



**HAL**  
open science

## L’empreinte antisémite sur YouTube

Charles De Dampierre, Andreï Mogoutov, Benjamin Tainturier, Bilel Benbouzid, Dominique Cardon, Jean-Philippe Cointet, Caterina Froio, Alexis Perrier.

► **To cite this version:**

Charles De Dampierre, Andreï Mogoutov, Benjamin Tainturier, Bilel Benbouzid, Dominique Cardon, et al.. L’empreinte antisémite sur YouTube : Contribution d’une équipe de recherche (médialab-CEE-LISIS) au Rapport sur la lutte contre le racisme, l’antisémitisme et la xénophobie 2020 de la CNCDH. Commission nationale consultative des droits de l’homme. 2022. hal-03938696

**HAL Id: hal-03938696**

**<https://hal.science/hal-03938696>**

Submitted on 13 Jan 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Public Domain

## L'empreinte antisémite sur YouTube

Contribution d'une équipe de recherche (médialab-CEE-LISIS) au Rapport sur la lutte contre le racisme, l'antisémitisme et la xénophobie 2020 de la CNCDH

Juin 2021

**Charles de Dampierre (Sciences Po – médialab), Andreï Mogoutov (Sciences Po – médialab), Benjamin Tainturier (Sciences Po – médialab), Bilel Benbouzid (université Gustave Eiffel – LISIS), Dominique Cardon (Sciences Po – médialab), Jean-Philippe Cointet (Sciences Po – médialab), Caterina Froio (Sciences Po – CEE), Alexis Perrier (université Gustave Eiffel – LISIS).**

**I. Le terrain de la recherche : YouTube**

**Erreur ! Signet non défini.**

YouTube, la participation politique et la modération du discours de haine

**Erreur ! Signet non défini.**

Le périmètre de l'enquête : un corpus des commentaires sur les chaînes d'information et d'actualité sur YouTube

**Erreur ! Signet non défini.**

Une cartographie des chaînes d'information et d'actualité sur YouTube

**Erreur ! Signet non défini.**

**II. Méthode : Apprendre l'antisémitisme à un algorithme**

**Erreur ! Signet non défini.**

Détecter "l'empreinte antisemite" dans les commentaires YouTube à partir d'exemples annotés

**Erreur ! Signet non défini.**

Échantillonnage des commentaires annotés manuellement et établissement d'une grille de codage en sept catégories

**Erreur ! Signet non défini.**

Annotation de l'échantillon d'apprentissage et de l'échantillon de test

**Erreur ! Signet non défini.**

Plongement de mots et reconnaissance automatique de l'empreinte antisemite en ligne

**Erreur ! Signet non défini.**

**III. Une empreinte antisemite faible et inégalement distribuée**

**Erreur ! Signet non défini.**

Quelle est l'empreinte de l'antisémitisme dans les commentaires des chaînes de YouTube ?

**Erreur ! Signet non défini.**

Comment se distribue l'empreinte antisémitisme au sein de la carte de l'information et de l'actualité ?

**Erreur ! Signet non défini.**

Comment les différentes connotations de l'empreinte antisemite se distribuent-elles en fonction des territoires informationnels ?

**Erreur ! Signet non défini.**

Quels thèmes de vidéo attirent les commentaires antisémites ?

**Erreur ! Signet non défini.**

**IV. Judéophobie, antisionisme et complots à toutes échelles**

**Erreur ! Signet non défini.**

Les deux antisémitismes de l'extrême droite

**Erreur ! Signet non défini.**

L'antisionisme de gauche, l'antisionisme de droite

**Erreur ! Signet non défini.**

Vers une convergence des complots ?

**Erreur ! Signet non défini.**

**V. Conclusion**

**Erreur ! Signet non défini.**

**Bibliographie**

**Erreur ! Signet non défini.**

**Annexes méthodologiques**

**Erreur ! Signet non défini.**

## Introduction

Le développement des discours de haine sur internet est l'objet d'une attention publique de plus en plus grande et se trouve aujourd'hui au cœur des enjeux de régulation du numérique. Après que les espaces numériques ont été considérés positivement comme supports de nouvelles formes d'expressivité et d'action politique<sup>1</sup>, ils suscitent aujourd'hui des inquiétudes et des craintes. Contrairement aux attentes les plus optimistes, l'essor d'internet n'a pas toujours entraîné une augmentation substantielle des niveaux de participation du public<sup>2</sup>. L'anonymat et la disparition des coûts d'accès à la parole publique ont aussi libéré des propos faux, moqueurs ou haineux<sup>3</sup>. Ceux-ci concentrent aujourd'hui une grande partie de l'attention des médias et des recherches académiques sans que ne soit véritablement évaluée la prévalence du phénomène<sup>4</sup>. Dans ce contexte, la question spécifique de l'antisémitisme en ligne a été l'une des premières à faire l'objet d'attention et de vigilance<sup>5</sup>. Compte tenu des spécificités des mondes numériques, il est difficile de mesurer la prévalence de l'antisémitisme en ligne et de mettre en place des dispositifs d'objectivation comme la mesure des actes ou des opinions antisémites<sup>6</sup>.

Dans cette étude, nous proposons une approche *exploratoire* de cette question en nous focalisant sur les chaînes d'information et d'actualité de YouTube en France. La plateforme vidéo de Google est désormais le deuxième site le plus visité au monde avec 1,86 milliards de visiteurs en 2021 et occupe une place privilégiée dans l'espace numérique français : 70 % des 60,4 millions d'internautes français se rendent régulièrement sur YouTube<sup>7</sup>. Cette plateforme constitue une source originale de données pour les chercheurs en sciences sociales et computationnelles étudiant le « discours de haine ». Elle rassemble d'abord un ensemble de producteurs de discours : médias, ONG, partis et acteurs politiques, institutions, simples youtubeurs développant leur propre chaîne. Cet espace de parole élargi a favorisé la désinhibition et le désenclavement de la parole publique en démocratisant l'expressivité et la créativité mais en rendant aussi possibles des formes de discours douteux ou haineux. À travers les nombreux commentaires que laissent les internautes sous les vidéos, YouTube rend accessible un ensemble de réactions exprimant dans des registres de discours extrêmement variés leurs opinions à

---

<sup>1</sup> BENKLER Yochai, *The Wealth of Networks. How Social Production Transforms Markets and Freedom*, New Haven, Yale University Press, 2006 ; INGELHART Ronald, *Cultural evolution: people's motivations are changing, and reshaping the world*, Cambridge University Press, 2018.

<sup>2</sup> MARGETTS Helen,, "Political behaviour and the acoustics of social media", *Nature Human Behaviour*, 1(86), 2017, pp. 1–3.

<sup>3</sup> PHILLIPS Withney, MILNER Ryan, *The Ambivalent Internet. Mischief, Oddity and Antagonism Online*, Cambridge, Polity Press, 2017.

<sup>4</sup> SIEGEL Alexandra, "Online Hate Speech", in PERSILY Nathaniel, TUCKER Joshua A., éd., *Social Media and Democracy. The State of the Field and Prospects for Reform*, Cambridge, Cambridge University Press, 2020, pp. 56-88.

<sup>5</sup> GAGLIARDONE Iginio, GAL Danit, ALVES Thiago, MARTINEZ Gabriela, *Countering On-Line Hate Speech*, Paris, UNESCO Publishing, 2015.

<sup>6</sup> GHILES-MEILHAC Samuel, « Mesurer l'antisémitisme contemporain : enjeux politiques et méthode scientifique », *Revue d'histoire moderne et contemporaine*, n°62, vol. 2/3, 2015, pp. 201-224.

<sup>7</sup> Source : Statista - <https://www.statista.com/statistics/805656/number-youtube-viewers-worldwide/> ; Source : DATAREPORTAL 2019 : <https://datareportal.com/reports/digital-2019-france>.

propos des vidéos. Ce matériel discursif singulier présente l'intérêt d'être spontané et d'éviter certains écueils dont pâtissent d'autres méthodes. Les approches plus classiques sur ces thématiques par entretiens ou par questionnaires achoppent à une difficulté : certains interviewés ne reconnaissent pas le fait qu'ils adhèrent à des idées réprouvées<sup>8</sup>, phénomène d'auto-correction qu'il est difficile de redresser. Les commentaires à connotation antisémite publiés sur YouTube sont intentionnels. Ils sont à la fois des représentations et des actes discursifs exposés à la vue des autres. L'étude de ces commentaires ne permet cependant pas de mesurer la prévalence de l'antisémitisme dans notre société. Ils ne sont en rien représentatifs de la population des internautes et, moins encore, de la population française<sup>9</sup>. Dans cette étude, nous proposons d'explorer ces commentaires comme l'expression d'un public numérique participatif en cherchant à mesurer relationnellement leur présence dans les commentaires de telle ou telle chaîne. Inscrite dans le champ d'études de la sociologie politique et des méthodes numériques, cette recherche présente deux caractéristiques. La première est la définition d'un corpus spécifique qui servira de fond de carte à cette recherche. Indépendamment de la question de l'antisémitisme, nous avons délimité un corpus de 628 chaînes d'information et d'actualité sur YouTube dont nous avons échantillonné les commentaires des vidéos publiées dans les trois derniers mois de 2020. Au plus, deux-cents commentaires par vidéo ont été indexés pour constituer un corpus de 1 925 717 commentaires en tout. La deuxième caractéristique est le développement d'une méthode algorithmique de détection de l'antisémitisme utilisant des techniques d'apprentissage automatique permettant de traiter un volume important de données. Cette méthode ne permet pas de définir de façon précise et tranchée le caractère antisémite d'un énoncé comme un propos discriminant implicitement ou explicitement les Juifs, individuellement ou collectivement, sur la base de la religion, de l'appartenance ethnique, de l'ascendance ou de l'appartenance à un groupe d'influence. Les algorithmes de YouTube n'y parviennent pas non plus. En revanche, cette méthode automatique permet de détecter des énoncés qui, en raison de l'usage de certains termes, d'une syntaxe particulière ou d'une coloration énonciative, ont une très forte chance d'avoir une connotation antisémite. C'est pourquoi nous parlerons dans ce rapport d'une *empreinte antisémite* dont nous mesurerons l'incidence au sein de presque deux millions de commentaires d'internautes.

Les données numériques ne permettent pas de donner une vue « représentative » d'un phénomène. En revanche, à partir du corpus des chaînes d'information constitué comme fond de carte, il est possible d'étudier dans quelles parties de la carte les commentaires à empreinte antisémite sont les plus nombreux et de conduire, relationnellement, un certain nombre d'analyses. Les questions auxquelles cette étude cherche à répondre sont les suivantes :

- Quelle est la prévalence des commentaires à empreinte antisémite au sein des médias d'information et d'actualité de YouTube ?
- Quels sont les groupes de médias d'information et d'actualité qui attirent le plus fréquemment des commentaires à empreinte antisémite ?
- Quels sont les différents types d'antisémitisme prévalents sur YouTube ?

---

<sup>8</sup> MARGETTS Helen, "Political behaviour and the acoustics of social media", *Nature Human Behaviour*, 1(86), 2017, pp. 1–3.

<sup>9</sup> KOTRAS Baptiste, *La voix du web. Nouveaux régimes de l'opinion sur Internet*, Paris, Seuil/République des idées, 2018.

- Peut-on identifier des formes d'antisémitisme différentes dans les commentaires des différentes chaînes de médias d'information et d'actualité de YouTube ?
- Le développement de théories complotistes à la faveur de la crise de la Covid-19 a-t-il encouragé l'antisémitisme sur YouTube ?

Pour aborder ces questions, le rapport présente d'abord le terrain de l'étude (I). Par la suite, nous discutons de la méthode d'apprentissage supervisé que nous avons mise en œuvre pour quantifier et qualifier le contenu des chaînes YouTube (II). Enfin, nous présentons les résultats de l'analyse du contenu qui apparaît sur ces chaînes, pour étudier l'empreinte antisémite dans l'espace numérique français (III). Les conclusions offrent un résumé des résultats principaux de la recherche, ainsi que quelques suggestions pour des recherches et actions futures dans ce domaine. Plus de détails sur la méthodologie et les données de la recherche sont détaillés dans les encadrés et en annexes.