



HAL
open science

Les robots-journalistes entre mythes et réalité : comment l'intelligence artificielle bouleverse les pratiques journalistiques dans les rédactions de presse écrite

Maëlle Lafond

► To cite this version:

Maëlle Lafond. Les robots-journalistes entre mythes et réalité : comment l'intelligence artificielle bouleverse les pratiques journalistiques dans les rédactions de presse écrite. Sciences de l'information et de la communication. 2018. dumas-02934359

HAL Id: dumas-02934359

<https://dumas.ccsd.cnrs.fr/dumas-02934359>

Submitted on 9 Sep 2020

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.

Copyright

Master professionnel

Mention : Information et communication

Spécialité : Journalisme

Option : Journalisme et innovation

Les robots-journalistes entre mythes et réalité

Comment l'intelligence artificielle bouleverse les pratiques journalistiques dans les rédactions de presse écrite

Responsable de la mention information et communication
Professeure Karine Berthelot-Guiet

Tuteur universitaire : Valérie Jeanne-Perrier

Nom, prénom : LAFOND Maëlle

Promotion : 2017-2018

Soutenu le : 23/11/2018

Mention du mémoire : Très bien

REMERCIEMENTS

J'adresse mes remerciements aux personnes qui ont participé à l'élaboration de ce mémoire, et qui m'ont aidé tout au long de ce processus. En premier lieu, je remercie le professeur Valérie Jeanne-Perrier, chercheuse et enseignante au Celsa, et Tristan Mendès-France, intervenant, pour leur accompagnement au cours de ce travail de recherche.

Je tiens également à remercier Damien Desbordes, journaliste et auteur de *Les robots vont-ils remplacer les journalistes ?* d'avoir pris le temps de partager ses recherches et d'en discuter avec moi.

Merci aussi à Robin Grovik, Neil Thurman et Claude de Loupy qui ont accepté de répondre à mes questions et dont les propos ont été la matière première de ce travail de recherche.

SOMMAIRE

Introduction	p.4
I – L'ARRIVÉE DE L'INTELLIGENCE ARTIFICIELLE DANS LES RÉDACTIONS ET SA RÉCEPTION	
<i>A) Panorama des cas d'usage en France et dans le monde</i>	
- De la recherche américaine aux précurseurs du vieux continent	p.7
- Une myriade de dénominations pour une même technologie	p.11
<i>B) Une réception en demi-teinte</i>	
- Des journalistes tentés mais inquiets	p.13
- Des lecteurs peu informés mais critiques	p.20
II – LES ENJEUX DE L'INTELLIGENCE ARTIFICIELLE APPLIQUÉE AU JOURNALISME	
<i>A) Un cadre moral et déontologique</i>	
- Qui parle quand une IA écrit ? Droit d'auteur et responsabilité	p.24
- Du besoin de transparence pour préserver la confiance et l'indépendance_.....	p.26
<i>B) Un coût économique et social mal évalué</i>	
- Le business de la data	p.30
- Des économies pour les médias mais à quel prix ? L'emploi_.....	p.31
Conclusion	p.36
Bibliographie indicative et sources	p.39
Résumé et mots-clés	p.40
Annexes	p.41
- Annexe n°1 : Entretien avec Robin Govik	p.44
- Annexe n°2 : Entretien avec Claude de Loupy	p.49
- Annexe n°3 : Entretien avec Neil Thurman	p.56
- Annexe n°4 : Questionnaire : le robot-journalisme et les lecteurs	p.58

Ce mémoire est né du constat que le secteur des médias, comme presque tous les pans de l'économie, n'échappe pas à la déferlante de l'intelligence artificielle. *Machine learning*, big data, algorithmes... Depuis quelques années, ces termes sont venus s'inviter dans toutes les conférences, dans toutes les rédactions et dans tous les esprits, sans que l'on sache bien ce qu'ils recouvrent. L'intelligence artificielle (IA), qui consiste à mettre en œuvre un certain nombre de techniques visant à permettre aux machines d'imiter une forme d'intelligence réelle, est en effet un vaste domaine qui touche non seulement à l'informatique mais aussi aux mathématiques, à la neuroscience et même à la philosophie. L'IA fascine depuis plus d'un demi-siècle les scientifiques, mais aussi les romanciers et cinéastes. Des cyborgs tueurs de *Terminator* aux androïdes de *Blade Runner* en passant par HAL 9000 de *2001, l'Odyssée de l'espace*, les humains semblent obsédés par la possibilité de répliquer leur comportement et communiquer avec des machines « pensantes ».

Des fantasmes qui n'ont pas épargné le journalisme, à mesure que s'est développée l'intelligence artificielle : entre les chatbots, des agents conversationnels capables de délivrer du contenu personnalisé aux lecteurs via les portails des titres de presse voire directement depuis Facebook, et les assistants vocaux, que les GAFAM (Google, Apple, Facebook, Amazon et Microsoft) annoncent déjà comme de véritables alternatives à la radio ou la télévision, le futur de la presse telle qu'on la connaît aujourd'hui s'annonce plus qu'incertain. D'autant que depuis le début des années 2010, les algorithmes ont aussi fait leur apparition au sein même des rédactions, apportant avec eux la crainte de voir un jour les journalistes disparaître. Au-delà des simples outils qui peuvent aider les correspondants et autres reporters dans leur travail au quotidien, nous avons choisi de nous intéresser plus spécifiquement aux robots-rédacteurs - bien que, comme nous le verrons plus tard, tous les acteurs sont loin d'être d'accord sur le nom qu'il convient de donner à cette branche de l'intelligence artificielle appliquée à la production de contenus journalistiques. Ce sont en effet eux qui cristallisent les appréhensions de tout un secteur aujourd'hui. Pourquoi ? Comment ces algorithmes fonctionnent-ils ? Qu'en pensent les premiers concernés, les journalistes ? Et les lecteurs ? Enfin dans quel cadre déontologique et juridique s'inscrivent ces technologies ?

L'objectif de ce travail de recherche est d'offrir une analyse des mythes qui entourent ces outils, entre technophilie et technophobie, tandis que l'enjeu est d'identifier les pistes qui se dessinent pour l'avenir de ce métier dans un secteur en crise.

Nous pourrions ainsi nous interroger sur la place des robots-journalistes dans les rédactions de presse écrite et sur ses enjeux

Afin d'éclairer la question de recherche ci-dessus, on s'appuiera sur trois hypothèses :

Hypothèse n°1 : les rédactions ont recours aux algorithmes pour des articles factuels, alors que les journalistes continuent à accomplir des tâches à forte valeur ajoutée ;

Hypothèse n°2 : les journalistes et les lecteurs restent partagés quant à l'arrivée de ces technologies dans les rédactions ;

Hypothèse n°3 : il existe encore un certain nombre de limites inhérentes à ces outils qui limitent leur développement.

1.2 Approche méthodologique et corpus

Bien avant de réfléchir à la problématique et au corpus, ce travail de recherche a pris forme dans un cadre assez restreint. Il nous a en effet fallu composer avec les deux autres mémoires ayant été rédigés sur des thématiques voisines (les chatbots en 2017 et les assistants vocaux en 2018) qui auraient toutes les deux eu leur place ici. Loin d'être un obstacle, cela nous a permis de centrer la réflexion sur les robots-rédacteurs, une part moins visible donc moins étudiée de l'intelligence artificielle appliquée aux journalistes. Or si un type de robot doit un jour remplacer les journalistes, ce seront bien ces nouveaux « producteurs » d'informations, là où les assistants vocaux et les chatbots, faisant office de canaux de distribution, mettraient plus en péril les médias eux-mêmes.

Car la question à laquelle nous allons tenter de répondre en filigranes est bien : est-ce que les robots vont remplacer les journalistes ? C'est en effet la question qui agite le champ journalistique et c'est donc celle-ci qui sera le véritable fil rouge du présent

mémoire, permettant aussi de dépasser l'effet catalogue qui pourrait naître si l'on se contentait de dresser un panorama (non exhaustif) des cas existants – bien que celui-ci soit nécessaire et permette de mettre en lumière les avantages que trouvent les rédactions à l'utilisation de tels outils, il est évidemment insuffisant pour nourrir une vraie réflexion sur le sujet. Mais la difficulté tient au fait que la littérature sur le robot-journalisme est relativement faible, à l'exception des nombreux articles de presse sur le sujet. On peut arguer que cela tient à la relative "nouveau" du phénomène, mais aussi comme nous le verrons plus loin, à la position des journalistes, qui sont à la fois sujets et objets.

C'est pourquoi le choix a été fait de ne pas réduire le corpus aux collaborations françaises, mais de prendre aussi en compte les exemples étrangers. Afin de limiter les biais induits par de tels écarts, nous prendrons bien le temps de mettre en avant le contexte culturel des pays étudiés, ainsi que tout autre facteur pouvant avoir un impact sur les résultats. Le corpus sera enfin complété par des entretiens avec des « experts » de ce secteur, chercheurs, professeurs et ingénieurs en IA, qui ont côtoyé au plus près ces outils de production automatisée de contenus. Leurs propos auront été récoltés dans le cadre d'entretiens téléphoniques, traduits si besoin, et retransmis intégralement. Ils seront analysés comme tels, à la lumière de nos recherches annexes. Il s'agira alors de tenter d'extraire de ces sources la matière nécessaire, au cas par cas, matière qui servira ensuite de socle à la réflexion et viendra appuyer notre propos¹.

Les différents cas d'usage de robots-rédacteurs dans le monde et leur réception représenteront donc le premier ensemble d'éléments de cette réflexion sur la place des robots dans les rédactions de presse écrite (I). Nous nous attacherons ensuite à identifier les enjeux soulevés par ces technologies (II).

¹ Cf bibliographie indicative et sources p.39

I – L'ARRIVÉE DE L'INTELLIGENCE ARTIFICIELLE DANS LES RÉDACTIONS ET SA RÉCEPTION

A) Panorama des cas d'usage en France et dans le monde

Alors que les journalistes et les rédactions de la presse écrite se remettent tant bien que mal du tremblement de terre que fut le journalisme web, une autre vague s'approche d'eux, menaçant de réduire à néants tous leurs efforts pour se relever : l'intelligence artificielle. Celle-ci a d'abord fait son apparition sous la forme d'algorithmes (un ensemble de lignes codées, qui sont autant d'ordres pour un ordinateur) dans les agences, agissant en tant qu'assistants, que ce soit pour l'archivage, le fact-checking ou le SEO (*Search Engine Optimization*, référencement naturel en français, soit ensemble de techniques visant à optimiser la visibilité d'une page web dans les résultats de recherche). À l'origine, ces outils sont donc censés faciliter le quotidien des journalistes en les débarrassant des tâches répétitives à faible valeur ajoutée. Mais depuis le début des années 2010, ces algorithmes d'assistance ont « muté », évoluant en une technologie autrement plus efficace : les « robots-rédacteurs ».

De la recherche américaine aux médias précurseurs du vieux continent

Cette année-là, aux Etats-Unis, le programme Stats Monkey, imaginé par Larry Birnbaum et Kris Hammond, deux chercheurs du laboratoire d'information intelligente (Infolab), installé sur le campus de l'université du Northwestern, à Evanston (Illinois) et développé par le journaliste John Templon et l'informaticien Nick Allen, s'illustre en étant l'un des premiers à rédiger des articles factuels sur des événements sportifs. Une première dans le monde de la Génération Automatique de Texte (GAT) appliquée au journalisme, qui ne tardera pas à essaimer dans les rédactions du monde entier. Dès 2012, le journaliste et développeur au *Los Angeles Times*, Ken Schwencke s'attèle à la création d'un programme similaire, qui voit le jour en grandes pompes deux ans plus tard : alors qu'un séisme secoue la Californie le 17 mars 2014, QuakeBot publie son compte-rendu moins de trois minutes après le tremblement de terre. « Le journaliste n'a eu qu'à relire et cliquer sur publier »

raconte Damien Desbordes dans *Les robots vont-ils remplacer les journalistes ?* (Le Point du Jour, 2018). Le fonctionnement du programme est simple : dès qu'une alerte au séisme (dépassant le seuil de magnitude 3 sur l'échelle de Richter) est émise par le bureau géologique des États-Unis (USGS), le robot réagit immédiatement. « QuakeBot est programmé pour extraire les données pertinentes du rapport de l'USGS et les insérer dans un gabarit pré-écrit », précisait *Slate Magazine*² qui a interrogé le journaliste du *LA Times*. Un outil particulièrement efficace dans une région qui connaît une activité sismique régulière, pour un coût limité puisque l'outil a été développé en interne.

Un choix fait deux ans plus tard par le *Washington Post* en développant son outil d'intelligence artificielle maison, The Heliograf. Comme les autres outils de GAT, celle-ci fonctionne grâce à une base de données que lui fournissent les journalistes, dont elle extrait et met en forme les informations utiles pour rédiger des simi-articles en suivant un modèle préalablement défini. D'abord chargé de couvrir les Jeux Olympiques de 2012 à Rio de Janeiro, au Brésil, au cours de laquelle il a rédigé plus de 300 brèves et alertes, l'IA du *Washington Post* s'est ensuite attaqué aux campagnes électorales en vue de l'Election Day (l'élection au suffrage universel pour les charges publiques au niveau national, fédéral et local) et les matchs de football universitaires dans la région. Au cours de sa première année, Heliograf a rédigé près de 850 articles pour le *Washington Post*, dont 500 sur les élections qui ont généré 500.000 cliques d'après le pure-player américain Digiday³. Un chiffre somme toute raisonnable, mais qui est loin d'être anodin si l'on considère que ces articles n'auraient jamais été aussi nombreux si Heliograf n'avait pas existé.

Contrairement au *Los Angeles Times* et au *Washington Post*, Associated Press a choisi de faire appel au leader américain du marché Automated Insight en 2015 pour son logiciel de GAT Wordsmith, afin de couvrir les événements sportifs universitaires (ce qu'elle ne faisait pas avant) et les rapports trimestriels des grands groupes. Avant Wordsmith, les journalistes étaient en effet chargés d'en tirer des milliers de dépêches sur la santé, les tendances et les perspectives des entreprises américaines - une tâche extrêmement chronophage qui « donnait aux journalistes

² « The First News Report on the L.A. Earthquake Was Written by a Robot » Slate, 2014

³ « The Washington Post's robot reporter has published 850 articles in the past year » Digiday, 2017

envie de s'arracher les yeux » rapporte Automated Insign sur son site dans un billet rédigé par ... Wordsmith lui-même.

La vague de la génération automatique de textes a frappé les côtes de l'Europe en 2015, simultanément par la France et la Suède. Au nord, le groupe de presse local MittMedia a, comme le *LA Times* et le *Washington Post*, développé son propre logiciel pour enrichir ses contenus sportifs, économiques et immobiliers. Créé en 2003, MittMedia est le deuxième acteur le plus important du paysage médiatique suédois en termes de revenus, avec 22 titres de presse local. Alors que le groupe entame un virage et prend conscience de la nécessité de développer une vraie stratégie digitale, il décide d'augmenter sa production d'articles à succès (immobilier, fait divers, entreprises locales) comme l'explique Robin Govik, responsable des contenus digitaux⁴ :

« Pour produire plus d'articles sur ces thématiques, nous pouvions évidemment demander aux journalistes, mais c'était plus logique de se lancer dans la production automatisée. Nous avons donc commencé à faire rédiger des articles pour toutes les transactions immobilières du pays, et le robot est vite devenu notre « employé » le plus productif, si l'on peut parler en ces termes. Il totalisait le plus grand nombre de pages vues derrière un paywall, et avait le taux de conversion lecteur-abonné le plus élevé. »

Comme en France, la presse écrite suédoise – qui a longtemps dominé le paysage médiatique du pays - a pâti de l'arrivée de ses concurrents digitaux, encaissant une baisse globale de sa diffusion. Mais la situation est loin d'être critique (25% de la population déclare lire un titre de presse papier généraliste quotidiennement et 50% de manière hebdomadaire⁵) et les médias locaux restent mieux lotis que leurs homologues hexagonaux qui ont subi d'innombrables coupes budgétaires ces dernières années. Ainsi MittMedia, pour le développement de son logiciel de GAT, a pu s'appuyer sur l'expertise de son département en recherche et développement

⁴ Cf annexe n°1 p.44

⁵ « Media Landscape: Sweden » European Journalism Centre

(R&D) composé d'une cinquantaine de personnes, dont 30 développeurs web de toutes sortes.

La même année, en France, c'est Luc Bronner (qui n'a pas donné suite à nos multiples demandes d'entretien), aujourd'hui directeur des rédactions au *Monde*, qui prend la décision de recourir pour la première fois à ce genre d'outils, pour couvrir les élections départementales. La rédaction fait donc appel à la société Syllabs, spécialisée dans la production de contenus automatisés, fondée par Helena Blancafort et Claude de Loupy⁶, chercheurs en informatique et sémantique :

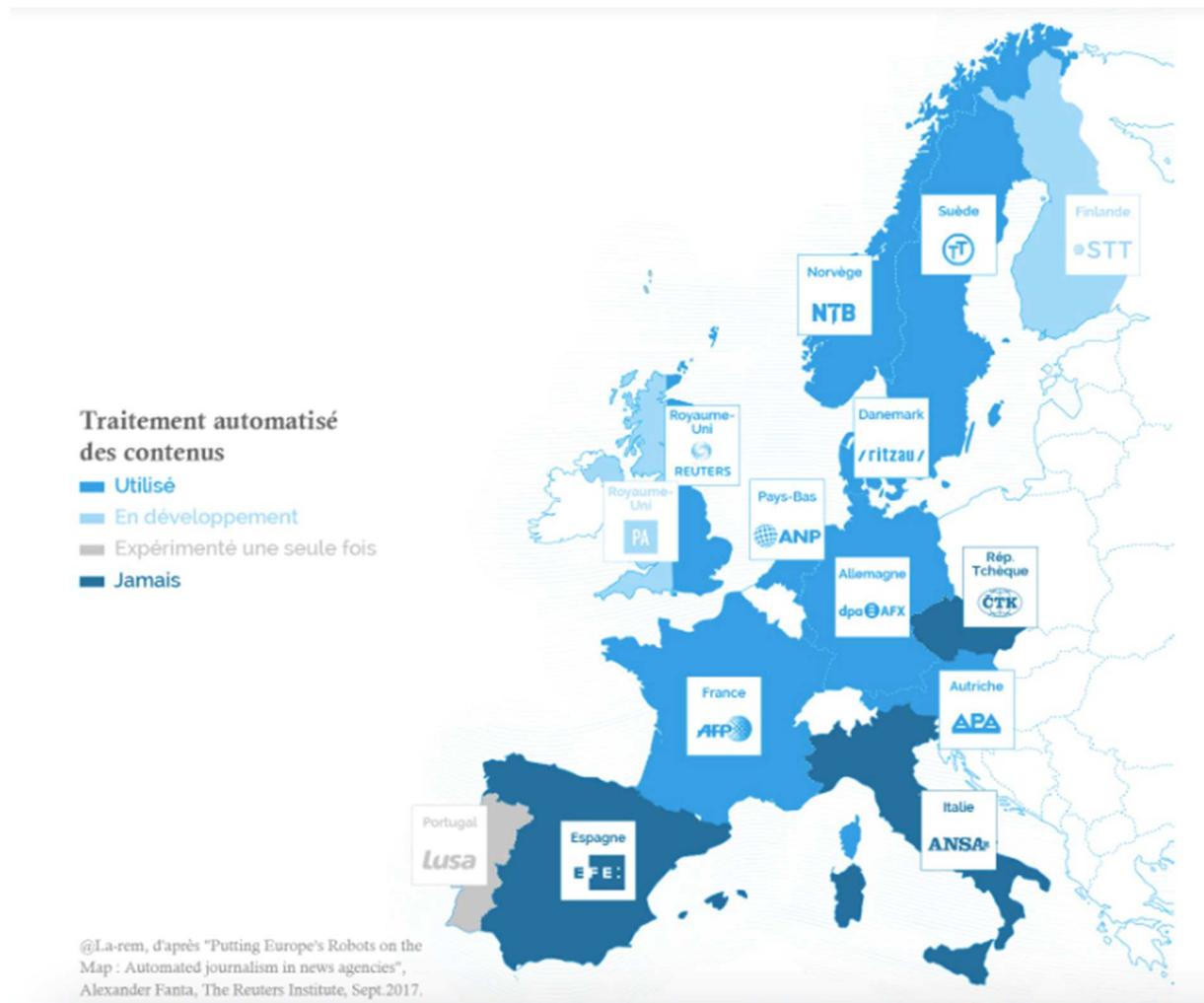
« *Le Monde est arrivé avec une idée très précise de ce qu'ils voulaient : c'était vraiment de produire un texte correspondant aux résultats au moment où ils tomberaient, le soir des élections en temps réel, et de le publier sur le site. C'était très précis, très cadré. On n'a pas développé un outil spécialement pour eux mais on a paramétré notre moteur de rédaction en fonction des directives données par Luc, pour qu'il produise les types de textes qu'il voulait. On était vraiment contraints dans ce qu'on pouvait faire, dans ce qu'on pouvait employer, l'ordre, la syntaxe des phrases.* »

Avec l'outil de GAT de Syllabs, *Le Monde* produit ce soir là 36 000 articles en temps réel sur les résultats des 36 000 communes et cantons concernés par le scrutin – sans aucune erreur. Une performance que le journal n'aurait probablement pas pu réaliser autrement, d'autant qu'elle a permis aux journalistes de se consacrer aux analyses des résultats.

Depuis, la pratique a essaimé. Une enquête menée auprès des agences de presse généralistes en Europe, y compris *Reuters*, devenue canadienne mais installée à Londres, et *Associated Press*, dont le siège social est à New York, a été publiée par le Reuters Institute for the Study of Journalism de l'Université d'Oxford. Sur un total de quinze agences qui ont accepté de répondre, trois seulement n'utilisent pas de

⁶ Cf annexe n°2 p.49

robots, sous aucune forme : la tchèque CTK, l'espagnole EFE et l'italienne ANSA. Au sein de l'agence de presse portugaise LUSA, on explique n'y avoir eu recours qu'à l'occasion des élections parlementaires de 2015, même si l'on est bien conscient que cette technologie représente l'avenir du journalisme. Neuf agences au total ont recours à des algorithmes pour produire des dépêches (sur le sport et la finance principalement), quand deux autres développent des projets de cette nature, comme le montre l'illustration ci-dessous :

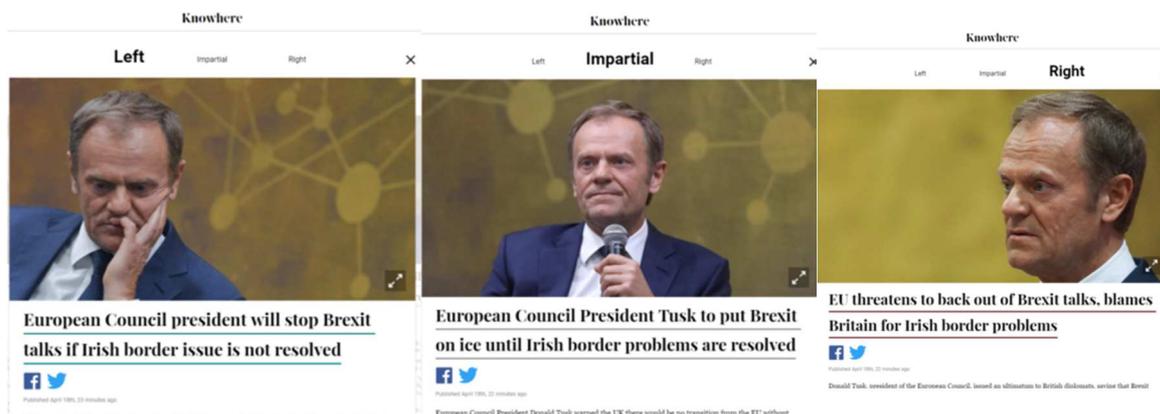


Une myriade de dénominations pour une même technologie

S'il fallait apporter un début de réponse à la question « Est-ce que les robots vont remplacer les journalistes ? » à l'aune de ces seuls résultats, il serait tentant de répondre par la négative. On imagine mal en effet de simples algorithmes produisant des articles factuels grâce à des bases de données remplacer toute la panoplie de qualités que les journalistes mettent en œuvre au quotidien. Mais d'autres cas

d'usage plus récents dans les rédactions de presse écrite pourraient bien inverser la tendance. Il s'agit pour le premier du lancement en 2018 du site Knowhere News, qui « mise sur l'intelligence artificielle pour enlever les biais humains » comme il le précise sur son site internet. Comme l'explique Eric Scherer sur son blog Méta-Media, Knowhere utilise l'intelligence artificielle un peu différemment.

Elle sert tout d'abord à identifier les sujets les plus populaires du moment sur la toile : une fois le thème sélectionné, l'IA va parcourir plus d'un millier de sources différentes (des sites vérifiés et priorisés selon un indice de confiance) et créer une base de données du sujet en question. Vient ensuite la phase de rédaction automatique de l'article « sans biais humains », qui se concentre sur les faits corroborés par plusieurs publications. Ce processus prend aujourd'hui entre une et quinze minutes. Deux éditeurs relisent ensuite le billet et en améliorent la grammaire (ce qui permet au passage à l'algorithme de progresser). La publication revient à Nathaniel Barling, cofondateur de la startup et éditeur en chef. L'IA va jusqu'à proposer trois perspectives – « gauche, droite, impartial » – pour les articles politiques et pour les sujets controversés, et un titre « positif, négatif ou impartial ». L'image de Une reflète parfois ces nuances :



Il ne s'agit ici donc pas de quelques sujets qui seraient traités par des algorithmes, mais de tout un média dont les publications sont rédigées par des programmes informatiques. Et si le résultat final de ces collaborations revêt à chaque fois une forme journalistique – les articles – peut-on pour autant parler de robot-journalisme ? C'est la question qui divise les différents acteurs, et les membres de notre corpus. Si pour Robin Grovik de MittMedia il s'agit bien de journalisme, pour Claude de Loupy,

fondateur de Syllabs en France, il conviendrait plutôt de parler de “rédaction automatisée”. Quant à Laurence Dierickx, journaliste indépendante et chercheuse, elle préfère parler sur son blog OhMyBox.info de « production automatisée d'informations » :

« Les « robots journalistes » n'ont rien de journalistes, sauf si on limite le travail de ces derniers à la rédaction d'articles. Ce ne sont pas non plus des robots au sens machine physique animée : ce ne sont pas humanoïdes, pas même un bras articulé qui taperait sur un clavier. En réalité, il s'agit davantage de logiciels ou d'algorithmes, d'une force de calculs programmée. »

« Plutôt que d'IA, il faudrait mieux parler d'outils intelligents permettant de mieux faire son travail, que ce soit dans la production d'informations ou dans leur diffusion », estime quant à lui le créateur du robot Flint Benoît Raphaël, dont le projet vise à sélectionner parmi les millions d'articles produits quotidiennement les informations pertinentes et de qualité en fonction des centres d'intérêt de chacun.

Robot-journaliste, algorithme de rédaction, GAT, journalisme automatique... Les dénominations varient mais le principe reste le même. Au départ simple assistants de rédaction, les algorithmes se font peu à peu une vraie place dans les rédactions. Avec un accueil pour le moins mitigé chez les humains.

B) Une réception en demi-teinte

Des journalistes tentés mais inquiets

Les expériences menées par MittMedia et *Le Monde* dans le domaine de l'automatisation l'ont montré : les outils de GAT peuvent servir de support aux journalistes dans leurs routines quotidiennes, en les libérant de tâches chronophages. Rien d'étonnant donc à ce que dans ces deux cas d'usage, la réception par les journalistes ait été plutôt bonne. Défendant l'automatisation de la

rédaction pour la libérer des tâches ingrates et mieux servir son audience, Robin Govik assure qu'après les avoir testés, les journalistes sont désormais ravis de travailler avec des robots :

« La plupart ne les voient pas comme une menace car pour eux, les robots ont un style ennuyeux et trop froid, presque clinique. Ils se sont beaucoup exprimés sur ces questions de style et d'écriture, mais pas tellement sur le phénomène en lui-même. Certains les considèrent même comme des collègues et leur donnent des images et des prénoms. Les avantages l'ont largement emporté sur les craintes qu'ils pouvaient avoir. »

Depuis 2016, les journalistes du groupe MittMedia collaborent donc quotidiennement avec leurs collègues robots. « Jusqu'ici nous pensions que les journalistes, traditionnellement conservateurs, détestaient les robots et les jugeaient non fiables. En fait, ce n'est pas vrai : en deux semaines, leurs a priori sont tombés », abondait la journaliste finlandaise Hanna Tuulonen, chercheuse à l'université d'Helsinki sur l'automatisation de l'information. On serait tentés d'expliquer cet optimisme par le fait que le journalisme automatisé est relativement répandu en Suède comme en Finlande, et que dans ces pays la technologie est globalement perçue positivement. Un phénomène déjà relevé par le *New-York Times* dans un article intitulé « Robots are coming, and Sweden is fine » qui citait une étude de la Commission Européenne menée en 2017 au sujet de l'accueil des robots et de l'intelligence artificielle dans toute l'Europe⁷ : en Suède, 80% des personnes interrogées avait donné un avis favorable. Robin Govik reprend les arguments de l'article pour expliquer l'accueil favorable fait à ces technologies :

« Je crois que les médias et les Suédois eux-mêmes s'adaptent vite aux nouvelles pratiques journalistiques, et voient facilement leur potentiel. Les pays nordiques se démarquent lorsqu'il s'agit des nouvelles technologies et leurs usages, et je crois que ça a beaucoup à voir avec notre société, notre code du travail et notre système de santé : nous prenons soin les uns des autres, la communauté est très forte, et

⁷ « Attitudes towards the impact of digitization and automation on daily life » Eurobarometer survey, 2017.

l'opinion selon laquelle si un robot est mieux placé qu'un humain pour effectuer une certaine tâche, alors qu'il le fasse, est assez répandue. À mon sens les autres pays d'Europe sont encore assez conservateurs à l'égard de la production automatisée de contenus. »

Mais il reste étonnant de constater que les journalistes n'ont relayé aucune critique ni aucune limite relativement au travail de ces « collègues » en deux ans à leurs côtés. On pourrait supposer que le fait que Robin Govik, rapporteur de l'expérience, soit un employé direct du groupe MittMedia n'y est pas entièrement étranger. En réalité, s'il n'aborde pas les limites de son outil, il reste plutôt lucide sur son avenir à court terme : « Oui je pense que les robots finiront par remplacer les journalistes, mais que cela n'arrivera pas de mon vivant. Les robots d'aujourd'hui ne sont pas capables de remplacer les journalistes : il y a encore trop de choses qu'ils ne savent pas faire – pour le moment. Ils ne peuvent pas encore se construire eux-mêmes, il faut des humains pour leur donner des instructions sur comment, pourquoi, etc. » Lui-même n'est pas journaliste, mais sa position stratégique en tant que chargé de la stratégie digitale au sein du deuxième plus gros groupe de presse suédoise reste intéressante pour notre travail : il est convaincu que les robots permettent d'alléger la charge de travail des journalistes, sans pour autant être en compétition avec eux. Au contraire, ils se complèteraient, les robots faisant ce pour quoi ils sont doués, et les humains aussi de leur côté. Il milite donc pour plus d'automatisation dans les rédactions, « afin de dégager du temps pour les journalistes ».

Sans aller jusque là, le constat que dresse Neil Thurman⁸, chercheur en journalisme computationnel à l'Institut de communication et des médias de l'Université de Munich, à l'issue de l'expérience qu'il a menée avec ses confrères Konstantin Dörr et Jessica Kunert, est relativement similaire. Regrettant le manque de données sur la réception que trouvait le robot-journalisme auprès des premiers concernés, ils ont présenté à 10 journalistes de différentes rédactions un logiciel de GAT (dont une clause de confidentialité l'oblige à taire le nom), qu'ils ont pu expérimenter pendant quelques heures. Qu'ils aient eu ou non une première expérience avec ce type d'outils, les participants se sont dits satisfaits de l'expérience, et comme au sein de

⁸ Cf annexe n°3 p.56

MittMedia, aucun des journalistes n'a exprimé de crainte vis-à-vis d'un éventuel remplacement :

« Ils voyaient bien les limites de l'outil et restaient persuadés qu'il ne serait pas prochainement en mesure de faire leur travail. Cela dit, ils ont quand même évoqué le cas des correspondants de Reuters en Inde, qui étaient plus à risques du fait de la nature de leurs tâches - des articles courts et factuels. Mais paradoxalement, la réception par les agences de presse a été plutôt positive, alors que les journalistes qui privilégient l'analyse et l'enquête dans leur travail au quotidien se sont montrés plus réfractaires. »

Si Neil Thurman relève le paradoxe d'une telle situation, celui-ci peut s'expliquer au regard des habitudes et des valeurs des journalistes : l'agencier qui rédige des dépêches se reconnaît inévitablement dans le résultat fourni par un outil de GAT (un article concis, factuel, au style sobre et direct). À l'inverse, il est aisément compréhensible qu'un rédacteur voire un journaliste d'investigation qui s'attachent à multiplier les sources, à choisir ses mots avec soin, à utiliser des effets de style quand il le juge nécessaire et qui tient à cette façon de travailler, puissent se montrer réfractaire à l'utilisation d'une telle technologie. Pourtant, si elle ne peut pas l'aider comme elle peut décharger son collègue agencier des tâches les plus ingrates, elle devrait, en théorie, être moins perçue comme une menace puisque pour l'instant incapable de reproduire son travail – et donc moins susceptible de le remplacer.

Car tous les participants à l'étude de Neil Thurman n'ont pas manqué de tomber d'accord sur un point : les limites techniques relatives à l'usage d'un outil de GAT pour rédiger des articles. Pour eux, le résultat final serait forcément limité puisque que la plateforme ne prenait en compte qu'une seule source – ce que ne font généralement pas les journalistes. Ils ont aussi pointé du doigt le besoin de devoir prédire l'angle de l'article en amont. En effet, l'outil de génération automatique de texte utilisé dans l'enquête, comme beaucoup, fonctionne sur un modèle dit « déterministe », qui exige que tous les cas de figure soient envisagés afin de définir son cadre de fonctionnement. Il fallait donc que les journalistes anticipent tous les angles possibles pour permettre au logiciel de rédiger l'article le moment venu. De

cette deuxième limite en découle une autre, autrement plus gênante pour les journalistes : l'impossibilité de rajouter des éléments de contexte. La plateforme ne le permettait pas, puisqu'elle ne peut intégrer que des informations pour lesquelles elles disposent de données. Style froid, absence de diversité des sources, rigidité du cadre technique et impossibilité de contextualiser : voilà les reproches que formuleraient donc les journalistes à l'égard de leurs collègues non-humains.

En France, *Le Monde* fait véritablement figure de pionnier dans le paysage médiatique français lorsqu'il décide de collaborer avec Syllabs pour couvrir les élections départementales de 2015. Lui qui travaille avec les médias sur les sujets de la sémantique et de l'informatique depuis le début des années 2000, constate le manque d'appétence de leur part pour toute collaboration. Comme ailleurs dans le monde, il semblerait que l'heure de l'IA ne soit pas encore venue, et les rédactions se montrent frileuses à l'idée d'embarquer un robot en leur sein. Pour Claude de Loupy, le journal fondé par Hubert Beuve-Méry a donc pris un vrai risque :

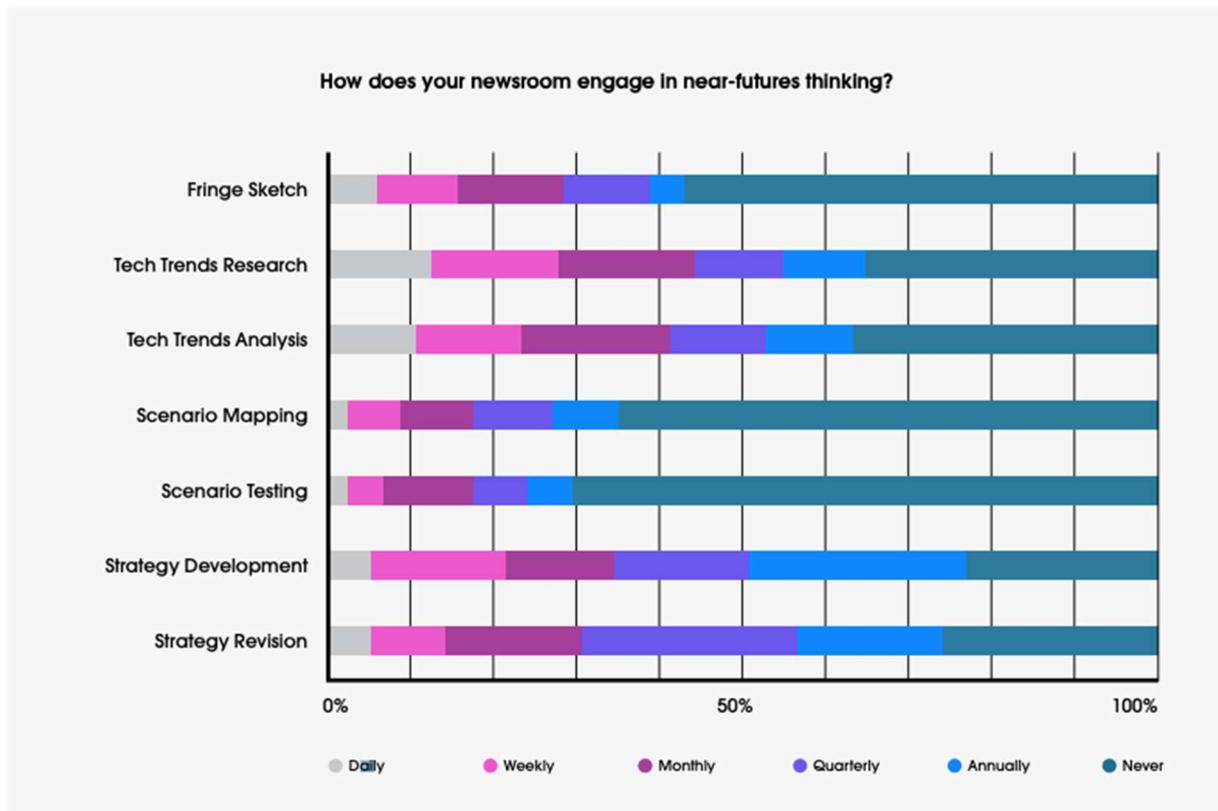
« Être les premiers en France à faire de la robot-rédaction, le premier au monde à rédiger des textes automatiquement sur un événement aussi sensibles que les élections, c'était vraiment très courageux de leur part. On dit toujours les médias ne sont pas innovants, mais c'est totalement faux, le problème des médias ce n'est pas d'innover - il y a souvent de très grand innovateurs dans les médias et ce sont même les premiers sur la plupart des technologies - le problème c'est l'exploitation de l'innovation après, c'est là où souvent c'est plus compliqué. »

Le fait qu'ils aient initié eux-mêmes la démarche prouve en tous cas qu'ils n'y étaient pas totalement réfractaires – au moins les journalistes responsables du projet, à commencer par Luc Bronner. Dans *Les robots vont-ils remplacer les journalistes ?* (D. Desbordes), le directeur des rédactions du *Monde* insiste sur le fait que ces articles ne sont ni des analyses ni des reportages, et n'ont en ce sens aucunement vocation à remplacer les journalistes qui fournissent un travail minutieux et de qualité. Pour lesdits journalistes, difficile de savoir ce qu'il en est. Mais l'étude des parutions journalistiques à ce sujet ces dernières années peut nous fournir quelques pistes sur l'évolution de l'opinion des journalistes. Au *Monde* par exemple, on écrivait

déjà sur l'arrivée de l'intelligence artificielle dans les rédactions avant l'expérience Syllabs, les premiers articles datant de 2014 (année marquée par le premier cas d'usage de robot-journaliste dans une rédaction avec la naissance de QuakeBot au *LA Times*). « Un robot pour contrôler la qualité du travail des journalistes », « Robots, les algorithmes prennent la plume », « Chez Associated Press, le robot repère et le journaliste analyse » ... On y décrit les premiers cas d'usage sans s'inquiéter outre-mesure des conséquences : manque de recul ou inconscience ? s'interroge Damien Desbordes dans son ouvrage. Selon lui, ils ne semblent pas prendre réellement la mesure du phénomène, « abordant cette peur sur le ton ironique, comme s'ils croyaient avant tout en leur irremplaçabilité technique, sinon ontologique (C. Fleury, 2015) ». Il cite pour preuve une analyse discursive réalisée sur plus de 200 textes concernant les robots-journalistes entre 2010 et 2017, avec un fil rouge : que pensent les journalistes ? « Dans la grande majorité des cas (63%), la tonalité du titre opte pour un traitement neutre, et les journalistes abordent la relation homme-machine. Loin derrière, dans seulement 2,67% des articles étudiés, on trouve la thématique de l'éthique. En France, le taux descend même à 1,16%. »

Les journalistes manqueraient-ils de lucidité sur l'avenir de ces technologies ? Pèchent-ils par manque de recul ou par excès d'orgueil ? Difficile de répondre, mais les chiffres sont significatifs, et le postulat de Damien Desbordes corroboré par une enquête du Future Today Institute menée en 2017⁹ : « Peu de rédactions se préoccupent du futur à long terme de l'information à l'ère prochaine de l'IA, de l'automatisation et de la reconnaissance vocale : 53% avouent penser rarement aux prochaines 10 à 20 années du journalisme. 78% confient ne jamais s'engager dans aucune réflexion ou prévision pour le futur de l'information de quelque manière que ce soit. »

⁹ The Global Survey on Journalism's Futurs, Future Today Institute.



Source : Future Today Institute, 2017

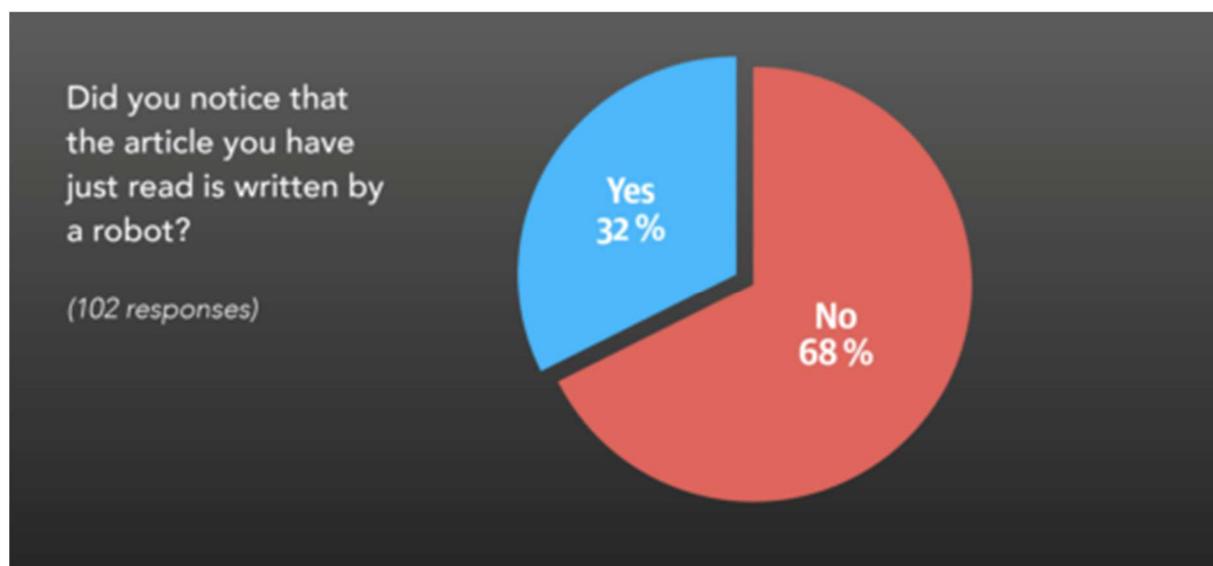
Au vu de l'explosion du nombre d'articles de presse portant sur les robots-journalistes sous toutes leurs formes, il semblerait qu'il y ait eu un début de prise de conscience. Il n'y a qu'à regarder notre bibliographie pour constater la multiplication de ces articles qui s'inquiètent de la disparition des journalistes au profit de l'IA. Comment expliquer un tel changement de mentalité en si peu de temps ? On peut supposer que le gain d'exposition dont ont bénéficié ces technologies – et l'intelligence artificielle plus généralement, avec une emphase sur ses effets – ait joué. Ainsi de l'élection hautement médiatisée (pour son profil atypique autant que ses recherches) du mathématicien spécialiste de l'IA Cédric Villani, qui a contribué à faire porter ces questions dans la sphère publique. Depuis, les études et les groupes de réflexion à ce sujet se sont multipliés, et les journalistes s'en sont logiquement fait l'écho. Mieux informés, ils auraient revu leur position. De là à envisager d'y avoir recours ? C'est en tous cas le constat que dresse Claude de Loupy dans un article de l'INA Global intitulé « Production automatique de textes : l'IA au service des journalistes » : « Nous vivons un très fort accroissement de notre activité et cela se ressent beaucoup dans le domaine des médias. Nos interlocuteurs sont de plus en plus au fait des enjeux du numérique et comprennent beaucoup mieux notre

proposition de valeur. » Avant de relativiser : « Certains journalistes nous voient comme une menace. Le terme même de robot rédaction (et encore plus de « robot journaliste ») suscite la méfiance. »

Des lecteurs peu informés mais critiques

Mais qu'en pensent les lecteurs ? Et d'abord, se sont-ils rendus compte de l'arrivée de ces robots-journalistes ? Pour le savoir, MittMedia a conduit sa propre étude il y a quelques mois :

« Nous faisons attention à l'opinion de nos lecteurs sur ces nouvelles pratiques. Il s'est avéré que malgré notre transparence, une grande majorité d'entre eux ne s'en rendaient pas compte. »



Source : « Des robots-journalistes déjà bien meilleurs que les vrais ! » Meta Media, 2018

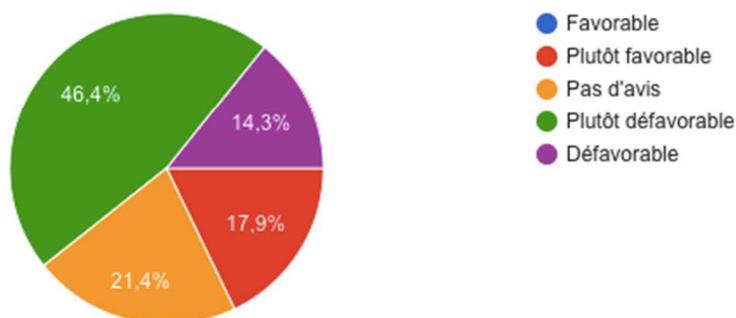
Comme les journalistes, ils se plaignent du style, qu'ils trouvent moins bon que celui des journalistes « humains ». Mais c'est bien là la seule critique dont fait état MittMedia. La « technophilie » des Scandinaves pourraient là encore être une des pistes à privilégier pour expliquer la bonne réception des robots-journalistes auprès du lectorat suédois. Et en France ? Faute d'avoir pu recueillir les retours de lecteurs du Monde suite à son expérimentation avec Syllabs à l'occasion des élections

départementales en 2015, et au vu de l'absence de littérature sur le sujet, aussi criante que sur la réception des journalistes, nous avons conduit notre propre questionnaire relayé sur Facebook et Twitter au mois d'octobre. Ici, le profil type du répondant est celui d'une étudiante de moins de 30 ans, habitant à Paris, qui consomme des médias de presse écrite (papier et web) au moins une fois par jour. Si l'échantillon est loin d'être représentatif, il permet tout de même de se faire une idée de l'opinion des lecteurs sur le robot-journalisme dans l'Hexagone. Et comme chez les journalistes, il existe un réel décalage entre la réception scandinave et celle qui est faite à l'IA en France. Sur les 28 répondants, plus de la moitié n'ont jamais entendu parler des robots-journalistes. Ce seul chiffre pourrait d'ailleurs peut-être expliquer leur position. Ceux qui connaissaient le concept en avaient entendu parler dans les médias ou dans un cadre lié au journalisme (cours sur le sujet, etc), mais seuls 17,9% avaient déjà lu un article écrit par un algorithme.

Cela dit, même sans avoir entendu le terme auparavant, les personnes interrogées semblent avoir très bien saisi le concept : « une IA qui va poser les questions de base OU QUAND QUI A QUI QUOI ? », « Des contenus d'"information" écrits non par des individus mais par des robots informatiques », « automatisation du travail des journalistes », « Données triées et hiérarchisées par IA plutôt que par [un] humain », « Génération automatique de texte, généralement pour présenter des statistiques ou des données structurées ». Il convient ici de prendre en compte le biais de méthodologie principal, à savoir qu'il a été relayé majoritairement par des journalistes, auprès d'un public probablement plus averti donc. Et si les sondés sont familiers avec le concept, ils sont très majoritairement opposés à l'idée, comme en témoigne le graphique ci-dessous :

Quelle est votre opinion sur les articles écrits par des robots ?

28 réponses



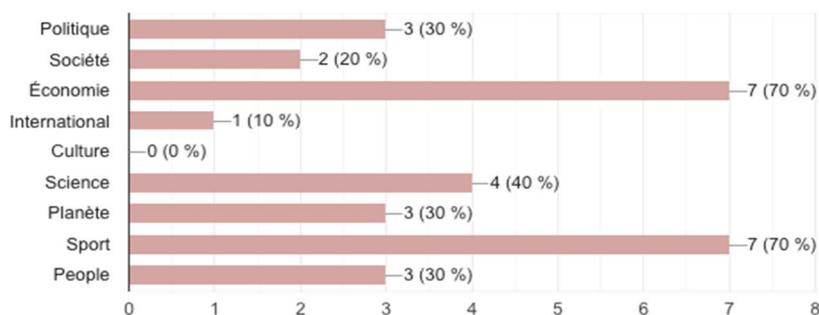
Source : Enquête menée auprès de 28 personnes sur Facebook et Twitter au cours du mois d'octobre 2018

Pourtant, ils semblent conscients des avantages que cela pourrait présenter : « gain de temps, d'argent, meilleure productivité, véracité et objectivité des informations, une masse salariale moindre et donc moins de frais, un orthographe parfait, une plus grande réactivité quand une news tombe ; des articles mieux ciblés pour plaire aux lecteurs, disparition de la hiérarchie, présenter des données sous un format plus lisible ... ». Mais visiblement, cela ne suffit pas à contrebalancer la – longue – liste des inconvénients qu'ils lui trouvent : « Moins fiable, problème de style rédactionnel, sélection des sujets biaisée, absence de subtilité, ne comprend pas nécessairement le contexte, ne recoupe pas les faits, informations basiques, faibles possibilités de travailler sur le long terme avec des sources, mort du métier, articles pas intéressants, perte d'audimat, rendre le média trop lisse, absence de légitimité, risque d'accroissement des fake news, incapacité à reproduire le ressenti des lieux en reportage, perte de contrôle de la rédaction sur l'éditorial si l'on fait appel à des prestataires externes... ». On notera que les questions de l'emploi des journalistes et du traitement des informations sont celles qui reviennent le plus. Encore une fois, le fait qu'un nombre certain de répondants ont un lien plus ou moins fort avec le secteur est à prendre en compte dans l'analyse de ces résultats, et peut expliquer la prépondérance des remarques concernant directement les enjeux déontologiques et économiques du robot-journalisme.

Si les trois quarts des répondants déclarent ne pas être plus enclins à lire des articles écrits par des robots, les autres privilégieraient visiblement des thématiques qui s’y prêtent d’un point de vue du traitement des données (comme nous le verrons plus loin) :

Si oui, sur quels sujets ?

10 réponses



Enfin les répondants concluent là encore à l'écrasante majorité (78%) que non, les robots ne vont pas remplacer les journalistes, pour toutes les raisons présentées comme des inconvénients plus tôt. Un chiffre qui devrait être rassurant pour les journalistes, mais qui trouve tout de suite sa limite étant donné le caractère binaire de la question. Par ailleurs si les avantages et les inconvénients inhérents au robot-journalisme sont bien connus des lecteurs, on peut tout à fait imaginer qu'une fois exposés à des productions variées de GAT, les lecteurs puissent modifier leur position sur le sujet.

Alors que l'IA assoit sa présence dans les rédactions de presse écrite à travers le monde, l'accueil réservé par les journalistes et les lecteurs reste mitigé – et sujet à des variations selon les pays, qui s'expliquent notamment par le contexte socio-culturel. Au-delà des limites techniques, le recours croissant aux algorithmes dans le journalisme soulève un certain nombre de questions déontologiques et juridiques, et ses conséquences économiques et sociales sont loin d'être anodines.

II – LES ENJEUX DE L'INTELLIGENCE ARTIFICIELLE APPLIQUÉE AU JOURNALISME

A) Un cadre juridique et déontologique flou

On l'a vu, l'arrivée des outils de génération automatisée de textes dans les rédactions bouleverse véritablement les pratiques journalistiques. Elle pose dans le même temps des questions cruciales pour le champ journalistique, que doivent prendre en compte les acteurs sous peine d'aggraver une situation déjà délicate. Car ses conséquences, que ce soit sur la relation avec le public ou avec les sphères économiques, sont loin d'être anodines.

Qui parle quand une IA écrit ? Droit d'auteur et responsabilité

C'est en effet l'un des enjeux fondamentaux soulevés par l'arrivée de l'IA dans les rédactions : la transparence sur l'identité de l'auteur. Faut-il ou non prévenir les lecteurs quand l'auteur de l'article n'est pas humain ? Chez MittMedia, on ne s'est pas posé la question longtemps : il a unanimement été décidé que chaque article écrit par leur outil de GAT serait signé comme tel, avec la mention « par le robot-rédacteur de MittMedia ». Une décision facile à prendre dans un contexte comme celui-ci, où la technologie est perçue plutôt positivement. Dans l'enquête qu'ont menée Neil Thurman et ses collègues à l'université de Londres, il y a aussi eu un consensus sur l'importance de déclarer l'identité de l'auteur. Mais le chercheur craint que face à des retours moins positifs, les rédactions ne fassent un choix moins éthique pour limiter les dégâts :

« Je pense que c'est effectivement la bonne chose à faire. Mais cela pourrait être un vrai problème, car une étude qui n'a pas encore été publiée a montré que plus un média est transparent au sujet de la production automatisée, moins les lecteurs sont

satisfaits. C'est une vraie question d'autant que le journalisme a un passé trouble en matière de droit à l'auteur et de transparence, notamment avec tous ces médias qui reprennent sans vergogne les brèves des agences de presse pour les publier telles quelles sur leur site, parfois en créditant même un de leur propre journaliste comme auteur. Peut-être que l'histoire sera amenée à se répéter si la production automatique de contenus par les rédactions décolle, qui sait ? »

La question de la transparence sur l'identité de l'auteur trouve aussi une résonance toute particulière en ces temps de défiance envers les médias et à l'heure des *fake news* : nul besoin de venir ajouter de l'huile sur le feu en laissant planer le doute à chaque article quant à l'identité de l'auteur, quand bien même les lecteurs seraient réfractaires à cette technologie. Si les médias veulent faire accepter à leurs lecteurs l'existence de robots-rédacteurs, c'est à eux de leur prouver en quoi cette technologie peut œuvrer au service d'une information de qualité, plus proche de leurs attentes. Pour la journaliste et chercheuse Laurence Dierickx, c'est un devoir des journalistes : "La pire pratique est de ne pas mentionner que le contenu (texte ou vidéo) est généré de manière automatique. Chaque jour, nous pouvons être exposés à des contenus automatisés sans le savoir. Il ne s'agit pas du droit de savoir. C'est une question de confiance entre les médias et leurs publics, entre les médias et leurs journalistes."

Et qui dit auteur dit aussi droit d'auteur et responsabilité juridique. Or n'existe pas à ce jour de cadre légal ou réglementaire dédié à l'intelligence artificielle, tant à l'échelle nationale qu'européenne ou internationale. Néanmoins, de nombreuses initiatives en France, en Europe et à l'étranger ont été prises afin de déterminer si le droit existant était applicable à l'intelligence artificielle ou s'il convenait de mettre en place un régime juridique dédié, parmi lesquelles la démarche Stratégie France IA, mise en place par le gouvernement en janvier 2017, ou la publication du rapport « Pour une intelligence artificielle maîtrisée, utile et démystifiée » par l'Office parlementaire d'évaluation des choix scientifiques et technologiques du Sénat deux mois plus tard. Au niveau européen, dans une résolution législative contenant des recommandations à la Commission européenne sur les « Règles de droit civil sur la robotique » et l'intelligence artificielle adoptée le 16 février 2017, les députés européens soutiennent qu'un projet législatif est urgent pour mettre en place des

normes éthiques et clarifier les règles juridiques. En effet, il ressort des considérants de la recommandation que « l'autonomie des robots pose la question de leur nature à la lumière des catégories juridiques existantes ou de la nécessité de créer une nouvelle catégorie dotée de ses propres caractéristiques et effets spécifiques ».

Et la problématique liée à la créativité de l'intelligence artificielle n'est pas oubliée des travaux actuels puisque, si les avancées technologiques rendent l'intervention humaine inutile au processus créatif, les réglementations applicables aux œuvres de l'esprit ne seront plus suffisantes à en assurer la protection. Ainsi, dès le 31 décembre 2012, euRobotics (European Robotics Coordination Action), action de coordination financée par le 7^e programme-cadre de l'Union européenne, identifiait dans son projet de « Livre vert sur des aspects juridiques de la robotique » une inadéquation des schémas juridiques du droit de la propriété intellectuelle à la protection des créations générées par une intelligence artificielle, au regard notamment de la notion d'originalité, attachée à la personne de l'auteur et à ses choix subjectifs.

À ce jour, seules les œuvres réalisées par l'homme assisté par ordinateur sont protégeables par le droit d'auteur, l'ordinateur étant appréhendé comme un outil n'étant pas exclusif de créativité humaine, dans laquelle doit être recherchée l'empreinte de la personnalité de l'auteur. En revanche, tel n'est pas le cas de l'œuvre réalisée de façon autonome par une intelligence artificielle bénéficiant d'une capacité d'analyse de l'environnement, d'apprentissage et de subjectivité qui lui permettent de faire des choix, comme les outils de GAT utilisés actuellement dans les rédactions.

De même, dans le rapport contenant des recommandations à la Commission européenne concernant des règles de droit civil sur la robotique du 21 janvier 2017, la commission des affaires juridiques du Parlement européen soutient que « la définition de critères de "création intellectuelle propre" applicables aux œuvres protégeables par droit d'auteur créées par des ordinateurs ou des robots est exigée », après avoir relevé qu' « il n'existe aucune disposition juridique qui s'applique spécifiquement à la robotique, mais que les régimes et doctrines existants peuvent s'appliquer en l'état à ce domaine ». Cette recommandation n'a toutefois

pas été retenue dans le rapport final adopté par le Parlement européen le 16 février 2017.

Du besoin de transparence pour préserver la confiance et l'indépendance

Dans leurs travaux, Tal Montal et Zvi Reich, de l'Université Ben-Gurion du Negev, soulignent l'importance pour les médias ayant recours à des algorithmes dans la production de l'information d'être transparent non seulement sur leur existence, mais aussi sur leur fonctionnement. En analysant la production automatisée de 12 sites de presse et en interrogeant leurs dirigeants, ils ont défini quatre niveaux de transparence : *full transparency* (avec note explicative et crédits appropriés à chaque fois), *partial transparency* (crédits appropriés mais pas de note explicative), *low transparency* (pas de note explicative, et les crédits sont vagues, laissent planer le doute) et *no transparency* (aucune information quelque'elle soit sur le sujet). Interrogée par la journaliste Madolyn Smith pour le European Journalism Center¹⁰, la chercheuse israélienne se montre pragmatique : « La transparence est un composant crucial du journalisme moderne, avec le débat actuel sur la nécessité d'être plus ouvert au public. Dans le champ du journalisme automatisé, où les "boîtes noires" des algorithmes sont utilisées à plusieurs étapes de la chaîne de production, et en considérant les conséquences potentielles de l'usage de tels algorithmes, mettre en place une vraie politique de transparence apparaît comme une nécessité. Cela permettra par ailleurs de diminuer les externalités négatives associées à l'utilisation des outils de GAT tout en aidant les journalistes et les lecteurs à comprendre leurs forces et leurs faiblesses pour définir au mieux les comportements à adopter. »

Car si les algorithmes produisent des contenus de manière automatisée, ils n'en sont pas pour autant autonomes. Derrière chaque action, chaque choix, il y a un ordre qui a été donné par un humain – avec ses opinions, ses valeurs, et autant de potentielles erreurs. Un phénomène amplifié si ces solutions ne fonctionnent pas en vase clos et fonctionnent sur la base du *machine learning* (apprentissage automatique), s'appuyant sur une sorte de réseau de neurones, qui ont la

¹⁰ « When data writes...who is the author? » Data Driven Journalism, 2016.

particularité de fonctionner de façon un peu organique. « On pourrait comparer ces réseaux de neurones au fonctionnement du cerveau d'un enfant. Il faut l'éduquer pour lui permettre de mieux comprendre », explique Benoît Raphaël¹¹, fondateur du robot Flint, une sorte d'assistant personnel que l'on va « éduquer », pour qu'il aille trouver des articles de qualité, fiables sur des sujets divers. Comme pour les journalistes humains, l'impartialité des algorithmes est un leurre : elle n'existe qu'entre les garde-fous installés par les développeurs, qui ne peuvent pallier totalement les biais d'apprentissage. Avec leur style froid et impersonnel, les productions des outils de GAT peuvent donner une fausse impression de neutralité et de distance - un danger supplémentaire que les journalistes doivent anticiper et expliquer, et auxquels le public doit être éduqué. Et ce n'est pas tout. Le recours à l'IA dans la production de l'information présente une autre limite déontologique : l'origine des données. Outre l'extrême dépendance des outils de GAT à des données de qualité – déjà évoquée par Robin Govik dans notre entretien – l'origine des données est un paramètre crucial dans l'établissement des risques et limites d'une telle technologie.

Or comme le rappelle très justement Damien Desbordes dans son ouvrage, « les données ne sont pas des informations au sens journalistique, mais elles sont devenues le nerf de la guerre de l'information et pose des questions auxquelles il est urgent de répondre. Qui contrôle les données ? Comment s'en empare-t-on ? À qui sont elles vendues et pour en faire quoi ? (...) Le journaliste est garant de la vérité dans cette étrange chaîne de l'information et ce rôle l'oblige à s'interroger sur les intérêts des fournisseurs de données. » Les journalistes doivent en effet faire un double travail pour s'assurer de leur qualité et leur authenticité avant de les utiliser, mais aussi être le plus transparent possible vis-à-vis du public afin de limiter les risques – qui se multiplient à mesure que s'allonge la chaîne de production. Et dans ce contexte, la question dépasse le seul champ de la déontologie ou même du journalisme : il s'agit là de s'interroger sur les implications démocratiques de ces technologies, qui entre les mauvaises mains pourraient faire des ravages bien supérieurs aux *fake news* - de par leur complexité et leur force de frappe. Evgeny Morozov, spécialiste des implications politiques et sociales du numérique, s'inquiétait

¹¹ « IA, des robots pour mettre de l'ordre dans le chaos de l'info » INA Global, 2018.

déjà, en 2012, de la possibilité de voir un jour les GAFAs investir dans le marché du 'journalisme automatique', et écraser tous les petits acteurs du secteur. Champions de la personnalisation et des algorithmes de recommandation, ces géants du Web pourraient être tentés de proposer aux médias des IA capables d'écrire et de modifier des articles instantanément, de "les personnaliser afin de les adapter aux intérêts et aux habitudes intellectuelles du lecteur"... Jusqu'à enfermer celui-ci dans des "bulles de filtres" - celles-là mêmes qui sont au coeur de la critique visant, depuis l'élection de Donald Trump, le réseau social Facebook.

Une crainte qui pourrait être devenue réalité, puisque Google vient de se lancer dans la GAT avec son projet RADAR (Reporters And Data And Robots). Porté par le groupe de presse Press Association et financé par la forme de Mountain View pour près d'un million de dollars, RADAR a pour objectif de produire quelque 30 000 articles d'information locale originaux chaque mois. RADAR s'inscrit dans un projet plus large de coopération entre l'Homme et la machine, PAIR (People + AI Research). « Par cette approche 'humaine' de l'intelligence artificielle, Google entend réconcilier les âmes et les algorithmes, les processeurs et les cerveaux » analyse Damien Desbordes dans son ouvrage. Éprise d'IA, la firme plébiscite les « initiatives d'innovation digitale pour l'actualité » à hauteur de 150 millions d'euros. Il s'agit d'un appel à projets lancé en 2015 pour soutenir la modernisation des médias en Europe. L'un des plus gros financements de ce fonds d'investissements s'est retrouvé, en juillet 2017, entre les mains de la startup britannique Urbs Media, avec l'instruction d'utiliser l'intelligence artificielle pour créer du contenu à peu de frais. C'est ainsi que le projet RADAR pour le traitement de l'actualité est né.

Alors qu'elle se fait de plus en plus présente dans les rédactions, l'IA manque cruellement d'un cadre juridique. Son utilisation par les médias pose aussi de sérieuses questions déontologiques, qui ne sont pas sans conséquences sur leur relation avec le public et leur rôle démocratique. Faute de moyens, les médias risquent de céder aux sirènes des acteurs privés qui voient dans la donnée un nouveau business – sans égard pour le coût social d'une telle technologie.

B) Un business lucratif mais pas sans conséquences

Peut-on vraiment parler de business alors que le marché émerge à peine ? Oui, au vu de l'explosion de l'offre de données et de leur intérêt croissant. Pourtant, une entreprise de presse qui souhaiterait faire appel aux services d'un robot pour enrichir sa plate-forme web ne pourrait choisir que parmi une vingtaine d'acteurs à l'échelle mondiale. Ce qui fait dire à Damien Desbordes qu'il s'agit d'un vrai oligopole : « [le marché] se partage entre AX Semantics en Allemagne, Narrativa en Espagne, Syllabs (et Yseops, NDLR) en France, sans oublier les deux pionniers américains, Narrative Science et Automated Insights, et quelques autres. Tous ces codeurs de robots ne s'expriment pas nécessairement dans toutes les langues. Ils ne s'intéressent pas aux mêmes types de produits (comptes rendus financiers, articles sportifs, rapports d'entreprises, etc.) et tous n'ont pas les médias pour clients, ce qui contribue à amenuiser une offre déjà limitée. »

Le business de la data

Tous, en revanche, fonctionnent avec le même carburant : la data. Sans elle, point de salut. Et on l'a vu, la question de l'accès à des données de qualité est cruciale. Le risque, pour les médias et pour la démocratie plus globalement, c'est donc que cette technique se développe entre les mains de quelques uns, qui seraient alors en mesure de fixer les tarifs en contrôlant à la fois l'offre et la demande. Une situation qui n'est pas si éloignée de la réalité : aujourd'hui, les GAFAMI contrôlent déjà le « robinet à data ». Et pour elles, ce flux est un jackpot. Le chiffre d'affaires du *big data* à l'horizon 2019 est estimé à 187 milliards de dollars soit 50% de plus qu'en 2015 (T. Roux, 2017). « Une économie naissante se structure autour de ce minerai, aussi abondant et précieux que l'eau. Des sociétés se spécialisent dans la récupération de données brutes et en font leur commerce », explique Damien Desbordes dans son ouvrage. L'industrie de la data a ainsi généré près de 21 milliards d'euros pour la seule année 2016, selon Gartner. Un environnement en or pour les robots-journalistes, et pour les entreprises qui les développent. La firme américaine Narrative Science a ainsi enregistré des financements de capital-risque à

hauteur de 17,5 millions de dollars en 2017, quant son concurrent français Syllabs levait en mars dernier 2 millions d'euros auprès de l'accélérateur et fonds brestois West Web Valley, du groupe Ouest-France et du fonds BNP Paribas développement. « Nous avons deux à trois ans d'avance sur la concurrence au niveau mondial et une fenêtre de tir que nous ne pouvons pas laisser passer », explique Claude de Loupy, président et co-fondateur de Syllabs¹².

Si leurs besoins en fonds de développement sont si importants, c'est que ces entreprises de production automatisée ont de grandes ambitions. Car les médias ne sont pas les seuls à s'intéresser à leur expertise : d'autres clients ont tôt fait de voir les multiples opportunités que présentaient les services de tels prestataires. Outre *Le Monde*, *Le Parisien*, *Les Echos*, *Slate* ou *France TV*, Syllabs travaille ainsi pour différents acteurs du tourisme, du e-commerce ou de l'immobilier qui ont des besoins en contenus pour leurs sites web comme *voyages-sncf.com*. L'année dernière, il s'est associé avec le groupe Guy Hoquet et l'agence 1min30 pour développer un robot-rédacteur et produire des contenus locaux et personnalisés pour les clients de ses 450 agences immobilières françaises.

Des économies pour les médias mais à quel prix ? L'emploi

S'ils ne sont pas les clients les plus aisés, les médias n'en sont pas moins fidèles. Car les outils de GAT leur permettent de faire de sérieuses économies tout en gagnant en productivité : sur son site, Automated Insights se vante de capacité de production prodigieuse (1 milliards de textes produits en 2014, 2000 productions à la seconde). Au service de l'agence américaine Associated Press (AP), il a permis de faire passer le nombre de dépêches financières de 300 à 5000 sans délai ni main d'œuvre supplémentaire. Un constat partagé par Robin Govik de MittMedia, qui ne cache pas son enthousiasme :

¹² « Syllabs lève 2 M€ pour déployer ses robots rédacteurs sur de nouveaux marchés » L'Usine Digitale, 2016.

« À l'époque, on dépensait 300.000 euros par an pour payer les freelances qui rédigeaient ces articles de sport, donc il y avait aussi un vrai potentiel d'économies si nous arrivions à industrialiser ce processus de production. Avec la chute de 25% des effectifs dans les rédactions suédoises ces dernières années, les machines produisent les contenus que les journalistes ont arrêté de fournir mais que les gens réclament. »

Alors qu'il ne nous coûte que la moitié du salaire d'un journaliste du même niveau, le robot-rédacteur de MittMedia dédié à l'immobilier est leur employé le plus performant : depuis son lancement, il a déjà publié plus de 10 000 articles et détient aussi le meilleur taux de conversion de pages vues en abonnement. L'opération est un tel succès que MittMedia a décidé de passer à l'échelle supérieure et d'industrialiser son outil de production. Le groupe suédois s'est associé avec un autre journaliste, Sören Karlsson, et une compagnie spécialisée dans les données sportives, pour fonder l'entreprise United Robots. « Nous avons décidé de construire un produit qui pourrait être développé à l'échelle du pays entier, et pas seulement au sein de MittMedia », raconte Robin Govik. Aujourd'hui, une grande partie des médias suédois sont des clients de United Robots, et publient les contenus produits par leur outil de génération automatique de texte. Au delà de la visibilité, ces partenariats sont une vraie source de revenus pour MittMedia, qui peut réinvestir les profits réalisés pour continuer à se développer.

Mais qu'en est-il des médias aux moyens plus limités ? Au vu des prix pratiqués par les fournisseurs d'outils de GAT (20 000 euros pour France Bleu avec Syllabs à l'occasion des élections régionales, soit l'équivalent de 3 personnes à temps plein pendant 2 mois), se pose la question de l'égalité et de l'accessibilité. D'autant qu'en plus des journalistes, il faut bien souvent compter avec une équipe d'ingénieurs, développeurs ou data-scientiste pour accompagner les journalistes dans l'exploitation des données. Le recours généralisé à ces technologies risque donc de creuser des inégalités de moyens déjà criantes entre les médias, les plus imposants étant les seuls en mesure de financer de tels investissements. Cela dit, il est probable qu'avec la généralisation de ces solutions, leur prix finisse par diminuer, et devenir plus abordable pour les rédactions. À tel point qu'elles préféreront investir dans ces outils plutôt que d'embaucher ? C'est la question à laquelle tous les acteurs

du champ tentent de répondre. Les conséquences de ces technologies sur l'emploi des journalistes reviennent régulièrement dans les inconvénients cités par les répondants de notre enquête.

Dans une logique de réduction des coûts, qui prime dans bon nombre d'entreprises de presse, il n'est pas inenvisageable que plus de rédactions fassent le choix d'automatiser ce qui peut l'être. S'il est risqué et difficile de faire de telles prédictions, certains s'y essaient tout de même. Ainsi pour Kris Hammond, directeur de Narrative Science, 90% des articles de presse seront rédigés par des robots d'ici 2030. Il allait même jusqu'à anticiper la victoire du prix Pulitzer, qui récompense chaque année les meilleures productions en journalisme, littérature, fiction et musique, par un robot-rédacteur en 2016. L'avis de son concurrent Kaleb Jesse, directeur commercial chez Automated Insight, n'est pas si éloigné : « 90% des articles rédigés par des robots en 2030 ? D'un point de vue technique, c'est réaliste. Mais il faut encore énormément d'intervention humaine pour atteindre ce résultat. Peut-être qu'un outil capable de remplacer complètement les journalistes verra le jour... dans une centaine d'années. » Gardons en tête que ces acteurs ont tout intérêt à croire que le recours aux outils de GAT se généralise, dans les rédactions ou ailleurs, puisque c'est leur source principale de revenus.

Pour autant, Peter Clifton, rédacteur en chef de RADAR, et Robin Govik de Mitt Media, se veulent rassurants : les journalistes humains qualifiés sont et resteront vitaux dans le processus de production de l'information. Chez RADAR, ils sont au nombre de cinq. C'est exceptionnellement peu au vu des 30 000 articles publiés chaque mois. Dans son ouvrage, Damien Desbordes cite le président de l'Union nationale des journalistes du Royaume-Uni, qui confiait ses inquiétudes au *Guardian* : « J'ai peur que les journaux se contentent de diffuser des articles sans panache générés par des machines pour avoir l'air intéressants, tout en se débarrassant de leurs vrais reporters. » De quoi craindre le pire pour le futur économique des journalistes ? Pas pour Claude de Loupy :

« Je n'ai pas d'espoir que cela se développe de manière réellement forte dans les 3 ans qui viennent, c'est une certitude - ni même dans les 5-10 prochaines années -

les freins sont trop nombreux. Il y évidemment le frein psychologique, la peur d'être remplacé par une machine, qui joue beaucoup. Il y a aussi la lourdeur même des médias qui sont de grosses machines. Et puis il y a évidemment le fait que les médias sont une industrie en crise, une crise très grave. Il y a une telle lutte pour la survie actuelle que ce n'est pas du tout évident d'aller regarder des innovations qui en plus sont compliquées d'un point de vue psychologiques à faire accepter en interne. »

Une étude belge de 2015 citée par Damien Desbordes dans son ouvrage ne classait pas les journalistes parmi les professions à risque (A. Baert et P. Ledent). Mais à cette époque les cas d'usage de production automatisée dans les rédactions étaient rares. L'apparition des robots-rédacteurs a-t-elle changé la donne ? Visiblement pas : les journalistes ne figuraient pas non plus dans la liste des cinq métiers susceptibles de disparaître au XXI siècle, d'après une étude menée par le think tank Institut Sapiens, fondé par le Dr Laurent Alexandre : « *Ce sont près de 2,1 millions d'actifs concentrés dans ces cinq métiers qui ont une forte probabilité de voir leur emploi disparaître dans les prochaines années* », est-il écrit dans l'étude¹³ publiée en août 2018.



Source : Institut Sapiens, 2018

¹³ « L'Impact de la révolution digitale sur l'emploi », Institut Sapiens, 2018.

Pour établir ce classement, l'Institut a pris en compte les métiers qui sont à la fois directement remis en question par une technologie et qui ont vu leurs effectifs diminuer depuis trente ans – ce qui explique pourquoi les journalistes n'y figurent pas, du moins pas encore. Car comme on le constate dans l'étude, les métiers les plus à risques ne sont pas que des métiers manuels. D'après Raja Chatila, professeur à l'université Paris-Sorbonne et directeur de l'Institut des systèmes intelligents et de robotique (ISIR), la situation actuelle n'a rien à voir avec la révolution industrielle, quand l'automatisation des mécanismes matériels a remplacé le travail à la chaîne : « Aujourd'hui, ce sont des systèmes qui sont plus performants en termes cognitifs, en termes d'intelligence artificielle, ce ne sont pas exactement les mêmes emplois qui risquent d'être transformés, supprimés, ce sont des tâches qui sont traditionnellement plus effectuées par des personnes qui ont un certain niveau d'éducation. Ce n'est pas nécessairement une question de perte d'emploi, c'est une question de transformation de tâches, donc parfois si l'emploi est exclusivement sur une tâche automatisable, il risque d'il y avoir une perte d'emploi. »¹⁴

L'autre risque de l'introduction de la technologie et de l'automatisation croissante des tâches que souligne Raja Chatila, c'est la baisse des salaires et la détérioration des conditions de travail : « la partie de l'automatisation effectuée par les machines va réduire la qualification du travail, et donc amener à une déqualification des personnes, et probablement à une baisse des salaires et, en même temps, le fait de mettre en place des systèmes automatisés risque de réduire les conditions de travail de la personne humaine, qui deviendrait guidée, esclave de la machine. » De quoi fragiliser encore la situation économique des journalistes, eux qui doivent déjà faire face à une grande précarité au quotidien, avec l'augmentation du recours aux contrats à durée déterminée (CDD) et aux piges (malgré la réduction de leur volume et de leur prix). Dans son ouvrage, Damien Desbordes ne dit pas autre chose : « Le journalisme est un métier où le salariat n'est plus la norme. L'indépendance éventuelle des pigistes se paie au prix de leur précarité. Difficile de s'insurger contre l'arrivée d'une machine dans l'équipe quand le rédacteur en chef n'a qu'à arrêter de vous passer des commandes pour ne plus entendre parler de vous. Laurence Diericx

¹⁴ « L'intelligence artificielle va-t-elle détruire des emplois ? » France Culture, 2018.

a étudié le dossier Syllabs pour ses recherches. Elle a rencontré un journaliste, collaborateur régulier du *Monde.fr*, dont le quota de piges a fondu. «Ils ont été prévenus deux jours avant. On n'est pas en position de force en tant que pigiste. » Les journalistes ne peuvent que craindre l'arrivée de ces nouveaux collègues corvéables à merci, dénués de besoins, de droits, de sentiments, d'ego – des employés que certains pourraient qualifier de parfaits en somme.

CONCLUSION GÉNÉRALE

Tout au long de ce travail, nous avons cherché à comprendre quelle était la place de l'intelligence artificielle dans les rédactions de presse écrite. En s'intéressant aux différents cas d'usages dans les rédactions françaises et étrangères, nous avons abordé d'autres thématiques (la réception par les journalistes et les lecteurs, le cadre juridique et déontologique, les conséquences sur l'emploi) qui permettent de mieux comprendre l'enjeu actuel du sujet. Nous avons aussi élaboré et cherché à vérifier trois hypothèses de recherche.

Nous l'avons vu, les outils de génération automatique de texte (GAT) utilisés par les rédactions de presse écrite leur permettent de rédiger des articles factuels sur des sujets bien délimités (l'immobilier, le sport, la finance) en raison notamment de leurs limites techniques. Pour le reste, ce sont toujours les journalistes qui sont chargés de réaliser les reportages, les enquêtes, ou tout autre article ne se présentant pas sous la forme d'une brève. L'hypothèse n°1 est donc validée.

Ensuite, les journalistes et les lecteurs restent partagés quant à l'arrivée de ces technologies dans les rédactions : si tous reconnaissent l'intérêt qu'elles peuvent présenter pour réduire les coûts et libérer les journalistes des tâches chronophages à faible valeur ajoutée, les lecteurs comme les rédacteurs regrettent la pauvreté du style, les capacités limitées et les conséquences sur l'emploi des journalistes. L'hypothèse n°2 est donc validée.

Enfin, entre les limites techniques, l'absence de cadre juridique et déontologique et les freins psychologiques, et malgré les investissements importants des acteurs du secteur, il semblerait que le recours généralisé aux outils de GAT dans les rédactions de presse écrite ne soit pas pour demain. L'hypothèse n°3 est donc validée

Si nos trois hypothèses ont été validées, cela ne signifie pas pour autant que la réponse à la question du remplacement des journalistes par les robots coule de source. On serait tentés de répondre par la négative, puisque les éléments avancés au fil de ce travail de recherche montrent bien qu'il existe encore un grand nombre de limites, que ce soit du côté des outils eux-mêmes ou de leur utilisation. Mais ces

considérations ne valent que pour le moment où nous parlons. Car la technologie progresse vite et l'IA pourra bientôt permettre bien plus que de simples brèves, mettant sérieusement en danger les postes de journalistes dans tous les types de médias. Si génération automatique de textes par des robots est, pour l'heure, la fonction la plus couramment développée, la détection des informations et la production vidéo sont également concernées par le développement des algorithmes. À l'agence Reuters, depuis septembre 2018, c'est Wibbitz qui met en ligne les petites vidéos. Cette startup israélienne, qui s'est également associé à TF1, transforme automatiquement les mots en images. Pour fonctionner, l'outil de Wibbitz ne requiert ni monteur, ni journaliste : il lui suffit de se connecter à une banque de données, et elle produit une courte vidéo d'une minute, qui se prête particulièrement aux réseaux sociaux (un format qui a fait le succès de médias comme Brut, et que beaucoup imitent depuis). Même les reporters ne sont pas à l'abri de se voir remplacer par des machines : des chercheurs de l'université de Tokyo ont développé un véritable robot-journaliste capable d'explorer son environnement de manière autonome, et de rendre compte de ce qu'il trouve sur le terrain. Le robot, qui détecte tout changement autour de lui pour décider si l'information est pertinente, peut prendre des photos, interroger les passants, et utiliser internet pour mieux comprendre son environnement, avant de publier automatiquement son article. Mais est-ce là tout le travail d'un reporter ? Un journaliste n'est-il qu'un canal, un tuyau qui délivre de l'information ? Sans parler des limites techniques – les robots ont encore du mal à trouver du sens dans le chaos informationnel, la plupart des liens qu'ils tissent sont prévisibles, ils ne savent pas profiter des concours de circonstances – qui pourraient disparaître à mesure que progresse la recherche, l'arrivée des robots-journalistes pose des questions qui ont trait au cœur même du métier, à sa raison d'être, à ses principes clés et à ses valeurs.

Et si justement, ce bouleversement à venir était l'occasion pour le champ journalistique de se remettre en question, de régler des problèmes qu'il rencontre depuis un moment déjà ? Anticiper la vague plutôt que de la subir permettrait aussi d'éviter des effets catastrophiques comme ont pu l'être pour certaines rédactions l'arrivée des journaux gratuits et l'essor du web. D'autant qu'entre peur, rejet et optimisme béat, il existe une voie du milieu : une collaboration saine et éclairée entre les journalistes et les robots-rédacteurs, où les rédactions ne prendraient que le

meilleur de ces technologies. C'est d'ailleurs ce que prêchent nombre d'acteurs, des chercheurs aux journalistes. Ensemble, robots et journalistes pourraient unir leurs forces pour produire un travail de qualité, dépasser les limites matérielles, techniques, chronologiques ou géographiques qui entravent parfois la pratique journalistique, et mettre au point de nouveaux formats. De quoi permettre in fine de sortir le journalisme de l'ornière dans laquelle il se trouve actuellement, entre *fake news*, manque de moyens et crise de confiance ? Cette collaboration potentiellement salvatrice supposerait un vrai travail en amont, notamment au niveau de la formation des journalistes à ces outils. Or à ce jour, à l'exception de quelques cursus spécialisés (à la Columbia Journalism School depuis 2012 et à l'Université de Cardiff depuis 2014), les possibilités pour les futurs journalistes de se former au journalisme computationnel restent maigres. Autant de pistes qu'il serait intéressant d'explorer dans un travail annexe.

BILIOGRAPHIE INDICATIVE ET SOURCES

Ouvrages académiques

Les robots vont-ils remplacer les journalistes ? Damien Desbordes, Edition du Plein du jour, 2018

Robot Journalism: Can Human Journalism Survive? Noam Lemelshtrich Latar, 2018

Articles universitaires et études

« Du statut d'objet technique à celui d'objet social dans le contexte de la production automatisée d'informations journalistiques » Laurence Dierickx, Panorama des innovations sociales numériques, *Terminal*, 2018

« When Reporters Get Hands-on with Robo-Writing: Professionals consider automated journalism's capabilities and consequences » Neil Thurman, Konstantin Dörr & Jesia Kunert, 2017

Articles de presse

« L'intelligence artificielle va-t-elle détruire des emplois ? » France Culture, 11/08/2018

<https://www.franceculture.fr/emissions/les-idees-claires/lintelligence-artificielle-va-t-elle-detruire-des-emplois>

« Une IA dans ma rédac » Delphine Soulas-Gesson, *Stratégies*, 13/07/2018

<http://www.strategies.fr/actualites/medias/4012754W/une-ia-dans-ma-redac.html>

« Les robots remplaceront-ils les journalistes ? » Romain Serre, *ECS Digital*, 27/05/2018

<https://ecs-digital.com/culture/les-robots-replaceront-les-journalistes/>

« Robot-journalisme : pour survivre, la profession a besoin d'un cadre déontologique » Nelly Lesage, Numerama, 20/05/2018

<https://www.numerama.com/tech/349630-robot-journalisme-la-profession-a-besoin-dun-cadre-deontologique.html>

« Des robots-journalistes déjà bien meilleurs que les vrais ! » Eric Scherer, Meta Media, 15/03/2018

<https://www.meta-media.fr/2018/03/15/des-robots-journalistes-deja-bien-meilleurs-que-les-vrais.html>

« Syllabs lève 2 M€ pour déployer ses robots rédacteurs sur de nouveaux marchés » L'Usine Digitale, 7/03/2018

<https://www.usine-digitale.fr/article/syllabs-leve-2-m-pour-deployer-ses-robots-redacteurs-sur-de-nouveaux-marches.N663224>

« IA et médias : une innovation plus qu'une révolution » Jean-Dominique Serval, INA Global, 12/02/2018

<https://www.inaglobal.fr/numerique/article/ia-et-medias-une-innovation-plus-qu-une-revolution-10084>

« IA, des robots pour mettre de l'ordre dans le chaos de l'info » Isabelle Didier et Philippe Raynaud, INA Global, 10/02/2018

<https://www.inaglobal.fr/numerique/article/ia-des-robots-pour-mettre-de-l-ordre-dans-le-chaos-de-l-info-10088>

« Production automatique de textes : l'IA au service des journalistes » Isabelle Didier et Philippe Raynaud, INA Global, 9/02/2018

<https://www.inaglobal.fr/numerique/article/production-automatique-de-textes-l-ia-au-service-des-journalistes-10092>

« L'intelligence artificielle est une menace sérieuse pour le journalisme » La Tribune, 8/10/2017

<https://www.inaglobal.fr/numerique/article/ia-et-medias-une-innovation-plus-qu-une-revolution-10084>

« #ProprieteIntellectuelle : Le robot-journaliste met-il les médias en danger ? »

Laetitia Levasseur, Jurilex Blog, 6/10/2017

<http://www.jurilexblog.com/proprieteintellectuelle-le-robot-journaliste-met-il-les-medias-en-danger-266089>

« Les robots font du journalisme » Nicolas Prouillac, L'Obs, 16/12/2016

[https://www.nouvelobs.com/rue89/rue89-nos-vies-](https://www.nouvelobs.com/rue89/rue89-nos-vies-connectees/20161216.RUE6025/les-robots-font-du-journalisme.html)

[connectees/20161216.RUE6025/les-robots-font-du-journalisme.html](https://www.nouvelobs.com/rue89/rue89-nos-vies-connectees/20161216.RUE6025/les-robots-font-du-journalisme.html)

« Ces journalistes entraînent des IA à démasquer la corruption », Ulyces, 13/12/2016

<https://www.ulyces.co/jonathan-stray/ces-journalistes-entraiment-des-intelligences-artificielles-a-demasquer-la-corruption/>

« Algorithmes et robots journalistes : l'avenir de la presse ? » Fabien Soyez, Cnet, 06/12/2016

<https://www.cnetfrance.fr/news/algorithmes-et-robots-journalistes-l-avenir-de-la-presse-39845608.htm>

« When Data Writes... Who Writes ? » Data Driven Journalism, 10/10/2016

http://datadrivenjournalism.net/news_and_analysis/when_data_writeswho_is_the_author

« Les quotidiens nationaux se sont-ils définitivement transformés en “médias numériques” ? » RFI, 16/09/2016

<http://atelier.rfi.fr/profiles/blogs/metamorphose-medias-nationaux-quotidiens-lemonde-letemps>

« Innovation : avec les robots, la fin des journalistes ? » Le Parisien, 11/12/2014

<http://www.leparisien.fr/magazine/grand-angle/innovation-avec-les-robots-la-fin-des-journalistes-11-12-2014-4366153.php>

« Non, messieurs les journalistes, les robots ne voleront pas vos emplois » Claude de Loupy, L'Obs avec Rue 89, 09/05/2014

<https://www.nouvelobs.com/rue89/rue89-tech/20140509.RUE3752/non-messieurs-les-journalistes-les-robots-ne-voleront-pas-vos-emplois.html>

« Robots contre journalistes : qui gagnera la guerre de l'information ? » Ph V-D, L'Obs avec Rue 89, 4/05/2014

<https://www.nouvelobs.com/rue89/rue89-tech/20140504.RUE3638/robots-contre-journalistes-qui-gagnera-la-guerre-de-l-information.html>

« Robot Journalist Takes Pictures, Asks Questions, Publishes Online » Aaron Saenz, Singularity Hub, 18/03/2010

<https://singularityhub.com/2010/03/18/robot-journalist-takes-pictures-ask-questions-publishes-online/#sm.00000jwua0jpsyf9tr88d0soebj7k>

« L'ère des robots-journalistes » Yves Eudes, Le Monde, 09/03/2010

https://abonnes.lemonde.fr/actualite-medias/article/2010/03/09/l-ere-des-robots-journalistes_1316608_3236.html?

Pages web

Knowhere News, <https://www.deepnews.ai>

United Robots, <http://unitedrobots.ai/>

Syllabs, <https://www.syllabs.com/fr/>

RÉSUMÉ ET MOTS-CLÉS

Mots-clés : Journalistes, intelligence artificielle, rédactions, presse écrite, technologie, robots

Résumé : *Machine learning*, big data, algorithmes... Depuis quelques années, ces termes sont venus s'inviter dans toutes les conférences, dans toutes les rédactions et dans tous les esprits, sans que l'on sache bien ce qu'ils recouvrent. L'intelligence artificielle, qui consiste à mettre en œuvre un certain nombre de techniques visant à permettre aux machines d'imiter une forme d'intelligence réelle, est en effet un vaste domaine qui touche non seulement à l'informatique mais aussi aux mathématiques, à la neuroscience et même à la philosophie. Depuis le début des années 2010, les algorithmes ont fait leur apparition au sein même des rédactions, apportant avec eux la crainte de voir un jour les journalistes disparaître. Pourquoi ? Comment ces algorithmes fonctionnent-ils ? Qu'en pensent les premiers concernés, les journalistes ? Et les lecteurs ? Quels sont les enjeux de l'utilisation de tels outils ?

Key words: Journalists, artificial intelligence, newsrooms, written press, technology, robots

Summary: Machine learning, big data, algorithms... For several years, these terms have grown to be wildly used at keynotes, within newsrooms and conversations, without a clear idea of what they imply. Artificial intelligence, which aims by various ways to make machines mimic some form of real intelligence, is a vast area that applies not only to computer sciences, but also to mathematics, neuroscience and even philosophy. Since the early 2010s, algorithms have started to make their way into the newsrooms of written press across the globe, and with them the fear that journalists will eventually disappear. Why? How do they work? What do journalists feel about them? And readers? What is at stake here?

ANNEXES

Annexe n°1 : Entretien avec Robin Govik, responsable de la stratégie digitale du groupe MittMedia (réalisé le 12 octobre 2018 et traduit de l'anglais)

Maëlle Lafond : Pouvez-vous revenir sur l'introduction des robots au sein de la rédaction de MittMedia ?

Robin Govik : Nous avons commencé à recourir à la production automatisée il y a quelques années, à toute petite échelle. Jusqu'à présent notre site Internet n'était qu'une copie de ce que l'on publiait en version papier, mais nous voulions que cela change. En 2016, nous avons donc pivoté vers une stratégie centrée sur l'acquisition de lecteurs via des abonnements en ligne, en adaptant nos contenus web, persuadés qu'il fallait voir chaque support comme ayant ses caractéristiques propres. Pour y parvenir, nous avons besoin de savoir quels articles transformaient les lecteurs en abonnés, et lesquels les faisaient rester sur le site.

Il faut savoir que nous avons un écosystème de différents services au sein de MittMedia, via notamment une plateforme qui enregistre les données relatives aux articles, aux lecteurs et à leurs interactions. Grâce à cela nous sommes capables de savoir quel type de contenus nous devons produire, il nous suffit d'extraire les données. Nous avons donc pu constater que les faits divers fonctionnaient bien, mais pas autant que l'immobilier ou les entreprises locales. Or pour produire plus d'articles sur ces thématiques, nous pouvions évidemment demander aux journalistes, mais c'était plus logique de se lancer dans la production automatisée. Nous avons donc commencé à faire rédiger des articles pour toutes les transactions immobilières du pays, et le robot est vite devenu notre « employé » le plus productif, si l'on peut parler en ces termes. Il totalisait le plus grand nombre de pages vues derrière un paywall, et avait le taux de conversion lecteur-abonné le plus élevé.

La météo était aussi un sujet fondamental pour nous, étant donnée la rudesse du climat nordique, et c'est pourquoi nous avons commencé à publier des articles tous les jours à ce propos. C'est du contenu facile à produire étant donné que les

données météorologiques sont faciles à traiter par les robots tout en étant peu intéressantes pour les journalistes – ce qui a permis qu'ils ne se sentent pas menacés. Aujourd'hui, nous avons donc deux formes de production automatisée : les contenus qui sont directement publiés sur notre plateforme, et ceux qui sont envoyés en brouillon aux journalistes. Parmi les premiers, nous disposons de trois robots : un pour les transactions immobilières, un pour les entreprises locales, et un pour le sport. Ce dernier est chargé de rédiger un article avant et après chaque match, avec 3 versions différentes à chaque fois : une version d'après le point de vue de l'équipe locale, une version neutre et une version d'après le point de vue de l'équipe adverse, car évidemment l'angle n'est pas le même selon le profil du lecteur.

Nous avons aussi des robots en beta qui ne sont pas encore prêts à être lancés auprès de nos lecteurs, sur la circulation, les transports en commun et la météo. Pour l'instant nous nous contentons de transmettre les contenus qu'ils produisent à nos journalistes, qui n'ont plus qu'à copier/coller avant de publier. Mais nous ne nous sommes pas arrêtés là : nous avons vite réalisé qu'il nous fallait améliorer nos articles sportifs, et avons décidé d'y ajouter des commentaires des entraîneurs et des joueurs. Pour cela nous leur envoyions automatiquement des messages textes à la fin des matchs pour obtenir leurs impressions, et leur réponse était automatiquement intégrée aux articles qui étaient postés dans la foulée. Évidemment nous avons dû mettre en place un système de modération pour vérifier les commentaires avant la publication, mais la relation de confiance tissée avec nos interlocuteurs permet de limiter les risques d'erreur.

Il est évident que nous faisons attention à l'opinion de nos lecteurs sur ces nouvelles pratiques. C'est pourquoi nous avons réalisé une enquête il y a quelques mois afin d'identifier la part de nos lecteurs qui se rendaient compte que les articles étaient écrits par des journalistes – sachant que nous sommes très transparents : lorsque c'est le cas il est clairement indiqué en lieu et place de la signature « par le robot-rédacteur de MittMedia ». Il s'est avéré que malgré cela une grande majorité d'entre eux ne s'en rendaient pas compte.

ML : À quelle initiative les robots-rédacteurs ont-ils intégré la rédaction de MittMedia, et comment cela s'est-il passé ?

RG : C'était une idée de mon équipe et moi-même pour répondre au problème que nous rencontrions alors : on voyait que la demande en information était bien plus grande que ce que nous étions capable de produire, et que ça n'allait pas aller en s'arrangeant puisque le nombre de journalistes n'allait certainement pas augmenter, bien au contraire. Il nous est donc apparu évident d'automatiser ce qui pouvait l'être. À l'époque on dépensait 300.000 euros par an pour payer les freelances qui rédigeaient ces articles de sport, donc il y avait aussi un vrai potentiel d'économies si nous arrivions à industrialiser ce processus de production.

ML : Envisagez-vous de continuer à recruter des journalistes malgré le développement du recours aux robots-rédacteurs ?

RG : Je crois que les robots permettent d'alléger la charge de travail des journalistes, mais qu'ils ne sont pas en compétition avec eux. Au contraire, ils se complètent : les robots continueront à faire ce pour quoi ils sont doués, et les humains aussi de leur côté. Je crois aussi qu'il faut automatiser tout ce qui est en mesure de l'être, afin de dégager du temps pour nos journalistes. Nous sommes un journal local, il est donc important que nous soyons sur le terrain auprès de nos lecteurs à faire tout le travail qui ne peut pas être automatisé.

ML : Comment avez-vous trouvé la technologie, et comment l'avez-vous implémentée ?

RG : Le département R&D, dont je suis à la tête, est assez important – de l'ordre d'une cinquantaine de personnes environ, dont 30 sont des développeurs web de différente sorte. Nous avons donc chargé l'un d'eux de se pencher pendant une semaine sur le développement de la première version de notre robot météo. C'était très expérimental, mais quand nous avons réalisé que c'était possible de passer à l'échelle supérieure, nous avons vite cofondé United robot, en partenariat avec son CEO actuel, Sören Karlsson, et une compagnie spécialisée dans les données sportives. Nous avons décidé de construire un produit qui pourrait être développé à l'échelle du pays entier, et pas seulement au sein de MittMedia. Et aujourd'hui, une grande partie des médias suédois sont des clients de United Robots, et publient nos

contenus. Ça a donc été un franc succès, nous permettant dans le même temps de nous développer grâce aux bénéfices réalisés.

ML : Quels ont été les freins à l'implémentation de ces robots-rédacteurs, et les réactions ?

RG : Au début les journalistes se disaient « *Ce n'est pas si grave, ils sont très basiques, mon emploi ne pourra jamais être remplacé par un robot* ». La plupart ne les voyaient pas comme une menace car pour eux, les robots avaient un style ennuyeux et trop froid, presque clinique. Ils se sont beaucoup exprimés sur ces questions de style et d'écriture, mais pas tellement sur le phénomène en lui-même. Pour ce qui est des freins techniques, la bonne nouvelle c'est qu'on ne peut pas faire la même erreur deux fois : comme pour un ordinateur, il suffit de réparer le « bug », et grâce au *machine learning* celui-ci ne se répétera pas.

Il reste cependant la question de la donnée : nous sommes extrêmement dépendants de jeu de données de qualité, et nous pouvons faire tous les efforts du monde, mais si les données sont fausses alors le robot se trompera inévitablement. Cela peut mener à des situations parfois absurdes, et nous a donc obligés à mettre en place un système d'alertes qui nous prévient que le contenu paraît problématique et limite ainsi un grand nombre d'erreurs.

ML : Prévoyez-vous de recourir plus largement aux robots-rédacteurs dans le futur, quitte à vous passer de journalistes ?

RG : Tout à fait, mais je ne pense pas que cela arrivera de mon vivant. Les robots d'aujourd'hui ne sont pas capables de remplacer les journalistes : il y a encore trop de choses qu'ils ne savent pas faire – pour le moment. Ils ne peuvent pas encore se construire eux-mêmes, il faut des humains pour leur donner des instructions sur comment, pourquoi, etc. Je parle de robot-journalisme parce que ces outils sont développés dans un esprit journalistique, avec des principes journalistiques, pour les journalistes que nous sommes. Mais je sais aussi qu'il est très difficile de prédire le futur, et qu'il vaut mieux s'en abstenir au risque de se tromper. Ce qui ne m'empêche pas de penser que si vous croyez que vous allez être remplacé par un robot, alors ce

sera probablement le cas : il faut être obtu pour ne pas voir en quoi ces outils sont sources d'opportunité plus que de dangers.

ML : Dans quelle mesure la culture suédoise facilite-t-elle l'adoption de ces technologies, au sein des médias et plus généralement dans la société ?

RG : Je crois que les médias et les Suédois eux-mêmes s'adaptent vite aux nouvelles pratiques journalistiques, et voient facilement leur potentiel. Le *New-York Times* en a même fait un article, intitulé « Les robots débarquent en Suède, mais tout va bien ». Un ressenti confirmé par une étude de la Commission Européenne en 2017 au sujet de l'accueil des robots et de l'intelligence artificielle dans le pays : 80% des personnes interrogées ont donné un avis favorable. Les pays nordiques se démarquent lorsqu'il s'agit des nouvelles technologies et leurs usages, et je crois que ça a beaucoup à voir avec notre société, notre code du travail et notre système de santé : nous prenons soin les uns des autres, la communauté est très forte, et l'opinion selon laquelle si un robot est mieux placé qu'un humain pour effectuer une certaine tâche, alors qu'il le fasse, est assez répandue.

ML : Avez-vous pu quantifier le coût de ces outils, ou les économies qu'ils ont permis ?

RG : C'est difficile de savoir exactement étant donné que l'entreprise qui les produit nous appartient, mais nous savons que le robot-immobilier est notre meilleur employé, alors qu'il ne nous coûte que la moitié de ce que coûte un journaliste du même niveau. Avec le recul je vois bien que la tendance se répand, notamment en Scandinavie, mais pas tant dans le reste de l'Europe, et je ne me l'explique pas. Certes, il y a un intérêt chez vous, mais personne ne semble prêt à prendre le train en marche : à mon sens les autres pays d'Europe sont encore assez conservateurs à l'égard de la production automatisée de contenus.

Annexe n°2 : Entretien avec Claude de Loupy, chercheur et fondateur de Syllabs, réalisé le 17 octobre 2018

Maëlle Lafond : Pouvez-vous revenir sur votre parcours et sur la naissance de Syllabs ?

Claude de Loupy : Je suis chercheur en informatique, spécialisé dans cette branche de l'IA qui est le traitement des langues, depuis 1994. Globalement j'ai plutôt été chercheur en industrie malgré quelques passages au CNRS et à l'université, et je ne retrouvais pas les choses que j'avais envie de faire dans la façon dont les choses étaient gérées dans notre domaine. Je travaillais avec Helena Blancafort, et on avait une façon d'aborder la relation de la linguistique à l'informatique qui n'était pas habituelle. Donc c'est par frustration qu'on a créé Syllabs. L'idée c'était de continuer à faire de la R&D pour nos clients sur la sémantique, tout en développant des outils d'analyse sémantique. On a commencé à regarder et on a vu qu'il y avait vraiment un marché, un besoin, donc on a développé cette technologie de moteurs de rédaction. Le premier date de 2011, et puis après on a essayé de chercher des clients, mais ça a été très, très compliqué. Il se trouve que je travaille avec les médias depuis le début des années 2000, surtout avec mon ancienne boîte, donc je leur proposais tout le temps des partenariats mais ça ne passait pas. Généralement la réaction c'était « C'est très intéressant mais ça va être compliqué côté rédaction donc on préfère ne pas y aller », jusqu'en 2015 où *Le Monde* nous a contacté pour faire les élections départementales.

ML : Comment s'est déroulé ce partenariat ? Avaient-ils une demande particulière ? Les avez-vous accompagné, ou créé un outil sur-mesure pour eux ?

CdL : *Le Monde* est arrivé avec une idée très précise de ce qu'ils voulaient, c'était vraiment de produire un texte correspondant aux résultats au moment où ils tomberaient, le soir des élections en temps réel, et de le publier sur le site. C'était très précis, très cadré. C'est Luc Bronner qui a piloté le projet, c'est vraiment lui qui en est à l'origine, qui a pris la décision, qui voulait nous voir, et c'est même lui qui nous a guidé dans la façon de rédiger les textes : il s'est vraiment énormément impliqué. On n'a pas développé un outil spécialement pour eux mais on a paramétré

notre moteur de rédaction en fonction des directives données par Luc, pour qu'il produise les types de textes qu'il voulait. La spécificité c'est que, pour pleins de raisons, et deux principales, il est nécessaire que les textes produits ne soient pas simplement des textes à trous : on a besoin d'avoir une variabilité plus ou moins importante dans les textes en questions. Les deux raisons c'est 1) le référencement dans Google, sinon ça fait du *duplicated content*, et les pages sont blacklistées donc c'est très mauvais pour le site, et de 2) pour le lecteur, parce que quand on lit un texte à trous une fois ça va, deux fois ça commence à faire beaucoup mais alors à la troisième on part vraiment mécontent, donc c'est important d'avoir de la variabilité.

Mais dans le contexte des élections départementales, la variabilité a été très limitée : Luc nous a vraiment contraints dans ce qu'on pouvait faire, dans ce qu'on pouvait employer, l'ordre, la syntaxe des phrases, et c'est tout à fait logique, parce que *Le Monde* a vraiment couru un gros risque. On dit toujours les médias ne sont pas innovants, mais c'est totalement faux, le problème des médias ce n'est pas d'innover - il y a souvent de très grand innovateurs dans les médias et ce sont même les premiers sur la plupart des technologies - le problème c'est l'exploitation de l'innovation après, c'est là où souvent c'est plus compliqué. Donc le risque était quand même grand, être le premier en France à faire de la robot-rédaction, le premier au monde à rédiger des textes automatiquement sur un événement aussi sensibles que les élections, c'était vraiment très courageux de leur part. Du coup comme il y avait un risque important - et c'est pour ça que c'était le directeur de publication qui gérait, ce qui n'a pas été le cas par la suite pour les élections régionales – on a effectivement paramétré quelque chose de spécifique pour eux.

Pour ce qui est de l'accompagnement, on a suivi ses instructions mais il fallait aussi lui donner les explications concernant les capacités du moteur, ce qu'on pouvait faire et ne pas faire, donc évidemment il y avait de l'accompagnement de notre part, ce n'était pas uniquement dans un sens. Il y a aussi eu un accompagnement technique, mais là en l'occurrence ils ont géré eux-mêmes la captation des données : pour les élections régionales, pour Radio France et les autres, c'est nous qui allions capter les données auprès du ministère de l'Intérieur, pour ensuite *pusher* les textes dans leur CMS. Là, *Le Monde* a préféré géré la data, c'était eux qui allaient chercher les

données au ministère de l'Intérieur, qui les ramenaient chez eux, les traitaient, nous les envoyaient sur notre plateforme et on renvoyait un texte en retour.

ML : Est-ce qu'il y a des soucis techniques lors de l'implémentation de ces outils ?

CdL : Alors sur les départementales il n'y a eu aucun soucis, vraiment c'était une première, on était hyper stressés, tout le monde était sur le pont jusqu'à 1h du matin, c'était vraiment intense, les deux tours, mais il n'y a eu aucun soucis technique ni de rédaction. Luc Bronner en a témoigné d'ailleurs, aucune erreur ne lui a été remontée, que ce soit par les lecteurs ou les journalistes. Nous-mêmes on ne nous a remonté aucune erreur – contrairement aux régionales pour *L'Express* où quand il y avait une égalité totale entre le deuxième et le troisième candidat, l'outil se trompait et publiait les résultats du premier et du troisième mais oubliait le deuxième. Ça on s'en est rendu compte très vite, ça a été corrigé en moins d'une demi-heure et on a régénéré les textes et c'était réglé.

ML : Vous avez donc collaboré avec d'autres médias par la suite ?

CdL : Oui, avec *L'Express*, *Radio France* et *Le Parisien*, suite à des demandes spécifiques de leur part à chaque fois qui devaient correspondre à leur ligne éditoriale.

ML : En tant que chercheur, comment vous positionnez-vous par rapport à ces technologies ? Parlez-vous d'IA, de robots ?

CdL : Déjà le problème c'est qu'on parle de robot : il se trouve que le terme était déjà dans la nature et que d'un point de vue marketing il est difficile de vendre quelque chose sans utiliser la terminologie déjà en usage, mais en fait robot ça ne correspond pas à ce que c'est. Les Américains pourraient dire *bot*, et ça correspond mieux, mais on ne peut vraiment pas parler de robot. Ce n'est pas Terminator, c'est juste un programme, un stupide bout de programme. Donc je ne lui donne aucun statut spécial, c'est juste un programme comme il en existe pour la traduction automatique, pour la retranscription de texte. On est dans le domaine de l'IA parce que selon la définition que Minski en donnait, c'est « *le domaine qui permet à des*

machines de faire une tâche qui si elle était faite par un humain, demanderait de l'intelligence. » Mais en fait quand on regarde on est entre l'intelligence de la fourmi et celle de l'abeille.

ML : Avez-vous des concurrents ?

CdL : Nous étions les premiers à le faire, depuis d'autres s'y sont mis : il y a Labson qui a été racheté, Igiop qui est spécialisé dans la finance, et plus récemment Deep Flow Realtor, spécialisé dans l'immobilier - comme on fait nous aussi, on a une branche dédiée – mais globalement nous ne sommes les 4 seuls à faire un travail un peu sérieux dans ce domaine là - de manière industrielle, parce qu'après il y a des chercheurs.

ML : Est-ce que vous pensez que les collaborations avec les médias sont amenées à se développer ?

CdL : C'est compliqué, c'est vraiment compliqué, l'industrie des médias est une industrie très compliquée. Quand on est sortis des élections départementales avec *Le Monde* où ça s'était super bien passé, on a eu que des retours positifs, *Le Monde* a resigné derrière, d'autres médias signent, on s'est dit « *C'est bon on va décoller* », et puis pas du tout : c'est extrêmement lent. C'est là où je dis que les médias comportent de très bons innovateurs qui testent des choses, et ils sont souvent les premiers, mais après pour que ça se diffuse au sien de la rédaction c'est extrêmement complexe. Donc je n'ai pas d'espoir que cela se développe de manière réellement forte dans les 3 ans qui viennent, c'est une certitude, ni même dans les 5-10 prochaines années.

ML : Quels sont les freins selon vous à ce développement ?

CdL : Il y en a beaucoup, je les ai certainement pas tous en tête, mais il y a évidemment le frein psychologique, la peur d'être remplacé par une machine, qui joue beaucoup : il y a une méfiance inhérente dans les relations qu'on a avec les journalistes qui est forte, c'est pour ça que je continue à parler, à expliquer, et je ne suis pas le seul. Il y a aussi la lourdeur même des médias qui sont soit de grosses

machines – *Le Monde* est une boîte exceptionnelle mais on a quand même 7 niveaux de hiérarchie dans une boîte de 350 personnes et quelques, c'est vraiment énorme. Donc c'est compliqué de faire avancer les choses et de diffuser au sein de structures aussi lourdes, et les petits médias n'ont généralement pas beaucoup de moyens. Et puis il y a évidemment le fait que les médias sont une industrie en crise, crise très très grave, et je suis le premier à dire aux gens « *Ça vous fait peur ? Et bien vous n'avez pas assez peur* », parce que vraiment on va dans le mur. J'étais à une réunion des agences de presse, il y avait 23 agences du monde entier, et l'agence du Canada montrait que 150 médias ont disparu en un temps incroyablement courts, c'est terrifiant, pour les médias et la démocratie c'est vraiment terrifiant. Et évidemment ce contexte extrêmement agressif pour les médias ne facilite pas ce genre de prise de décisions - certains parlent de « lapin dans les phares », et il y a peut-être un peu de ça mais je pense pas que ce soit seulement ça : il y a une telle lutte pour la survie actuelle que c'est pas du tout évident d'aller regarder des innovations qui en plus sont compliquées d'un point de vue psychologiques à faire accepter en interne. Donc c'est vraiment difficile.

ML : Mais est-ce que justement ça ne serait pas une solution pour les médias que d'emprunter cette route de l'innovation pour se sortir de la crise ?

CdL : Pour moi c'est une évidence, et je milite en ce sens. Alors je sais que vu les technos que je vends, on considère que je suis un peu le diable qui va uniformiser l'info mais au contraire, ces machines là sont un moyen pour redonner de la valeur au vrai journalisme et pas simplement à l'information. Pour moi le problème des médias de manière générale dans tout ce que j'ai pu voir et suivre depuis 18 ans que je travaille avec eux, c'est qu'en fait on se retrouve dans une course contre les moteurs de recherche, les réseaux sociaux, et c'est une course qui est perdue d'avance. S'il s'agit de donner une information, Twitter va plus vite, même Facebook peut aller plus vite dans certains cas. S'il s'agit juste de dire « Il s'est passé ça » sans une vraie analyse, alors les utilisateurs lambda de ces réseaux vont devenir producteurs d'informations, avec toutes les erreurs et les mensonges que ça comporte. Courir après l'information et vouloir couvrir toute l'information parce que ça peut rapporter des clics, c'est une folie, vraiment. Et ça se résume en une chose : demander à des humains journalistes de lutter contre des machines comme Google

en faisant des textes optimisés pour des mots et des requêtes, ça ne tient pas la route : il faut que ce soit des machines qui luttent contre des machines parce que les humains vont perdre à ce jeu-là. On dégrade le niveau de l'humain pour le ramener au niveau de la machine, pour lutter contre une machine. D'autre part, demander à une machine de lutter contre des millions de gens qui sont sur les réseaux sociaux, avec les mêmes armes, c'est absurde également.

Donc pour moi, le seul moyen pour les médias de se sauver eux-mêmes et de sauver notre modèle de démocratie c'est d'aller sur du vrai journalisme et de se concentrer sur la valeur ajoutée. Et là notre technologie va apporter quelque chose d'un peu violent, parce qu'on vend le même type d'articles qu'on produit pour *Le Monde* etc à des entreprises qui ne sont pas des médias et donc on a des gens qui vont publier de l'information, qui ne sont pas des médias : ça peut être des agences de communication, des agences immobilières, qui nous achètent des textes pour faire du volume de média, et on en aura de plus en plus. Ça veut dire quoi ? Ça veut dire que l'information de base – prenons les résultats d'un match de foot – sera partout sur le web : donc quelle est la valeur ajoutée de L'Équipe à donner cette même information ? Aucune. Certes, c'est le site de référence donc les gens vont continuer à y aller, mais quand ils vont voir que l'information se trouve ailleurs, et bien ils iront de moins en moins voir L'Équipe. Alors que derrière ils ont de vrais professionnels qui sont capables d'apporter beaucoup, beaucoup plus que cette information là. Donc s'ils veulent survivre ils n'auront pas le choix ils vont devoir se concentrer sur du journalisme et non pas de l'information – en tous cas pas facilement captable par la technologie.

ML : L'idée ce serait donc d'utiliser la technologie pour libérer les journalistes des tâches ingrates, afin qu'ils puissent se concentrer sur les contenus à forte valeur ajoutée ?

CdL : Exactement. Et encore une fois je ne suis pas en train de décrire un monde idéal où l'IA n'a pas d'impact sur nous : elle aura un impact fort sur l'emploi, et même si les études se contredisent parfois il est certain que des gens seront dans de vrais difficultés parce qu'un pourcentage important des tâches qu'ils font sera automatisé. Je ne sais pas exactement lesquels mais c'est comme sur les métiers à tisser : les

machines plus ou moins sophistiquées sont arrivées et quasiment plus personne ne fait ça à la main, et là ce sera pareil. Il y a donc un énorme travail à faire sur l'accompagnement, voire même un intérêt à des mesures comme la taxe robot. On est dans une période de mutation profonde du marché du travail, qui sera longue mais qu'il faudra accompagner, et dans le domaine du journalisme, pour l'instant je ne vois rien qui va faire qu'on aura moins de journalistes.

ML : Comment expliquez-vous que d'autres pays (les Etats-Unis, la Scandinavie) aient moins de freins que nous à l'utilisation de ces technologies ?

CdL : Il y a une donnée purement sociétale : aux US il y a beaucoup plus d'ouverture à ce type de choses. Si vous montrez le ROI d'une technologie, les gens y vont. Et il y a beaucoup, beaucoup moins de questions existentielles qui sont posées, c'est impressionnant, alors qu'on parle de business. Et c'est la grande différence entre la France et les US, ça fait partie de nos sociétés latines, c'est comme ça. Après il y a évidemment aussi, et ça c'est plutôt en notre faveur, le côté « Si ça doit supprimer des emplois, tant mieux », ils n'ont aucun scrupule là-dessus aux US, donc forcément les peurs que je dis non-fondées sur la suppression des emplois de journalistes et qui existent là-bas aussi, ne sont pas un frein pour autant : ils y vont quand même. Alors qu'en France on ne s'en fout pas.

Pour terminer, j'insisterai sur le fait que c'est très difficile pour nous mais pour tout le monde, même ceux qui ont ces peurs là, de différencier le fantasme de la réalité. Et pour moi il y a un point qui est vraiment important, c'est qu'il faut arriver à parler de l'IA en des termes simples, que tout le monde puisse comprendre. Sinon on va se retrouver avec cette incompréhension qui pollue la vie de tout le monde alors que ces outils pourraient simplifier la vie de ceux qui en ont peur.

Annexe n°3 : Entretien avec Neil Thurman, chercheur en journalisme computationnel à l'Institut de communication et des medias de l'Université de Munich (réalisé le 26 octobre et traduit de l'anglais)

Maëlle Lafond : Qu'est-ce qui vous a mené à conduire cette étude ?

Neil Thurman : Il existait déjà quelques études sur le journalisme automatisé qui étaient traitées dans les médias, et d'autres sur la réception des lecteurs, mais rien sur celle des journalistes qui en avaient fait l'expérience. Nous avons donc décidé de recruter nous-mêmes des journalistes pour avoir leur retour sur le sujet. Nous aurions pu choisir des professionnels qui avaient déjà travaillé avec cette technologie – d'ailleurs c'était le cas de certains de nos cobayes – mais nous avons préféré les étudier en condition, en leur apprenant directement à s'en servir. Pour cela nous avons choisi une plateforme de production automatisée de contenus qui a pignon sur rue, mais dont nous ne pouvons révéler l'identité, et nous avons formé 10 journalistes à l'utiliser, avant de recueillir leurs impressions.

ML : Comment cela s'est-il passé ?

NT : Très bien : c'était plutôt amusant d'apprendre aux journalistes à se servir de la plateforme, et de préparer la session en amont, parce que ça nous a demandé de nous y former nous aussi, en préparant les jeux de données. Pour cela, nous avons choisi de nous concentrer sur des données sportives, criminelles et économiques. Et il s'est avéré plutôt compliqué de travailler avec les données de sport, parce qu'un simple match de foot par exemple inclue une multitude de variables : qui a marqué, quand, comment, pour la combienième fois, qui est l'arbitre... La difficulté était de devoir intégrer toutes ces composantes dans un modèle pour que la plateforme puisse automatiser la production de l'article, du choix du titre à celui de l'angle. Or ces paramètres sont fondamentalement contextuels, il faut donc se préparer à tous les scénarios, même ceux qui sont contre-intuitifs. C'était déjà plus simple avec les données économiques, comme les résultats d'entreprises par exemple - telle compagnie a enregistré tant de profits sur telle période, que l'on peut comparer avec la période précédente – ou bien le chômage, car il y a bien moins de variables en jeu. Donc le sport a vraiment été un défi en ce sens, alors que c'est un cas d'usage répandu dans la production de contenus automatique.

ML : Quelle expérience les journalistes avaient de ce genre de technologie ?

NT : L'un d'entre eux était en poste à la BBC, mais avait travaillé auparavant avec le New-York Times au moment où ils avaient produit du contenu automatisé pendant les JO 2012. Il y avait aussi un datajournaliste qui travaillait pour un journal de PQR. D'ailleurs il y avait aussi bien des managers que des correspondants, car nous voulions pouvoir comparer les résultats : en théorie on aurait pu s'attendre à ce que ces derniers soient plus susceptibles de se sentir menacés que les managers, qui pourraient y voir une opportunité en termes de coût. Mais ça n'a pas vraiment été le cas, et aucun des journalistes n'a exprimé de crainte vis-à-vis d'un éventuel remplacement : ils voyaient bien les limites de l'outil, et restaient persuadés qu'il ne serait pas prochainement en mesure de faire leur travail. Cela dit ils ont quand même évoqué le cas des correspondants de *Reuters* en Inde, qui étaient plus à risques du fait de la nature de leurs tâches - des articles courts, factuels. Mais paradoxalement, la réception par les agences de presse a été plutôt positive, alors que les journalistes qui privilégient l'analyse et l'enquête se sont montrés plus réfractaires.

ML : Les journalistes ont-ils fait état d'autres limites en utilisant la plateforme ?

NT : Nous n'avions que 3 heures, mais cela a suffi pour qu'ils remarquent ses limites, notamment que le résultat serait limité puisque que la plateforme ne prenait en compte qu'une seule source – ce que ne font généralement pas les journalistes. Ils ont aussi pointé du doigt le besoin de devoir prédire l'angle de l'article en amont, qui venait du fait qu'il fallait prévoir un cadre pour toutes les données. Or c'est théoriquement possible, pour le sport par exemple, de prévoir tous les scénarios, mais ça prend du temps et donc ça a un coût. Il faut donc se demander si on sera en mesure de rentabiliser ces technologies, ça peut difficilement être un outil auquel on a recours ponctuellement. L'autre point évoqué était l'importance du contexte, un élément que ne prenait pas en compte la plateforme : or pour tous les journalistes c'était quelque chose de nécessaire, par exemple un commentaire après un match de foot. Or ils ne pouvaient pas le faire puisqu'ils n'avaient pas les données pour. Evidemment il y a d'autres solutions qui offrent ce genre de possibilités, mais cela souligne que ça n'a rien de répandu, et que toute information ne se prête pas à un tel traitement.

ML : Ont-ils fait état de préoccupations éthiques, déontologiques ou juridiques ?

NT : Nous avons en effet abordé la question de la transparence, qui a à voir avec l'éthique : faut-il mentionner que l'article a été produit de manière automatique, à quel point faut-il expliquer la démarche ? Il y a eu un consensus sur l'importance de déclarer l'identité de l'auteur, et je pense que c'est effectivement la bonne chose à faire. Mais cela pourrait être un vrai problème, car une étude qui n'a pas encore été publiée a montré que plus un média est transparent au sujet de la production automatisée, moins les lecteurs sont satisfaits. C'est une vraie question d'autant que le journalisme a un passé trouble en matière de droit à l'auteur et de transparence, notamment avec tous ces médias qui reprennent sans vergogne les brèves des agences de presse pour les publier telles quelles sur leur site, parfois en créditant même un de leur propre journaliste comme auteur. Peut-être que l'histoire sera amenée à se répéter si la production automatique de contenus par les rédactions décolle, qui sait ?

Annexe n°4 : Questionnaire à destination des lecteurs sur leur opinion au sujet du robot-journalisme (relayé sur Facebook et Twitter en octobre 2018)

« Dans le cadre de mon mémoire de fin d'étude au Celsa, je m'interroge sur l'opinion du public sur les robots-journalistes dans la presse écrite, et votre avis m'intéresse ! »

- *Avez-vous déjà entendu le terme "robot-journalisme" ?*

Oui/non/ne sait pas

- *Si oui, où ça/dans quel contexte ?*

Champ libre

- *Que signifie le terme "robot-journalisme" pour vous ?*

Champ libre

- *Avez-vous, à votre connaissance, déjà lu un article écrit par un robot ?*

Oui/non/ne sait pas

- *Si oui, lequel/dans quel média ?*

Champ libre

- *Quelle est votre opinion sur les articles écrits par des robots ?*

Champ libre

- *Quels seraient selon vous les avantages pour un média à utiliser des robots pour rédiger les articles ? (3 minimum)*

Champ libre

- *Et les inconvénients ? (3 minimum)*

Champ libre

- *Seriez-vous plus enclins à lire des articles écrits par un robot ?*

Oui/non/pas d'avis

- *Pourquoi ?*

Champ libre

- *Si oui, sur quels sujets ?*

Champ libre

- *Pensez-vous que les robots vont remplacer les journalistes ?*

Champ libre

- *Pourquoi ?*

Champ libre

Profil

- *Vous êtes :*

Un homme/une femme/autre

- *Votre âge :*

Champ libre :

- *Votre département (en chiffres) :*

Champ libre

- *Votre activité professionnelle :*

Indépendant et profession libérale/cadre et profession intellectuelle
supérieure/employé/ouvrier/étudiant/retraité/en recherche d'emploi

- *À quelle fréquence consommez-vous des médias de presse écrite (papier et web) ?*

Au moins une fois par jour/par semaine/par mois/par an/jamais

- *Lesquels ?*

Champ libre