

# Des journalistes enfin libérés de leurs sources ?

## Promesse et réalité du « journalisme de données »<sup>1</sup>

*« Les journalistes ont besoin d'être familiers avec les données. Avant, vous dénchiez des idées d'article en discutant avec des gens dans les bars, et vous le ferez peut-être encore de temps à autre. (...) Mais maintenant, vous allez aussi devoir vous pencher sur les données et vous équiper d'outils pour les analyser et y trouver ce qui est intéressant. Et vous devrez les mettre en perspective, aider les gens à vraiment comprendre la signification de tout ceci, à comprendre ce qui se passe dans le pays. Tim Berners-Lee »*

(Arthur, 2010)

**SYLVAIN PARASIE**  
Maître de conférences  
Université Paris-Est Marne-la-Vallée  
LATTS  
sylvain.parasie@univ-paris-est.fr

**ÉRIC DAGIRAL**  
Maître de conférences  
Université Paris Descartes  
CERLIS  
eric.dagiral@parisdescartes.fr



Depuis le milieu des années 1990, l'expérimentation de nouvelles formes de journalisme en ligne a pris appui sur les critiques qui ont longtemps été exprimées à l'égard de la presse et des médias traditionnels. Plusieurs chercheurs ont ainsi analysé les liens entre l'essor du journalisme participatif et la critique de l'asymétrie entre producteurs et récepteurs de l'information (Domingo et al., 2008 ; Aubert, 2009 ; Rebillard, 2011). Mais on a moins souvent identifié le rôle qu'a joué la critique des rapports entre les journalistes et leurs sources dans la conception de nouveaux produits journalistiques. La dépendance et l'asymétrie entre les journalistes et leurs sources, institutionnelles notamment, ont ainsi constitué des griefs récurrents dans l'esprit de ceux qui ont imaginé de nouvelles manières de faire du journalisme grâce aux technologies en ligne.

### **Pour citer cet article**

#### Référence électronique

Sylvain Parasie, Éric Dagiral, « Des journalistes enfin libérés de leurs sources ? Promesse et réalité du « journalisme de données », *Sur le journalisme, About journalism, Sobre jornalismo* [En ligne], Vol 2, n° 1 - 2013, mis en ligne le 15 avril 2013.  
URL : <http://surlejournalisme.com/rev>

Ce fut le cas du mouvement altermondialiste qui est apparu à partir de 1999 en développant une forte critique des médias. Il a donné naissance à de nombreux sites indépendants qui ont voulu produire une information libérée des formats médiatiques traditionnels (Cardon et Granjon, 2003). Il s'agissait notamment de contrer la partialité des journalistes à l'égard de certaines sources et de contester les formes de distanciation auxquels ils ont recours (Estienne, 2010). Dans un genre très différent, ce fut aussi le cas de *Wikileaks*, qui défraya la chronique en 2010 en publiant des dizaines de milliers de documents secrets sur les guerres en Afghanistan et en Irak. À l'origine de ce site fondé par des programmeurs informatiques, on trouve l'idée selon laquelle les journalistes traditionnels ne peuvent garantir au lanceur d'alerte que son identité ne sera pas révélée et que ses informations seront réellement diffusées. Quelles que soient les pressions politiques et commerciales exercées sur les journalistes, les technologies utilisées par les concepteurs de *Wikileaks* étaient supposées garantir l'anonymat de la source et la diffusion des documents secrets (Domscheit-Berg, 2011 : 196-211).

Depuis le milieu des années 2000, de nouveaux liens se sont noués entre les mondes informatiques et les mondes journalistiques sur l'autel d'une volonté de renouveler les rapports entre les journalistes et leurs sources. D'abord aux États-Unis et en Grande-Bretagne, des programmeurs ont soit intégré les rédactions de grands journaux (*New York Times*, *Washington Post*, *Chicago Tribune*, *Guardian*, etc.) ou d'organisations indépendantes produisant de l'information (*ProPublica.org*), soit encore créé des start-ups à vocation journalistique (*EveryBlock.com*). Au-delà de la variété de leurs trajectoires et leurs compétences, ces individus ont en commun de concevoir des programmes informatiques qui sont présentés comme constituant intrinsèquement des produits journalistiques. Qu'ils s'appellent « data-journalistes », « journalistes-programmeurs » ou « journalistes hackers », ils revendiquent un positionnement à la fois technique et éditorial (Parasie, 2011). Souvent désignés sous le terme de « *news applications* », les contenus qu'ils conçoivent sont aussi bien des cartes interactives que des infographies, des bases de données interrogeables ou d'autres formes de présentation en ligne. Leur fabrication repose sur une variété de compétences informatiques permettant de collecter, de traiter, de combiner et de visualiser des données – qu'il s'agisse de chiffres, de textes, de photographies ou de contenus audiovisuels disponibles sur des supports numériques. Le terme qui semble s'imposer aujourd'hui pour désigner cet ensemble de pratiques hétérogènes est celui de « journalisme de données » (*data journalism*) (Gray et al, 2012), même si le traitement informatique de

données a profondément modifié la connaissance journalistique nord-américaine depuis les années 1960, soit bien avant l'essor du Web (Parasie et Dagiral, 2012a).

Mais l'idée selon laquelle un journalisme reposant sur le traitement de données serait porteur d'une transformation des relations entre les journalistes et leurs sources a de quoi surprendre. La grande majorité de ces projets s'appuient en effet sur des données qui sont enregistrées, stockées et diffusées par les pouvoirs publics eux-mêmes ou par des organisations qui leur sont proches : les crimes commis dans une ville ; l'activité des hôpitaux d'une région ; les transformations du réseau routier d'une métropole, etc. Dès lors, comment ce type de journalisme n'est-il pas susceptible de renforcer au contraire la dépendance des journalistes aux sources institutionnelles ? Dans cet article, nous montrons que ces pratiques journalistiques s'identifient à une promesse formulée au croisement des mondes informatiques et journalistiques – affaiblir la dépendance de l'information à l'égard des relations asymétriques entre les journalistes et leurs sources – qui n'est qu'en partie satisfaite dans les réalisations actuelles.

Une telle question pose néanmoins plusieurs difficultés. En premier lieu, nous avons aujourd'hui à notre disposition plusieurs cadres théoriques expliquant l'influence que les sources exercent sur les journalistes. Mais lorsque les programmeurs-journalistes conçoivent ces nouveaux contenus, ils le font en partie à partir d'un ensemble de jugements portant précisément sur la manière dont ils pensent être influencés par les sources. C'est la raison pour laquelle nous avons privilégié une perspective pragmatique – au sens où il s'agit d'intégrer l'analyse des prétentions des acteurs à la justice à l'étude de ces pratiques émergentes (Lemieux, 2000). La seconde difficulté tient aux artefacts socio-techniques que ces programmeurs-journalistes manipulent et conçoivent – des applications interactives, des bases de données, des scripts informatiques, etc. – et qui ne sont pas familiers des études sur le journalisme. Nous avons donc opté pour un cadre théorique inspiré de la théorie de l'acteur-réseau, attentif à analyser de quelle manière des acteurs proposent des définitions portant sur les usages d'une nouvelle technologie (Akrich et Latour, 1992 ; Weiss et Domingo, 2010).

Cet article s'appuie sur une enquête réalisée à l'automne 2010 à Chicago, à l'occasion de laquelle nous avons conduit des entretiens avec une quinzaine de personnes. Premièrement, nous avons interrogé sept journalistes du *Chicago Tribune* : deux « programmeurs-journalistes » rattachés à la « *news applications team* » ; quatre journalistes d'investigation qui travaillent quotidiennement avec

des bases de données ; un ancien vice-président du journal en charge de la stratégie et du développement (entre 2000 et 2008). Deuxièmement, nous avons interrogé l'un des fondateurs de *EveryBlock.com*, un site web d'information locale reposant intégralement sur le traitement automatique de données. Troisièmement, nous avons interrogé cinq personnes profondément impliquées dans le militantisme pour l'ouverture des données publiques à Chicago. Quatrièmement, nous avons conduit deux autres entretiens avec des enseignants de l'école de journalisme de Medill, très impliqués à établir des liens entre les mondes informatiques et celui de la presse à Chicago. Enfin, nous avons constitué des corpus composés des différentes applications en ligne conçues dans le cadre de ces différentes organisations de presse.

Dans une première partie, nous montrons que les formes actuelles du journalisme de données reposent sur la promesse d'une transformation profonde des rapports entre les journalistes et leurs sources grâce aux technologies informatiques. Dans une seconde partie, nous montrons à partir du cas de l'information par les données criminelles à Chicago, que cette promesse n'est qu'en partie satisfaite.

---

#### DES DONNÉES POUR REPRENDRE L'INITIATIVE FACE AUX SOURCES

---

Au-delà des divergences théoriques entre les chercheurs, il existe un consensus pour considérer que les relations asymétriques entre les journalistes et leurs sources d'information jouent un rôle important dans la sélection des sujets traités et les modalités de ce traitement (Schlesinger, 1992). Les sources qui sont les plus proches du pouvoir politique et économique seraient ainsi structurellement en position d'imposer plus souvent leurs stratégies aux organisations médiatiques. Or c'est en partie autour de la promesse de contrer cette asymétrie entre les journalistes et leurs sources que se sont établis de nouveaux liens entre les mondes journalistiques et informatiques depuis le milieu des années 2000.

#### Des journalistes dépendants des stratégies de leurs sources

Même si les travaux portant sur les sources des médias sont relativement peu nombreux parmi l'ensemble des recherches sur le journalisme (Schlesinger, 1992), plusieurs travaux ont montré à quel point la pratique journalistique favorise généralement les intérêts des sources faisant autorité, c'est-à-dire des acteurs qui représentent les principales institutions politiques et sociales.

Dans un article fondateur, Harvey Molotch et Marilyn Lester expliquent que la grande majorité des nouvelles publiées dans les journaux s'identifient à ce qu'ils appellent des « événements de routine », autrement dit des événements qui sont promus par les acteurs qui en sont à l'origine (Molotch et Lester, 1974). La plupart du temps, les promoteurs de ces événements – élus, fonctionnaires, responsables institutionnels, etc. – jouissent d'un accès régulier et habituel dans les médias. Selon les auteurs, cet accès privilégié s'explique par le fait qu'il va de soi pour les journalistes qu'un responsable politique de tout premier plan ne peut dire que des choses importantes – même si ces acteurs doivent faire attention à la concurrence qu'ils subissent nécessairement du point de vue de l'attention publique. Les acteurs dotés d'un pouvoir important ont souvent davantage de chance de susciter l'attention des journalistes, en usant de méthodes brutales (pressions, menaces de boycott, etc.) ou plus subtiles (en encourageant les routines journalistiques qui paralysent l'investigation).

La capacité plus ou moins grande des sources institutionnelles à imposer leurs propres définitions des événements a fait l'objet de discussions au sein des sciences sociales (Schlesinger, 1992). Ainsi, selon Stuart Hall et ses collègues, les journalistes tendent structurellement à reprendre les définitions de la réalité sociale que leur fournissent les porte-parole de l'institution (Hall et al, 1978). Selon eux, c'est parce que les journalistes valorisent la plus grande neutralité professionnelle qu'ils ouvrent leur porte de façon préférentielle aux acteurs qui disposent d'un large pouvoir institutionnel. Cette préférence structurelle conduit alors à ce que les porte-parole de l'institution deviennent les « premiers définisseurs » (*primary definers*) et parviennent à imposer leurs propres interprétations aux journalistes. Mais d'autres analyses ont mis en avant l'existence de luttes entre les différentes sources officielles, nuancant ainsi la capacité de ces acteurs à imposer systématiquement leurs propres définitions. Par exemple, dans une perspective bourdieusienne, Dominique Marchetti explique que la plupart des révélations journalistiques à prétention anti-institutionnelle faites par le journalisme dit d'investigation depuis les années 1970 sont directement dues aux luttes internes aux mondes politique et judiciaire (Marchetti, 2000).

Mais c'est plus souvent par les routines du travail journalistique que les chercheurs rendent compte de l'asymétrie des relations entre les médias et leurs sources. La nécessité de se plier à des formats de production et de diffusion particulièrement contraignants expliquerait leur tendance, dans certaines situations, à valider le traitement qui leur est proposé par les sources (Lemieux, 2000). Par exemple, si les journalistes ont tendance à privilégier les sources ins-

titutionnelles, c'est que celles-ci permettent souvent de réduire considérablement les risques juridiques associés à la diffusion d'une information (Lemieux, 2004).

Notre objectif n'est pas ici de rouvrir ce débat théorique, mais d'abord d'expliquer de quelle façon les nouvelles formes de journalisme appuyé sur les données sont apparues aux yeux de certains journalistes nord-américains comme un moyen pour eux de reprendre l'initiative face aux sources, institutionnelles notamment.

---

#### DE NOUVEAUX LIENS ENTRE LES MONDES JOURNALISTIQUES ET INFORMATIQUES

---

Le renouvellement des formes du journalisme de données à partir du milieu des années 2000 s'explique par deux mouvements distincts. D'abord, plusieurs grands journaux américains embauchent des journalistes pour concevoir des « applications web » à valeur journalistique. Il s'agit d'un ensemble varié de contenus interactifs publiés en ligne : infographies, bases de données, cartes, etc. Ces journalistes les plus connus sont Adrian Holovaty (*Washington Post*), Aron Pilhofer (*New York Times*), Matt Waite (*St-Petersburg Times*) ou Ben Welsh (*Los Angeles Times*). Disposant eux-mêmes de larges compétences dans les différents langages de programmation (Python, Ruby, PHP, etc.), ils mettent en avant l'idée selon laquelle les organisations de presse doivent mettre en place des équipes constituées d'individus qui s'y connaissent autant en journalisme qu'en programmation (Royal, 2010). Pour la plupart, ces individus ont été formés au traitement de données au sein même du monde journalistique américain, au sein de la tradition du « computer-assisted reporting » (Parasie et Dagiral, 2012a). Mais ils sont également investis dans les communautés du logiciel libre. C'est tout particulièrement le cas d'Adrian Holovaty, qui est à la fois diplômé d'une école de journalisme et l'un des créateurs de Django – une plateforme de programmation web largement reconnue dans les communautés du logiciel libre.

Le second mouvement est largement extérieur au monde de la presse. Il s'inscrit dans la mobilisation pour l'ouverture des données publiques qui commence vers 2005 dans les grandes métropoles américaines<sup>2</sup> avant de s'étendre au niveau fédéral après l'élection de Barack Obama en 2009. Un assemblage hétérogène d'entrepreneurs du Web, de militants politiques, de programmeurs informatiques, de passionnés d'informatique et de journalistes joignent leurs forces pour réclamer la libération des données produites par les pouvoirs publics. Partageant les

principes de liberté de l'information défendus par le mouvement pour le logiciel libre (Holtgrewe et Werle, 2001), ils investissent les technologies du Web pour concevoir un grand nombre de projets dont ils attendent qu'ils augmentent la transparence des gouvernements et la participation des citoyens (Lathrop et Ruma, 2010). Tous ces projets reposent sur le traitement, l'analyse et la visualisation de données ayant trait à une grande variété de domaines (criminalité, démographie, transports, infrastructures, dépenses publiques, environnement, emploi, etc.) ainsi qu'aux responsables politiques (financement des campagnes, dépenses réalisées, votes, décisions publiques, etc.) à tous les niveaux de gouvernement (villes, États, gouvernement fédéral). Pour ne donner qu'un exemple parmi de nombreuses réalisations, citons le cas de MapLight.org, un site web lancé en 2007 par Daniel Newman – qui est à la fois programmeur, entrepreneur et militant politique – et qui permet de croiser les votes des membres du Congrès avec les financements qu'ils ont reçus des groupes d'intérêt (Lathrop et Ruma, 2010 : 223-232).

C'est ainsi qu'à partir du milieu des années 2000, plusieurs arènes se sont progressivement constituées dans les grandes métropoles américaines, autour du militantisme pour l'ouverture des données publiques. S'y sont notamment établies de nouvelles relations entre les mondes informatiques et journalistiques, lesquelles ont favorisé l'approfondissement de la figure du « journaliste-programmeur ». À Chicago, nous avons pu observer l'existence de réseaux informels réunissant des programmeurs travaillant pour des entreprises privées (finance, marketing, etc.), des relations avec des acteurs impliqués dans les communautés du logiciel libre, ainsi que des militants, des journalistes et des spécialistes de la gestion urbaine.

#### Corriger l'inégale distribution sociale et spatiale des sources

Devant le volume croissant de données disponibles à l'échelle des métropoles américaines, plusieurs journalistes ont considéré qu'il s'agissait là d'une opportunité unique pour couvrir plus en profondeur l'actualité du territoire urbain (Parasie et Dagiral, 2012b). Dès lors qu'il peut accéder à toutes les données relatives aux crimes et délits (vols, homicides, cambriolages), à la gestion de l'infrastructure urbaine (attribution des permis de construire, réparations effectuées sur les canalisations, signalisation urbaine), à l'économie de la ville (attributions des permis de vente d'alcool) et à l'ensemble des services offerts par celle-ci (évaluation des écoles ou des hôpitaux), le journaliste disposerait en quelque sorte de nouvelles « sources » également réparties sur l'ensemble du territoire métropolitain. Habituellement, les reporters tendent à couvrir davantage l'actualité

de certains espaces urbains — le centre-ville où se trouvent les principales institutions — du fait qu'ils bénéficient d'un accès routinier aux sources qui sont attachées à ces espaces. En conséquence, l'actualité des autres quartiers n'est pas couverte avec la même intensité parce qu'il n'existe pas de sources établies susceptibles de donner des informations au journaliste. Selon ces journalistes-programmeurs, le traitement de données permettrait de corriger cette inégale distribution des sources sur le territoire urbain en faisant apparaître de nouvelles sources plus également réparties sur ce territoire.

Aux États-Unis, Adrian Holovaty a particulièrement incarné cette posture. Jouissant déjà d'une certaine réputation dans le monde du logiciel open source, ce journaliste de 24 ans connaît une grande popularité auprès des journalistes et programmeurs de Chicago lorsqu'il lance en 2005 le site *Chicago-Crimes.org*. Combinant les données diffusées sur le site du *Chicago Police Department* avec une carte Google, le site permet de prendre connaissance de l'ensemble des crimes qui ont été constatés à l'échelle d'un quartier. En 2007, Holovaty et O'Neil forment une petite équipe de programmeurs afin de lancer le site *EveryBlock.com*, qui étend le principe de *Chicagocrime.org* à de nombreux jeux de données provenant des services de la ville de Chicago. Désormais, le site permet aux habitants d'une dizaine de grandes villes américaines de prendre connaissance d'une grande variété d'informations à l'échelle de leur quartier (crimes, inspections des restaurants, permis de construire, etc.). Selon Holovaty, ce site permet de corriger la profonde inégalité territoriale de la couverture journalistique qu'assurent les organisations de la presse locale. Il applique en quelque sorte le principe de la « longue traîne » à l'information locale par l'identification de nouvelles sources d'information jusque-là ignorées (Anderson, 2006).

Dès lors que les données fournies par la ville sont granulaires — c'est-à-dire disponibles à l'échelle la plus locale possible —, complètes et régulièrement mises à jour, une organisation de presse a tout intérêt, selon Holovaty, à exploiter ces données. Cela signifie en effet que les arrangements d'humains et de non humains qui produisent ces données deviennent de véritables sources d'information. Ainsi, lorsqu'un policier constate un incident sur la voirie et l'enregistre dans son ordinateur de bord, ou lorsqu'un employé de la mairie enregistre la demande de construction d'un immeuble, ils sont constitués comme des sources d'autant plus intéressantes qu'elles étaient jusque-là souvent inaccessibles pour le journaliste. Ces multiples arrangements couvrent l'ensemble du territoire urbain, pour un coût négligeable en comparaison de l'ensemble des moyens qu'une organisation de presse serait dans l'obligation de déployer pour

produire les mêmes informations. C'est en quelque sorte le caractère localement très situé des sources plus traditionnelles de la presse métropolitaine — liées aux institutions du centre-ville de la métropole — qui s'en trouve ainsi corrigé.

### Contre les stratégies des sources

Mais aux yeux des journalistes-programmeurs que nous avons interrogés, l'ouverture des données publiques est porteuse d'une autre promesse plus grande encore. Le fait de concevoir des produits journalistiques à partir de données doit permettre non seulement de transformer en « sources » des policiers, des inspecteurs ou des employés des services publics transformant des incidents en « données », mais cela doit aussi permettre au journaliste de révéler des réalités que les sources gouvernementales préféreraient le plus souvent cacher. Ainsi, en dépit du fait que les données soient enregistrées, stockées et diffusées par les pouvoirs publics, leur traitement informatique permettrait aux journalistes de contrer les stratégies de ces sources institutionnelles.

Dans le cadre de la mobilisation pour l'ouverture des données publiques, de nombreuses applications web ont été développées dans le but d'accroître la transparence des gouvernants. Quoique discuté (Swartz, 2010 ; Birchall, 2011), le postulat de ces démarches est que l'enregistrement continu d'un ensemble d'actions opérées par des gouvernants — et plus largement par toutes les personnes qui ont des responsabilités — dans leurs activités quotidiennes permet d'accéder à une « vérité » qui diffère des justifications que les responsables sont susceptibles de produire publiquement. Voici de quelle manière Daniel X. O'Neil, cofondateur du site *EveryBlock.com*, présente la « vérité » exprimée par les données :

*« Je crois beaucoup à la vérité [truth] et à la responsabilisation [accountability]. Si quelque-chose est vrai, c'est juste vrai et ça ne pose pas de problème. Cela ne signifie rien d'autre que la vérité. (...) Je vais vous donner un exemple. Vous savez qu'aux États-Unis, il y a depuis deux ou trois ans cette mode des cupcakes. Et il y avait cette boutique appelée « Bleeding heart bakery » dans le North Side. C'était une pâtisserie végétarienne, elle était très cool : le chef avait des tatouages et passait à la télévision. Hé bien, quelle qu'en soit la raison, un jour l'eau a été coupée. Et un contrôleur est précisément venu ce jour-là, et ils ont dû fermer pour violation du règlement en matière d'hygiène. Le chef est alors entré dans une colère noire, expliquant que c'était une coïncidence, qu'on ne comprenait pas à quel point ils étaient cool, etc. « Clic, vous avez*

*échoué au contrôle d'hygiène ce jour-là ». C'est la vérité, et il se peut qu'il y ait de bonnes raisons qui expliquent cela ou que ce soit la faute de quelqu'un d'autre, mais la vérité c'est que vous avez échoué au contrôle d'hygiène et vous devez faire avec la vérité<sup>3</sup> »*

Selon cette explication, la vérité exprimée par la donnée tiendrait principalement à la séparation radicale entre ce qui relève, d'une part de la qualification d'une situation, de l'enregistrement et de la diffusion en ligne de cette qualification, et ce qui concerne d'autre part l'ensemble des jugements d'imputation causale ou d'évaluation morale de la situation. Dès lors que l'enregistrement des données repose sur une règle claire, celles-ci expriment une vérité qui échappe à la source considérée — ici, non seulement la pâtisserie, mais aussi les services d'hygiène de la ville. Lorsqu'on qualifie une situation selon des règles générales, et qu'on enregistre ensuite automatiquement cette qualification pour ensuite la diffuser, on fait en sorte de pouvoir mettre les acteurs devant leurs propres responsabilités : ils doivent faire quelque chose pour corriger la situation rendue ainsi visible.

Si tous les programmeurs-journalistes ne partagent pas cette perspective objectiviste répandue dans le mouvement pour l'ouverture des données publiques, tous pensent néanmoins que les données ont une valeur en dehors de l'organisation qui les a produites, et qu'elles peuvent donc nourrir des projets journalistiques révélant des réalités que les sources institutionnelles souhaiteraient cacher. Tout particulièrement, la combinaison de données provenant de plusieurs sources différentes est considérée comme devant permettre de mettre au jour des réalités que certaines de ces sources ignorent ou voudraient cacher. Aux États-Unis, il est ainsi relativement courant de voir des journalistes combiner des données provenant d'institutions locales, nationales et fédérales afin de pointer les dysfonctionnements d'une administration particulière. La mobilisation des technologies informatiques est ainsi perçue comme permettant de déconnecter largement les données des intentions ou des stratégies qui animent les sources institutionnelles pourtant productrices de ces données.

### **Des robots pour résister à l'influence des sources**

Les historiens Lorraine Daston et Peter Gallison ont montré qu'à la fin du 19<sup>e</sup> siècle, une nouvelle conception de l'objectivité était apparue dans les sciences naturelles (Daston et Gallison, 2012). Selon cette objectivité qu'ils appellent « mécanique », il s'agissait alors de concevoir des machines qui rejetteraient certains traits de la personnalité des chercheurs (le dogmatisme) tout en valorisant d'autres éléments

de la personnalité du scientifique (la rigueur). Or, tout se passe aujourd'hui comme si certains journalistes cherchaient à concevoir des machines rejetant certains traits du journaliste — sa prédisposition à être influencé par les sources, par exemple — tout en valorisant d'autres éléments de sa personnalité — le désir de révéler ce que les sources ne veulent rendre public. Le témoignage de Ben Welsh, journaliste au *Los Angeles Times*, est particulièrement révélateur de cette forme d'objectivité :

*« Chaque jour, vers 2 h 30 du matin, je reçois automatiquement un email du Los Angeles Police Department. Dans cet email, il y a un fichier csv qui comprend l'ensemble des personnes qui ont été arrêtées le jour d'avant. Donc j'ai un flux de données simple, répétitif et constamment mis à jour qui arrive tous les jours dans ma boîte aux lettres. J'ai écrit un script informatique qui cherche automatiquement dans ma boîte aux lettres, récupère ce fichier, en sort les données et alimente une base de données. Et c'est là que ça devient intéressant : je peux rédiger des scripts informatiques de façon à répondre automatiquement aux questions qu'un reporter se pose habituellement (...). Par exemple, j'ai écrit un script qui me permet de savoir automatiquement quels sont les motifs d'arrestation les plus graves, en sélectionnant les cautions les plus élevées. (...) Et je peux aussi automatiquement savoir quels sont les prêtres, les officiels, les enseignants qui ont été arrêtés (...). L'année dernière, un fan des Giants a été battu. La police a d'abord arrêté le mauvais gars, et ils ont fait un battage médiatique sans précédent. Mais ils ont totalement raté leur coup. Et quand ils ont finalement réussi à arrêter le bon gars, ils voulaient tenir ça secret. Mais moi, j'avais automatiquement reçu un email qui me donnait l'identité du gars en question, et nos reporters ont pu immédiatement aller interroger ses voisins, pour savoir de qui il s'agissait.<sup>4</sup> »*

Prenant appui sur une objectivité mécanique, des dispositifs informatiques sont ainsi mobilisés au sein de certaines rédactions nord-américaines afin d'étendre l'exercice de certains aspects de la démarche journalistique — chercher à identifier une réalité derrière les explications officielles — en réduisant les coûts pour l'organisation de presse. De ce fait, les sources seraient moins en mesure d'influencer le journaliste par le biais des contraintes qui enserrrent son activité quotidienne (la pression du temps, les impératifs de production, etc.).

On mesure donc à quel point ces journalistes-programmeurs mobilisent les technologies associées au

traitement de données afin de réduire la dépendance des journalistes à l'égard de leurs sources. Cela s'incarne dans plusieurs promesses : (1) accroître le nombre de sources disponibles ; (2) déconnecter l'information des stratégies des sources ; (3) réduire les coûts associés à l'enquête. Examinons maintenant comment ces promesses sont traduites en pratique, à partir du cas de l'information criminelle à Chicago.

## LE CAS DE L'INFORMATION CRIMINELLE À CHICAGO

À Chicago comme dans la plupart des grandes métropoles américaines, l'information criminelle prend une importance considérable. Pour les pouvoirs publics locaux et la mairie en place, le nombre d'homicides enregistré annuellement constitue un indicateur particulièrement surveillé<sup>5</sup>. À Chicago, le nombre d'homicide a même été au cœur du volontarisme municipal des années 1990, et il reste aujourd'hui une préoccupation centrale de la part d'une partie du monde associatif local, du grand public et de la quasi totalité des médias, qui lui consacrent une attention considérable. D'une manière générale, l'information criminelle a ceci de particulier qu'elle dépend plus étroitement des sources officielles. Stuart Hall expliquait d'ailleurs que ce domaine est celui dans lequel les sources officielles parviennent le plus facilement à imposer leurs propres interprétations, à la fois parce que les institutions disposent d'un monopole de fait sur la production des informations criminelles, et parce que ces informations sont structurellement moins ouvertes aux interprétations élaborées par les acteurs extérieurs au pouvoir (Hall, 1978 : 69-70). Dès lors que le *Chicago Police Department* dispose d'un monopole sur la production de l'information criminelle, la mobilisation des technologies informatiques a-t-elle permis à ces programmeurs-journalistes de se libérer de leur dépendance à cette source ?

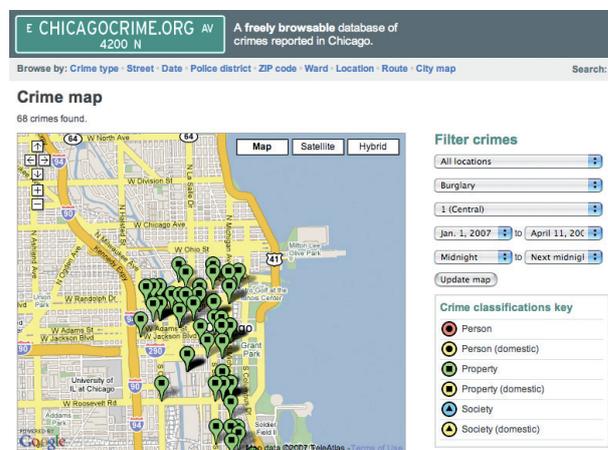
### La fin du monopole de la police sur la diffusion des données

Le premier point marquant est que la police de Chicago a progressivement perdu le monopole qu'elle exerçait sur la diffusion des données criminelles. Cela résulte d'abord de plusieurs initiatives mises en place par la mairie de Chicago dans la première moitié des années 1990. À cette époque, les services de police des grandes métropoles américaines utilisent de plus en plus l'informatique pour produire des cartes criminelles à l'échelle des villes (Chamard, 2006). La ville de Chicago a été pionnière dans ce domaine, alors que les autorités mettent en place un nouveau modèle de police fondé sur la proximité avec les habitants<sup>6</sup>. En 1995, elle équipe les policiers de la ville d'un système leur permettant d'élaborer leurs

propres cartes des crimes, en sélectionnant les zones ou les types de crimes concernés. Les cartes générées sont ainsi directement articulées aux systèmes d'information de la police de Chicago, qui reposent sur les données saisies par les agents depuis leur voiture – sur un PC équipé du logiciel MapInfo –, sur les lieux d'enregistrement des crimes<sup>7</sup>. Dès l'année suivante, la police permet à n'importe quel habitant de Chicago de prendre connaissance, sur un site web dédié<sup>8</sup>, des crimes qui ont été commis. L'accès aux données demeure alors néanmoins limité : uniquement disponibles pour les deux dernières semaines, elles ne peuvent pas être téléchargées ni agrégées selon les souhaits de l'utilisateur, et encore moins faire l'objet de tris ou de croisements en dehors d'une courte période donnée.

Au milieu des années 2000, l'essor du mouvement pour l'ouverture des données publiques met définitivement fin au monopole de la police sur la circulation des données criminelles dans les grandes métropoles américaines. À Chicago, le journaliste et programmeur Adrian Holovaty joue un rôle majeur en élaborant le site *ChicagoCrimes.org* en mai 2005. Combinant les données du site de la police de Chicago avec une carte Google – dont la technologie n'est alors pas encore ouverte –, Holovaty est alors largement salué – dans les mondes informatiques et au-delà – à la fois pour sa virtuosité et pour permettre aux habitants de prendre connaissance des crimes individuels avec leur localisation précise, en rendant possible un ensemble de tris variés. La source des données demeure le *Chicago Police Department* – qui conserve un monopole absolu sur la production des données –, mais les données sont alors beaucoup plus accessibles, triables et réutilisables.

Figure 1 : ChicagoCrimes.org (2005)



À Chicago comme dans les autres grandes métropoles nord-américaines, le déclin du monopole policier sur la diffusion des données a constitué une

opportunité pour les organisations de presse. Dès 1996, le site du *Chicago Tribune* publie une carte interactive reposant sur les données criminelles diffusées par le *Chicago Police Department*<sup>9</sup>, mais le mouvement s'approfondit au milieu des années 2000, alors que de nombreuses organisations de presse diffusent, sur leurs sites web, des bases de données policières indiquant tous les crimes commis dans une zone. Aux yeux des journalistes, l'accès aux données permet surtout d'offrir une autre représentation de la criminalité à l'échelle de la métropole. D'une part, il est désormais possible d'offrir une couverture des crimes qui soit plus complète : le traitement de données permet de couvrir tous les crimes, et non plus seulement ceux que les autorités voudront faire connaître ou ceux qui présentent une plus grande valeur journalistique<sup>10</sup>. D'autre part, il s'agit pour les journalistes de couvrir les crimes d'une façon plus locale : le lecteur du journal pourra connaître les détails des crimes associés à une adresse précise, et non plus seulement les taux de crime associés à un quartier plus vaste.

Dans une certaine mesure, il est possible d'affirmer que le déclin du monopole de la police de Chicago sur la diffusion des données a permis aux journalistes de la ville de gagner en liberté à l'égard de la source policière. Ceux-ci sont désormais en mesure de couvrir un nombre plus important de crimes en dépendant moins du choix de la police de diffuser les informations relatives à tel ou tel crime.

### **Contester le monopole policier sur la production des données**

La police de Chicago conserve toutefois un monopole presque total sur la production des données criminelles<sup>11</sup>, et on pourrait craindre que l'intérêt croissant des journalistes pour ces données ne se traduise par une dépendance accrue à la source policière. Un débat a d'ailleurs émergé dans une partie du monde journalistique nord-américain au sujet de l'utilisation des données provenant de la police (Welsh, 2009) : était-ce raisonnable, pour les journalistes, d'exploiter ces données sans chercher à les vérifier ? Parmi les journalistes que nous avons interrogés, Darnell Little — qui officia au *Chicago Tribune* jusqu'en 2009 — critique l'intervention de la police dans la production des données, et la dépendance des journalistes à l'égard de la police :

*« Le département de police revient toujours en arrière et modifie la catégorisation de ces meurtres qui deviennent autre chose. C'est la raison pour laquelle on a vraiment besoin d'un meilleur état des choses, surtout au niveau local, avec la création d'une sorte d'agence gouvernementale. L'accès*

*du public aux données criminelles est l'une des choses les plus importantes, parce que nous sommes en quelque sorte pris en otage devant la moindre information qu'ils nous donnent. Ils diront que la criminalité diminue d'autant — et on ne peut pas vérifier les données ou la méthodologie utilisée, puisque c'est eux qui collectent l'information, qui détiennent l'information, et que nous dépendons entièrement de ce qu'ils décident de nous distribuer.<sup>12</sup> »*

Mais il serait toutefois inexact de conclure à un renforcement unilatéral de la dépendance des journalistes de Chicago à la source policière. Afin de contrer cette dépendance à une source unique, plusieurs organisations de presse ont cependant cherché à combiner les données provenant de sources institutionnelles différentes tout en maintenant la granularité et la complétude des données criminelles recueillies. Le projet « Homicide Report » du *Los Angeles Times* a été avant-gardiste en la matière : systématiquement, les données policières étaient comparées ou complétées avec les données provenant des médecins légistes. En 2009, les responsables du *Chicago Tribune* suivent ce modèle en chargeant une jeune journaliste, Tracy Swartz, du décompte quotidien des homicides à Chicago. Intitulée « Tracking homicides in Chicago »<sup>13</sup>, sa rubrique hebdomadaire s'appuie sur un patient travail de vérification et de recoupement systématique de sources policières, médico-légales et des archives du journal :

*« Ce qui est différent avec les homicides, c'est qu'il y a un cadavre ; un médecin légiste doit faire une autopsie. La police a ses propres informations. Et le médecin légiste a ses propres informations. Donc, j'ai pensé qu'il serait plus rigoureux de dépeindre cette réalité à partir de ces deux sources d'information.<sup>14</sup> »*

Le fait de recouper des données provenant de sources différentes permet ici aux journalistes de jouer de la pluralité des sources institutionnelles, et finalement de regagner en autonomie à l'égard du *Chicago Police Department* :

*« Si nous avons accès à une information de meilleure qualité de la part de la police, nous ne serions pas obligés d'aller en chercher ailleurs. Mais la police peut sans cesse recatégoriser n'importe lequel de ces meurtres, et c'est pourquoi à la fin de l'année, vous pourrez aller sur le site de Tracy et compter le nombre de meurtres, et vous verrez que cela ne correspondra pas aux chiffres de la police.<sup>15</sup> »*

**Figure 2 : « Chicago sees 448 homicides in 2010<sup>16</sup> »**  
(6 janvier 2011)

Based on preliminary data from the police and Cook County Medical Examiner's Office, RedEye determined there were 448 homicides in 2010, which was 13 fewer homicides than in 2009.

Chicago Police Supt. Jody Weis said this week the department determined there were 435 homicides in 2010, a record low since 1965, based on their own investigations.

Twenty-five homicides were recorded in December, including fatal shootings in Bridgeport and Belmont Cragin and a fatal assault in South Deering in the last week, according to RedEye data.

There were 37 homicides in December 2009, RedEye determined.

On voit donc que le traitement de données — ici la constitution d'une base des homicides à partir des données provenant des coroners et de la police — permet aux journalistes du *Chicago Tribune* de contester le monopole du *Chicago Police Department* sur la production des données criminelles. L'appui sur une norme journalistique bien établie — le recoupement des sources — permet ici de gagner en autonomie à l'égard de la source institutionnelle<sup>17</sup>.

**La police impose-t-elle plus difficilement ses interprétations ?**

À Chicago, la conception de produits journalistiques en ligne reposant sur des données criminelles permet de faire émerger de nouvelles interprétations qui échappent en partie à la source policière. Il faut ici distinguer les interprétations qui concernent le nombre de crimes commis dans la métropole ou dans un quartier précis, de celles qui sont relatives à un crime singulier (a-t-on identifié le bon coupable ? La police consacre-t-elle des moyens suffisants à l'enquête ?). La plus large circulation des données criminelles à Chicago permet d'abord à des citoyens et à des militants de se faire leur propre idée du nombre de crimes commis dans une zone précise de la ville :

*« C'est toujours intéressant de voir comment les gens utilisent les données. (...) Des groupes s'en servent pour créer des banderoles. J'ai entendu que des associations utilisent les données pour faire du porte-à-porte dans certains quartiers et susciter la prise de conscience en quelque sorte. « Savez-vous qu'il y a eu X meurtres à Chatham cette année ? Vous vivez ici. Vous devriez être bouleversé par ça ! » Je sais que plusieurs conseillers municipaux en ont pris conscience et sont allés trouver la police pour demander plus de policiers dans leur quartier.<sup>18</sup> »*

Cela concerne notamment des associations de lutte contre la violence par armes à feu, qui utilisent les données à la fois pour sensibiliser les habitants

d'un quartier et organiser leur action. La diffusion des données parmi les habitants de la ville conduit donc à l'émergence et à la circulation d'interprétations qui peuvent, sinon totalement s'opposer à celles qui sont mises en avant par la police, du moins permettre la constitution de nouvelles formes de dénonciation autour de la question de la quantification des crimes.

Il en va autrement des interprétations concernant des crimes singuliers. Notre enquête montre d'abord que les journalistes du *Chicago Tribune* peuvent décider d'enquêter sur un crime — en interrogeant les proches de la victime ou la police chargée de l'enquête — à partir de l'analyse des données criminelles. Le traitement de données leur permet alors d'identifier un crime qui est révélateur d'un phénomène plus général et qui mériterait à leurs yeux une plus grande investigation. Mais il est beaucoup plus rare de voir des citoyens analyser les données pour se forger leur propre interprétation sur un crime. D'abord parce que les données sont lacunaires — elles ne disent rien des moyens déployés par la police ou de l'évolution de l'enquête — et parce que la police conserve la mainmise sur la catégorisation des actes criminels et la fabrication des données. Mais plus généralement, on constate que le seul accès aux données ne saurait automatiquement susciter des interprétations alternatives au sein de la population de Chicago. C'est ce qui explique une partie des réticences qu'a pu rencontrer le site EveryBlock.com auprès des journalistes de Chicago. Ainsi Brian Boyer, manager de la « News Applications Team » du *Chicago Tribune* de 2009 à 2012, juge que les données brutes ne racontent pas une histoire :

*« Je ne voudrais pas taper trop fort sur Adrian et sur EveryBlock, mais je pense vraiment que ce site manque cruellement d'âme. On le consulte et on se dit : « oh, regarde, il y a un crime là ». Adrian croît fermement à cette idée que les données racontent une histoire, mais je ne le pense pas. Je pense que ce que nous faisons, c'est mettre en perspective les données, et je ne veux pas simplement diffuser un jeu de données brutes. Je veux essayer de raconter une histoire avec ces données.<sup>19</sup> »*

Les journalistes demeurent en première ligne pour contester les interprétations véhiculées par la police au sujet des crimes commis dans la ville. Et souvent, le traitement de données ne constitue que l'un des éléments de l'enquête du reporter. Considéré sous cet angle, l'asymétrie demeure importante entre les organisations de presse et le *Chicago Police Department*, qui dispose toujours de ressources considérables pour imposer ses interprétations.

---

## CONCLUSION

---

Alors que se développent aujourd'hui les recherches sur les nouvelles pratiques journalistiques reposant sur le traitement de données (Parasie et Dagiral, 2012a ; Anderson, 2012 ; Turner et Hamilton, 2009), il serait à la fois naïf et présomptueux de juger définitivement dans quel sens celles-ci affectent le rapport des journalistes à leurs sources. Ces pratiques connaissent aujourd'hui un essor dans plusieurs pays du monde — principalement aux États-Unis, en Grande-Bretagne et dans les pays d'Europe du Nord — et nous manquons aujourd'hui de travaux empiriques qui mettraient en lumière la diversité des pratiques, notamment sous l'angle du rapport aux sources. En outre, ce phénomène est étroitement lié aux liens qui s'établissent depuis quelques années entre les mondes journalistiques et informatiques, et qui sont là encore trop peu documentés.

Il est cependant possible de formuler plusieurs conclusions à partir de notre enquête. En premier

lieu, il apparaît qu'à travers la mobilisation des technologies associées au traitement de données, les journalistes cherchent à gagner en liberté dans le rapport aux sources, institutionnelles notamment. Cela repose sur un ensemble de promesses dont nous avons esquissé les principales composantes, mais qu'il serait important d'explorer de façon plus profonde. En second lieu, l'analyse des pratiques journalistiques en matière criminelle dans le contexte de Chicago indique que ces promesses ne se traduisent pour le moment que de façon partielle dans les réalisations. Le déclin du monopole policier sur la diffusion des données — encouragée par l'institution policière elle-même — a incontestablement permis aux journalistes de gagner en liberté. Cependant, l'élaboration des données restant une prérogative de la police, l'accroissement des marges de manœuvre journalistiques demeure relativement restreinte. On peut néanmoins faire l'hypothèse que l'un des enjeux majeurs lié à la mobilisation de l'informatique dans les rédactions dans les prochaines années résidera dans l'expérimentation d'autres relations avec les sources journalistiques.

---

## NOTES

---

<sup>1</sup> Cette enquête a été menée dans le cadre de l'ANR BASICOM (bases informatiques et coopération entre mondes sociaux) et d'une recherche sur la presse locale en ligne et ses publics, financée par le ministère de la Culture (département des études, de la prospective et des statistiques).

<sup>2</sup> Washington est la première ville américaine à s'être dotée d'un portail diffusant les données de l'administration urbaine en 2004. Plusieurs autres grandes villes ont suivi ce mouvement, dont Chicago en mai 2010.

<sup>3</sup> Daniel X. O'Neil, entretien réalisé le 8 septembre 2010.

<sup>4</sup> Ben Welsh, « Human-assisted reporting : how to create robots reporters in your own image ? », 13th International Symposium on Online Journalism, Austin, Texas, 20-21 avril 2012. Conférence consultable à cette adresse : <http://www.youtube.com/watch?v=iP-On8PzEy8> (consulté le 1er décembre 2012).

<sup>5</sup> On estime à 448 le nombre de personnes qui ont été assassinées à Chicago en 2010. Cette année-là, le taux d'homicide à Chicago a été de 15 homicides pour 100 000 habitants, soit le double du taux qui prévalait à New York et à Los Angeles (source : FBI).

<sup>6</sup> Il s'agit du programme CAPS (*Chicago Alternative Police Strategy*), mis en place en 1993 par le maire Richard Daley.

<sup>7</sup> Source : « The Chicago Police Department's Information Collection for Automated Mapping Program », Program Focus, National Institute of Justice, 1996.

<sup>8</sup> Via le site ClearMap Chicago, toujours disponible aujourd'hui à cette adresse : <http://gis.chicagopolice.org>.

<sup>9</sup> Cf. Steve Outing, 1995.

<sup>10</sup> Les homicides touchant des personnes qui ne vivent pas dans des quartiers pauvres et qui ne sont pas membres de gang sont ceux qui présentent la plus grande valeur journalistique pour les organisations de la presse locale.

<sup>11</sup> Tempéré par le fait que l'État de l'Illinois et le FBI produisent également des données, plus limitées, sur la criminalité à Chicago.

<sup>12</sup> Darnell Little, entretien réalisé le 10 septembre 2010.

<sup>13</sup> Disponible à cette adresse : <http://homicides.redeyechicago.com>

<sup>14</sup> Tracy Swartz, entretien réalisé le 2 septembre 2010.

<sup>15</sup> Darnell Little, entretien réalisé le 10 septembre 2010.

<sup>16</sup> Disponible à cette adresse : <http://homicides.redeyechicago.com/blog/page/14/>

<sup>17</sup> À certaines occasions, le jeu de la pluralité des sources en matière de données criminelles ne se limite pas au recouplement, mais passe par la combinaison de données dans le but de contraindre la source à révéler certains faits qu'elle ne souhaite pas faire connaître. En septembre 2010, Tracy Swartz participe ainsi à la conception d'une application qui permet de croiser la base des homicides qu'elle élabore quotidiennement avec les données diffusées par la compagnie de transports urbains — faisant ainsi apparaître le nombre et le type de crimes commis selon les lignes de métro et de bus de Chicago. Suite à la contestation de la compagnie, l'application est mise hors ligne quelques jours plus tard.

<sup>18</sup> Tracy Swartz, entretien réalisé le 2 septembre 2010.

<sup>19</sup> Brian Boyer, entretien réalisé le 9 septembre 2010.

## RÉFÉRENCES BIBLIOGRAPHIQUES

---

- Akrich, M. et Latour, B., 1992, « A Summary of a Convenient Vocabulary for the Semiotics of Human and Nonhuman Assemblies », in Bijker, W. and Law, J. (Eds), *Shaping Technology-Building Society: Studies in Sociotechnical Change*, Cambridge, MA, MIT Press, pp. 205-224.
- Anderson, C., 2006, *The Long Tail. How the Future of Business is Selling Less of More*, New York, Hyperion Books.
- Anderson, C. W., 2012, « Towards a sociology of computational and algorithmic journalism », *New Media and Society*, mis en ligne avant publication, 10 décembre 2012.
- Arthur, C., 2010, « Analyzing data is the future for journalists, says Tim Berners-Lee », *The Guardian*, 22 novembre 2010.
- Aubert, A., 2009, « Le paradoxe du journalisme participatif. Motivations, compétences et engagements des rédacteurs des nouveaux médias (enquête) », *Terrains & travaux*, n° 15, pp. 171-190.
- Birchall, C., 2011, « Introduction to “secrecy and transparency” : the politics of opacity and openness », *Theory Culture Society*, vol. 28, n° 7-8, pp. 7-25.
- Cardon, D., Granjon, F., 2003, « Peut-on se libérer des formats médiatiques ? Le mouvement alter-mondialisation et l’Internet », *Mouvements*, n° 25, pp. 67-73.
- Chamard, S., 2006, « The history of crime mapping and its use by American police departments », *Alaska Justice Forum*, vol. 23, n° 3.
- Cohen, S., Hamilton, J., Turner, F., 2011a, « Computational journalism », *Communications of the ACM*, vol. 54, n° 10, pp. 66-71.
- Cohen, S., Li, C., Yang, J, Yu, C., 2011b, « Computational Journalism : A Call to Arms to Database Researchers », *5th Biennial Conference on Innovative Data Systems Research*, January 9-12 2011, Asilomar, California, USA.
- Daston, L., Gallison, P., 2012, *Objectivité*, Paris, Les Presses du Réel.
- Dodier, N., 1995, *Les hommes et les machines*, Paris, Métailié.
- Domingo, D., Quandt, T., Heinonen, A., Paulussen, S., Singer, J. B., Vujanovic, M., 2008, « Participatory journalism practices in the media and beyond. An international comparative study of initiatives in online newspapers », *Journalism Practice*, vol. 2, n° 3, pp. 326-342.
- Domscheit-Berg, D., 2011, *Inside Wikileaks. Dans les coulisses du site Internet le plus dangereux du monde*, Paris, Grasset.
- Estienne, Y., 2010, « Indymedia aujourd’hui : la critique en acte du journalisme et les paradoxes de l’open publishing », *Mouvements*, vol. 1, n° 61, pp. 121-131.
- Gray, J., Bounegru, L, Chambers, L., 2012, *The Data Journalism Handbook*, O’Reilly.
- Hall, S., Critcher, C., Jefferson, T., Clarke, J., Roberts, B., 1978, *Policing the Crisis. Mugging, the State, and Law and Order*, London, MacMillan.
- Hamilton, J. T., Turner, F., 2009, « Accountability through Algorithm : Developing the Field of Computational Journalism », Stanford, CA, Center for Advanced Study in the Behavioral Sciences Summer Workshop, 27-31 juillet.
- Holtgrewe, U., Werle, R., 2001, « De-commodifying software ? Open source software between business strategy and social movement », *Science Studies*, vol. 14, n° 2, pp. 43-65.
- Lathrop, D., Ruma, R., 2010, *Open Government : Collaboration, Transparency and Participation in Practice*, Sebastopol, CA, O’Reilly Media.
- Lemieux, C., 2000, *Mauvaise presse : une sociologie compréhensive du travail journalistique et de ses critiques*, Paris, Métailié.
- Marchetti, D., 2000, « Les révélations du “journalisme d’investigation” », *Actes de la recherche en sciences sociales*, vol. 131-132, pp. 30-40.
- Molotch, H., Lester, M., 1996, « Informer : une conduite délibérée de l’usage stratégique des événements » [1974], *Réseaux*, n° 75, pp. 23-41.
- Outing, S., 1995, « Newspapers Adding Databases to the Web », *Editor & Publisher*, octobre 1995.
- Parasie, S., 2011, « Le journalisme hacker : une nouvelle utopie pour la presse ? », *La vie des idées*, 21 juin 2011. <http://www.laviedesidees.fr/Le-journalisme-hacker.html>
- Parasie, S., Dagiral, É., 2012a, « Data-driven journalism and the public good. “Computer-assisted reporters” and “programmer-journalists” in Chicago », *New Media and Society*, mis en ligne avant publication, 18 novembre 2012.
- Parasie, S., Dagiral, É., 2012b, « Quand le Web colle au territoire : l’exploration de l’information hyperlocale à Chicago », *Sciences de la société*, n° 84-85, pp. 81-101.
- Pilhofer, A., 2010, « Programmer-Journalist ? Hacker-Journalist ? Our Identity Crisis », *Mediashift IdeaLab*, [www.pbs.org/ideallab/2010/04/programmer-journalist-hacker-journalist-our-identity-crisis107.html](http://www.pbs.org/ideallab/2010/04/programmer-journalist-hacker-journalist-our-identity-crisis107.html)
- Rebillard, F., 2011, « Création, contribution, recommandation : les strates du journalisme participatif », *Les Cahiers du journalisme*, n° 22-23, pp. 28-41.
- ROYAL, C., 2010, « The Journalist as Programmer : A Case Study of the New York Times Interactive News Technology Department », *International Symposium on Online Journalism*, Austin, The University of Texas, avril 2010.
- Schlesinger, P., 1992, « Repenser la sociologie du journalisme. Les stratégies de la source d’information et les limites du média-centrisme » [1990], *Réseaux*, n° 51.
- Schudson, M., 2010, « Political observatories, databases & news in the emerging ecology of public information », *Daedalus*, n° 139, pp. 100-109.
- Swartz, A., 2010, « When is Transparency Useful ? », in Lathrop, D. et Ruma, R. (Eds.), *Open Government*, Sebastopol, CA, O’Reilly Media, pp. 267-272.
- Weiss, A., Domingo, D., 2010, « Innovation processes in online newsrooms as actor-networks and communities of practice », *New Media & Society*, vol. 12, n° 7, pp. 1156-1171.
- Welsh, B., 2009, « Q&A : EveryBlock’s Adrian Holovaty on the LAPD and Accuracy Standards ». <http://latimesblogs.latimes.com/lanow/2009/04/adrian-holovaty-everyblock.html>

**Fr.** Depuis plusieurs années, le traitement de données apparaît à certains journalistes comme une façon de réduire leur dépendance aux sources. En élaborant des bases de données et des algorithmes au sein des rédactions, ils seraient moins tributaires des tuyaux et des interprétations que leur fournissent régulièrement leurs sources, notamment institutionnelles. S'inspirant de la sociologie pragmatique et de la théorie de l'acteur-réseau, cet article montre que les technologies de traitement de données sont actuellement mobilisées afin de permettre aux journalistes de reprendre l'initiative face aux sources. À partir d'une enquête de terrain menée à Chicago en 2010, nous examinons d'abord le contenu de cette promesse apparue au croisement des mondes journalistiques et informatiques. Puis, en étudiant le cas de l'information criminelle à Chicago, nous montrons que cette promesse n'est qu'en partie satisfaite dans les réalisations actuelles.

**Mots-clés :** data-journalisme, journalisme de données, artefacts sociotechniques, sources, information criminelle

**En.** Over the past few years, some journalists have expressed that data processing is a way of reducing their dependence to sources. Elaborating databases and algorithms in the newsrooms is considered to reduce journalists' need for tips and interpretations from sources, including those institutional. By drawing from pragmatic sociology and the network-actor theory, this article shows that data-processing technologies are currently mobilized to encourage journalists to reinitiate their relation to sources. Based on a study that took place in Chicago in 2010, first we examine the content of the promise that appeared at the intersection of the journalistic and computer science worlds. Then, by analysing criminal news in Chicago, we show that currently, the promise is only partly fulfilled.

**Keywords:** interview, sociology, journalism, profession, jurisdictional claim

**Pt.** Há vários anos que alguns jornalistas acreditam que o processamento de dados pode ser uma forma de reduzir sua dependência em relação às fontes de informação. Ao desenvolverem bases de dados e algoritmos no interior das redações, eles seriam menos dependentes das dicas e interpretações fornecidas regularmente pelas fontes, sobretudo as institucionais. Tendo como inspiração a sociologia pragmática e a teoria do ator-rede, este artigo mostra como as tecnologias de tratamento de dados são atualmente mobilizadas para permitir que os jornalistas possam assumir novamente a iniciativa frente às fontes. A partir de uma pesquisa de campo realizada em Chicago, em 2000, examinamos, em um primeiro momento, o conteúdo dessa promessa, advinda do cruzamento dos mundos do jornalismo e da informática. Em seguida, por meio de um estudo de caso sobre a informação criminal em Chicago, nós mostramos que essa promessa não foi totalmente cumprida.

**Palavras-chave:** data journalism, artefatos sociotécnicos, fontes, notícias sobre crimes

