Algorithmic Accountability

David Horneber · Sven Laumer

1 Introduction

Advancements in technology have led to the widespread adoption of machine learning (ML) algorithms in almost all areas of society (e.g., shaping customer experiences through recommendations, and supporting organizational activities through automating tasks). Despite their potential benefits (e.g., personalizing experiences, enhancing productivity, improving decision-making) various examples have shown that systems based on such algorithms can lead to negative consequences. For instance, use cases in healthcare (Obermeyer et al. 2019) or finance (Blattner and Nelson 2021), have demonstrated how ML systems can undermine fairness and discriminate against minorities by generating statistical or reproducing societal biases (Mitchell et al. 2021). Other examples in marketing have illustrated how individuals' privacy can be compromised by inferring intimidating knowledge about individuals and using it for targeted advertising (Hill 2012; Mattu and Hill 2022). With the development from simple rule-based to complex probabilistic algorithms, the potential for harmful effects gets even more pressing, as their operation is increasingly opaque and more and more automated. While it has been already widely investigated how such issues can be addressed (e.g., Liu et al. 2022; Mehrabi et al. 2021), many organizations still fail to mitigate the often unintended, negative outcomes of the ML systems they are developing, providing, and using. Thus, academics (e.g., Novelli et al. 2023; Wieringa 2020) and policymakers (e.g., Mökander et al. 2022; Smuha 2021) have put an increasing emphasis on the topic of algorithmic accountability to ensure the ethical development and use of such systems (Donia 2022).

Although algorithmic accountability is mentioned as an important principle in almost all guidelines and regulatory documents for the ethical development of ML systems, there is still considerable uncertainty on what constitutes it and how it can be accomplished (Jobin et al. 2019). This is due to the fact, that algorithmic accountability is an ambiguous concept that deals with many different questions (Bovens 2010; Poechhacker and Kacianka 2021; Wieringa 2020). Many of these questions arise from the existence of different accountability types. To shed light on these accountability types and to bring together the different questions around algorithmic accountability, our article aims to introduce the topic to the BISE community. Future work on algorithmic accountability promises to build the foundation for organizations and developers to effectively implement and manage accountability measures and to inform policy-makers on how to appropriately regulate ML systems. Furthermore, it could provide insights into how individuals deal with accountability-related concerns and how it affects their interaction with ML systems.

D. Horneber (✉) · S. Laumer
Friedrich-Alexander-Universität Erlangen-Nürnberg, Fürther
Straße 248, 90429 Nuremberg, Germany
e-mail: david.horneber@fau.de

2 Conceptual Foundations

Algorithmic accountability focuses on the question of who takes the obligation to justify the design, use, and outcome of machine learning systems and who assumes responsibility for the negative consequences of these systems (Bovens 2007; Wieringa 2020). By assigning responsibility and demanding justifications one could oversee and restrain the behaviors of others. Hence, demanding algorithmic accountability can be understood as a governance function to either proactively avoid the negative impacts of the provision and use of ML systems or to reactively sanction accountable actors, if there have been any adverse effects caused by the systems (Novelli et al. 2023). While previous research has extensively studied the requirements that are necessary to achieve the ethical development and use of ML systems (e.g., fairness (Feuerriegel et al. 2020), interpretability (Lipton 2018), privacy (Liu et al. 2022), robustness (Tocchetti et al. 2022)), algorithmic accountability is focused on the demands that shape these requirements and the governance measures that responsible actors can take to fulfill them.

Answering the question of who has to take responsibility for the actions and outcomes of ML systems, the term algorithmic accountability has the potential to mislead. ML systems as technical artifacts cannot be held accountable (Bryson et al. 2017; Martin 2022). However, organizations and users that develop or use these systems could (Martin 2022). As Martin (2019) argued, ML systems are not value-neutral, instead, they integrate the norms and intentions of those developing, operating, and using them. While organizations can design ML systems in a way that gives full control to the users to put the responsibility on them, typically, firms remain in control of the systems they provide. Thus, the normative obligation to be accountable for these systems resides with them and their employees developing and operating ML systems. Concerning the motivation of organizations to act accountable, the accountability concept implies that stakeholders (e.g., regulators) who demand accountability, can put consequences (e.g., economic damage) on accountable actors if they fail to provide justifications and do not assume responsibility (Bovens 2007; Wieringa 2020). Therefore, if stakeholders can successfully enforce accountability demands (e.g., through regulations), firms should have an incentive to avoid the negative effects of their systems and adequately justify their design and use.

Focusing on the question of who are the stakeholders that demand accountability there are typically two types of external stakeholders (Bovens 2007; Wieringa 2020). First, there are users and other individuals that are affected by ML systems (e.g., applicants). These can execute a form of social accountability (e.g., public protest), individually or collectively, to express their concerns and demand justifications and consequences. Second, there are regulators. These are administrative or legal institutions that can implement institutional control mechanisms (e.g., legislation) to require explanations and enforce sanctions.

Following these accountability demands and mechanisms, typically requirements are set that firms have to fulfill to avoid negative consequences. Organizational and technical measures can be used to accomplish this. From an organizational perspective, companies can implement different governance instruments (e.g., establishing and enforcing development guidelines) to hold their employees accountable who are dealing with ML systems (Donia 2022; Schneider et al. 2022). As such, algorithmic accountability can be part of a corporate digital responsibility strategy that is focused on taking responsibility for the ML systems that are developed or used in an organization (Lobschat et al. 2021; Mueller 2022). In addition, developers need to consider certain technical requirements and properties of ML systems that need to be technically managed and addressed to work toward algorithmic accountability by design (i.e., the fulfillment of all requirements that are set by different accountability demands).

Figure 1 illustrates the different stakeholders and the external accountability demands (institutional and social accountability) and internal accountability measures (organizational and technical accountability) that exist around a ML system.

In Fig. 1, we distinguish between provider and operator organizations, since the firm that develops and deploys an ML system is not necessarily the same organization that runs and employs it. Similarly, there are typically different types of practitioners within an organization (e.g., developers, product owners, quality managers, specialists who use ML systems) who deal with ML systems and are in control of their design and use. Whenever this is the case, companies or practitioners among each other hold distributed accountabilities to be responsible for the consequences of the system they develop, operate, and employ. In the following sections, we discuss the different accountability demands and measures in detail.

2.1 Social Accountability

Machine learning systems pose several risks to users and other individuals who are affected by these systems. To enable this group to effectively demand accountability whenever they have used or been influenced by an algorithmic decision is vital to ensure that the power of these systems can be controlled. Furthermore, these stakeholders need to be enabled to proactively assess the actions and impact of ML systems.
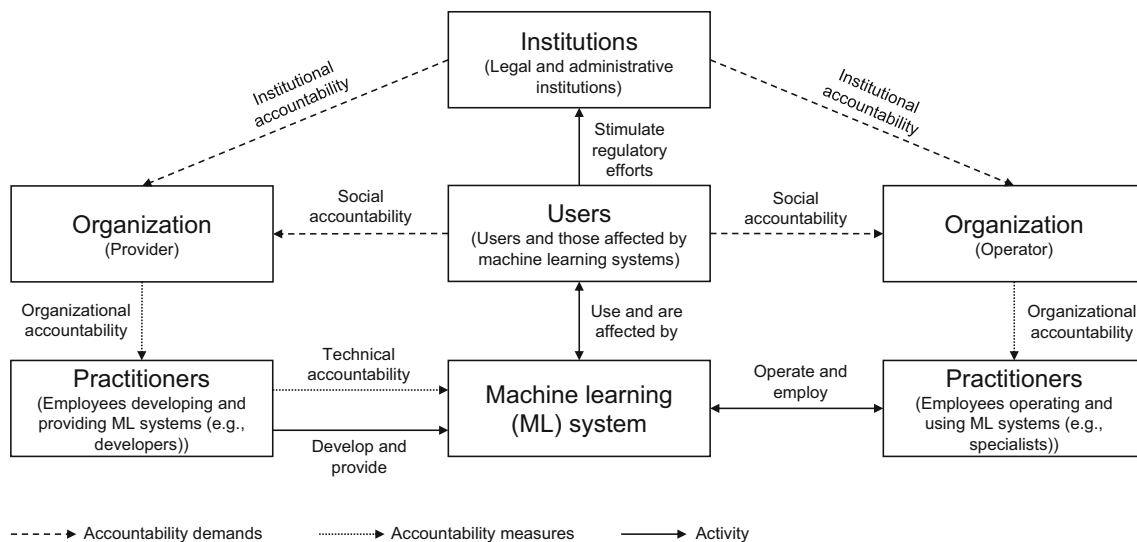
Fig. 1 Algorithmic accountability types and stakeholders

Research has shown that higher perceptions of individuals that providers of ML systems have implemented accountability measures (perceived system accountability) can positively influence users' trust and satisfaction (Shin 2021; Shin and Park 2019). In addition, it can strengthen the willingness of users to accept and follow algorithmic advice (Adam 2022). Perception of system accountability is thereby influenced by the interpretability of the systems (e.g., providing evidence that the system is regularly reviewed) (Adam 2022; Shin 2021). Corporations consequently risk losing trust and stimulating reputational concerns if they fail to provide explainable and transparent systems (Buhmann et al. 2020).

Users have different possibilities to claim and execute accountability based on the context in which they are using ML systems: consumer or professional. In the professional context, users can perform individual (e.g., noncooperation) and collective (e.g., striking) user resistance behaviors to create awareness of negative effects and attempt companies to adapt their business practices (Kellogg et al. 2020; Möhlmann et al. 2021). Consumers on the other hand can make use of different demand- (e.g., discontinuing ML services), rating- (e.g., negative ratings), or discourse-based (e.g., negative word-of-mouth) mechanisms to denounce misconduct and impose consequences on organizations (Grégoire and Fisher 2008; Labrecque et al. 2013). In both contexts, consumers and professionals can also express their concerns to policymakers to seek support and regulation from legal or administrative institutions. Journalists can support such efforts by providing algorithmic accountability reports in public outlets and creating and framing a public discourse about algorithmic issues (Diakopoulos 2015; Kellogg et al. 2020).

While initial results have shown that individuals can effectively trigger regulatory efforts or discipline organizations by such methods (e.g., Benson et al. 2020), several challenges can prevent consumers and professionals from enforcing accountability. For example, many collective approaches involve mobilizing a large number of users to accomplish an effect. This in turn requires that a considerable amount of individuals are affected and that they use the same set of standards to evaluate the conduct of the accountable actor. In addition, firms can take countermeasures (e.g., framing a public discourse) to mitigate the negative reactions of users. In the subsequent section, we explain the role of legal and administrative institutions to safeguard individuals and to support them to enforce their accountability demands.

2.2 Institutional Accountability

The call for regulation and governance of ML systems from legal institutions is becoming increasingly intense (Donia 2022; Mökander et al. 2022; Smuha 2021). While existing laws such as the GDPR regulate ML practices to some extent, it can be expected that in the future more extensive legislation will be implemented (Koniakou 2022). Hence, organizations need to be prepared for stronger liability requirements. The term liability is often mentioned alongside accountability and deals with justification when confronted with legal institutions and possible juridical sanctions (Slota et al. 2021). Legal accountability controls organizations' ML activities and ensures that interventions can be enforced (Stahl 2021). For organizations, two goals are relevant when it comes to legal accountability. First, they need to be prepared to provide information about their ML systems whenever legal

institutions request them. Second, legal compliance needs to be accomplished from a management as well as from a technical perspective. Establishing appropriate governance measures and considering legal requirements in the design of ML systems can be a means to accomplish these goals.

Besides legal accountability, firms can expect that the accountability demands raised by administrative institutions will also play a more important role in the future (Matus and Veale 2022). Private or public regulators such as auditors or certification bodies can implement different mechanisms to hold organizations accountable. The most frequently discussed is algorithm audits (e.g., Brown et al. 2021; Raji et al. 2022). Algorithm audits can be either focused on technical aspects such as reliability and robustness or on problematic behaviors and the societal consequences of ML systems (Raji et al. 2020). Auditing for societal effects, often some type of impact assessment is proposed to assess the potential negative outcomes of ML systems (Metcalf et al. 2021). Similar to algorithm audits, certifications can be used to assess the ML practices of an organization and to make them meet certain requirements (Matus and Veale 2022). Another frequently suggested accountability mechanism, that can be implemented via independent external stakeholders, is the collection and reporting of critical incidents (Brundage et al. 2020). Anonymized disclosure of unexpected algorithmic behaviors can help researchers and practitioners to identify shareable lessons to avoid such behaviors in the future. It also can support auditors to guide their inspections toward specific, repeatedly occurring problems (Raji et al. 2022).

While in the case of legal accountability, organizations are faced with solely negative outcomes if they fail to provide the right information and justifications, successfully fulfilling administrative accountability requirements could enable provider and operator organizations to achieve legitimacy and trust in their ML systems (Adam 2022; Kordzadeh and Ghasemaghaei 2022; Martin and Waldman 2022). Therefore, if implemented correctly, companies can have an incentive to voluntarily conduct third-party inspections and adhere to accountability requirements set by such stakeholders. Next, we elaborate on the measures that organizations can take to proactively counter accountability claims.

2.3 Organizational Accountability

Already several companies have taken steps to implement a form of organizational accountability. Organizational accountability can be understood as the internal governance measures of an organization to align their ML practices with the organizational values and external requirements (Mäntymäki et al. 2022). As a part of this, companies can hold their employees accountable who are

dealing with ML systems to achieve internal control (Donia 2022). Implementing accountability among practitioners developing, operating, and managing ML systems, presents organizations with the challenge to monitor responsibilities across departments and roles (Feuerriegel et al. 2022). While the purpose of accountability demands by users and regulators is to proactively prevent as well as reactively sanction misbehaviors, the goal of organizational accountability is mainly proactive. Firms can implement internal accountability measures, to promote awareness of potential harms within their organization and continuously monitor and improve their ML systems to follow internal and external policies.

Currently, most organizations rely on a principle-based approach to govern their ML systems (Jobin et al. 2019; Mittelstadt 2019). This approach has often been criticized as insufficient, due to the issues arising when trying to translate ethical principles and guidelines into practice (Mittelstadt 2019). While often criticized, defining and implementing internal policies can be considered a strategic first step to achieving successful governance of ML systems (Schneider et al. 2022; Seppälä et al. 2021). Building upon these guidelines, companies can specify and apply standardized processes and rules to develop, test, deploy, and operate their ML systems (e.g., coding, documentation, testing, and architectural guidelines) (Schneider et al. 2022). These can then be used in internal audits to assess whether the algorithmic practices adhere to these regulations (Raji et al. 2020). Furthermore, internal impact assessments can be conducted to identify risks and manage them proactively. Similar to incident systems set up by external stakeholders, firms can implement such a system internally, to ensure that problems are resolved promptly and lessons learned can be drawn (Schneider et al. 2022). Establishing such accountability measures allows organizations to be aware of and mitigate the potential negative consequences of their ML systems. Moreover, it enables them to proactively create accountable ML systems and achieve legal compliance. Whenever corporations are faced with external accountability demands, providing information gathered through their internal governance measures supports them in fulfilling these demands.

In addition to the regulation-oriented approach, companies can also attempt to empower their employees through training or communication to address their responsibilities when working with ML systems (Ryan et al. 2022; Schneider et al. 2022). Making employees aware of their duty to act accountable is an essential first step to ensuring that ethical guidelines and procedures are translated into practice. Subsequently, we describe the technical requirements that developers face in the course of accountability demands.

2.4 Technical Accountability

An important technical requirement that needs to be considered by developers to work toward accountability by design is the interpretability of ML systems. The need for interpretability arises from the obligation to explain and justify the design and outcomes of ML systems (Wieringa 2020). What information is needed to be interpretable is always dependent on the stakeholder and context of accountability demands (Berente et al. 2021). In general, developers and providers should reflect on the following interpretability objectives to make their ML systems accountable: design transparency (Kroll 2021; Loi et al. 2021) and auditability (Felzmann et al. 2020; Kroll 2021). Design transparency requires developers and providers to disclose information about the design goals, input data, and mechanisms of construction and operation of their ML systems (Kroll 2021; Loi et al. 2021). Auditability enables third parties to probe, understand and review ML systems before deployment and during operation (Felzmann et al. 2020; Kroll 2021). Achieving these two interpretability goals should enable different stakeholders to assess and understand the actions and intentions of the systems.

Another technology-related factor that must be reflected by designers and developers is the ability to monitor and control ML systems. While it may not always be necessary and reasonable that humans oversee the actions and outcomes of a system, firms and practitioners must be aware that there are high-risk applications (e.g., medical treatment recommendation) that pose controllability requirements that must be met (Methnani et al. 2021). Accordingly, designers and developers need to carefully assess the impact and risk of their systems to decide on the right level of human oversight. Keeping humans in the loop is often seen as insurance for organizations to avoid negative consequences. However, companies need to ensure that individuals are informed of their role and are supported by the systems (e.g., through interpretability mechanisms) to act by the relevant accountability standards.

Finally, firms and developers should aim to technically address the well-known sources for negative outcomes of ML systems, such as robustness and safety, bias, and privacy. For each of these issues, researchers and practitioners have already started to propose technical measures to test and protect the algorithmic system against them (e.g., Liu et al. 2022; Mehrabi et al. 2021; Tocchetti et al. 2022). Developers should make use of these to mitigate the unintended detrimental effects of ML systems and to avoid unnecessary accountability demands.

3 Implications for BISE Research and Future Work

Despite its increasing relevance, the topic of algorithmic accountability has not received much attention in the BISE community so far. The different accountability demands and measures motivate several empirical and design-oriented research opportunities. Table 1 gives an overview of potential areas for future research organized around the different accountability types introduced in the previous sections.

As explained in Sect. 2.1, social accountability is about users and individuals affected by ML systems holding the providers or operators of the system they use accountable. Thus, research in this area should focus on understanding how different perceptions of accountability arise and how they can influence the interaction and usage of ML systems. Thereby, different perceptions of accountability can be distinguished. Individuals can either perceive that the provider or operator organization is willing or unwilling to be accountable for the system they offer, or they can perceive that they are accountable for certain outcomes or actions when interacting with the system. In both cases, scholars can investigate what are the individual, organizational, and environmental characteristics that shape such perceptions. Furthermore, it can be examined how these perceptions influence diverse dependent variables such as algorithmic acceptance, intention to delegate tasks to an ML system, or organizational trust in the provider and operator. Additionally, further work is required that focuses on the strategies that consumers and professionals use to demand accountability. An important issue is thereby the effectiveness of different strategies and the challenges that consumers and professionals face when demanding accountability. Results in this domain can help to better understand how accountability claims are raised and determined by distinct individual and environmental factors. Moreover, it sheds light on the effects of algorithmic accountability on human behaviors.

Current legislation on ML systems is relatively weak (Mittelstadt 2019). BISE researchers can support regulators in specifying algorithmic policies in several ways. First, scholars can theorize about the long-term consequences of ML systems to advise legal and administrative authorities about upcoming implications. Additionally, it can be systematically examined, which values individuals hold that should guide algorithmic accountability policies (Mason 2021). Furthermore, empirical work can be conducted to analyze the effectiveness of future and existing regulations. Focusing on the impact of algorithmic regulations on organizations, future work can investigate how firms adopt certain legal and administrative requirements and how this affects the development and provision of ML systems. How practitioners act on regulations and inspections and if

Table 1 Overview of future research opportunities

Accountability type	Sample research questions
Social	What individual, organizational, and environmental characteristics shape users' perceptions of corporate/personal accountability when interacting with ML systems?
	How do perceptions of corporate/personal accountability influence individual behavior when interacting with ML systems?
	What procedures can individuals and society use to hold providers of ML systems accountable?
Institutional	What policies can effectively prevent the negative societal ramifications of ML systems?
	How can algorithmic incidents be publicly tracked and reported?
	How can policymakers incentivize the implementation of accountability measures?
Organizational	What governance measures can be effectively used by organizations to identify and mitigate the negative consequences of their ML systems?
	How does the perceived accountability of practitioners dealing with algorithmic decisions influence their business activities?
	What are the drivers and challenges for the adoption of accountability measures?
Technical	What are the design principles to develop accountable ML systems?
	How to develop and provide interpretable ML systems that allow to pro- and reactively address accountability demands?
	How does the design of ML systems influence corporate/personal accountability perceptions?

they see it as a threat or opportunity, are additional questions that can be studied. While policy-related research may not be the core of the BISE community, analyzing such issues from a sociotechnical lens can help to inform regulators and practitioners dealing with algorithmic accountability.

Managing algorithmic accountability within an organization presents the challenge of translating abstract ethical principles into actionable requirements and practical governance measures (Mäntymäki et al. 2022; Mittelstadt 2019). Scholars can support practitioners to tackle this challenge by examining how to successfully establish an ethical-aware culture within a company. Moreover, it can be studied which governance instruments enhance the perceptions of practitioners to be accountable for their work and how this translates to technical decisions and business activities. Taking away the skepticism of employees and enabling them with the necessary resources and capabilities to develop ethically-aligned ML systems, many knowledge- and awareness-related barriers to more accountable organizational practices can be removed (Tomilova 2021). Additionally, researchers can analyze which governance measures are most effective in mitigating risks and how they can create actual business value for corporations. Furthermore, the effectiveness of organizational response strategies to accountability claims can be studied to provide guidelines for practitioners on how they can mitigate the economic and reputational damage that often follows accountability demands. Considering the fact, that typically multiple actors are involved in creating and providing ML systems, future work can investigate how distributed accountabilities are organized between

different actors and how this influences their perceived accountability. In this vein, it can be also analyzed how different roles (e.g., management vs. developer) are affected by and deal with (distributed) accountability demands. Generating insights on how firms can implement and manage accountability measures is crucial to support practitioners in managing the potential negative consequences of ML systems.

Although there can be many reasons for algorithmic accountability claims, the technical design is often one of the main factors contributing to the adverse outcomes of ML systems. Researchers have already started to theoretically discuss interpretability requirements that are necessary to develop accountable ML systems (e.g., Felzmann et al. 2020; Kroll 2021). However, it has not been defined and evaluated how they can be implemented to allow different stakeholders to assess the business practices of organizations and practitioners. Taking a design-oriented lens, scholars can build and assess explainability mechanisms and interaction interfaces that fulfill the interpretability needs of distinct institutional or social accountability demands. Focusing on the controllability of ML systems, scholars can examine and propose accountability design principles for different levels of autonomous systems. Moreover, additional work on the different sources of negative consequences of ML systems and the development of technical measures to address them is necessary. This can help to tackle the currently existing robustness, fairness, and privacy challenges. Another interesting area for future research can be to analyze the development practices around the implementation of interpretability and controllability methods. Understanding

the challenges of the technical implementation can help to develop better tools and methods to work toward accountability by design.

4 Conclusion

Overall, it can be summarized that algorithmic accountability is much more than the mere question of who takes responsibility for the impacts of ML systems. Instead, it is about how institutions, organizations, and individuals can govern ML systems and how developers and provider of ML systems can fulfill their accountability obligations. While researchers have already started to propose measures and define requirements to achieve interpretable and controllable ML systems, their concrete implementation and the effects on the ecosystem around them have not gained much attention yet. By defining algorithmic accountability as a governance issue and introducing it from different perspectives we provide a foundation for future research on the topic.

References

Adam M (2022) Accountability-based user interface design artifacts and their implications for user acceptance of AI-enabled services. In: European conference on information systems, Timisoara

Benson A, Sojourner A, Umyarov A (2020) Can reputation discipline the gig economy? Experimental evidence from an online labor market. Manag Sci 66(5):1802–1825. https://doi.org/10.1287/mnsc.2019.3303

Berente N, Gu B, Recker J, Santanam R (2021) Managing artificial intelligence. MIS Q 45(3):1433–1450

Blattner L, Nelson S (2021) How costly is noise? Data and disparities in consumer credit. https://doi.org/10.48550/ARXIV.2105.07554

Bovens M (2007) Analysing and assessing accountability: a conceptual framework. Eur Law J 13(4):447–468. https://doi.org/10.1111/j.1468-0386.2007.00378.x

Bovens M (2010) Two concepts of accountability: accountability as a virtue and as a mechanism. West Eur Politics 33(5):946–967. https://doi.org/10.1080/01402382.2010.486119

Brown S, Davidovic J, Hasan A (2021) The algorithm audit: scoring the algorithms that score us. Big Data Soc 8(1):205395172098386. https://doi.org/10.1177/2053951720983865

Brundage M, Avin S, Wang J, et al (2020) Toward trustworthy AI development: mechanisms for supporting verifiable claims. https://doi.org/10.48550/ARXIV.2004.07213

Bryson JJ, Diamantis ME, Grant TD (2017) Of, for, and by the people: the legal lacuna of synthetic persons. Artif Intell Law 25(3):273–291. https://doi.org/10.1007/s10506-017-9214-9

Buhmann A, Paßmann J, Fieseler C (2020) Managing algorithmic accountability: balancing reputational concerns, engagement strategies, and the potential of rational discourse. J Bus Ethics 163(2):265–280. https://doi.org/10.1007/s10551-019-04226-4

Diakopoulos N (2015) Algorithmic accountability: journalistic investigation of computational power structures. Digit J 3(3):398–415. https://doi.org/10.1080/21670811.2014.976411

Donia J (2022) Normative logics of algorithmic accountability. In: ACM conference on fairness, accountability, and transparency, pp 598–598. https://doi.org/10.1145/3531146.3533123

Felzmann H, Fosch-Villaronga E, Lutz C, Tamò-Larrieux A (2020) Towards transparency by design for artificial intelligence. Sci Eng Ethics 26(6):3333–3361. https://doi.org/10.1007/s11948-020-00276-4

Feuerriegel S, Dolata M, Schwabe G (2020) Fair AI: challenges and opportunities. Bus Inf Syst Eng 62(4):379–384. https://doi.org/10.1007/s12599-020-00650-3

Feuerriegel S, Shrestha YR, von Krogh G, Zhang C (2022) Bringing artificial intelligence to business management. Nat Mach Intell 4(7):611–613. https://doi.org/10.1038/s42256-022-00512-5

Grégoire Y, Fisher RJ (2008) Customer betrayal and retaliation: when your best customers become your worst enemies. J Acad Mark Sci 36(2):247–261. https://doi.org/10.1007/s11747-007-0054-0

Hill K (2012) How target figured out a teen girl was pregnant before her father did. Forbes. https://www.forbes.com/sites/kashmirhill/2012/02/16/how-target-figured-out-a-teen-girl-was-pregnant-before-her-father-did/?sh=3da7bbd56668

Jobin A, Ienca M, Vayena E (2019) The global landscape of AI ethics guidelines. Nat Mach Intell 1(9):389–399. https://doi.org/10.1038/s42256-019-0088-2

Kellogg KC, Valentine MA, Christin A (2020) Algorithms at work: the new contested terrain of control. Acad Manag Ann 14(1):366–410. https://doi.org/10.5465/annals.2018.0174

Koniakou V (2022) From the "rush to ethics" to the "race for governance" in artificial intelligence. Inf Syst Front. https://doi.org/10.1007/s10796-022-10300-6

Kordzadeh N, Ghasemaghaei M (2022) Algorithmic bias: review, synthesis, and future research directions. Eur J Inf Syst 31(3):388–409. https://doi.org/10.1080/0960085X.2021.1927212

Kroll JA (2021) Outlining traceability: a principle for operationalizing accountability in computing systems. In: Proceedings of the ACM conference on fairness, accountability, and transparency, pp 758–771. https://doi.org/10.1145/3442188.3445937

Labrecque LI, vor dem Esche J, Mathwick C, Novak TP, Hofacker CF, (2013) Consumer power: evolution in the digital age. J Interact Mark 27(4):257–269. https://doi.org/10.1016/j.intmar.2013.09.002

Lipton ZC (2018) The mythos of model interpretability. Commun ACM 61(10):36–43. https://doi.org/10.1145/3233231

Liu B, Ding M, Shaham S, Rahayu W, Farokhi F, Lin Z (2022) When machine learning meets privacy: a survey and outlook. ACM Comput Surv 54(2):1–36. https://doi.org/10.1145/3436755

Lobschat L, Mueller B, Eggers F, Brandimarte L, Diefenbach S, Kroschke M, Wirtz J (2021) Corporate digital responsibility. J Bus Res 122:875–888. https://doi.org/10.1016/j.jbusres.2019.10.006

Loi M, Ferrario A, Viganò E (2021) Transparency as design publicity: explaining and justifying inscrutable algorithms. Ethics Inf Technol 23(3):253–263. https://doi.org/10.1007/s10676-020-09564-w

Mäntymäki M, Minkkinen M, Birkstedt T, Viljanen M (2022) Defining organizational AI governance. AI Ethics. https://doi.org/10.1007/s43681-022-00143-x

Martin K (2019) Ethical implications and accountability of algorithms. J Bus Ethics 160(4):835–850. https://doi.org/10.1007/s10551-018-3921-3

Martin K, Waldman A (2022) Are algorithmic decisions legitimate? The effect of process and outcomes on perceptions of legitimacy of AI decisions. J Bus Ethics. https://doi.org/10.1007/s10551-021-05032-7

Martin K (2022) Algorithmic bias and corporate responsibility: how companies hide behind the false veil of the technological imperative*. In: Martin K (ed) Ethics of data and analytics. Auerbach, pp 36–50. https://doi.org/10.1201/9781003278290-7

Mason RO (2021) Policy for ethical digital services. J Assoc Inf Syst 22(3):571–578. https://doi.org/10.17705/1jais.00672

Mattu S, Hill K (2022) How a company you've never heard of sends you letters about your medical condition*. In: Martin K (ed) Ethics of data and analytics. Auerbach, pp 107–111. https://doi.org/10.1201/9781003278290-17

Matus KJM, Veale M (2022) Certification systems for machine learning: lessons from sustainability. Regul Gov 16(1):177–196. https://doi.org/10.1111/rego.12417

Mehrabi N, Morstatter F, Saxena N, Lerman K, Galstyan A (2021) A survey on bias and fairness in machine learning. ACM Comput Surv 54(6):1–35. https://doi.org/10.1145/3457607

Metcalf J, Moss E, Watkins EA, Singh R, Elish MC (2021) Algorithmic impact assessments and accountability: the co-construction of impacts. In: Proceedings of the 2021 ACM conference on fairness, accountability, and transparency, pp 735–746. https://doi.org/10.1145/3442188.3445935

Methnani L, Aler Tubella A, Dignum V, Theodorou A (2021) Let me take over: variable autonomy for meaningful human control. Front Artif Intell 4:737072. https://doi.org/10.3389/frai.2021.737072

Mitchell S, Potash E, Barocas S, D'Amour A, Lum K (2021) Algorithmic fairness: choices, assumptions, and definitions. Ann Rev Stat Appl 8(1):141–163. https://doi.org/10.1146/annurev-statistics-042720-125902

Mittelstadt BD (2019) Principles alone cannot guarantee ethical AI. Nat Mach Intell 1(11):501–507. https://doi.org/10.1038/s42256-019-0114-4

Möhlmann M, Zalmanson L, Henfridsson O, Gregory RW (2021) Algorithmic management of work on online labor platforms: when matching meets control. MIS Q 45(4):1999–2022. https://doi.org/10.25300/MISQ/2021/15333

Mökander J, Juneja P, Watson DS, Floridi L (2022) The US algorithmic accountability Act of 2022 vs The EU artificial intelligence act: what can they learn from each other? Minds Mach 32(4):751–758. https://doi.org/10.1007/s11023-022-09612-y

Mueller B (2022) Corporate digital responsibility. Bus Inf Syst Eng 64(5):689–700. https://doi.org/10.1007/s12599-022-00760-0

Novelli C, Taddeo M, Floridi L (2023) Accountability in artificial intelligence: what it is and how it works. AI Soc. https://doi.org/10.1007/s00146-023-01635-y

Obermeyer Z, Powers B, Vogeli C, Mullainathan S (2019) Dissecting racial bias in an algorithm used to manage the health of populations. Sci 366(6464):447–453. https://doi.org/10.1126/science.aax2342

Poechhacker N, Kacianka S (2021) Algorithmic accountability in context Socio-technical perspectives on structural causal models. Front Big Data 3:519957. https://doi.org/10.3389/fdata.2020.519957

Raji ID, Smart A, White RN, Mitchell M, Gebru T, Hutchinson B, Smith-Loud J, Theron D, Barnes P (2020) Closing the AI accountability gap: defining an end-to-end framework for internal algorithmic auditing. In: Proceedings of the conference on fairness, accountability, and transparency, pp 33–44. https://doi.org/10.1145/3351095.3372873

Raji ID, Xu P, Honigsberg C, Ho D (2022) Outsider oversight: designing a third party audit ecosystem for AI governance. In: Proceedings of the 2022 AAAI/ACM conference on AI, ethics, and society, pp 557–571. https://doi.org/10.1145/3514094.3534181

Ryan M, Christodoulou E, Antoniou J, Iordanou K (2022) An AI ethics 'David and Goliath': value conflicts between large tech companies and their employees. AI Soc. https://doi.org/10.1007/s00146-022-01430-1

Schneider J, Abraham R, Meske C, Vom Brocke J (2022) Artificial intelligence governance for businesses. Inf Syst Manag. https://doi.org/10.1080/10580530.2022.2085825

Seppälä A, Birksted T, Mäntymäki M (2021) From ethical AI principles to governed AI. In: Proceedings of the international conference on information systems, Austin

Shin D (2021) The effects of explainability and causability on perception, trust, and acceptance: implications for explainable AI. Int J Hum-Comput Stud 146:102551. https://doi.org/10.1016/j.ijhcs.2020.102551

Shin D, Park YJ (2019) Role of fairness, accountability, and transparency in algorithmic affordance. Comput Hum Behav 98:277–284. https://doi.org/10.1016/j.chb.2019.04.019

Slota SC, Fleischmann KR, Greenberg S, Verma N, Cummings B, Li L, Shenefiel C (2021) Many hands make many fingers to point: challenges in creating accountable AI. AI Soc. https://doi.org/10.1007/s00146-021-01302-0

Smuha NA (2021) From a 'race to AI' to a 'race to AI regulation': regulatory competition for artificial intelligence. Law Innov Technol 13(1):57–84. https://doi.org/10.1080/17579961.2021.1898300

Stahl BC (2021) Addressing ethical issues in AI. In: Stahl BC (ed) Artificial intelligence for a better future. Springer, pp 55–79. https://doi.org/10.1007/978-3-030-69978-9_5

Tocchetti A, Corti L, Balayn A, Yurrita M, Lippmann P, Brambilla M, Yang J (2022) A.I. robustness: a human-centered perspective on technological challenges and opportunities. https://doi.org/10.48550/ARXIV.2210.08906

Tomilova A (2021) Barriers to improving algorithmic accountability: an elaborated action design research. In: Proceedings of the Pacific Asia conference on information systems, Dubai

Wieringa M (2020) What to account for when accounting for algorithms: a systematic literature review on algorithmic accountability. In: Proceedings of the conference on fairness, accountability, and transparency. https://doi.org/10.1145/3351095.3372833