
Intégration des outils liés à l'intelligence artificielle en journalisme : usages et initiatives

Nicolas St-Germain, Université du Québec à Montréal

Patrick White, Université du Québec à Montréal

RÉSUMÉ

Le domaine journalistique évolue au gré des changements technologiques. Les outils liés à l'intelligence artificielle s'inscrivent dans cette tendance pour les salles de rédaction à se réapproprier les innovations. Cet article fait un tour d'horizon des initiatives liées à cette technologie mises en place un peu partout dans le monde.

ABSTRACT

Journalism is evolving with technological changes. The tools related to artificial intelligence are part of these changes that media organizations have to face. This paper presents an overall view of this technology which includes its benefits, its limits and a few initiatives developed by certain newsrooms all around the world.

Les outils liés à l'intelligence artificielle (IA) ne sont pas la solution aux difficultés que connaît le secteur des médias : perte de revenus publicitaires au profit des géants du web ; diminution de la consultation des médias traditionnels à la faveur des médias et plateformes numériques (CEM, 2020 ; Newman *et al.*, 2021). Il ne s'agit que d'une façon pour le journalisme de faire ce qu'il a fait à de nombreuses reprises dans son histoire : se réapproprié l'innovation technique et la technologie afin de l'adapter à ses besoins (Delporte, 2005 ; Pavlik, 2000).

En effet, d'un point de vue historique, le lien entre le développement du journalisme et les nouvelles technologies est fort. Les journalistes se sont régulièrement approprié les innovations pour en faire un usage propre aux visées de leur domaine, les intégrant dans leurs routines (Robinson, 2007, p. 308). Pour Christian Delporte, la nature même du journalisme oblige les organisations médiatiques à « se saisir des derniers progrès dans le domaine de la technologie de la communication » (2005, p. 204).

Il en va de même avec les outils liés à l'IA. Cet article compte donc présenter quelques exemples de l'intégration de ces outils dans diverses salles de rédaction en Europe et en Amérique du Nord dans trois sphères du domaine journalistique, soit (1) dans la collecte et l'analyse de l'information, (2) dans la production de l'information et (3) dans la distribution des nouvelles (3) (Peretti, 2019).

Dans un second temps, les résultats d'une étude publiée et effectuée à l'automne 2020 seront intégrés pour parler de la situation propre au Canada (St-Germain et White, 2021). Cette étude nous permet notamment de constater que les outils liés à l'IA développés en journalisme proviennent principalement de grandes organisations médiatiques causant une inégalité quant à l'accès à ces outils (Beckett, 2021 ; Keefe *et al.*, 2021). Il y a ainsi un besoin de collaboration entre les diverses instances œuvrant dans le secteur journalistique que ce soit entre les médias ou bien avec les universités et les *startups* dans le but de mettre en commun les connaissances et rendre accessible les outils au plus grand nombre d'acteurs possibles.

L'intelligence artificielle comme discipline

L'expression intelligence artificielle englobe un énorme lot de technologies dont l'apprentissage automatique (*machine learning*), l'automatisation et l'analyse de données (Beckett, 2019). À proprement parler, cette discipline du domaine informatique ne compte pas de définition claire. Au contraire, il s'agit d'un « terme un peu fourre-tout pour désigner les nombreuses possibilités offertes par les récents développements technologiques » (Dierickx, 2021b).

Cette absence de définition se répercute dans celles qui sont données par les salles de nouvelles qui se servent de l'IA. En effet, elles divergent d'une salle à l'autre selon l'utilisation qui en est faite, selon le rapport *New Powers, New Responsibilities. A Global Survey of Journalism and Artificial Intelligence* (Beckett, 2019). Toutefois, cela ne signifie pas que la définition que l'on confère à cette branche de l'informatique n'est pas importante. En effet, Peretti estime que le fait de trouver une définition propre à son organisation médiatique permet de mieux établir la stratégie liée à l'usage des outils liés à l'IA (2019, p. 11).

Par conséquent et pour faciliter la compréhension de cet article, lorsqu'il sera question d'« intelligence artificielle », nous ferons référence à la définition donnée par Nicholas Diakopoulos pour qui « l'IA est un programme informatique en mesure d'effectuer des

tâches qui nécessitent habituellement un certain niveau d'intelligence humaine¹» (Kelly, 2020). Dans la même veine, l'expression « *outils liés à l'IA* » fait référence à des programmes informatiques automatisés venant aider le journaliste dans ses tâches au quotidien.

Ces deux définitions viennent illustrer ce qui est attendu de cette technologie, soit une assistance au travail journalistique et non pas un remplacement des effectifs. C'est d'ailleurs le constat des chercheurs Laurence Dierickx et Carl-Gustav Linden, qui précisent que « *la plupart des expériences en matière d'automatisation témoignent non pas d'un remplacement du travail humain, mais bien d'une transformation du travail* » (2021a). L'outil vient donc accomplir un travail qui peut être fait par l'humain, mais qui prendrait beaucoup plus de temps.

Série d'initiatives avec des outils liés à l'IA

Collecte et analyse d'information

La collecte et l'analyse d'information inclut, par exemple, la mise sur pied de veilles automatiques ou d'alertes à partir de signaux faibles sur des sujets dans un champ donné de spécialisation. Les outils s'insérant dans cette catégorie s'appliquent plus souvent qu'autrement à des initiatives permettant d'utiliser l'apprentissage automatique avec de grands ensembles (images, documents, données), pour y détecter des tendances, ainsi que l'automatisation de la récolte de données par l'entremise de programme informatique.

C'est le cas d'un projet qui a été mis sur pied par une agence ukrainienne de journalisme de données du nom de *Texty* (Herasymenko *et al.*, 2018). En somme, le projet cherchait à répertorier tous les endroits au nord-ouest de l'Ukraine où une mine illégale d'ambre avait été exploitée. Présentée sous forme de carte interactive, des images satellites des lieux ont été fournies au système pour qu'il détermine la présence ou non de site d'exploitation minière. Pour lui permettre de prendre les décisions, l'outil a reçu des images étiquetées *néгатif* lorsqu'il n'y avait pas présence de site minier et *positif* lorsque c'était le cas.

Ici, le logiciel ne fait rien que l'humain n'est pas en mesure de faire puisque l'étiquetage préalable des images satellites a été effectué manuellement par les journalistes. La différence, c'est que le système le fait à bien plus grande échelle et en moins de temps : 450 000 images ont été analysées en une centaine d'heures selon la méthodologie du projet (Bondarenko, 2018). Il faut cependant préciser qu'il demeure un risque d'erreur provenant de la qualité des données ayant été fournies au système.

Le deuxième exemple faisant usage de l'apprentissage automatique a été produit par l'*International Consortium of Investigative Journalists* (ICIJ) en partenariat avec l'Université Stanford. Dans ce cas, l'outil utilisant l'apprentissage automatique a été conçu pour déterminer le genre de 340 000 personnes blessées à la suite de bris de dispositifs médicaux aux États-Unis (Guevara, 2019). Au total, une douzaine de personnes a travaillé sur ce projet qui a permis de conclure que 67% des cas, les blessures ont été infligées à une femme.

Le projet s'inscrit dans la lignée des « *Implant Files* », une série d'articles produits par l'ICIJ sur des patients ayant été blessés ou tués à la suite de bris d'implants médicaux. Avant de concevoir l'outil, l'équipe a vérifié dans la base de données « *Manufacturer and User Facility Device Experience* » (MAUDE) de la Food and Drug Administration (FDA), aux

¹ Traduction libre de la citation suivante : « *AI is a computer system that can perform a task that would typically require some level of human intelligence* ».

États-Unis, mais en vain. En fait, la FDA récolte les données sur le sexe des patients, mais ne les affiche pas en ligne « pour des raisons de confidentialité ».

Étant donné la pertinence de cette donnée dans le cadre de cette enquête journalistique (les hommes et les femmes ne répondent pas de la même façon aux équipements médicaux), les experts du consortium et l'Université Stanford ont décidé de concevoir un outil pour déterminer le sexe des patients. Ils ont commencé par l'analyse de rapports produits par les autorités après un accident dans lesquels on retrouvait soit les pronoms « *she* », « *he* », ou des expressions comme « *a male patient* », par exemple.

Avant de laisser le programme informatique travailler seul, un certain travail manuel a dû être accompli par l'équipe de recherche. Trois journalistes ont répertorié 1 000 dossiers disponibles dans la base de données MAUDE et indiqué manuellement le sexe du patient : femme, homme ou inconnu. Cette dernière catégorie permettait d'inclure des cas où les informations produites par le rapport ne suffisaient pas à indiquer le genre du patient (Guevara, 2019).

L'ICIJ a répertorié quatre étapes dans la mise en place d'un outil utilisant l'apprentissage automatique. Premièrement, il faut accomplir un premier tri manuel des données où les journalistes indiquent les paramètres qui seront fournis au logiciel. Deuxièmement, une première récolte automatique est faite par l'outil. Troisièmement, les humains reviennent dans le processus pour évaluer la première récolte automatique et corriger les erreurs de l'outil. Ces trois premières étapes constituent la phase d'« entraînement ». La dernière consiste, pour l'outil, à analyser l'ensemble de la base de données.

Production de nouvelles

La production de nouvelles peut inclure des outils qui permettent l'écriture automatisée de nouvelles routinières dans les sports ou en finance, qui automatisent la traduction ou encore qui facilitent la transcription d'entrevues (Journalism AI et al., 2019). Pour ce qui est de la transcription, toutefois, la majorité de ces outils sont en anglais. L'Agence France-Presse dispose aussi d'un outil baptisé *Transcriber* (AFP, s. d.) qui offre une transcription dans plus de 20 langues. Ce dernier est toutefois réservé aux journalistes de l'agence.

Le premier exemple que nous souhaitons présenter provient de l'agence de presse La Presse Canadienne (PC) (St-Germain et White, 2021). Elles font usage d'un outil nommé *Ultrad* pour traduire automatiquement de l'anglais vers le français les dépêches provenant de la Canadian Press (CP), le pendant anglophone de la PC, ou de l'Associated Press (AP). La traduction est par la suite revue et corrigée par un journaliste publiant le contenu sur le fil de presse du service français de la PC. Cet outil permet de réduire le temps passé sur chacune des dépêches et ainsi augmenter la productivité. Pour fonctionner, l'outil intègre l'apprentissage automatique lui permettant d'apprendre de ses erreurs et de s'améliorer au fur et à mesure de son utilisation.

Le deuxième exemple faisant usage d'une production automatique de contenu provient de la Suède où l'entreprise Öst Media a lancé en 2016 le site internet Klackspark. Le seul et unique objectif de cette plateforme est de couvrir tous les matchs de soccer de la province d'Östergötland, et ce, jusqu'à la division six (Diakopoulos, 2019, p. 96-97 ; Simone, 2020). Il se sert d'une technologie mise au point par l'entreprise suédoise United Robots. Au cours de la saison 2019, le site a publié environ 850 articles par mois, dont 70% étaient automatisés (United Robots 2018). Le média intègre aussi un autre outil appelé « Q&A » qui permet d'ajouter une certaine profondeur au contenu en envoyant automatiquement des questions par message texte aux entraîneurs des équipes et en intégrant leurs citations aux articles (Simone, 2020). Une option d'alerte a aussi été créée permettant aux

14 journalistes attirés au site d'être informés lorsque des événements rares se produisent (une victoire de 11 à 0, par exemple) nécessitant qu'ils recueillent des réactions. Pour « entraîner » l'outil d'automatisation du contenu, l'organisation médiatique a d'abord analysé 1 000 des articles sportifs les plus lus dans la dernière année.

Le dernier outil intègre des techniques informatiques associées aux *deepfakes*, ces vidéos manipulées qu'on appelle aussi « hypertrucages ». Produit par l'agence Reuters en partenariat avec Synthesia, une start-up londonienne, l'outil permet de produire des vidéos montrant un présentateur qui résume des matchs de la Première ligue anglaise (Chandler, 2020). Contrairement aux hypertrucages, cependant, l'outil génère du contenu à partir d'événements qui ont réellement eu lieu.

Pour concevoir le projet, l'agence de presse a d'abord utilisé un algorithme qui produit des résumés textuels des rencontres. Ces résumés sont ensuite intégrés à l'outil mis au point par la firme Synthesia où un présentateur lisant le résumé est simulé. Le présentateur est en fait un humain filmé au préalable en train de prononcer le nom de tous les joueurs et de toutes les équipes ainsi que pratiquement toutes les situations de jeu. Ces enregistrements sont ensuite conservés et intégrés aux vidéos lorsque la situation l'exige. Cette initiative soulève toutefois de nombreuses questions. Qu'arrive-t-il si le présentateur quitte l'agence, ou s'il meurt ? Est-il remplacé ? Est-il rémunéré à chaque utilisation de son image et de sa voix, ou lui a-t-on versé un montant fixe dès le départ ?

Cette sous-section, selon nous, démontre que les outils liés à l'IA ne fonctionnent pas sans intervention humaine, ce qui vient accentuer l'importance d'intégrer ces outils en collaboration avec les journalistes.

Distribution de nouvelles

La sphère de la distribution des nouvelles s'inscrit dans la façon dont les articles sont mis de l'avant par une entreprise de presse, que ce soit sur son site web, sur les réseaux sociaux ou par l'entremise d'une application mobile. Cette sous-section présentera de façon plus ou moins détaillée deux initiatives mises en place par deux médias différents.

Le premier exemple est celui du quotidien canadien anglais *The Globe and Mail*. Depuis deux ans, l'entreprise a remporté de nombreux prix au sujet de son outil *Sophi* qui automatise la mise en valeur des contenus publiés sur son site web (The Globe and Mail Inc., 2021, 2020a, 2020b). Toutes les 10 minutes, l'outil met à jour toutes les pages du site du journal en fonction des probabilités de rentabilité des contenus. Il détermine également si un contenu devrait ou non être disponible derrière le mur payant (*paywall*). Selon le rédacteur en chef du journal, cela permet aux journalistes de ne plus se consacrer à la mise en ligne de leurs articles et de se concentrer sur la production de journalisme de qualité (Turvill, 2021).

Le journal prétend que depuis l'intégration de *Sophi*, le nombre de personnes qui s'abonnaient après avoir visité son site web a augmenté de 51 %. *The Globe and Mail* ajoute que le nombre de ses abonnés numériques avait atteint 170 000 en avril 2021. Ces abonnements représentent désormais 70 % de ses revenus, alors que dans le passé, c'est la publicité qui représentait 70 % de son chiffre d'affaires (*ibid.*).

Notre deuxième exemple est celui du diffuseur public finlandais Yle. Il a conçu un assistant virtuel baptisé Voitto pour distribuer son contenu. L'outil fonctionne par l'entremise de notifications envoyées directement sur l'écran verrouillé du téléphone de l'utilisateur qui peut répondre par la voix (Koponen, 2018). Dans un article décrivant Voitto, le diffuseur finlandais décrit une situation où l'outil propose un contenu audio à un utilisateur qui répond

qu'il préfère l'écouter quand il sortira faire son jogging. Ainsi, lorsqu'il quitte la maison et se met à courir, le document audio se met en marche sans intervention.

Voitto est intégré à un agrégateur des contenus du diffuseur public appelé NewsWatch. Celui-ci est personnalisable en fonction des intérêts de l'utilisateur. L'outil se base donc sur ce profil et tous les autres intérêts de l'utilisateur pour recommander du contenu. Le média finlandais ajoute que l'outil « apprend » même quand l'appli NewsWatch n'est pas active sur l'appareil de l'utilisateur (*ibid.*). L'outil semble fonctionner de la même façon que n'importe quelle notification *push*, mais avec une couche d'apprentissage automatique qui se fait par l'entremise du profil du lecteur et de ses habitudes.

Intégration des outils liés à l'IA au Canada

En réponse à toutes ces initiatives, nous étions curieux de connaître les usages des technologies liées à l'IA au Canada et de mesurer les connaissances des salles de rédaction canadiennes à propos de ces outils. Pour ce faire, cette section présentera de façon sommaire les résultats d'un questionnaire envoyé à l'automne 2020 à 13 grands médias au Canada.

Séparé sous trois constats, ce premier portrait réalisé au Canada note une grande disparité dans les usages, disparité également observée aux États-Unis (Keefe *et al.*, 2021). Cette disparité semble par ailleurs influencée par la situation financière des médias sondés. Au Canada, les entreprises de presse vivent une crise financière sans précédent. En 2019, 58 % des dépenses publicitaires dont dépendent traditionnellement les médias privés étaient effectuées sur le numérique (CEM, 2020), un marché lui-même dominé à 78 % par Facebook et Google (Roy, 2020).

Les réponses que nous avons obtenues suggèrent également que les médias interrogés possèdent une connaissance de base des technologies liées à l'IA sans toutefois posséder l'expertise pour développer des outils en ce sens. De même, certaines craintes face à la technologie sont aussi formulées par nos répondants. Ces deux points seront traités dans le deuxième constat.

Enfin, il est important de préciser que neuf des 13 médias qui ont répondu à notre sondage ont demandé que leurs réponses soient anonymes. Cet aspect de l'anonymat sera discuté dans le cadre de notre troisième constat.

Notre recherche comporte très certainement des limites. Le fait de n'avoir pu compter que sur la participation de 13 médias sur les centaines qui existent au Canada ne permet d'élaborer qu'un portrait assez sommaire des usages faits au sein de ce pays. Cependant, cette recherche n'avait pas pour but de faire le tour de toutes les innovations mises en place, mais plutôt de brosser un premier portrait de la situation auprès de 13 grands médias du pays. Par conséquent, les chiffres présentés ne concernent que les médias répondant, qui font partie des plus grands organes de presse du pays.

Premier constat

La situation financière des entreprises médiatiques semble jouer un rôle dans l'intégration des outils liés à l'IA. En effet, les deux médias en faisant un plus grand usage sont le *Globe and Mail* et La Presse Canadienne/The Canadian Press. Ils font, par ailleurs, partie des répondants à notre questionnaire.

Outre ces deux organisations, sept autres répondants, possédant tous une portée moindre que les deux principaux, font un usage somme tout limité des technologies liées à l'IA. De plus, quatre répondants mentionnent ne pas faire du tout usage de ces outils. Dans la

version de cette recherche mise en ligne sur le serveur de prépublication SSRN (St-Germain et White, 2021), le journal montréalais *Métro*, propriété de Métro Média, avait indiqué ne pas faire usage d'outils liés à l'IA. Toutefois, l'entreprise a précisé par courriel, quelques mois après sa réponse à notre questionnaire, qu'elle comptait quatre usages de cette technologie.

Sur le plan des usages, outre les initiatives abordées précédemment concernant le *Globe and Mail* et la PC/CP, les autres répondants précisent faire usage d'outils liés à l'IA pour concevoir des robots conversationnels tout en les combinant à divers outils comme DataMinr et CrowdTangle qui permettent de surveiller les réseaux sociaux et d'en faire émerger les tendances ou les contenus populaires.

Enfin, plusieurs de nos répondants nous ont dit qu'ils se servaient du profil de leur lectorat pour proposer des articles en fonction de ses intérêts, une pratique qui ressemble, dans une certaine mesure, à l'exemple du diffuseur public finlandais, présenté ci-dessus. Ce profil se fait à partir de données de première main² générées directement par l'utilisateur lorsqu'il navigue sur le site du média, qu'il en utilise les applications mobiles ou qu'il interagit avec du contenu diffusé par le média sur les réseaux sociaux. L'entreprise médiatique peut emmagasiner ces données afin d'offrir une expérience plus personnalisée... ou pour proposer un ciblage publicitaire plus précis à ses annonceurs. Le *Washington Post* a justement mis au point un outil appelé *Zeus* se basant sur ce type de récolte de données dans le but avoué d'offrir une solution plus complète aux publicitaires sur ses sites et ainsi augmenter ses revenus (Davies, 2019).

Deuxième constat

Le deuxième constat que nous pouvons tirer de notre sondage est le suivant : les différentes réponses fournies par nos répondants canadiens nous donnent à penser que s'ils possèdent une connaissance de base des outils liés à l'IA ainsi que de leur champ d'applicabilité, ils ne détiennent pas, toutefois, l'expertise nécessaire à la conception de ces outils. C'est ici que la collaboration entre journalistes, universitaires et experts en informatique pour permettre la mise en place d'initiatives utilisant les technologies liées à l'IA peut s'avérer intéressante.

Un exemple : le programme *Collab JournalismAI* mis sur pied par le *think tank* Polis, de la London School of Economics and Political Science, en collaboration avec la Google News Initiative. Plus de 40 journalistes provenant d'une vingtaine de médias dans le monde se sont réunis lors de la première édition, en 2020. L'année suivante, en 2021, trois différentes *Collabs* ont réuni des universités en journalisme et en ingénierie avec des organisations médiatiques : la *Collab Americas* a compté sept médias provenant de six pays différents qui ont été jumelés avec l'équipe du *Knight Lab* de l'Université Northwestern ; 12 médias en provenance de neuf pays d'Europe, du Moyen-Orient et de l'Afrique ont pour leur part été affiliés aux News Labs de la BBC ; et la *Collab Asia-Pacific* a, de son côté, réuni cinq médias provenant de quatre pays avec la *Times School of Media* et le département de génie informatique de l'Université Bennett (Peretti, 2021 ; Sivadas, 2021).

S'inspirer de ce type d'initiative au Canada permettrait ultimement de pallier l'absence d'expertise que notre sondage a mise en relief. Cela pourrait aussi répondre à différentes

² Traduction française du terme *first-party data* qui se veut une technique de récolte de données où les informations sont prélevées directement de l'utilisateur par le site internet visité. Concrètement, cela peut se produire lorsque l'utilisateur s'authentifie sur le site, répond à des questions sur ses préférences ou encore navigue tout simplement sur les différentes pages du site.

craintes exprimées par nos répondants. Par exemple, certains se sont demandé si des outils qui présentent du contenu sur mesure aux utilisateurs ne risquaient pas d'homogénéiser les contenus en fonction de leurs seuls intérêts, sans proposer d'autres points de vue. D'autres se sont interrogés quant aux risques d'une surabondance d'articles générée par la rédaction automatisée des contenus.

La collaboration est efficace dans la mesure où les outils liés à l'IA sont développés en fonction des recommandations de la salle de nouvelles. Dans le but de mieux comprendre les visées potentielles de l'outil ainsi que les étapes à automatiser, le chercheur Nicholas Diakopoulos recommande le processus de « déconstruction »³ où le but est de déconstruire les différentes étapes qui devront être effectuées par un éventuel outil (Knight Center Courses, 2019a). Il faut donc bien connaître le secteur qui sera automatisé, mais aussi la ligne éditoriale du média ainsi que son public cible pour que l'outil s'inscrive en continuité avec les visées de l'entreprise et non pas en opposition.

Ces craintes sont donc toutes légitimes et pourraient être amoindries par la collaboration entre journalistes et personnes concevant ces outils, mais aussi en abordant cette thématique lors de rencontres entre les médias, les universités et autres acteurs de l'IA.

Troisième constat

Le cas de l'anonymat est intéressant à aborder dans le cadre de cette recherche au Canada. En effet, ce désir chez la majorité de nos répondants de taire leur identité semble s'inscrire en opposition avec ce qui est rapporté dans la littérature scientifique citée jusqu'à maintenant dans cet article et dans laquelle la majorité des initiatives sont partagées afin de permettre un avancement des connaissances. Cette demande de la part de nos répondants semble montrer que la concurrence est toujours rude entre les médias canadiens, surtout avec la baisse des revenus publicitaires au profit des géants du web. Cependant, il faut admettre que nous n'avons pas sondé les médias sur une possible collaboration entre eux pour trouver des idées innovantes.

Conclusion

Somme toute, cet article permet, nous semble-t-il, de mieux comprendre ce qui est attendu et entendu par les outils liés à l'intelligence artificielle ainsi que leurs potentialités dans les salles de rédaction un peu partout dans le monde. En donnant quelques exemples de la technologie, il est plus facile de concevoir des initiatives en ce sens et de développer des outils. Par ailleurs, il est difficile à ce stade de déterminer l'impact qu'aura l'intégration des outils liés à l'IA dans le domaine journalistique. Selon Charlie Beckett, le futur de la technologie est incertain. En revanche, elle possède le potentiel de transformer en profondeur la façon dont le journalisme est produit et consommé (Beckett, 2019). Dans une certaine mesure, les outils liés à l'IA utilisés par les géants du web ont déjà un impact sur le journalisme alors que les médias doivent prendre en compte les divers paramètres pour tenter de rejoindre leur public en adaptant, par exemple, titres, descriptions de photos ou métadonnées de leurs articles pour être bien référencés.

Si l'on revient sur le cas de l'automatisation de certaines pratiques et des craintes de pertes d'emploi qu'elle avait suscitées chez certains de nos répondants, les chiffres sont plutôt rassurants. Effectivement, Frey et Osborne estiment à 8% pour les éditeurs et à 11% pour les reporters et les correspondants le pourcentage de travailleurs à risque de voir leurs postes automatisés (2017). Par ailleurs, l'objectif, lorsqu'un poste est automatisé, est de voir l'employé effectuant préalablement la tâche être déplacé vers ce

³ Traduction libre du terme anglais *decomposition*.

que les gestionnaires de salles de nouvelles appellent un travail « à forte valeur ajoutée » (Linden, 2017, p. 132). On comprend, ici, qu'il s'agit d'enquêtes qui autrement n'auraient jamais vu le jour, de journalisme de solutions, d'infolettres spécialisées ou de longs balados (*podcasts*).

Par ailleurs, comme nous l'avons mentionné en introduction, le développement du journalisme est intrinsèquement lié au développement des nouvelles technologies (Delporte, 2005 ; Pavlik, 2000). De ce fait, « *les journalistes n'ont pas le choix de se mettre les deux mains dans le cambouis des technologies s'ils ne veulent pas devenir les dindons de la farce numérique* » (Roy, 2015, p. 3). L'auteur reformulait à ce moment la citation quelque peu prophétique du journaliste Yves Leclerc qui, dès 1978, avait prédit que le journalisme allait devenir davantage technique au fil des ans et que le « gros bon sens et le flair instinctif ne suffiront plus à [empêcher de] nous faire manipuler par les techniciens de la propagande qui, eux, auront une connaissance de plus en plus profonde des possibilités et des effets des nouvelles techniques » (*ibid.*).

Cette vision est exactement ce que l'on retrouve, en 2021, en cette ère de « post-vérité » et où les fausses nouvelles ainsi que le manque de transparence des algorithmes relativement à la vie privée sont omniprésents (Nocetti, 2017). C'est ainsi que d'investir dans les technologies liées à l'IA permet aussi de développer une compréhension de celles-ci dans le but de mieux couvrir ses faits et gestes. C'est en ce sens que les nouvelles pratiques journalistiques telles que le journalisme algorithmique (Dierickx, 2021a) ou la reddition de compte algorithmique (Diakopoulos, 2014) s'inscrivent. Carl-Gustav Lindén résume bien ce changement dans la pratique journalistique :

Il est important de penser aux systèmes [informatiques] [...] sans nécessairement avoir la prétention de devenir un expert en développement. Vous devez devenir un expert dans la compréhension de ces outils et de cette technologie. Vous devez comprendre leur but, la logique derrière leur fonctionnement. [...] Je crois que c'est quelque chose qui doit être appris aux journalistes et dans les cours de journalisme pour qu'ils comprennent comment fonctionnent les systèmes⁴. (Knight Center Courses, 2019b)

Il serait utopique de penser que l'intégration des outils liés à l'IA se fasse du jour au lendemain et sans heurts. Au contraire, c'est un jeu d'essais et d'erreurs pour lequel il n'existe aucune recette magique. Cependant, notre survol de diverses expériences menées dans le monde et de la littérature, ainsi que notre propre sondage, font ressortir certaines tendances relativement à l'intégration des outils liés à l'IA en journalisme.

Si la collaboration en fait partie, il existe aussi deux autres éléments à traiter. Tout d'abord, il convient d'avoir un objectif clair lorsqu'on souhaite intégrer ces outils. Pour cela, il existe différentes techniques dont le processus de déconstruction développé par Nicholas Diakopoulos. Ensuite, il est recommandé de ne pas voir trop grand dès le départ. Il est préférable de viser plus simple et ensuite d'étendre l'outil à diverses pratiques. De plus, l'embauche d'une équipe multidisciplinaire semble aussi recommandée. Finalement, il ne faut pas négliger la qualité et la quantité de données nécessaires à la conception de ces outils pour s'assurer des performances de la machine développée. Comme le précise l'Association mondiale des journaux dans un rapport publié en 2019 : « *la qualité du*

⁴ Traduction libre de la citation suivante : « *It's important to think about the systems [...] but you might not need to have the ambition [of] becoming an expert on developing this stuff. You need to be an expert in understanding what they do. What they're for. What's the logic behind. [...] I think that [is] something we need to teach journalists and journalism students to understand how these systems work.* »

« système automatisé dépend de la disponibilité et de la qualité des données qui lui sont fournies⁵ » (Linden et al., 2019).

Au fond, on se rend compte que l'intégration des outils liés à l'IA dans les salles de rédaction n'a rien de plus normal pour un domaine qui se réinvente au gré des nouvelles technologies. ■

Nicolas St-Germain est étudiant à la maîtrise en communication à l'Université du Québec à Montréal.

Patrick White est professeur à l'École des médias de l'Université du Québec à Montréal.

Références

- AFP (s. d.). AFP Transcriber. [En ligne] afp.com.
- Beckett, Charlie (2019). New powers, new responsibilities. A global survey of journalism and artificial intelligence. *Polis*. [En ligne] blogs.lse.ac.uk/polis, 18.11.2019.
- Beckett, Charlie (2021). Introducing the AI Academy for Small Newsrooms. *Polis*. [En ligne] blogs.lse.ac.uk/polis, 06.07.2021.
- Bondarenko, Anatoliy (2018). Search for illegal amber mining on satellite images, HOWTO. *Github*. [En ligne] github.com/texty.
- CEM. (2020). Publicité. Dans *Centre d'études sur les médias*. [En ligne] cem.ulaval.ca.
- Chandler, Simon (2020). Reuters Uses AI To Prototype First Ever Automated Video Reports. *Forbes*. [En ligne] forbes.com, 07.02.2019.
- Davies, Jessica (2019). The Washington Post is preparing for post-cookie ad targeting. *Digiday*. [En ligne] digiday.com, 16.07.2019.
- Delporte, Christian (2005). Jules Verne et le journaliste. Imaginer l'information du XX^e siècle. *Le Temps des médias*, n° 4(1), 201-213.
- Diakopoulos, Nicholas (2014). *Algorithmic Accountability Reporting: On the Investigation of Black Boxes*. Tow Center for Digital Journalism, Columbia University. [En ligne] academiccommons.columbia.edu.
- Diakopoulos, Nicholas (2019). *Automating the news: how algorithms are rewriting the media*. Harvard University Press.
- Dierickx, Laurence (2021a). Journalisme algorithmique: un état de l'art de la recherche. *Les carnets du LaPIJ*. [En ligne] lapij.ulb.ac.be, 31.03.2021.
- Dierickx, Laurence (2021b). Intelligence artificielle et journalisme: une course avec les machines. *Equal Times*. [En ligne] equaltimes.org, 06.04.2021.
- Frey, Carl Benedikt et Osborne, Michael A. (2017). The future of employment: How susceptible are jobs to computerisation? *Technological Forecasting and Social Change*, 114, 254-280.

⁵ Traduction libre de la citation suivante : « *The quality of output of an automated journalistic system is highly dependent on the availability and quality of the data that is fed into it* ».

- Guevara, Marina Walker. (2019). We Used AI to Identify the Sex of 340,000 People Harmed by Medical Devices. *ICIJ*. [En ligne] www.icij.org.
- Herasymenko, Vlad, Kelm, Nadia, Bondarenko, Anatoliy, Drozdova, Yevheniia et Tymoshchuk, Yaroslava (2018). Leprosy of the land. *TEKCTH.ORG.UA*. [En ligne] texty.org.ua.
- Journalism AI, VRT NWS, et LSE (2019). Introduction to Machine Learning. Dans *Google News Initiative Training Center*. [En ligne] newsinitiative.withgoogle.com.
- Keefe, John, Zhou, Youyou et Merrill, Jeremy B. (2021). The present and potential of AI in journalism. *Knight Foundation*. [En ligne] knightfoundation.org, 12.05.2021.
- Kelly, Tara (2020). "Episode 10: Conversation With Prof Nicholas Diakopoulos (Northwestern University)". *Data Journalism*. [En ligne] soundcloud.com/datajournalism.
- Knight Center Courses (2019a, 8 janvier). *Computational Thinking*. [Vidéo YouTube]. 08.01.2019.
- Knight Center Courses (2019b, 28 janvier). *News Algorithm guest speaker Carl-Gustav Lindén, Media Journalism Researcher*. [Vidéo YouTube]. 28.01.2019.
- Koponen, Jarno (2018). First in the world: Yle's smart news assistant Voitto ensures that you don't miss the news you want to read. *Yle News Lab – Future of Journalism*. [En ligne] newslab.yle.fi, 10.10.2018.
- Linden, Carl-Gustav (2017). Decades of Automation in the Newsroom. Why are there still so many jobs in journalism? *Digital Journalism*, 5(2), 123-140.
- Linden, Carl-Gustav, Tuulonen, Hanna, Bäck, Asta, Diakopoulos, Nicholas, Granroth-Wilding, Mark, Haapanen, Lauri, ... Toivonen, Hannu (2019). *News Automation The rewards, risks and realities of "machine journalism"*. [En ligne] jyx.jyu.fi.
- Newman, Nic, Fletcher, Richard, Schulz, Anne, Andi, Simge, Robertson, Craig T. et Nielsen, Rasmus Kleis (2021). Reuters Institute Digital News Report.
- Nocetti, Julien (2017). Comment l'information recompose les relations internationales. La faute à Internet ? Dans *La guerre de l'information aura-t-elle lieu ?* 138-143. Institut français des relations internationales.
- Pavlik, John (2000). The Impact of Technology on Journalism. *Journalism Studies*, 1(2), 229-237.
- Peretti, Mattia (2019). 7 things to consider before adopting AI in your news organisation. [En ligne] blogs.lse.ac.uk/polis, 12.12.2019.
- Peretti, Mattia. (2021). Kicking off the 2021 JournalismAI Collab in the Americas and EMEA. *Polis*. [En ligne] blogs.lse.ac.uk/polis, 28.04.2021
- Robinson, Sue (2007). "Someone's Gotta Be in Control Here". *Journalism Practice*, 1(3), 305-321.
- Roy, Jean-Hugues (2015). Le grand dérangement numérique: plaidoyer pour un journalisme « hacker ». Dans Robert Maltais et Pierre Cayouette (dirs), *Les journalistes. Pour la survie du journalisme*, 125-141. Québec-Amérique.
- Roy, Jean-Hugues (2020). Google et Facebook seuls ne sauveront pas les médias de leur marasme financier. *The Conversation*. [En ligne] [Theconversation.com](https://theconversation.com), 06.11.2020.

Simone, Carlo (2020). "Look at the robot as your new colleague": what automation can do for sports reporting. *Journalism.co.uk*, section Media news. [En ligne] journalism.co.uk, 21.07.2021

Sivadas, Lakshmi (2021). Kicking off the JournalismAI Collab Challenge in Asia-Pacific. *Polis*. [En ligne] blogs.lse.ac.uk/polis, 19.05.2021.

St-Germain, Nicolas et White, Patrick (2021). Les impacts de l'intelligence artificielle sur les pratiques journalistiques au Canada. [Pré-publication] *SSRN*.

The Globe and Mail Inc. (2020a). The Globe and Mail's Sophi Wins Best Digital News Start-Up Award. *Cision*. [En ligne] newswire.ca, 20.05.2020.

The Globe and Mail Inc. (2020b). Sophi.io, du *Globe and Mail*, remporte un prestigieux prix de journalisme en ligne. *Cision*. [En ligne] prnewswire.com, 20.10.2020.

The Globe and Mail Inc. (2021). The Globe and Mail's Sophi.io Wins Digiday Media Award. *GlobeNewswire News Room*. [En ligne] globenewswire.com, 28.06.2021.

Turvill, William (2021). Phillip Crawley: How AI helped Globe and Mail reach 170,000 digital subs. [En ligne] pressgazette.co.uk, 29.04.2021.

United Robots. (2018). *Robot texts underpin regional football site*. [En ligne] unitedrobots.ai.