



POINT DE VUE

Journalisme automatisé : remettre l'humain au goût du jour

*Au cours des dernières années, la production de textes automatisés a gagné du terrain dans les salles de nouvelles. Si pour le moment son usage reste limité à des sujets simples, les progrès réalisés dans le domaine de l'intelligence artificielle pourraient bien changer la donne. En attendant, journalistes et responsables médiatiques peuvent apprendre à apprivoiser cette technologie en la mettant au service de l'humain. Par **Samuel Danzon-Chambaud***

C'est une idée qui fait frémir bien des journalistes : des systèmes automatisés, capables de digérer puis de retranscrire l'information en une fraction de seconde. Loin d'un scénario de science-fiction, cela est maintenant bien du domaine du possible puisque plusieurs rédactions s'essayent au « journalisme automatisé ». Aussi connu sous le nom de « robot journalisme » ou de « journalisme algorithmique », ce procédé permet de créer des textes de nouvelles sans intervention humaine, mis à part la programmation initiale du script informatique. Devant la multiplication rapide de ses applications et la variété des questions qu'elles soulèvent, un

état des lieux permet de distinguer certains enjeux essentiels pour le journalisme¹.

Le *Los Angeles Times* a ouvert le bal en 2010 avec une nouvelle version de son Homicide Report, un blog ambitionnant de couvrir tous les homicides dans le comté de Los Angeles. Cette nouvelle formule s'appuie en effet sur des articles automatisés pour couvrir tous les homicides de manière systématique. Le

¹ Ce projet a bénéficié d'un financement au titre du programme-cadre pour la recherche et l'innovation « Horizon 2020 » dans le cadre de la convention de subvention Marie Skłodowska-Curie n° 765140.

journal renchérit plus tard avec Quakebot, un programme informatique taillé pour rapporter les alertes de tremblements de terre.

L'agence américaine Associated Press emboîte le pas au *Los Angeles Times* en 2014, en s'associant à la startup Automated Insights pour automatiser une partie de sa rubrique économique. Deux ans plus tard, c'est au tour des résumés sportifs de la Ligue mineure de baseball, et depuis l'an dernier, d'articles de basketball universitaire, d'être automatisés. De son côté, le *Washington Post* a développé son propre logiciel, Heliograf, d'abord pour couvrir les Jeux olympiques d'été de Rio de 2016, puis les résultats de courses électorales américaines. Au Canada, l'agence Canadian Press a déployé son nouveau « bot électoral »

Le journalisme automatisé fonctionne essentiellement en s'appuyant sur une technologie somme toute assez datée, puisqu'elle était déjà évoquée dans les années 1980.

lors de récents scrutins en Ontario, au Québec et en Alberta, tandis qu'en France, c'est le quotidien *Le Monde* qui a collaboré avec l'entreprise Syllabs, lors des élections départementales de 2015.

Une technologie simple, mais limitée

Le journalisme automatisé fonctionne essentiellement en s'appuyant sur un système répondant au nom de NLG, pour *Natural Language Generation* (Génération automatique de textes, en français), une technologie somme toute assez datée, puisqu'elle était déjà évoquée dans les années 1980 et 1990 pour produire des rapports financiers, des bulletins météo ou des résumés sportifs².

² Konstantin Dörr, « Mapping the field of algorithmic journalism », *Digital Journalism*, 4(6), 2016, p. 700-722.

De plus, selon les mots du journaliste-programmeur du *Los Angeles Times* à l'origine de la nouvelle version du Homicide Report, la formule algorithmique permettant au blog de fonctionner est « *tellement simple que cela en devient presque embarrassant* »³.

Pour simplifier, la génération automatique de textes peut effectivement fonctionner à la manière d'un exercice de cahier de vacances : des phrases toutes faites avec des espaces laissés pour vide sont à compléter. Pour résoudre cela, un jeu d'algorithmes va chercher les réponses à même le web, ou directement à partir d'une base de données. Il peut s'agir notamment d'informations disponibles sur les portails de données ouvertes des administrations, des entreprises, des clubs sportifs, etc. Une fois remplis et éventuellement assemblés, ces modèles de texte donnent à leur tour naissance aux articles créés par ordinateur.

L'essor de la génération automatique de textes est ainsi étroitement lié à la quantité de données disponibles. Mais leur qualité rentre aussi en jeu. Si les données s'avèrent défaillantes, le contenu journalistique le sera tout autant : à cause d'erreurs qui se sont glissées dans les données de l'Institut d'études géologiques des États-Unis, Quakebot a signalé des secousses en Californie quand il s'agissait en fait du large du Japon et de l'Alaska ; il a aussi mis en garde la population contre un tremblement de terre survenu... en 1925.

Remplacer les journalistes ?

Pour autant, les journalistes devraient-ils se faire du mauvais sang ? À en croire une expérience menée à l'Université de Karlstad, en Suède, il y a en effet de quoi s'inquiéter. Cette première étude du genre⁴ visait à demander à un groupe de lecteurs (des étudiants dans ce cas-ci) de comparer des articles écrits par des humains à ceux produits par ordinateur. Les résultats ont démontré que, d'une part, les lecteurs avaient du mal à faire la différence entre les deux et, d'autre part, qu'ils accor-

³ Mary Lynn Young et Alfred Hermida, « From Mr. and Mrs. Outlier to central tendencies », *Digital Journalism*, 3(3), 2015, p. 388.

⁴ Christer Clerwall, « Enter the robot journalist », *Journalism Practice*, 8(5), 2014, p. 519-531.

daient davantage de crédibilité aux articles automatisés. Ce dernier aspect a été confirmé par d'autres expériences du même type, la crédibilité des articles automatisés étant tantôt jugée similaire à celle des journalistes, ou bien, à défaut, recevant des avis particulièrement positifs.

Si, pour le moment, les chercheurs reconnaissent que les journalistes ont généralement une plus belle plume, le vent pourrait vite tourner.

Une étape importante a été franchie l'an dernier lors de la controverse ayant entouré un générateur de texte produit par OpenAI, un organisme dédié à ce que l'intelligence artificielle soit bénéfique pour l'humanité. Contrairement à d'autres projets, le code informatique de ce générateur de texte n'a pas été entièrement dévoilé. La raison évoquée ? Son réalisme serait tel qu'il pourrait être utilisé à mauvais escient, comme pour créer des *fake news* par exemple. Cet épisode illustre toutefois les progrès réalisés dans le domaine de l'intelligence artificielle. De là, il se pourrait qu'il n'y ait qu'un pas à franchir pour que le « robot journaliste » devienne aussi talentueux que ses homologues humains.

Il se pourrait qu'il n'y ait qu'un pas à franchir pour que le « robot journaliste » devienne aussi talentueux que ses homologues humains.

En attendant, les salles de nouvelles peuvent apprendre à tirer parti des avantages du journalisme automatisé, en le mettant au service de l'humain. Prenons par exemple le cas d'Associated Press, qui veut l'utiliser pour alléger le fardeau des journalistes. Si l'agence a automatisé une partie de sa rubrique économique et certains résumés sportifs, c'est pour donner plus de temps à ses reporters pour creuser leurs histoires. Des essais sont aussi menés pour retranscrire automatiquement leurs scripts, afin de les rendre compatibles à d'autres formats comme ceux des panneaux numériques ou des assistants intelligents. Là encore, il s'agit de gagner du temps en évitant d'avoir à écrire deux fois la même histoire.

Nouvelles possibilités

D'autres stratégies découlant du journalisme automatisé sont également à l'œuvre, par

exemple la transformation des rôles et des fonctions. Au Royaume-Uni, l'agence RADAR emploie des journalistes pour écrire les modèles de texte servant à la production d'articles automatisés. Cela permet à l'agence de produire, chaque mois, des milliers d'histoires locales qu'elle vend par la suite à ses clients. De la même manière, lors des résultats d'un référendum en Suisse, le groupe Tamedia est parvenu à créer 40 000 articles personnalisés grâce au travail de ses journalistes. Leur rôle était de prédire les scénarios possibles et, là encore, d'écrire les modèles de texte permettant à la technologie de fonctionner. Dans un marché en proie aux contrats

précaires et aux réductions d'effectifs, le journalisme automatisé a donc la capacité d'amener de nouveaux emplois, comme le rôle d'« automation editor » pour Associated Press le démontre.

Enfin, grâce au journalisme automatisé, une nouvelle forme de collaboration entre l'homme et la machine permettrait de mieux équiper les journalistes. Selon le site *Digiday*, le magazine Forbes envisage de fournir à ses rédacteurs une première ébauche d'article automatisée, qu'ils peuvent ensuite retravailler⁵. Dans le même ordre d'idées, un logiciel de l'agence Reuters va même jusqu'à suggérer des phrases toutes faites à ses journalistes, selon ce que rapporte le site *Wired*⁶. Ces façons de travailler pourraient se révéler pratiques lors de délais serrés, lorsqu'un journaliste doit publier un texte le plus rapidement possible, quitte à le réécrire de manière plus créative après.

Le journalisme automatisé bouleverse donc le paysage médiatique, mais brouille aussi les cartes parmi les professionnels de l'industrie.

⁵ Max Willens, « Forbes is building more AI tools for its reporters » [En ligne], *Digiday.com*, 03.01.2019.

⁶ Nicole Kobie, « Reuters is taking a big gamble on AI-supported journalism » [En ligne], *Wired.co.uk*, 10.03.2018.

*Il amène surtout le
journaliste à s'interroger
sur ce qui le démarque
vraiment d'un texte
pondu par l'ordinateur...*

Il amène surtout le journaliste à s'interroger sur ce qui le démarque vraiment d'un texte pondu par l'ordinateur : peut-être avoir l'esprit critique, ou encore être en mesure de donner un sens et du contexte aux histoires... Autant d'ingrédients qui font en sorte qu'en se concentrant sur leur cœur de métier, les journalistes humains resteront bien d'actualité. ■

*Samuel Danzon-Chambaud est doctorant
à la Dublin City University.*