

# DATA JOURNALISM : POURQUOI LES MÉDIAS FRANÇAIS NE S'Y METTENT PAS ?

LE 8 MARS 2010 CAROLINE GOULARD

Contrairement aux pays anglo-saxons, ce type de journalisme est quasiment absent en France. Plusieurs facteurs se combinent pour expliquer ce fait, certains relèvent des rédactions, d'autres de leur environnement.



Pourquoi les médias français se sont-ils si peu saisis du data journalism, à la différence des médias anglo-saxons ? Quelques éléments de réponses ont déjà été apportés : par **Valérie Peugeot** sur [www.lavoixdudodo.info](http://www.lavoixdudodo.info) et par **Elodie Castelli** sur [www.journalismes.info](http://www.journalismes.info). Après **les études de cas**, je vous livre ici ma synthèse. L'occasion de vous faire partager les enseignements tirés de cinq entretiens, réalisés en janvier dernier avec **Hubert Guillaud**, **Jean-Marc Manach** et **Charles Népote** de la **Fing**, avec **Fabrice Epelboin** de **RWW France** et avec **Nicolas Voisin** de **Owni.fr**.

Comment expliquer le peu d'empressement des rédactions françaises à s'emparer du journalisme de données ? Plusieurs facteurs se combinent, certains relèvent des rédactions, d'autres de leur environnement

**Côté rédactions :**

## DES RÉDACTIONS EN MANQUE DE MOYENS FINANCIERS...

Tout d'abord, côté rédactions traditionnelles, la plupart consacrent très peu de ressources à la R&D, et donc à du journalisme d'expérimentation, comme de la visualisation de données. La presse quotidienne n'en a tout simplement pas les moyens, les pure players difficilement. La presse magazine ou le secteur audiovisuel pourraient peut-être parier sur le journalisme de données, mais la crise économique ne les incite pas à de tels investissements.

Quelques exceptions néanmoins : l'Express.fr a recruté deux documentalistes pour réfléchir sur la structuration de données (**plus d'info sur le blog d'Eric Mettout**) ; France 24 mène un gros travail autour du Web sémantique au sein de son Lab (**plus d'info sur le blog de Mikiane**)

## ... EN MANQUE DE MOYENS HUMAINS

Les rédactions ne sont pas seulement appauvries sur le plan financier, elles manquent aussi

de ressources humaines. Car le data journalism nécessite du temps et des compétences : en datamining, en statistiques, en développement, en web-design, en interaction design, en sémiologie visuelle...

Actuellement, personne en France n'a réussi à réunir le bon mix de compétences. Pourtant, c'est sans doute ce qui fait le succès des visualisations du **nytimes.com** depuis deux ans : le titre fait travailler ensemble des ingénieurs, des infographistes et des journalistes, tous payés le même salaire et sur un même pied d'égalité. Rien à voir avec l'état d'esprit des rédactions françaises, dans lesquelles les « informaticiens » sont déconsidérés.

Ce cloisonnement des rédactions est sans doute un peu moins prégnant lorsqu'on s'intéresse aux rédactions web, mais il n'en reste pas moins un frein au développement du data journalism en France.

## ... EN MANQUE DE CULTURE WEB

Tout simplement, les rédactions traditionnelles n'ont souvent pas l'intuition du data journalism. La plupart du temps, elles ont un train de retard par rapport aux développements du web. Les écoles de journalisme commencent juste à intégrer le journalisme d'innovation et le web dans leurs enseignements. Pour beaucoup des acteurs de ce secteur, cela reste encore un truc de « geek ».

## ... EN MANQUE D'APPROCHE STATISTIQUE

Ce manque d'intuition n'est pas sans rapport avec une culture journalistique très française. Une certaine hagiographie du journalisme made in France prend racine dans l'opposition « faits vs fiction » : opposition entre le journalisme de faits à l'anglo-saxonne et le journalisme littéraire et d'opinion du pays d'Albert Londres. La mythologie journalistique française sacralise la belle plume et le subjectivisme. Sur ce terreau pousse la défiance de nombreux journalistes envers tout ce qui pourrait paraître trop rationaliste, trop technophile ou trop américain.

A ceci s'ajoute la faible culture mathématique, statistique et scientifique de bien des rédacteurs de presse généraliste.

Aversion à mettre les mains dans les données brutes, malaisance avec les valeurs chiffrées, crainte de voir les techniciens commander les rédactions : autant de sensations diffuses qui ne facilitent pas la reconnaissance du data journalism en France.

Pour trouver quelques affinités entre la visualisation de données et l'histoire française, il faut sortir du champ journalistique et se pencher sur celui de la sémiologie. En particulier, la **sémiologie graphique**, inventée en France par **Jacques Bertin**, aborde les problématiques de visualisation d'informations géographiques.

## DES JOURNALISTES AMÉRICAINS AU SERVICE DES COMMUNAUTÉS LOCALES ?

Enfin, une dernière hypothèse pourrait expliquer l'affinité des médias anglosaxons avec le data journalism. Les journalistes américains se considèrent peut-être plus comme étant au service d'une communauté.

Aux États-Unis, le journalisme de données s'est beaucoup développé à l'échelon local avec du crimemapping et des services pratiques (les horaires d'ouvertures des magasins, par exemple). La référence en la matière reste **EveryBlock** d'**Adrian Holovaty** : un « agrégateur-visualiseur » de données micro-locales (critiques de restaurants, prix de l'immobilier, etc.).

Les données jouent un rôle important dans la valorisation des territoires. Le journalisme de données, au niveau hyperlocal, peut ainsi être utilisé par les rédactions pour générer de la proximité avec les communautés d'habitants pour lesquelles elles travaillent.

### Côté environnement :

Une autre dimension doit être prise en compte : le journalisme de données ne dépend pas uniquement des journalistes, mais également des données à leur disposition.

# UNE CULTURE DE LA TRANSPARENCE DIFFÉRENTE ENTRE LA FRANCE ET LES PAYS ANGLO-SAXONS

Et, là aussi, la France est à la traîne par rapport aux Anglo-Saxons. Les États-Unis et la Grande Bretagne se sont illustrés par leurs mouvements d'ouverture des données : avec les sites gouvernementaux **data.gov** et **data.gov.uk**, mais aussi avec de puissants militants de la cause de l'opendata, la **Sunlight Foundation** aux États-Unis, et le **datablog du Guardian** en Grande Bretagne.

Ici encore, on pourrait invoquer un fossé culturel : la culture anglo-saxonne de la transparence dans la gestion publique n'a pas d'équivalent en France. La campagne « **Give us back our crown jewels** », portée par le Guardian en 2006, ne pourrait pas avoir d'écho sur nos terres républicaines. Pourtant elle a joué un rôle important dans la libération des données publiques en Grande Bretagne. Le Guardian a ainsi activement milité pour que les données collectées grâce à l'argent du contribuable britannique soient accessibles gratuitement, afin de stimuler l'innovation. Il a joué un rôle d'exemplarité en ouvrant ses propres bases de données (**DataStore**) et en organisant un concours d'applications basées sur ces données libérées. (Voir à ce sujet l'**article de Jean Marc Manach « Les joyaux de la couronne n'appartiennent à personne »**)

## PAS DE CONSENSUS SUR LA VALEUR ÉCONOMIQUE DE L'OUVERTURE DES DONNÉES EN FRANCE

Dans son plaidoyer pour l'ouverture des données, le Guardian insistait sur l'enjeu économique de l'ouverture des données publiques : une meilleure valorisation des données stratégiques, plus de services, plus d'opportunités commerciales, plus d'innovation, moins d'asymétrie informationnelle et donc des marchés plus efficaces, et au final, un plus grand dynamisme économique.

En France, il n'y a pas de consensus sur la valeur économique des données publiques. Les entreprises dont l'activité repose sur la privatisation de données n'ont pas intérêt à leur ouverture. L'avantage économique de la libération des données ne s'est pas imposé face aux gains espérés de leur monétisation via la vente de licences. C'est ainsi, par exemple, que l'**IGN** continue à faire payer l'accès à une importante partie de ses bases de données. (**voir ce précédent post pour plus d'info**)

## LES CONDITIONS JURIDIQUES DE LA RÉUTILISATION DES BASES DE DONNÉES

Sans aller jusqu'à dire que l'appareil juridique français est un frein à l'ouverture des données, il faut garder en tête certaines particularités de notre droit des bases de données.

Premier point : le droit d'auteur. Les données brutes ne sont pas, en elles-mêmes, soumises au droit d'auteur mais une base de données peut-être protégée par le droit d'auteur si elle découle d'un acte de création intellectuelle, c'est à dire si elle témoigne d'une originalité caractérisée. L'auteur de la base de données jouit alors d'un monopole d'exploitation de son œuvre (droits patrimoniaux) ainsi que de droits au respect de l'intégrité de son œuvre et au respect de sa paternité sur l'œuvre (droits moraux).

Deuxième point : le droit des producteurs de bases de données. Lorsque que la création d'une base de données nécessite un investissement financier, humain et matériel substantiel, le droit des producteurs reconnaît au créateur un droit de protection analogue à celui de l'auteur sur son œuvre. Cette disposition est destinée à protéger l'investissement des personnes qui prennent l'initiative et le risque économique de créer une base de données. Cette protection garantie au producteur de la base de données un monopole d'exploitation, il peut interdire toute extraction substantielle de données de sa base.

Pour plus de détails voir la très bonne **synthèse de Didier Frochot sur les-[infostrateges.com](http://infostrateges.com)** et l'**étude de la jurisprudence par Bernard Lamon**.

Troisième point : la CNIL et les dispositions relatives à la protection de la vie privée. Toute base de données impliquant des données nominatives et/ou personnelles doit faire l'objet d'une déclaration à la CNIL. La collecte et la conservation des données d'une telle base

sont également soumises à conditions (**voire le site de la CNIL pour plus d'info**). De même, doit être soumis à la CNIL tout croisement de bases de données qui aboutirait à qualifier des données personnelles.

## L'ENJEU DE LA STRUCTURATION DES DONNÉES

Enfin, l'ouverture des données repose avant tout sur un enjeu d'accessibilité. Certes, on trouve aujourd'hui de nombreuses données chiffrées sur des organismes publics ou privés. Mais bien souvent ces données sont perdues au milieu de fichiers pdf, éparpillées entre des masses de texte, scannées en format image... et lorsqu'il s'agit de croiser de bases de données, on se retrouve face à des formats disparates et peu malléables... bref, les données sont rarement structurées.

D'accord, la **loi du 17 juillet 1978** reconnaît à tout citoyen français le droit d'obtenir communication des documents détenus par une administration.

D'accord, une autorité administrative dédiée, la **CADA** (commission d'accès aux documents administratifs), veille au bon respect de ce droit d'accès aux documents administratifs.

Mais rien n'oblige les administrations à communiquer leurs données sous format numérique, encore moins de façon structurée.

Sur ce sujet, l'expérience de **Nicolas Kayser-Bril** est édifiante (voir l'article « **On l'a pas en format ordinateur** »).



**Billet initialement paru sur Database journalism**

**Illustration :** <http://www.sxc.hu/>



### KATSEYE

le 8 mars 2010 - 13:20 &bullet; SIGNALER UN ABUS - PERMALINK



*Excellente synthèse.*

*La caste journalistique française a peur de se faire dépecer de ce qu'elle croit être encore son "aura".*

VOUS AIMEZ



0

VOUS N'AIMEZ PAS



0

LOUI RÉPONDRE

### PIGEONPIGEON

le 8 mars 2010 - 16:40 &bullet; SIGNALER UN ABUS - PERMALINK



*Ils ne s'y mettent pas vite, mais ils s'y mettent : pour exemple récent, une carte interactive du chômage publiée sur LeMonde.fr, et réalisée en externalisant récolte de données, développement et mise en forme à un journaliste et un développeur : [http://www.lemonde.fr/societe/visuel/2010/03/03/le-chomage-en-france-depuis-1982\\_1310908\\_3224.html](http://www.lemonde.fr/societe/visuel/2010/03/03/le-chomage-en-france-depuis-1982_1310908_3224.html)*

VOUS AIMEZ



0

VOUS N'AIMEZ PAS



0

LOUI RÉPONDRE

### CAROLINE GOULARD

le 8 mars 2010 - 21:39 &bullet; SIGNALER UN ABUS - PERMALINK



*@PigeonPigeon : merci, je n'avais pas vu passer cette carte interactive du monde.*

*Pouvez vous m'en dire plus : qui l'a réalisée ? comment ? Vous dites "en externalisant" : il ne s'agissait pas de journaliste du Monde.fr ?*

VOUS AIMEZ



0

VOUS N'AIMEZ PAS



0

LUI RÉPONDRE

**NICOLAS VOISIN**

le 9 mars 2010 - 1:21 &amp;bullet; SIGNALER UN ABUS - PERMALINK



@Caroline : débattre avec @PigeonPigeon est un privilège rare que seul le web ou la folie procurent ;)

et enchanté, cher contributeur ailé (je "plussois" à la question de Caroline ;-)

VOUS AIMEZ



0

VOUS N'AIMEZ PAS



0

LUI RÉPONDRE

**ALAIN CHAUMET**

le 9 mars 2010 - 8:33 &amp;bullet; SIGNALER UN ABUS - PERMALINK



2 commentaires, "Côté environnement" tout d'abord.

La phrase citant l'IGN, "c'est ainsi, par exemple, que l'IGN..." illustre une phrase précédente affirmant que "les entreprises dont l'activité...". Illustration à mon avis mal à propos, l'IGN étant non une "entreprise" ayant pour objectif de réaliser "des gains", mais un établissement public auquel ses tutelles demandent d'équilibrer un budget tout en assurant des missions clairement définies. La tarification des services est dans ce cadre une conséquence logique comme l'est la tarification d'une série de services publics qu'il est ici inutile de rappeler.

Ensuite sur la demande d'une meilleure "structuration des données" proposées par l'administration, il est sans doute plaisant de faire sourire en proposant en référence un article un peu moqueur à l'endroit d'un agent public peu au fait de la chose informatique, toutefois, il me semble qu'une telle présentation du sujet appelle deux remarques :

- premièrement, ranger les rieurs derrière soi pour moquer son prochain, fut-il coupable d'être fonctionnaire, ne met nullement à l'abri d'un effet boomerang, éventuellement dévastateur; pour ma part, j'adopte à l'égard de toutes les professions éloignées de ma spécialité la plus grande réserve, surtout lorsqu'il s'agit de lancer des passerelles pour échanger plus et mieux,

- en second lieu, s'il est incontestable qu'il est préférable de disposer de bases de données et non de documents numériques, mais sourds aux interrogations, je n'en remarque pas moins que la quantité de documents lisibles par tous, et non seulement par les seuls spécialistes, est de plus en plus importante et inversement massivement sous exploitée.

Pour fournir sur ce dernier point un exemple malheureusement d'actualité, je constate que le DDRM de Vendée :

[http://www.vendee.pref.gouv.fr/sections/thematiques/securite/securite\\_civile/risques\\_naturels\\_et/dossier\\_departement/nocache=1204652305.04](http://www.vendee.pref.gouv.fr/sections/thematiques/securite/securite_civile/risques_naturels_et/dossier_departement/nocache=1204652305.04)

qui signale page 23 de manière éminente la situation de certaines communes du littoral à l'égard du risque d'inondation par submersion marine, ne semble pas avoir été examinée plus par les commentateurs que par les promoteurs, pétitionnaires ou instructeurs de permis de construire.

Hors la voie d'avenir de services en ligne qui permettront de croiser à volonté des sources de données pour nourrir ou inspirer des commentaires, je vois donc un présent bien réel permettant au plumet héritière d'Albert Londres, de Zola et de John Reed d'appuyer leurs propos sur des informations fiables, vérifiées et accessibles, sinon mathématiquement exactes, traduisant au moins en toute transparence l'état des connaissances sur de très nombreux sujets.

VOUS AIMEZ



0

VOUS N'AIMEZ PAS



0

LUI RÉPONDRE

**CAROLINE GOULARD**

le 9 mars 2010 - 14:32 &amp;bullet; SIGNALER UN ABUS - PERMALINK



@alain chaumet : Bonjour et merci

1/pour vos éclairages sur l'IGN. Bien sûr, l'IGN est dans une démarche de service public. Ceci dit, le financement d'un service public peut être assumé par l'impôt (gratuit pour tous) ou par l'utilisateur final (tarifié), selon l'utilité qui lui est reconnue par le plus grand nombre. C'est une question de préférence sociale finalement. Disons que les préférences sociales exprimées dans les pays anglosaxons ne sont pas les mêmes qu'en France.

2/pour votre très bon exemple d'usage journalistique de données publiques. Je partage tout à fait votre enthousiasme pour un journalisme d'investigation basé sur les données.

VOUS AIMEZ  0 VOUS N'AIMEZ PAS  0

LUI RÉPONDRE

### BIBI

le 9 mars 2010 - 18:44 &bullet; SIGNALER UN ABUS - PERMALINK



Robot journalisme, next step.

[http://www.lemonde.fr/actualite-medias/article/2010/03/09/1-ere-des-robots-journalistes\\_1316608\\_3236.html](http://www.lemonde.fr/actualite-medias/article/2010/03/09/1-ere-des-robots-journalistes_1316608_3236.html)

VOUS AIMEZ  0 VOUS N'AIMEZ PAS  0

LUI RÉPONDRE

### PIGEONPIGEON

le 10 mars 2010 - 11:33 &bullet; SIGNALER UN ABUS - PERMALINK



@Caroline ; @Nicolas

Désolée pour le délai de réponse. Il s'agit, comme l'indiquent les crédits, si vous fouillez un peu l'animation Flash, de Pierre Bance, développeur, et David Castello-Lopes, journaliste, tous deux indépendants, puisqu'aucune rédaction française ne semble encore être capable d'intégrer ce genre de montage professionnel, comme le dit très bien l'article. Pour plus d'information vous pouvez contacter Pierre Bance ici : [bance.pierre@jgmail.com](mailto:bance.pierre@jgmail.com)

VOUS AIMEZ  0 VOUS N'AIMEZ PAS  0

LUI RÉPONDRE

### LENARD YIP

le 29 août 2010 - 2:35 &bullet; SIGNALER UN ABUS - PERMALINK



hiya all, i'm really bored on the net so you all should contact me if you are too, strike up a convo :). or possibly facebook, my name on there is chelsey ramos

VOUS AIMEZ  0 VOUS N'AIMEZ PAS  0

LUI RÉPONDRE

### ALISON AKSOY

le 30 août 2010 - 5:41 &bullet; SIGNALER UN ABUS - PERMALINK



hello all, i'm really listless on the net so you all should mail me if you are also, strike up a convo :). or possibly facebook, my name on there is angela walloch

VOUS AIMEZ  0 VOUS N'AIMEZ PAS  0

LUI RÉPONDRE