



HAL
open science

La reconfiguration du pluralisme de l'information opérée par les réseaux socionumériques

Dario Compagno, Arnaud Mercier, Julien Mésangeau, Kamel Chelghoum

► To cite this version:

Dario Compagno, Arnaud Mercier, Julien Mésangeau, Kamel Chelghoum. La reconfiguration du pluralisme de l'information opérée par les réseaux socionumériques. *Réseaux : communication, technologie, société*, 2017, 5 (205), pp.91-116. 10.3917/res.205.0091 . hal-01723704

HAL Id: hal-01723704

<https://hal.univ-lorraine.fr/hal-01723704>

Submitted on 9 Mar 2021

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.

LA RECONFIGURATION
DU PLURALISME DE L'INFORMATION
OPÉRÉE PAR LES RÉSEAUX
SOCIONUMÉRIQUES

Dario COMPAGNO
Arnaud MERCIER
Julien MÉSANGEAU
Kamel CHELGHOUM

Twitter est un instrument d'infomédiation sociale de l'actualité (Smyrnaiois et Rebillard, 2011). En offrant à ses usagers la possibilité de partager librement des contenus d'actualité depuis leur fil d'actualité ou depuis les sites d'information, cette plateforme sociale contribue à reconfigurer l'équilibre des contenus offerts sans que les usagers soient conscients du résultat produit globalement. Via notre étude sur le partage des articles journalistiques français, nous sommes capables de caractériser de manière globale une tendance majeure de cette reconfiguration : l'offre d'information perd son pluralisme originaire pour se concentrer sur un nombre très restreint de médias. Le passage à travers Twitter permet aux médias qui ont une plus forte réputation et identité de prendre une place majeure dans la structure des partages, conformément à une logique qui résume un adage bien connu : « l'argent va à l'argent ». Néanmoins, cet effet n'est pas dû exclusivement à la productivité du média (nombre d'articles mis en ligne) et à son stock d'audience acquise (nombre de visites quotidiennes sur son site). Dès lors, pour mieux comprendre tous les facteurs explicatifs de ce phénomène, nous proposons d'analyser les logiques relationnelles des partages sur le réseau socionumérique.

En effet, l'offre d'information de chaque média est pensée à la source comme un tout complet qui entend construire un regard, un lectorat, une audience, dans une relation de fidélisation et de satisfaction de leurs attentes. Le média d'information se donne à voir comme un tout, certes composé d'articles, de reportages, mais dont l'agencement est organisé thématiquement (les rubriques) et éditorialement (la fameuse hiérarchie de l'information, avec les événements poussés en une, ceux traités seulement en brèves, etc.). Or la mise en circulation des articles sur les réseaux socionumériques par les internautes s'apparente à une décomposition de l'agencement rédactionnel. Cette décomposition se fait à chaque fois qu'un usager de Twitter (mais c'est vrai aussi de Facebook, de LinkedIn ou Reddit) choisit de publier un message contenant l'URL d'un article qu'il souhaite partager, extrayant d'un tout (le journal du jour, le site d'information) un seul morceau, et l'ajoutant sur son fil d'actualité à d'autres morceaux venus d'ailleurs. Car il peut partager librement des articles venant de plusieurs médias, sans préoccupation de fidélité à

un média particulier. Le partage d'une veille autour d'une thématique (l'élection, le sport, le foot, un pays, une ville...) ou le coup de cœur face à un article spécifique l'emportent sur un lien exclusif avec un média avec lequel on entretiendrait un rapport de confiance quotidienne (comme c'est souvent le cas du lectorat âgé avec « son » titre de PQR, par exemple). Par ailleurs, il existe des logiques d'usage en sérendipité dans la façon d'être confronté à un article sur ces réseaux, qui nous donne envie de le (re)partager : par l'alchimie des recommandations de ses « amis », par un algorithme qui filtre ce qu'il nous est permis de voir, dans le cas de Facebook ; par le hasard du moment où l'on se connecte à son fil d'actualité alors même que personne ou presque ne le remonte jamais dans sa totalité, surtout si on suit beaucoup de comptes, dans le cas de Twitter ; par l'existence ou non d'une stratégie systématique de veille. La consommation médiatique sur les réseaux socionumériques s'apparente donc largement à un picorage de l'information où la décomposition de l'offre d'information de chaque média est la règle et où la reconfiguration est le résultat d'un travail collectif. Cory Haik, responsable éditoriale du digital au *Washington Post*, nomme ce changement « the great unbundling of journalism »¹, qu'on peut traduire par le grand découplage, la grande dissociation.

L'éditorialisation de l'offre médiatique est déterminée par les choix des rédacteurs des médias, ces choix procédant de logiques de production d'information connues. Depuis les travaux pionniers de David Mann White, les journalistes sont considérés comme ceux qui filtrent et trient, parmi les centaines d'occurrences quotidiennes, ce qui mérite d'être considéré comme une information, et donc d'être mis en forme et en récit pour s'adapter aux formats médiatiques. Ce rôle de *gatekeeper* (White, 1950) confère aux journalistes une maîtrise sur ce qui fait l'actualité et, par extension, le fond de la conversation sociale, voire de l'action publique. De fait, les médias exercent une fonction d'*agenda building* et d'*agenda setting* (McCombs et Shaw, 1972), du moins tant qu'ils apparaissent comme les détenteurs d'un certain monopole de l'accès à l'espace public envisagé en priorité comme l'espace public médiatique. Or précisément, l'ère du numérique octroie ce rôle de *gatekeeper* à bien d'autres acteurs que les journalistes professionnels. Internet est pour l'individu un espace de diffusion potentiellement massif de sa parole, qui pourra selon les circonstances avoir un retentissement au-delà de la sphère des proches. Les réseaux socionumériques en tant que puissants outils de désintermédiation

1. Interview avec Ravi Somaiya, « How Facebook Is Changing the Way Its Users Consume Journalism », *New York Times*, 27 octobre 2014.

journalistique bouleversent l'écosystème de l'information et la position dominante et oligopolistique des journalistes dans l'accès à l'espace public. Cette déstabilisation de la position des journalistes touche aussi bien le monopole de production de l'information, que les circuits de diffusion, et même la maîtrise de la circulation des contenus d'actualité pourtant produits par les journalistes. En organisant une veille et une curation d'informations qu'ils repara-tagent via leurs réseaux sociaux numériques, les internautes ont le pouvoir de rééditorialiser les contenus médiatiques. Cette rééditorialisation diverge de la traditionnelle revue de presse (collection composite de photocopies d'articles émanant de divers supports), car c'est une mise en forme et en circulation des contenus à destination d'un très large public potentiel, et avec ajouts de commentaires, d'interpellations, de photos, de mentions d'autres comptes...

Bien évidemment, les usagers de Twitter ont des profils sociologiques assez spécifiques, et il ne faut donc pas généraliser ce que l'on peut dire des partages sur ce réseau social numérique à l'ensemble de la consommation d'informations sur Internet. Cependant, Twitter reste un objet important, étant un réseau social numérique souvent utilisé pour le partage de l'information journalistique, par les citoyens tout comme par des professionnels. La sociologie des usagers de Twitter en France nous apprend que 54 % des utilisateurs français de Twitter s'intéressent à l'actualité et à l'économie (source : sondage Harris Interactive, 2015). Et sur les plus de 10 millions d'utilisateurs actifs, l'utilisateur moyen français a moins de 34 ans (58 %), où les diplômés sont largement surreprésentés : 40 % des utilisateurs de Twitter en France ont un diplôme supérieur, 43 % sont issus des CSP+ (sources : Nielsen, 2015 et GlobalWebIndex, 2015). Et si une étude sur la consommation médiatique des jeunes en France montre que Facebook est le moyen largement majoritaire pour accéder aux informations (Mercier *et al.*, 2017), Twitter n'est néanmoins pas à négliger dans ce nouvel écosystème, car il est utilisé justement pour suivre l'actualité. Twitter (comme Facebook) devient pour certains internautes une sorte de kiosque, donnant à voir une sélection d'articles issus de divers médias. Ce kiosque n'offre donc pas accès à la totalité des contenus publiés par les médias de la même manière. Et surtout, le contenu du kiosque est le résultat de la somme des actions individuelles de chaque détenteur de compte, sans conscience collective du résultat produit. Dès lors que Twitter s'apparente à un kiosque d'articles, on peut parler d'offre médiatique reconfigurée, collectivement. Cette réalité repose sous un jour nouveau la question du pluralisme des informations disponibles sur Internet.

Les observations empiriques que nous avons conduites dans une étude précédente montrent que les personnes identifiées comme de « gros » partageurs d'informations sur Twitter tendent à remettre en circulation des articles issus de plusieurs médias, représentant des sensibilités et des lignes éditoriales n'étant pas nécessairement homogènes. À partir d'un panel d'utilisateurs de Twitter issus d'un premier corpus de millions de tweets constitué en 2014, nous avons étudié leurs stratégies de partage de liens d'actualité alors qu'ils étaient parmi ceux en partageant le plus (Mercier *et al.*, 2017). Durant quatre mois, de novembre 2015 à mars 2016, nous avons collecté tous les tweets de 117 comptes pour étudier (et comparer) leurs pratiques de partage d'information. Globalement, ces comptes ont remis en circulation des articles issus *a minima* de 4 médias et au maximum de 26 médias. La dispersion des liens vers les médias étudiés est donc très variable. Ce qui recoupe une autre disparité, puisque la proportion de tweets diffusant des articles provenant de médias dans le total des tweets pour chaque compte varie de 20 % à 98 %. Malgré cette diversité et cette hétérogénéité, on repère six sources dominantes mobilisées. Par ordre : *Le Figaro*, *Le Monde*, *Le Point*, *BFM-TV*, *Le Parisien*, *20 Minutes*.

Notre démarche actuelle doit nous permettre d'avoir un aperçu global de cette nouvelle intermédiation où les individus participent à l'éditorialisation des contenus. Ce que nous souhaitons étudier est une *rééditorialisation distribuée*, déterminée de manière collective par les utilisateurs. Les formes d'appropriation pouvant être diversifiées, nous allons observer une rééditorialisation reposant sur des régimes pluriels de consommation de l'information, stratégiques pour certains, exploratoires pour d'autres.

MÉTHODOLOGIE ET CORPUS

La reconfiguration, entre analyse du discours et analyse de la réception

Le partage des contenus journalistiques sur Twitter passe par *deux moments* : le premier est celui de la lecture des tweets (même minimale et fugace) et le second est quant à lui le temps de la rediffusion volontaire (manuelle ou assistée, voir Compagno et Pignard-Cheynel, 2017, à paraître). La rediffusion quant à elle peut se réaliser de deux manières : une première, « fruste », en partageant des tweets ne contenant aucun autre message que l'URL et le titre de l'article (souvent le message est généré automatiquement à partir du bouton de partage sur le site du média) ; une seconde manière, réellement

« active », où les tweets contiennent les informations de base (URL et titre) et des commentaires et ajouts personnels, ce que l'on appelle les « discours d'es-corte » (Simon *et al.*, 2017). Dans l'approche que nous développons ici, les tweets – « frustes » ou « actifs » – sont vus comme des *traces interprétatives*, qui témoignent d'une influence subie, mais aussi d'une volonté d'influencer. Chaque tweet est le résultat d'une action et est susceptible de produire des effets sur les autres usagers dans le même environnement numérique. Plus spécifiquement, les tweets qui partagent ou repartagent des informations journalistiques contribuent à redéfinir l'offre d'information elle-même. Il n'y a plus de seuil net entre production et réception de l'information, et chaque usager peut jouer un rôle dans les deux.

Cela a une conséquence pour la pratique de la recherche en SHS : en traitant l'information en ligne, nous sommes obligés de remettre en question la distinction nette entre analyse du discours et études de réception. D'un point de vue méthodologique, s'il y a ici un réel enjeu à repenser les domaines traditionnels de la recherche avec les données disponibles sur Internet (Rebillard, 2011), il reste à identifier le procédé qui soit susceptible de couvrir, d'une part, les préoccupations de la sémiotique et de l'analyse du discours et, d'autre part, celles de la sociologie de l'information. En fait, l'analyse du contenu et l'analyse de la réception *convergent* là où la réception met elle-même en forme le contenu. Dans notre travail, nous cherchons à développer une approche qui rend l'analyse du sens et des processus relationnels complémentaires. Pour cette raison, nous étudions à la fois les tweets comme des textes (en relation intertextuelle avec les articles journalistiques auxquels ils se réfèrent) et comme résultat de l'action d'énonciation des usagers qui les produisent. Ce qui nous semble intéressant est le processus dit de *sémiose sociale*, les chaînes de production et d'interprétation de traces signifiantes. Lire un article pousse à partager ce contenu, et les liens que l'utilisateur entretient avec d'autres usagers (sa position dans la structure du réseau) activent des effets de partage ultérieur du même contenu.

Paul Ricœur (1983) a formulé un célèbre modèle de l'activité de production et interprétation qui distingue trois moments. Il y a un « avant » du discours, la *préfiguration* des événements et des valeurs qui sont partagées par les groupes sociaux. Par exemple, si les journalistes s'accordent sur ce qui est digne de couvertures médiatiques, c'est parce qu'ils partagent ce fond de savoir et de sentir commun. Deuxièmement il y a « l'ici » du discours, la *configuration* opérée par les producteurs de discours, y compris le discours médiatique (Charaudeau, 1997). La production discursive donne un sens aux

événements parce qu'ils se trouvent structurés dans des récits, éventuellement dans des récits distribués comme dans le cas du discours médiatique (Compagno, 2014). Troisièmement, il y a un « après » du discours qui ne peut pas être oublié. Ricœur l'appelle la *refiguration* de ce qui est raconté. Cette appropriation des récepteurs met en rapport ce qui est écrit avec ce qui est actuel au moment de l'interprétation. Il n'y aurait pas de sens sans refiguration. Remarquons que la refiguration n'est jamais « sauvage » ou aléatoire, elle n'est pas une simple *utilisation* des textes (Eco, 1992). Le concept de refiguration sert justement à rendre compte de l'activité interprétative respectant les contraintes textuelles, relevant des systèmes et des normes sémiotiques immanentes aux pratiques sociales.

Nous verrons que notre étude montre les traces de la refiguration des usagers, d'une refiguration *active, collective et distribuée* que l'on peut bien appeler *reconfiguration*. Il s'agit en effet d'une activité de production de discours, où les usagers deviennent des instances co-énonçantes. Cela ne veut pas dire que les usagers réalisent un « deuxième discours » ou même un « contre-discours ». L'activité des usagers est bien encadrée par la production institutionnelle d'information. Plusieurs études nous montrent désormais que les réseaux sociaux numériques s'appuient fortement sur les médias et ne s'y substituent aucunement (voir par exemple Rebillard et Loicq, 2013). Les « textes » sont encore bien là, les internautes ne se passent pas de l'écriture médiatique et de ses « auteurs ». Néanmoins, nous verrons que l'offre est transformée par les logiques de partage du réseau. Il y a une liberté que le dispositif de Twitter accorde à ses usagers, liberté qui a un sens par rapport au discours médiatique et à une possibilité limitée de lui donner un nouvel équilibre. Deuxièmement, l'activité des usagers produit bien une nouvelle configuration de l'offre médiatique, mais celle-ci est un *effet émergent*, il n'y a pas de stratégie unitaire responsable du résultat. Il s'agit de la somme des comportements individuels qui aboutissent à un résultat qui n'a été recherché par personne, selon une logique d'effets d'agrégation débouchant sur un résultat global. En étudiant « ce que Twitter fait aux médias », donc, nous pouvons à la fois apprendre des choses sur les liens et les pratiques sociales de consommation de l'information (traditionnellement étudiées par analyse de la réception) et sur cette nouvelle forme de discours que nous pourrions nommer « l'information distribuée ». D'un point de vue technique, nous allons combiner les résultats d'une analyse textuelle assistée par ordinateur ainsi qu'une analyse de réseaux (d'une manière proche de celle que l'on peut retrouver par exemple dans les travaux de Smyrniotis et Ratinaud, 2013).

Corpus

L'absence de filtrage ou de hiérarchisation des partages sur les comptes des usagers (en tout cas, pas au moment de la collecte des données) est une de nos raisons profondes d'étudier le partage d'informations sur Twitter. Tous les contenus partagés (« tweetés ») semblent avoir la même possibilité d'être lus et partagés ultérieurement (« retweetés ») que les autres². Il s'agit donc d'une logique transparente de partage qui nous permet d'étudier plus simplement la transformation de l'offre journalistique eu égard à ce que l'on peut observer sur Facebook ou sur d'autres réseaux sociaux numériques qui filtrent les contenus partagés, avec en plus dans le cas de Facebook un algorithme qui impose ses choix en toute opacité (Cardon, 2015).

Nous allons nous concentrer sur la rediffusion d'informations produites par des médias reconnus, et souhaitons donc exclure la production de contenus originaux opérée par les usagers eux-mêmes. Notre démarche consiste à travailler sur la base d'un groupe d'usagers ne « faisant population », au sens méthodologique du terme, qu'à partir d'une famille précise de pratiques sémiotiques : la consommation et la rediffusion d'informations issues de quelques grands quotidiens et magazines. Nous ne travaillons pas prioritairement la dimension linguistique des partages, par exemple dans son *framing* sélectif de l'information (Huang, 2015) réalisé à travers le « discours d'accompagnement » consistant pour l'internaute à placer dans le tweet, autour des liens hypertextes et du titre de l'article, des contenus qui peuvent être un mot, une courte phrase, des émoticônes, un hashtag... (Simon *et al.*, 2017). Nous ne prenons pas en compte les variables géographiques structurant en amont le réseau des partages (Takhteyev *et al.*, 2012). Enfin, nous ne prenons pas en compte non plus le partage d'informations sur la base des seules structures de retweets (Chelghoum, Hoang et Kacem, 2016).

En particulier, nous étudions un ample spectre de l'offre d'information institutionnelle française : un périmètre de 23 sites web d'information sélectionnés sur la base de leur importance et de leur représentativité par rapport aux différents types de presses : presse militante ou non, médias nés en ligne ou mis en ligne à partir d'un support préexistant, médias locaux ou nationaux, quotidiens ou magazines, gratuits ou payants, généralistes ou spécialisés, représentant

2. Quand les données de cette étude ont été collectées, en 2014, Twitter ne réorganisait pas l'ordre dans lequel les tweets apparaissent dans le flux des usagers.

les divers supports³. Notre corpus de travail pour cette étude est constitué de tous les articles publiés par un des médias du périmètre pendant la première semaine du mois de septembre 2014, articles mentionnés (avec un hyperlien) par des tweets créés au plus un mois après la publication de l'article⁴. Le corpus de travail inclut également tous ces tweets. Le mois de septembre a été choisi parce que nous avons pu déterminer, à partir d'études exploratoires sur la base, que les données collectées présentaient le moins d'anomalies. Nous avons décidé de ne prendre en compte que les usagers avec au moins 5 abonnés, que les articles partagés au moins 5 fois et que les médias dont on a mis en partage au moins 50 articles. En plus, nous n'avons pris en compte que les médias pour lesquels nous disposons de données sur leur fréquentation (classement OJD septembre 2014). Cela donne un périmètre réduit à 23 médias⁵ et un total de 12 949 articles et de 550 652 tweets partagés par 135 305 comptes. Nous estimons avoir accès à une sélection hautement représentative de la totalité des informations médiatiques partagées sur Twitter pendant la période étudiée pour ces 23 médias. Notre première assumption est que notre collection d'articles est aussi une représentation hautement représentative de l'offre d'information des 23 médias pendant la première semaine de septembre 2014. En fait, les seuls articles non récoltés seraient ceux qui ont été twittés moins

3. Les données sont celles de la base créée lors de la mise en œuvre du projet Info-RSN, financé par l'ANR et porté par le CREM (Centre de recherche sur les médiations) à l'Université de Lorraine, en partenariat avec le LCOMS (Laboratoire de conception, optimisation et modélisation des systèmes) et la société privée Semicast. L'étude a notamment consisté à collecter, traiter et analyser deux récoltes de données sur Twitter en lien avec le partage d'informations (Mercier *et al.*, 2017). La première phase de collecte s'est faite de mi-mai à mi-octobre 2014 en collectant systématiquement tous les tweets qui comportaient un lien URL pointant vers 31 sites d'information prédéfinis. Cette base a ensuite été longuement traitée, par des spécialistes en information-communication et par des informaticiens associés à l'ANR Info-RSN, pour limiter les possibles sources de bruit (par exemple des articles diffusés avec plusieurs URL ou des liens qui ne faisaient pas référence à des vraies informations journalistiques). Finalement, plusieurs indicateurs ont été calculés pour synthétiser l'activité de partage des informations et évaluer leur prédictivité.

4. Nous avons déjà pu vérifier que sur Twitter on retrouve des articles partagés des mois et même des années après la publication (Compagno et Mercier, à paraître, 2017). Pour cette étude, nous limitons notre collection aux tweets créés au plus un mois après la publication de l'article mentionné, écart « court » qui tient compte de la réception récente des contenus. Ce choix donne à chaque article la même opportunité *a priori* d'être partagé.

5. Liste des médias retenus : leparisien.fr; courrierinternational.com; france24.com; franceinter.fr; huffingtonpost.fr; lesechos.fr; lequipe.fr; mety.fr; lemonde.fr; 20minutes.fr; nouvelobs.com; francetvinfo.fr; franceinfo.fr; ouest-france.fr; lexpress.fr; lepoint.fr; lefigaro.fr; bfmtv.com; latribune.fr; rue89.nouvelobs.com; liberation.fr; ladepeche.fr; la-croix.com.

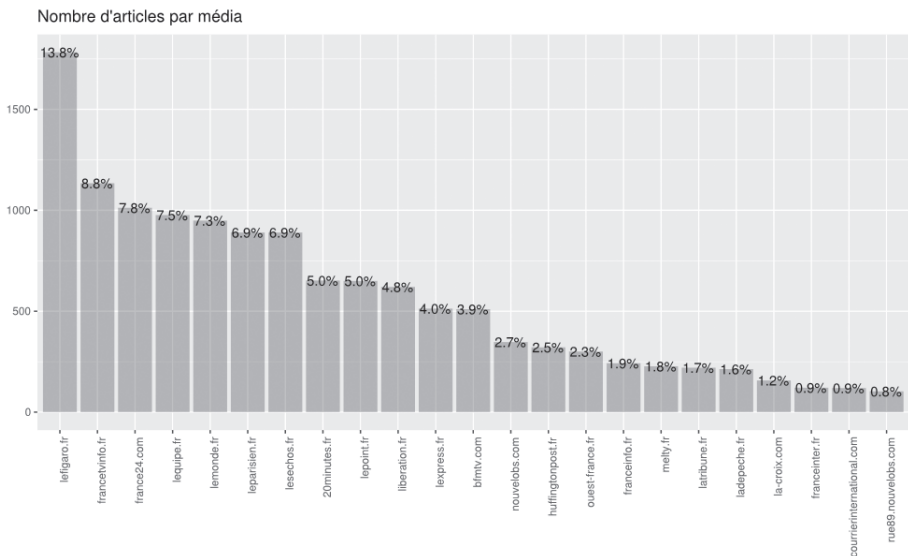
de 5 fois pendant le mois de septembre 2014. Ces articles sont forcément en nombre très limité et peuvent bien être considérés comme secondaires dans l'offre globale d'information.

LE PLURALISME DE L'INFORMATION RECONFIGURÉ PAR TWITTER

L'offre médiatique (OM), la fréquentation médiatique (FM) et l'offre reconfigurée (OR)

Nous appelons offre médiatique (OM) la totalité des articles publiés par un ensemble de médias pendant une période donnée. L'OM de notre étude inclut la totalité des articles publiés entre le 1^{er} et le 7 septembre 2014 par les 23 médias de notre périmètre, et partagés au moins cinq fois sur Twitter pendant ce mois. Les médias ne produisent pas le même nombre d'articles, ce qui détermine une distribution non uniforme de l'OM par rapport aux médias qui la structurent. La figure 1 montre la répartition des articles selon les médias qui les ont publiés.

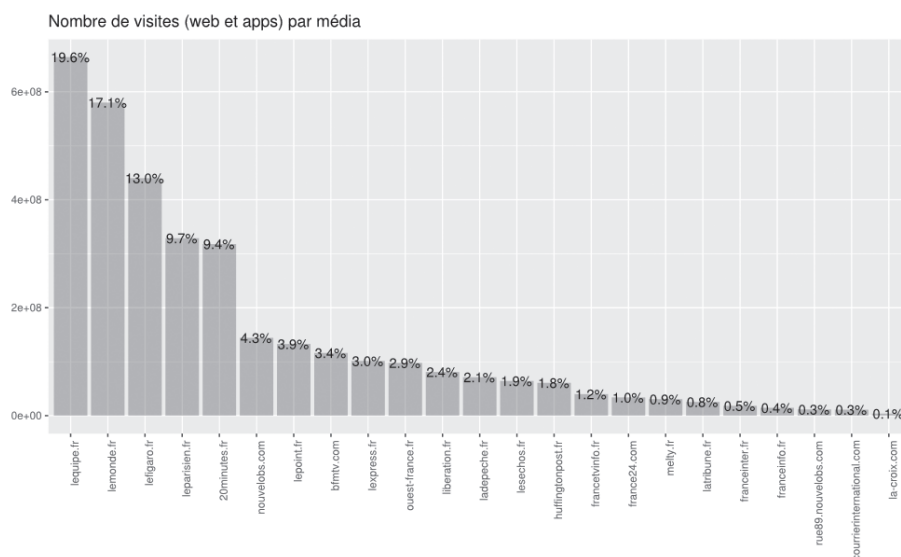
Figure 1. Nombre d'articles par médias



Source : auteurs.

Le but de cette étude est de voir comment l'offre médiatique, déterminée exclusivement par des choix éditoriaux propres aux médias, est affectée par les possibilités de rééditorialisation propres à Twitter. Ce qui compte pour notre étude est donc le nombre d'articles produits (OM) et la reconfiguration sociale de l'offre. Cependant, nous avons pris en compte des données sur la consommation de l'offre médiatique, ou *fréquentation médiatique* (FM). Il s'agit des données sur la fréquentation des sites web et des applications pour dispositifs mobiles publiées par l'Alliance pour les chiffres de la presse et des médias (ACPM, 2014a, 2014b). Ces données nous permettent d'avoir un aperçu de la fréquentation médiatique indépendamment du filtrage opéré par les réseaux sociaux numériques. Les données publiées par l'ACPM sont agrégées par médias. La figure 2 montre les médias classés par leur fréquentation relative.

Figure 2. Nombre de visites par média

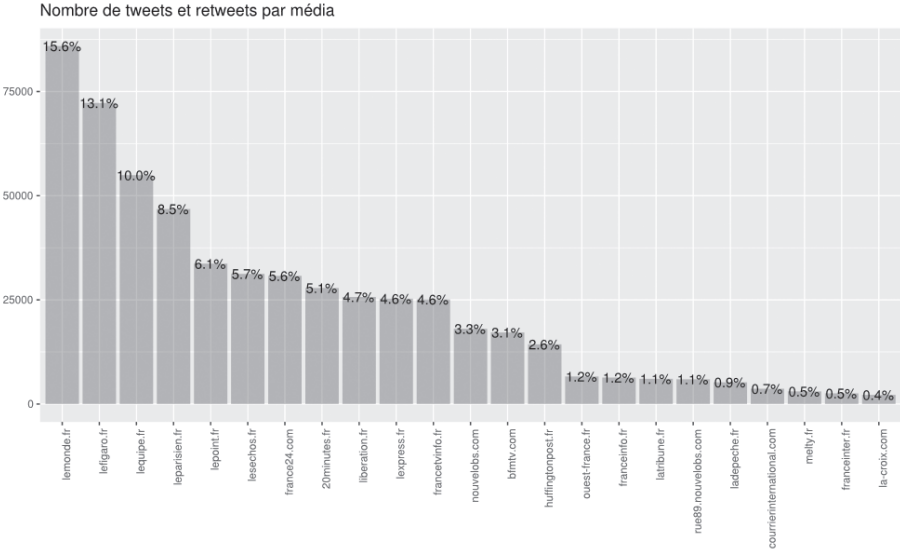


Source : auteurs.

Enfin, nous appelons offre reconfigurée (OR) la totalité des partages d'articles sur un ou plusieurs réseaux sociaux numériques pendant une période donnée.

Le « partage » est l'acte qui permet à d'autres usagers du même réseau social de recevoir l'article. L'offre reconfigurée est essentielle à prendre en compte désormais, puisque la fréquentation de nombreux sites devient de plus en plus dépendante du trafic généré par les liens postés sur les réseaux sociaux. Sur Twitter, l'acte de partage est le tweet. La figure 3 montre les médias classés par nombre de tweets partageant leurs URL.

Figure 3. Nombre de tweets et retweets par média



Source : auteurs.

Nous remarquons d'ores et déjà que l'ordre des médias a changé : le nombre de tweets obtenus par un média ne correspond pas nécessairement au nombre d'URL qu'il produit (figure 1) ni à la fréquentation de son site ou application (figure 2). Si *Le Figaro* produit le plus gros nombre d'articles et si *L'Équipe* est le média le plus fréquenté, ce sont les contenus du *Monde* qui sont les plus tweetés. En fonction de l'échelle de classement construite, la hiérarchie entre les médias diverge donc, faisant apparaître, dans le top 6, des médias qui n'étaient pas classés sur un autre critère, et faisant progresser le journal *Le Monde*, comme on le voit dans le tableau récapitulatif suivant.

Tableau 1. Comparaison des hiérarchies médias sur les trois échelles

Offre médiatique (OM)	Fréquentation médiatique (FM)	Offre reconfigurée (OR)
<i>Le Figaro</i>	<i>L'Équipe</i>	<i>Le Monde</i>
<i>France TV Info</i>	<i>Le Monde</i>	<i>Le Figaro</i>
<i>France 24</i>	<i>Le Figaro</i>	<i>L'Équipe</i>
<i>L'Équipe</i>	<i>Le Parisien</i>	<i>Le Parisien</i>
<i>Le Monde</i>	<i>20 Minutes</i>	<i>Le Point</i>
<i>Le Parisien</i>	<i>Nouvel Obs</i>	<i>Les Echos</i>

Source : auteurs.

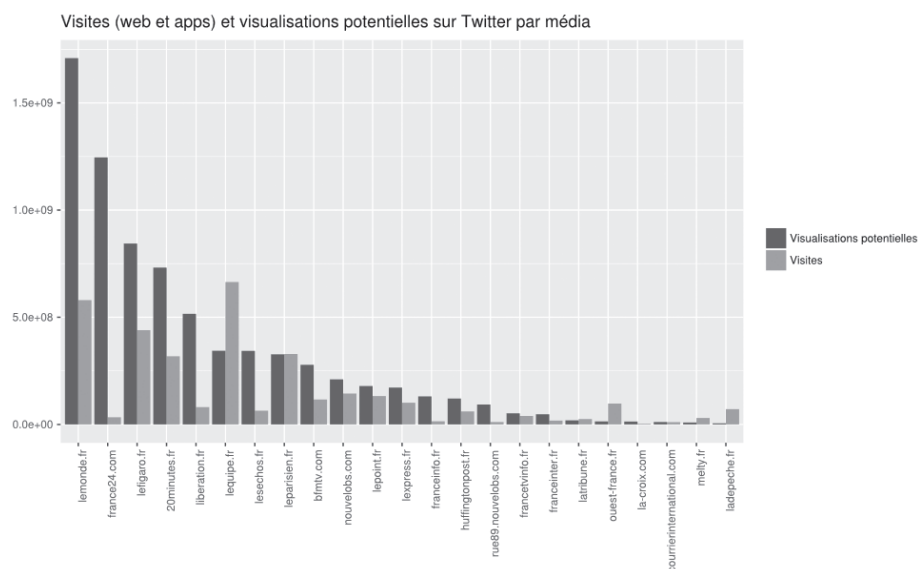
La prise en compte du nombre brut de tweets n'est cependant pas suffisante pour saisir l'impact d'un média sur la totalité des partages. En effet, les tweets n'ont pas tous le même « poids », puisqu'en fonction du nombre d'abonnés de chaque compte le potentiel de visibilité de chaque tweet diffère. Et par voie de conséquence, les URL partagées sur Twitter *n'ont pas non plus le même potentiel de visibilité*. La probabilité de voir apparaître certains contenus sur nos fils d'actualité n'est ni constante ni identique. Cela dépend du dispositif de Twitter – qui partage des logiques communes, bien que non parfaitement équivalentes, avec la plupart des réseaux socionumériques contemporains. Sur Twitter, chaque tweet produit par un usager est affiché sur le flux de tous ses abonnés. Cela veut dire qu'une URL partagée par un compte ayant beaucoup d'abonnés sera potentiellement vue par beaucoup d'utilisateurs. En parallèle, une URL partagée par beaucoup de comptes sera potentiellement vue par tous leurs abonnés, et donc par beaucoup d'autres comptes. Sur la base de ces considérations, nous avons établi de mesurer le potentiel de lecture d'un contenu journalistique sur le réseau socionumérique en mobilisant *la somme du nombre d'abonnés de tous les comptes partageant son URL*. Cela rend compte du nombre exact de comptes susceptibles de recevoir l'URL dans leur flux. Cette mesure prend bien sûr en compte également les utilisateurs qui utilisent Twitter de manière « passive », en lisant les tweets, mais sans tweeter à leur tour. En fait, ces lecteurs reçoivent quand même les tweets dans leurs flux, et sont donc comptabilisés tout comme les utilisateurs « actifs ».

Afin de calculer cette valeur, nous avons construit un réseau dit *biparti* (Faust, 1997 ; Borgatti et Everett, 1997 ; Latapy *et al.*, 2008)⁶. Dans ce réseau

6. Le traitement, l'analyse et la visualisation des données ont été réalisés avec *R* (R Core Team, 2015), *tidyverse* (Wickham et Grolemund, 2017) et *igraph* (Csardi et Nepusz, 2006). Le code est disponible pour consultation en contactant les auteurs.

sociosémiotique, les sommets sont soit des URL soit des usagers ; chaque tweet partageant une URL est représenté par un arc qui va du sommet de l'utilisateur à celui de l'URL. Chaque arc est pondéré sur la base du nombre d'abonnés de l'utilisateur publiant le tweet. Le potentiel de visualisation de chaque article est alors identifié avec son *degré entrant pondéré*. La figure 4 compare la fréquentation de chaque média avec le potentiel de visibilité global de leurs articles.

Figure 4. Visites et visualisations potentielles sur Twitter par média



Source : auteurs.

En prenant ainsi en compte le potentiel de visualisation, on obtient un résultat qui renforce certaines tendances que l'on avait déjà pu remarquer avec juste le nombre brut de tweets (figure 3). *Le Monde* est bien le média qui a la plus grande portée, mais maintenant on peut apprécier la valeur de cette assertion : *Le Monde* se démarque grandement du *Figaro* puisque son potentiel de visibilité est presque double.

L'Équipe fait beaucoup moins bien sûr Twitter que les autres médias de notre périmètre. Il faut dire que nous avons montré par ailleurs que les résultats sportifs publiés par *L'Équipe* faisaient l'objet d'un intensif travail de publication

automatisée par des *bots*, sans que l'on soit bien certain de l'efficacité d'audience de pareil procédé (Compagno, Pignard-Cheynel, 2017, à paraître). Et *France 24* apparaît comme le deuxième média ayant globalement la plus grande portée. Ce résultat est impressionnant vu l'écart avec les fréquentations de son site et le nombre relativement modeste de tweets bruts. Mais les usagers qui tweetent les contenus de *France 24* (souvent en anglais ou en arabe) ont un nombre moyen très élevé d'abonnés. On voit là que *France 24*, par son ambition internationale, évolue dans un écosystème d'information très spécifique, incomparable aux autres médias français étudiés, ce qui produit ce « désordre » dans le classement qui aurait pu être attendu puisqu'il devance largement *Le Figaro* et *20 Minutes*, pourvoyeurs bien connus de nombreuses publications quotidiennes.

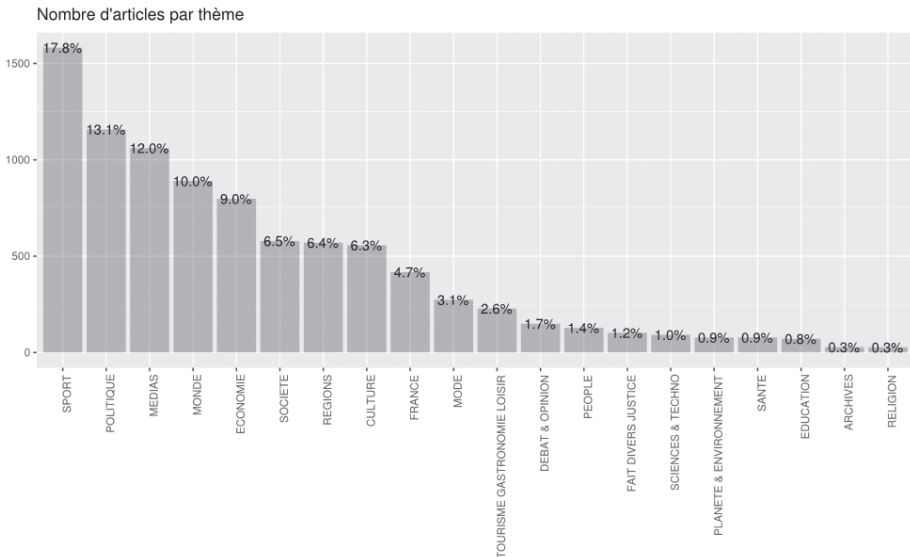
Comparaison d'OM, FM et OR par thème

Si chaque contenu médiatique était partagé sur Twitter exactement le même nombre de fois, par des comptes qui ont exactement la même importance dans le réseau, l'OR et l'OM correspondant seraient équivalents. Le pluralisme de l'offre ne serait donc pas affecté par la structure des partages sur Twitter. Cependant, nous observons que cette structure de partage joue un rôle déterminant pour filtrer l'offre initiale d'information médiatique et la reconfigurer radicalement en fonction, notamment, des structures relationnelles dans lesquelles le processus de partage de l'information est encastré. Pour avoir un aperçu plus précis de la reconfiguration de l'offre d'information opérée par Twitter, nous allons maintenant segmenter l'offre thématiquement.

Le pluralisme de l'information en ligne a déjà fait l'objet de plusieurs travaux. En particulier, Franck Rebillard étudie l'information selon le prisme de son *cadre primaire* et *secondaire* (Rebillard *et al.*, 2012). Le cadre primaire comptabilise le nombre de *sujets d'actualité* traités pendant une période. Chaque sujet est un événement qui a retenu l'attention médiatique, indépendamment des choix idéologiques des différents médias. L'étude des sujets d'actualité permet de rendre compte de « l'ampleur » de l'offre médiatique, c'est-à-dire du nombre d'événements traités au même moment. Le cadre secondaire prend lui en compte la manière dont des articles différents traitent d'un même sujet d'actualité. L'étude du cadre secondaire vise donc à rendre compte de « la profondeur » de l'offre, c'est-à-dire de la complexité des perspectives interprétatives associées au traitement d'un même événement.

L'offre médiatique est structurée par les différences « en ampleur et en profondeur » entre les articles publiés. Nous posons d'abord que chaque média a sa spécificité idéologique (sa vision cohérente du monde), correspondant à un cadrage secondaire général. Nous allons alors différencier les articles sur la base des médias qui les ont publiés. Pour le cadrage primaire, il s'agit de mesurer la variété de l'offre d'information en termes d'agrégation de sujets d'actualité (*thèmes*). Certaines OM pourraient être caractérisées par une concentration sur un nombre réduit de thèmes, d'autre part une distribution plus dispersée de thèmes. Nous nous attendons à ce que certains sujets (par exemple la politique) soient traités plus que d'autres. La figure 5 montre la répartition des articles selon leurs thèmes.

Figure 5. Nombre d'articles par thème



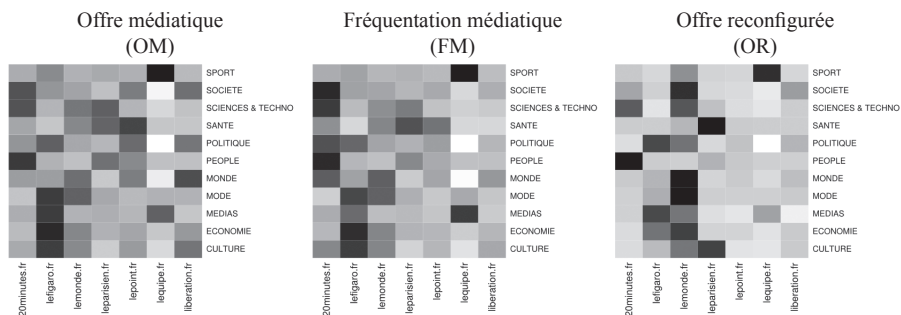
Source : auteurs.

Ces thèmes ont été identifiés en plusieurs étapes. Nous avons d'abord recherché au sein des URL la présence de rubriques éditoriales, ce qui a permis de dresser une première liste. Nous avons également consulté les médias afin d'identifier les rubricages constitutifs de leurs sites, ce qui a complété la liste initiale (puisque toutes les URL de tous les médias ne portent pas forcément trace de la rubrique). Nous avons ensuite réalisé une classification manuelle

de ces rubriques utilisées par les médias, en produisant les regroupements nécessaires pour aboutir à une liste de 20 thèmes. Cette classification a enfin été utilisée pour catégoriser de manière automatique les articles à partir de l'occurrence du nom de la rubrique apparaissant dans les URL de notre base de tweets. Pour catégoriser les articles produits par des médias qui ne mentionnent pas nativement un rubricage, nous avons pris en compte les tags ou d'autres métadonnées accompagnant le descriptif pour les moteurs de recherche d'un article. Quelques rares articles échappent aux thèmes identifiés qui couvrent les grands domaines de l'information. D'autres sont à l'intersection et n'ont pas été classés (refusant le double arbitraire de les classer dans un seul thème, ou de les faire compter deux fois en les classant dans deux rubriques). Cela donne comme résultat une classification que nous jugeons satisfaisante, car en mesure de couvrir un peu plus de 70 % des articles de notre base de données. Cette procédure s'inscrit dans une approche quali-quantitative, où la classification des articles n'est ni entièrement manuelle, ni totalement automatisée. Or la procédure utilisée pour identifier les thèmes fait qu'ils ne sont pas homogènes et directement comparables : par exemple, les thèmes majeurs sont traités par la plupart des médias, par contre certains des thèmes mineurs peuvent être abordés de manière différenciée par un nombre très limité de médias. Pour cette raison nous allons nous concentrer exclusivement sur ces thèmes que l'on retrouve à travers plusieurs médias.

videmment, le fait que chaque article est décrit par deux variables (thème et média), fait que OM, FM et OR peuvent être représentées comme des espaces bidimensionnels où chaque article occupe un point. La figure 6 permet de comparer les OM, FM et OR des 7 plus gros médias segmentés par thème.

Figure 6. Comparaison d'OM, FM et OR par thème et média



Source : auteurs.

Les cellules des trois parties de la figure 6 « contiennent » exactement les mêmes articles. Cependant, les articles sont soit tout simplement comptabilisés de manière brute (OM), soit pondérés par rapport à la fréquentation médiatique des médias (FM) soit par rapport à leur potentiel de visualisation sur Twitter (OR). Chaque cellule de la partie gauche de la figure 6 (OR) représente les articles produits par un média (axe des abscisses) sur un certain thème (axe des ordonnées). Pour chaque ligne, nous avons calculé la *distance euclidienne* entre les cellules. Cela nous permet de voir quels médias ont le plus traité chaque thème. Une cellule à la couleur plus foncée signifie que le thème de la ligne a été beaucoup plus traité par un média par rapport aux autres. La fréquentation médiatique par thème (FM, partie centrale de la figure 6) est approximée : on ne disposait que de données agrégées par média. Chaque cellule représente le nombre d'articles produits par un média multiplié par le nombre moyen de visiteurs de ces pages. Il s'agit d'une approximation puisque nous ne pouvons pas identifier pour chaque média, si les articles sur un thème sont plus populaires que d'autres. La partie droite de la figure 6 représente l'OR étudiée, avec la somme du potentiel de lecture de tous les articles. Une couleur plus foncée représente alors un plus grand potentiel de lecture sur Twitter des articles produits sur un thème par un certain média.

Il apparait que le critère de la fréquentation des sites des médias ne modifie pas l'équilibre et le pluralisme de l'offre médiatique. Bien qu'un examen statistique précis ne soit pas possible (vu que le nombre absolu d'articles dans l'OM n'est pas directement comparable au nombre de visites), on peut l'affirmer grâce à une comparaison proportionnelle, rendue possible par le calcul et la visualisation des distances euclidiennes. De plus, le passage de l'OM à la FM ne montre pas une concentration de l'offre. Parmi les différences que l'on observe toutefois, *20 Minutes* prend plus d'importance dans les thèmes « Politique » et « Monde », en dépit de *Libération* et du *Point* plus attendus sur ces rubriques.

Au contraire, la comparaison avec l'OR laisse transparatre un nouvel équilibre qui a pour forme une concentration de l'offre éditoriale potentielle vers *Le Monde*. Ce média a un grand ou même un très grand potentiel de visibilité sur Twitter, par rapport aux autres grands médias, sur tous les thèmes sauf pour le « Sport », la « Santé » et les articles « People », *Le Figaro* est plus visible que *Le Monde* en « Politique » et « Médias », ce qui prouve que si *Le Monde* prend une place dominante il n'en devient pas pour autant hégémonique.

Reconfiguration et centralité

Le visage de l'information, passé au filtre de Twitter, devient un portrait déformé. Bien sûr, ces résultats sont à mettre en perspective par rapport à l'identité sociodémographique des usagers de Twitter. Cependant on ne peut que remarquer une réduction du pluralisme de l'information dans son passage à travers Twitter.

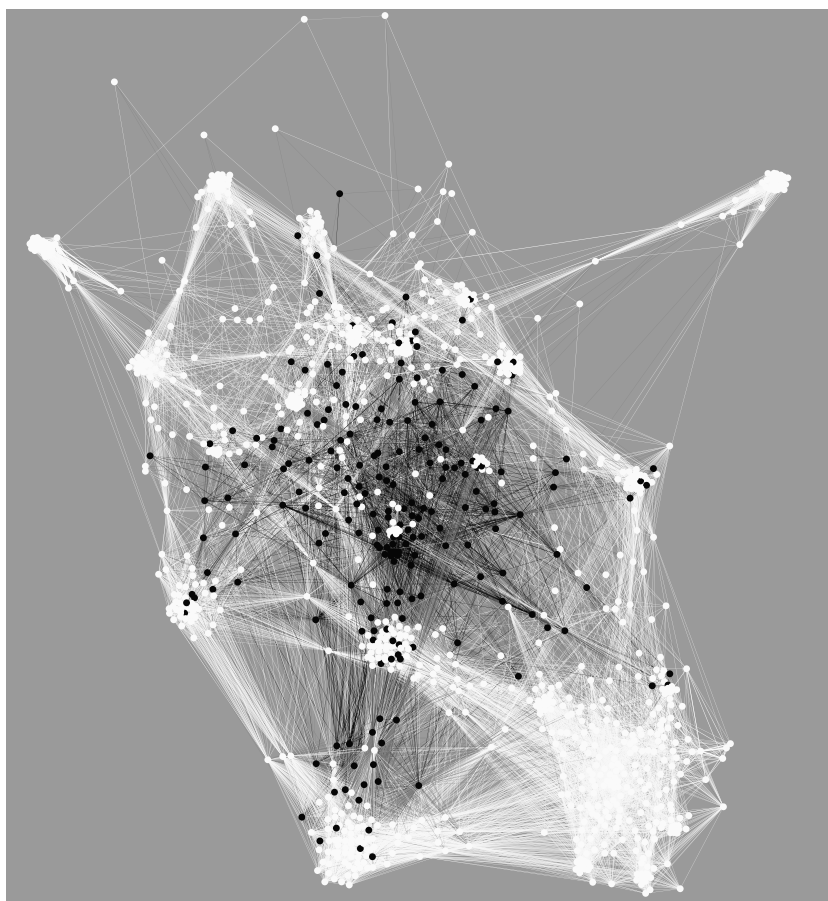
Comment expliquer ces résultats ? Il n'est pas suffisant de faire référence à des facteurs externes au réseau socionumérique, tels que le nombre d'URL créées par les rédactions ou la notoriété et fréquentation des médias hors Twitter. Nous avons justement mis en relation le potentiel de visualisation sur Twit-ter avec ces facteurs externes et nous ne trouvons pas de rapport direct qui pourrait inviter à se passer d'une analyse spécifique de ce réseau socionumérique. Des médias qui devraient avoir des résultats assez proches (*Le Figaro*, *L'Équipe* et *Le Monde*), voient leurs résultats sur Twitter diverger.

Afin de mieux comprendre, nous avons calculé une *projection* de notre réseau biparti entre URL et usagers, à partir d'un échantillon aléatoire de 20 000 articles (Borgatti *et al.*, 2013). Ce procédé transforme le graphe *biparti* initial en un graphe *uniparti* où les sommets ne sont que des articles et où un arc entre deux sommets représente le fait que ces deux URL ont été partagées par un même usager. Le poids des arcs est alors déterminé par le nombre d'usagers qui ont partagé les couples d'articles. Cette nouvelle représentation nous permet de formuler des hypothèses sur les différents résultats des médias sur Twitter.

Nous remarquons par exemple que les articles de *L'équipe* occupent une région unique et périphérique du réseau : les lecteurs de ce quotidien ont tendance à ne partager que ses articles. *Le Monde* au contraire, est très central : ses articles sont connectés aux articles des autres médias par un grand nombre d'usagers. C'est vraisemblablement la raison qui donne au *Monde* un tel « pouvoir de reconfiguration » de l'offre d'information. Ses contenus sont réappropriés par des usagers hétérogènes, parmi lesquels un grand nombre de comptes très suivis, qui les partagent amplement, leur donnant une grande visibilité potentielle. *Le Figaro* n'arrive pas à réaliser cela : bien que moins isolé de *l'équipe*, ses contenus ont tendance à « faire cluster », partagés par des usagers assez fidèles et qui ont un régime médiatique plus spécialisé que les lecteurs du *Monde*. Cette « non-spécialisation » du *Monde*, qui n'a pas un effet évident sur la fréquentation de son site (en comparaison avec les autres

grands quotidiens) paie sur Twitter, et lui donne un rôle phare dans l'écosystème de l'information distribuée. Cette hypothèse d'explication est compatible avec l'importance des « weak ties » pour la diffusion de l'information sur les réseaux socionumériques (Bakshy *et al.*, 2012). *Le Monde* bénéficie de sa réputation de « journal de référence », ce qui lui donne la chance d'être tweeté et retweeté en priorité ou disons par confiance, par crédibilité supposée. La figure 7 montre la projection des articles : les articles du *Monde* sont en noir.

Figure 7. Projection des articles partagés



Source : auteurs.

CONCLUSION

Nous avons étudié la reconfiguration émergente de l'offre d'information opérée par la twittosphère française. À travers la construction d'un graphe d'articles et d'utilisateurs, nous avons observé la concentration de l'offre vers un nombre assez réduit de médias. On obtient donc un aperçu d'une des propriétés de ce qu'on peut appeler un « nouvel écosystème socio-numérique de l'information ».

Traditionnellement, l'écosystème médiatique d'information se compose d'échanges entre des journalistes qui prélèvent dans le monde vécu, des faits et des données, afin de les mettre en forme (les « in-former ») sur des supports de diffusion spécifiques (radio, presse écrite, télévision) régis par des logiques économiques de marché, car une transaction s'opère. Ils fournissent aux citoyens un bien public : des informations triées, vérifiées, hiérarchisées, sur ce qui se passe dans le monde afin de faire de chacun des citoyens infor-més donc éclairés. Ce modèle vole en éclat avec l'introduction des technologies numériques et de l'Internet. Les internautes sont devenus des producteurs d'information via la miniaturisation et la simplification des techniques de capture des faits (mini caméras, smartphones, etc.). Via leurs blogs, mais de plus en plus via les réseaux socio-numériques, ils assurent la diffusion de leurs autoproductions. Et par la veille qu'ils mettent en place et partagent sur leurs réseaux, en recommandant des contenus, les internautes deviennent les promoteurs des contenus d'information, que celle-ci soit créée par des journalistes ou par des amateurs. L'accès direct aux médias peut commencer à devenir une exception, en gommant l'idée de stratégie ordonnée dans sa quête d'information, au profit d'une circulation plus chaotique, d'une entrée sur un site de média par un clic sur son Facebook ou son Twitter.

Nos recherches conduisent au constat d'une situation paradoxale : la liberté de partage des internautes fait émerger une voix dominante. Mais il serait possible d'aller encore plus loin dans l'analyse de cette reconfiguration de l'information par les réseaux socio-numériques. Une analyse plus détaillée du réseau des partages pourrait mettre mieux en avant le rôle des communautés périphériques, à l'intérieur desquelles éventuellement identifier des effets de « chambre d'écho » capables de renforcer l'existence d'accès parallèles et non communicants à l'information. Car dans le fil d'actualité des infomédiaires règne souvent le mélange des genres, entre informations sérieuses et divertissement, entre faits recoupés et rumeurs. Des nouveaux régimes sémiotiques naissent donc, couplant *news* et *fake news*, vérité et post-vérité, terrain fertile

pour l'émergence de *faux authentiques* (Eco, 1985). Les chercheurs doivent prendre acte de ce changement et viser la compréhension des potentialités et des risques de l'information distribuée. On a l'impression d'assister à une plongée profonde dans l'*hyperréel* annoncé par Jean Baudrillard (1981), où les signes se détachent de plus en plus des structures institutionnelles de leur production. On ne peut savoir si cette « popularisation » de la construction consensuelle du réel, résultat d'une reconfiguration émergente (jamais vue) dont nous avons ici montré une image, libérera le discours des contraintes oligarchiques propres au système de l'information, en réduisant de ce fait la distance entre les mots et les choses, ou au contraire videra de sens les distinctions structurant la communication – vrai et faux, important et secondaire, pertinent et impertinent, uniforme et plural – en retournant finalement à la « masse amorphe » aurorale de la pensée sociale. Au risque, finalement, d'une réduction des interprétations du réel vers une seule autorité, effet si éloigné des souhaits des inventeurs d'Internet.

RÉFÉRENCES

- ACPM (2014a), « Classement OJD des Sites Internet & Hybrides Grand Public Septembre 2014 », accès : <http://www.acpm.fr/index.php/content/download/6658/43285/version/1/file/Classement%2Bde%2Bla%2Bfr%25C3%25A9quentation%2Bdes%2BSites%2BOJD%2Bsur%2Bseptembre%2B2014.pdf> (consulté le 28 juillet 2017).
- ACPM (2014b), « Classement OJD des Applications Septembre 2014 », accès : <http://www.acpm.fr/content/download/6659/43291/version/1/file/Classement%2Bde%2Bla%2Bfr%25C3%25A9quentation%2Bdes%2BApplications%2BOJD%2Bsur%2Bseptembre%2B2014.pdf> (consulté le 28 juillet 2017).
- BAKSHY E., ROSENN I., MARLOW C., ADAMIC L. (2012), « The Role of Social Networks in Information Diffusion », in *Proceedings of the 21st International Conference on World Wide Web*, ACM, New York.
- BAUDRILLARD J. (1981), *Simulacres et simulation*, Paris, Galilée.
- BORGATTI S.P., EVERETT M.G. (1997), « Network analysis of 2-mode data », *Social Networks*, n° 19, pp. 243-269.
- BORGATTI S.P., EVERETT M.G., JOHNSON J.F. (2013), *Analysing Social Networks*, London, Sage.
- CARDON D. (2015), *À quoi rêvent les algorithmes. Nos vies à l'heure des big data*, Paris, Seuil.
- CHARAUDEAU P. (1997), *Le discours d'information médiatique. La construction du miroir social*, Paris, Nathan.
- CHELGHOU M., HOANG B.-T., KACEM I. (2016), « A Learning-Based Model for Predicting Information in Social Networks », in *Proceedings of IEEE CODIT'16*.
- COMPAGNO D. (2014), « A Quali-Quantitative Narrative Analysis of the 2012 Fukushima Nuclear Accident in the French Media », *Actes des Journées internationales d'analyse statistique des données textuelles (JADT)*, Paris, INALCO, pp. 161-173.
- COMPAGNO D., MERCIER A. (2017), « Flux de tweets et articles d'archives : comprendre les logiques sociales d'un apparent paradoxe temporel », in V. Shafer (dir.), *Temps et temporalité du web*, Paris, Éditions Paris Ouest, à paraître.
- COMPAGNO D., PIGNARD-CHEYNEL N. (2017), « Robots, algorithmes et partages d'information sur les réseaux socionumériques », in A. Mercier et N. Pignard-Cheynel (dir.), *Partager et commenter l'info sur les réseaux sociaux*, Paris, Éditions de la Fondation MSH, à paraître.
- CSARDIG., NEPUSZ T. (2006), « The *igraph* software package for complex network research », *InterJournal Complex Systems*, 1695.
- ECO U. (1985), *La guerre du faux*, Paris, Grasset.

- ECO U. (1992), *Les limites de l'interprétation*, Paris, Grasset.
- FAUST K. (1997), « Centrality in affiliation networks », *Social Networks*, n° 19, pp. 157-191.
- GLOBAL WEBINDEX (2015), *GWI Social Summary Q1 2015*, mai.
- HARRIS INTERACTIVE (2015), *Social Life 2015*, février.
- HUANG Y. (2015), « A Semantic Network Analysis of Twitter Reaction to the Ferguson Grand Jury Decision: Implications for Framing and Deliberation », présentation à la *65th ICA (International Communication Association) Preconference Bringing together Social and Semantic Networks in Communication Research*, San Juan, 10-15 mai.
- LATAPY M., MAGNINE C., DEL VECCHIO N. (2008), « Basic notions for the analysis of large two-mode networks », *Social Networks*, n° 30, pp. 31-48.
- MERCIER A., PIGNARD-CHEYNEL N. OUKRAT A. (2017), « Facebook pour s'informer ? Actualité et usages de la plateforme par les jeunes », in A. Mercier, N. Pignard-Cheynel (dir.), *Partager et commenter l'info sur les réseaux sociaux*, Paris, Éditions de la Fondation MSH, à paraître.
- MERCIER A., PIGNARD-CHEYNEL N. (dir.) (2017), *Partager et commenter l'info sur les réseaux sociaux*, Paris, Éditions de la Fondation MSH, à paraître.
- NIELSEN (2015), *Twitter Consumer Deep Dive Survey*, juillet.
- R CORE TEAM (2015), « R: A language and environment for statistical computing », R Foundation for Statistical Computing, Vienne, Autriche.
- REBILLARD F. (2011), « L'étude des médias est-elle soluble dans l'informatique et la physique ? », *Questions de communication*, n° 20.
- REBILLARD F., FACKLER D., MARTY E. (2012), « L'offre d'informations est-elle plus diversifiée sur le web qu'à la télévision ? Une comparaison exploratoire entre sites d'actualité et journaux télévisés », *Réseaux*, n° 176, pp. 143-173.
- REBILLARD F., LOICQ M. (2013), *Pluralisme de l'information et media diversity*, Bruxelles, De Boeck.
- RICŒUR P. (1983), *Temps et récit*, tome 3, Paris, Seuil.
- MCCOMBS M.E., SHAW D.L. (1972), « The Agenda Setting Function of Mass Media », *The Public Opinion Quarterly*, vol. 36, n° 2, pp. 176-187.
- SIMON J., TOULLEC B., BADOUARD R., BIGEY M., COMPAGNO D., MERCIER A., PIGNARD-CHEYNEL N., SEBBAH B. (2017), « Identifier et analyser les discours d'escorte sur Twitter. L'influence des discours d'accompagnement sur le partage social », in C.R. Wigham, G. Ledegen (dir.), *Corpus de communication médiée par les réseaux : construction, structuration, analyse*, Paris, L'Harmattan.
- SMYRNAIOS N., REBILLARD F. (2011), « Entre coopération et concurrence : les relations entre infomédiaires et éditeurs de contenus d'actualité », *Concurrences*, n° 3.

SMYRNAIOS N., RATINAUD P. (2013), « Comment articuler analyse de réseaux et des discours sur Twitter », *Tic et société*, vol. 7, n° 2.

TAKHTEYEV Y., GRUZD A., WELLMAN B. (2012), « Geography of Twitter networks », *Social Networks*, n° 34, pp. 73-81.

WICKHAM H., GROLEMUND G. (2017), *R for Data Science*, Sebastopol (CA), O'Reilly.

WHITE D.M. (1950), « The *gate keeper*: A case study in the selection of news », *Journalism Quarterly*, n° 27, pp. 383-391.