

# **LE DATAJOURNALISM PAR SES ACTEURS, AUTODÉFINITION D'UNE PRATIQUE ÉMERGENTE EN FRANCE**

**Valérie Croissant et Annelise Touboul<sup>1</sup>**

Cet article se propose d'interroger les formes d'interventions des acteurs dans l'espace public et médiatique pour promouvoir l'émergence d'un nouveau genre de production journalistique, le *datajournalism* ou journalisme de données. L'étude souligne les filiations du *datajournalism* qui se trouve à la croisée des chemins de l'informatique, du graphisme et du journalisme. L'analyse porte dans un premier temps sur les discours des acteurs pour mettre au jour les présupposés du nouveau contrat de communication dessiné par les promoteurs du journalisme de données. Dans un second temps, l'étude sémiotique de quelques réalisations primées par les professionnels du secteur permet d'interroger une autre modalité de promotion du *datajournalism*.

## **Le datajournalism par ses acteurs, autodéfinition d'une pratique émergente en France**

L'objet d'étude de cet article présente une difficulté terminologique dans la mesure où il s'inscrit dans de multiples domaines dont la situation d'émergence empêche la mobilisation de catégories stables. Le terme anglo-saxon de « *datajournalism* » traduit en français par journalisme de données, est lié à d'autres termes qui traduisent une diversité de pratiques sociales et professionnelles ne se limitant pas au seul secteur du journalisme et des médias. En effet, infographie,

---

<sup>1</sup> Valérie Croissant est Enseignant-Chercheur à l'Université Lumière Lyon 2. Annelise Touboul est Maître de Conférences en sciences de l'information et de la communication à l'Université Lumière Lyon 2.

*fact-checking*, *visual mining*, *data visualization*, *open data* dépassent largement le domaine du journalisme. Si par commodité, nous utilisons les vocables de *datajournalism* ou journalisme de données dans le cadre de cet article, cela ne signifie pas qu'il s'agit d'une catégorie stable et construite. Ainsi, l'existence du *datajournalism* en tant que nouvelle pratique journalistique en France ne sera pas mise en question dans la mesure où notre objet d'analyse est moins le journalisme de données en soi que ce qui le fait advenir, ce qui le rend visible à travers différentes formes discursives. Les travaux de Sylvain Parasio et Eric Dagiral (2011) sur le développement du *datajournalism* aux Etats-Unis montrent de quelle manière les pratiques se structurent à partir notamment de la figure du « *hacker* » mais aussi du poids des compétences techniques. Dans la lignée de ces travaux, nous souhaitons nous intéresser à l'apparition du phénomène en France dont l'histoire du journalisme crée un contexte fort différent de celui des Etats-Unis. Il s'agit alors d'observer la manière dont les acteurs font exister le phénomène dans l'espace public en le nommant et en le qualifiant. En référence aux travaux conduits par Flichy (2001), nous considérons que les discours d'acteurs et les expérimentations portent la trace de valeurs, d'imaginaires collectifs qui participent à la définition du *datajournalism*, à sa légitimation en tant que pratique journalistique identifiée. Sur quels arguments cet ensemble de discours et de productions s'appuie-t-il pour faire exister, rendre visible et légitimer la catégorie émergente nommée *datajournalism* ? Dans le but de saisir les processus de construction d'une pratique innovante nous faisons le choix d'une approche sémi-discursive pour révéler les présupposés et les stratégies à l'œuvre dans les discours ainsi que les significations construites par les réalisations du *datajournalism*. Ainsi, tout en traçant les liens du journalisme de données avec d'autres champs sociaux, nous présenterons trois étapes dans notre travail : une analyse des discours faisant la promotion du journalisme de données autour de deux points récurrents : la puissance de la donnée et celle l'image. Dans un second temps, le travail portera sur un concours de *datajournalism* qui fera l'objet de deux traitements : une analyse de l'événement en tant qu'objet discursif et une étude des productions primées.

## **1. Le datajournalism dans les discours : filiations et promesses**

A travers un corpus d'articles publiés sur Internet, nous avons relevé les principaux arguments qui assoient la pertinence et la valeur

du journalisme de données. Le choix s'est opéré sur des individus, des institutions, des lieux qui se sont positionnés rapidement et énergiquement sur le sujet. La diversité des intervenants n'est pas très importante, les mêmes personnes étant souvent sollicitées sur le sujet<sup>1</sup>. Le mode de collecte des discours par *navicrawling* (navigation en suivant les liens entre les sites) permet l'identification des positions influentes, des récurrences ou des variations. Ainsi la composition du corpus, partiel et hétérogène, s'est attachée à cueillir une écume de discours médiatiques, une surface agitée par une actualité qui ne rend pas compte de l'ampleur ou de la réalité d'un phénomène, mais seulement de la volonté de certains acteurs à rendre le phénomène présent et à y occuper une place centrale. Les textes (une quarantaine de documents) ont été publiés entre 2010 et 2012 sur des sites d'information ou des blogs de professionnels (journalistes, formateurs, consultants). Cette dimension d'exposition induit leur rôle structurant dans l'installation du journalisme de données comme une pratique sociale et professionnelle avérée. Ils constituent des discours d'accompagnement du changement, le changement portant ici sur la nature de l'information et au-delà, sur la définition professionnelle du journalisme. Les arguments légitimant le journalisme de données comme une nouvelle voie du journalisme portent principalement sur deux dimensions, le traitement des données issues du monde réel et la mise en scène visuelle de ces informations pour parvenir à une compréhension de phénomènes complexes. Ces dimensions ancrent le journalisme de données en France dans d'autres champs sociaux connexes, parmi lesquels l'informatique et le graphisme, mais le positionnent également comme une pratique journalistique spécifique. Avant de présenter notre analyse, deux précisions s'imposent : les discours analysés font état d'une certaine diversité dans les arguments et prises de positions. Si notre article se focalise sur les éléments saillants de la définition du journalisme de données, nous ne souhaitons en aucun cas donner une image monolithique de la situation. Deuxièmement, les discours analysés sont portés par des individus plus que par des institutions. Même lorsqu'un journaliste s'exprime en tant que membre d'un journal, d'une équipe éditoriale ou d'un site, il s'exprime surtout à titre individuel, eu égard à son expérience, ses réalisations dans le domaine. Si certains lieux deviennent des références à l'étranger (comme le *Guardian*), le journalisme de données en France se présente comme un secteur en

1 Les personnalités récurrentes du corpus sont Caroline Goulard, Nicolas Kayser-Bril, Alain Joannes.

émergence, donc faiblement institutionnalisé, où les innovateurs et créateurs ont toute leur place.

### **La donnée comme référence scientifique ?**

Le premier champ sémantique associé au journalisme de données est bien sûr celui de l'informatique. Ce domaine étant lui-même parent d'Internet, la filiation est évidente, les processus d'informatisation de la production journalistique ayant grandement transformé le paysage médiatique depuis une trentaine d'années. L'informatique intervient dans la collecte et le traitement des données ou encore l'accès via Internet ; données qu'il faut savoir trier, traiter et faire parler grâce à des logiciels, des applications. Le programmeur est alors indispensable au journaliste. C'est la manifestation de la rencontre entre les mondes des données, au sens de « *data* » et des informations (*news*) ; l'association entre l'informaticien et le journaliste permettant de passer des unes aux autres.

Concernant la définition professionnelle du journaliste de données, deux tendances se distinguent : celle de la double compétence et celle de la coopération. Dans la perspective d'une double compétence, certains professionnels affirment la nécessité pour le journaliste d'être un codeur. Il doit posséder des compétences en programmation, tout en étant capable de lire, comprendre et mettre en formes des données mathématiques et statistiques complexes, être un « journaliste-hacker » (Dagiral & Parasie, 2011). C'est dans cet esprit que certaines formations spécialisées ont été créées dans différentes écoles de journalisme<sup>1</sup>. L'autre option organisationnelle du journalisme de données s'appuie sur la coopération entre les compétences journalistiques et informatiques, portées par des professionnels distincts. La coopération entre des professionnels d'univers socioprofessionnels différents peut ainsi, tout en conservant l'identité professionnelle de chacun, permettre la réalisation de projets dans le journalisme de données, que ceux-ci soient portés par des organisations de presse ou des *start-up* plus proches de l'informatique. La référence culturelle qui trace la filiation entre les techniques informatiques et le journalisme de données est la croyance commune dans la véracité des données. Dans les discours, la donnée est entendue comme un élément objectif extrait directement des activités

1 Une formation à destination des développeurs a été créée par l'École de journalisme de Medill à la Northwestern University. À la Columbia University, le journalisme et l'informatique sont enseignés dans le même cursus.

du monde ; l'objectivation des faits étant ce qui rapproche le journaliste et le scientifique. L'importance des statistiques et du traitement des données s'oppose à la tradition française du journalisme de commentaire dont les fondements sont politiques et littéraires (Neveu, 2009). Le *datajournalism* reposant sur des méthodes scientifiques et non plus sur le seul talent d'écriture, défend une conception du journalisme révélateur de vérité puisque basé sur des données.

Les discours d'acteurs posent de manière récurrente comme toile de fond le phénomène inflationniste de l'information pour souligner l'importance des formes de médiation dont le *datajournalism* constitue une possible réponse. Les sociétés produisent en permanence quantité de données publiques, accessibles en ligne mais invisibles et illisibles sans l'existence de médiations, ce qui permet au journalisme de justifier sa place, avec une dimension pédagogique dominante. L'enjeu de l'information en ligne semble désormais moins celui de l'accès que celui de sa lecture, de sa signification. Se construit alors un discours de l'herméneutique de la donnée, du chiffre qui devrait nous révéler la vérité du monde qui nous entoure. On n'est pas loin de la croyance, d'une révélation comme l'atteste le titre d'un article publié par *Owni.fr* « Le Data journalism notre religion »<sup>1</sup> ou le sous-titre du *Datablog* du *Guardian* « Facts are sacred »<sup>2</sup>.

Dans les discours, on retrouve l'ambiguïté soulignée par Y. Jeanneret concernant le terme « information », selon qu'on lui attribue une définition sociale ou mathématique. L'auteur précise que ces définitions s'opposent puisque la première s'attache à attribuer du sens par l'interprétation, alors que la seconde s'attache à l'évincer. (Jeanneret, 2007, p.63). Cette opposition anime les débats à propos du *datajournalism* qui tournent parfois à l'affrontement entre les partisans de la donnée et ceux de la narration comme mode d'accès au monde. Majoritairement, les discours prônent la coopération entre les compétences, le journaliste restant au centre des projets. La narration, le *storytelling* ou l'éditorialisation ne sont pas éliminés du journalisme de données, mais c'est l'approche s'appuyant sur les mathématiques et la programmation informatique qui légitime le *datajournalism* et en fait une pratique médiatique innovante. Comme le résume Adrian Holovaty<sup>3</sup>,

---

1 *Owni.fr*, article du 30/08/ 2011, dernière consultation le 15/09/2012

2 <http://www.guardian.co.uk/data>

3 Adrian Holovaty, « A fundamental way newspaper sites need to change » publié le 6.09. 2006 sur [www.holovaty.com](http://www.holovaty.com), dernière consultation le 05.05.2012.

considéré comme un des fondateurs du domaine : « *Newspapers need to stop the story-centric worldview* ».

Du fait de la diversité des acteurs impliqués et des positions défendues, le journalisme de données remet à l'ordre du jour les questions relatives au flou de l'identité professionnelle journalistique (Ruellan, 2007). Il nécessite la mobilisation de compétences variées et témoigne de préoccupations présentées comme complémentaires au journalisme : la science pour appuyer le travail d'enquête et de veille ; ou encore le graphisme afin de rendre intelligibles les informations.

### **Le graphisme comme référence artistique ?**

La traduction visuelle des données est une dimension essentielle des discours de promotion du *datajournalism*. Bien plus qu'une simple mise en forme intervenant à la fin du processus dans un seul souci de lisibilité, il s'agit d'une démarche qui conditionne l'intelligibilité de l'information. Dans les contenus journalistiques, l'infographie n'est pas nouvelle mais sa dimension interactive, animée, appliquée à des données complexes est en revanche une caractéristique du journalisme de données. La visée esthétique de ces images est souvent revendiquée et donne d'ailleurs lieu à des productions se présentant parfois dans des lieux ou des formes qui sont celle de l'œuvre d'art<sup>1</sup>. Cependant, la dimension esthétique des infographies se trouve limitée dans sa référence artistique au fait qu'elle est asservie à l'intelligibilité. La visée pédagogique reste forte et la mise en image doit produire, selon les acteurs du journalisme de données, une clarté permettant d'accéder de manière plus efficace à la compréhension des phénomènes sociaux complexes<sup>2</sup>.

### **« Une image vaut mille mots »**

Les discours analysés font fréquemment référence à la puissance de l'image et à ses qualités face au texte. Celui-ci, associé aux notions de commentaire et d'opinion, est opposé aux infographies du *datajournalism* qui traduiraient une forme de vérité directe.

---

1 Robert Klanter, Gilles Berton, *DataFlow : design graphique et visualisation de l'information*, Thames and Hudson, 2004. Le volume 2, 2010.

2 L'exposition itinérante « Expoviz » rend visible une sélection de *datavisualisations* dans un souci de connaissance et de vulgarisation auprès du grand public. [www.expoviz.fr](http://www.expoviz.fr)

Les infographies animées, résultats de calculs et d'algorithmes informatiques, affirment leur rapport avec le réel basé sur la donnée ce qui justifierait le primat de l'image sur le texte. Mais d'autres acteurs, ceux notamment qui inscrivent pleinement leur pratique dans le journalisme, prônent une complémentarité indispensable entre texte et image. Ils envisagent le message linguistique dans sa fonction d'ancrage du message iconique (Barthes, 1964), permettant de limiter, cadrer le sens des visuels. Le texte est perçu comme étant au service de l'image : si la tradition européenne du journalisme littéraire l'a placé au centre des préoccupations, le *datajournalism* lui réassigne désormais une place plus modeste.

Considérant les impératifs de productivité des journalistes, c'est le gain de temps sur l'énoncé verbal qui prime ; l'affirmation « une image vaut mille mots » étant à ce titre souvent mobilisée<sup>1</sup>. C'est dans l'acte d'embrasser du regard et du cerveau les données que la rapidité doit opérer. Cette affirmation de l'efficacité de la représentation graphique s'appuie sur des travaux en psychologie cognitive (Treisman, 1985); d'autres prolongent ceux de Jacques Bertin (Cleveland & McGill, 1984) et soulignent les économies permises par la représentation graphique en comparaison du texte : économies de temps, de place mais aussi d'efforts. L'image, ici, a pour fonction de rendre simple ce qui est complexe. A. Joannès (2010) dans son manuel sur le journalisme de données, utilise le terme de « simplicité » pour qualifier ce pouvoir quasiment magique de l'image. C'est bien la séduction de l'image qui est censée agir, principe déjà appliqué dans la presse imprimée qui veut que le texte rebute et soit chronophage, alors que l'image attire et peut être embrassée d'un seul coup d'œil.

### **Les promesses de l'animation et de l'interactivité**

L'animation et l'interactivité ne sont pas des conditions nécessaires du journalisme de données ; mais si les données ne sont pas animées, elles seront au moins modulables. La notion d'interactivité qui revient au centre des préoccupations de la production d'information constituait le point de mire de nombreux projets du journalisme en ligne à ses débuts. Elle s'est ensuite épuisée dans la recherche d'un modèle économique viable de la presse en ligne et sa définition fut quelque peu galvaudée lorsque devint interactif tout ce qui propose quelques liens vers d'autres

1 Phrase communément attribuée à Confucius, elle est entrée dans le langage commun comme un principe.

pages web. Avec le journalisme de données, l'interactivité retrouve une seconde jeunesse en permettant à l'utilisateur, une entrée singulière dans les données. L'image infographique semble alors atteindre le statut d'image totale, celle qui peut répondre à toutes les questions, être un potentiel infini, un disponible inépuisable. Le dispositif place l'utilisateur au centre, co-producteur non seulement du sens mais de la forme même de l'image. Il agit sur elle, change l'échelle « *zoome ou dézoome* » pour reprendre l'expression de D. Cardon (2011) qui souligne les dangers de cette approche auto-centrée et ses conséquences politiques. Cette prise en compte du lecteur-internaute se présente comme une dimension du journalisme de donnée qui ne considère plus un public, mais chaque membre de ce public. La promesse de communication est que « chacun doit y trouver son compte » dans la mesure où chaque individu peut adapter la forme et le contenu informationnel à sa situation en fonction de critères démographiques, géographiques, sociologiques, compétences, intérêts, etc.

Le consensus n'a pas lieu d'être dans le journalisme de données, pas plus qu'il n'existe dans la définition même du journalisme en général. Ces discours confirment ainsi que le journalisme n'est pas un domaine fermé sur lui-même, mais que cette perméabilité à d'autres pratiques sociales lui est inhérente (Ringoot & Utard, 2006). Si les discours d'acteurs analysés diffèrent par le statut de leurs énonciateurs (journalistes, développeurs, blogueurs..), ils témoignent d'une conscience partagée des enjeux liés au journalisme de données. Il est moins question de compétences, d'outils ou de modèle économique, que d'une posture, un « état d'esprit » qui réaffirme la médiation journalistique entre le monde et le lecteur-internaute. Les acteurs du journalisme de données participent ainsi à la grande et incessante bataille sur la manière de percevoir, rendre compte et construire le monde. Cependant, la récurrence des personnes qui s'expriment et des lieux qui font la promotion du *datajournalism* témoignent de la faible ampleur du phénomène en France (le site *Owni.fr* a fermé depuis notre enquête) qui semble peu entrer en résonance avec l'identité professionnelle historique du journalisme français. Dans ce contexte, l'élaboration d'espaces collectifs de valorisation est un moyen complémentaire important pour la reconnaissance du domaine. Ainsi, l'organisation d'un premier concours international constitue une action de promotion et de structuration d'un champ professionnel qu'il nous a semblé intéressant d'étudier de plus près.



## 2. Le concours de datajournalism, un outil stratégique

Les 30 et 31 mai 2012, s'est tenu à Paris le premier « *Data Journalism Awards* » qui a présenté d'après *Owni.fr* « *la crème de la crème des applications de data journalism produites à travers le monde* »<sup>1</sup>. Les 57 projets nominés (sur près de 300 reçus selon les organisateurs) ont été répartis en trois catégories : le journalisme d'investigation conduit à partir de données (*data driven investigation*), la visualisation de données et la narration (*data visualisation and storytelling*) et enfin, les applications web et mobiles autour des données (*data driven applications*). Dans chacune des catégories, deux projets ont été récompensés, le premier d'envergure nationale ou internationale et le second d'envergure locale ou régionale.

### Dessiner un territoire et fixer un cap

À travers la définition du périmètre de chacune des trois catégories, il apparaît que l'utilité sociale des réalisations est un critère indispensable et commun à toutes (« *The work should cover matters relevant to society and should aim to have results or an impact on a societal level* »<sup>2</sup>). S'il est demandé à la catégorie *investigation* de présenter un travail d'éditorialisation, la catégorie *visualisation* doit proposer une traduction visuelle de données complexes et enfin, la catégorie *application* récompense la modularité et l'adaptation aux usages individuels. Les objectifs assignés au concours tiennent en quatre points<sup>3</sup> : contribuer à établir des normes dites de qualité, proposer des modèles considérés comme stimulants pour les journalistes, convaincre les dirigeants des entreprises médiatiques et enfin promouvoir les collaborations pluridisciplinaires. Les cibles sont donc avant tout professionnelles : journalistes, patrons de presse et tous les acteurs qui pourraient apporter leurs compétences à des collaborations futures. Enfin, la composition du jury<sup>4</sup> permet d'affirmer l'importance de la dimension internationale (chaque continent est représenté), de

---

1 <http://owni.fr/2012/05/07/les-data-en-forme-episode30/>

2 Phrase reprise pour définir toutes les catégories du concours, <http://datajournalismawards.org/prizes/>

3 voir la page de présentation du projet, <http://datajournalismawards.org/about-the-dj-awards/>. Consulté le 24/02/2013.

4 Biographies détaillées des membres du jury sur <http://datajournalismawards.org/jury/>

revendiquer une légitimité (avec des personnalités déjà récompensées par divers concours) et une priorité accordée à l'indépendance et au travail d'enquête (choix de Paul Steiger, PDG d'une structure qualifiée de « *non-profit investigative newsroom* » pour présider le jury). Les critères de cette manifestation internationale semblent indiquer en outre que les organisateurs ne veulent pas imposer un centre au journalisme de données, que ce centre soit géographique (dans une seule langue ou un seul pays) ou professionnel (pas d'exclusivité des entreprises de presse).

Le site du concours publie les commentaires et justifications du jury concernant les prix attribués. Il ressort ainsi que chaque production a été évaluée dans sa singularité, et que le concours est moins un classement qu'une distinction de quelques productions pour servir d'exemple. Les commentaires reprennent les éléments de définition du *datajournalism* identifiés dans notre étude des discours d'acteurs. Les arguments valorisent la profondeur de l'enquête et de l'analyse pour enrichir les données brutes qui ne constituent pas en elles-mêmes des informations suffisantes (« *very dry data* »). L'importance des interfaces et des modélisations est réaffirmée car elles donnent accès à l'information et la rendent pertinente pour les usagers. C'est bien le travail de médiation qui est ainsi valorisé et qui justifie l'ancrage dans le journalisme. Le *datajournalism* se distingue ainsi des autres secteurs qui mobilisent données et applications informatiques à des fins marketing, industrielles, sécuritaires etc. et affirme son rôle dans l'espace public.

### **Proposer un palmarès soucieux de diversité**

Avant de procéder à l'analyse, il est nécessaire de présenter succinctement les réalisations lauréates.

Dans la catégorie enquête :

- *Mother Jones*, mensuel de presse écrite (USA), présente un dossier sur les informateurs du FBI dans la lutte antiterroriste sur le sol étasunien, à partir de rapports d'enquêtes et de comptes-rendus de jugements de 2001 à 2011,
- *The Seattle Time*, quotidien régional (USA), publie son enquête sur les décès liés à la Méthadone, médicament anti-douleur bon marché, à partir de rapports publics, d'enregistrements de colloques de médecine, de données socio-professionnelles concernant les

2173 personnes décédées entre 2003 et 2010 et consommant de la Méthadone.

Dans la catégorie visualisation des données :

- *The Guardian*, quotidien national (GB), tente de représenter l'origine, le développement et la fin de sept rumeurs sur *Tweeter* pendant les manifestations de violences de l'été 2011 à Londres, à partir de l'analyse de 2,6 millions de tweets.
- *NGS*, (Russie) projet de trois concepteurs, publié sur le site personnel de l'un d'entre eux (*web designer*) qui cartographie, à partir des données de la police, les accidents de la route impliquant des piétons dans la ville de Novosibirsk en 2011.

Dans la catégorie des applications :

- *Politnet.ag* structure suisse indépendante, non médiatique regroupe de nombreuses données sur l'activité du parlement suisse.
- *Chicago Tribune*, quotidien régional (USA), propose une application sur les écoles de Chicago à partir de données publiques fournies par le conseil d'État de l'Illinois.

Catégorie : *data driven investigation*

National/international	Local/régional
<a href="#">Terrorists for the FBI</a> , Mother Jones, USA	<a href="#">Methadone and the politics of pain</a> , The Seattle Times, USA

Catégorie : *data visualisation and storytelling*

National/international	Local/régional
<a href="#">Riot rumours : how misinformation spread on Twitter during a time of crisis</a> , The Guardian, Royaume-Uni	<a href="#">Pedestrian Crashes in Novosibirsk 2011</a> , NGS, Russie

Catégorie : *data driven applications*

National/international	Local/régional
<a href="#">Transparent Politics</a> , Politnetz AG, Suisse	<a href="#">2011 Illinois school report cards</a> , Chicago Tribune, US

**Figure 1.** Liste des lauréats du « Data Journalism Awards 2012 » dans chaque catégorie<sup>1</sup>

1 Palmarès et réalisations primées visibles sur <http://datajournalismawards.org/>

Cette présentation rapide appelle d'ores et déjà quelques commentaires. Seulement quatre des six réalisations ont été publiées par des sites se revendiquant du journalisme conformément aux intentions des organisateurs du concours de fédérer par-delà la seule sphère professionnelle des médias. Par ailleurs, les projets ne s'inscrivent pas dans l'urgence de l'actualité immédiate et manifestent des préoccupations sociales (justice, santé, vie politique, sécurité, écoles...). Enfin, le palmarès rend compte de l'avance étasunienne dans le domaine puisque la moitié des réalisations provient des États-Unis d'Amérique ce qui s'explique aisément par l'antériorité du pays en matière de CAR (*Computer Assisted Reporting*) et de *datajournalism* (Dagiral & Parasio, 2011).

### **Des projets collaboratifs et engagés**

Pour articuler les discours aux réalisations concrètes, nous avons procédé à l'analyse des six projets récompensés au *Data Journalism Awards* en mai 2012. Compte tenu de la diversité des projets, nous avons renoncé à adopter une approche comparative pour privilégier le repérage systématique des éléments relatifs à la méthodologie employée, au travail de mise en forme visuelle et de mise en récit des données. Ainsi, nous avons tout d'abord cherché ce qui pouvait attester de la mise à distance de l'objet étudié et la mise en œuvre d'une méthode rigoureuse à la fois explicitée et justifiée. Sur ce point, l'étude des réalisations montre que les acteurs du *datajournalism* consacrent au moins quelques lignes, voire un article complet pour présenter la conduite du projet : provenance des données, noms et compétences des personnes impliquées, temps nécessaire à la réalisation. Ainsi, le *Guardian* mentionne par exemple son partenariat avec la *London school of Economics* et la collaboration des journalistes avec une équipe pluridisciplinaire de chercheurs des universités de Manchester, des doctorants et des développeurs. Cependant, ce type de précisions ne permet pas toujours de vérifier le mode de collecte des données secondaires (souvent obtenues auprès de divers organismes en amont des projets) ou de valider la pertinence de la méthode qui donne naissance à la modélisation. Les articles se contentent de mobiliser des arguments d'autorité et de promotion (une équipe pluridisciplinaire de l'université, des données publiques, des compétences particulières en informatique, une durée importante). Malgré ces réserves, il faut reconnaître que le

fait de rendre visible au lecteur une partie du squelette de la production est une pratique nouvelle qui rompt avec les habitudes des journalistes.

Concernant la mise à distance, c'est-à-dire le travail d'objectivation du réel, l'étude des sites valide l'idéal de transparence et la croyance par les professionnels, dans la vérité intrinsèque des données. Ces dernières semblent garantir à elles seules une posture objective vis-à-vis des sujets traités. La donnée est mobilisée en tant que preuve (dans le cadre d'une démonstration notamment dans la catégorie investigation) ou comme élément factuel indiscutable. Dans leur croisade pour la vérité et la transparence, les concepteurs/réalisateurs se répartissent selon deux attitudes distinctes. Pour certains (*The Seattle Time*, *The Chicago Tribune*), les faits parlent d'eux-mêmes et ce qu'ils révèlent ne nécessite pas d'engagement énonciatif dans le discours. L'énonciation ne présente aucune trace de subjectivité et le registre s'apparente aux conventions normatives de l'écriture journalistique du compte-rendu factuel. Pour d'autres (*Mother Jones*), la donnée est la base d'un engagement énonciatif assumé. Loin de la neutralité énonciative et de la rigueur revendiquées par la science, le *datajournalism* adopte une posture scientifique sans renoncer à la liberté de ton et d'invention de formes. Cette posture scientifique mêle vigilance et propositions de lectures du monde, les données constituant l'ossature des réalisations et leur raison d'être.

### **Privilégier l'efficacité**

Considérant la dimension graphique et esthétique des réalisations, seul le *Guardian* ose une traduction visuelle réellement novatrice en représentant la propagation des rumeurs sur *Twitter* sous forme de cellules qui se contaminent, se multiplient, s'opposent pour finalement laisser apparaître la domination de tel ou tel camp puis la disparition parfois brutale des tweets sur le sujet étudié. On peut bien entendu questionner et critiquer la métaphore biologique et ses présupposés idéologiques mais la dimension innovante de la mise en forme pour un média d'information ne fait aucun doute. Les autres projets privilégient un graphisme clair et efficace à partir d'histogrammes, de cartes et de tableaux. La dimension esthétique voire artistique, pourtant présente dans les discours ne semble donc pas constituer un critère essentiel pour le jury, à moins que les réalisations en compétitions n'aient pas intégré l'inventivité graphique comme une contrainte du travail.

Même classiques et simples, les images censées rendre les données significatives ne sont jamais autosuffisantes. Ce constat s'impose notamment au contact des réalisations dans une langue étrangère ou dans le cas de créations graphiques complexes qui ne peuvent se passer de légendes et d'explications explicites. Si une image vaut mille mots, le texte reste généralement nécessaire pour comprendre le sens des infographies proposées. Ces dernières s'appuient sur un double codage et par conséquent sur une double abstraction : celle de la saisie du monde par des données (géographiques, économiques, juridiques, médicales) puis celle de leur traduction graphique. Chaque étape est le résultat de choix, de priorités, d'intentions qui ont deux conséquences importantes. En premier lieu, comme toute traduction, la modélisation des données n'est pas neutre. En second lieu, compte tenu de la nécessaire activité de décodage, les formes graphiques produites ne sont pas forcément intelligibles « intuitivement ».

### **Promouvoir la variété des mises en récits**

Une des premières observations que l'on peut faire en parcourant les diverses réalisations de *datajournalism* concerne l'absence de normes et de règles qui permettraient de les situer au sein d'un genre discursif identifié. La diversité des traitements et des mises en récits semblent plutôt prévaloir, les réalisations se situant entre deux pôles opposés : celui d'une mise en récit qui emprunte au genre fictionnel et celui d'une modélisation de données qui semble se suffire à elle-même. À titre d'exemple, le site de *Mother Jones*, primé dans la catégorie « investigation », ne se contente pas de laisser parler les données. L'internaute est invité à parcourir le dossier avec des formules chocs et des illustrations dans le style des couvertures de romans policiers des années 50 ou d'affiches de films d'action. Ainsi, le sous-titre du dossier est une invitation à pénétrer dans un réseau secret chargé de surveiller et piéger les américains. Quant au procédé graphique, il rappelle les méthodes des journaux de la fin du XIX<sup>ème</sup> siècle qui donnaient à voir, grâce à des dessins, des scènes de crimes et de faits divers dramatiques impossibles à photographier; une manière de rendre spectaculaire ce à quoi il n'est pas possible d'avoir accès<sup>1</sup>. Les illustrations de *Mother Jones* mettent

---

1 L'exposition « Du sang à la Une. La belle époque du fait divers dans la presse illustrée » organisée par l'Espace Patrimoine la Bibliothèque Municipale de Lyon en 1998-99 en présente quelques exemples significatifs visible sur <http://www.bm-lyon.fr/expo/virtuelles/sang/sang.htm> Dernière consultation le 24/02/2013.

en scène de façon volontairement caricaturale et stéréotypée les acteurs ainsi que les ingrédients classiques du film de guerre ou d'espionnage : des hommes en uniformes, des victimes brutalisées, des personnages « exotiques » et inquiétants, des armes, de l'argent, etc. Le traitement graphique transforme l'effet de réel habituellement produit par l'image réaliste (Tétu, 2008) en effet fictionnel. La lecture des articles montre que le texte contribue aussi à construire une mise en intrigue grâce à différentes stratégies narratives (focalisation sur des individus dont on décrit le caractère, le comportement, l'apparence comme de véritables personnages de roman, écriture de dialogues dont il est impossible de savoir si ce sont des retranscriptions d'enregistrements ou de pures digressions romanesques...). Pour *Mother Jones*, média associatif sans recettes publicitaires, la mise en intrigue permet non seulement d'enrichir les données peu bavardes par elles-mêmes et d'attirer des internautes mais aussi de renforcer son positionnement de média d'investigation qui se doit de dénoncer et de révéler des vérités cachées. A l'opposé de cette stratégie, le site *Politznet.ch* fait le choix d'un apparent effacement énonciatif et se contente de traduire visuellement les votes sur les questions débattues au parlement. Il croise différents types de données concernant les députés (prises de paroles, présence aux sessions parlementaires, historique des votes, biographies, etc.) et incite les internautes à publier commentaires et questions à destination des politiques.

À mi-chemin entre la mise en intrigue et l'absence de remédiation journalistique, plusieurs projets publiés par des médias présentent autour du traitement des données, des articles illustrés par des photographies et des vidéos élaborées dans le respect des conventions professionnelles du journalisme. L'énonciation semble neutralisée mais l'engagement se manifeste grâce à la mise en avant de faits soigneusement sélectionnés dont on amplifie le caractère scandaleux grâce à des stratégies de mise en contiguïté : rapprochement de situations contrastées, d'acteurs aux positionnements antagonistes, de propos contraires aux faits attestés...

L'observation montre que le *datajournalism* ne peut se résumer à la traduction graphique d'une grande quantité de données. Sans format ou sujet imposé, sans règle d'écriture ou d'usage, ses productions exigent d'être explorées, expérimentées pour en comprendre le fonctionnement : les parcours, les différentes entrées dans les données, les tableaux, les cartes, etc. L'utilisateur y est incité à « lire », « visionner », « sélectionner », « questionner », « cliquer ». Loin des genres discursifs conventionnels du journalisme, le *datajournalism* conçoit des

environnements qui empruntent à différents domaines et se rapprochent des univers multimédias, scénarisés et interactifs des jeux vidéos.

Avant de conclure, nous souhaitons souligner que malgré la notoriété de certains médias primés, les réalisations de *datajournalism* affichent un faible nombre d'appréciations positives et de partages sur les réseaux sociaux. Ces critères quantitatifs, bien que discutables, questionnent l'accueil réservé par les internautes aux produits du *datajournalism*. Loin de la simplicité promise, de l'utilité pour chacun et de la séduction des visuels, le journalisme de donnée semble constituer une forme expérimentale et complémentaire à l'offre d'actualité ; l'engouement auprès d'un large public reste pour l'heure très hypothétique.

### **Eléments de conclusion**

En se présentant comme une nouvelle forme journalistique, le *datajournalism* participe du dynamisme perpétuel du secteur mais probablement pas d'une redéfinition stabilisée du champ professionnel. Les acteurs n'exercent pas de pression centripète qui concentrerait les énergies vers les seuls médias et journalistes. Bien au contraire, les discours comme les réalisations sont alimentés par un souci permanent d'ouverture et de coopération entre différents domaines, professions et pratiques. L'analyse dessine l'espace social du journalisme de données déclinant la logique de projet à tous les niveaux, mobilisant le principe de la coopération. Ces modalités de fonctionnement traduisent la présence de valeurs politiques (au sens du politique et non de la politique) réinscrivant assez nettement ces formes médiatiques dans l'espace public. La dimension économique y est absente, ou en tous cas non centrale dans les projet (financés par des fondations, des dons, des programmes de recherche), au profit de l'utilité sociale, citoyenne des projets. L'étude vérifie le rôle des discours et des expérimentations qui comme le souligne Flichy (2001, p. 71) contribuent à construire « *des imaginaires collectifs, communs à des groupes de concepteurs, puis un jour aux concepteurs et aux utilisateurs. L'imaginaire est ainsi une des façons de construire une identité collective, de rompre tout d'abord avec les modèles existants, puis de légitimer la nouvelle technique et de mobiliser les différents acteurs.* » Le journalisme de données tente ainsi de donner corps au projet qui a animé le journalisme multimédia dans les années 2000 et se présente comme le mariage réussi des compétences techniques avec des préoccupations citoyennes (Croissant & Touboul, 2013). À bien y réfléchir, à l'heure des réductions de personnels dans les



salles de rédactions et de l'augmentation des cadences, le *datajournalism* apparaît comme une tentative de résister aux pressions patronales en investissant de nouveaux territoires présentés comme prometteurs. En prenant le contre-pied d'un journalisme conventionnel, les promoteurs du journalisme de données semblent adresser un message aux dirigeants d'entreprises médiatiques, les invitant à faire le pari de la qualité, du temps long et de la coopération.

### Références

- Barthes, R. (1964). Rhétorique de l'image. *Communications*, 40-51. Seuil.
- Bertin, J. (1967). *Sémiologie graphique*. Paris : Mouton/Gauthier-Villars.
- Bertin, J. (1975). *La Graphique et le traitement graphique de l'information*. Paris : Flammarion.
- Cardon, D. (2011). « Zoomer ou dézoomer ? Les enjeux politiques des données ouvertes ». Disponible à : <http://owni.fr/2011/02/21/zoomer-ou-dezoomer-les-enjeux-politiques-des-donnees-ouvertes/>
- Cleveland, W. S., & McGill, R. (1984). Graphical Perception: Theory, Experimentation, and Application to the Development of Graphical Methods. *Journal of the American Statistical Association*, 79(387), 531-554.
- Croissant, V., & Touboul, A. (2013). Journalisme et internet, 15 ans de discours d'ajustements. Dans A. Levrier et A. Wrona (Éd.), *Matière et esprit du journalisme. Du Mercure Galant à Twitter*. (pp. 253-263). Paris : PUPS.
- Dagiral, E., & Parasie, S. (2011). Portrait du journaliste en programmeur : l'émergence d'une figure du journaliste « Hacker ». *Les Cahiers du journalisme*, 22/23, 144-154.
- Flichy, P. (2001). La place de l'imaginaire dans l'action technique. Le cas de l'internet. *Réseaux*, 109, 51-73.
- Jeanneret, Y. (2007). *Y a-t-il (vraiment) des technologies de l'information*. Lille : Presses Universitaires du Septentrion.
- Joannes, A. (2010). *Data Journalism*. Paris : Editions du CFPJ.
- Neveu, E. (2009 réédition), *Sociologie du journalisme*. Paris : La Découverte.
- Parasie, S. (2011). Le journalisme « hacker ». Une nouvelle utopie pour la presse ? *La Vie des idées*. Disponible à : <http://www.laviedesidees.fr/Le-journalisme-hacker.html>
- Ringoot, R., & Utard, JM. (2006). *Le journalisme en invention. Nouvelles pratiques, nouveaux acteurs*. Rennes : PUR.
- Ruellan, D. (2007 réédition). *Le journalisme ou le professionnalisme du flou*. Grenoble : PUG.
- Treisman, A. (1985). Preattentive Processing in Vision. *Journal Computer Vision, Graphic and Image Processing*, 31(2), 156-17.
- Trédan, O. (2011). Quand le journalisme se saisit du Web : l'exemple du datajournalism. 1er colloque international - *Les mutations structurelles du journalisme* (Mejor – 2011), Université de Brasilia. Disponible à : <http://florlecam.com/slj/wp-content/uploads/2011/05/Datajournalism-Tr%C3%A9dan.pdf>