

Johnson's Algorithmic Accountability and Corporate Accountability Dissonance.

Kirsten Martin, PhD

University of Notre Dame, kmarti33@nd.edu

Abstract In this chapter, I illustrate the impact of Johnson's scholarship on the study of corporate responsibility. I extend Johnson's accountability-as-practice to begin to scope (a) the normative grounding for *why* tech firms are accountable to their stakeholders and (relatedly) (b) *what* tech firms are accountable for, and (c) to *whom* firms are accountable. Firms are accountable for their design and deployment decisions about AI because firms have the power to make different decisions to elicit different moral implications in use. Firms are accountable for the decisions they make that impact others – whether those impacts are positive, as when creating value for stakeholders, or negative, as when firms destroy value for other stakeholders. Currently, industry has *accountability dissonance* where scholars and firms take credit for their ability to design algorithms that create value for key stakeholders while simultaneously shying away from the negative consequences, rules being broken, value being destroyed, or rights being diminished *for those same decisions*. I use online platforms to illustrate the importance of Johnson's approach to algorithmic accountability. In each case, firms have been slow to embrace accountability in the moral implications of their decisions; and attribution is made more complicated with the use of AI on a platform.

Keywords Accountability, practices, accountability dissonance, online platforms

Introduction

Tech firms struggle to acknowledge the scope of their accountability in the technologies they develop and produce. For decades, the industry appeared to embrace Meta CEO Mark Zuckerberg's "Move fast and break things" mantra: negative consequences are a natural side effect for the best technologies (Hamilton, 2022). One cannot know, they claimed, how anything worked (Calvello, 2017; Knight, 2017). AI technologies perhaps were sentient or just super, super complicated (Huckins, 2023).

But then we kept hearing about social media platforms that targeted teens with mental health problems or that recommended an insurrection (Mac & Kang, 2021). The same technologies that were lauded as progress were suddenly seen as tools harming vulnerable communities. As an alternative, executives (with the help of a few philosophers) came up with a new theory known as longtermism or effective altruism (Ackermann, 2022; Torres, 2021). While firms may break things, and perhaps could do things differently, this new technology (whatever it is) is in long term interest of society – to include an indeterminate number of future sentient robots, future humans, etc. People may be harmed in the here-and-now, but AI technologies are going to make more people and robots happy in the future.¹

Firms are dodging and weaving to avoid being seen as accountable for the technologies they produce. The shapeshifting narrative morphs to fit whatever counter argument or experience is reported: technology is neutral,²

¹ Unfortunately, those future humans and robots will then be sacrificed for even-more-future humans and robots if longtermism is applied at that time.

² Algorithms, like all technology, are not neutral (Johnson, 2023; Martin, 2022a)

accurate,³ efficient, and we all benefit from its existence either now or in the long, long term. Firms either had no say in the matter (technological determinism) or their brilliance led the design decisions to create the best technology for everyone (Martin, 2022a). Or, firms make mistakes but the magic of the market ‘picks’ the best technology. The ebb and flow of this industry narrative has only one constant theme: firms are not accountable for any downsides of their technology.

Enter Deborah Johnson’s article on Algorithmic Accountability (2021). Johnson’s accountability in the making reframes accountability as a relationship between an actor (e.g., firm) and a forum of others who believe an actor has an obligation to account to them: “accountability is a social practice involving actors, forums, shared beliefs and norms, performativity, and sanctions, and the practice is aimed at putting constraints on the exercise of power.” Importantly, whether firms acknowledge their role and accountability is not particularly important to holding firms accountable. In fact, firms’ denials may just be a part of a process or practice of accountability.

Johnson’s algorithmic accountability allows us to see industry’s shape shifting narrative about their (lack of) accountability as part of an accountability process: a forum demands actors to be accountable for their actions, actors may reflexively deny their role or accountability (or the harms), yet the accountability discourse should continue (see auto safety, manufacturing pollution, opioid epidemic, tobacco industry, etc.).

In this chapter, I focus on the important implications of Johnson’s approach to technology and accountability for research in management and business ethics. Where the AI accountability discourse can fixate on adequate notification to subjects or users, Johnson rightly (and consistently (Johnson, 2011, 2015; Johnson & Miller, 2006; Johnson & Powers, 2005)) refocuses accountability discussions on the value-laden decisions of developers of technology as creating technology with moral implications. As such, technology, including AI, are moral entities within their larger socio-technical system, and developers are accountable for their decisions in the creation of those moral entities - including when their decisions have moral implications for stakeholders of the AI in use (Johnson, 2011). Johnson, in this manner, leads a group of scholars arguing that technology is value-laden through design, and firms are accountable to a forum for these value-laden decisions (Abebe et al., 2020; Akrich, 1992; Birhane et al., 2021; Cummings, 2006; Johnson, In Press; Latour, 1992; Martin, 2019; Rosenblat & Stark, 2016).

The goal of this chapter is to explore the implications of Johnson’s accountability in the making for firms, markets, and industries. I illustrate the impact of Johnson’s scholarship on the study of corporate responsibility and extend her accountability-as-practice to begin to scope (a) the normative grounding for *why* tech firms are accountable to their stakeholders and (relatedly) (b) *what* tech firms are accountable for, and (c) to *whom* firms are accountable. These are related – identifying *why* firms are accountable to their stakeholders will shine a light on not only *what* firms are accountable for but also *to whom* they are accountable. Firms make value-laden decisions in the technologies they bring to market or technology they adopt. And firms are accountable to their stakeholders – those impacted by the firm’s decisions or impacting the firm – as to the moral implications of their decisions (Martin, 2019). Specifically, I argue:

- Accountability is a check on those in power (Johnson). Firms are accountable for their design and deployment decisions about AI because firms have the power to make different design/deployment decisions to elicit different moral implications in use. Firms are accountable to those impacted by those decisions because (a) firms have power to choose their decisions (b) which have consequences for others.
- Johnson’s approach to algorithmic accountability broadens *to whom* firms are accountable by emphasizing that firms are accountable to those that are impacted by decisions made in the design, development, and implementation of algorithms. Vulnerable stakeholders to the firm – legitimate stakeholders impacted by a firm’s decisions about AI but without power in the market or with the firm – feel the sharp edge of AI since the tool is disproportionately used to predict and categorize less powerful individuals in society (Benjamin, 2019; Eubanks, 2018; Martin, 2022b).
- Similarly, Johnson’s approach bring clarity as to *what* decisions firms are accountable for. Firms are accountable for the decisions they make that impact others – whether those impacts are positive, as when creating value for stakeholders or negative, as when firms destroy value for other stakeholders. Currently, firms have a form of *accountability dissonance* where scholars and firms take credit for their ability to design algorithms that create value and a positive impact on key stakeholders while simultaneously shying away from the negative consequences, rules being broken, value being destroyed, or rights being diminished *for those same decisions*.

I use online platforms to illustrate the importance of Johnson’s approach to algorithmic accountability in piercing through a *fog of accountability*.⁴ In each case, firms have been slow to embrace accountability in the moral

³ Firms may create that accuracy for profit (Martin, 2023a) However, industry still claims it is accurate and efficient (MIT Technology Review Insights, 2023).

⁴ Where Johnson and Powers note the added complexity with computer systems, with hardware and software and the many organizations that design and develop different parts of the technology, here I focus on the many human actors on a platform and the difficulty in parsing out who is accountable for what. According to actor-network

implications of their decisions; and attribution is made more complicated with the use of AI on a platform. These platforms allow other actors to transact, and those actors sometimes act in ways that are unethical. However, as I show below, Johnson's approach to algorithmic accountability sheds light on the specific accountability of platforms – to whom they are accountable, why they are accountable, and for what actions are they accountable – even with the creation of what we may call a fog of accountability created not only by the use of AI but also by the involvement of other market actors contributing to the problem.

Johnson's Accountability in the Making.

In "Algorithmic Accountability in the Making," Deborah Johnson takes a sharp turn away from algorithmic accountability as an issue of opacity (2021). Johnson argues this current focus on opacity and transparency "reflects a narrow and inadequate understanding of accountability" (p. 113). Similar to Nicholas Diakopoulos, transparency may promote accountability, but transparency does not ensure accountability.⁵ While a lack of opacity may be an issue for AI (as it is for many organizational decisions), Johnson argues that solving the opacity problem through 'transparency' or 'explainability' does not solve the issue of accountability: *who is accountable for the AI being used and to whom?*

Johnson begins with Mark Bovens' broader account of accountability as a starting point where Bovens describes accountability as "a relationship between an actor and a forum, in which the actor has an obligation to explain and to justify his or her conduct, the forum can pose questions and pass judgement, and the actor may face consequences" (Bovens, 2007).

Johnson identifies the important element of this definition of accountability as being within a relationship between an actor and a 'forum' where a forum is "any group of individuals who believe an actor has an obligation to them" (p. 118). To understand accountability, according to this approach, begs the question *accountable to whom?* For business, which will be important for AI, Johnson notes, "A company, for example, may be accountable (for different types of behavior) to a board of directors, regulatory agencies, employees, and customers" (p. 188).

Within the field of management, strategy, business ethics, etc, we refer to these members of the forum as *stakeholders* (Freeman, 1984, 2010; Freeman et al., 2007, 2020). While the field has shifted since, traditionally stakeholders are defined broadly as those who are impacted by a firm's decisions or impact the firm (Freeman, 1984). And, firms are accountable to stakeholders: "managers must account for the interests of these stakeholders when making decisions" (Phillips et al., 2003, p. 485).

Freeman, the seminal stakeholder scholar within the field, uses the terms responsibility or obligation in a manner similar to Bovens' and Johnson's use of accountability: both phrases are used to describe how an actor (a firm) must answer to others (stakeholders) and, for Freeman, how decisions the firm makes should take into consideration those stakeholders, since the firm will be accountable to them. Key to all is that the accountability emanates from the relationship between the actor and the forum (Bovens, 2007; Johnson, 2021) or between the firm and stakeholders (Freeman, 2010).

For Bovens, and differing from Johnson, both actor and forum *share* the belief that the actor is accountable to others.⁶ Johnson rightly points out (and business ethicists would agree!) that "there are situations in which actor and forums disagree. This situation typically occurs when an actor does not believe that an explanation is owed while a forum does" (2021, p. 118). Johnson notes this is where she and Bovens diverge:

On Bovens's account it would seem that we would have to say that there is no accountability because the actor and forum do not share the belief that an explanation is owed. This, however, doesn't seem right. That is, it seems wrong to say that when an actor doesn't believe he or she owes an explanation, the actor is, therefore, not accountable (p. 118).

Johnson identifies the key implication for where we are with AI: "it is important to note that when actors and forums disagree, it does not mean necessarily that there is no accountability, rather it may be a sign of accountability practices *in the making*." This approach to accountability should resonate with business scholars. Firms are accountable to many stakeholders – actors in a forum for Bovens and Johnson – with whom they have

theory, we would making similar arguments about different actors (material and nonmaterial) (Johnson & Powers, 2005).

⁵ Transparency is "an enabling factor that can support the monitoring of behavior with respect to ethical expectations" (Diakopoulos, 2020, p. 197).

⁶ "A central aspect of the relationship between actor and forum is that both parties share the belief that the actor has an obligation to explain particular types of behavior" (D. G. Johnson, 2021).

a relationship (Freeman, 1984, 2010). Disagreement as to the scope of what firms are accountable for does not mean firms have no accountability but only that accountability practices are *in the making*.

We can therefore take as a starting point to understand firms' accountability to stakeholders of their decisions about AI:

accountability is a social practice involving actors, forums, shared beliefs and norms, performativity, and sanctions, and the practice is aimed at putting constraints on the exercise of power" (Johnson, 2021, p. 123).

Johnson's approach, however, will broaden to whom firms are accountable (and for what) since Johnson's approach does not require the actor in the forum (or stakeholder) to have an established relationship with the firm. Key stakeholders for algorithmic accountability are those "who are or might be affected by algorithmic decision-making."⁷

Broadening the scope and grounding of business ethics

Why firms are accountable – AI and Power.

Power – who has it, how it is deployed and to what effect – is critical to understand AI. First, as noted by Jonne Maas, AI evokes a "power-dependence relation between those who shape a system (i.e., developers and users) and those affected by a system (i.e., end-users)" (Maas, 2023). Second, firms operate in markets where not all actors have equal power (Poole et al., 2021). Firms are accountable because their decisions, made possible through their power in the market and within this power-dependent relation between developers and users, impacts others. For Johnson, accountability is a practice "aimed at putting constraints on the exercise of power."⁸

We should expect to see accountability-as-practice in situations where power is not currently being held to account or where firms are acting in a manner that negatively impacts less powerful stakeholders. These marginalized stakeholders – by definition – do not have the power to hold the firm accountable. In fact, this lack of power is why the firm is able to make opportunistic decisions, where the harms fall on marginalized stakeholders.

For those that study firms and markets, we can take from Johnson:

- Firms should reasonably expect to be held accountable for their actions and decisions by those who are or might be affected by algorithmic decision-making
- Firms will be held accountable *in particular* in those situations where they exert power over stakeholders through the design and deployment of algorithms.
- Ignoring or denying accountability for actions is one step in being held to account – such denials are not the final word but an expected part of the process.
- We will see accountability in the making in situations where firms abuse power to harm vulnerable stakeholders since more powerful stakeholders are able to push back on firms when they are harmed and the 'negotiation' is done in private.

This last point is important to understand how *little* reliance we should place on firms' denial of accountability for their actions. Since accountability is a tool to place restraints on the abuse of power, calls for accountability will occur if and only if firms' exercise of power is not being adequately restrained by the immediate stakeholders. When we see accountability in the making, we should assume that a more powerful actor is currently not being held to account for the harms inflicted on less powerful members of the forum. Their denial should be expected.

⁷ "When it comes to the forums for algorithmic accountability, perhaps the largest forum is all of those who are or might be affected by algorithmic decision-making" (Johnson, 2021 p. 125).

⁸ See also Maas, accountability is meant to place a check on domination or superior and unaccountable power (Maas, 2023).

Clarifying what decisions firms are accountable for

For business ethics and management generally, Johnson’s approach to algorithmic accountability has implications for the scope of what firms will be accountable for. In other words, in correctly identifying firms as being accountable for their decisions about AI because (a) these firms are in power to make decisions that (b) impact others, we can then think through the scope of what firms should be held accountable for. Specifically, firms have the power to design, develop, and use AI. And firms are able to make different design and deployment decisions with different moral implications in use. Firms decide the training data to use, how data is labeled, the assumptions and models to use, the outcome variable to optimize, the types of testing before and after deployment, the criteria for ‘success’ in testing, etc, etc.

Any decisions that impact stakeholders – including value-laden design decisions of AI or the decisions to adopt and deploy AI within an organization – are within the scope of what firms are accountable for. This is not different from stakeholder theory that says that firms are accountable for the decisions they make, for the value they create and destroy, for the rights they enable or diminish, for the rules they break or respect, and for the values they uphold (Freeman et al., 2020).

And, firms take credit for design decisions or adoption decisions for the value these decisions create for themselves and some stakeholders. Their design decisions create value for the firm and are touted as a competitive advantage. People should use their product, work at their firm, and invest in their vision due to their strategic decisions to design and develop and adopt the ‘right’ AI technology. The executives and managers take credit for their decisions around AI and are happy to be held accountable for the positive impacts of their decisions.

Simultaneously, firms deny accountability for the negative implications for these same decisions. Firms have claimed a type of *algorithmic accountability dissonance*: firms make value-laden decisions in regards to the design, development, or adoption of AI and want to be accountable for the good outcomes, but claim they are not accountable for negative outcomes *for the same decision*.

Fig . Algorithmic Accountability Dissonance



Firms waive away the moral implications of the same design decisions for which they wish to take credit. For example, cheating detection software exclaims the value their design creates for schools and educators while brushing past the negative impacts to falsely accused students. Resume readers boast their ability to find the right candidate while blaming others for the same technology discriminating against women.

In general, if a firm claims to create value from their decision and wishes to be held accountable for the creation of value from their actions, then the firm is also accountable for the moral implications of those decisions as well. The accountability dissonance emerges when a firm wishes to be considered the ‘decider’ or ‘actor’ held accountable for the *good* consequences or moral implications of their decision while, perhaps, shrugging their shoulders as to the negative moral implications of that *same* decision.

Broadening To Whom Firms are Accountable.

So far, I have argued that Johnson’s approach to algorithmic accountability augments the stakeholder approach to corporate responsibility. Similar to stakeholder theory, firms are accountable to stakeholders for the decisions over which they have power (that are in their control). And more focus and more accountability will be thrust upon those firms with more power since, according to Johnson, accountability is a practice meant to constrain the exercise of power. Johnson applies this approach specifically to algorithms. Firms are accountable for the value laden decisions they make in regards to technology in the same way firms are accountable to stakeholders for their decisions around the design of cars and auto safety or are accountable to stakeholders for their decisions around how they treat employees.

Within management, the stakeholder approach to corporate responsibility, in its current incarnation, tends to focus on those stakeholders with whom firms have an immediate relationship or contract (customers, suppliers, communities, employees, financiers, etc) which is a subset of stakeholders firms are accountable to for their actions. For Johnson, a key stakeholder – or member of the forum – to whom firms are accountable are those impacted by AI decisions, such as subjects and users of AI. However, these stakeholders do not have a direct or market relationship to the firm developing or deploying AI.

In fact, the original approach to stakeholder theory within business ethics and strategy defined stakeholders broadly as *those impacted by or impacting firm decisions* (Freeman, 1984). However, more recent scholarship has focused on those stakeholders who are voluntary, with whom the firm has a relationship or agreement, or who previously provided value to the firm (Freeman & Phillips, 2002; R. A. Phillips, 1997; Van Buren III, 1999). This more limited definition focuses on those stakeholders that are obvious to managers.

This is a problem (Dawkins, 2015; Greenwood & Mir, 2019; Martin, 2023b). By limiting the scope of firm accountability to those stakeholders with whom the firm already has a relationship, who are voluntary, and who – in some way – have provided value to the firm, scholars studying firms are missing the less powerful, vulnerable stakeholders who have legitimate claims and interests. These vulnerable stakeholders - such as data subjects, victims of hate speech online, users or targets of AI programs -- are too easily ignored in the current version of management theory. Where management theory tells Facebook to pay attention to regulators and advertisers, users as well as the subjects of hate speech, targets of online trolls, and victims of a recommendation algorithm pushing non-consensual porn are left stranded with their legitimate claims but no market power.

Johnson’s approach should remind stakeholder theorists that firms are accountable to actors and organizations that are impacted by their decisions *regardless* if there is a contractual relationship or if the stakeholder is in a voluntary transaction with the firm. In fact, for AI, subjects and users of AI – e.g., students subject to cheating detection programs, individuals scanned at the border using FR – have no direct relationship with the firm developing AI but are considered key actors in the forum to whom the firm is accountable.

Within management theory and practice, this narrow definition of stakeholder, limiting obligations to those with whom a firm has a relationship and who previously benefited the firm, may be descriptive in what firms regularly do but lacks a focus on what firms *should* do in terms of legitimate stakeholders who are directly impacted by an AI decision but who have little power. In fact, those legitimate, marginalized stakeholders, impacted by a firms’ algorithmic decisions, are the people most in need of a firm being held accountable to them given the power-dependent relation between those who shape a technology and the end users who are subjects of the technology (Maas, 2023). In other words, Johnson’s approach pushes stakeholder theory to return to its normative roots and ask *to whom should a firm be accountable* and answer *those who are impacted by the firms’ decisions*.

Platforms and the Fog of Accountability

Johnson’s approach to accountability in the making provides clarity as to not only why firms are accountable but also for what and to whom firms are accountable. Firms are accountable because they use their power to make AI design and development decisions impacting others within the forum. Firms are accountable for decisions over which they have power (and could have decided otherwise) and to stakeholders, broadly defined, impacted by those decisions.

I turn now to demonstrate the utility of this approach to algorithmic accountability through digital platforms, where the actions of the platforms are sometimes minimized in comparison to the actions of the actors on the platform. Platforms act as an exchange and create a market for other organizations, individuals, and market actors to transact (Martin, 2024). The NYSE, Tinder, eBay, Facebook, and Google Search are all platforms or exchanges. Platforms have become more and more prevalent in our lives and, at the same time, have proven slippery when we look for who is accountable for what happens on a platform. Tech firms, reporters, and scholars appear to frame platforms as ‘neutral’ or as even having the goal of neutrality – in that platforms only facilitate the actions of others on their platforms.⁹ For platforms, the difficulty in attributing accountability in the design and deployment of AI is exacerbated by the difficulty in identifying the actions attributable to the platform versus the actors on the platform.¹⁰

⁹ “the myth that online platforms are neutral pervades the tech industry” (Chander & Krishnamurthy, 2018).

¹⁰ Similar to the complexity due to hardware versus software as noted by Johnson and Powers, and the problem of many hands as noted by Villegas-Galaviz, here I am arguing that platforms introduce an accountability fog due to many different actors (material and non-material) with not-clear lines of accountability (Johnson & Powers, 2005; Villegas-Galaviz & Martin, 2023).

For example, normally prices of goods or services sold on the platform are not the responsibility of the platform, since actors on the platform set their own prices for the goods and services they sell. However, Amazon recently designed a program to set and move prices on their own platform. Amazon’s use of a program to set prices on a platform illustrates the fog of accountability introduced by the use of an algorithm to facilitate a decision and the vagueness of the role of a platform in facilitating the transactions of other market actors. The normal ambiguity around accountability is made more complicated – creating a fog of accountability – if the firm can hide behind (a) the algorithm and (b) the mistaken notion that platforms are ‘neutral’ facilitators.

Social media platforms, for example, point to users and content creators as the bad actors who should be held accountable for the content created – e.g., dehumanizing, anti-LGBTQ hate speech. However, social media companies are accountable for their decisions about not only what content to keep on their site but also what content to recommend. Social media platforms are accountable for those decisions where they have power and to individuals impacted by those decisions. So, both can be true: the user is accountable for creating the hate speech the social media platform is accountable for promoting and recommending the hate speech by those impacted by the promotion – whether current users on the platform or not.

Table 1: Platforms, Algorithms, and Accountability.

	Actors on the Platform		Platform Decisions
Travel Platforms	Airline Decisions: pricing, routes, services	Customer Actions: payment, rating,	Travel Platform: recommendation, formatting, vetting, etc. what user data necessary to collect, use, share
Dating Apps	User Actions: feedback, truthfulness in profile, choosing potential partner, communication with other users.		Dating Platform: Factors of users to include, prioritize, exclude; recommendation system, ability to block/protect, vetting, etc; what user data necessary to collect, use, share.
Social Networks	Whether to engage with content online, what content to post.		Social Networks: What posts and groups to recommend and prioritize for users; which posts to remove; what engagement options to make available (like, dislike, post, etc); what reminders to provide; what feedback to allow (reporting of content); what user data necessary to collect, use, share
Search Engine	User Actions: Prompts, search terms, choice of link for content, engagement with content provider	Content Providers: proving relevant content for users.	Search Platforms: what content to include in the search index; prioritizing search results; matching users to relevant content; deprioritizing or blocking content could be fraudulent; what user data necessary to collect, use, share
LLM (ChapGPT)	User Actions: Prompts to provide, how to use the answers, whether to attribute the LLM in the use of the answer provided, etc	Content Providers: Whether to allow their work to be used in model development.	LLM: which sites or sources to use for training, design decisions as to model assumptions, which sites or sources to <i>exclude</i> from model development, which prompts to exclude from model development, how to apply the model in use, what answers to provide for prompts.

Travel Reservation Platform

Take, for example, a travel reservation platform. The platform allows travel agents (in the old days) or consumers (recently) to find and reserve flights from many different airlines. Such a travel platform does not dictate the routes that airlines offer or the types of seats or even the price. The platform may require that all prices be in USD or that the type of information provided by airlines is consistent to make sure that consumers can easily compare. These types of design decisions help facilitate the exchange or market for flights and decrease the cost to transact for consumers and airlines by being able to find and compare flights. So, we can think of platforms as making some design decisions and taking some actions for which they should be accountable.

Table 1 shows how we may start to think about the actions, decisions, and behaviors of different actors associated with digital platforms – for which they are accountable. In each case, airlines and travel exchanges claim to create value from these actions – a travel exchange will claim to show you the best routes for your preferences or make it easy for you to find a new travel route – but also are held accountable for the negative impacts or moral implications of those same decisions. Sabre, for example, was held accountable for their recommendation of American Airlines above all other airlines even when American was not the best route or price for the consumer (Friedman & Nissenbaum, 1996; Martin et al., 2024). Travel platforms consider their recommendation system to be a competitive advantage or differentiator for users, and the same platform must also be held accountable for the negative implications of those same recommendations.

Dating Apps

Dating apps, such as Tinder, eHarmony, or Bumble, offer a platform for individuals to find romantic matches. The actors on the platform are seen as the primary decision makers – deciding who they wish to engage with, what information is provided, how each individual is ‘portrayed’ on the app, and whether or not they wish to continue to communicate or see someone they meet.

However, the dating platform makes design decisions as to the type of information that is required or even possible to share. In doing so, these platforms decide which factors are important to consider when looking for a particular mate: whether height or education is important; whether marital status is required, whether truthfulness is even a goal. These platforms also decide which individuals on the site are in control of who they hear from or who users can block or mute. In addition, these dating apps decide the order in which potential partners are recommended to the users, which factors to highlight as particularly important, which factors to prioritize in the recommendation system. Dating platforms differentiate and compete based on these types of decisions.

Importantly, decisions as to what information is allowed to be important in choosing a potential romantic partner, in addition to how potential partners are sorted and prioritized, are value-laden decisions for which dating apps are accountable to their users. For example, dating apps use of race as a factor has come under criticism as prioritizing intra-racial dating and even segregating people by race or ethnicity – which has a long, problematic history in the United States (Bedi, 2019, 2022; Williams, 2024). These same value-laden recommendation and sorting decisions are posited as a competitive advantage and differentiator – dating apps claim to be *better* at facilitating a match – and it holds that these same dating apps are accountable for the moral implications of those same decisions.

Generative NLP Models (v. Search).

Bert and Chat GPT offer another example of how firms may mistakenly not acknowledge accountability for their design decisions – particularly when used to augment search. These tools create new content using generative language processing models. So, where a search engine allows users to enter prompts and then provides related links to relevant content (websites, government documents, etc), Chat GPT *creates* new content to answer the same type of prompt provided by the user while also providing links to original content. The new content is generated by Chat GPT, and Chat GPT was created by ‘learning’ different types of content deemed acceptable by Open AI (owners of Chat GPT). Open AI had to decide which type of content was appropriate to include in training data and input to Chat GPT’s models. And Open AI decided the use cases for their technology as well as how ‘answers’ are presented to users. Also, Open AI decided whether source material was also provided to users or not.

While the use of LLM in search is positioned as mere facilitator, similar to a search platform, LLM design decisions are above and beyond the decisions that a search engine would have to make. As such, Open AI, through

the design and use of Chat GPT, makes more value-laden decisions than a search platform and even claims to have a competitive advantage based on those previous and current decisions. OpenAI should also be accountable for these same value-laden decisions.

Conclusion

This chapter has sought to illustrate how Johnson's algorithmic accountability as practice sheds light on why, to whom, and for what firms are accountable when it comes to AI. Firms are accountable to stakeholders because their decisions have moral implications for stakeholders. Design and development decisions around AI, digital platforms, facial recognition, etc. create value for stakeholders including the firm. These decisions are why users flock to their technology, the rationale for charging users and customers for their technology, and the argument for why people should invest in their company: because these firms are making hard, important decisions in the design and development of their technologies and those decisions lead to positive outcomes. Firms are not reluctant to being accountable for the good outcomes of design and development decisions and make sure that those decisions are attributable to their unique capabilities. However, these decisions also undermine rights, marginalize vulnerable stakeholders, foster abuse, have negative consequences, break norms, rules, and laws, and create unfair systems and outcomes. Within this practice of algorithmic accountability, more work needs to be done to (a) not take seriously tech firms claims they are not accountable for the moral implications of their decisions and (b) help those same firms work through the value-laden design decisions as well as identify the moral implications of those decisions.

References

- Abebe, R., Barocas, S., Kleinberg, J., Levy, K., Raghavan, M., & Robinson, D. G. (2020). Roles for computing in social change. *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*, 252–260.
- Ackermann, R. (2022, October 17). The growing Influence of Effective Altruism. *MIT Technology Review*. <https://www.technologyreview.com/2022/10/17/1060967/effective-altruism-growth/>
- Akrich, M. (1992). The De-Description of Technological Objects. In W. Bijker & J. Law (Eds.), *Shaping Technology/Building Society: Studies in Sociotechnical Change* (pp. 205–224). MIT Press.
- Bedi, S. (2019). Online Dating Sites as Public Accommodations: Facilitating Racial Discrimination. In S. Brison & K. Gelber (Eds.), *Free Speech in the Digital Age*. Oxford University Press.
- Bedi, S. (2022). Sexual Racism. In *The Palgrave Handbook of Sexual Ethics* (pp. 407–419). Springer.
- Benjamin, R. (2019). *Race After Technology: Abolitionist Tools for the New Jim Code*. John Wiley & Sons.
- Birhane, A., Kalluri, P., Card, D., Agnew, W., Dotan, R., & Bao, M. (2021). The Values Encoded in Machine Learning Research. *2022 ACM Conference on Fairness, Accountability, and Transparency*, 173–184. <https://arxiv.org/abs/2106.15590>
- Bovens, M. (2007). Analysing and assessing accountability: A conceptual framework 1. *European Law Journal*, 13(4), 447–468.
- Calvello, A. (2017). We Will Never Fully Understand How AI Works—But That Shouldn't Stop You From Using It. *Institutional Investor*. . <https://www.institutionalinvestor.com/article/2bstr1aauoix25yto45c/opinion/we-will-never-fully-understand-how-ai-works-but-that-shouldnt-stop-you-from-using-it>
- Chander, A., & Krishnamurthy, V. (2018). The myth of platform neutrality. *Geo. L. Tech. Rev.*, 2, 400.
- Cummings, M. L. (2006). Integrating ethics in design through the value-sensitive design approach. *Science and Engineering Ethics*, 12(4), 701–715.
- Dawkins, C. (2015). Agonistic Pluralism and Stakeholder Engagement. *Business Ethics Quarterly*, 1–28.
- Diakopoulos, N. (2020). Accountability, Transparency, and Algorithms. *The Oxford Handbook of Ethics of AI*, 17(4), 197.
- Eubanks, V. (2018). *Automating inequality: How high-tech tools profile, police, and punish the poor*. St. Martin's Press.
- Freeman, R. E. (1984). Stakeholder management: Framework and philosophy. *Pitman, Mansfield, MA*.
- Freeman, R. E. (2010). *Strategic management: A stakeholder approach*. Cambridge University Press.
- Freeman, R. E., Harrison, J. S., & Wicks, A. C. (2007). *Managing for stakeholders: Survival, reputation, and success*. Yale University Press.
- Freeman, R. E., Parmar, B. L., & Martin, K. (2020). *The Power of And: Responsible Business Without Trade-Offs*. Columbia University Press.
- Freeman, R. E., & Phillips, R. A. (2002). Stakeholder theory: A libertarian defense. *Business Ethics Quarterly*, 331–349.
- Friedman, B., & Nissenbaum, H. (1996). Bias in computer systems. *ACM Transactions on Information Systems (TOIS)*, 14(3), 330–347.

- Greenwood, M., & Mir, R. (2019). Critical management studies and stakeholder theory: Possibilities for a critical stakeholder theory. In *The Cambridge Handbook of Stakeholder Theory* (pp. 35–52). Cambridge University Press.
- Hamilton, I. A. (2022, February 16). Mark Zuckerberg’s new values for Meta show he still hasn’t truly let go of “move fast and break things.” *Business Insider*. <https://www.businessinsider.com/meta-mark-zuckerberg-new-values-move-fast-and-break-things-2022-2>
- Huckins, G. (2023, October 16). Minds of machines: The great AI consciousness conundrum: Philosophers, cognitive scientists, and engineers are grappling with what it would take for AI to become conscious. *MIT Technology Review*. <https://www.technologyreview.com/2023/10/16/1081149/ai-consciousness-conundrum/>
- Johnson, D. G. (2011). Computer Systems Moral Entities but Not Moral Agents. *Machine Ethics*, 168.
- Johnson, D. G. (2015). Technology with No Human Responsibility? *Journal of Business Ethics*, 127(4), 707–715. <https://doi.org/10.1007/s10551-014-2180-1>
- Johnson, D. G. (2021). Algorithmic accountability in the making. *Social Philosophy and Policy*, 38(2), 111–127.
- Johnson, D. G., & Miller, K. W. (2006). *A dialogue on responsibility, moral agency, and IT systems*. 272–276.
- Johnson, D. G., & Powers, T. M. (2005). Computer systems and responsibility: A normative look at technological complexity. *Ethics and Information Technology*, 7(2), 99–107.
- Johnson, G. (2023). Are algorithms value-free? Feminist theoretical virtues in machine learning. *Journal Moral Philosophy*. <https://philpapers.org/rec/JOHAAV>
- Knight, W. (2017, April 11). The Dark Secret at the Heart of AI: No one really knows how the most advanced algorithms do what they do. That could be a problem. *MIT Technology Review*. <https://www.technologyreview.com/2017/04/11/5113/the-dark-secret-at-the-heart-of-ai/>
- Latour, B. (1992). Where are the Missing Masses? The Sociology of a Few Mundane Artifacts. In W. Bijker & J. Law (Eds.), *Shaping Technology/Building Society: Studies in Sociotechnical Change* (pp. 225–258). MIT Press.
- Maas, J. (2023). Machine learning and power relations. *AI & SOCIETY*, 38(4), 1493–1500.
- Mac, R., & Kang, C. (2021, October 3). Whistle-Blower Says Facebook ‘Chooses Profits Over Safety.’ *The New York Times*. <https://www.nytimes.com/2021/10/03/technology/whistle-blower-facebook-frances-haugen.html>
- Martin, K. (2019). Ethical Implications and Accountability of Algorithms. *Journal of Business Ethics*, 160(4), 835–850.
- Martin, K. (2022a). Algorithmic Bias and Corporate Responsibility: How companies hide behind the false veil of the technological imperative. In K. Martin (Ed.), *Ethics of Data and Analytics*. Taylor & Francis.
- Martin, K. (2022b). *Ethics of Data and Analytics*. Taylor & Francis.
- Martin, K. (2023a). Predatory predictions and the ethics of predictive analytics. *Journal of the Association for Information Science and Technology*, 74(5), 531–545.
- Martin, K. (2023b). Who Counts in Business Ethics. *Business Ethics Quarterly*, 33(1), 216–243.
- Martin, K. (2024). Platforms, Privacy, and the Honeypot Problem. *Harvard Journal of Law & Technology*.
- Martin, K., Hong, G., & Easley, R. (2024). When Platforms Act Opportunistically: The Ethics of Platform Governance. *Working Paper*. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4202821
- MIT Technology Review Insights. (2023, October 5). “Transformative” AI will lead to rapid 25% efficiency gain, according to 81% of tech execs surveyed for new MIT Technology Review Insights research report. *PR Newswire*. <https://www.prnewswire.com/news-releases/transformative-ai-will-lead-to-rapid-25-efficiency-gain-according-to-81-of-tech-exec-surveyed-for-new-mit-technology-review-insights-research-report-301948000.html>
- Phillips, R. A. (1997). Stakeholder theory and a principle of fairness. *Business Ethics Quarterly*, 7(1), 51–66.
- Phillips, R., Freeman, R. E., & Wicks, A. C. (2003). What stakeholder theory is not. *Business Ethics Quarterly*, 479–502.
- Poole, S. M., Grier, S. A., Thomas, K. D., Sobande, F., Ekpo, A. E., Torres, L. T., Addington, L. A., Weekes-Laidlow, M., & Henderson, G. R. (2021). Operationalizing critical race theory in the marketplace. *Journal of Public Policy & Marketing*, 40(2), 126–142.
- Rosenblat, A., & Stark, L. (2016). Algorithmic labor and information asymmetries: A case study of Uber’s drivers. *International Journal of Communication*, 10, 27.
- Torres, E. (2021, October 21). Against longtermism. *Aeon*. <https://aeon.co/essays/why-longtermism-is-the-worlds-most-dangerous-secular-credo>
- Van Buren III, H. J. (1999). If Fairness is the Problem, Is Consent the Solution? Integrating ISCT and Stakeholder Theory. *Academy of Management Proceedings*, 1999(1), C1–C6.
- Villegas-Galaviz, C., & Martin, K. (2023). Moral distance, AI, and the ethics of care. *AI & Society*, 1–12.
- Williams, A. (2024). Not My Type: Automating Sexual Racism in Online Dating. In *Not My Type*. Stanford University Press.

