DEFINING AND IMPLEMENTING ALGORITHM REGISTERS: AN ORGANIZATIONAL PERSPECTIVE

Completed Research Paper

Martijn van Vliet, Utrecht University, Utrecht, Netherlands, m.vanvliet@uu.nl
Nena Schuitemaker, Utrecht University, Utrecht, Netherlands, n.schuitemaker@students.uu.nl
Sergio España, Utrecht University, Utrecht, Netherlands, and Universitat Politècnica de València, Valencia, Spain, s.espana@uu.nl
Inge van de Weerd, Utrecht University, Utrecht, Netherlands, g.c.vandeweerd@uu.nl
Sjaak Brinkkemper, Utrecht University, Utrecht, Netherlands, s.brinkkemper@uu.nl

Abstract

The transformative impact of AI on organizations is becoming more apparent, as organizations are required to develop new capabilities to deal with the challenges that come with the implementation of AI. Recent developments within the legal field add to these challenges, as public organizations in The Netherlands will be required by mandate to publish information about their use of algorithms to the public. The lack of scientific attention, combined with the lack of governmental guidance, further complicates the implementation of algorithm registers for organizations. We conducted six case studies at public organizations to assess the internal impact of the algorithm register on these organizations. Based on these case studies, we propose a reference process for the implementation of algorithm registers. Additionally, we provide a refined definition on algorithm registers that more accurately reflects its possible contributions towards current challenges in the fields of AI Innovation, AI governance and responsible AI.

Keywords: Algorithm Register, AI Governance, Responsible AI, Organizational Change

1 Introduction

Artificial Intelligence (AI) is already significantly altering the way that consumers, industries, businesses, organizations and information systems (IS) act and operate (Akinsola et al. 2022; Collins et al. 2021; Girasa 2020; Jahić and Roitsch 2020). As part of the ongoing digital transformation of organizations, AI has spread to the public sector, as governments worldwide have committed major investments towards the research and development of AI-related technologies (Khatib et al. 2022; Wirtz, Weyerer and Geyer 2018). It is expected that AI-driven innovation will have a profound impact on public sector employees, citizens, and societies (Madan and Ashok, 2023). The increasing prevalence of AI systems within the public sector is already increasingly determining whether and how individuals have access to social goods and rights (Martin, 2018). Citizens are increasingly expressing their concern towards the influence of AI on their daily lives (Ingrams, Kaufmann and Jacobs, 2021).

Specific examples that justify these concerns are not hard to find. A childcare-benefits scandal caused by the use of a discriminative algorithm by the Dutch Tax Authorities wrongfully accused approximately 26,000 parents of fraud, government systems that were initially designed to help turned out to profile and discriminate the poor in the United States, and algorithms used for exam grading in the UK turned out to be biased against students with a lower socioeconomic status background, causing public outrage and legal action (Konaté and Pali, 2023; Wieringa, 2020; Eubanks, 2018; Kelly, 2021).

The increasing concerns of the public regarding the inherent susceptibility of the technology to privacy and ethical issues have resulted in a push for more regulations and ethics in AI (Akinsola et al., 2022). Most recently, the AI Safety Summit led to the signing of the Bletchley Declaration for AI Safety, where 28 countries agreed on international co-operation towards the management of challenges and risks of AI and to build policies for the latest technologies (UK Government, 2023). Secondly, the recently passed European AI act introduces, among others, prohibitions of certain technologies, declare different requirements for systems based on risk levels and facilitates social scrutiny through the means of a public database (Veale and Borgesius, 2021). Furthermore, the Dutch parliament passed a motion on the 28th of October 2021 to initiate the development of a publicly accessible national algorithm register. In the near future, it will become mandatory for public organizations in The Netherlands to publish information about their use of algorithms to the public through that register. Being compliant with legislations coming from such initiatives will be crucial for organizations if they want to use AI as a technology in the future.

In this research, we study the current state of affairs regarding the development, the implementation, and the organizational impact of an algorithm register. Our goals are to catch the current perspectives of organizations towards the concept of an algorithm register. We aim to discover what efforts these public organizations have undertaken towards the implementation of an algorithm register and we set out to capture what the contribution of the register is towards the realization of responsible AI principles in practice. We summarized our goals in the following three research questions: **(RQ1)** What is the definition of an algorithm register? **(RQ2)** How can an algorithm register be implemented within an organization? **(RQ3)** In what way does the algorithm register contribute towards the realization of responsible AI principles in practice? Our goal is to provide more guidance for organizations that are currently working on or are starting to work towards the realization of their algorithm register.

Using the achieved insights from this research, we propose the following definition for an algorithm register: An **algorithm register** *is a governance mechanism that allows organizations to be transparent and to provide accountability to society by providing an overview of **(1)** the documentation about algorithms, **(2)** the organization or organizational department responsible for their use, and **(3)** the goals pursued with their use.*

We present our findings of the multivocal literature study in the next section. In section 3, we elaborate on our research method. In section 4, we present the results of the case studies. Section 5 shows the core elements of the reference process that we created. In section 6, we discuss the results and in the last section, we draw the conclusions of our research questions and present options for future work.

2 Related Work

2.1 Organizational change to tackle emerging challenges

The alluring benefits that AI as a technology offers, such as the reduction in required production costs and time, increased performance and customer satisfaction, achieving more accurate predictions and decision making are important drivers for the adoption of AI in organizations (Cubric, 2020). Global and US focused surveys show that 91,5% of financial and life science organizations report investments in AI initiatives, illustrating that AI is currently one of the main driving forces behind the digital transformation of organizations (Benbya, Davenport and Pachidi, 2020; Siebel, 2019).

Organizations engage in digital transformation to keep up with digital developments in their industry and technological progress (Osmundsen, Iden and Bygstad, 2018). Innovation plays a crucial role for organizations, as their survival, growth and competitiveness depend on their ability to change (Sartori et al., 2018; Weick and Quinn, 1999). However, the contribution of a new technology towards economic growth can only be realized when and if the new technology is widely diffused and used (Hall and Khan, 2003). Rogers (1995) refers to diffusion as the process by which new innovations spread among members of a social system. Innovation and diffusion are multi-stage continuous processes, where organizations transform ideas into new or improved products, services, or processes (Baregheh, Rowley

and Sambrook, 2009). What we think of as a single innovation is often the result of a lengthy process involving many interrelated innovations (Fagerberg, 2006). The eventual rate of diffusion determines the pace of economic growth and the change of productivity in organizations (Hall and Khan, 2003).

The disruptive characteristics of AI are cause for challenges that act as barriers for the diffusion of the technology (Nagy, Schuessler and Dubinsky, 2016). A vast majority of all AI and Big Data projects fail to achieve significant positive impact and are therefore rarely deployed into widespread production (Benbya, Davenport and Pachidi, 2020; Weiner, 2022). The challenges encountered during the implementation of AI systems are classified as being of (1) organizational and managerial, (2) data, (3) political, legal and policy, (4) ethical, (5) technical, (6) social and (7) economical nature (Dwivedi et al., 2021). The causes for these challenges are attributed to factors such as AI initiatives being hard to implement with existing processes and systems, technological immaturity, lack of available knowledge and market expectations that are too high (Brock and Von Wangenheim, 2019). The encountered challenges lead to a higher chance of risky and unsuccessful projects (Reim, Åström and Eriksson, 2020). As a result, it is unlikely that the biggest opportunities that AI has to offer have been tapped already (Brynjolfsson and Mcafee 2017). Regardless of the existing challenges, organizations continue their innovating efforts as the alternative, the underuse of the technology, could become costlier due to opportunity costs (Floridi et al., 2018). Dealing with these challenges accordingly requires capabilities that organizations currently do not possess, creating the drive towards organizational change.

Current research efforts towards the implementation of AI are asymmetrical, as few researchers study the literature about AI from a managerial viewpoint (Kitsios and Kamariotou, 2021). Even though results show that both organizational and technological issues can cause project failure (Westenberger, Schuler and Schlegel, 2022), researchers tend to focus on technical aspects of AI (Alsheibani, Cheung and Messom, 2018). Even though there is a clear need for more research towards the impact of AI on organizational activities, boundaries goals, mechanisms, and processes (Holmström, 2022), managers are currently left with little support from academia when they aim to implement AI in their firms (Reim, Åström and Eriksson, 2020).

2.2 Applying trustworthy AI principles in practice

The scientific community is already aware of the large impact that AI is having on society and the daily lives of people (Floridi and Cowls, 2022). With this impact, AI raises the bar compared to previously established standards regarding trustworthy computing, therefore birthing the area of trustworthy AI (Wing, 2021). Trustworthy AI is called the reflection of the recognition that maintaining trust in AI is critical for ensuring its acceptance and successful adoption (Toreini et al. 2020).

The calls for research towards trustworthy AI resulted in a sharp increase in available documents each endorsing a set of principles for AI, originating from different governments, organizations and sectors combined (Floridi, 2019a; Fjeld et al., 2020). A mapping of the available corpus for principles and ethical AI showed 84 different documents, with 88% of them appearing after 2016 (Jobin and Ienca, 2019). Floridi and Cowls (2022) express concerns that all these different declarations generate inconsistency and are cause for confusion among stakeholders. This confusion provides the opportunity of a choice for stakeholders to only address the most appealing ones or for 'ethics bluewashing' where superficial measures are taken to make misleading claims in an effort to appear more trustworthy (Floridi and Cowls, 2022; Floridi, 2019b). To address these issues, Floridi and Cowls (2019b) created a framework for ethical AI by comparing six high-profile policy and principle initiatives. The framework consists of the four traditional bio-ethics principles of beneficence, non-maleficence, autonomy, and justice, complemented with the principle of explicability. The latter refers to both providing understanding on how the system works (transparency) and providing clarity about the entity that is responsible for the way the system works (accountability). The Fairness, Accountability and Transparency in Machine Learning initiative declared principles that are in line with the framework of Floridi and Cowls (2019), but added auditability to the mix, referring to the authorization of third parties to check and criticize the algorithm as an extra outside source of control (Spiegelhalter, 2020). A comprehensive study that included thirty-six prominent documents referencing trustworthy AI

principles resulted in the emergence of eight themes: (1) privacy, (2) accountability, (3) Safety and Security, (4) Transparency and Explainability, (5) Fairness and Non-Discrimination, (6) Human Control of Technology, (7) Professional Responsibility (8) Promotion of Human Values (Fjeld et al., 2020). These themes represent the normative core of trustworthy AI, as their results indicate the convergence of the discussions regarding the principles among the communities that are responsible for drafting them.

This convergence can generally be interpreted as a step in the right direction, but there is still criticism. McNamara, Smith and Murphy-Hill (2018b) showed that the effectiveness of guidelines or ethical codes is almost zero and that they do not influence the behavior of professionals in the software industry (Hagendorff 2020). Hagendorff (2020) stresses that initiatives have thus far largely produced vague, high-level principles and value statements, which promise to be action-guiding, but in practice provide few specific recommendations. Mittelstadt (2019) concludes on basis of the aforementioned arguments that AI development lacks proven methods to translate principles into practice. As a result, the added value of the AI ethics principles is questioned, with some scholars breaking them down into meaningless, isolated, and toothless principles (Munn, 2022). Vakkuri et al. (2019) explicitly call for tools and methodologies to help industry experts implement AI ethics in practice.

2.3 The call for governance mechanisms for AI

Corporate governance concerns internal and external checks and balances to ensure accountability towards all stakeholders of an organization, so that they act in a socially responsible way (Solomon, 2020). Legal compliance is an important part of this social responsibility, but governance can go further than what the law requires (Mäntymäki et al., 2022). Appropriate governance has a strong positive connection with firm performance (Bhagat and Bolton, 2008). Corporate governance has changed and evolved together with new technological innovations and trends, to adapt to new challenges and risks that come with innovation. The increasing presence and impact of information systems called for specific IT governance, while the ever-increasing presence and importance of data in organizations called for specific data governance (Khatri and Brown, 2010; Korac-Kakabadse and Kakabadse, 2001).

The transformative effect that AI is currently having on organizations means that the role of AI and decision-making algorithms within organizations continues to grow. Software components, software design and development, data management, and software system deployment are amongst the areas that are affected by the introduction of AI (Jahić and Roitsch, 2020). Furthermore, the many currently experienced challenges surrounding the implementation of AI allows for the chance for unwanted consequences of the use of the technology to occur. As a response to these developments, calls for appropriate AI governance are getting louder. Mäntymäki et al. (2022b) provide the following definition of AI Governance: "*a system of rules, practices, processes, and technological tools that are employed to ensure an organization's use of AI technologies aligns with the organization's strategies, objectives, and values; fulfills legal requirements; and meets principles of ethical AI followed by the organization.*"

Benbya, Davenport and Pachidi (2020) describe the area of management and governance mechanisms as one of the most important directions for organizations to focus on, in order to remove or reduce existing barriers that are preventing organizations from implementing the technology. Mäntymäki et al. (2022b) call for research that applies design science to complement their definition with practical descriptions of tools, processes, mechanisms, structures and frameworks.

2.4 The role of public registers

Public registers are referred to as lists or databases of (personal) information, to which the public has a right of access by statute (November, 2008). Examples of existing registers that are used to realize compliance with existing laws or as a measure to help prevent unwanted consequences from happening. are plenty. Electronic Tax registers adopted in countries worldwide, the European Clinical Trials Register and the Swedish register for the registration of prenatal babies exemplify their uses and prove their effectiveness as tools (Naibei and Siringi, 2011; Goldacre et al., 2018; Holmström et al., 2012).

A public register for (AI) algorithms, however, is a new phenomenon that has so far received little scientific interest. Floridi (2020) was the first to react to the presentation with the launch of the registers of the municipalities of Amsterdam and Helsinki and reiterates the conceptual description of what an algorithm register can potentially become (Haataja, Fliert and Pasi, 2020). Furthermore, Floridi (2020) projects the perceived benefits of the register from his perspective as contributing towards the normalization of AI, helping AI becoming a public service, contributing to the field of AI ethics and by achieving trust of the public through transparency. The statements of Floridi (2020) were met with criticism by Cath and Jansen (2021) and called the perceived benefits too early to assess. Cath and Jansen (2021) cite further concerns regarding the creation of an 'ethics theater', highlight concerns regarding the decontextualization of AI and encourage the critical interrogation of the presumed existing relationship between AI registers, governance, and accountability. Cammers-Goodwin and Van Stralen (2021) mention algorithm registers in their research towards transparent cities but refrain from going into further detail. Nouws, Janssen and Dobbe (2022) briefly mention some possible goals and purposes by referring to discussions held within a consortium of municipalities within The Netherlands, but only mention that the algorithm register is still in development. Murad and Hafrey (2021) conducted a single case study at the municipality of Amsterdam at a very early stage and described only the initial development process that took place there. We have to conclude that insufficient recent literature on the topic of algorithm registers exists, none of them address potential transformative internal impact of the implementation of an algorithm register on organizations.

3 Research Method

We performed theory-building research with the objective of formulating new propositions based on evidence drawn from observation (Dul and Hak, 2007). We constructed a method in accordance with the first cycle of design science, focusing on problem investigation, treatment design, and treatment validation (Wieringa, 2014).

3.1 Problem investigation

We conducted a multivocal literature study to include grey literature from sources such as government reports and whitepapers, on top of available academic papers to achieve a knowledge base that is as broad as possible. Furthermore, we attended workshops organized by the Dutch Ministry of the Interior and Kingdom Relations on algorithm registers, where representatives of multiple organizations congregated to share their experiences, expectations, and concerns towards algorithm registers. The on first-hand basis experienced discourse on the topic contributed towards the initialization of our research.

3.2 Treatment design

Our treatment design comprises the creation of the algorithm register definition and reference process. We conducted a multi-case study in which we conducted interviews and carried out a survey.

Case selection. By searching government websites, news items on algorithms, and attending workshops, we identified several organizations that are currently working on developing an algorithm register or that have already published in the national register. Of the six organizations that agreed to participate, three were actively working on the implementation while the others were in the orientation phase. Other organizations declined due to the sensitivity of the topic or because they had not started yet.

The organizations are all Dutch government organizations that vary significantly from each other in size and in the type of tasks they carry out. The people who participated in the interviews and survey varied a lot in their backgrounds, experience in IT, and their current roles within their organizations. All participants were in some way involved with the initiative towards the realization of an algorithm register in their organization. Some were very actively involved, such as from the perspective of a project manager, while the involvement from others was more from an advisory perspective. Table 1 provides an overview of organizations and interviewees.

Organization	Size (Fte)	Participant	Experience (years)	Role
O1	> 10,000	P1	7.5	Ethics advisor
O2	> 1,000	P2	-	Project manager innovation
O2	> 1,000	P3	2.5	Legal policy advisor
O3	> 10,000	P4	6.5	Product owner, algorithm expert
O4	< 1,000	P5	8.5	Policy advisor data science
O5	> 1,000	P6	7	Senior project manager data science
O6	< 1,000	P7	26	IT Lawyer

Table 1. Detailed demographic information of the participants.

Data collection. We used two techniques to gather data in our case study: a survey and interviews. Our survey addressed the eight themes of responsible AI principles defined by Fjeld et al. (2020). The participants were presented with the same five questions for each of the eight principles. Participants could answer through the means of a 5-point Likert scale, ranging from Strongly Disagree to Strongly Agree. For each question and discussed theme, the participant was provided the opportunity to add their reasoning and explain their motivation for their answers. The survey could either be completed by the participant beforehand in an online environment or in person during the sessions. The following five statements were presented to the participant for each of the eight responsible AI themes:

1. I agree with the presented definition of this responsible AI theme.
2. In my personal opinion, I think that our organization adheres to this principle.
3. Our organization is of the opinion that we adhere to this principle.
4. Citizens have the opinion that our organization adheres to this principle.
5. The algorithm register contributes towards the adherence to this principle.

Our semi-structured interviews consist of four parts: The first part focuses on gaining insight into the current role of the participant, their expertise, experience, interest in the topic, and their involvement with initiatives for the algorithm register within their organization. The second part focused on the tasks and responsibilities of the organization, the role of algorithms and AI in their organization, the presence of an organization-wide mission or vision statement regarding the use of algorithms or AI in general, and awareness surrounding the importance of responsible use of the technology. The third part focused solely on the organizational perception and progress towards the implementation of their algorithm register. The last part of the interviews consisted of a free format part where the participants were asked about their personal visions and experiences with an algorithm register. Participants were provided with an empty piece of paper and were asked to draw what a process for the implementation of an algorithm register could or should look like from their perspective. They were instructed to think out loud and describe what they were drawing and to constantly provide explanations for their actions. The participants were allowed to use, share, or refer to any existing internal documentation present within their organization. The researchers prepared follow-up questions about possible elements that could come up, based on expectations and knowledge gained from the literature study.

Data analysis and creation of artifacts. The definition process for the algorithm register commenced with the finalization of the literature study, survey, and interviews. We started from a publicly available definition and iteratively modified the definition using our results as input until a version was created that was satisfactory with our shared understanding of the topic. Additionally, we used Process-Deliverable Diagrams (PDDs), a meta-modeling technique often used for analyzing, storing, selecting, and assembling different method fragments in the field of method engineering, to model the processes provided by the participants in the interviewees (van de Weerd and Brinkkemper, 2009). Schuitemaker et al. (2024) describe in detail how the final reference process is built from the resulting artifacts.

3.3 Treatment validation

Finalizing our last step within the design cycle, we conducted follow-up interviews with the previously involved participants to validate the algorithm register management process. These interviews were structured in a way to facilitate a discussion about the resulting reference process, with the intention to confirm, substantiate, verify and correct our findings (Buchbinder, 2010). This would allow us to confirm that the perspectives of the participant were accurately captured.

4 Results

Between the 24th of May 2023 and the 12th of June 2023, we conducted six interviews with seven participants that were at the time employed at a total of six different public organizations within the Netherlands. Less than one third of the approached organizations eventually were willing to participate in our research. The interviews all took around 90 minutes to complete, with one interview being split into two sessions of 45 minutes due to limited availability of the participant. All but one interview was conducted in person and on location, except for a single interview taking place through Microsoft Teams. The following sections provide a summary of the results that are most relevant to our research objectives. Full details are in the technical report (Van Vliet et al., 2024).

4.1 Defining the algorithm register

Using the process described in section 3.2, we have built our definition of an algorithm register through the steps described in Table 2. The table shows how the definition was carefully developed over multiple rounds of discussion. Italics are used to indicate the differences between a definition-version and its predecessor. Ultimately, we created a definition that describes all aspects of the register and its goals.

Version	**Definition**	**Motivations for change**
0.1	An algorithm register is both "a journal of algorithms used by, in this case, a public organization that has some level of direct impact on its citizens" (Murad and Hafrey, 2021) and "an overview of the algorithms used to provide municipal services" (Haataja, Fliert and Pasi 2020).	Both definitions insinuate that an algorithm register is nothing more than a descriptive list of individually used algorithms. It can be said that algorithm registers serve transparency and accountability because they make decision-making more transparent to individuals (Landsbergen, 2004). Furthermore, our research indicates alignment with the definition of AI governance as defined by Mäntymäki et al. (2022) due to the mentioned presence of an internal register.
0.2	*An algorithm register is an AI governance mechanism for organizations to be transparent to*	Transparency constitutes one of the main goals of an algorithm register, but it does not cover the entire purpose. The explicability principle by Floridi and Cowls (2019) covers accountability as well.
	society about the deployment of AI algorithms.	Our results show that an algorithm register can also contain information on the responsible entity for the algorithm, therefore causing accountability.
0.3	An algorithm register is an AI governance mechanism for organizations to be transparent to society about the deployment of AI algorithms *and the entity responsible.*	Within software engineering, the word "deployment" can mean that the system is ready for users (Dearie, 2007). As there is a possibility where an algorithm is already part of the algorithm register before it is deployed, the choice has been made to replace "deployment" with "use" in the definition.
0.4	An algorithm register is a governance mechanism that allows organizations to be transparent to society about the	The algorithm register is not mandated by law yet, so the word "allow" was used to reduce the overall impact of an algorithm register. This way, an algorithm register solely gives the opportunity to

	use of AI algorithms and the entity responsible.	provide transparency and accountability, rather than ensuring it.
1.0	An algorithm register is a governance mechanism that allows organizations to be transparent *and to provide accountability to society by providing an overview of (1) the documentation about algorithms, (2) the organization or organizational department responsible for their use, and (3) the goals pursued with their use.*	The final definition was created by making it more tangible through describing what actual form the governance mechanism will take in practice. Calling it an overview and describing its elements make it more descriptive towards what information the overview should consist of.

Table 2. Representation of our progress towards the definition of the algorithm register.

4.2 Interview results

Algorithm usage. Three out of six participants indicated that the role that algorithms play in their organization has already become very large. Two out of the remaining three classified the role of algorithms as medium, with the last one classifying it as small respectively. This question was often cause for discussion, due to the uncertainty of what definition of *algorithm* should be used to make this assessment. Participants noted that when a very broad definition was used, that algorithms would be present almost everywhere and that therefore their role within the organization would be huge. However, when using a narrower definition that often included the terms of "self-learning" or "complex", the role of algorithms was indicated to be much smaller. Some participants specified that it was a hard question to answer from their position within the organization in relation to the indicated size of the organization. We specified that the participants should use the narrower definition for this question.

Regarding the use of algorithms, participants often mentioned that algorithms currently are mostly used to support human-performed tasks. All participants acknowledged that the influence of technology and algorithms has been growing and that their future role is likely to change. In most cases however, specific emphasis was placed on the fact that this development would have to happen in a responsible manner. On the question where the calls for responsible use were mostly coming from, 4 out of 6 times this came from the bottom-up. In the remaining two cases, the calls for responsible use came from both bottom-up and top-down perspectives. In addition to this, 4 out of 6 participants mentioned the presence of an organizational wide mission or vision statement that addressed the responsible use of (AI) algorithms.

Algorithm register contribution. When referring to the algorithm register itself, an apparent divide appeared. Some organizations were confident in their assessment of what the algorithm register would mean for their organization, while others were still figuring this out. Concerns were raised that insufficient thought has been given to the various forms of transparency, predominantly with different statutory tasks in mind. Furthermore, concerns were raised regarding accidentally enabling citizens to 'game the system', as they would be provided with information about characteristics of algorithms that could help them to circumvent the intended purpose of a system. Overall, the algorithm register was predominantly described as an endpoint, and referred to 'a bow around the package'. Participants described the algorithm register as a window through which the public can be informed, to ensure the deployment of algorithms was achieved in a responsible and ethical manner. One participant went as far as saying that the algorithm register does not change their internal processes and that the register would have no impact on the implementation of safeguards or the performance of risk-mitigating measures during the development and deployment of algorithmic systems. Other participants however argued that the obligation to publish information to the public did add value, as it would force organizations to have all matters in order, which they admitted has not always been the case in the past. Furthermore, some organizations especially recognized and highlighted the internal value that the algorithm register could

bring, as it would help to create an overview of which algorithms are used in their organization. This overview was more than once referred to as an internal register and was mentioned to function as a collection point for all necessary information. However, most organizations did not acknowledge the existence of an already present internal register in their organization yet.

4.3 Responsible AI principles survey results

Definition agreement. The survey questions were almost equally split between instances where they were answered beforehand in an online setting or in person during the interview sessions. For both instances, the provided answers were discussed and the participants were asked for additional motivation and substantiation. The results of the survey show three interesting aspects. As indicated by Figure 1, the presentation of the definitions of the responsible AI principles by Fjeld et al. (2020) resulted in many diverse answers. Participants showed indifference regarding the degree of agreement with the definition of said principles, as shown in Figure 1. Additional analysis unveiled both diversity in the levels of agreement with the principles, and a large degree of variability between the organizations. We suspect that the different roles that offer different perspectives on the matter played a role here, for us highlighting the complexity and the multidisciplinary nature of the subject matter.

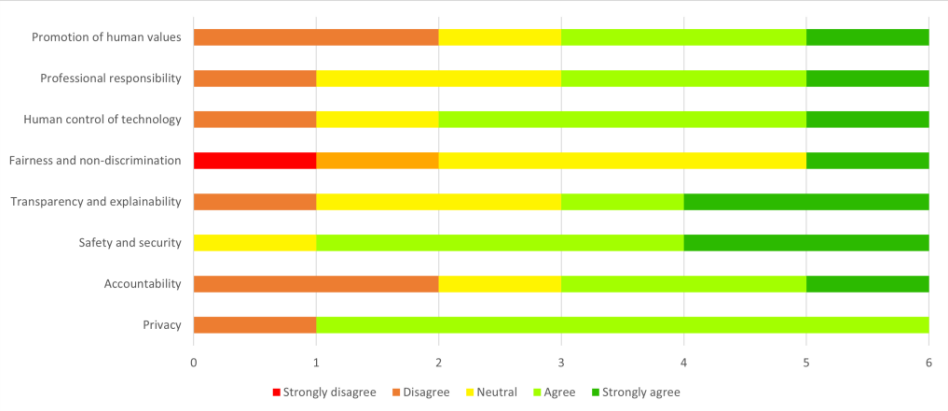

Figure 1. Levels of agreement with the definition of the responsible AI principles.

Adherence to principles. The middle part of the survey resulted in mixed results, as shown in Figure 2. Most participants found it difficult to answer the questions either from the perspective of the citizens or from their own perspective as well. Furthermore, for some participants, the answers might have become skewed towards the neutral option, as their indicated disagreement with the earlier presented definition of the principles created a tendency towards the choice for a bit more negative option. While it is hard to draw significant conclusions from these three questions, a slight indication of a lack of confidence towards the familiarity with the principles could be observed.

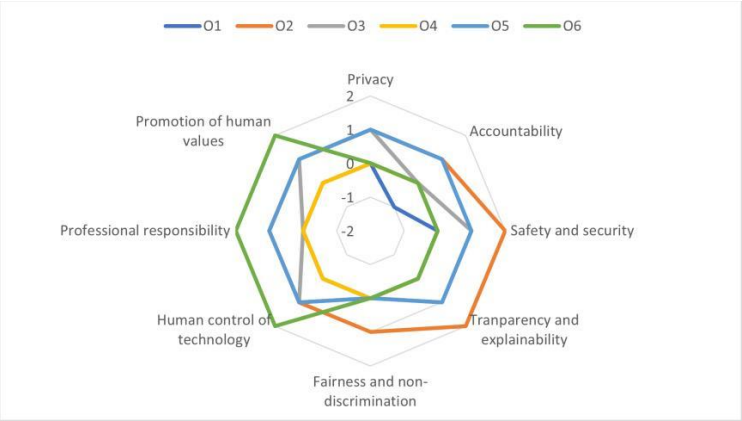

Figure 2. Perception on the organizational adherence to the responsible AI principles.

Algorithm register contribution. For the final part of the survey, contrary to our expectations, some participants had negative feelings about the contribution of the algorithm register towards the realization of a more responsible use of AI in its current form. Participants were often clearly opinionated on the register's contribution towards the principles, therefore revealing two different perspectives from their reasoning. Firstly, some participants argued that the algorithm register forces organizations to reassess everything that they have been doing so far, which can be a positive effect. However, other participants argued that people within the organizations think about the principles beforehand regardless, and then use the register in its current form only to report on these principles.

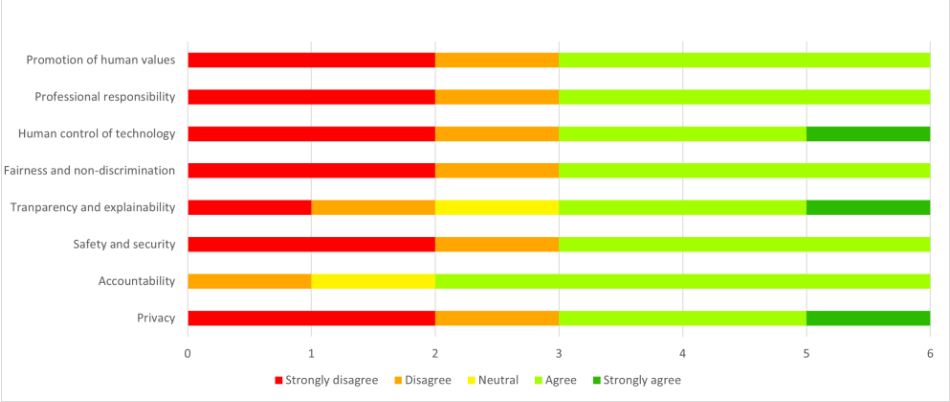

Figure 3. Opinions on the contribution of the algorithm register towards the principles.

4.4 Organizational process

Some participants provided a clear step by step process, while others were openly brainstorming and writing down concepts, roles, and activities that they thought were relevant. One of the resulting PDD artifacts is shown in Figure 4, the others are found in the technical report (Van Vliet et al., 2024). The fragments are based on direct input from the participants and therefore leave room for improvement.

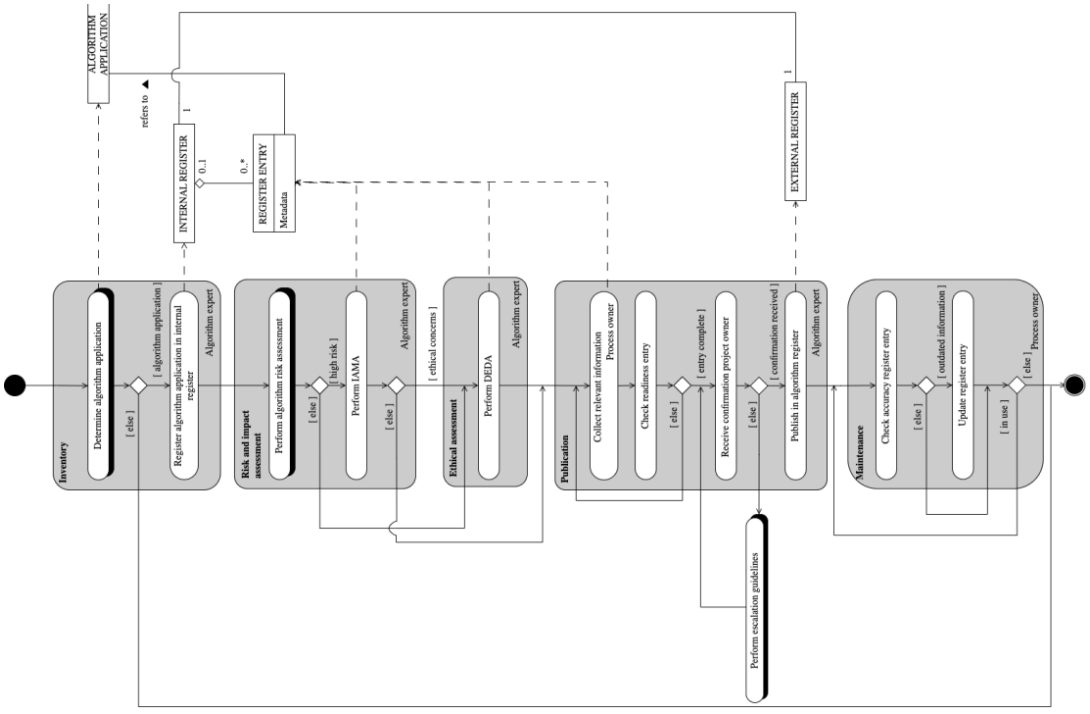

Figure 4. Method fragment represented as a Process-Deliverable Diagram.

5 Algorithm Register Management Reference Process

Based on the PDDs of the organizations, a reference process was engineered. This reference process showcases how organizations can implement an algorithm register, depicted in Figure 5. Each phase is further discussed and, where possible, linked to responsible AI principles it helps realizing.

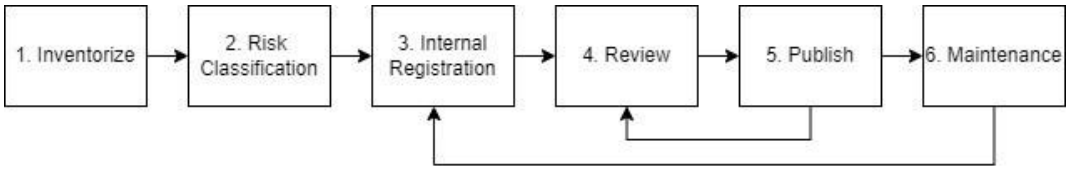

Figure 5. Simplified representation of the constructed reference process.

(1) Inventorize: Algorithms have been part of an organization for a long time already which have never before needed to be published to the public. Most organizations therefore lack a complete overview of their current use of algorithms. Necessary information and documentation required to create this overview is often scattered throughout the organization. The first step that should be taken revolves around systematically mapping the necessary aspects of all relevant operational algorithms. We observed two different approaches to achieve this. The first approach revolved around creating a task force to start searching for algorithms by interviewing employees of different departments. The goal was to raise awareness and to identify algorithms themselves by running into them. The second strategy is the creation of a central point of contact where employees can report algorithms that they encounter.

(2) Risk Classification: Assessing and classifying algorithms based on perceived risk produces valuable information and determines what concrete steps should be taken. We observed four criteria where the classification was based on: (1) risk level as defined by the AI Act or Dutch government, (2) the presence of personal data, (3) whether the algorithm or the process that it is used in can cause legal consequences for data subjects, and (4) gut feeling. It is important to create a multi-stakeholder collaboration, as defined within the professional responsibility principle by Fjeld et al. (2020). Within the case studies, ethicists, legal advisors, data scientists, and privacy and security experts were mentioned. Existing risk assessment tools are recommended to be used during this step. The assessment and all considerations should be documented during this step, criteria for high or low risk can change over time. Documentation of the considerations means that it becomes possible to explain decisions on a substantive level, either to take accountability or to provide additional transparency when needed.

(3) Internal registration: All organizations part of the case study had some internal documentation of the algorithms. However, in many cases, this was decentralized and not necessarily standardized. The value of the internal register grows as the amount of impactful algorithms in the organization increases. It can provide value as it provides a centralized point of information and creates a single point of truth for the characteristics of the algorithm register. One case already pointed out that the internal register's value was huge as it gave them an overview of the algorithms in the organization, also helping in coordinating development efforts by preventing similar initiatives from being started. The internal register also turned out to be helpful for answering any questions about the use of algorithms and for responding to requests under the Freedom of Information act. Furthermore, as future expectations are that the use of algorithms will continue to grow, the amount of information about these algorithms will grow as well. We consider the internal register therefore an essential part for information management. The internal register should contain all information that is necessary and relevant for external publication as well. Incomplete register entries can be queued here before they are ready for publication.

(4) Review: Reviews must have taken place before entries in the internal register, before they can be pushed towards external publication. The text and content should both be reviewed at least once to make sure that the information in the internal registry entry is accurate. A communication department is recommended to be responsible for checking whether the text is written in a way that it is optimally understandable for the target audience. Furthermore, extra checks can be added to this step to prevent

the (accidental) publication of characteristics of an algorithm that people with malicious intent can misuse and to try to game the system.

(5) Publish: In this step, the information from the internal registry entry is pushed to an external medium suitable for presentation. The case studies described two examples, either an organization's own medium or the National Algorithm register. Apart from compliance with the law, some case studies showed additional reasons for publication even when not strictly necessary: (1) if an algorithm that is not necessarily scary or sensitive has received attention in the media or has started a public debate, (2) for the purpose of demystification, when publishing can contribute towards the creation of a more realistic view of algorithms in the public eye or (3) algorithms that previously have drawn attention from a regulatory body that passed the inspection. Mentioned reasons to refrain from publication were (1) exceptions to the law, (2) concerns regarding gaming the system, (3) intellectual property rights and (4) no apparent impact on citizens or organizations. In these cases, organizations can even report on why the choice was made to refrain from full publication. This way they can be transparent about why they cannot be fully transparent, which may be viewed as more trustworthy from the public.

(6) Maintenance: Until the end of the lifecycle, information in the internal register, the external register, and the algorithms themselves will require maintenance. Documentation relevant for the register should be updated regularly, to prevent loss of data when employees leave the organization. Maintenance should be done periodically to ensure continuous quality. It is recommended to reassess high risk or high impact algorithms more frequently in order to increase the chance of detecting unwanted consequences before they become reality. It is important that experts with relevant knowledge are responsible for the maintenance of each item.

5.1 Outcome of the validation interviews

Validation interviews were held with half of the original participants as described in section 3.3. Although the participants agreed with our representation of the reference process, a total of six relatively small suggestions for change were made. A single activity and three sub activities were added in total. One sub-activity was removed, and one participant role was changed from 'process owner' to 'expert'.

6 Discussion

The scarce scientific literature on algorithm registers highlights its novelty. This paper pioneers by delving into the organizational aspects of algorithm registers and proposing a theoretical framework on how to develop them. As such, we see several scientific implications. First, Cath and Jansen (2021) emphasize the superficial nature and the lack of contextualization in existing algorithm registers. Our study reveals that public organizations are actively addressing this concern. Some of our case organizations, constrained by the need to protect sensitive information, are implementing their own algorithm registers in addition to the national register. This approach aims to eliminate superficiality and provides more meaningful explanations. Another concern raised in the literature was that of transparency, cf. Houtzager, Verbeek, and Terlouw (2022). To answer this call, we elaborated on the considerations organizations make regarding this issue. Our study revealed that it might not be possible to expect the same level or transparency from each organization. Finally, regarding the development of algorithm registers, the literature is even scarcer. Existing studies briefly mention or compare algorithm registers, but there is a notable gap in understanding how they should be established. Floridi (2020) discusses algorithm registers in Helsinki and Amsterdam but does not provide any lessons learned or best practices. Our study extends this work by inventorying the methods that organizations are using or planning to use to ultimately guide organizations in setting up an algorithm register.

Our study also resulted in several practical implications. First, the presence of an internal register shifts the required efforts from external publication more towards internal documentation. It is here where the possibilities lie for the creation of an internal overview that can be used to align the use of AI technologies with the organization's strategies, objectives, and values, therefore creating a governance mechanism. The internal register also enables internal auditing, so that organizations dealing with

concerns regarding gaming the system can still work towards providing meaningful transparency. Secondly, the necessity of the review step in the reference method indicates a difference in the required language for internal use and external presentation of the information towards the public. In practice, filling and maintaining the different registries may require different efforts and activities, so that the external register can focus on the presentation towards the general public in a way that the vast majority of the people use the provided information in an optimal way.

During the process, we have ensured the validity and reliability of the findings by following a rigorous interview protocol, triangulating the evidence to increase the confidence in the conclusions, and applying a systematic method engineering approach to construct both the organizational PDDs and the reference process. However, there are some remaining threats to the validity of the results. First, the limited number of participating organizations and our focus on Dutch government organizations may impact the broader applicability of our findings. Furthermore, a potential selection bias exists, given that organizations participated only if they were open to discuss their algorithmic practices, potentially skewing the results. However, the interviewees that participated in this study worked at various types of organizations in a range of different roles, strengthening the credibility of our findings. While the gathered data is in certain areas too limited to draw conclusive statements, the case studies provide many interesting insights that helps shaping the implementation of an algorithm register in an organization, still contributing towards the creation of a shared understanding. In future work, we suggest extending this research by involving more organizations from more diverse industries and countries. We have proposed a reference process for the implementation of an algorithm register, but we have not rigorously evaluated its effectiveness in practice. The Dutch Police is using the method during the ongoing implementation of their algorithm register and we plan to further address this limitation in the future.

7 Conclusion and Future Work

With the creation of our definition of an algorithm register and the constructed reference method, we achieved the first two goals of this research captured by **RQ1** and **RQ2**. The definition broadens the perspective of the topic and extends the meaning from just a simple list to the possibility of a usable mechanism. Furthermore, we believe that our reference process for the algorithm register provides an accurate description of how organizations are currently approaching the implementation of the algorithm register. We think that these two contributions help towards the creation of a shared understanding of the topic and aid in the prevention of the development of sub-optimal solutions. Furthermore, we managed to identify possible ways that the algorithm register can contribute towards the realization of AI principles, as our third research objective captured by **RQ3**. We do however not assume that the current version of the reference process immediately leads to the perfect implementation of an algorithm register when an organization starts from scratch. Furthermore, we do not assume the current version addresses all current opportunities or challenges in the fields of AI implementation, responsible AI, and AI governance. Our research identified ample opportunity for the improvement of the reference method, opening up opportunity for future work and additional research towards the implementation of an algorithm register, or other methods that help applying responsible AI principles in practice. We plan to validate the reference method and update it according to empirical evidence about its effectiveness. We have performed theory-building case studies to inform our definition of algorithm register and the creation of a reference process for its implementations. To increase the depth and breadth of data collection, we plan to conduct theory-testing case studies (Dul and Hak, 2007) that involve a wider range of countries, organization types, as well as distinct roles within each organization. Also, longitudinal studies will allow monitoring changes over time. All in all, the algorithm registers and other AI challenges will continue to exert their transformative power on (public) organizations. This calls for an investigation into the long-term impacts of algorithm registers on organizational practices and public trust. The phenomenon gives a tangible demonstration of all challenges, considerations and efforts that need to be undertaken to properly implement, adopt and use (AI) algorithms in an organization and illustrates how AI is affecting the way organizations and society function and interact. Because of these reasons, we strive for more awareness on the topic of algorithm registers.

Acknowledgements

Sergio España is supported by a María Zambrano grant of the Spanish Ministry of Universities, co-funded by the Next Generation EU European Recovery Plan.

References

Akinsola, J., M. Adeagbo, K. Oladapo, S. Akinsehinde, and F. Onipede (2022). "Artificial intelligence emergence in disruptive technology." In: *Comput. intellig. and data sciences*. CRC Press, pp. 63–90.

Alsheibani, S., Y. Cheung, and C. H. Messom (2018). "Artificial Intelligence Adoption: AI-readiness at Firm-Level." *PACIS* 4 (2018), 231–245.

Baregheh, A., J. Rowley, and S. Sambrook (2009). "Towards a multidisciplinary definition of innovation." *Management decision 47* (8), 1323–1339.

Benbya, H., T. H. Davenport, and S. Pachidi (2020). "Artificial intelligence in organizations: Current state and future opportunities." *MIS Quarterly Executive* 19 (4).

Bhagat, S. and B. Bolton (2008). "Corporate governance and firm performance." *JoCf* 14 (3), 257–273.

Brock, J. K. - U. and F. Von Wangenheim (2019). "Demystifying AI: What digital transformation leaders can teach you about realistic artificial intelligence." *California man. review* 61 (4), 110–134.

Brynjolfsson, E. and A. Mcafee (2017). "Artificial intelligence, for real." Harvard Bus. Rev. 1, 1–31.

Buchbinder, E. (2011). "Beyond checking: Experiences of the validation interview." *QSW* 10, 106–122.

Cammers-Goodwin, S. and N. Van Stralen (2021). "Making data visible in public space." *McGill GLSA Research Series* 1 (1), 1–32.

Cath, C. and F. Jansen (2021). "Dutch Comfort: The limits of AI governance through municipal registers." *arXiv preprint arXiv:2109.02944*.

Collins, C., D. Dennehy, K. Conboy, and P. Mikalef (2021). "Artificial intelligence in information systems research: A systematic literature review and research agenda." *IJoIM* 60, 102383.

Cubric, M. (2020). "Drivers, barriers and social considerations for AI adoption in business and management: A tertiary study." *Technology in Society* 62, 101257.

Dearle, A. (2007). "Software deployment, past, present and future." In: *FOSE'07*. IEEE, pp. 269–284.

Dul, J. and T. Hak (2007). *Case study methodology in business research*. Routledge.

Dwivedi, Y. K., L. Hughes, E. Ismagilova, G. Aarts, C. Coombs, T. Crick, Y. Duan, R. Dwivedi, J. Edwards, A. Eirug, et al. (2021). "Artificial Intelligence (AI): Multidisciplinary perspectives on emerging challenges, opportunities, and agenda for research, practice and policy*." International Journal of Information Management* 57, 101994.

El Khatib, M. M., H. M. Alzoubi, G. Ahmed, H. H. Kazim, S. A. A. Al Falasi, F. Mohammed, and M. Al Mulla (2022). "Digital transformation and SMART-the analytics factor." In: 2022 *IC on Business Analytics for Technology and Security (ICBATS)*. IEEE, pp. 1–11.

Eubanks, V. (2018). Automating inequality: How high-tech tools profile, police, and punish the poor. *St. Martin's Press*.

Fagerberg, J. (2006). *"Innovation: A guide to the literature."*

Fjeld, J., N. Achten, H. Hilligoss, A. Nagy, and M. Srikumar (2020). "Principled artificial intelligence: Mapping consensus in ethical and rights-based approaches to principles for AI." *Berkman Klein Center Research Publication* (2020-1).

Floridi, L. and J. Cowls (2019). *A unified framework of five principles for AI in society*. Har. Dat. SciRev.

Floridi, L. (2019). "Translating principles into practices of digital ethics: Five risks of being unethical." *Philosophy & Technology* 32 (2), 185–193.

Floridi, L. (2020). "Artificial intelligence as a public service: Learning from Amsterdam and Helsinki." *Philosophy & Technology* 33 (4), 541–546.

Floridi, L. (2021). "Establishing the rules for building trustworthy AI." *Ethics, Governance, and Policies in Artificial Intelligence*, 41–45.

Floridi, L. and J. Cowls (2022). "A unified framework of five principles for AI in society." *Machine learning and the city: Applications in architecture and urban design*, 535–545.

Floridi, L., J. Cowls, M. Beltrametti, R. Chatila, P. Chazerand, V. Dignum, C. Luetge, R. Madelin, U. Pagallo, F. Rossi, et al. (2018). "AI4People—an ethical framework for a good AI society: opportunities, risks, principles, and recommendations." *Minds and machines* 28, 689–707.

Girasa, R. (2020). *Artificial intelligence as a disruptive technology: Economic transformation and government regulation.* Springer Nature.

Goldacre, B., N. J. DeVito, C. Heneghan, F. Irving, S. Bacon, J. Fleminger, and H. Curtis (2018). "Compliance with requirement to report results on the EU Clinical Trials Register: cohort study and web resource." *bmj* 362.

Haataja, M. (2020). "Public AI Registers: Realising AI transparency and civic participation in government use of AI: Whitepaper written by Meeri Haataja, Linda van de Fliert and Pasi Rautio." *City of Amsterdam, City of Helsinki: Saidot Ltd.*

Hagendorff, T. (2020). "The ethics of AI ethics: An evaluation of guidelines." *Minds and machines* 30 (1),99–120.

Hall, B. H. and B. Khan (2003). *Adoption of new technology.*

Holmström, G. E., A. Hellström, P. G. Jakobsson, P. Lundgren, K. Tornqvist, and A. Wallin (2012). "Swedish national register for retinopathy of prematurity (SWEDROP) and the evaluation of screening in Sweden." *Archives of Ophthalmology* 130 (11), 1418–1424.

Holmström, J. (2022). "From AI to digital transformation: The AI readiness framework." *Business Horizons* 65 (3), 329–339.

Ingrams, A., W. Kaufmann, and D. Jacobs (2022). "In AI we trust? Citizen perceptions of AI in government decision making." *Policy & Internet* 14 (2), 390–409.

Jahíc, J. and R. Roitsch (2020). "State of the practice survey: Predicting the influence of ai adoption on system software architecture in traditional embedded systems." In: *European Conference on Software Architecture*. Springer, pp. 155–169.

Jobin, A., M. Ienca, and E. Vayena (2019). "The global landscape of AI ethics guidelines." *Nature machine intelligence* 1 (9), 389–399.

Kelly, A. (2021). "A tale of two algorithms: The appeal and repeal of calculated grades systems in England and Ireland in 2020." *British Educational Research Journal* 47 (3), 725–741.

Khatri, V. and C. V. Brown (2010). "Designing data governance." *Comm. of the ACM* 53 (1),148–152.

Kitsios, F. and M. Kamariotou (2021). "Artificial intelligence and business strategy towards digital transformation: A research agenda." *Sustainability* 13 (4), 2025.

Konaté, S. and B. Pali (2023). "You have to talk with us, not about us": Exploring the harms of wrongful accusation on those affected in the case of the Dutch 'childcare-benefit scandal." *Revista de Victimología/Journal of Victimology* (16), 139–164.

Korac-Kakabadse, N. and A. Kakabadse (2001). "IS/IT governance: Need for an integrated model." *Corporate Governance: The international journal of business in society* 1 (4), 9–11.

Landsbergen, D. (2004). "Screen level bureaucracy: Databases as public records." *Government Information Quarterly* 21 (1), 24–50.

Madan, R. and M. Ashok (2023). "AI adoption and diffusion in public administration: A systematic literature review and future research agenda." *Government Information Quarterly* 40 (1), 101774.

Mäntymäki, M., M. Minkkinen, T. Birkstedt, and M. Viljanen (2022). "Defining organizational AI governance." *AI and Ethics* 2 (4), 603–609.

Martin, K. (2019). "Ethical implications and accountability of algorithms." *Journal of business ethics* 160 (4), 835–850.

Munn, L. (2023). "The uselessness of AI ethics." *AI and Ethics* 3 (3), 869–877.

Murad, M. (2021). "Beyond the" Black Box": Enabling Meaningful Transparency of Algorithmic Decision- Making Systems through Public Registers." PhD thesis. Mass. Institute of Technology.

Nagy, D., J. Schuessler, and A. Dubinsky (2016). "Defining and identifying disruptive innovations." *Industrial marketing management* 57, 119–126.

Naibei, I. and E. Siringi (2011). "Impact of Electronic Tax Registers on VAT Compliance: A Study of Kenyan Private Business Firms." *International Journal of Finance & Policy Analysis* 3 (1).

Nouws, S., M. Janssen, and R. Dobbe (2022). "Dismantling Digital Cages: Examining Design Practices for Public Algorithmic Systems." In: *ICEG*. Springer, pp. 307–322.

November, J. (2008). "New Zealand: Public Registers Report." *Comm. Law Bull.* 34 (3), 697–701.
Osmundsen, K., J. Iden, and B. Bygstad (2018). "Digital transformation: Drivers, success factors, and implications."
Reim, W., J. Åström, and O. Eriksson (2020). "Implementation of artificial intelligence (AI): a roadmap for business model innovation." *Ai* 1 (2), 11.
Rogers, E. M. (1995). "Diffusion of Innovations: modifications of a model for telecommunications." *Die diffusion von innovationen in der telekommunikation*, 25–38.
Rudko, I., A. Bashirpour Bonab, and F. Bellini (2021). "Organizational structure and artificial intelligence. Modeling the intraorganizational response to the ai contingency." *Journal of Theoretical and Applied Electronic Commerce Research* 16 (6), 2341–2364.
Sartori, R., A. Costantini, A. Ceschi, and F. Tommasi (2018). "How do you manage change in organizations? Training, development, innovation, and their relationships*." Front. in psych.* 9, 320628.
Schuitemaker, N., Van Vliet, M., Brinkkemper, S., Van De Weerd, I., España, S., Realizing the Accountability of Algorithms in the Public Sector: a Reference Method for Algorithm Registers (2023). *International Conference on Business Process Modeling, Development and Support* [Preprint]
Siebel, T. M. (2019). *Digital transformation: survive and thrive in an era of mass extinction.* RosettaBooks.
Solomon, J. (2020). *Corporate governance and accountability.* John Wiley & Sons.
Spiegelhalter, D. (2020). "Should we trust algorithms." *Harvard Data Science Review* 2 (1), 1.
Toreini, E., M. Aitken, K. Coopamootoo, K. Elliott, C. G. Zelaya, and A. Van Moorsel (2020). "The relationship between trust in AI and trustworthy machine learning technologies." In: *Proceedings of the 2020 conference on fairness, accountability, and transparency*, pp. 272–283.
UK Government, (2023). The Bletchley Declaration by Countries Attending the AI Safety Summit, 1-2 November 2023. UK.gov. URL: https://www.gov.uk/government/publications/ai-safety-summit-2023-the-bletchley-declaration/the-bletchley-declaration-by-countries-attending-the-ai-safety-summit-1-2-november-2023/ (visited on March 31, 2024).
Vakkuri, V., K.- K. Kemell, J. Kultanen, M. Siponen, and P. Abrahamsson (2019). "Ethically aligned design of autonomous systems: Industry viewpoint and an empirical study." *arXiv preprint arXiv:1906.07946.*
Van Vliet, M. et al., *Defining and Implementing Algorithm Registers: An Organizational Perspective*. Technical Report. Available: https://www.doi.org/10.17632/cy4vspgn53.1
Veale, M. and F. Zuiderveen Borgesius (2021). "Demystifying the Draft EU Artificial Intelligence Act—Analysing the good, the bad, and the unclear elements of the proposed approach." *Computer Law Review International* 22 (4), 97–112.
Weerd, I. van de and S. Brinkkemper (2009). "Meta-modeling for situational analysis and design methods." In: *Handbook of research on modern systems analysis and design technologies and applications*. IGI Global, pp. 35–54.
Weick, K. E. and R. E. Quinn (1999). "Organizational change and development." *Annual review of psychology* 50 (1), 361–386.
Weiner, J. (2022). *Why AI/data science projects fail: how to avoid project pitfalls*. Springer Nature.
Westenberger, J., K. Schuler, and D. Schlegel (2022). "Failure of AI projects: understanding the critical factors." *Procedia computer science* 196, 69–76.
Wieringa, M. (2020). "What to account for when accounting for algorithms: a systematic literature review on algorithmic accountability." In: *Proceedings of the 2020 conference on fairness, accountability, and transparency*, pp. 1–18.
Wieringa, R. J. (2014). *Design science methodology for information systems and software engineering*. Springer.
Wing, J. M. (2021). "Trustworthy ai." *Communications of the ACM* 64 (10), 64–71.
Wirtz, B. W., J. C. Weyerer, and C. Geyer (2019). "Artificial intelligence and the public sector—applications and challenges." International Journal of Public Administration 42 (7), 596–615.