

Sélection de contenu algorithmique en Suisse – une étude sur Google et YouTube

Résumé

Mykola Makhortykh*, Maryna Sydorova*, Aleksandra Urman**, Franziska Keller*, Silke Adam*

* Institut des sciences de la communication et des médias, Université de Berne

** Social Computing Group, Université de Zurich

Mars 2023

Les systèmes de sélection du contenu algorithmiques, comme la recherche Google ou les recommandations YouTube, façonnent l'écosystème de médias actuel. En filtrant et en priorisant automatiquement les informations fournies en réponse aux demandes de l'utilisateur, ils aident les utilisateurs à naviguer dans l'abondance de contenus numériques disponibles. Cependant, les algorithmes à la base de ces systèmes sont susceptibles de fournir des performances erratiques et donc de faciliter la propagation d'informations factuellement fausses ou trompeuses, ce qui peut affaiblir la confiance envers les institutions démocratiques ou convaincre des personnes à prendre des décisions dangereuses pour leur bien-être personnel ou celui de la collectivité.

Pour examiner l'influence des systèmes algorithmiques sur l'exposition à la désinformation en Suisse, le projet s'est intéressé à Google et YouTube, deux plateformes monopolistes sur les marchés de la recherche et de l'hébergement de vidéos. Il a cherché à déterminer quelles sources sont priorisées par ces deux plateformes sur les sujets du Covid et de l'holocauste, souvent ciblés par la désinformation. Il a également analysé comment la performance des algorithmes est influencée 1) par la formulation de la demande (p. ex. si les demandes exprimant un intérêt pour la désinformation entraînent une exposition plus prononcée à celle-ci), 2) par des profils algorithmiques (p. ex. si la consultation de sites web ayant une inclinaison politique spécifique avant la recherche entraîne une exposition plus prononcée à la désinformation) et 3) par le temps (p. ex. si l'exposition change avec le temps).

Le projet a recouru à un algorithme d'audit basé sur un agent virtuel, une technique de recherche simulant le comportement humain afin de générer des inputs pour les systèmes algorithmiques et enregistrer les résultats. Trois collectes de données ont été menées entre mai et juin 2022. Dans le cas de la recherche Google et YouTube, elles ont porté sur les 10 premiers résultats affichés en réponse aux 28 demandes effectuées en allemand: 14 pour l'holocauste et 14 pour le Covid, dont la moitié exprimaient un intérêt pour la désinformations (p. ex. "vaccin coronavirus micropuce") et l'autre moitié non (p. ex. "vaccination coronavirus"). En ce qui concerne le système de recommandation de YouTube, le projet s'est penché sur une série de recommandations faites pour des vidéos ayant un intérêt explicite pour la désinformation liée à l'holocauste et au Covid, et pour des vidéos n'en ayant pas. Afin d'analyser les résultats, l'équipe d'assistants spécialement formés a examiné manuellement la totalité du contenu des résultats de la recherche Google ainsi que les trois premières minutes des vidéos proposées par YouTube et par le système de recommandation.

Selon les conclusions du projet, pour répondre aux demandes concernant le Covid et l'holocauste, tant Google que YouTube ont tendance à prioriser des contenus provenant de sites et de canaux journalistique, généralement issus de sources en dehors de la Suisse (notamment l'Allemagne). Les sources journalistiques suisses apparaissent moins fréquemment dans les premiers résultats (à l'exception des résultats des systèmes de recommandation de YouTube). De même, les algorithmes tendent à privilégier les sources éducatives étrangères, contrairement aux sources gouvernementales ou administratives qui sont, elles, suisses.

Relativement peu de résultats Google et YouTube alimentent la désinformation, en particulier s'agissant de l'holocauste, même si la désinformation concernant le Covid est plus fréquente (jusqu'à 40% pour les demandes "dictature coronavirus" et "mensonges coronavirus" dans Google). Ce constat indique que les algorithmes ont plus de difficulté à gérer les sujets de désinformation en développement. La désinformation est plus présente dans la recherche Google que sur YouTube, bien que cette dernière ait la réputation de favoriser l'exposition aux fausses informations par le biais de ses algorithmes. Il convient toutefois de noter que notre analyse risquait davantage de trouver de la désinformation sur Google que sur YouTube, car elle portait uniquement sur les trois premières minutes de chaque vidéo.

Le projet n'a trouvé que peu d'influence de la part des profils algorithmiques sur l'exposition à la désinformation, ce qui indique que les sites web visités avant la recherche ne sont pas nécessairement pris en compte par les algorithmes. Les facteurs les plus influents sont la demande (les demandes avec un intérêt explicite pour la désinformation donnent lieu à plus d'exposition à celle-ci) et le moment de la demande, comme le montrent les variations dans la visibilité des contenus liés à la désinformation.

Ces résultats prouvent que dans certaines circonstances (en particulier si les utilisateurs manifestent un intérêt dans ce sens), les systèmes algorithmiques peuvent faciliter l'exposition des citoyens suisses aux fausses informations. Ils soulignent l'importance de sensibiliser davantage la population suisse à la désinformation algorithmique ainsi que la nécessité de créer une infrastructure chargée de suivre comment les algorithmes gèrent la désinformation au fil du temps et d'encourager un débat normatif sur le rôle des algorithmes dans le contexte suisse.