



Les sciences sociales en question :
grandes controverses épistémologiques et méthodologiques

Compte-rendu de la 57^e séance
**Comment mesurer l'antisémitisme
sur les réseaux sociaux ?**

16 novembre 2021

Nonna Mayer ouvre la 57^e séance du séminaire¹ en rappelant que Fariba Adelkhah est toujours retenue en Iran et ce depuis deux ans et demi.

Cette séance porte sur la mesure de l'antisémitisme sur les réseaux sociaux. Leur essor a ouvert un nouvel espace de communication et de débat, sans frontières, où propos racistes et antisémites se multiplient de manière inquiétante. Deux experts du numérique et des réseaux sociaux interviendront durant cette séance. Dominique Cardon et son équipe ont cartographié les manifestations d'antisémitisme sur YouTube en 2020, à partir d'un corpus de 628 chaînes d'information et de près de deux millions de commentaires. Cette enquête répond à une commande de la CNCDH

¹ Compte rendu de Justine Brisson, relu par les deux intervenants.

(Commission nationale consultative des droits de l'homme). Leur rapport peut être lu dans son intégralité soit sur le site du Médialab, soit sur celui de la CNCDH². Dominique Boullier discutera les problèmes méthodologiques que pose ce type d'enquête.

Dominique Cardon (Sciences Po, Médialab) est directeur du Médialab de Sciences Po. Il a notamment publié *Culture numérique*³ et coordonné l'enquête collective évoquée plus haut sur « L'empreinte antisémite dans l'espace YouTube français »⁴. Dominique Boullier (Sciences Po, CEE) a notamment publié *Sociologie du numérique*⁵ et « Comment sortir de l'emprise des réseaux sociaux : L'ère du réchauffement médiatique »⁶.

Dominique Cardon

Dominique Cardon commence par rappeler que l'enquête est le fruit d'un travail collectif, au sein d'une équipe comprenant Jean-Philippe Cointet (Sciences Po, Médialab), Andrei Mogoutov (Sciences Po, Médialab), Caterina Froio (Sciences Po, CEE, CNRS), Bilel Benbouzid (LISIS), Alexis Perrier (LISIS), Benjamin Tainturier (Sciences Po, Médialab), et Charles de Dampierre (ENS, Sciences Po, Médialab). Cette équipe a tenté de mettre en œuvre une technique relativement nouvelle dans les sciences sociales du numérique, qui correspond aux méthodes d'apprentissage automatique.

² Les liens vers le rapport : https://www.cncdh.fr/sites/default/files/rapport_racisme2020_basse_def.pdf et <https://medialab.sciencespo.fr/actu/cartographier-lempreinte-antisemite-sur-youtube/>.

³ Dominique Cardon, *Culture numérique*, Paris, Presses de Sciences Po, 2019.

⁴ Dominique Cardon et al. in CNCDH, *La lutte contre le racisme, l'antisémitisme et la xénophobie*, Année 2020, La documentation française, 2021 (Medialab/ CEE/Université Gustave Eiffel).

⁵ Dominique Boullier, *Sociologie du numérique*, Paris, Armand Colin, 2019.

⁶ Dominique Boullier, « Il est temps de réguler les réseaux sociaux : l'ère du réchauffement médiatique », *Cogito*, <https://www.sciencespo.fr/research/cogito/home/il-est-temps-de-reguler-des-reseaux-sociaux/>

Dominique Cardon justifie d'abord le choix de la plateforme numérique YouTube pour étudier l'empreinte antisémite sur Internet. Habituellement, on se tourne vers Twitter lorsque l'on s'intéresse aux données numériques mais YouTube est devenu absolument central en termes de fréquentation et d'audience (près de 46 millions de visiteurs chaque mois en France). Par ailleurs, YouTube joue un rôle majeur dans la production de désinformation ou la diffusion de messages à caractère antisémite. Le site a d'ailleurs été accusé il y a quelques années de produire de la radicalisation du fait de ses algorithmes de recommandation (effet *Rabbit Hole*)⁷.

Par ailleurs, Dominique Cardon précise un point important : YouTube a récemment fait disparaître un certain nombre de chaînes (25 000 au total), dont des chaînes problématiques (suprémacistes, antisémites, racistes...). Il y a eu ainsi une migration massive des producteurs de vidéos antisémites vers Odysee et d'autres plateformes vidéo d'extrême-droite. Par exemple, les chaînes liées à la galaxie autour d'Alain Soral et de « la dissidence » ne se trouvent plus sur YouTube. Il faut bien garder ce point en tête pour l'interprétation des résultats. Sur le plan méthodologique, Dominique Cardon rappelle qu'ils ont fait le choix de traquer l'antisémitisme non pas dans le contenu des vidéos mais plutôt dans les commentaires sous les vidéos. Il y a toutefois un angle mort dans cette étude : les chercheurs n'ont pas eu accès aux commentaires se trouvant sur les vidéos *live*, puisque ces derniers arrivent sur le site en temps réel. Cela est regrettable car les propos les plus virulents sont justement ceux présents dans le *chat*.

⁷ Sur cette question, voir Manuel Horta Ribeiro, Raphael Ottoni, Robert West, Virgilio Almeida, Wagner Meira, « Auditing radicalization pathways on YouTube », *FAT'20 Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*, janvier 2020, pp. 131-141.

Dans le but de cartographier un espace de l'information sur YouTube, Dominique Cardon et Bilel Benbouzid ont essayé de faire ce qu'ils appellent un « espace à la main ». Ils se sont intéressés à 628 chaînes françaises sur YouTube (actualité, information, commentaire et analyse politique, humour sur l'actualité, vulgarisation, etc.). Ils ont fait le choix de retenir aussi bien des chaînes liées à des grands médias, que des chaînes plus confidentielles. Ils se sont servis des classements déjà produits par des institutions de classements des chaînes en fonction de l'audience (Wizdeo). Au total, le corpus se compose donc d'environ 2 millions de commentaires.

Le deuxième temps de l'enquête a porté sur la formation de « territoires médiatiques » sur YouTube. Il s'agissait de cartographier cet espace de 628 chaînes d'information, en essayant de rapprocher celles où l'on retrouvait des commentaires de visiteurs d'autres chaînes. Cela permettait d'obtenir une sorte de visualisation des territoires de l'information à l'intérieur de ce vaste corpus. À droite du graphique se trouvent toutes les grandes chaînes de médias d'informations (du type *C à vous*) ; à gauche, toute la galaxie des chaînes liées de près ou de loin au parti France insoumise. En bas de la carte graphique, on trouve davantage de chaînes à vocation humoristique. La partie haute du graphique montre deux polarités : l'une renvoie à l'extrême droite traditionnelle (du type Action française ou Agora Vox) et une autre, plus difficile à qualifier, à ce qu'est la droite « nationale-populiste ». Cette dernière comporte une critique plus systémique, accompagnée de raisonnements économiques de type marxiste. Enfin, tout en haut de la carte on trouve un groupe parfois qualifié dans les médias de « complotiste ». Il recoupe aussi bien des chaînes sur la santé ou la médecine naturelle que d'autres sur les francs-

maçons, les martiens ou encore les Illuminati. Cette cartographie constitue donc la toile de fond sur laquelle l’empreinte antisémite est mesurée.

Dominique Cardon indique qu’il existe différentes formes d’antisémitisme. Une première exploration a consisté à retenir les commentaires autour d’un certain nombre de termes communs (antisionisme, nationalisme antisémite, numérologie sataniste, complots et lobbies internationaux, etc.). Le but était de voir se dessiner des géographies de sens associées au mot « juif ». Ensuite, six membres de l’équipe ont mis en place un jeu de données : il s’agissait pour chacun de traiter manuellement mille commentaires pour décider du caractère antisémite du commentaire. Cet exercice collectif a permis d’entraîner un algorithme à détecter les commentaires à « empreinte antisémite ». Dominique Cardon diffuse ensuite quelques exemples de commentaires afin de montrer comment les algorithmes attribuent à chaque commentaire une probabilité d’empreinte antisémite.

Le résultat de l’enquête révèle une proportion de propos antisémites dans les commentaires YouTube du corpus étudié de 0,41%. L’enquête montre que l’antisémitisme est bel et bien présent sur YouTube mais il n’est pas d’une virulence aussi importante que l’on aurait pu le croire en écoutant les médias, par exemple. Toutefois, 0,41% n’est pas non plus un résultat anodin ou mineur. Les territoires qui possèdent le pourcentage d’empreinte antisémite le plus élevé sont, sans surprise, en premier lieu ceux de l’extrême droite identitaire, puis ceux de la contre-information et des médecines et pratiques de santé alternatives, et enfin, en troisième position, ceux de la droite nationale-populiste. Afin d’obtenir une lecture plus fine des résultats, Dominique Cardon et son équipe ont comparé non seulement le pourcentage de l’empreinte antisémite par chaîne mais aussi

la visibilité de l’empreinte antisémite par chaîne. Il s’agissait de tenir compte de l’audience des chaînes.

Afin d’obtenir une compréhension qualitative des résultats, les membres de l’équipe ont tenté de former des clusters sémantiques à partir des principaux thèmes abordés par les commentateurs antisémites (soit une analyse à partir de 12 756 commentaires). Certains thèmes reviennent particulièrement souvent, comme la « dépravation des élites », le « complot mondial sataniste » ou le « nouvel ordre mondial ». Ensuite, il est possible de projeter l’intensité des thématiques en fonction des territoires médiatiques évoqués précédemment. Les chaînes de contre-information et de médecines et pratiques de santé alternatives, par exemple, recoupent beaucoup de thèmes liés au « grand complot mondial sataniste ».

Dominique Cardon conclut en mentionnant les prochaines étapes auxquelles lui et son équipe souhaiteraient s’atteler : densifier la cartographie des chaînes YouTube, associer le contenu des vidéos aux commentaires et travailler sur d’autres cibles de discours haineux que les juifs.

Dominique Boullier

Le premier point que souhaiterait rappeler Dominique Boullier, c’est que le travail qu’a fourni Dominique Cardon et son équipe est absolument énorme. Il considère que partir des commentateurs constitue une démarche originale, même si cela peut poser des problèmes méthodologiques.

Dominique Boullier estime que le résultat de 0,41% d'empreinte antisémite est plutôt rassurant. En même temps, il faut bien garder en tête qu'une grande partie du public antisémite a migré sur d'autres plateformes. L'espace médiatique bouge en permanence et produit un effet de « renforcement de la bulle de filtre », puisque les personnes qui sont sur des environnements trop « ouverts » ou tolérants préfèrent se retrouver sur des médias plus fermés où ils pourront reconstituer une communauté affinitaire avec des personnes ayant des opinions proches des leurs. L'espace public est donc quelque peu nettoyé des propos les plus extrêmes, ce qui n'est pas plus mal.

Dominique Boullier fait deux remarques d'ordre méthodologique. La première concerne la notion de fond de carte et la seconde celle d'empreinte antisémite. Il se demande si l'antisémitisme est une opinion comme les autres. Il rappelle qu'une partie des discours antisémites est illégale, de sorte que l'antisémitisme ne peut pas être considéré comme une opinion parmi d'autres. Dominique Boullier se demande si, d'une certaine façon, Dominique Cardon ne cherche pas à identifier une « disposition » plus qu'une « opinion »⁸. Dominique Boullier questionne ensuite la notion d'« empreinte » antisémite. Parler d'« empreinte » antisémite plutôt que de « discours » antisémite permet d'inclure ce qui est latent par rapport à ce qui est manifeste. Pourtant, il existe des situations où les indices sont très faibles. Dans ce cas, les algorithmes vont avoir tendance à positiver la valeur des mots, notamment lorsqu'ils fonctionnent en mode boîte noire, ce qui devrait questionner les sciences sociales. Or, dans les commentaires, il s'agit bien de situations de conversations. Il y a donc nécessairement une dimension

⁸ Une « disposition » est davantage liée à une position dans l'espace médiatique cartographié qu'une « opinion » qui est obtenue en respectant les conventions des sondages.

communicationnelle qui complexifie l'identification de l'empreinte antisémite par les algorithmes. Que faire des déclarations dans les commentaires qui ne sont pas directement antisémites mais qui le deviennent par interprétation ? Il y a là un véritable enjeu politique et méthodologique.

Pour illustrer son propos, Dominique Boullier donne l'exemple d'une personne qui tenait une pancarte « Mais qui ? » dans une manifestation antivax⁹. Cette photo a beaucoup circulé dans les médias d'informations qui ont expliqué qu'il y avait là une connotation antisémite. Or, au premier abord, Dominique Boullier explique qu'il ne l'avait pas perçu ainsi, car il ne connaît pas l'histoire associée à ce « qui ? ». Il n'avait donc pas perçu la connotation antisémite avant que les médias la lui signalent. Certes, dans un réseau de références partagées, dans un réseau de connivences, ce « qui ? » fait sens en tant que vocable antisémite mais cela n'est pas automatique et cela peut poser un problème pour le traitement des données par les algorithmes.

Réponse de Dominique Cardon

Dominique Cardon remercie Dominique Boullier pour ses remarques très éclairantes. Plus que des opinions, sans doute faut-il parler d'« énoncés ». Il en convient, les structures de communications ne sont jamais évidentes. Toutefois, l'intérêt de ces dispositifs, c'est que les communications des utilisateurs y sont « spontanées », dans le sens où elles ne sont pas sollicitées par une demande extérieure. Toutefois, le chercheur ne prétend

⁹ Manifestation de Metz le 7 août 2021. La pancarte était brandie par une enseignante ex-membre du Front national, Cassandre Fristot. Le « Mais qui ? » était entouré d'une dizaine de noms dont ceux de George Soros, Jacques Attali, Patrick Drahi, Bernard-Henri Lévy, Laurent Fabius, Jérôme Salomon mais lors d'autres manifestations, ces noms n'apparaissaient pas.

en aucun cas que les résultats reflètent une « opinion » du type de celle que l'on cherche à mesurer dans une enquête par sondage.

Concernant l'implicite dans les commentaires, Dominique Cardon précise que la grande majorité des commentaires est transparente et que par conséquent, les annotateurs n'ont aucun problème à identifier l'empreinte antisémite. Toutefois, il est vrai que dans les discours de haine plus « lettrés », le ton est beaucoup plus euphémisé, ce qui complique probablement la détection de l'empreinte antisémite par les logiciels algorithmiques. Ainsi, le « qui ? » auquel Dominique Boullier faisait allusion n'aurait certainement pas été détecté.

Question de Martine Cohen

Martine Cohen (sociologue émérite au Groupe sociétés, religions, laïcités) souhaite savoir s'il faut inférer du pourcentage faible (0,41% d'empreinte antisémite) que l'antisémitisme n'est pas si répandu que cela en France. Deuxièmement, elle se demande si on observe des différences qualitatives entre les commentaires des chaînes plus confidentielles et celles plus *mainstream* (comme le *Huffington Post*, par exemple). Troisièmement, elle souhaite savoir si des chaînes de grands médias ont répondu à ces commentaires antisémites et s'il existe des modérateurs.

Réponse de Dominique Cardon

Au premier abord, Dominique Cardon a lui aussi estimé que 0,41% d'empreinte antisémite était un résultat plutôt faible. Par rapport à ce qu'il avait anticipé, il lui semblait que les résultats relativisaient l'idée selon laquelle l'antisémitisme était très présent sur YouTube. Pourtant, avec du

recul, il réalise que le phénomène est loin d'être négligeable. En l'état, Dominique Cardon ne peut répondre à la deuxième question de Martine Cohen car l'étude s'est faite à une échelle encore trop globale. Adopter une échelle plus fine est nécessaire pour répondre à cette question, ce qui n'a pas encore été fait mais le sera peut-être. Concernant les modérateurs, il ne sait pas vraiment s'il en existe, y compris sur les chaînes des grands médias. Dans tous les cas, s'il y en a, ils ne sont pas très efficaces puisqu'ils laissent filtrer de nombreux commentaires antisémites. Le chercheur précise que sur la chaîne de Thierry Casasnovas en revanche, ils n'ont trouvé aucun commentaire antisémite contrairement à ce qu'ils avaient anticipé. Ce n'est qu'une hypothèse mais on peut supposer qu'il y a eu de la modération.

Question de Marie-Anne Matard-Bonucci

Historienne de formation, Marie-Anne Matard-Bonucci (Paris 8, CHRIPA) se demande s'il existe des travaux permettant de mieux comprendre quel est le profil des commentateurs. Ensuite, elle fait remarquer que si 0,41% d'empreinte antisémite peut sembler être un faible pourcentage, il ne faut pas oublier qu'il ne s'agit que de la partie émergée de l'iceberg. Tous les visionneurs ne postent pas de commentaires. Enfin, elle se demande s'il y a des pistes pour comprendre ce retour de la figure de Satan dans les discours antisémites, alors même que Satan avait quasiment disparu de la rhétorique antisémite depuis la fin du XIX^e siècle.

Réponse de Dominique Cardon

Il existe très peu de travaux sur les commentateurs. Dominique Cardon est tout à fait d'accord avec Marie-Anne Matard-Bonucci concernant le fait

que les commentateurs ne représentent qu'une petite partie des visionneurs.

Réponse de Benjamin Tainturier à Marie-Anne Matard-Bonucci

Benjamin Tainturier (Sciences Po, Medialab) revient sur la figure énigmatique de Satan qui imprègne largement la rhétorique antisémite. Il considère que la figure de Satan joue un rôle important dans la façon dont fonctionnent narrativement les théories du complot. Elle fonctionne en mettant une sorte de cran d'arrêt face aux différentes théories complotistes. De plus, les théories du complot possèdent fréquemment un arrière-fond apocalyptique et la figure de Satan apparaît comme une puissance permettant de former narrativement un discours total, à la fois totalisant et totalisateur.

Questions de Nonna Mayer

Nonna Mayer considère qu'il faudrait comparer l'empreinte antisémite de YouTube à celle d'autres types de discours haineux (par exemple concernant les Roms, les asiatiques, etc.). Il pourrait être intéressant de s'intéresser à des médias beaucoup moins souvent censurés, comme 4chan ou TikTok chez les jeunes utilisateurs. Nonna Mayer se demande aussi comment la distinction délicate entre antisémitisme et antisionisme a été traitée. Et, rappelant Gabriel Tarde, elle indique qu'il lui semble que nous sommes bien face à des conversations, là où précisément se forment des opinions.

Réponse de Dominique Cardon

Dominique Cardon rappelle qu'il faut être très prudent avec la prétention des énoncés du web à représenter l'opinion. Toutefois, y chercher l'empreinte antisémite reste intéressant car elle mesure les conditions de licéité qui sont offertes à un certain type de discours et l'engagement des personnes à occuper ces positions discursives. Cela dit nécessairement quelque chose de l'opinion, du moins dans une acception tardienne. Toutefois, il ne faudrait pas considérer pour autant les énoncés du web comme des réponses à des enquêtes d'opinion. En outre, nous n'avons aucune information disponible sur les informateurs. Comparer l'empreinte antisémite avec des discours de haine visant une autre population permettrait d'avoir des effets de contraste tout à fait stimulants. La distinction antisémite/antisioniste est quant à elle très délicate. Dans leur annotation, les chercheurs ont dû faire des réglages pour que les deux concepts ne soient pas automatiquement amalgamés.

Question de Jean-Pierre Raoult

Jean-Pierre Raoult (IREM) est membre de la CNCDH comme Nonna Mayer. Il revient sur le taux de 0,41% d'empreinte antisémite dans les commentaires. Il note que dans les discussions, un glissement semble s'opérer entre « commentateurs » et « commentaires ». Or, précise-t-il, 46 millions de commentaires, ce n'est pas 46 millions de commentateurs. Certaines personnes sont coutumières des commentaires et d'autres qui n'en postent jamais. Aussi, il se demande combien de commentateurs se cachent réellement derrière les commentaires. Jean-Pierre Raoult revient aussi sur la distinction faite par Dominique Boullier entre « acte de communication » et « acte de conversation ». En l'occurrence, il se

demande si les commentaires sur YouTube ne se situent pas précisément dans une zone grise et ne constituent pas à la fois des « actes de communication » et des « actes de conversation ».

Réponse de Dominique Cardon

Dominique Cardon explique qu'une partie des débats qui se développent dans les médias s'appuient sur des phénomènes qui ont émergé préalablement sur Internet. Cela pose un problème car ces discours, très proliférants, déforment parfois les préoccupations réelles des Français. Dominique Cardon est d'accord avec le fait que certains commentateurs produisent de nombreux commentaires et qu'il peut donc y avoir beaucoup d'énoncés provenant d'un seul et même commentateur. Il n'existe pas de chiffres précis pour quantifier le ratio commentateur/commentaire.

Question de Mathieu Brugidou

Mathieu Brugidou (EDF/R&D, Pacte) s'interroge sur le pourcentage d'erreurs (c'est-à-dire le désaccord entre les algorithmes et les annotateurs) dans les commentaires antisémites. Ce pourcentage d'erreur varie-t-il selon le type de médias ? Est-il possible d'extraire une connaissance supplémentaire de ces énoncés mal classés en les analysant à la main, qualitativement ?

Réponse de Dominique Cardon

Dominique Cardon explique qu'ils n'ont malheureusement pas travaillé sur les faux positifs et qu'il ne peut par conséquent pas en tirer d'enseignement. Il apparaît souvent que les faux positifs sont le fait d'un

problème de valence : tous les commentaires qui mentionnent le terme juif ne sont pas antisémites. Certains propos positifs sont parfois codés négativement par l'algorithme.

Question de Manon Berriche

Manon Berriche (Sciences Po, Medialab) s'interroge sur l'idée d'une « réunification » des complots. Elle se demande si le propre de l'énonciation complotiste n'est pas de désigner plusieurs cibles sous un même énoncé. Cet aspect « pêle-mêle » des discours complotistes ne vient-il pas complexifier l'analyse des discours de haine ?

Réponse de Dominique Cardon

Dominique Cardon rappelle que l'hypothèse de la « réunification » des théories complotistes ne vient pas spécifiquement de leur propre étude. C'est une hypothèse mise en avant par une série d'articles de presse. Ce sont les médias qui ont considéré qu'il y avait une réunification. En reprenant l'expression de Luc Boltanski, Dominique Cardon considère des chaînons de causalité sont manquants lorsque Bilderberg, Bill Gates, Israël, les extraterrestres, les Templiers, etc. sont réunis dans un seul et même sac. D'où la figure de Satan comme grand unificateur de ces discours hétéroclites.

En clôturant la séance, Nonna Mayer remercie Dominique Cardon et Dominique Boullier pour leurs interventions riches et précises ainsi que toutes les personnes qui ont posé des questions et qui ont contribué à dynamiser cette 57^e séance.