

Quand le journalisme se saisit du Web : l'exemple du *datajournalism*

Olivier Trédan
Doctorant

CRAPE – Université de Rennes 1 / IUT de Lannion / M@rsouin / REJ

Le journalisme se trouve confronté à un paradoxe. « *Jamais dans l'histoire autant de données n'ont été disponibles pour la masse des audiences. Jamais les responsables de la production et de l'analyse des informations – le journalisme – n'avaient été aussi fortement affaiblis par l'effondrement de leurs sources de financement.* » (Neveu, 2010 : 39). Ce paradoxe semble partiellement se résorber dans la capacité du journalisme à incorporer des innovations et dans la réaffirmation de ses spécificités : vérification et analyse des informations ; pouvoir des mythes professionnels.

Dans cette perspective, cette combinaison entre appropriation d'innovations et réaffirmation du pré carré journalistique a donné naissance à de nouveaux acteurs professionnels, soit chargés de la gestion de la relation à l'audience – les « *community managers*¹ » ont fait leur apparition dans les salles de rédaction françaises en 2009 –, soit de la production de nouveau support de narration – le webreportage et le webdocumentaire combinent techniques du reportage, narration hypertextuelle et maîtrise des langages informatiques, en particulier le *Flash*. Ces activités émergentes n'ont pas pour le moment été saisies par la sociologie des médias. Dans cette perspective, nous avons choisi d'analyser le *datajournalism*². Cette pratique journalistique est fondée sur le recueil, le traitement statistique et la visualisation de données sur Internet. Ce choix est justifié pour sa dimension révélatrice des enjeux liés aux transformations structurelles du journalisme. D'une part, elle se nourrit et se légitime en tant qu'activité professionnelle par la mobilisation de mythes, le rôle démocratique du journalisme ou la nature intrinsèquement participative du Web. D'autre part, elle nécessite d'observer les dynamiques d'acteurs et de saisir les nouvelles coopérations à l'œuvre entre journalistes et acteurs professionnels issus du Web, développeurs logiciels, infographistes et webdesigners, dans la production des réalisations *datajournalistiques*. Elle vient alors questionner des enjeux identitaires et de formation professionnelle.

Pratique émergente, il est pour l'heure impossible d'en mesurer l'impact. Elle demeure une pratique marginale, autant par la faiblesse numérique de ses promoteurs que celle de leurs réalisations. Elle nous intéresse toutefois parce qu'elle symbolise les passerelles jetées entre mondes du journalisme et du Web, et parce qu'elle interroge le rôle de l'innovation dans l'évolution de la pratique journalistique.

1. Une pratique collective associant journalisme et technologie

1.1 Eléments de définition

Le *datajournalism*, appelé également « *database journalism* » et francisé par « *journalisme de données* » est une pratique fondée sur le recueil, l'analyse statistique, la visualisation de données sur Internet et leur traitement journalistique. « *Le datajournalism consiste à rechercher dans les bases de données des éléments statistiques, quantitatifs, chronologiques qui, en eux-mêmes, ne sont pas des informations mais dont les relations peuvent faire apparaître des informations latentes.* » (Joannès, 2010a : 146). Autrement dit, le *datajournalism* se décompose en une chaîne d'opérations qui va du recueil de données

1 Notons ici qu'il ne s'agit que de la reprise d'une activité issue du monde du marketing. « *Dès 1997, Netgain, le livre phare de John Hagel et Arthur G. Armstrong, fixe les règles du « community management ». [...] Les « community managers » ont fortement joué sur le besoin de reconnaissance comme d'un levier afin de développer le sentiment d'appartenance des usagers.* » (Galibert, 2010 : 97-99)

2 Méthodologiquement, pour ce travail de défrichage du *datajournalism* et de ses enjeux, nous avons procédé à la lecture systématique des articles liés à cette pratique journalistique et à l'ouverture des données de deux principaux sites : *InternetActu*, la lettre d'information de la FING, *Owni.fr*, sorte d'agrégateurs des articles plus pertinents sur ces questions. Ces deux sites, par les liens proposés, ont constitué une porte d'entrée pour une investigation plus large.

à leur mise en visualisation via une interface Web, en passant par leur exploitation, au centre de laquelle se trouverait le journaliste chargé du croisement des données susceptible de faire émerger un éclairage nouveau sur un événement donné. Cette chaîne d'opérations implique des pratiques coopératives entre journalistes et non-journalistes (développeurs Web, statisticiens, etc.).

L'originalité de la démarche repose sur le recours à l'informatique autant pour le traitement algorithmique des données que pour la visualisation de l'information. Anna Daniel et alii regroupent ces évolutions sous le terme de « *computational journalism* », pour signifier « *l'application de logiciels et de technologies aux activités journalistiques, [l'appui sur] l'informatique, les sciences sociales et de la communication.* » (Daniel, Flew, Spurgeon, 2009 : 1) Si le *datajournalism* s'inscrit dans une logique ancienne de recours à l'informatique chez les journalistes, il se singularise avant tout par la médiation d'une interface Web³. Elle permet de visualiser de manière simplifiée un grand nombre de données éparses. L'interface étant manipulable par le lecteur, elle permet une personnalisation de la lecture. Ce recours aux interfaces de visualisation de données suggère de manière implicite les dynamiques de coopération qui se jouent entre les mondes du Web et du journalisme, et, à travers elles, de nouvelles coopérations à l'œuvre entre journalistes, développeurs Web, webdesigners, infographistes, statisticiens, qu'Alain Joannès désigne de manière métaphorique dans le cas du *rich média* comme un « *travail en essais* ».

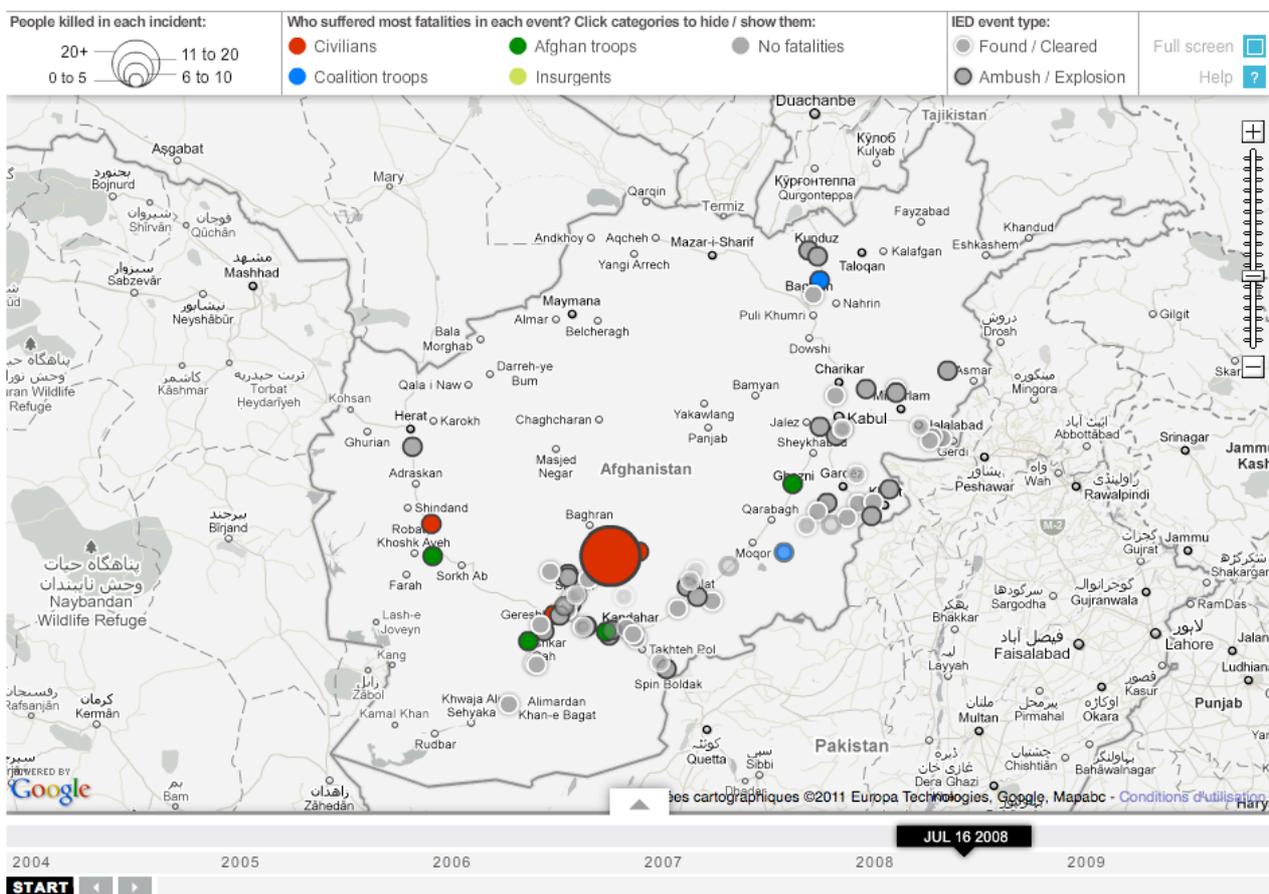


Illustration 1: Carte des 16 000 attentats commis en Afghanistan entre 2004 et 2009 montrant leur recrudescence, passant de 308 en 2004 à 7 155 en 2009 (Mark McCormick, Paddy Allen and Alastair Dant, « Afghanistan war logs: IED attacks on civilians, coalition and Afghan troops », Guardian, <http://www.guardian.co.uk/world/datablog/interactive/2010/jul/26/ied-afghanistan-war-logs>, mis en ligne le 27 juillet 2010, consulté le 15 janvier 2011.)

3 Caroline Goulard, « Quatre voies du *datajournalism* », *Le blog d'ActuVisu*, <http://blog.actuvisu.fr/quatre-voies-du-datajournalism/>, mis en ligne le 7 avril 2010, consulté le 3 février 2011.

En cela, le *datajournalism* apparaît comme une activité reposant sur la constitution d'un travail de coopération entre acteurs professionnels issus de mondes différents. Celle-ci se traduit par la mise en place de partenariats entre industriels de l'informatique et entreprises de presse. Le *New York Times* a travaillé avec *IBM* autour de sa plateforme de traitement et de visualisation de données *ManyEyes*⁴, destinée aux lecteurs⁵. À l'interne, sous l'impulsion de deux salariés, un Département des Technologies interactives a été créé. En 2010, il regroupait onze salariés, dont neuf informaticiens. Toutefois, les associations au sein des rédactions entre journalistes et développeurs Web restent une configuration marginale et ne valent que pour quelques grands titres de presse nord-américains et anglo-saxons. Ailleurs, et en France notamment, le travail de conception d'interface de visualisation est externalisée auprès de petites sociétés innovantes. La conception de l'interface de visualisation du *Monde.fr*⁶ des mémos diplomatiques fournis par *Wikileaks* a été confiée à une start-up spécialisée dans l'« analyse sociale du Web », *Linkfluence*. *La Gazette des communes*, mensuel spécialisé à destination des cadres de la fonction publique territoriale, a mené une expérimentation de *datajournalism*, dont la réalisation a été confiée à une start-up, 22 Mars, éditrice du site *Owni.fr*, espace de discussion consacré au « *digital journalism* ».

1.2 Le *datajournalism* dans la lignée du Web 2.0

Le *datajournalism* s'inscrit à l'articulation de plusieurs mondes, entre journalisme et Web. Il s'inscrit dans la filiation des évolutions technologiques regroupées sous le terme de *Web 2.0*, puis de *Web Squared*⁷, notions idéologiquement connotées. Sans rentrer dans le détail de cette vision, elle est décrite sous la forme d'un point de bascule dans l'histoire de l'Internet, caractérisée par un internaute dont la nature aurait changé sous l'effet de la technique, passant d'un état de passivité à une (sur)activité en ligne, et par une ouverture progressive de la programmation.

La perception du « Public » comme producteur de contenu a ainsi fait miroiter la possibilité d'une émergence d'un journalisme sans journaliste, où l'internaute serait le nouveau producteur de contenu, logique popularisée par les *User Generated Contents*. Une nouvelle rhétorique de la participation a vu le jour à partir de 2006, le *Crowdsourcing*⁸, désignant le recours à un grand nombre de personnes pour externaliser la réalisation de certaines tâches. Dans le cas du *datajournalism*, il ne s'agit plus d'une remise en cause du journaliste. Au contraire, il se trouve renforcé par la délégation de tâches à la foule – la lecture et le classement de documents bruts.

Par ailleurs, les promoteurs du Web 2.0 ont cherché à mettre en exergue de nouvelles modalités de production de services Web. À un modèle privé et corporatiste, s'opposerait une nouvelle « culture » de l'innovation issue du monde du logiciel libre, caractérisée par l'ouverture et l'interopérabilité. Il s'agit là

4 <http://manyeyes.alphaworks.ibm.com>

5 Selon cette logique, le *New York Times* ajoute une nouvelle brique à la rhétorique de la participation des lecteurs-contributeurs. « *The Times estime que les utilisateurs peuvent contribuer à l'interprétation de données et découvrir de façons innovantes de les présenter. Tout comme les commentaires des lecteurs renforcent notre journalisme, ces visualisations peuvent ajouter une nouvelle valeur et devenir un catalyseur pour les discussions.* » « The New York Times Data Visualization Lab », <http://open.blogs.nytimes.com/2008/10/27/the-new-york-times-data-visualization-lab/>, mis en ligne le 27 octobre 2008, consulté le 4 février 2011.

6 « WikiLeaks : naviguez dans les mémos diplomatiques », *Lemonde.fr*, http://www.lemonde.fr/documents-wikileaks/visuel/2010/12/06/wikileaks-lire-les-memos-diplomatiques_1449709_1446239.html, mis en ligne le 6 décembre 2010, consulté le 5 février 2011.

7 À l'origine des deux termes se trouve Tim O'Reilly, gourou de l'Internet et fondateur d'*O'Reilly Media*, une maison d'édition spécialisée dans l'informatique. En 2005, il définissait le Web 2.0 (Tim O'Reilly, « What is Web 2.0? », O'Reilly, <http://oreilly.com/web2/archive/what-is-web-20.html>, mis en ligne le 30 septembre 2005). En 2009, il arrive à la conclusion que « *l'ère du Web 2.0 [est] une course à l'acquisition et au contrôle de fonds de données* », sans toutefois définir de quelles données il est question, ni se préoccuper des enjeux économiques, politiques ou sociaux sous-jacents (Tim O'Reilly, John Battelle, « Web Squared: Web 2.0 Five Years On », *Web 2.0 Summit*, <http://www.web2summit.com/web2009/public/schedule/detail/10194>, mis en ligne).

8 Le terme a été proposé pour la première fois par la revue *Wired*, en 2006, sur une variation de l'outsourcing (externalisation), pour signifier la collaboration des masses extérieures aux organisations à but lucratif dans la co-production de valeur ajoutée. Howe Jeff, « The rise of crowdsourcing », *Wired*, <http://www.wired.com/wired/archive/14.06/crowds.html>, mis en ligne le 14 juin 2006, consulté le 11 juin 2009.

à notre sens une innovation dont le *datajournalism* tire pour partie son origine. Ce terme technique désigne une pratique d'ouverture partielle du code informatique des services en ligne permettant la coopération des acteurs de l'informatique et de l'Internet pour développer des passerelles entre services, voire proposer de nouveaux services. Ces derniers sont baptisés « *mashups* » et reposent sur des interfaces de programmation (API). Parmi ceux-ci, un site Web revient sans cesse dans les écrits des technophiles et observateurs du Web : *housingmaps*, service de géolocalisation de petites annonces immobilières sur Internet. Il est né de la fusion entre deux services préexistants. Le premier est un site de petites annonces en ligne, *Craigslist* ; Le second est le service de cartographie *Google Maps*. *Craigslist* forme ainsi une base de données de petites annonces que *Google Maps* permet de visualiser. Le *datajournalism* s'inscrit dans cette filiation technologique, entre contribution des internautes et programmation ouverte.

2. La réactivation des mythes par les données

Le *datajournalism* repose sur deux ressorts principaux : d'un côté, un savoir-faire technologique qui permet le développement des interfaces de visualisation des données ; de l'autre, un savoir-faire dans l'identification et l'exploitation de données disponibles. L'accès à ces données et de son traitement est devenu un enjeu autour duquel s'est cristallisé le *datajournalism* : actions de lobbying destinées à la libération des données publiques, investigation dans les données réactivant des mythes professionnels, mobilisation des internautes dans l'exploitation de volumes de données importants.

2.1 Vers la libéralisation des données publiques?

En janvier 2009, l'administration de Barack Obama a décidé l'ouverture des données administratives. Cette démarche s'inscrit dans le contexte d'un mouvement visant à la libération des données publiques, l'« *Open Data* ». Ce mouvement a pris naissance en Angleterre, en 2006. Le *Guardian* a accompagné une campagne « *Free Our Data* » visant à libéraliser les données publiques. À l'origine, se trouve la publication d'une tribune⁹ de deux journalistes britanniques, Charles Arthur et Michael Cross. Ceux-ci y dénonçaient le double coût porté par le contribuable : la vente des données produites à partir des fonds publics. En guise de réponse, le gouvernement britannique a mis en place en 2008 une commission chargée d'organiser un concours sur la meilleure manière d'utiliser les données publiques, « *Show us the better way*¹⁰ », doté de 20 000 £, pour récompenser les cinq meilleures réalisations parmi les 450 proposées. En juin 2009, le gouvernement britannique a fait appel à l'inventeur du Web, Tim Berners-Lee¹¹, pour réfléchir à l'ouverture des données publiques gouvernementales. Le portail *data.gov.uk* a été ouvert au public en janvier 2010, soit quelques mois après celui de l'administration Obama¹².

En France, le mouvement de l'« *Open Data* » n'est pas porté par les médias, mais par quelques think tanks. Un portail de base de données *Data.gouv.fr* a été annoncé, mais demeure toujours à l'état de projet. Une des pierres d'achoppement repose sur le format des licences des données publiques et leur possible commercialisation. L'administration et les collectivités locales sont incitées à participer à ce mouvement. Un think tank spécialisé dans l'Internet, la *FING* (*Fédération Internet Nouvelle Génération*), a publié en janvier 2011 un livre blanc concernant l'ouverture des données publiques territoriales¹³. Les auteurs y vantent les bénéfices d'une telle libération autant en termes de

9 Charles Arthur, Michael Cross, « Give us back our crown jewels », *Guardian*, <http://www.guardian.co.uk/technology/2006/mar/09/education.epublic>, mis en ligne le 9 mars 2006, consulté le 2 février 2011.

10 <http://webarchive.nationalarchives.gov.uk/20100807004350/http://showusabetterway.co.uk>

11 Charles Arthur, « Web inventor to help Downing Street open up government data », *Guardian*, <http://www.guardian.co.uk/technology/2009/jun/10/berners-lee-downing-street-web-open>, mis en ligne le 10 juin 2010, consulté le 10 février 2011.

12 L'initiative du portail *data.gov* de l'administration américaine a été suivie d'autres démarches similaires : Grande-Bretagne, qui propose 5 600 bases de données au 1er février 2011, la Nouvelle-Zélande (*data.govt.nz*), l'Australie (*data.australia.gov.au*), le Brésil (*ipeadata.gov.br*).

13 Amandine Brugière, Charles Népote, « Guide pratique de l'ouverture des données publiques territoriales », FING,

développement économique, qu'en termes d'efficacité administrative et de transparence politique. Pour le moment, seules quelques municipalités se sont lancées dans l'« *Open data* ». Face à cette ouverture tâtonnante, les rares *datajournalistes* se heurtent à la difficulté de l'exploitation des données publiques. « *Force est de constater que les administrations ont du mal à passer à l'informatique. [...] La palme de l'inefficacité revient à celles qui ont pris le temps de scanner des documents Word imprimés. [...] Le nouveau débat concerne [le] format et [la] facilité d'accès [des données].*¹⁴ »

2.2 Du journalisme d'investigation au crowdsourcing

Le *datajournalism* est donc émergent, notamment en France, où la question de l'ouverture des données n'est pas effective. Pratique en devenir, elle est toutefois âprement discutée, ayant réussi à se constituer une légitimité à partir de quelques faits d'armes notoires, au carrefour des mythes journalistiques et de l'Internet.

L'une des premières enquêtes de *datajournalistes* (et non pas de *computer-assisted reporters*¹⁵) a été menée par Charles Ornstein et Tracy Weber entre 2008 et 2009, pour le compte du *Los-Angeles Times* et de *Propublica*, un pure-player californien, dont le titre « *Journalism in the interest public* » résume le positionnement. L'enquête des deux journalistes s'est concentrée sur l'absence de sanction à l'égard d'infirmières coupables de fautes professionnelles. Ils se sont procuré une liste de 2 371 infirmières ayant fait l'objet d'une procédure disciplinaire depuis 2002. L'enquête a nécessité dix-huit mois d'investigation¹⁶. Les deux enquêteurs, après construction et exploitation d'une base de données, ont démontré que plus de 80 infirmières fichées avec la mention « risques pour la santé publique » avaient continué à exercer. Ils ont pointé ainsi les dysfonctionnements du *California nursing board*, organisme chargé de la régulation de la profession, dont le président a démissionné après publication de l'enquête. L'originalité de la démarche des deux journalistes est par ailleurs d'avoir proposé à leurs lecteurs les recettes utilisées pour mener leur enquête, allant jusqu'à organiser une conférence téléphonique avec leurs lecteurs¹⁷. Les deux reporters ont été nommés pour leur enquête au prix Pulitzer de 2010.

Dans le même temps, le *Guardian* a proposé une enquête *datajournalistique* qui a bénéficié d'un large écho médiatique. Début mai 2009, le quotidien *Daily Telegraph* révèle le scandale des notes de frais des parlementaires britanniques à partir de documents achetés, recueillis par des employés du *Stationery Office*. L'ensemble des données fournies regroupant plus de 700 000 documents, le « *coup de génie*¹⁸ » du *Guardian* a été de proposer une application permettant de visualiser les notes de frais et autres documents récupérés, soit 458 000 documents. 223 000 ont été observés par quelque 23 000 lecteurs du quotidien. Si la méthode peut être saluée comme permettant de découvrir « *plus de cas de corruption que tout un tas de journalistes d'investigation n'aurait pu le faire [...], donnant naissance à des analyses fouillées et critiques menées par des journalistes.* » (Neveu, 2010 : 46), elle s'inscrit dans le renouvellement d'un discours sur les formes de participation des internautes de manière périphérique à

<http://doc.openfing.org/RDPU/GuidePratiqueDonneesPubliquesv1beta.pdf>, mis en ligne le 19 janvier 2011, consulté le 5 février 2011.

14 Nicolas Kayser-Bril, « On l'a pas en format ordinateur », *Database Journalism*, <http://databasejournalism.wordpress.com/2009/12/15/%C2%AB-on-l%E2%80%99a-pas-en-format-ordinateur-%C2%BB/>, mis en ligne le 15 décembre 2009, consulté le 31 janvier 2011.

15 Sylvain Parasio et Eric Dagiral soulignent que « *depuis la fin des années 1960, des journalistes d'investigation nord-américains utilisent des ordinateurs. Cela s'est développé sous l'impulsion de Philip Meyer autour de l'expression Computer-Assisted Reporter. Bien avant l'informatisation des rédactions, il s'agissait d'appliquer des méthodes statistiques, issues des sciences sociales, et d'utiliser l'outil informatique dans le cadre de l'enquête journalistique.* » (Parasio, Dagiral, 2011 : 3-4)

16 Il est nécessaire de souligner l'originalité du modèle économique de *Pro Publica*, « *lancé par Paul Steiger, ancien directeur du Wall Street Journal, grâce aux fonds d'un couple de milliardaires de la finance à la retraite, les Sandler.* » (Couve et alii, 2010 : 79)

17 Charles Ornstein et Tracy Web, « Track Your State Nursing Board », <http://projects.propublica.org/state-boards/subjects/nursing#recipe>, mis en ligne le 30 juin 2010, consulté le 2 février 2011.

18 Eric Mettout, « Les renforts du Guardian », *Nouvelle formule*, http://blogs.lexpress.fr/nouvelleformule/2009/06/19/attention_sujet_complexe_sensi/, mis en ligne le 19 juin 2009, consulté le 3 février 2011.

la production de l'information, qui légitime un savoir-faire dans la relation au « Public », désignée sous le terme de Crowdsourcing.

Investigate your MP's expenses: Home → Conservative → Douglas Hogg → Incidental Expenses Provision 2007/08

- Pick yourself a username to see all of your contributions on a page (privacy policy) [Get username](#)

Page 36 of Douglas Hogg's Incidental Expenses Provision 2007/08

This document has 71 pages (see all) [← Previous](#) | [Next →](#)

your vodafone bill

Quick breakdown

Service charges	
Anytime 75 (Apr-2008)	£17.02
Reserved Billing	£1.28
Video Calling	£0.00
Usage charges up to 75 Aug	
Calls	£18.50
Calls	£61.29
Total before VAT	£78.69
VAT on 20% bill number 301074460	£6.44
Total	£104.53

by Direct Debit on or after 09 October 2007

Your price plan

Anytime 75 minutes (Anytime 75)

+ 150 inclusive minutes to use this month

75 rolled over from your last bill plus 75 in your price plan

What kind of page is this?

Claim

An expense form

Proof

Receipt, invoice or purchase order

Blank

Nothing to see here

Other

Something we haven't thought of

Is this page interesting? Should we investigate further?

Not interesting

e.g. a coversheet or stationery

Interesting

It's significant expenses data

Interesting but known

e.g. it's a duckhouse

Investigate this!

I would like to know more

[Go to next unreviewed page](#)

Enter individual line items here

You need to indicate if this page contains **claims** or **proofs** before you can add line items.

Page 1 of 4

Observations and coverage

Illustration 2: Copie d'écran de l'application du Guardian permettant aux lecteurs de lire et trier les notes de frais des parlementaires britanniques. Deux possibilités sont offertes : renseigner la nature des documents et leur intérêt parmi quatre choix proposés.

Investigation dans les données, publiques ou volées, et appel à la contribution des lecteurs constituent deux ressorts de légitimation de cette nouvelle pratique, dont un des enjeux peut être lu comme une volonté de ne pas réduire le *datajournalism* au seul développement d'interfaces de visualisation de données, et par là réaffirmer la dimension plus journalistique qu'infographique de l'activité.

3. Une pratique d'avenir?

Telle que présentée jusqu'à maintenant, l'ampleur du *datajournalism* demeure difficile à cerner. Il se légitime autour de quelques expériences, mises en avant pour affirmer les potentialités d'une nouvelle pratique journalistique. Nous voulons maintenant souligner, à travers l'exemple français, les enjeux d'une pratique portée par quelques jeunes journalistes.

3.1 Des jeunes et des start-ups

La pratique et la promotion du *datajournalism* est limitée pour le moment en France à trois espaces. Le premier est formé par un couple journaliste/développeur Web, David Castello-Lopes et Pierre Bance, auteurs de la « carte interactive¹⁹ » du chômage en France depuis 1982, publié par *Lemonde.fr*, vendue pour moins de 3 000 euros. Le premier, âgé de 29 ans, est diplômé de l'Institut Français de Presse, et s'est spécialisé en journalisme multimédia à l'Université de Berkeley, où il a « *appris à construire un sujet multimédia, depuis sa conception graphique jusqu'à l'animation en Flash en passant par la prise de vue et le montage.*²⁰ » Quant à Pierre Bance, après des études d'agronomie, il s'est reconverti dans le développement Web. La formation de ce duo est liée à une amitié, tous deux scolarisés dans un même lycée parisien.

Un autre espace de développement de la pratique du *datajournalism* a pris corps dans un projet étudiant, autour d'une étudiante diplômée de l'IEP de Rennes de 24 ans, Caroline Goulard, achevant un cycle de formation par un Master de Management et Nouvelles Technologies à HEC. Dans ce cadre, elle pilote le site de *datajournalism Actuvisu.fr*, lancé le 2 juillet 2010. Ce dernier présente la particularité de regrouper trois étudiants d'un Master Management des médias écrits de l'IEP de Rennes et sept étudiants en quatrième année à l'HETIC, école supérieure spécialisée dans l'Internet. *ActuVisu* propose pour le moment trois réalisations de *datajournalism*, dont une est consacrée à la mise en corrélation des performances des coureurs du Tour de France et des cas de dopages. D'autres réalisations ont été effectuées dans le cadre d'un stage au sein de la rédaction de *L'Express*. Le projet étudiant *DataVisu* a évolué fin 2010 vers une start-up *Dataveyes*, fondée par quatre des participants au projet initial, comprenant une journaliste, un statisticien, un développeur Web et un gestionnaire.

Enfin, troisième espace identifié, le site *Owني.fr* se démarque tant pour ses réalisations que comme un espace de discussion autour des enjeux du *datajournalism* et comme le suggère, le sous-titre du site, du « digital journalism ». Le site est un espace de ressources en ligne, reprenant les publications jugées intéressantes de quelques 900 blogueurs. Il revendique 250 000 visiteurs uniques par mois. *Owني.fr* s'adosse à une jeune start-up lancée en février 2008, baptisée *22 Mars*. Elle est l'oeuvre d'un « *entrepreneur, pas un journaliste*²¹ » de 32 ans, Nicolas Voisin, qui s'est fait connaître à partir de 2006 dans le monde des blogs politiques, en lançant le *PolITIC'show*, blog consacré à l'interview des hommes politiques durant la campagne de 2007 à l'élection présidentielle. L'équipe de *22 Mars* est composée d'une vingtaine de membres, répartie entre journalistes, développeurs Web et infographistes.

Owني.fr est devenu en l'espace de quelques mois la référence en termes d'innovations journalistiques, mobilisée par quelques sociologues de l'Internet²². Il a tiré son épingle du jeu dans l'affaire *Wikileaks*.

19 David Castello-Lopes, Pierre Bance, « Le chômage en France depuis 1982 », *Lemonde.fr*, http://www.lemonde.fr/societe/visuel/2010/03/03/le-chomage-en-france-depuis-1982_1310908_3224.html, mis en ligne le 3 mars 2010, consulté le 5 février 2011.

20 Patrick Sarradeil, « Comment devenir journaliste plurimédias en 480 heures de cours », Communiqués de Presse, <http://www.communique-de-presse.info/devenir-journaliste-plurimedias/>, mis en ligne le 15 juin 2010, consulté le 15 février 2011.

21 Frédérique Roussel, « Owني secoue la galaxie de l'info », *Liberation.fr*, <http://www.liberation.fr/medias/01012300945-owni-secoue-la-galaxie-de-l-info>, mis en ligne le 8 novembre 2010, consulté le 25 janvier 2011.

22 Nous pensons ici à Dominique Cardon (Cardon, 2010) et Patrice Flichy (Flichy, 2010) qui mentionne *Owني.fr* dans la promotion de leurs essais respectifs, et non dans ces derniers. Ainsi, le premier affirme que « *le monde du journalisme avec Médiapart, Rue 89, OWNI ou l'univers émergent du data journalism sont en train de montrer que de ce mélange peut naître des formes éditoriales, des contenus et des ressorts critiques nouveaux.* » (Hubert Guillaud, « Dominique Cardon : Pourquoi l'internet n'a-t-il pas changé la politique? » *InternetActu*, <http://www.internetactu.net/2010/10/07/dominique-cardon-pourquoi-linternet-na-t-il-pas-change-la-politique/>, mis en ligne le 7 octobre 2010, consulté le 25 janvier 2011) Patrice Flichy évoque « *le journalisme de données, qui se*

Owni.fr est intervenue en développant une application permettant de visualiser les données récupérées, et en s'associant au pure-player *Slate.fr*²³ et au mensuel *Le Monde Diplomatique*²⁴. Permettant à tout internaute de consulter les données fournies, la démarche est présentée sous le titre d'une « *enquête participative européenne* ». Sur les 75 000 documents fournis, 1 200 ont été lus. Toutefois, la mise en ligne de cette application a contribué à la consécration d'*Owni.fr*. En octobre 2010, il a été primé par la *Online News Association*. Le même mois, l'équipe a été contactée par Julian Assange pour une nouvelle version de l'application permettant de visualiser les mémos diplomatiques²⁵.

3.2 Vers un « journaliste hacker » ?

Les expérimentations de *datajournalism* sont si peu nombreuses que les quelques conclusions énoncées le sont sans doute trop rapidement. La première concerne « *une espèce hybride est en train d'apparaître sur un marché de l'emploi déjà sous tension*²⁶ », un journaliste jouissant de savoir-faire, voire maîtrisant le codage informatique. C'est une récurrence que l'on peut trouver dans les différents profils croisés dans les trois espaces d'expérimentation français : il s'agit de jeunes journalistes caractérisés par leur goût pour l'informatique et le Web.

Derrière ces quelques pionniers, se profile un enjeu en termes de formation professionnelle. Les quelques *datajournalistes* évoqués ont également comme caractéristique, à l'image des autodidactes du *New York Times* (Joannès, 2010b : 149) ou de Nicolas Kayser-Bril, responsable du pôle *datajournalism* à *Owni.fr*, d'avoir « *en autodidacte appris à faire un site Web, puis à coder. Je me suis aperçu que cela pouvait être une compétence utile quand on veut devenir journaliste. [...] J'ai fait Science Politique et je n'ai jamais eu de cours d'informatique.*²⁷ » Quelques écoles de journalisme commencent à se saisir de la question. Aux avant-postes se trouvent la *Columbia University* qui enseigne de concert le journalisme et l'informatique et la *Medill Scholl of Journalism* qui se propose de former des développeurs Web au journalisme. Cette dernière formation est proche d'un think tank « Hacks-Hacker », qui oeuvre à la mise en place de coopération entre développeurs et journalistes. Cette question, dans les écoles de journalisme françaises (Le Cam, Trédan, 2008), demeure encore peu soulevée. La voie empruntée diffère toutefois sensiblement du modèle du « journaliste hacker », jouissant à la fois de la double casquette de journaliste et de programmeur. Elle passe moins par l'acquisition de compétences que par une sensibilisation à la technique, pour que puisse s'engager « *un dialogue avec les gens de la technique pour voir comment on peut faire. Et le dialogue est intéressant quand il est équilibré.* » (Le Cam, Trédan, 2008 : 106) C'est en particulier cette capacité qui caractérise le *datajournalism*, une aptitude à jouer sur les frontières par des dynamiques de coopération entre des acteurs professionnels issus de mondes distincts.

spécialise dans le traitement journalistique de grosses bases de données, comme le fait, en France le site Owni.fr, offre des sources brutes à l'amateur d'informations » (Patrice Flichy, « La réhabilitation du journalisme d'expertise », *Lemonde.fr*, http://www.lemonde.fr/idees/article/2010/12/03/la-rehabilitation-du-journalisme-d-expertise_1448556_3232.html, mis en ligne le 3 décembre 2010, consulté le 25 janvier 2011).

23 Slate.fr, « Une application pour enquêter sur les documents Wikileaks », <http://www.slate.fr/story/25467/wikileaks-afghanistan-application-enquete-participative>, mis en ligne le 27 juillet 2010, consulté le 25 janvier 2010.

24 Philippe Rivère, « Rapports explosifs sur la guerre en Afghanistan », *Le Monde Diplomatique*, <http://www.monde-diplomatique.fr/carnet/2010-07-26-Rapports-explosifs-sur-la-guerre>, mis en ligne le 26 juillet 2010, consulté le 25 janvier 2011.

25 Nicolas Kayser-Bril, « Allô, c'est Julian Assange », *Owni.fr*, <http://owni.fr/2010/10/22/julian-assange-wikileaks-irak/>, mis en ligne le 22 octobre 2010, consulté le 25 janvier 2011.

26 Sabien Blanc, « Journaliste programmeur, le mutant parfait? », *Owni.fr*, <http://owni.fr/2010/02/23/journaliste-programmeur-le-mutant-parfait/>, mis en ligne le 23 février 2010, consulté le 25 janvier 2011.

27 Tatiana Kalouguine, « Nicolas Kayser-Bril, data-journaliste », *La voix du Dodo*, <http://www.lavoixdudodo.info/2009/11/05/interview-nicolas-kayser-bril-datajournaliste/>, mis en ligne le 5 novembre 2009, consulté le 21 janvier 2011.

Conclusion

Nous avons choisi le *datajournalism* pour illustrer certaines dynamiques de changement à l'œuvre dans l'espace des pratiques journalistiques, lorsque l'on regarde le journalisme au prisme de ses frontières avec le Web. Pratique très récente, et s'affichant comme l'avenir du journalisme²⁸, nous avons cherché à en délimiter ses contours sans prophétie. Une difficulté, et non des moindres, demeure l'accès aux sources, aux données publiques.

L'enjeu, à notre sens, se situe du côté des dynamiques d'innovation. Saisir les changements qui affectent le journalisme, dans la dynamique d'expansion qui semble le caractériser, nous paraît intelligible au prisme des outils conceptuels fournis par l'interactionnisme symbolique et par une sociologie de l'innovation qui s'inspire de ce courant. Le *datajournalism* apparaît comme un « *objet frontière* » (Flichy, 1995), c'est-à-dire placé à l'intersection de plusieurs mondes sociaux, celui du journalisme et de la programmation notamment. Le *datajournalism* est suffisamment flexible – parce que mal défini – et suffisamment robuste – parce qu'il se réclame du journalisme, voire entend incarner son avenir –, il facilite la coopération entre acteurs aux vues et aux savoir-faire différents, sans que chacun ne renonce à ses compétences propres, tout en adoptant une approche commune. Le *datajournaliste* s'allie ainsi au statisticien chargé d'extraire les données pertinentes, au développeur Web chargé de la production d'une interface ergonomique, voire cherche à faire coopérer le lecteur dans l'analyse des documents bruts. Le journalisme apparaît comme suffisamment flexible pour parvenir à s'augmenter en captant de nouvelles médiations, de nouveaux outils, de nouveaux acteurs œuvrant à leur manière à la collecte et la mise en scène de l'information, et ainsi à redéfinir en permanence ses contours. Cette redéfinition n'est pas le seul fait de nouvelles formes de coopération au sein des rédactions. Pour se légitimer dans l'espace journalistique, ces nouvelles pratiques doivent être discutées, faire l'objet de controverse. L'essor du *datajournalism* s'accompagne aussi de nouveaux lieux de débat, de réflexion et de nouveaux porte-parole. Ceux-ci réactivent les mythes professionnels autour de l'investigation, tout en y juxtaposant des revendications et des potentialités portées par les acteurs d'autres mondes. Les promoteurs du *datajournalism* s'appuient sur le mouvement de l'*Open Data*, militant en faveur de la libéralisation des données publiques et de la transparence politique. Face aux masses de données ouvertes, ils réinvestissent le thème de la participation des lecteurs au tri de l'information, déjà largement débattu sur le Web²⁹.

En élargissant la focale, nous voyons se profiler un mouvement de légitimation des espaces de production de l'information en ligne, mouvement auquel participe le *datajournalism*. Pour saisir les changements structurels du journalisme au prisme de la légitimation de ces nouvelles pratiques, il est nécessaire d'observer conjointement les différentes coopérations à l'œuvre entre journalistes et non-journalistes aux frontières du Web et du journalisme et les controverses qui accompagnent leur développement. Incidemment, une telle approche appelle à se saisir des acquis de la sociologie de l'innovation pour l'étude du journalisme.

28 Eric Mettout, « Pourquoi le data-journalisme, c'est l'avenir en marche », *Nouvelle formule*, http://blogs.lexpress.fr/nouvelleformule/2010/04/07/elle_sappelle_caroline_goulard/, mis en ligne le 7 avril 2010, consulté le 3 février 2011.

29 Nous pensons ici, parmi la littérature technophile abondante sur le thème de l'intelligence des foules, à l'essai de Don Tapscott and Anthony D. Williams, consacré à « *l'art et la science de la collaboration* » sur Internet (Tapscott, Williams, 2006 : 18). Les premières pages de l'ouvrage sont dédiées à la démarche de Rob Mc Ewen, PDG de *Goldcorp*, société d'extraction minière gérant une mine d'or au Canada, qui a sollicité l'aide des internautes pour mener des prospections en ouvrant ses données géologiques (Tapscott, Williams, 2006 : 6-10).

Références bibliographiques:

Cardon Dominique (2010), *La démocratie Internet. Promesses et limites*, Paris, Seuil, col. La République des idées, 101 p.

Couve Philippe, Kayzer-Bril Nicolas, Senant Marion, « Médias: nouveaux modèles économiques et questions de déontologie », Rapport financé par l'Alliance internationale de journalistes et avec l'appui de l'Initiative pour repenser l'économie, novembre 2010, 107 p. [http://www.journaliste-entrepreneur.com/wp-content/uploads/2010/11/medias_nouveaux_modeles_PhCouve.pdf]

Daniel Anna, Flew Terry, Spurgeon Christina (2010), « The promise of computational journalism », Media, Democracy and Change: Refereed Proceedings of the Australian and New Zealand Communications Association Annual Conference, Canberra, 7-9 Juillet 2010, [http://www.anzca.net/component/docman/doc_download/416-the-promise-of-computational-journalism.html]

Flichy Patrice (1995), *L'innovation technique, récents développements en sciences sociales, vers une nouvelle théorie de l'innovation*, Paris, La Découverte, coll. Sciences et société, 250 p.

Flichy Patrice (2010), *Le sacre de l'amateur. Sociologie des passions ordinaires à l'ère numérique*, Paris, Seuil, col. La République des idées, 96 p.

Galibert Olivier (2010), « Le marketing tribal électronique : vers une rationalisation sociale généralisée de la socialisation en ligne ? », Actes du colloque « Culture et identités dans les communications marketing », 78e congrès de l'ACFAS, UQAM, 11 mai 2010, p. 94-103

Joannès Alain (2010a), « La formation des journalistes face à l'innovation technique », Les Cahiers du Journalisme, n° 21, p. 146-155

Joannès Alain (2010b), *Data journalism. Bases de données et visualisation de l'information*, Paris, CFPJ Editions, 169 p.

Le Cam Florence, Trédan Olivier, (2008), « Journalisme et web : quels outils de formation ? », in *MédiaMorphoses*, octobre 2008, pp. 105-112

Neveu Erik (2010), « News without journalists. Real threat ou horror story? », *Brazilian Journalism Research*, vol. 6, n° 1, p. 29-54

Parasie Sylvain, Eric Dagiral, « Des « journalistes hackers » dans la rédaction. Ce qu'une identité professionnelle émergente fait à la production de l'information au *Chicago Tribune* », communication présentée au colloque *Le journalisme, une activité collective : formes, acteurs, pratiques, enjeux, 1er colloque du GIS Journalisme*, Paris, 16-18 mars 2011

Tapscott Don, Williams Anthony D., *Wikinomics, how mass collaboration changes everything*, New York, Portfolio, 2006, 324 p.