



HAL
open science

Intelligence artificielle et mesures d'audience du journalisme en ligne : pour quelle intelligence des publics ?

Muriel Béasse

► **To cite this version:**

Muriel Béasse. Intelligence artificielle et mesures d'audience du journalisme en ligne : pour quelle intelligence des publics ?. Les cahiers du journalisme - Recherches, 2021, 2 (7), pp.R25-R38. 10.31188/Ca-Jsm.2(7).2021.R025 . hal-03709484

HAL Id: hal-03709484

<https://hal.science/hal-03709484>

Submitted on 29 Jun 2022

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.

Intelligence artificielle et mesures d'audience du journalisme en ligne : pour quelle intelligence des publics ?

Muriel Béasse Laboratoire Interuniversitaire des Sciences de l'Éducation et de la Communication (LISEC) - Université de Strasbourg

Résumé

De plus en plus de salles de rédaction numériques s'appuient sur des mesures d'audience développées par l'intelligence artificielle (IA). Ces évaluations sophistiquées, obtenues en temps réel, ne satisfont pourtant pas nécessairement les acteurs de la production journalistique. Cette étude s'appuie sur des entretiens auprès de producteurs et responsables éditoriaux de médias d'information français. Elle souligne l'influence des mesures d'audience sur les stratégies de mise en forme de l'information en ligne. Les données d'audience générées par l'IA tendent à encourager la production d'objets médiatiques favorisant une consommation passive de l'information au détriment de dispositifs stimulant l'intelligence des publics, entendue au sens étymologique de « discernement », « faculté de percevoir » et de « comprendre ».

Abstract

More and more digital newsrooms are relying on audience measurements developed by artificial intelligence (AI). These sophisticated evaluations, obtained in real time, do not necessarily satisfy the actors of journalistic production. This study is based on interviews with producers and editorial managers of French news media. It underlines the influence of audience measurements on the strategies of online news formatting. The audience data generated by AI tend to encourage the production of media objects that favour a passive consumption of information to the detriment of devices that stimulate the intelligence of audiences, understood in the etymological sense of "discernment", "ability to perceive" and "understand".

Introduction

L'intelligence artificielle (IA), comprise comme métaphore de l'ensemble des théories et techniques offrant la possibilité aux machines de simuler l'intelligence humaine (Castets-Renard, 2018) imprègne de plus en plus d'activités propres à la pratique du journalisme, que ce soit dans l'écriture, l'enquête ou encore la vérification d'information. L'IA s'impose également dans la relation qui unit un site d'information en ligne à ses destinataires et dans l'interprétation de ses modes de consultations et pratiques de lecture. Elle est un outil d'identification et de prédiction des activités et engagements de l'audience d'un site d'information au rythme de ses publications. Avec ses capacités d'analyse et de traitement automatique d'une vaste quantité d'indicateurs, aussi bien sur le trafic que sur le comportement des internautes, l'IA perfectionne une nouvelle génération de « métriques du web »¹, ces données synthétisant le volume et l'activité du public d'un site (Parasie et Dagiral, 2013, p. 212). Ces mesures se sont généralisées et diversifiées en quelques années dans les médias en ligne. Elles prennent désormais le pas sur des données d'audience qualifiées, exigeant beaucoup plus de temps², et tendent à influencer le contenu et la forme de l'information proposée. Elles sont pourtant loin d'établir une représentation objective des usages d'un site d'information puisque ces données sont souvent générées pour des intérêts spécifiques. Il peut s'agir, notamment, d'accroître les revenus des entreprises de presse qui misent sur les médias sociaux et, plus généralement, de comprendre les attentes et préférences du public dans un espace médiatique extrêmement compétitif.

Les mesures d'audience font l'objet de nombreuses recherches pour ce qui concerne la diversité des technologies déployées et le contrecoup des données obtenues sur la production des sites examinés (Christin, 2020 ; Barbeau, 2017 ; Beauvisage, 2013). Des études ont montré que ces évaluations engendrent des transformations non négligeables dans l'organisation du travail des journalistes et sur le contenu éditorial de la presse en ligne (Eustache, 2020 ; Ouakrat, 2016) avec des impacts plus ou moins grands selon les cultures de travail (Grossi, 2020). Notre approche consiste à interroger l'incidence de ces données sur la compréhension des modes de réception de l'information, c'est-à-dire sur les façons dont une production peut être accueillie³. Les mesures d'audience développées par l'IA, aussi perfectionnées soient-elles, offrent-elles suffisamment d'informations sur les différentes formes de mobilisation et d'activités des publics ? Sont-elles satisfaisantes pour les responsables éditoriaux des médias d'information ?

La question de la représentation des publics et de leurs usages se dote en effet d'enjeux multiples qui affectent les différents acteurs du processus de production-réception de l'information. Notre réflexion s'inscrit plus spécifiquement au cœur d'un triangle d'analyse dynamique qui prend en compte les interactions entre producteurs, messages d'information et publics. Elle questionne l'impact des profilages de l'audience obtenue à l'aide d'intelligence artificielle sur l'usage projeté de l'utilisateur numérique (Merzeau, 2010). Notre postulat est double : les données obtenues tendent, d'une part, à renforcer la perception d'utilisateurs d'abord

1 Le terme « métrique » est couramment utilisé dans le marketing numérique et la production web. Tout en lui laissant une place dans cet article, nous lui préférons l'expression générique de « mesure d'audience » pour désigner les données et analyses ayant trait aux différentes performances d'un site.

2 Comme le définit parfaitement Flores (2016) les mesures qualitatives cherchent « à mieux cerner le profil de l'individu exposé, ses attentes, la qualité de l'expérience interactive que l'exposition au message lui procure, mais aussi les perceptions engendrées par le message » (Flores, 2016, p. 53).

3 Cette définition s'inspire d'une formule de Soulez (2006) pour qui la réception est « une façon dont une pensée est accueillie par une autre » (p. 53).

consommateurs d'informations plutôt que lecteurs actifs⁴. L'intelligence des publics, entendue au sens étymologique de « discernement », de « faculté de percevoir » et de « comprendre », échappe aux logiques de l'intelligence artificielle et affecte, d'autre part, les formats et formes discursives de l'information. Qui plus est, la diversité des méthodologies utilisées pour ces mesures laisse encore de nombreux usages difficilement identifiables (Beauvisage, 2013). L'interprétation non distancée de ces données accentue donc les angles morts dans la compréhension des modèles complexes de la réception.

Le présent texte s'attache d'abord à préciser les différents statuts, rôles et activités prêtés aux publics d'information dans un contexte médiatique concurrentiel où le journalisme tente de survivre à la fois économiquement et symboliquement, en revendiquant son mandat démocratique et citoyen. L'analyse prend appui sur le point de vue de responsables éditoriaux de grands médias français interrogés à l'occasion d'une recherche plus générale qui porte sur les stratégies et choix de productions de longs reportages journalistiques. Les représentations des décideurs sur un processus de mesures d'audience qui leur échappe en partie. Quand bien même ces professionnels n'ont pas été questionnés frontalement sur les métriques perfectionnées par l'intelligence artificielle, l'analyse de leurs propos montre que ces mesures prennent une place prépondérante dans l'univers de la production avec une incidence sur les mises en forme de l'information.

Quel lecteur d'information ?

Loin de se référer à un groupe homogène, ce qui nous amène à parler « des publics » plutôt que « du public », le lecteur d'information est envisagé comme instance symbolique de référence et figure essentielle du journalisme. Les projections pour se représenter cet acteur clef dans les médias d'information ont animé une bonne part de l'histoire de la pratique (Guéneau, 2006 ; Dayan, 1992 ; Verón, 1983, 1984). La prise en considération du lecteur type, de ses usages et de ses intérêts de lecture est un motif ancien et intrinsèque à l'écriture journalistique qui accorde, par tradition, une place importante au lecteur⁵. Les journalistes ont intégré dans leur pratique professionnelle la nécessité d'« écrire pour être lu » (Brisson, Maltais et al., 2016). Ce réflexe est tout autant nourri par la nécessité économique de l'entreprise de presse qui a besoin de lecteurs pour exister que par le mandat social et citoyen qui porte et légitime la pratique journalistique. Les journalistes se préfigurent des publics modèles avec des représentations souvent floues et plus ou moins fantasmées tandis que les entreprises médiatiques projettent des cibles de lecture et se consacrent à entretenir une communauté de lecteurs en cherchant à faire correspondre les intérêts de ce groupe avec l'information proposée (Legrave, 2006). Ce « public imaginé » (Gans, 1979), qui permet d'appréhender symboliquement les destinataires du discours journalistique, s'appuie sur une variété de sources d'interprétation pour attribuer plus ou moins de rationalité aux lecteurs d'information (Demers, 2012). Ces échafaudages abstraits façonnent considérablement la communication journalistique (Coddington, Lewis et al., 2021).

4 Dans le cadre de cet article, nous privilégions le terme « lecteur » pour signifier, au sens large, l'instance de réception de médias tout autant écrits qu'audiovisuels.

5 Le terme « lecteur » marque également l'approche retenue pour cette étude qui est de considérer la lecture de l'information comme une activité minimale d'interprétation. Étymologiquement, en effet, le terme ne présuppose pas d'une quelconque passivité et renvoie à l'idée de prendre connaissance ainsi qu'à l'action de discerner, déchiffrer et comprendre (Centre national des ressources textuelles et lexicales).

Avec le développement du numérique et du journalisme en ligne, les questions qui portent plus spécifiquement sur la place et le rôle du lecteur d'information au cœur du système médiatique se sont intensifiées. Le lecteur, plus largement envisagé comme utilisateur, voire acteur du web (Weissberg, 1999), devient une figure centrale d'un réseau où les frontières entre activités de production et de réception s'entremêlent (Vitali-Rosati, 2012). Le web accentue la dimension participative de la lecture en s'appuyant sur la curiosité des individus comme moteur de la connaissance par l'intermédiaire de structures hypertextuelles. Cette architecture de liens renforce de nouvelles façons d'accéder au savoir « *de façon à la fois libre et guidé* » (Lambert, 1991, p. 62). Dans l'environnement numérique, l'utilisateur du web est donc généralement appréhendé comme un individu actif (Donnat, 2007 ; Merzeau, 2010), mais cette perception demeure souvent équivoque. L'activité de lecture, qu'elle soit sur écran ou non, est un phénomène complexe qui oscille dans une série d'indéterminations (Eco, 1962), entre reconnaissance et ignorance, participation, distanciation (Dufays, 1994). L'IA apparaît alors comme une opportunité de démythifier le phénomène à travers un ensemble de chiffres matérialisant les trajectoires, mouvements et formes d'engagements de l'utilisateur du web. Les données tirées des logiciels de l'intelligence artificielle participent à la schématisation d'une activité de lecture qui reste toutefois complexe, voire mystérieuse. Pour exemple, l'IA affine les modèles d'activité de lecture en examinant le trafic d'un site sur de multiples traces et dimensions d'utilisations, mais cette technologie parvient plus difficilement à différencier interactivité (les propriétés dynamiques d'un site web) et interaction (relation énonciateur et coénonciateur, au sens de « *quelqu'un qui raconte et quelqu'un qui écoute* » (Bouchardon, 2008, p. 81)). Les données chiffrées par l'IA marquent ainsi le produit d'un processus, mais ignorent ce processus lui-même.

Précisons que l'engouement suscité par ces indicateurs quantitatifs est aussi quelque peu modéré par des scandales mis à jour en 2018 par le *New York Times* (Keller, 2018) et le *New York Magazine* (Read, 2018). Des enquêtes menées par ces deux médias révélaient des données truquées sur les statistiques de fréquentation de grandes plateformes d'Internet, accentuant le doute sur la fiabilité de ces systèmes chiffrés. Les données d'audience du web, loin d'être neutres, continuent toutefois de prendre une place prépondérante dans les médias numériques où elles font écho à un web extrêmement marchand et publicitaire. Face à la montée en puissance des infomédias, la percée du mobile ou encore la concurrence et les modèles de performance des plateformes de vidéo en ligne, la production journalistique s'inscrit dans des logiques de survie à l'intérieur d'un écosystème compétitif aux audiences fragmentées. Depuis plusieurs dizaines d'années le développement du numérique a suscité de multiples bouleversements dans le journalisme aussi bien comme pratique professionnelle que comme activité discursive (Antheaume 2013 ; Bouquillon et Matthews 2010 ; Charon 2011 ; Rebillard, 2006). Les conditions économiques et techniques de production et réception des messages offrent un nouveau régime de consommation de l'information où la captation de l'attention du public est un enjeu crucial. Le volume de lecteurs, de clics réalisés, comme la durée de visite d'un site, composent désormais la véritable valeur marchande de l'information. Les entreprises médiatiques sont incitées à une course à l'innovation qui les amène à concevoir des contenus toujours plus attractifs afin de séduire des publics soumis à une abondance de messages (Brin, Charron et de Bonville, 2004). Dans cette perspective, les données fournies par l'IA sont souvent appréhendées comme des outils d'amélioration et de perfectionnement d'un site. Appliquées au marketing numérique, elles influencent la production web selon l'interprétation qu'en font les décideurs, en particulier au sein des rédactions d'informations qui misent sur les médias sociaux pour accroître leur audience.

Au-delà des problématiques liées à la pertinence des critères de recueil et d'analyse de ces données, aux enjeux et conséquences qu'ils revêtent sur la pratique journalistique, notre démarche consiste à questionner les représentations favorisées par ces chiffres générés par l'IA. Notre intérêt porte sur le point de vue des responsables éditoriaux et producteurs de médias d'information afin de comprendre comment ces données affectent leurs choix de production de l'information. Il s'agit, en somme, d'interroger deux phénomènes : 1. l'incidence de ces mesures d'audience chiffrées sur les représentations que se font les responsables de presse de leur lectorat, 2. La répercussion de ces représentations sur les stratégies de production de l'information⁶.

Dans une approche sociologique du discours médiatique, toute production des médias peut être envisagée comme « *l'élément symbolique d'un processus associant des acteurs sociaux disparates* » (Esquenazi, 2014, p. 9). Les productions journalistiques relèvent ainsi des négociations en jeu entre les différents acteurs de ce processus. L'intelligence artificielle appliquée aux mesures d'audience peut alors s'appréhender comme une nouvelle logique professionnelle venant s'insérer dans la relation producteur-lecteur, puisqu'elle participe à la représentation d'une figure emblématique du lecteur d'information.

Méthodologie et recueil de données

L'analyse effectuée s'appuie sur des entretiens semi-dirigés menés auprès de responsables éditoriaux de médias français d'information. Seize responsables éditoriaux de médias en ligne et quatre producteurs de sociétés de productions partenaires de ces médias ont été interviewés, individuellement, sur leur lieu d'exercice ou par téléphone entre novembre 2019 et février 2020⁷.

L'enquête initiale, inscrite dans une recherche exploratoire plus vaste, visait à comprendre les conditions de production de longs reportages journalistiques explorant les nouveaux formats du numérique (*newsgame*, webdocumentaire, datajournalisme, réalité virtuelle, etc.). Ces objets d'étude étaient appréhendés en tant que dispositifs médiatiques, c'est-à-dire comme des structures de production à lire et à interpréter intégrant une « modélisation du processus de communication » (Souchier, Candel et al., 2019, p. 126). Les entretiens menés pour cette recherche se sont articulés autour d'un canevas de questions visant à connaître les méthodes de travail des producteurs et responsables éditoriaux ainsi qu'à percer les motivations et contraintes de productions journalistiques numériques⁸. Les discours des répondants ont permis, au final, de révéler, en creux, les représentations que les professionnels du milieu se font du web, de ses usages, et par extension de ses usagers. Au cours de ces échanges, il n'a pas

⁶ Par le terme « stratégies », nous désignons les choix réalisés par un sujet (individuel ou collectif), conscient ou non, sur un certain nombre d'opérations langagières (Charaudeau et Maingueneau, 2002, p. 549).

⁷ Les responsables interrogés portent les titres différents de directeur de l'information, responsable de programme numérique, rédacteurs en chef, producteur, etc. Nous les regroupons sous les termes génériques de « producteurs » ou « responsables éditoriaux » afin de préserver leur anonymat. Leurs responsabilités s'exercent au sein des rédactions numériques du *Monde*, *Arte*, *Francetv*, *France 24*, *RFI*, *Médiapart*, *L'Équipe*, *TV5*, *Les Jours*, *Disclose* ainsi que des sociétés de production partenaires : *Upian*, *Hans Lucas*, *Pixel Hunt*, *We do data*.

⁸ Les entretiens se sont déroulés très librement autour des quatre grandes questions générales suivantes : depuis quand et pour quels motifs votre rédaction produit ou diffuse des longs reportages journalistiques ? Qui sont les principaux acteurs, collaborateurs qui participent à la réalisation ces productions ? Quelles sont les principales étapes et procédures de travail pour la réalisation de ces productions ? Quel est le public ciblé par ces modes d'information ?

été posé de questions directes sur les mesures d'audience, ou sur l'impact des algorithmes et de l'utilisation de l'intelligence artificielle dans les rédactions. Quinze répondants ont fait mention, d'eux-mêmes, de ces différentes réalités, ce qui nous a poussés à retraiter notre analyse de contenu sous l'angle de cette thématique. Une analyse thématique des discours des répondants nous a amenés à nous concentrer sur le contenu manifeste et latent des propos enregistrés en faisant émerger différentes catégories d'éléments sémantiques fondamentaux (Bardin, 1977). L'analyse thématique de contenu est mue par une double visée qui consiste à relever, d'une part, la signification d'un énoncé, sans ignorer sa dimension subjective, et à dégager, d'autre part, des éléments pertinents et objectifs pour le groupe social analysé (Negura, 2006). Il ne s'agit donc pas de s'en tenir aux seules opinions exprimées, mais de rapprocher les représentations repérées au contexte socio-économique d'un milieu, aux relations symboliques entre les différents acteurs sociaux de ce milieu professionnel (Dany, 2016).

En reprenant les entretiens sous l'angle des contenus ayant trait aux mesures d'audience des sites web des médias et aux représentations faites des publics ciblés par les données générées par ces mesures, trois niveaux de représentations discursives ont pu être mis en évidence : 1. Les représentations que se font les répondants sur les données d'audience; 2. Les représentations des répondants sur la figure projetée du lecteur d'information à partir de ces données; 3. Leurs représentations de l'impact de ce système d'évaluation sur les produits d'information.

Une compréhension contrastée

L'analyse effectuée montre d'abord que l'ensemble des répondants, producteurs et responsables éditoriaux, ont intégré les logiques socioéconomiques des médias numériques que plusieurs d'entre eux nomment « *logique de flux* ». Comme le rapporte un producteur : « *on essaie de capter l'attention des gens et surtout leur engagement, qu'ils passent un maximum de temps pour qu'ils explorent un maximum de nos objets (de production)* » (Responsable Francetv, le 25.11.2019). Les répondants font état de leurs stratégies de production qui consistent, pour l'essentiel, à être au plus près des usages et des modes de consommation passant désormais par le mobile et les réseaux sociaux. Dans cette perspective, les mesures d'audience du web, perfectionnées par les technologies issues de l'intelligence artificielle, sont appréhendées comme une assistance précieuse visant à cerner les pratiques pour les articuler à la ligne éditoriale de l'entreprise : « *ce qu'on essaie de faire c'est de vraiment comprendre quels sont les usages et de toujours raconter des histoires qui nous tiennent à cœur* » (Responsable Arte France, 28.11.2019)

Pour chercher à comprendre ces usages, des répondants expriment leur fascination vis-à-vis de la comptabilité des traces recueillies et du maillage effectué autour du lecteur. La technologie est tout entière au service du décryptage de la présence, des déplacements et de toutes les actions des usagers des médias d'information : « *Aujourd'hui, pour un objet web, quel qu'il soit, il y a une myriade de capteurs qui nous permettent de mesurer ça !* » (Productrice We do data, 12.02.2020)

Toutefois, de façon générale, l'analyse des discours tenus par ces acteurs de la production française d'information met en évidence l'insatisfaction de nombres de producteurs interrogés face à l'opacité des processus d'évaluation issus de l'intelligence artificielle. À la question de savoir, par exemple, comment un responsable se projette dans la figure de ceux qui vont découvrir les reportages proposés, la réponse met en exergue un sentiment de dépossession :

C'est super compliqué parce qu'on est incapable de savoir comment les gens nous lisent. On a des données sur le temps passé et, en fait, le temps moyen que les gens passent sur un reportage c'est très, très, court, c'est quatre minutes ou quelque chose comme ça, mais un reportage si on veut le lire du début jusqu'à la fin il faut passer au moins une demi-heure grosso modo. Donc, ces chiffres, on ne sait pas trop ce que ça veut dire même si les spécialistes de ces questions sont contents. (Responsable *L'Equipe.fr*, 29.11.2019)

Un autre responsable fait un constat similaire en évoquant des reportages multimédias mis en ligne :

On ne sait pas comment les gens les consomment. Il y en a sans doute qui les lisent jusqu'au bout et probablement pas tous. Pour le moment on n'a pas de moyen d'affiner le comportement de ceux qui nous regardent. On sait juste qu'ils sont assez nombreux et c'est juste ce qui suffit aux gens du marketing. (Responsable *L'Equipe.fr*, 29.11.2019)

Même quand ces retours quantifiés semblent positifs, la prudence est de mise et ils peuvent être très facilement relativisés comme l'énonce une responsable : « *on le voit bien les vidéos sont consultées des milliers de fois, mais ça ne veut pas dire que les gens sont allés jusqu'au bout de la vidéo !* » (Responsable TV5, 27.02.2020). Le flou enveloppant la figure du lecteur d'information tend alors à se confondre avec l'opacité de fonctionnement des classements de contenus d'infomédiaires comme *Facebook* ou *YouTube* que des responsables « *cherchent à comprendre* ».

Ouakrat (2016) a déjà montré comment les métriques d'audience participent au cadre de la production éditoriale des rédactions du web en renforçant notamment les liens entre les rédacteurs en chef et les responsables du marketing ou encore en stimulant la compétition interne au sein de l'équipe éditoriale. Face à ces mesures d'audience, le chercheur note que l'attitude des journalistes oscille entre cynisme, défiance, ambivalence (Ouakrat, 2016, p. 111-112) ce qui fait écho à ce que révèle aussi la présente analyse. Au-delà des ressentis, les propos des répondants permettent alors de distinguer par contrastes ce que ces responsables perçoivent du lecteur imaginé, mis en visibilité par ces métriques.

Persistence d'un lecteur insaisissable

L'analyse met en évidence une forme de paradoxe : en même temps que les producteurs ont accès à de plus en plus de données sur les usages des destinataires de leurs produits d'information ils ne trouvent pas d'éléments sur leurs publics qui les satisfassent. Des répondants font valoir la nécessité de mesures et recherches complémentaires. Comme l'exprime une productrice :

Aujourd'hui on peut mesurer le temps passé sur chaque bouton, sur chaque page, on peut prévoir quel trajet il (le lecteur d'information) a fait, d'où il vient [...]. On devrait passer plus de temps à analyser pourquoi il reste et faire par exemple des interviews qualitatives pour comprendre ça. On reste suspendu à du métrique, de la donnée qui n'est pas qualifiée. (Productrice *We do data*, 12.02.2020)

Cette même productrice se réfère à « *l'époque où les analyses n'étaient pas aussi poussées* » et où, sachant peu de choses sur le lecteur, les producteurs n'avaient pas d'autres choix d'être confiant sur ses engagements de lecture :

Je pense qu'on prend un peu plus en compte, peut-être trop, je ne sais pas, cette notion de l'attention de l'internaute. Je trouve qu'avant on lui faisait confiance en se disant puisqu'on lui offre quelque chose de qualité il va rester ; et maintenant avant même qu'on produise quelque chose on est complètement obsédé du fait qu'il n'a qu'une micro seconde à nous accorder ! Peut-être que ça a un impact un peu négatif sur nos productions où on essaie de faire court, court, court... On lui fait moins confiance qu'avant. (Productrice *We do data*, 12.02.2020)

La confiance qui étymologiquement se réfère à la croyance d'une valeur partagée s'oppose ainsi à l'ambiguïté de l'interprétation des chiffres et à l'autorité scientifique qu'ils véhiculent. Le terme de confiance, loin d'être anodin, révèle en particulier l'importance portée à la construction de la relation producteur-lecteur et à la charge symbolique que cette relation revêt dans le processus de production d'information. Les mesures d'audience obtenues à l'aide d'intelligence artificielle font figure d'instances tierces qui s'immiscent entre les acteurs fondamentaux du journalisme.

L'obsession de l'attention du lecteur côtoie alors de très près le préjugé de son éventuelle passivité ou du moins de son peu de disponibilité, de sa volatilité dans un univers médiatique rempli de sollicitations :

C'est vrai qu'on reste un peu obsédé par une masse de gens qui effectivement vont vite ; après il y a des gens qui restent aussi, donc on devrait passer plus de temps à analyser pourquoi ils restent. (Productrice *We do data*, 12.02.2020)

Dans les discours des responsables, une distinction s'opère souvent entre les notions d'attention captée (la présence du lecteur) et engagement réel (l'implication cognitive du lecteur) qui se cristallise dans la question des formes courtes ou longues de l'information : le court attire et le long engage. Ce à quoi des responsables s'insurgent :

On proposait jusque-là sur les réseaux sociaux et sur le site des formats courts parce qu'on se disait que l'internaute est incapable de se concentrer plus de trois minutes. C'est moyennement vrai parce que pour les programmes qui sont vraiment de qualité les gens regardent tout ! (Responsable *RFI*, 6.12.2019)

Face à la difficile fidélisation des publics, dans un réseau caractérisé par le flux et la circulation rapide de contenus, les responsables éditoriaux et producteurs interrogés font valoir la nécessité d'offrir des productions de qualité. Cette qualité s'établit aussi bien dans le contenu que dans la forme : longueurs de texte, visuel privilégié, recours aux hypermédias, etc. Autant d'éléments associés au design d'information qui constitue un enjeu supplémentaire dans l'interprétation des métriques.

Uniformisation de la production ?

Les mesures transmises aux décideurs leur permettent de dessiner, bon an, mal an, un profil de consommateur pressé et à l'attention éphémère, saturé de multiples incitations et messages d'informations. Ce type d'interprétation a des conséquences directes sur la dynamique production-réception puisque, selon les répondants, ces données et leurs analyses finissent par affecter les formes de l'information : « *On va vers des formes plus simples, plus accessibles [...] en gros, on ne va pas trop en demander aux gens* ». (Responsable *LeMonde.fr*, 13.01.2020)

Concrètement, les vidéos courtes, perçues comme des formats moins exigeants en termes d'implication cognitive sont mises en avant et préférées aux messages à dominance verbale privilégiant la dimension rationnelle et réfléchie de l'information (Meunier et Peraya, 2010, p.362). Ces stratégies d'écritures ne sont pas nouvelles dans la pratique journalistique, mais se retrouvent cautionnées ou même renforcées par les bons scores d'audiences réalisés par ces formats. Une responsable fait ainsi valoir l'intérêt de la création de vidéos courtes d'information sur une chaîne de vidéos en ligne :

On s'est aperçu qu'entre une vidéo *YouTube* qui peut faire d'un coup sur une journée 50 000 vues et un long format où on est content quand on arrive à 10 000 vues sur une semaine, quelque part, il n'y a pas photo ! En plus on ne sait pas si les gens lisent tout. (Responsable France24.fr, 28.02.2019)

La plupart des responsables interrogés sur leurs décisions éditoriales affirment faire des choix pragmatiques, selon leurs propres mots, en pointant aussi de multiples contraintes de production qui ne se limitent pas aux seules métriques. En effet, ces comptabilités d'audience sont loin d'être les seules sources d'influence sur la production-consommation d'information. Comme le synthétise un des répondants, les stratégies de production sont certes dictées par « *l'appétit de l'audience* », mais aussi par « *les coûts de production, les investissements, aussi bien que les compétences de l'équipe journalistique* » (Responsable *Médiapart*, 18.11.2019).

Les choix de production laissent, quoi qu'il en soit, des impressions mitigées aux répondants pour ce qui concerne l'inventivité des dispositifs d'information proposés au regard des différentes contraintes de production. :

J'ai l'impression qu'il y a plutôt une uniformisation des écritures journalistiques au niveau mondial, mais ça n'engage que moi [...] il y a des spécificités sur le ton sur la manière de traiter l'information, peut-être, mais sur les formats j'ai un doute. (Responsable *France Média Monde*, 10.12.2019)

Ou encore :

On a l'impression qu'il n'y a pas autant de créativité qu'il y a quatre ou cinq ans où il y avait des formes qui portaient dans tous les sens avec des choses ultras longues ou ultra-courtes... (Responsable *LeMonde.fr*, 13.01.2020)

Ces remarques laissent alors penser que le profilage du lecteur d'information participe à la généralisation de dispositifs identiques et monotones favorisant une consommation passive de l'information. Ce constat sévère est à nuancer en tenant compte du contexte de fragilité financière que connaissent les grands médias d'information⁹ où travaille la majorité des professionnels que nous avons interrogés. Le système contraignant de l'interprétation des mesures d'audience du web se superpose dans les faits aux contraintes initiées par la crise des médias.

Selon un des producteurs interrogés, le phénomène d'uniformisation des contenus numériques d'information serait attribuable à une méconnaissance de la culture d'écriture du web qui perdure chez les responsables éditoriaux du numérique venant plutôt de la télévision ou de la presse écrite :

Il faut évidemment adapter sa proposition au support et donc aux usages. Les entreprises qui viennent de l'audiovisuel n'ont pas l'habitude de faire du design interactif parce que ce n'est pas comme ça que ça marche dans le linéaire dans l'audiovisuel, donc ce n'est pas forcément une habitude qui est bien répandue. (Producteur *Pixel Hunt*, 10.12.2019)

L'adaptation de l'écrit journalistique à l'écran (Catroux, 2000) n'a pas encore acquis sa pleine maturité. Sur ce point, les mesures d'audience du web participent à une prise de conscience ou pourraient jouer un rôle de repère dans le développement des compétences d'écriture numérique

⁹ Une bonne part des responsables éditoriaux interrogés appartient au secteur de l'audiovisuel public français (*France tv, France Media Monde, Arte.fr*), secteur confronté à une crise structurelle chronique et soumis à de nombreux mouvements de réformes (Bellon, 2016).

des journalistes. Toutefois, la mise en exergue des compétences d'écriture des producteurs d'information interroge, par jeu de miroir, les compétences de lecture prêtées à ceux pour qui sont destinées ces productions. Les témoignages recueillis mettent en lumière un nivellement dans la production des dispositifs d'information : des formats inventifs sont bridés par crainte de perdre l'attention du lecteur ou d'exiger de lui trop d'effort, mais cet effort de lecture n'est-il pas minimalement nécessaire pour parvenir à l'engagement souhaité par les producteurs ?

La question qui mérite de plus longues investigations fait écho aux arguments de Merzeau qui souligne « *l'intelligence de l'utilisateur numérique* » et interroge en 2010 l'autonomie et les marges de manœuvre qui sont réellement accordées à cet usager dans les dispositifs numériques :

Surfer, chercher, évaluer, relayer, coopérer, s'exposer, organiser, créer... L'environnement numérique autorise bien un large éventail d'opérations, mais à la condition que le degré de maîtrise et d'implication de l'utilisateur ne soit pas rabattu par la recherche obsessionnelle d'une prétendue facilité, qui équivaut en fait à déléguer l'intelligence au prestataire. (Merzeau, 2010, p. 14).

Conclusion

En cherchant à distinguer quelle compréhension du lecteur est favorisée par les données d'audience mesurées et analysées par l'IA, cette contribution s'est focalisée sur les témoignages de responsables éditoriaux et producteurs français confrontés à ces données multiples dans les choix de production de leurs rédactions. L'analyse des propos recueillis met en évidence l'impact de ces données d'audience sur des représentations du lecteur déjà bien ancrées dans ce milieu professionnel ainsi que leur influence sur la création des dispositifs d'information. Il s'agit, selon les propres mots des répondants de ne pas « *faire trop compliqué* », ne « *pas trop demander* » au lecteur dans un écosystème médiatique extrêmement compétitif surchargé de sollicitation de messages d'informations. Traditionnellement présenté comme actif dans les environnements numériques, le lecteur d'information figuré par les mesures d'audience du web est appréhendé, plus spécifiquement, comme un consommateur pressé et volatile. Par conséquent, le système médiatique privilégie pour lui un usage passif des dispositifs d'information, ce qui peut constituer une boucle inextricable.

Les professionnels interrogés font une distinction entre l'attention captée du lecteur et les conditions de son véritable engagement au cœur des dispositifs d'information tout en faisant valoir que les statistiques leur fournissent finalement peu d'éléments sur les implications et les différentes formes d'engagement des publics. Quoique ce ne soit pas son principal objectif, les mesures d'audience enrichies par l'IA peuvent être abordées comme des outils pratiques afin de rendre moins abstrait un lecteur d'information source de nombreux fantasmes et projections. L'interprétation des données générées par l'IA offre l'opportunité d'améliorer la production de modèles de consommations efficaces et attractifs (regarder plutôt que lire), mais avec le risque de limiter la création de formes journalistiques stimulantes pour l'intelligence du lecteur.

À l'issue de leur enquête sur les métriques du web, Parasie et Dagiral (2013) préconisaient d'étudier les publics en couplant les données statistiques d'audience avec des résultats d'enquêtes sociologiques d'usages menées auprès des internautes. Cette proposition conserve son intérêt pour le monde des médias et du journalisme confronté à des outils de mesure et d'analyse d'audience innovants. L'intelligence artificielle réactive les débats liés aux problématiques de la réception de l'information et invite à questionner plus précisément à quel

point l'irruption de cette technologie modèle la production-réception en favorisant une certaine idée de l'information et une figure spécifique du lecteur.

Muriel Béasse est enseignante à l'Université de Strasbourg, chercheure associée au LISEC.

Références

Antheaume, Alice (2013). *Le journalisme numérique*. Presses de Sciences Po.

Barbeau, Jean-Sébastien (2017). Mesures d'audience et rémunération bonifiée. *Communication*, 34/2.

Beauvisage, Thomas (2013). Chapitre 9 - Compter, mesurer et observer les usages du web : outils et méthodes. Dans C. Barats, *Manuel d'analyse du web en Sciences Humaines et Sociales*, 187-211. Armand Colin.

Bellon, Anne (2016). La politique de concurrence au secours de l'exception culturelle : L'arrivée de Netflix en France depuis le ministère de la Culture. *Gouvernement et action publique*, vol. 5(4), 119-140.

Bouchardon, Serge (2008). Le récit littéraire interactif : une valeur heuristique. *Communication & Langages*, 155, 81-97.

Bouquillon, Philippe et Matthews, Jacob. (2010). *Le Web collaboratif dans le capitalisme : mutations des industries de la culture et de la communication*. Presses Universitaires de Grenoble.

Brin Colette, Charron Jean et de Bonville, Jean. (2004). *Nature et transformation du journalisme. Théorie et recherches empiriques*. Les Presses de l'Université Laval.

Brisson, Pierre, Maltais, Robert, Parent, André, Asselin, Yvan (2016). *L'écriture journalistique sous toutes ses formes*. Presses de l'Université de Montréal.

Castets-Renard, Céline (2018). Quels impacts de l'intelligence artificielle sur les métiers du droit et du journalisme ? *Cahiers de la propriété intellectuelle*, 30(3), 987-1001.

Catroux, Michèle (2000). L'écrit à l'écran : contraintes et libertés de l'écriture journalistique sur la page Web. *Revue Asp*, 27-30.

Charaudeau, Patrick et Maingueneau, Dominique (2002). *Dictionnaire d'Analyse du Discours*. Seuil.

Charon, Jean-Marie (2011). *La presse en ligne*. La Découverte

Christin, Angèle (2020). *Metrics at work. journalism and the contested meaning of algorithms*. Princeton University Press.

Coddington, Mark, Lewis, Seth C. et Belair-Gagnon, Valerie. (2021). The imagined audience for news: where does a journalist's perception of the audience come from?. *Journalism Studies*, 22(8), 1028-1046.

- Dany, Lionel (2016). Analyse qualitative du contenu des représentations sociales. Dans G. Lo Monaco, S. Delouée et P. Rateau (Eds.), *Les représentations sociales*, 85-102. De Boeck.
- Dayan, Daniel (1992). Les mystères de la réception, *Le Débat*, 71, 141-157.
- Demers, François (2012). Le défi de la construction d'un public pour le journalisme au temps de Google Analytics. *Tic&Société*, 6(1), 9-26.
- Donnat, Olivier (2007). Pratiques culturelles et usages d'internet. *Culture études*, 3, 1-12.
- Dufays, Jean-Louis (1994). *Stéréotype et lecture. Essai sur la réception littéraire*. Mardaga.
- Eco, Umberto (1962). *L'œuvre ouverte*. (C. Roux de Bézieux et A. Boucourechliev trad.). Seuil.
- Esquenazi, Jean-Pierre (2014). *L'écriture de l'actualité*. Presses universitaires de Grenoble.
- Eustache, Sophie (2020). *Bâtonner, comment l'argent détruit le journalisme*. Éditions Amsterdam
- Flores, Laurent (2016). Chapitre 3. Métriques et KPI disponibles : le quantitatif et le qualitatif. Dans L. Flores, *Mesurer l'efficacité du Marketing digital : Estimer le ROI pour optimiser ses actions*, 53-90. Dunod.
- Gans, Herbert J. (1979). *Deciding what's news: A study of CBS Evening News, NBC Nightly News, Newsweek, and Time*. Evanston, Northwestern University Press.
- Grossi, Valentina (2020). L'autonomie journalistique face aux chiffres. L'interprétation des métriques dans deux entreprises de presse. *Sociologies pratiques*, 1(1), 25-37.
- Guéneau, Catherine (2006). Du spectateur à l'interacteur ? *INA, Médiamorphoses*, 18. Dossier *Le spectateur introuvable*, 68-73.
- Keller, Michael H. (2018). The flourishing business of fake YouTube views. *The New York Times interactive*, 11.08.2018.
- Legavre, Jean-Baptiste (2006). Qu'est-ce qu'un « bon » lecteur de presse ? Les lecteurs du *Monde* à travers les billets du médiateur. *Communication et langages*, 150, 3-15.
- Merzeau, Louise (2010). L'intelligence de l'utilisateur. Dans L. Calderan, B. Hidoine et J. Millet (dir.), *L'Usager numérique. Actes du Séminaire INRIA 2010*. ADBS éditions, 9-37.
- Meunier, Jean-Pierre et Peraya, Daniel (2010). *Introduction aux théories de la communication*. De Boeck.
- Negura, Lilian (2006). L'analyse de contenu dans l'étude des représentations sociales. *Sociologies, Theory and research*.
- Ouakrat, Alan (2016). Les métriques d'audiences comme agents de transformation du journalisme en ligne ? », *Études digitales*, 2(2), *Le gouvernement des données*, 107-120.

Parasie, Sylvain et Dagiral, Éric (2013). Chapitre 10 - Le site et son audience. Enquêter sur les « métriques du web ». Dans C. Barats, *Manuel d'analyse du web en Sciences Humaines et Sociales*, 212-227. Armand Colin.

Read, Max (2018). How much of the Internet is fake? Turn out, a lot of it, actually. *New York Magazine*, 26.12.2018.

Rebillard, Franck (2006). Du traitement de l'information à son retraitement: La publication de l'information journalistique sur l'internet. *Réseaux*, 137(3), 29-68.

Salles, Chloë et Schmitt, Laurie (2017). Les webdocumentaires, un terrain d'expérimentation numérique. *Sur le journalisme, About journalism, sobre jornalismo*, 6(1), 158-171.

Souchier, Emmanuël, Candel, Etienne, Gomez-Mejia, Gustavo et Jeanne-Perrier, Valérie. (2019). *Le numérique comme écriture. Théories et méthodes d'analyse*. Armand Colin.

Soulez, Guillaume (2006). La réception : traduction, rencontre et passage. *INA, Médiamorphoses*, 18. *Dossier Le spectateur introuvable*, 52-57.

Verón, Eliseo (1985). L'analyse du contrat de lecture dans *Les médias : expériences et recherches actuelles*, IREP.

Verón, Eliseo (1984). Quand lire, c'est faire : l'énonciation dans le discours de la presse écrite Dans *Sémiotique II, Paris, Institut de recherches et d'études publicitaires*, 33-56.

Vitali-Rosati, Marcello (2012). Une éthique appliquée ? *Éthique publique*, 14(2), 13-32.

Weissberg, Jean-Louis. (1999). *Présences à distance*. L'Harmattan.

Référence de publication :

Béasse, M. (2021). Intelligence artificielle et mesures d'audience du journalisme en ligne : pour quelle intelligence des publics ? *Les Cahiers du journalisme - Recherches*, 2021, 2(7), R25-R38.