



*Extended Abstract*

## **Big Data, Corporate Governance, and the Limits of Algorithmic Accountability**

**Gernot Rieder** <sup>1,2,\*</sup>, **Judith Simon** <sup>1,2</sup>

<sup>1</sup> University of Vienna, Universitätsstraße 7, AT-1010 Vienna

<sup>2</sup> IT University of Copenhagen, Rued Langgaards Vej 7, DK-2300 Copenhagen S

E-Mails: [grie@itu.dk](mailto:grie@itu.dk); [jusi@itu.dk](mailto:jusi@itu.dk)

\* Author to whom correspondence should be addressed; Tel.: +4369910222204

*Accepted:*

---

In our increasingly datafied societies, algorithms play an ever more important role. Private companies such as Google, Facebook, or Amazon use algorithmic operations to steer information flows, rank content, strategically place product ads, and predict future user behavior. As scholars have argued, such algorithms are neither neutral nor objective, but the result of subjective interpretations and decisions, choices and classifications, potentially giving way to conscious or unconscious discrimination and bias (see boyd and Crawford 2011; Barocas and Selbst 2015). However, if algorithms are formative rather than descriptive, not just depicting realities but actively producing them by choosing "which information is "best"" (Mayer 2010), shouldn't the modalities of these realities, i.e. their ingrained assumptions and actual social consequences, be subject to scrutiny and critical reflection? Put differently: If people's actions and experiences are to a growing extent defined by and mediated through algorithmic processes, shouldn't we be interested in establishing ways to hold these algorithmic systems accountable?

As recent legal (Lunden 2015) and regulatory (Scott 2015) disputes have demonstrated, achieving algorithmic accountability that increases transparency and renders online intermediaries' automated decision-making procedures answerable is a challenging undertaking, for several reasons:

First, there is the problem of *secrecy*. Internet companies tend to be extremely tight-lipped about details on their algorithmic formulas, arguing that any public disclosure would allow spammers to "game the system" (Mayer 2010) and manipulate results. What's more, however, is that algorithms such as Google's Page Rank are core elements of a platform's product, which is why the high level of secrecy is also designed to maintain a competitive advantage and prevent rival companies from copying methods and building upon them (see Pasquale 2013).

Second, even if the details of a specific algorithm were made accessible and the necessary technical expertise to investigate could be mustered, chances are that a 'smoking gun', i.e. evidence of 'hardcoded' bias or discrimination, could not be readily found. This is because algorithmic systems do not function as standalone boxes, but as networked sociotechnical assemblages that include a multitude of human and non-human actors, with people debating models, setting target goals, cleaning training data, adjusting parameters, and choosing the specific context of application (see Gillespie 2014). Hence, algorithmic accountability is also difficult to achieve because algorithmic systems are fundamentally *complex*.

A third concern is the issue of *speed*. Online companies are continuously engaged in updating and tuning existing algorithms, testing and implementing new ones, abolishing those that have proven ineffective. Much of this goes unnoticed by the user, some of it, such as Facebook's 'emotional contagion' study (Kramer et al. 2014), gains public attention. If certain lines have been crossed and there is a backlash, organizations are usually quick to apologize, conceding that they "did a bad job" and "really messed up" (Isaac 2014). Features then disappear, but the tinkering with similar products often continues. In such an experimental space of fast-paced hit-and-miss, accountability concerns seem downright outdated. After all, how to oversee and regulate what is so elusive?

All of these factors contribute to the notion of algorithms as opaque, inscrutable artifacts and have arguably led to a veritable crisis of three major actors that can demand accountability, namely the media, government bodies, and jurisprudence. Hampered by a lack of data, expertise, and technical skills, and substantially affected by the progressing computer- and datafication themselves, these institutions have not only been largely toothless in controlling and governing rapidly changing digital markets, but have also, voluntarily or not, contributed to the spread of algorithmic opacity. The result is a private sector dominated by a few increasingly powerful organizations that are capable of monopolizing vast parts of the available online data, thereby further strengthening their market position and becoming obligatory passage points (Callon 1986) in a networked world.

Given this somewhat troubling state of affairs, is there any possibility to obtain or at least support some form of algorithmic accountability? How can we begin to address and solve these issues in a productive manner?

While it is certainly possible to improve the practices of each of the three watchdog institutions referred to above – the continuing antitrust investigations of Google by U.S. and European authorities and attempts to improve skills in "algorithmic accountability reporting" (Diakopoulos 2014) seem to point in the right direction – there may indeed be need to consider some more far-reaching solutions and alternatives. The paper will address both small- and large-scale solutions, seeking to provide ideas for a more effective and democratic governance of digital (data) markets.

## Acknowledgments

The authors wish to acknowledge the financial support of the Austrian Science Fund (P-23770).

## References and Notes

Barocas, Solon; Selbst, Andrew D. (2015) *Big Data's Disparate Impact*. Draft Version. Available at: <http://ssrn.com/abstract=2477899>.

- Boyd, Danah; Crawford, Kate (2012) "Critical Questions For Big Data: Provocations for a Cultural, Technological, and Scholarly Phenomenon." In: *Information, Communication & Society*, Vol. 15, No. 5.
- Callon, Michel (1986) "Elements of a sociology of translation: Domestication of the Scallops and the Fishermen of St Brieuc Bay." In: Law, John (Ed.) *Power, Action and Belief: A New Sociology of Knowledge?* London: Routledge, pp. 196-233.
- Diakopoulos, Nicholas (2014) "Algorithmic Accountability." In: Digital Journalism. Available at: <http://www.tandfonline.com/doi/abs/10.1080/21670811.2014.976411>
- Gillespie, Tarleton (2014) *Algorithm*. Draft Paper. Available at: <http://culturedigitally.org/2014/06/algorithm-draft-digitalkeyword/>.
- Kramer, Adam D. I.; Guillory, Jamie E.; Hancock, Jeffrey E. (2014) "Experimental evidence of massive-scale emotional contagion through social networks". In: *Proceedings of the National Academy of Sciences of the United States of America*, Vol. 111, No. 29.
- Lunden, Ingrid (2015) *Facebook's European Privacy Class Action Hearing Set For April 9*. Techcrunch article. Available at: <http://on.tech.ch/1/a5Uh>.
- Marissa, Meyer (2010) *Do Not Neutralise the Web's Endless Search*. Financial Times article. Available at: <http://on.ft.com/NDYPwz>.
- Pasquale, Frank (2013) "Paradoxes of Digital Antitrust: Why the FTC Failed to Explain Its Inaction on Search Bias." In: *Harvard Journal of Law & Technology Occasional Paper Series*. Available at: <http://jolt.law.harvard.edu/antitrust/articles/Pasquale.pdf>.
- Scott, Mark (2015) *E.U. Official Urges Google to Offer Greater Concessions in Antitrust Inquiry*. New York Times article. Available at: <http://nyti.ms/1wnrWdW>.