

Algorithmic content selection in Switzerland – a study of Google and YouTube

Summary

*Institute of Communication and Media Studies, University of Bern

**Social Computing Group, University of Zurich

March 2023

Algorithmic content selection systems, such as Google search or YouTube recommendations, shape today's media ecosystem. By automatically filtering and prioritising information in response to user queries, these systems help individuals to navigate the abundance of available digital content. However, algorithms powering these systems can be subjected to the erratic performance that may result in them facilitating the spread of factually incorrect or misleading information, which can undermine trust in democratic institutions or convince individuals to make decisions harmful to their personal or collective well-being.

To examine the impact of algorithmic systems on exposure to misinformation in Switzerland, the project looked at Google and YouTube, two monopolist platforms in the search and video-hosting markets. It investigated what sources are prioritised by Google and YouTube in relation to COVID and the Holocaust, two topics which are often targeted with misinformation. The project analysed how algorithms' performance is influenced by 1) search query formulation (i.e. whether queries expressing interest in misinformation result in more misinformation exposure); 2) algorithmic personas (i.e. whether visiting websites with specific political leanings before searching leads to more misinformation exposure); and 3) time (i.e. whether exposure to misinformation changes over time).

The project used a virtual agent-based algorithm audit, which is a research technique that simulates human behaviour to generate inputs for the algorithmic system and then record outputs. It consisted of three waves of data collection in May and June 2022. In the case of Google Search and YouTube search, it collected the first 10 search results in response to 28 queries in German: 14 for the Holocaust and 14 for COVID, with half of the queries expressing interest in misinformation (e.g. "corona-impfung mikrochip") and half not expressing such interest (e.g. "corona-impfung"). In the case of the YouTube recommendation system, the project examined a set of recommendations for videos with an explicit interest in Holocaust/COVID misinformation and without it. To analyse the results, the trained team of assistants manually examined the full content of Google Search results and the first three minutes of videos from YouTube Search and recommendation system outputs.

The findings of the project suggest that both Google and YouTube tend to prioritise content coming from journalistic websites and channels in response to COVID- and Holocaust-related queries. Such content also usually comes from media outlets from outside Switzerland (usually Germany), whereas Swiss journalistic outlets appear in top outputs less commonly (with the exception of YouTube recommendation outputs). Similarly, educational sources prioritised by algorithms usually originate outside of Switzerland, unlike government/administrative sources which tend to be Switzerland-based.

Relatively few Google and YouTube outputs support misinformation, in particular in relation to the Holocaust; however, for COVID, such outputs appear more often (up to 40% of outputs for the queries "corona-diktatur" and "corona-lüge" for Google Search). This observation indicates that it is more difficult for algorithms to deal with developing misinformation topics. Misinformation was more present in Google Search compared with YouTube, despite YouTube having a reputation as a platform facilitating exposure to false information through its algorithms. It is important to note, however, that our analysis

is more likely to identify misinformation on Google than on YouTube due to YouTube content analysis focusing only on the first three minutes of each video.

In terms of specific factors influencing exposure to misinformation, the project found little influence of algorithmic personas, thus indicating that websites visited before searching are not necessarily taken into consideration by the algorithms consistently. The factors which matter more are the selection of the search queries (with queries with an explicit interest in misinformation resulting in more exposure to it) and the time at which the input was generated, as shown by the variation in the visibility of misinformation-related content.

Together, these findings demonstrate that under certain circumstances (in particular, when users signal interest in it), algorithmic systems can facilitate the exposure of Swiss citizens to false information. They stress the importance of increasing awareness of algorithmic misinformation in the Swiss population, the need to establish monitoring infrastructure to track how algorithms deal with misinformation over time, and fostering normative discussion on what the role of algorithms should be in the Swiss context.