European elections and algorithm influences: call to action

2019: we will be ready!   *document version 1.3 (4th November 2018)*

Summary

This is a pan-european effort aimed at one goal: observe narratives around the EU electoral campaigns to support and foster a deeper understanding of data exploitation models and algorithms by social media platforms such as Facebook.

Who this document is for

- **Political groups** that wish to develop and increase awareness on online political advertisements and targeting (We started with Facebook, but other social media platforms can be covered too, e.g. Youtube or Twitter);

- **NGOs and Internet freedom activists** who wish to enhance the impact of their activities, both locally and globally, by adding a powerful tool to their asset as to plan; GDPR compliance monitoring (e.g. Art 22?), strategic litigation, communication/awareness campaigns about social media monopolies and the exploitation of user profiling.

- **Journalists and researchers** who want to improve their understanding of the social dynamics developing outside of their immediate network of knowledge ("filter bubble")

The concept

facebook.tracking.exposed (fbTREX) is a project that adopts a **collective approach**.

Current social networks give each of us a personalized experience, which prevents an understanding of the complexity of the public debate, the extent of media manipulation, and the real impact of targeted advertising on our perception and, ultimately, our choices.

For these reasons, we are asking individuals to share what information Facebook feeds them with: the goal is to study social media and how they exert influence - not the individuals willing to use our tool.

Still, we recognise this information could contain [Personally identifiable information](), Since the only goal of this project is the collective interest, participants will always have control of their data. We don't have a business to develop, and there is no user profiling scheme behind us. Transparency and fairness are at the core of our vision, and we pledge to

collect and process PII with utmost care; exclusively in accordance with our Ethical code and data policy that you can find below.

Introduction to the "black box" approach

fbTREX is meant to be used by individuals or groups who want to investigate and attribute the appropriate responsibilities to the social network they use.

The logic and mechanics of social networks are secret, therefore we can only observe them from the outside and estimate them. This methodology, the so-called *black box analysis*, is one of the approaches we can use for algorithm analysis[1].

In fbTREX, we started observing how Facebook automatically disseminates information thanks to its social reach and footprint. We are currently focussing our analysis on Facebook timelines because they represent the product of algorithm prioritization and the leading place where Facebook exerts its influence[2].

An analysis of algorithm influence cannot be achieved by only observing one's own personal newsfeed/timeline, because the sequence and selection of contents displayed is the result of many variables. Among the main ones, we can identify:

1. the public figures and pages he/she follows (and what kind of content is published on those pages)

2. the profile of the user (who they are, and their centres of interest)

3. the algorithm and the platform logic (content moderation, engagement, advertising, experiments)

If on the one hand the black box analysis still allows us to observe variable 1 and 2, to get hold of variable n. 3, there are only two possible approaches: either gather a statistically significant sample of variables 1 and 2, and from their dynamics infer 3; or run synthetic tests, with variables 1 and 2 under our control. In either case, the goal is to isolate and analyse 3. By doing so, we will be able to assess and attribute the appropriate responsibility to the social network.

---

[1] Hamilton, K., Karahalios, K., Sandvig, C., & Eslami, M. (2014, April). A path to understanding the effects of algorithm awareness. In *CHI'14 Extended Abstracts on Human Factors in Computing Systems* (pp. 631-642). ACM,
[2] Hazelwood, K., Bird, S., Brooks, D., Chintala, S., Diril, U., Dzhulgakov, D., ... & Law, J. (2018, February). Applied Machine Learning at Facebook: A Datacenter Infrastructure Perspective. In *High Performance Computer Architecture (HPCA), 2018 IEEE International Symposium on* (pp. 620-629). IEEE

Campaign narratives

1. Collectively "free the data back to society."Once the copy gets done, we can try to apply our logic, algorithm, and politics.

2. "Unionize the users" This be a method to guide collective power? free their data from the Facebook walled garden as a form of civil disobedience, a possible beginning to claim/demands new collective rights.

3. Analyze, understand, visualize and explain the Facebook algorithm.

4. Experiment with privacy-preserving methods to access a database of public interest.

5. Get a grasp of the ongoing debate around "filter bubbles".

6. Monitor real-time advertising on Facebook, and do a qualitative assessment of what political advertising is. Facebook announced changes after the US presidential elections backlash, but they'll be inevitably biased (see our opinion piece[3], the VICE experiments prior to mid term US elections[4], and other issues[5])

Outputs

The project itself will produce data and engagement. The outputs are listed in order of likelihood: the more resources, partnerships and advocacy we will be able to gather, the more likely it is to tick off the whole list.

1. Analyzing how algorithms affect our perception.
   We have the expertise, we have funds: at the very least, this will be achieved.

2. Building a community of analysts and reporters.
   Algorithms are a complex topic which requires translation, visualization, and engaging stories. We aim at providing accurate data and tools to analysts, reporters, and expert users.

3. Analyze, re-publish and index the content published by key figures of the public debate. Likely it would include public figures, institutions, maybe a selected amount of content published by media and journalists.

4. Measuring misinformation around European Union elections and related hot topics, and identifying how and which political groups exploit the phenomenon to foster anti-European narratives.

---

[3] [Facebook implement Advertising API for researcher: spot the bias!](#)
[4] https://news.vice.com/en_us/article/wj9mny/facebooks-political-ad-tool-let-us-buy-ads-paid-for-by-mike-pence-and-isis
[5] https://twitter.com/ShashiChi/status/1059184113851215872

5. Monitoring political targeted advertising in the EU Member States and abroad.

Our Background

fbTREX has been successfully used in the last two years., and is intended to help users understand algorithm influence and agenda. Although it can be used to analyse any day to day activity, we organize tests and want to enable external research groups in doing the analysis..

The most recent success story concerns the Italian elections in March 2018; you can read more at the following publications:

Italy: Personal Data and Political Influence: a paper about the influence of the Italian electoral campaign. The algorithm analysis is on the final part of the document.

Biases in the Facebook News Feed: a Case Study on the Italian Elections: a peer-reviewed scientific article which explains how our tool can be used to keep the algorithm accountable through external measurement

The invisible curation of content | Facebook's News Feed and our information diets: co-authored with the Web Foundation, a report on tests conducted in Argentina (but replicable anywhere else).

Scientific and Amateur Analysis of the Facebook Algorithm: our speech at HOPE Conference, a detailed explanation of the test made during the electoral campaign and some of the insights gained

fbTREX, under the generalized name of Algorithms Exposed (ALEX), has been sponsored by the European Research Council with the University of Amsterdam (official press release). We were interviewed for Datactive to explain some of the individual rights we want to fight for.

Ethical code and data policy

We are asking individuals to share some of the data - not their personal data, but what Facebook gives them: the goal is to study how social media influences, not the subject participating.

Still, this information can contain a lot of PII. We don't have a business to develop, and there is no user profiling scheme behind us. Since the only goal of our data collection is the collective interest, transparency and fairness are two essential values. Project ethics clearly define what limits we impose on ourselves.

First limitation: we observe only the timelines, not individual profiles or pages. This is the difference between self-assessment and enabling social media intelligence[6], which could be an abusive practice and we do not want to enable it carelessly.

Second limitation: our servers only store public posts.

Third limitation: users who install the extension have full control on their data. They can delete what they submitted whenever they want.

Fourth limitation: third party access to user data is granted by each user on a case by case and strictly opt-in basis.

Fifth limitation: analysis that are run on the dataset will be aimed at understanding social phenomenons, not individual profiles. Since this can't be formally verified, we'll formulate and implement safeguards during the development of ALEX, and publish updates on this process.

We are writing a proper Data protection policy and DPIA. Get in touch to know more.

Contacts: claudio@tracking.exposed, make it circulate among allies!

---

[6] [Social media intelligence, the wayward child of open source intelligence](#).