 2018, pp. 221-232.

<sup>34</sup> However, software robots could cause physical harm indirectly, for example by exhibiting discriminatory behavior in loan decisions. See Huang, J. Y., Gupta, A. and Youn, M. "Survey of EU Ethical Guidelines for Commercial AI: Case Studies in Financial Services," *AI and Ethics* (1:4), March 2021, pp. 569-577.



or emotional harm did arise during the RPA implementation at ESP. Not surprisingly, coping with the fear of possible job losses affected back-office employees' well-being, as mentioned by one employee: "I think the fear of being rationalized away was very strong at the beginning. And I believe the emotions are still mixed." These issues and concerns arose despite leaders continuously repeating that there would be no layoffs, as stressed by one of them: "[RPA] will definitely be [about] relieving workload ... staff reduction is not our goal. Rather, our goal is to free employees from mundane processes ... and to enable them to have more time for valuable tasks." However, despite ESP making it clear from the outset that no jobs were at risk and even struggled to fill vacant positions, a few doubts apparently remained at the back of employees' minds.

The ability of RPA to free employees from routine tasks, thus allowing them more time for complex, meaningful work, was both a relief and a source of stress. On the one hand, employees felt that the RPA bots helped them deal with the large volume of work, thereby preventing stress during peak times in particular. One employee declared: "I'm a friend of this technology because it gives me more time for things that need deeper investigation." One leader reiterated this sentiment: "I haven't heard anything negative. On the contrary: where it works, our employees are happy about the elimination of annoying, tedious work." Thus, ESP's RPA implementation helped to reduce psychological harm and supported mental well-being.

On the other hand, employees were stressed by disruptions to their established work routines and by the need to reskill or upskill when accepting unknown and new tasks. The possible impact of these disruptions caused concerns among members of the works council, as one of them explained: "If you suddenly come to [employees] with new robots and tell them, 'we're changing all processes,' the world changes overnight for them. If you don't prepare them, involve them, and also address their concerns and worries in advance, they might shut down and possibly become ill."

Not surprisingly, 53% of ESP employees who responded to the company-wide survey reported a desire for additional training to better cope with RPA. Some even used the survey to express

the view that the company should take their concerns more seriously. One said: "[ESP should] address employee fears/concerns honestly, not just [provide] platitudes, [but] a survey like this is ... a good start." Though leaders acknowledged the concerns about needing to learn new skills, they also emphasized that there was really no alternative: "That's simply the change of time, which cannot be stopped, and, from my perspective, shouldn't be stopped either. There will always be opportunities to engage people and also to qualify them in a way that they can take on other tasks."

Another issue related to the AI ethics principle of non-maleficence arose from realizing that the RPA implementation would not only affect ESP's employees but also the company's customers. For example, one leader mentioned that the automated meter reading entry process triggered when a customer contacted customer service via phone was not universally appreciated: "There are customers of a certain age who mostly do not find it great ... because they just want confirmation from a human."

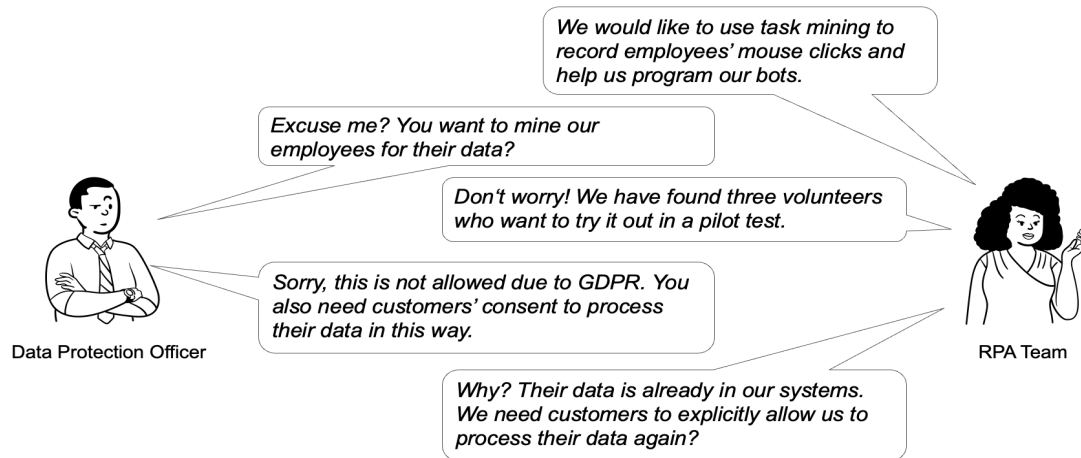
However, many of ESP's RPA bots were designed to reduce the need for human interactions, such as customer service calls between employees and customers. As a result, customers who preferred to talk to a human employee might have been unable to do so. One employee described how some of her customers deliberately visited ESP's local customer service center because they appreciated the personal interaction with an employee. Though ESP did not have an answer to this problem, it raised the awareness that implementing RPA might have unexpected side effects inside and outside the organization that are easy to overlook.

### Issues Concerned with Privacy

The protection of people's *privacy* is another key AI ethics principle. During our study of ESP's RPA implantation journey, we realized that the works council and data protection officer had an acute awareness of privacy issues, as the following example illustrates. To reduce the effort of finding processes suitable for RPA, the RPA team wanted to introduce a technique called "task mining," which provides in-depth insights into how employees perform their tasks by recording their actions and collecting user interaction data.



**Figure 4: Discussion About Task Mining Between the Data Protection Officer and the RPA Team**



These insights are then used to automatically identify work routines suitable for RPA and to provide ideas for programming the bots.<sup>35</sup>

Three back-office employees volunteered to try out task mining on their computers. Even though they could stop the recording at any time without having to provide an explanation or face consequences, the data protection officer ultimately did not agree to the pilot test because of concerns about privacy risks. The RPA team was puzzled when it discovered that these risks did not stem from the recording of employee behavior but from the (indirect) processing of customer data during the task mining phase—the data protection officer stated that customers needed to consent to their data being processed for this specific purpose (The discussion between the RPA team and the data protection officer is depicted in Figure 4.).

As a result, the task mining initiative was stopped before it even started, causing confusion among the volunteers and frustration in RPA team, which believed that task mining would have made the team members' lives much easier. This example demonstrates the power of data protection officers and highlights

the importance of their early involvement in discussing and resolving privacy issues during RPA implementation.

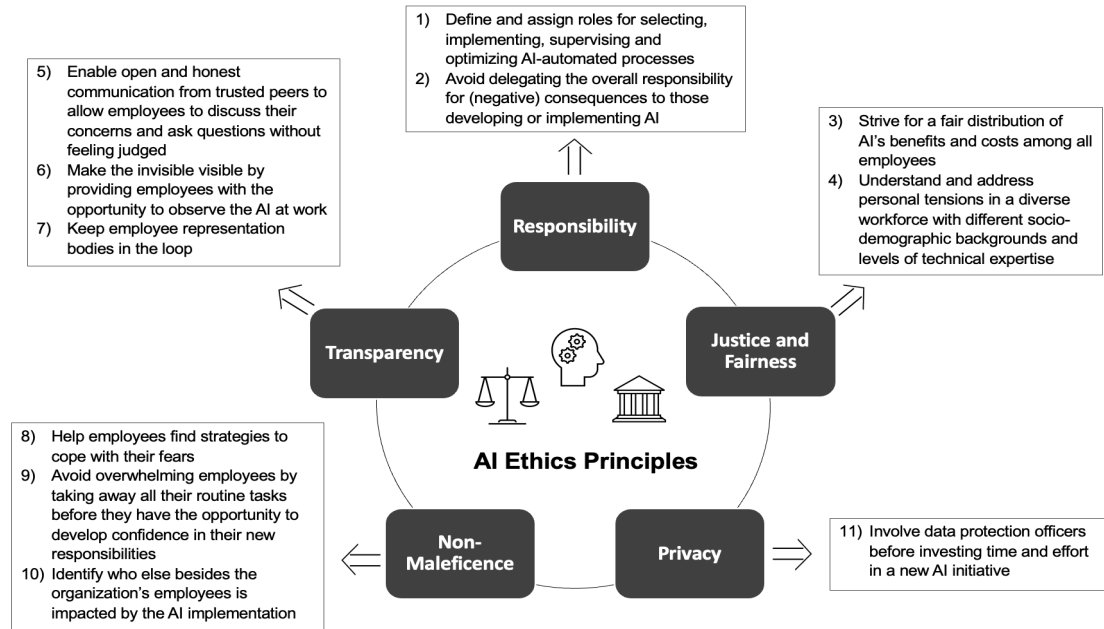
Apart from this particular task mining issue, the AI ethics principle of privacy seemed to be less of a concern during RPA implementation at ESP. Because ESP operates in a country with strong data privacy laws—i.e., The European Union's General Data Protection Regulation (GDPR)—protecting the privacy of customers and employees has always been a key priority. To ensure GDPR compliance, the RPA software was installed on-premise and customer data remained on local servers. Moreover, none of the RPA-automated processes involved the processing of employee data.

## Recommendations for Translating AI Ethics Principles into Practice

Based on the ESP case, we provide 11 recommendations for translating AI ethics principles into practice—two for the *responsibility* principle, two for the *justice and fairness* principle, three for the *transparency* principle, three for the *non-maleficence* principle and one for the *privacy* principle. Though these recommendations were derived from the analysis of a large-scale RPA implementation at

<sup>35</sup> For more information, see Leno, V., Polyvyanyy, A., Dumas, M., La Rosa, M. and Maggi, F. M. "Robotic Process Mining: Vision and Challenges," *Business & Information Systems Engineering* (63:3), September 2020, pp. 301-314.

**Figure 5: Summary of Recommendations for Translating AI Ethics Principles into Practice**



an energy company, we are convinced that they are not only applicable to other industries but also to the implementation of more advanced AI applications that cause uncertainty and fear among employees. Our recommendations are summarized in Figure 5 and described in detail below.

### Addressing Issues Concerned with Responsibility

**Recommendation 1: Define and assign roles for selecting, implementing, supervising and optimizing AI-automated processes.** Responsibility gaps are a common phenomenon when implementing AI in organizations. These gaps can be problematic because holding others responsible for actions and their consequences is essential for building a trusting environment. Leaders must therefore carefully define and assign roles for selecting, implementing, supervising and optimizing AI-automated processes. As the ESP case illustrates, these roles and the people who perform them will need to change over time as expertise grows and the organization's AI implementations become

more advanced. While only a few roles might be essential in the proof-of-concept phase, additional and more clearly defined roles and responsibilities become necessary when AI is rolled out on a large scale. Moreover, certain roles might be assigned to an experienced partner (e.g., an external consultant) prior to training in-house personnel.

**Recommendation 2: Avoid delegating the overall responsibility for (negative) consequences to those developing or implementing AI.** Another key recommendation related to the AI ethics principle of responsibility is that leaders higher up in the hierarchy should not delegate responsibility to those developing or implementing AI.<sup>36</sup> Instead, senior leaders should be aware of the real impact AI will have on their employees and accept full responsibility. Doing so will not only show employees that senior leadership takes their needs and concerns seriously but will also help the AI team to focus on the actual implementation rather than on

<sup>36</sup> For an in-depth discussion of the responsibility of AI developers, see Martin, K. "Designing Ethical Algorithms," *MIS Quarterly Executive* (18:2), June 2019, pp. 129-142.

broader discussions about possible consequences for employees.

### Addressing Issues Concerned with Justice and Fairness

**Recommendation 3: Strive for a fair distribution of AI's Benefits and costs among all employees.** The concepts of justice and fairness are highly subjective. Though the ESP case suggests that it might be difficult to please everyone, we recommend that leaders pay special attention to a fair distribution of AI's benefits and costs among all employees. All individuals, teams and departments should benefit from AI—for example, by being freed from “boring” repetitive tasks. At the same time, everyone should be required to learn something new (i.e., by upskilling and reskilling) when tasks are reallocated after the implementation of AI.

**Recommendation 4: Understand and address personal tensions in a diverse workforce with different sociodemographic backgrounds and levels of technical expertise.** AI implementations can create or exacerbate tensions in a diverse workforce characterized by varying backgrounds and levels of expertise. Though some employees might not (want to) engage with the new technology (e.g., those nearing retirement), others might be enthusiastic about it and see it as an opportunity to accelerate their careers. Leaders therefore need to ensure that they address these tensions and consider everyone's perspective, rather than focusing only on highly motivated employees.

### Addressing Issues Concerned with Transparency

**Recommendation 5: Enable open and honest communication from trusted peers to allow employees to discuss their concerns and ask questions without feeling judged.** Transparency from the very beginning is essential for a safe and secure working environment and can help avoid many misunderstandings when implementing AI. Communication, of course, plays a key role in increasing transparency. The ESP case suggests that, though employees generally seem to appreciate town hall meetings and top-down communication from senior management, face-to-face meetings with peers who have already gained some AI experience are

a more effective way of increasing transparency. We therefore recommend that leaders provide employees with a “safe space” where they can discuss concerns or fears with trusted peers and ask questions without feeling judged.

**Recommendation 6: Make the invisible visible by providing employees with the opportunity to observe the AI at work.** Another key aspect of (non)transparency relates to the uncertainty and lack of understanding of how the AI technology works. As indicated by the ESP case, simple remedies such as setting up a computer to allow everyone to watch the “robot” performing its work can foster transparency and turn uncertainty into curiosity. We therefore recommend that organizations create opportunities for employees to observe the AI at work—for example, in an area they often frequent.

**Recommendation 7: Keep employee representation bodies in the loop.** Our findings highlight the importance of providing transparency to all stakeholders who have a say in the AI implementation. For example, a works council can intervene if there are possible privacy risks for employees. We therefore recommend that from the outset, employee representation bodies such as works councils are kept in the loop.

### Addressing Issues Concerned with Non-Maleficence

**Recommendation 8: Help employees find strategies to cope with their fears.** While most AI implementations without physical robots are unlikely to cause any direct physical harm, possible psychological or emotional harm can easily be overlooked. The ESP case suggests that even when job security is guaranteed, the fear of possible job losses can remain at the back of employees' minds. Rather than trying to dismiss these (perhaps irrational) fears, leaders should help employees identify strategies to cope with their fears. For example, leaders could involve employees more directly in the AI implementation to show them how and where their expertise and skills are indispensable.

**Recommendation 9: Avoid overwhelming employees by taking away all their routine tasks before they have the opportunity to develop confidence in their new**

**responsibilities.** Though AI can reduce stress by giving employees more time for complex tasks, it can also increase stress and thus cause psychological harm because it disrupts existing work routines and forces employees to learn new tasks and assume additional responsibilities. It is therefore important to acknowledge that not every employee wants to be freed from routine tasks because such work offers a sense of predictability and security. We therefore recommend that leaders avoid overwhelming employees by removing all their routine tasks before they have had the opportunity to develop confidence in their new responsibilities. Instead, leaders should provide employees with the time and support needed to acquire the skills for more meaningful albeit more demanding tasks, allowing them to gradually transition into their new roles.

**Recommendation 10: Identify who else besides the organization's employees is impacted by the AI implementation.** Our findings underscore the need for a holistic evaluation of the impact of AI. In addition to the organization's workforce, various stakeholder groups, including customers and employees at partner companies, may experience direct or indirect impacts resulting from the AI implementation. We recommend that leaders identify who else besides their own employees is impacted by the AI implementation and in what way. This will enable leaders to assess possible unexpected side effects that might cause harm outside their own organization.

### Addressing Issues Concerned with Privacy

**Recommendation 11: Involve data protection officers before investing time and effort in a new AI initiative.** In contrast to other AI ethics principles, many countries have already introduced legislation to protect their citizens' privacy. As a result, it is less difficult to translate the AI ethics principle of privacy into practice because rules exist that regulate how organizations can collect, use and store data. The need to comply with the GDPR played a key role in ESP's decisions to deploy the AI technology on-premise, to keep customer data on local servers and to refrain from automating processes involving employee data. We believe that these

practices will also be valuable in countries lacking strict privacy laws.

However, our findings also show that opinions on privacy issues can differ between those responsible for ensuring privacy (e.g., data protection officers) and those developing or implementing AI. In the ESP case, for example, a promising AI initiative for which several employees volunteered to record all activities on their computers was stopped due to data privacy concerns. This caused frustration in the AI implementation team, which had already invested effort in exploring the technology and setting everything up. A key recommendation, therefore, is to involve data protection officers before launching a new AI initiative. Doing so will avoid wasting time and energy on initiatives that subsequently fall foul of privacy issues.

## Concluding Comments

Leveraging AI to automate processes previously performed by human employees presents organizations with many ethical challenges. As a result, leaders often have to deal with difficult issues, including responsibility (e.g., Who is responsible for the actions and possible errors in the programming of AI?), fairness (e.g., How can benefits and costs of AI be distributed fairly?), and transparency (e.g., How can invisible AI systems be made visible to employees?). To help leaders navigate the minefield of AI ethical challenges, we analyzed a large-scale RPA implementation journey through the lens of five common AI ethics principles and derived actionable recommendations for putting these abstract, high-level principles into practical use. Our recommendations highlight that leaders at all levels of the organization need to take employees' needs and concerns seriously, even before the first process is automated using AI.

RPA is a fairly basic form of AI, but as organizations adopt and deploy generative AI solutions, with their potential to revolutionize various industries and reshape the nature of work, they will inevitably face new and more complex ethical challenges, and the importance of AI ethics will only grow. Leaders who are involved in AI implementations can leverage our insights to not only maximize productivity and increase employee acceptance of AI but also to foster a



safe and secure working environment for all employees.

## Appendix: Research Method

This article is based on a longitudinal case study of the large-scale implementation of robotic process automation (RPA) at a German energy service provider, referred to anonymously as ESP. Our study was guided by the overall research question: How can leaders address ethical challenges that arise during the implementation of AI by seeking guidance from AI ethics principles? We followed a case study approach, rather than action research, because our emphasis was on observing, analyzing and interpreting ESP's efforts in addressing the ethical challenges instead of solving them immediately.<sup>37</sup>

Over a period of two and a half years, we collected data through interviews, a survey and observations of ESP employees in different roles at different levels of the organization. More specifically, we first conducted 19 semi-structured interviews with back-office employees, works council members, process owners, leaders and consultants when RPA was still in the proof-of-concept phase. The main goal of these interviews was to assess the first impressions of RPA and to identify possible challenges related to its introduction at ESP. After the RPA implementation had reached a more advanced stage, we conducted an online survey of 85 ESP employees, asking closed and open questions about their experiences with and involvement in the RPA implementation. Finally, we conducted a second round of 10 semi-structured interviews, primarily with leaders and works council members to capture their reflections on the overall RPA implementation journey and to discuss the role of AI ethics principles. In addition, one of the authors was extensively on-site at ESP for the entire RPA implementation journey.

We analyzed the data using an iterative approach through the lens of five common AI ethics principles: *responsibility, justice and fairness, transparency, non-maleficence and privacy*. Our focus on these five principles was

grounded in research by Jobin et al.,<sup>38</sup> who conducted a systematic meta-analysis of 84 AI ethics documents published by government agencies, private companies and research institutions from different geographic regions. They found that these five principles were referenced in more than half of all documents, indicating an emerging global convergence across stakeholders on their importance. Finally, we shared our findings with ESP and presented them at a practitioner event to collect feedback and verify our recommendations.

## About the Authors

### Dörte Schulte-Derne

Dörte Schulte-Derne (doerte.schulte-derne@hhl.de) is a part-time Ph.D. student at HHL Leipzig Graduate School of Management, Germany, supervised by Professor Dr. Claudia Lehmann, who holds the Chair of Digital Innovation in Service Industries at HHL. Dörte is also managing director of WVV GmbH in Würzburg, Germany. In addition to studying political science, public law and business, she has 20 years of business experience, most recently leading sustainable transformations in the energy and transport sectors, where digital transformation acts as a key enabler. Since 2020, she has researched how leaders can foster digital transformation, especially through RPA or AI implementations.

### Ulrich Gnewuch

Ulrich Gnewuch (ulrich.gnewuch@uni-passau.de) is a professor of information systems at the University of Passau, Germany, where he holds the Chair of Explainable AI-based Business Information Systems. His research focuses on the design, use and impact of AI in business and society. His industry experience includes managing and developing customer service chatbots and RPA solutions for energy companies.

<sup>37</sup> Myers, M. D. *Qualitative Research in Business & Management*, Sage Publications, 2020.

<sup>38</sup> For details, see Jobin, A., Ienca, M. and Vayena, E., op. cit., September 2019.