

L'information à tout prix, Julia Cagé, Nicolas Hervé et Marie-Luce Viaud

1. Auteurs

Julia Cagé est professeur d'économie à Sciences Po Paris et chercheuse affiliée au *Centre for Economic Policy Research*

Nicolas Hervé est ingénieur de recherche à l'INA. Ses travaux portent sur l'analyse automatique et l'indexation de contenus multimédias.

Marie-Luce Viaud est chercheuse à l'INA, spécialisée en fouille et visualisation des données. Elle coordonne avec Nicolas Hervé L'Observatoire TransMédia (OTMedia)

Ce livre est issu de deux articles scientifiques des mêmes auteurs : « The production of information in an online world: is copy right ? » (2016) et « Estimating the production and demande for online news: micro-level evidence from the universe of French new media. » (2017)

2. Introduction

Les auteurs partent du constat que les médias en ligne n'ont toujours pas trouvé le bon modèle économique.

Ils ont utilisé les outils du *big data*, du *machining learning* (apprentissage automatique) et de l'économie pour créer et analyser une base de données inédite : le contenu produit en ligne par les médias d'information en France sur l'année 2013, combinée avec les caractéristiques des différents médias et l'audience des différents sites.

Il s'agit donc d'une analyse quantitative fine de la totalité des textes produits par les médias français sur leur site internet : journaux, TV, radios et pure player (médias uniquement présents sur Internet). Un détecteur de copie a été créé pour distinguer ce qui relève de l'information originale et de sa diffusion.

Le modèle économique sur Internet est complexe : la gratuité attire une audience importante (entre 1,5 et 2 millions de lecteurs du Monde par jour), les articles originaux sont repris très rapidement par les autres médias (64 % de copié-collé pur parmi les contenus en ligne). Les internautes ont donc toujours le choix de lire un média gratuit plutôt que payant. Sous forme papier, l'information avait une valeur commerciale car sans possibilité de reprise immédiate par les journaux concurrents. Sur Internet, la question de la monétisation d'une information originale se pose encore. Pourquoi continuer à alimenter des rédactions importantes quand on peut se contenter de faire du copié-collé ? Aux États-Unis, on comptait 55000 journalistes pour la presse quotidienne en 2007 contre 35000 en 2015. La taille des rédactions diminue fortement (1 % par an en France depuis 2013). Il y a un risque d'appauvrissement de l'information, en quantité et en qualité.

Malgré ce constat, la réputation des médias joue un rôle important : quand pratiquement 50 % des internautes parviennent sur un site d'information via un

agrégateur de lien (réseaux sociaux...), l'autre moitié va directement sur son média favori. Il y a une corrélation entre l'originalité des contenus et l'audience.

Les auteurs reviennent sur l'importance de la citation : la source de l'information intéresse le lecteur. Pourtant, la mention de la source n'a lieu que dans 8% des cas. Les médias préfèrent citer l'AFP, plutôt qu'un média concurrent. Cela soulève la question de la propriété intellectuelle.

3. Chapitre 1 : l'information sur Internet : une approche transmedia

L'Observatoire TransMedia (OTMedia) est un programme de recherche subventionné par l'Agence nationale de la recherche et porte sur « l'étude et la traçabilité des événements médiatiques sur tous les supports de diffusion ». Elle est actuellement développée par l'INA. Son objectif est de fournir un outil de recherche pour analyser et comprendre l'information aujourd'hui.

La base de données constituée par les auteurs permet une analyse de 86 médias, soit 2,5 millions de documents (73 % issus des sites internet de la presse papier).

Cela représente un total de 25000 événements médiatiques (« un ou des faits qui sont considérés comme assez importants pour un public visé. ») Les domaines les plus représentés sont dans l'ordre : la politique, l'économie et la finance, ce qui a trait à la justice et le sport. Malgré Internet, les articles sont mis en ligne entre 6h et 18h, avec une baisse à la pause méridienne.

Cette étude met en avant l'importance du travail de veille des journalistes sur les autres médias concurrents : être réactif pour proposer de l'information à ses lecteurs, même si elle n'est pas originale. Les auteurs pointent un risque d'homogénéisation de l'actualité.

4. Chapitre 2 : la propagation de l'information sur Internet : illustrations

Les auteurs s'intéressent dans ce chapitre trois événements médiatiques pour analyser comment l'information se diffuse sur Internet :

- Les écoutes de la NSA (*Le Monde*, octobre 2013)
- Le micro-parti de Marine Le Pen (*Mediapart*, novembre 2013)
- La candidature de François Fillon à la présidentielle (AFP, février 2013).

La réputation des médias est mise en avant : l'article du *Monde* et la dépêche de l'AFP sont repris très rapidement par les autres journalistes, alors que l'article de *Mediapart* a nécessité des vérifications avant d'être repris. Au final, l'article du *Monde*, une information inédite, aura fait l'objet de 115 autres articles, quand *Mediapart* aura été repris 29 fois et l'AFP 35 fois.

La copie entre les journalistes est très importante, puisque plus de 70% du contenu des articles sont entièrement copiés pour être repris sur les autres médias.

5. Chapitre 3 : Les déterminants économiques de la production d'information

Ce chapitre interroge la corrélation entre la taille des rédactions des différents médias et la quantité d'information produite.

La taille moyenne d'une rédaction est de 122 journalistes en 2013, avec une variation de 1 à 800. Cela ne prend pas en compte l'expérience, les moyens, et l'environnement des journalistes, ni la pluralité des articles (investigation, information...).

En moyenne, un journaliste supplémentaire dans une rédaction permet de produire 28 articles originaux par an. Les plus grandes rédactions couvrent aussi plus d'événements (AFP : 93 % des événements couverts).

Les auteurs évoquent les robots producteurs d'articles très utilisés par les différents médias. Lors des élections municipales, cela a permis au *Monde* d'être exhaustif dans la transmission des résultats de chaque commune. De même, pour traiter les résultats sportifs, l'information économique et les bilans d'entreprise.

Une autre technologie transforme du texte en vidéo. La publicité est en effet plus rentable sur les formats vidéos que textuels, et est plus facilement diffusable sur les réseaux sociaux.

6. Chapitre 4 : Les agences de presse : des acteurs à part

Une agence d'information a pour mission « de rechercher des nouvelles et d'une façon générale des documents d'actualité ayant exclusivement pour objet l'expression ou la représentation des faits et de les distribuer à un ensemble d'entreprises d'information et exceptionnellement à des particuliers en vue de leur assurer, contre paiement d'une redevance et dans des conditions conformes aux lois et usages du commerce, un service d'information aussi complet et impartial que possible. » (UNESCO, 1953). La France compte en 2013 une centaine d'agences de presse, dont la principale l'Agence France Presse (3^e agence de presse du monde).

L'AFP compte 780 journalistes (sous contrat français) en 2013, Reuters 2500 et Chine nouvelle 6000. Elle est l'héritière de l'agence Havas créée en 1835, puis AFP en 1944. Son statut de 1957 précise qu'il s'agit d'un « organisme autonome doté de la personnalité civile et dont le fonctionnement est assuré suivant les règles commerciales ». L'AFP n'a pas de statut coopératif, contrairement à Associated Press aux Etats-Unis.

Les revenus de l'AFP proviennent à 35% des subventions de l'Etat, dans un objectif d'intérêt général, et des abonnements des autres médias. Quand un média reprend une information originale d'un média concurrent, il peut le faire par le biais de l'AFP de manière légitime puisque payant un abonnement.

7. Chapitre 5 : la monétisation de l'information

Les auteurs cherchent dans ce chapitre à trouver des solutions financières à la production d'information qui est un bien public. Ils distinguent quatre sources de revenus : les utilisateurs, les publicitaires, l'état et les actionnaires.

La première partie s'intéresse à l'audience des différents médias. Logiquement, ils constatent que plus il y a d'informations sur un média, plus l'audience est importante. L'information originale attire finalement relativement peu : pour 1% d'information

originale en plus, il y a une augmentation de seulement 0,018% des visiteurs. De plus, cette croissance est de courte durée : produire de l'information originale ne suffit pas à fidéliser les lecteurs, d'où l'importance de la réputation.

Seuls 27% de l'audience sont constitués par les sites internet des médias. En moyenne, 46 % de l'audience en ligne provient des moteurs de recherche.

Les revenus publicitaires ne représentaient en 2013 que 5% des revenus des sites en ligne. Ce marché est capté par les principaux acteurs du web (notamment Google et Facebook). Les bloqueurs de publicité ont aussi un impact sur ce modèle économique (une publicité bloquée ne génère pas de revenu pour le site internet), puisqu'en 2017, plus de 30% des internautes en ont installé un.

8. Chapitre 6 : L'importance des citations et de la réputation

La production d'information originale entretient la réputation d'un média, sous forme papier ou numérique. Par exemple, l'article du *Monde* sur les Panama Papers a triplé le nombre de journaux vendus en kiosque la semaine suivante.

Un des enjeux actuels est donc de moduler la réputation des médias pour correspondre au plus de lecteurs possibles, afin de les conserver sur leur site. Un internaute reste en moyenne 4 minutes sur un site de médias, avec une moyenne de 3 articles parcourus. Plus de la moitié des lecteurs ne lisent qu'un seul article avant de se rendre sur un autre site Internet.

Une des stratégies est de multiplier les informations sous forme de vidéo.

9. Chapitre 7 : quelles solutions ?

Des *pure player* sont créés et favorisent les articles d'investigation (ex : [Les jours](#), [l'imprévu](#)). Des sites Internet diffusent aussi des informations thématiques de manière payante (ex : [Contexte](#)). Le modèle économique de ces nouveaux médias est cependant fragile et leur nombre nuit à leur visibilité.

Une des pistes est de créer des sites Internet entièrement payant : c'est le cas de 75% des journaux américains qui diffusent à plus de 50000 exemplaires. Facebook a développé Instant articles : les médias peuvent y déposer gratuitement leurs articles et les lecteurs les lisent directement depuis Facebook. Cela nuit bien sûr aux sites payants mais permet d'être mieux référencé. Les sites Internet payants souffrent également du non-respect de la propriété intellectuelle.

En France, la propriété intellectuelle ne protège pas l'idée, ni les faits. Il est donc difficile de protéger les articles des journalistes. En France, existe l'exception de courte citation, à condition de citer la source. Aux Etats-Unis, cette mesure est beaucoup plus large avec la notion d'utilisation loyale. Les problèmes de respect du droit d'auteur se posent principalement avec les agrégateurs de médias, notamment Google News. Google référence en effet les articles (titres, chapeau et photos) en utilisant l'exception de courte citation, sans jamais donner accès à l'intégralité des articles. Si cela génère des revenus pour Google, l'audience des médias référencés est également augmentée (environ 25 % du trafic sur ces sites Internet vient de Google). Les réflexions juridiques menées au niveau européen sur le droit voisin pourraient

solutionner en partie ces problèmes de droit d'auteur : rémunération lorsqu'il y a copie et bénéfice.

Contre la copie sauvage des articles, des logiciels ont été développés pour déceler ces copies. Des journaux revendent également aux autres médias leurs informations. La protection des contenus numériques doit aussi être adaptée à la production journalistique (Digital right management).

Pour réduire le coût de production de l'information, deux pistes sont envisageables : la mutualisation des moyens (partenariats journal/agence de presse) et les consortiums de journalistes. Mutualiser la diffusion de l'information est aussi envisageable : mettre en place des plateformes numériques où les articles de plusieurs médias seraient vendus à l'unité.

D'un point de vue politique, les auteurs suggèrent une politique transmedia, et non plus séparée entre la presse papier et le numérique. Soustraire les médias à la seule du marché, sur les fondements de l'intérêt général et de l'information comme bien public, permettrait d'envisager un autre rapport à la production et à la lecture d'informations. Le modèle juridique et économique d'une société de média à but non lucratif est une piste.

10. Conclusion

En guise de conclusion, les auteurs citent un extrait de *Sur la télévision* (Bourdieu, 1996) : « Pour être le premier à voir et à faire quelque chose, on est prêt à peu près à n'importe quoi, et comme on se copie mutuellement en vue de devancer les autres, de faire avant les autres, ou de faire autrement que les autres, on finit par faire tous la même chose ; la recherche de l'exclusivité, qui, ailleurs, dans d'autres champs, produit l'originalité, la singularité, aboutit ici à l'uniformisation et à la banalisation. »