**LT-03122018-La lutte contre les «fake news» passera par le «deep learning»**

Des projets utilisant le «machine learning» s’attachent non au contenu des fausses nouvelles mais à leurs structures de diffusion sur le Net pour les combattre

«Il faudrait pouvoir étiqueter les nouvelles comme on étiquette les aliments. Pouvoir dire «Attention, risque de fausse information, traces de manipulation» comme on dit «Attention, risque d’allergie, traces de cacahuètes». Sinan Aral était à Genève ces derniers jours, orateur invité à la conférence [TEDxCERN](https://tedxcern.web.cern.ch/%22%20%5Ct%20%22_blank) (organisée en partenariat avec *Le Temps*) pour présenter ses découvertes sur les *fake news* [qui ont fait la une de la revue](http://science.sciencemag.org/content/359/6380/1146%22%20%5Ct%20%22_blank) *[Science](http://science.sciencemag.org/content/359/6380/1146%22%20%5Ct%20%22_blank)* [en mars](http://science.sciencemag.org/content/359/6380/1146%22%20%5Ct%20%22_blank). Avec deux autres chercheurs du Media Lab du Massachusetts Institute of Technology (MIT), il y montrait qu’une fausse nouvelle met six fois moins de temps qu’une vraie pour atteindre 1500 personnes sur Twitter et qu’elle est 70% plus susceptible d’être partagée.

Plus vite, plus loin, touchant plus de monde: on ne pourra jamais empêcher la création de fausses nouvelles ou de rumeurs, un tweet rageur ou une photo retouchée étant si faciles à produire et les intérêts si grands. En revanche, il doit être possible de réduire leur contagion, leur potentiel de nuisance. Le prochain livre de Sinan Aral s’intitulera *[The Hype Machine, The Madness of Crowds in the New Social Age](https://www.curtisbrown.co.uk/client/sinan-aral/work/the-hype-machine%22%20%5Ct%20%22_blank)*(«La machine à pub, la folie des foules dans la nouvelle ère sociale»). Tout un programme...

Rencontré quelques heures avant d’ouvrir la conférence, le chercheur note cinq pistes de lutte contre les *fake news*, «qui ont chacune ses limites». Des étiquettes donc, pour identifier les sites tendancieux ou sincères. «Mais qui décidera des étiquettes?» Facebook a d’ailleurs commencé à qualifier ses sources en se fondant sur Wikipédia (le petit «i» qui apparaît sous une publication), et il n’est que de consulter les étiquettes du *Matin* ou du *Temps* pour comprendre que la partie n’est pas gagnée.

Deuxième piste, réduire l’incitation économique qui a poussé par exemple de jeunes Macédoniens à inonder le web de publications spectaculaires pendant la campagne présidentielle américaine, des posts qui ont attiré beaucoup de publicité et leur ont donc rapporté beaucoup d’argent – mais cela irait contre le modèle d’affaires des réseaux sociaux: pas réaliste.

Haut du formulaire

Bas du formulaire

Troisième piste, la régulation. Aux Etats-Unis, toute publicité politique doit être clairement identifiée, et l’ingérence de pays étrangers est interdite. «Mais davantage de régulation peut devenir dangereux aux mains de pouvoirs autoritaires. En Malaisie la diffusion de *fake news* est désormais punie de 6 ans de prison!»

La quatrième piste est celle de la transparence. «Facebook et Twitter doivent être plus ouverts et montrer leurs algorithmes, leurs données. Mais c’est contradictoire avec notre volonté de protéger nos données, qui doivent être en sécurité, inaccessibles…»

**Geometric deep learning**

Dernière piste, la plus prometteuse: un mélange d’intelligence artificielle et d’intelligence humaine, pour identifier les schémas de diffusion de *fake news*. C’est ce que Sinan Aral a utilisé pour son article dans *Science*, étudiant la structure des cascades de tweets, ce qui peut permettre par la suite de ralentir la viralité d’un contenu problématique en le masquant, en l’isolant sur le web.

C’est aussi ce que développe dans sa start-up [Fabula](https://fabula.ai/%22%20%5Ct%20%22_blank) le professeur à l'Université du Tessin et [à l'Imperial College de Londres](https://www.imperial.ac.uk/people/m.bronstein/page/honours.html%22%20%5Ct%20%22_blank) Michael Bronstein, qui vient d’être [primé à Lausanne par la Fondation Dalle Molle](https://portal.klewel.com/watch/webcast/labels-2018-dalle-molle/%22%20%5Ct%20%22_blank) pour ses travaux sur la détection des *fake news* dans les réseaux sociaux grâce au *geometric deep learning*. Il s’intéresse non au contenu mais à la façon dont il se propage, les *fake news* se diffusant différemment des vraies nouvelles, comme l’a montré l’étude du MIT. Michael Bronstein utilise les modèles de réseau neuronal pour identifier les caractéristiques les plus significatives des cascades de *fake news*. «J’ai été surpris de la polarisation. Fondamentalement, sur Twitter, les gens qui partagent surtout des *fake news* et ceux qui partagent de vraies nouvelles sont deux communautés qui ne se parlent quasiment pas. On savait que Twitter était politiquement polarisé, mais c’est la première fois qu’on le voit dans le contexte des *fake news*.»

La lutte contre les *fake news* est capitale. Le projet de Michael Bronstein a reçu [d’importantes bourses de la Commission européenne](https://twitter.com/ERC_Research/status/999632938936479746%22%20%5Ct%20%22_blank)… et de Google et Facebook.

Le réseau aux plus de 2 milliards d'utilisateurs multiplie de fait les tentatives de mieux lutter contre les fake news, qui mettent en danger son fonctionnement et son modèle d'affaires. Facebook conclut des partenariats par exemple avec le site de fact checking américain [Snopes](https://www.snopes.com/%22%20%5Ct%20%22_blank), qui va labelliser des contenus et aider les ingénieurs de Facebook à entraîner leurs algorithmes. Parodie ou diffamation, le diable est dans les détails, et souvent seul un esprit humain peut aujourd'hui encore faire la différence. «Les fack checkers contribuent à corriger les problèmes, à mieux flaguer les contenus. Mais ils donnent gratuitement leur savoir-faire», regrette le mathématicien spécialiste de la protection des données et basé à Genève Paul-Olivier Dehaye.