Regulation of algorithmic errors in digital commons platforms: managing the conformity-creativity conundrum

Erwan Joud LEGO, Univ. Bretagne Occidentale Erwan.Joud@univ-brest.fr
Nicolas Jullien LEGO, IMT-Atlantique Nicolas.Jullien@imt-atlantique.fr
Marine Le Gall LEGO, Univ. Bretagne Occidentale Marine.LeGall@univ-brest.fr

Abstract

Algorithmic management is blamed for its errors which would have discriminatory effects. Does Wikipedia do better in that matter? Through the analysis of the management of the most important bot fighting vandalism for the Fr-Wikipedia, we show that 1) over-standardization and discrepancies are hardly avoidable on the long run; 2) it is an issue for any platform, as it decreases its creativity and thus its attractiveness; 3) counterbalancing this is not one of technical limitations, but of socio-technical arrangements, on developing the human control and analysis of algorithmic decisions.

Keywords: Algorithmic management, Socio-technical Regulation, Conformity-creativity, Case study, Wikipedia

Introduction

In a digital platforms, the need of instantaneous and personalized interaction with thousands of users/contributors makes the automation of decision making a necessity (van Knippenberg et al., 2015). But, even if an algorithm may be efficient and even effective at a point in time, errors may always occur: either by misinterpreting the actions of the users, or because new actions, not learned because absent from the data occurs. It is difficult for the organizations and especially open, online platforms to detect and document its errors, because users may not make the effort to do so (they may simply drop out). From the management point of view, the reliance on algorithmic management may also leads to under-estimate the problems, because of two human factors (Markus, 2017): complacency (e.g., over-reliance on the program and its results due to low level of suspicion) and bias (i.e., tendency to ascribe greater power and authority to automated aids over other sources of information). The consequences is that algorithm based organizations are said to be less creative (Fügener et al., 2021), less able to detect new, interesting proposition for them. Have digital knowledge commons, and more precisely the one we have chosen to study, Wikipédia (with an 'é' as it looks at the Wikipedia's French project), found a way to solve, or at least mitigate the bad consequences of this algorithm management? This is what we studied here.

Methods

Considering the previous discussion, we performed a longitudinal case study the management of Salebot's management of contributions within Wikipédia. Active since 2007, Salebot detects and reverts "vandalism" on the French Wikipedia. It is one of the oldest and most influential bots of its kind out of all Wikipedia versions and has one of the richest user interfaces. Through the study of Salebot, we wanted to investigated how vandalism was defined by this algorithm, its developer, but also the community which has to analyze Salebot's actions, how a discrepancy between a rule and the algorithmic decision can occur (detected, signaled, analyzed and corrected). To do so we performed a digital ethnography (Kozinets, 2015) of the discourses, the exchanges about the bot, its activities, its decisions collecting the traces of exchanges regarding Salebot's actions in its page, in the regulated pages, but also in its developer's page and in the community's spaces. These discourses had three sources(de Vaujany et al., 2018): the questions, or the protest against the regulation put forward by the regulated; the answer to the regulated from the project managers (Salebot's developer or other wikipedians); the discussions which may follow-up. These elements were collected until saturation was reached (Glaser and Strauss, 2017), i.e. no new, different type or regulating episodes were found, thus over a period of ten years (2006 – 2015). We chronologically organized these elements and presented the text to a heterogeneous group of contributors for corrections and confirmations. The text, almost 70,000 characters long in French, was considered stable when no more modifications were applied by this group.

Results

Humans have delegated the "dirty work" of eliminating suspect contributions to a machine (Salebot means "malicious bot"). This organizational necessity because of the volume of vandalism, but also removes the need for policy makers to question whether this control goes against

the principle of freedom to contribute. But, as the decision space is so narrow and incapable of dealing with the unexpected, this algorithmic management increases organizational blindness to the evolution of needs, or the discrepancy between the rules and their algorithmic implementation. And errors or discrepancies may occur, i.e. good faith mistake (instead of vandalism), or even unexpected but valid new, original contributions. Also, the algorithmic control makes it difficult for the contributors to make sense of why they have been reverted. It is only when they make an extra effort to denounce a problem that the organization may be able to fix it.

But even when the effort is made, there are also traces of complacency and bias regarding the efficiency of the tool: negotiations focus on the role of a machine to control contribution over the rules and their justification ('Salebot was wrong', or 'Salebot was too rude', instead of 'I made a mistake, but you could have helped me improve my contribution instead of deleting it', for the users, or 'you did something the wrong way, but we can discuss about what you wanted to do and how to do it' for the supervisors). As far as Wikipedia in general is concerned, this echoes the ongoing debate between deletionists and inclusionists about project scope, and the algorithmic management reinforces the deletionist structural behaviors within the socio-technical organization.

Discussion/Conclusions

All this creates a situation in which, as is the case for private platforms, the algorithmic management reinforces the gap between simple contributors and those who are allowed to control them, the "policy-makers". This reinforces an organizational hierarchy characteristic of the commons, whether or not in digital form (Hess and Ostrom, 2007a).

However, saying that algorithmic management only enforces control, hierarchy and standardization would be to overlook the construction of a dual two-level control and autonomy system. First, more freedom to test the system gives newcomers more opportunities to learn the rules, although this may require more effort on their part. Second, because policy-makers are freed from the time-consuming basic control tasks, they still have time to collectively monitor the algorithmic management and its consequences, and to use it to discuss the evolution of the rules and find solutions to increase contributor, and thus contribution number and variety. It is also in these spaces where the tuning between conformity and creativity, control and autonomy is discussed.

These online discussion spaces represent the most innovative aspect of Wikipedia's algorithmic management (See Figure 1): the existence, within the information system, of places where controlees can provide feedback on any system – and therefore algorithmic control-related difficulties encountered to real humans, something that the participants in such systems hold dear(Newlands, 2022). By learning by discussing the rules they gain recognition as valuable participants and may give them access to the policy-making which remains open to anyone.

It may be less efficient as it requires a high level of involvement of both users and policy-makers, but we make the assumption that it is a necessity for any project based on algorithmic management, for it to maintain its exploration capacity, especially its capacity to welcome and retain newcomers, and thus its long-lasting effectiveness.

References

[de Vaujany et al.2018] François-Xavier de Vaujany, Vladislav V Fomin, Stefan Haefliger, and Kalle Lyytinen. 2018. Rules, practices, and information technology: A trifecta of organizational regulation. *Information Systems Research*, 29(3):755–773.

[Fügener et al.2021] Andreas Fügener, Jörn Grahl, Alok Gupta, and Wolfgang Ketter. 2021. Will humans-in-the-loop become borgs? merits and pitfalls of working with ai. *Management Information Systems Quarterly (MISQ)-Vol*, 45.

[Glaser and Strauss2017] Barney G Glaser and Anselm L Strauss. 2017. *Discovery of grounded theory: Strategies for qualitative research*. Routledge.

[Hess and Ostrom2007a] Charlotte Hess and Elinor Ostrom. 2007a. Introduction: An Overview of the Knowledge Commons. In (Hess and Ostrom, 2007b), editor, *Understanding Knowledge as a Commons. From Theory to Practice*, pages 3–26.

[Hess and Ostrom2007b] Charlotte Hess and Elinor Ostrom, editors. 2007b. *Understanding Knowledge as a Commons. From Theory to Practice*. MIT Press, Cambridge, MA.

[Kozinets2015] Robert Kozinets. 2015. Management netnography: axiological and methodological developments in online cultural business research. *Handbook of qualitative business and management research methods*.

[Markus2017] M Lynne Markus. 2017. Datification, organizational strategy, and is research: what's the score? *The Journal of Strategic Information Systems*, 26(3):233–241.

[Newlands2022] Gemma Newlands. 2022. Anthropotropism: Searching for recognition in the scandinavian gig economy. *Sociology*.

[van Knippenberg et al.2015] Daan van Knippenberg, Linus Dahlander, Martine R. Haas, and Gerard George. 2015. Information, attention, and decision making. *Academy of Management Journal*, 58(3):649–657.

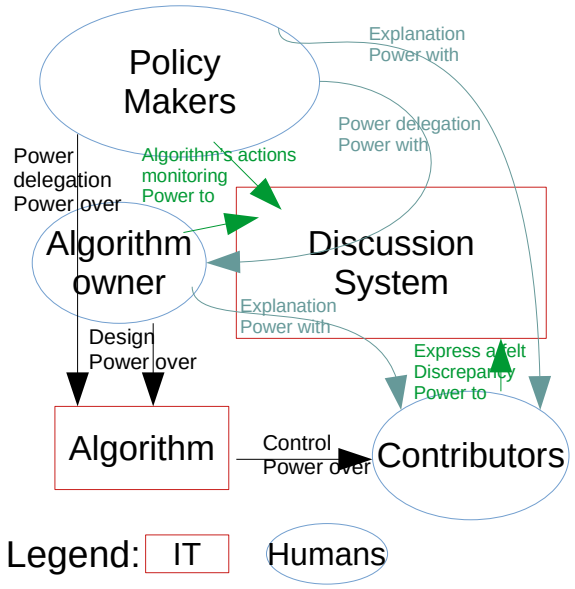

Figure 1: Wikipédia's Information System for Algorithmic management regulation