Personalisation algorithms and elections: Breaking free of the filter bubble

Milan, S.; Agosti, C.

Publication date
2019
Document Version
Final published version

Personalisation algorithms and elections: breaking free of the filter bubble

Stefania Milan, *University of Amsterdam, Amsterdam, Netherlands*
Claudio Agosti, *Algorithms Exposed, University of Amsterdam, Netherlands*

PUBLISHED ON: *07 Feb 2019*

Personalisation algorithms allow platforms to carefully target web content to the tastes and interests of their users. They are at the core of social media platforms, dating apps, shopping and news sites. They make us see the world as we want to see it. By forging a specific reality for each user, they silently and subtly shape customised "information diets", including around our voting preferences. We still remember Facebook's CEO Mark Zuckerberg testifying before the US Congress (in April 2018) about the many vulnerabilities of his platform during election campaigns. With the elections for the European Parliament scheduled for May 2019, it is about time to look at our information diets and take seriously the role of platforms in shaping our worldviews. But how?
Personalisation algorithms are kept a closely guarded secret by social media platform companies. The few experiments auditing these algorithms rely on data provided by platform companies themselves. Researchers are sometimes subject to legal challenges by social media companies who accuse them of violating the Terms of Services of their utility. As we speak, technological fencing-offs are emerging as

the newest challenge to third-party accountability. Generally, auditing algorithms fail to involve ordinary users, missing out on a crucial opportunity for awareness raising and behavioural change.

The Algorithms Exposed (ALEX) project [1], funded by a Proof of Concept grant of the European Research Council, intervenes in this space by promoting an approach to algorithms auditing that empowers and educates users. ALEX stabilises and expands the functionalities of a browser extension - *fbtrex* - an original idea of lead developer Claudio Agosti. Analysing the outcomes of Facebook's news feed algorithm, our software enables users to monitor their own social media consumption, and to volunteer their data for scientific or advocacy projects of their choosing. It also empowers advanced users, including researchers and journalists, to produce sophisticated investigations of algorithmic biases. Taking Facebook and the forthcoming EU elections as a test case, ALEX unmasks the functioning of personalisation algorithms on social media platforms.

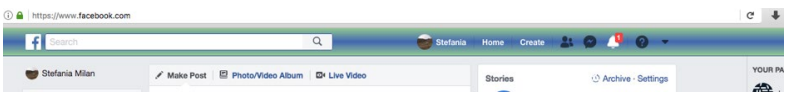

FIGURE 1: When active, fbtrex changes the colour of your Facebook bar.

OUR EVIDENCE: FACEBOOK AND THE ITALIAN ELECTIONS

In December 2018, the *Guardian* published a

report on "how Italy's populists used Facebook to win power". Using data from the MediaLab at the University of Pisa, the journalists showed how the two front-runners, namely Matteo Salvini (Lega) and Luigi Di Maio (Five Star Movement) managed to bypass mainstream media coverage: together, they totaled 7,8 million Facebook likes and shares during the two-month electoral campaign. Most of the content consisted in "viral videos and personal, off-the-cuff live broadcasts", reaching users in the intimate space of a personal Facebook page.

However, the analysis published by the *Guardian* is based on the measured engagement of users with a given set of content. While this might be helpful to describe a political outcome, in our opinion it provides only a partial picture. In fact, user engagement is a metric produced by the mash-up of three diverse factors.

The first factor refers to how much a page (or a political candidate) produces. Often candidates like Matteo Salvini, who can count on a communication campaign and a team managing their social media accounts, produce more content than would be normally available to an individual alone. The second factor concerns Facebook's prioritisation algorithm, which has unknown reasons to prefer a given content over others, and to recursively proposing it to an individual—or on the contrary, hiding it from a certain user. The functioning of this algorithm is unclear, but we have reasons to believe it has a

strong influence on the outcome. The third factor refers to the actual engagement metric, namely what we can observe within a certain segment of the electoral body which interacts with its preferred candidate. These factors have distinct political implications—respectively, a candidate's accountability, intervention of opaque algorithms, and legitimate political interest—and cannot be reduced to engagement metrics alone.

We, too, conducted an experiment in the wake of the Italian general elections (March 2018). We used *fbtrex* and bots created *ad hoc* to collect evidence on social media manipulation. How did we proceed?

We selected 30 Facebook pages across the political spectrum, identifying six amongst the most active ones in five distinct political orientations, recording content by means of the now discontinued Facebook API. This allowed us to collect the posts selected by Facebook as evidence of the ongoing algorithmic curation. At the same time, we created six profiles (our "bots") which started following the pages, accessing the platform 13 times per day at given moments and automatically scrolling over the timeline, and distributing likes differently on content from distinct political orientations. All public posts in the "wall" of these profiles were collected to show users what Facebook selects from each of them. Comparing data from the different profiles, we got a glimpse of the

algorithmic selection. We could thus generate dedicated metrics to compare content types and post repetition, also in relation to content sources.

In a controlled environment, we could isolate emerging patterns among the timelines. For example, the percentage of content shown is quite stable for every profile: different users have a recurring and unique ratio. The individual with more pictures than text, for example, kept being fed more pictures. The percentage of posts displayed more than once in a given timeline, varies by profiles, and appears to be influenced by the "informative variety" a profile is exposed to.

For a more detailed visual account on how we did it, watch Claudio Agosti's presentation at last Chaos Computer Conference and/or check out the slides here. *The presentation was designed with a hacker audience in mind.*

RECLAIMING ALGORITHMIC SOVEREIGNTY

FIGURE 2: fbtrex slogan

In times of elections more than ever, it is of paramount importance to empower users to break free of their filter bubble. But reclaiming our own algorithmic sovereignty is no easy task. Here we suggest some key features algorithmic sovereignty projects like *fbtrex* should implement.

- First of all, algorithmic accountability efforts should move out of the realm of the experts - academics and the industry above all - to involve users. Self-determination should be a community endeavour, if we are to counteract the current algorithm hegemony, centralised in a handful of corporations.
- Algorithm literacy is key: users should be empowered to independently test the contours of their own filter bubble, to find out for themselves how algorithmic personalisation affects their digital experience. Putting easy-to-use algorithmic accountability tools in the hands of users means they can move away from the hypothetical into the real life, and judge by themselves. Enabling self-awareness, algorithm literacy promotes a healthy information diet and fosters a responsible use of social media.
- Any algorithmic accountability tool should analyse the algorithm and not track individual behaviour. In addition, because algorithmic personalisation becomes visible only by comparing individual data, such tools should be able to aggregate data from various users. It is paramount that data collection and data reuse protocols protect users and their data, while at the same time supporting data analysis in the public interest.

Users should have full control of data extraction patterns, and be able to decide at any time whether they intend to volunteer their data. They should also be able to withdraw their participation whenever they want.

- Finally, any algorithmic sovereignty tool should be open-source, in order to promote transparency in its functioning and enable others to check its functioning, evolve its functions, modify or customise it.

You can read how we implemented this within *fbtrex* here. We are currently looking for test groups in various EU countries. If you want to join the study, please write to info@algorithms.exposed. To find out more, visit our website and https://facebook.tracking.exposed.

Acknowledgement: This project is funded by the European Research Council (ERC) under the European Union's Horizon2020 research and innovation programme (grant agreement No 825974-ALEX, awarded to Stefania Milan as Principal Investigator). See algorithms.exposed. For additional information, contact info@algorithms.exposed

---

Footnotes

1. ALEX is a spin-off of the DATACTIVE project, investigating the evolution of citizen participation in the age of datafication.