



HAL
open science

Opposition politique et polarisation de groupe dans la discussion des traitements de la Covid-19 sur Twitter

Liana Ermakova, Diana Nurbakova, Irina Ovchinnikova

► To cite this version:

Liana Ermakova, Diana Nurbakova, Irina Ovchinnikova. Opposition politique et polarisation de groupe dans la discussion des traitements de la Covid-19 sur Twitter. Extraction et Gestion des connaissances EGC 2021, Jan 2021, Montpellier, France. hal-03140862

HAL Id: hal-03140862

<https://hal.archives-ouvertes.fr/hal-03140862>

Submitted on 14 Feb 2021

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.

Opposition politique et polarisation de groupe dans la discussion des traitements de la Covid-19 sur Twitter

Liana Ermakova*, Diana Nurbakova**
Irina Ovchinnikova***

*Université de Bretagne Occidentale, HCTI - EA 4249, Brest, France
liana.ermakova@univ-brest.fr

**LIRIS UMR 5205 CNRS, INSA de Lyon, Villeurbanne, France

***Institute of Linguistics and Intercultural Communication, Sechenov University, Moscow, Russia

Résumé. Durant la pandémie de Covid-19, de nombreux utilisateurs ont considéré Twitter comme une source fiable d'information médicale. Toutefois, cette information est souvent altérée lorsqu'elle est partagée. A travers l'étude qualitative de tweets portant sur la controverse des traitements de la Covid-19, nous montrons que malgré la qualité de l'information dans le tweet initial, les flux de personnes publiques créent des cascades de désinformation via leurs abonnés. Nous clarifions le mécanisme de distorsion de l'information médicale en cascade, découvrons les principaux acteurs de la discussion sur les traitements controversés, et montrons l'effet de polarisation du groupe pendant la discussion. L'information médicale est souvent altérée accidentellement et inconsciemment en raison d'une mauvaise compréhension, qui provoque une substitution des avis relatifs aux prescriptions controversées par un positionnement politique.

1 Introduction

Les opinions des personnes publiques (PP) sont très demandées dans tous les médias sociaux et affectent l'idéologie des citoyens, leurs préférences et attitudes sur des sujets différents. Les citoyens adoptent les regards des PP surtout en période d'instabilité sociale comme la pandémie actuelle. La motivation pour partager l'information dépend des caractéristiques démographiques et sociales des utilisateurs (Majmundar et al., 2018). Le retweet est donc un acte de validation et d'engagement avec les autres (Boyd et al., 2010). Le nombre de retweets dépend de la personnalité de l'auteur du tweet initial, sa longueur et sa difficulté linguistique (Kupavskii et al., 2012). Pendant la pandémie de la Covid-19, de nombreux utilisateurs considèrent Twitter comme une source fiable d'information médicale (Bavel, 2020). Ils cherchent des actualités dans les fils des PP et partagent des preuves anecdotiques de leurs guérisons. Contrairement aux rumeurs (Zubiaga et al., 2016), la désinformation médicale est souvent propagée inconsciemment comme une infection lors d'une épidémie. La désinformation est souvent disséminée en *cascades d'information* (Jasser, 2019) quand les utilisateurs partagent un tweet en le citant ou en le retweetant successivement (Galuba et al., 2010). La désinformation apparaît dans les paraphrases ou les commentaires du tweet d'origine quand l'information est partagée (Ribeiro et al., 2019).

Opposition politique et polarisation de groupe dans la discussion de la Covid-19 sur Twitter

Les cascades se distinguent selon leur structure (Leskovec et al., 2007) : les cascades larges se produisent quand le tweet d'origine est retweeté en même temps par plusieurs utilisateurs, les cascades profondes apparaissent via les citations et les commentaires successifs des messages (Zang et al., 2017). *Les cascades superficielles ne modifient pas l'information d'origine (e.g. retweets, reposts) mais le terme n'est appliqué qu'à la propagation de l'information, tandis que les cascades profondes amènent de nouvelles informations éventuellement altérées (citations, replies)*. La structure dépend des auteurs, du sujet et du contenu ; les politiciens et les sujets politiques provoquent des cascades profondes et virales tandis que les célébrités sont à l'origine des cascades larges et superficielles (Zang et al., 2017). Donc, dans les cascades provoquées par les politiciens, il y a un risque que l'information médicale soit altérée dans les citations et les commentaires (Ribeiro et al., 2019). Le raccourcissement ou la paraphrase du tweet d'origine dans les commentaires, la substitution de la terminologie médicale par les non-médecins diminuent la fiabilité de l'information médicale sensible à toutes sortes de modifications de la terminologie et de la structure du texte faites par des non-spécialistes (Nye et al., 2018; Wilce, 2009). Même la communauté scientifique produit des opinions controversées sur le traitement de la Covid-19 à cause du manque de données vérifiées. Actuellement, il n'existe pas de système automatique de détection de l'information médicale altérée, car le domaine est très sensible à la substitution de la terminologie et aux paraphrases, et la compréhension de l'information médicale demande une acculturation scientifique très solide. Les systèmes de ce type peuvent être conçus à la base des ontologies ou des ressources de données ; mais les outils comme InfoPath (Gomez Rodriguez et al., 2013) ne sont pas en mesure de reconnaître cette altération.

Souvent les citoyens regardent les postes des PP qu'ils suivent et partagent tous les messages liés à la pandémie sans hésiter. L'information vérifiée publiée par un PP peut être altérée en cascade par ses followers. Les données et les faits d'origine sont parfois évalués à travers le prisme de l'attitude de l'utilisateur envers la source de l'information au détriment de la compréhension de la science ou de la logique. Vu que les PP attirent beaucoup de followers, leur influence directe ou indirecte sur la désinformation est un facteur majeur à étudier. Ainsi les objectifs de notre recherche sont : (1) **identifier les personnes publiques** les plus impliquées dans la discussion de l'information médicale sur Twitter, (2) **examiner** les transformations du contenu de leurs tweets d'origine qui amènent à **l'altération de l'information**, et (3) **tracer la dynamique des attitudes** en cascades d'information. Les recherches sont conduites dans le domaine de la diffusion d'information et de la prédiction de la véracité du contenu (cascades superficielles) sans altération (e.g. reposts, retweets) (Kong et al., 2020; Li et al., 2017; Lamprier, 2019; Kupavskii et al., 2012). Cependant elles ne considèrent pas comment la désinformation apparaît en cascade malgré la qualité du tweet d'origine. Le travail le plus proche est celui de Huyghues-Despointes et al. (2019) où la similarité cosinus TF-IDF entre les éléments d'une chaîne de documents est calculée en se basant sur deux corpus (1) une collection des résumés de papiers scientifiques issue de Citation Network Dataset V1 AMINER5 et (2) un ensemble d'articles de Huffington Post. Contrairement aux travaux précédents sur la désinformation pendant les sinistres et ceux de la pandémie en cours (Pennycook et al., 2020; Pennycook et Rand, 2018), nous explorons les mécanismes de l'altération de l'information médicale en lien avec des caractéristiques socio-démographiques de l'initiateur de la cascade. Ceci est fait grâce à une profonde analyse linguistique. Une attention particulière est portée sur les politiciens et les soignants. Derrière la discussion des traitements de la Covid-19, nous avons découvert une opposition politique des utilisateurs et leur confiance dans les décisions gouvernementales. À

notre connaissance, c'est la première analyse des types d'altération de l'information médicale sur les traitements de la Covid-19 qui montre le changement d'attitude en cascades, même quand le tweet d'origine est neutre.

2 Méthodologie

Jeu de données : Nous avons utilisé l'API de Twitter pour collecter 10M de tweets en anglais sur les traitements controversés de la Covid-19 publiés entre 30/03/20 et 13/07/20. La requête était construite par les termes ["chloroquine", "hydroxychloroquine", "Raoult", "remdesivir", "tocilizumab", "favipiravir", "Avigan", "azithromycin", "HCQ", "Axemal", "Dolquine", "Quensyl", "Hydroxychloroquinum", "Montagnier", "Hydroquin", "Quinoric"] parmi d'autres. La liste a été modifiée selon les actualités pendant la pandémie¹ (Noel et al., 2020). La plupart des tweets collectés sont des retweets (cf. Tab. 1). À noter qu'un retweet ne modifie pas le contenu du tweet d'origine. Ainsi, nous n'avons sélectionné que des tweets qui incluent le contenu original généré par les utilisateurs (141,866 tweets). Pour rechercher les tweets qui forment des cascades, d'abord, nous avons choisi 1,000 tweets les plus retweetés et 1,000 tweets les plus cités, en total 1,356 IDs de tweets avec des textes différents. Ces tweets représentent des *tweets initiaux (ou tweets d'origine) des cascades*. Après, nous avons recherché des *paliers des cascades correspondantes* comme suit. Nous avons sélectionné tous les tweets si un des champs `retweeted_status.id`, `quoted_status.id` ou `in_reply_to_status_id` était présent dans l'ensemble des IDs des tweets initiaux. Ces tweets ont été ajoutés aux tweets initiaux. Alors, nous avons obtenu des cascades de profondeur maximale 9. Pour l'analyse, nous avons utilisé les champs `text` ou `extended_tweet.full_text` s'ils étaient présents dans les tweets des cascades.

TAB. 1: Exemples de tweets de PPs les plus partagés (6 avril et 19 mai 2020)

	Utilisateur	Texte	#RTs	#quotes	#replies
6 avril 2020	HillaryClinton	Please do not take medical advice from a man who looked directly at a solar eclipse.	422,447	57,326	78,621
	realDonaldTrump	Be prepared, there is a small chance that our horrendous leadership could unknowingly lead us into World War III.	219,116	87,035	16,789
	realDonaldTrump	HYDROXYCHLOROQUINE & AZITHROMYCIN, taken together, have a real chance to be one of the biggest game changers in the history of medicine. The FDA has moved mountains - Thank You! Hopefully they will BOTH (H works better with A, International Journal of Antimicrobial Agents).....	103,609	30,058	70,802
19 mai 2020	realDonaldTrump	REOPEN OUR COUNTRY!	100,162	23,078	58,082
	RossFairchild	A NY Doctor shared with Hannity his Hydroxy Chloroquine/Azithromycin results. 200mg 2x daily Hydroxy Chloroquine 500mg 1x daily Azithromycin 220mg 1x daily Zinc sulfate 350 patients *Breathing restored 3-4 hours *Zero deaths *Zero hospitalizations *Zero intubations	40,552	4,074	1,879

Tout d'abord, nous avons identifié et catégorisé les auteurs des tweets qui ont initié les cascades, s'agissant souvent de PP. Ensuite, nous avons fait l'analyse sémantique des textes dans les cascades. Nous avons examiné le changement d'attitude des participants des cascades.

1. Nous ne réclamons pas l'exhaustivité de cette liste, mais ces termes correspondent aux études lancées par l'OMS fin mars 2020 (cf. <https://bit.ly/3jouj01>, <https://bit.ly/37KS2pd>) (remdesivir, HCQ, lopinavir/ritonavir et interferon). Le 17 mars, le microbiologiste français D. Raoult a annoncé que selon l'étude sur 24 patients HCQ a montré son efficacité dans le traitement de la Covid-19 (cf. <https://bit.ly/3ji3h9z>). Le 4 juillet, l'OMS a décidé d'arrêter les études de HCQ et lopinavir/ritonavir car ils n'ont montré aucun avantage clinique dans le traitement de la Covid-19 (cf. <https://bit.ly/3kq6iam>). Notre analyse s'arrête 10 jours plus tard.

Identification des personnes publiques les plus impliquées : Nous avons identifié les PP qui ont beaucoup contribué à la discussion de l’information médicale et qui ont initié les cascades. Nous avons généré une liste de fréquence des mots des tweets les plus retweetés et cités (2, 5M mots). Nous avons analysé les mots avec une fréquence supérieure à 100 pour un million de mots et nous avons vérifié la présence de nos mots-clés et des noms propres (`user_screen_name`, `user_name`). Nous avons également considéré la fréquence des mentions des PP dans les hashtags (e.g. #Trump, #donalddtrump, #drfauci, #billgates) comme un indicateur de l’importance de ces personnes dans la discussion de la Covid-19 sur Twitter. La liste de PP utilisée dans l’analyse des cascades contenait les noms les plus fréquents dans le jeu de données. Notez que les PP ont beaucoup de followers, leurs messages sont partagés plus souvent que les messages des citoyens. Mais ce sont des citoyens qui, en citant les tweets des PP, provoquent la désinformation dans les commentaires du message d’origine. Ainsi, nous analysons l’altération de l’information médicale générée par les citoyens car notre étude est focalisée sur la circulation de l’information altérée involontairement en cascades évoquées par les tweets initiaux des PP. Parmi les personnes qui ont plus de 150 mentions, nous avons identifié 36 PP (cf. Tab. 2). La majorité d’entre elles sont issues du domaine ‘Politique’. Les politiciens sont chargés d’informer les gens, d’organiser la prévention et de contrôler la propagation de la Covid-19, en fournissant des possibilités de traitement et en régulant la vie quotidienne pendant la pandémie en cours.

TAB. 2: Les PPs les plus mentionnées dans les tweets cités

Catégorie	#membres	Effectif	La PP la plus mentionnée dans la catégorie	#entries	#hashtags
Politique	17	41,974	Donald Trump, president of the U.S.A.	33,821	1,218
Santé	7	12,327	Dr. A. Fauci, US National Institute of Allergy and Infectious Diseases chief	7,196	484
Journalisme	9	2,565	Nail Cavuto, business journalist for Fox News	1,051	26
Business	3	2,201	Bill Gates, business magnate	1,896	100

Analyse d’altération de l’information : Notre expert (PR en linguistique à l’Université de médecine Setchenov de Moscou) a conduit une analyse sémantique du contenu de tweets en cascades pour identifier les types d’altération de l’information. L’analyse se basait sur la distribution des termes clés en cascades, l’étude de leurs contextes dans les tweets et la vérification de la relation logique entre les termes médicaux. L’analyse distributive permet de reconnaître les substitutions des termes et des mots du tweet d’origine. L’analyse distributive permet de détecter l’information altérée associée avec les tweets d’origine des PP.

Changement d’attitude : Nous avons conduit l’analyse de sentiment sur tous les tweets en cascades pour déterminer les émotions et les attitudes de leurs auteurs. Ainsi, nous avons attribué une polarité à chaque tweet dans notre jeu de données en utilisant la bibliothèque `TextBlob`. Cette polarité des tweets de la cascade a été utilisée pour définir l’état du palier i comme $state_i = sgn(polarity_i)$. Nous définissons la changement d’attitude qui se produit dans chaque palier de la cascade, dénoté Δ , comme $\Delta_i = state_i - state_{i-1}$.

3 Résultats

Nous discutons l’altération de l’information médicale observée en cascades initiées par les tweets de PP et rapportons les résultats de l’analyse du changement de l’attitude.

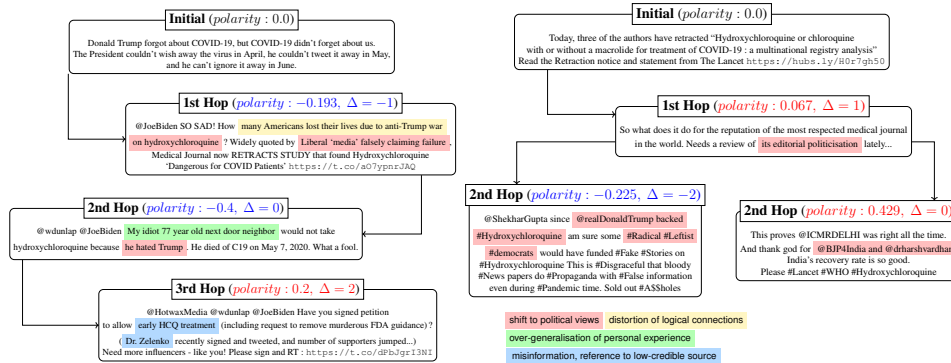


FIG. 1: Exemples de cascades et altération de l'information dans celles-ci

Altération de l'information dans les cascades évoquées par les PP : La liste des mots les plus fréquents dans les tweets cités commence par *Hydroxychloroquine* (HCQ) (62,521 occurrences), tandis que le hashtag le plus fréquent est #Hydroxychloroquine (7,144 occurrences). Donc, l'effet de HCQ pour le traitement de la Covid-19 attire l'attention des utilisateurs de Twitter et il est devenu un sujet important en cascades d'information générées par des commentaires sur les tweets des PP.

Étude de cas de deux cascades profondes liées du point de vue de changement de sujet et altération de l'information : Un tweet de Donald Trump en Avril, 2020 sur l'effet de HCQ a provoqué 103,637 retweets et a généré une cascade profonde (4 paliers). Cette dernière contenait des messages de soutien ainsi que de partage d'expérience de prévention et traitement de la Covid-19 à HCQ et des clarifications des professionnels de santé sur les effets du médicament. Les professionnels de santé essayaient de réparer les dégâts causés par le tweet initial et les supporters du président qui exagéraient l'efficacité de HCQ. Cependant, de nombreux politiciens et personnes des médias critiquaient l'efficacité de HCQ et ridiculisaient le conseil de D.Trump (2M retweets en avril-mai). Le tweet le plus partagé appartient à la membre du Congrès, Hillary Clinton. Il a provoqué une autre cascade, changeant le sujet des débats du traitement à la politique. Ainsi, les discussions contenaient l'information altérée et révélaient l'opposition Républicains vs. Démocrates. Le sujet sur l'effet de HCQ était évoqué dans presque toutes les cascades quand un politicien postait un tweet sur la gestion de la pandémie par D.Trump. Les cascades provoquées par les politiciens contenaient l'information médicale altérée indépendamment du sujet de la discussion; leurs participants utilisaient l'effet de HCQ (positif ou négatif) comme argument dans *les luttes politiques entre libéraux et conservateurs*.

Mécanismes d'altération de l'information médicale dans les cascades provoquées par des tweets des politiciens : L'information médicale était altérée dans les cascades provoquées par les tweets des politiciens par le biais de la vision simplifiée des procédures de traitement de la maladie. Ainsi, la prescription de HCQ était considérée comme une étape importante du processus de guérison. Son retard lié aux procédures bureaucratiques, approvisionnement insuffisant ou délai de livraison était donc présenté comme une violation des droits de l'Homme et/ou une manifestation de la théorie du complot de Big Pharma. L'exagération de l'impact du médicament sur le traitement de la Covid-19 mène à l'altération de l'information, car les corrélations compliquées entre les médicaments et les résultats du traitement de cas différents

sont souvent omises. Des exemples de cascades, de types de l'altération de l'information, des polarités et des changements d'attitudes à travers les paliers de cascades sont présentés à la figure 1. Les mécanismes d'altération de l'information médicale dans les cascades provoquées par des tweets des politiciens comprennent *la simplification, l'omission, l'addition, l'insertion de fausses bases et pré-requis et la surgénéralisation de l'expérience personnelle*. Cette altération est souvent liée à la confrontation des Républicains et Démocrates.

Mécanismes d'altération de l'information médicale dans les réponses et commentaires sur les tweets des professionnels de santé : Contrairement aux tweets des politiciens, ceux *des professionnels de santé* étaient rarement à l'origine des cascades informationnelles. En revanche, les spécialistes étaient plutôt mentionnés dans les cascades existantes provoquées par d'autres, y compris des utilisateurs ordinaires et autres médecins. Souvent, leurs réponses provoquent une mauvaise interprétation, et les présentations des résultats de leurs études sont erronées. Notons que les tweets des professionnels de santé ont réussi à désamorcer une cascade informationnelle. Les présentations erronées apparaissaient principalement dans les commentaires sur les tweets quand les utilisateurs *comprenaient mal la terminologie médicale et remplaçaient les concepts médicaux*, e.g. : le fait que la plupart des personnes qui contractent la Covid-19 en guérissent s'est transformé à *'not such a deadly virus'*, la Chloroquine était souvent remplacée par sa dérivée HCQ. En outre, l'altération de l'information médicale apparaît via le changement de sujet de discussions des résultats des études à *l'encouragement des expériences non-vérifiées de certains praticiens* (e.g. Dr. Vladimir Zelenko) par des utilisateurs ordinaires.

Changement d'attitude : L'analyse de sentiment a montré que les tweets d'origine sont neutres en moyenne, cependant le changement d'attitude apparaît entre les paliers des cascades (voir Fig. 2). Le premier palier est souvent peu différent du tweet d'origine tandis qu'à partir du deuxième palier nous remarquons la polarisation d'attitudes (alternance de colonnes rouges et bleues). Ce phénomène peut être expliqué par la polarisation de groupe, i.e. l'attitude du groupe vis-à-vis de la situation change en intensifiant les attitudes initiales des individus après la discussion, ce qui est cohérent avec la confrontation politique.

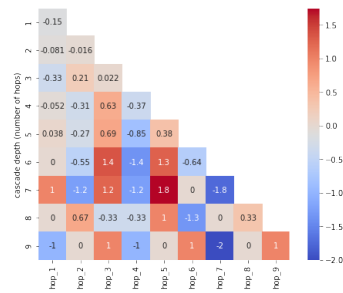


FIG. 2: Changement moyen d'attitude dans les cascades

4 Conclusion

Nous avons présenté une étude qualitative de la circulation de la désinformation médicale sur les médias sociaux pendant la pandémie de la Covid-19. Contrairement aux travaux précédents sur la diffusion de l'information et la prédiction de la véracité du contenu (Kong et al., 2020; Gomez Rodriguez et al., 2013; Li et al., 2017; Lamprier, 2019; Kupavskii et al., 2012), nous avons analysé les cascades profondes où l'information peut être altérée par rapport au message d'origine. Les recherches précédentes sont plutôt focalisées sur la propagation de fake news initiales (Ahmed et al., 2020), tandis que nous avons clarifié le mécanisme d'altération de l'information médicale en cascades, évoqué les acteurs principaux de la discussion des traitements controversés et découvert la substitution de la discussion du traitement de la Covid-19 par l'opposition politique. À notre connaissance, nous sommes les premiers à montrer la polarisation de

groupe sur Twitter pendant la pandémie en conduisant une analyse automatique du changement d'attitude en cascades, même si le tweet d'origine est neutre. L'impact des tweets de PP sur la circulation de la désinformation est révélé dans l'altération du contenu de leurs tweets d'origine en cascades. Les followers des politiciens se comportent comme des partisans en soutenant les différents médicaments contre la Covid-19. La discussion de la pandémie correspond à la confrontation Républicains-Démocrates. L'altération de l'information médicale apparaît à cause de la spécificité du domaine qui exige une analyse des actualités basée sur les connaissances profondes dans les domaines de la santé et la science. Les citoyens simplifient l'information sur la Covid-19, omettent des détails essentiels et des connexions logiques en commentant ou citant les tweets des PP. L'analyse future des cascades d'information peut aider à développer des méthodes pour la classification automatique des altérations d'information.

Références

- Ahmed, W., J. Vidal-Alaball, J. Downing, et F. Segui (2020). COVID-19 and the 5G Conspiracy Theory : Social Network Analysis of Twitter Data. *Journ. of Medical Internet Research* 22(5).
- Bavel, J. V. e. a. (2020). Using social and behavioural science to support COVID-19 pandemic response. *Nature Human Behaviour* 4(5), 460–471.
- Boyd, D., S. Golder, et G. Lotan (2010). Tweet, Tweet, Retweet : Conversational Aspects of Retweeting on Twitter. In *2010 43rd Hawaii International Conference on System Sciences*, pp. 1–10. ISSN : 1530-1605.
- Galuba, W., K. Aberer, D. Chakraborty, Z. Despotovic, et W. Kellerer (2010). Outtweeting the twitterers - predicting information cascades in microblogs. In *Proceedings of the 3rd Wconference on Online social networks, WOSN'10, USA*, pp. 3. USENIX Association.
- Gomez Rodriguez, M., J. Leskovec, et B. Schölkopf (2013). Structure and dynamics of information pathways in online media. In *Proc. of the sixth ACM international conference on Web search and data mining, WSDM '13, New York, NY, USA*, pp. 23–32. ACM.
- Huyghues-Despointes, C., L. Khouas, J. Velcin, et S. Loudcher (2019). Weaving information propagation : Modeling the way information spreads in document collections. In *Proc. of 32nd Canadian Conference on AI, 2019, Volume 11489 of LNCS*, pp. 394–399. Springer.
- Jasser, J. (2019). Dynamics of misinformation cascades. In *Companion Proc. of The 2019 World Wide Web Conference, WWW'19, New York, NY, USA*, pp. 33–36. ACM.
- Kong, Q., M.-A. Rizoïu, et L. Xie (2020). Modeling Information Cascades with Self-exciting Processes via Generalized Epidemic Models. In *Proc. of the 13th International Conference on Web Search and Data Mining, WSDM '20, New York, NY, USA*, pp. 286–294. ACM.
- Kupavskii, A., L. Ostroumova, A. Umnov, S. Usachev, P. Serdyukov, G. Gusev, et A. Kustarev (2012). Prediction of retweet cascade size over time. In *21st ACM International Conference on Information and Knowledge Management, CIKM'12, Maui, HI, USA*, pp. 2335–2338.
- Lamprier, S. (2019). A Recurrent Neural Cascade-based Model for Continuous-Time Diffusion. In *International Conference on Machine Learning*, pp. 3632–3641. PMLR.
- Leskovec, J., M. McGlohon, C. Faloutsos, N. Glance, et M. Hurst (2007). Patterns of Cascading Behavior in Large Blog Graphs. In *Proc. of the 2007 SIAM International Conference on*

Opposition politique et polarisation de groupe dans la discussion de la Covid-19 sur Twitter

- Data Mining*, Proceedings, pp. 551–556. Society for Industrial and Applied Mathematics.
- Li, C., J. Ma, X. Guo, et Q. Mei (2017). DeepCas : An End-to-end Predictor of Information Cascades. In *Proc. of the 26th International Conference on World Wide Web, WWW '17*, Republic and Canton of Geneva, CHE, pp. 577–586.
- Majmundar, A., J.-P. Allem, T. Boley Cruz, et J. B. Unger (2018). The why we retweet scale. *PLOS ONE* 13(10), 1–12.
- Noel, M., L. Ermakova, P. Rammaciotti, A. Perrier, et B. Benbouzid (2020). Controverse scientifique.
- Nye, B., J. Li, R. Patel, Y. Yang, I. Marshall, A. Nenkova, et B. Wallace (2018). A Corpus with Multi-Level Annotations of Patients, Interventions and Outcomes to Support Language Processing for Medical Literature. In *Proc. of the 56th Annual Meeting of the Association for Computational Linguistics (Vol. 1 : Long Papers)*, Melbourne, Australia, pp. 197–207. ACL.
- Pennycook, G., J. McPhetres, Y. Zhang, J. Lu, et D. Rand (2020). Fighting COVID-19 Misinformation on Social Media : Experimental Evidence for a Scalable Accuracy-Nudge Intervention. *Psychological Science* 31(7), 770–780. Publisher : SAGE Publications Inc.
- Pennycook, G. et D. Rand (2018). Lazy, Not Biased : Susceptibility to Partisan Fake News Is Better Explained by Lack of Reasoning Than by Motivated Reasoning.
- Ribeiro, M., K. Gligorić, et R. West (2019). Message distortion in information cascades. *Proc. of the 2019 World Wide Web Conference (WWW '19)*. ACM, 12 pages.
- Wilce, J. (2009). Medical Discourse. *Annual Review of Anthropology* 38(1), 199–215. :eprint : <https://doi.org/10.1146/annurev-anthro-091908-164450>.
- Zang, C., P. Cui, C. Song, C. Faloutsos, et W. Zhu (2017). Quantifying Structural Patterns of Information Cascades. In *Proc. of the 26th International Conference on World Wide Web Companion - WWW '17 Companion*, Perth, Australia, pp. 867–868. ACM Press.
- Zubiaga, A., M. Liakata, R. Procter, G. Wong Sak Hoi, et P. Tolmie (2016). Analysing how people orient to and spread rumours in social media by looking at conversational threads. *PLOS ONE* 11(3), 1–29.

Summary

During the Covid-19 pandemic, many users consider Twitter as a reliable source of medical information, which is often distorted unintentionally while shared. Through the qualitative study of 10^6 tweets about controversial treatments, we show that despite the information quality in the initial tweet, the public figures' feeds evoke cascades contributing to misinformation circulation. We clarified the mechanism of the medical information distortion within cascades, discovered the main actors in the discussion on the controversial medical treatment, and showed group polarisation effect. Oversimplification of the medical issue allows lay users to conform controversial and unclear prescriptions of different doctors to the very clear political opposition.