

# next cetera





Chaque mois, *Next* invite des spécialistes à décrypter les nouveaux comportements et les mutations de l'époque.



SYLVAIN PARASIE

## hackers et à chiffres

Propos recueillis par Cécile Daumas

**Si aux États-Unis le « *data journalism* » fait partie intégrante des grands médias, la France est à la peine, et les banques de données ne sont pas assez utilisées. État des lieux et perspectives avec un spécialiste des nouvelles technologies.**

**D**es députés britanniques qui se font payer leur « *home cinema* » avec les deniers publics ; un dépassement d'honoraires de médecin plus élevé à Nice qu'à

Laval ; le nombre de vols dans une rue de Chicago. Aussi disparates soient-ils, ces faits chiffrés sont le résultat d'une nouvelle manière de produire de l'information : sélectionner, analyser et visualiser des bases de données disponibles sur Internet, pour sortir des infos inédites, qui vont du scoop mondial (cf. les documents confidentiels sur les guerres en Irak et en Afghanistan publiés par WikiLeaks) au renseignement pratique de quartier. Le journalisme de données, ou « *data journalism* » en anglais, est la combinaison entre le savoir-faire traditionnel de Rouletabille et la puissance du web. Né dans les pays anglo-saxons au milieu des années 2000, il se propage – doucement – en France. Alors que sort en anglais un manuel compilé

par les meilleurs experts de la planète, *The data journalism handbook* <sup>(1)</sup>, le sociologue Sylvain Parasie, spécialiste de l'innovation dans les médias, explique pourquoi la France a pris du retard dans ce domaine. Maître de conférences à l'université Paris-Est Marne-la-Vallée, il est chercheur au Laboratoire Techniques, Territoires et Société (Latts) de l'École nationale des ponts et chaussées. Il a publié en 2010 *Et maintenant, une page de pub ! Une histoire morale de la publicité à la télévision française (1968-2008)*, aux éditions de l'INA.

**Le journalisme de données n'est pas un pur produit des médias traditionnels. Où est-il né ?**

Cela fait plusieurs décennies qu'on utilise les données informatiques pour faire du journalisme. Mais ce qu'on appelle aujourd'hui le « *journalisme de données* » est la conséquence des nouveaux liens qui se forment depuis le milieu des années 2000 entre le monde de la presse et les mondes informatiques. Des entrepreneurs du web, des militants du logiciel libre ou d'ONG conçoivent, à partir des bases de ►





► données consultables sur le web, des applications permettant aux citoyens d'exercer un contrôle plus pertinent sur le débat public. Par l'analyse des votes de chaque membre du Congrès américain, par exemple, ils voient si leurs représentants favorisent ou non les donateurs privés qui avaient financé leur campagne. D'autres sites répertorient les circuits de financements internationaux. Pour ces militants de l'« *Open government* », il s'agit de favoriser la transparence et la démocratisation de la vie politique, et lutter contre la corruption. À partir des années 2000, la généralisation de l'ouverture des données publiques sur le web donne ainsi aux citoyens le moyen d'exercer une plus grande vigilance démocratique.

#### **Comment le lien se crée-t-il avec les médias traditionnels, les journaux quotidiens notamment ?**

Comme la presse reste une des instances fondamentales pour la circulation de l'information, ces militants du web s'y intéressent. Certains programmeurs associent leur engagement dans les médias à la question démocratique. « *Rejoignez le journalisme et aidez à*

met aux internautes d'examiner plusieurs centaines de milliers de documents concernant les notes de frais des parlementaires britanniques. La participation fut très importante, et à l'origine de plusieurs révélations sur certaines dépenses incongrues de ces députés.

#### **Avec ces codeurs venus de l'Internet, a-t-on affaire à un nouveau type de professionnels ?**

Les rédactions américaines voient débarquer un nouveau type de professionnels : des « *journalistes hackers* » ou « *programmeurs-journalistes* ». Ils allient une double compétence, à la fois technique et éditoriale. En 2007, Aron Pilhofer, un des pionniers de la discipline, met en place au *New York Times* une cellule de journalistes capables de coder et produire de nouvelles informations. Ce profil de journaliste-programmeur est neuf, même si depuis les années 1970, les Américains font du journalisme d'investigation à partir de données informatiques. Inconnue en France, la tradition dite du « *computer-assisted reporting* » (journalisme assisté par ordinateur) a connu un certain succès en portant de nom-

les sites d'information, qui ont conçu de nombreux services en ligne à destination des internautes. En France, les statistiques criminelles ne sont pas disponibles au niveau du quartier. Lors de sa venue en France, Adrian Holovaty, qui est également créateur d'everyblock.com, site de référence sur les infos locales (de la criminalité au meilleur resto), s'est étonné qu'un tel service n'existe pas à Paris. La criminalité donnée artère par artère susciterait sans doute d'énormes débats dans notre pays.

#### **Quand la France se met-elle réellement au journalisme de données ?**

En France, plusieurs initiatives sont apparues depuis quelques années. Lancé en 2009, le site owni.fr est rapidement devenu le site de référence du « *data* » journalisme en France. Proposant régulièrement des applications en ligne, il s'est fait connaître en 2010 alors qu'il nouait un partenariat avec Wikileaks. Quelques petites sociétés sont apparues récemment, qui produisent des visualisations de données, notamment pour le compte de journaux. Mais l'essentiel de ces projets français sont menés en dehors des grands médias, contrairement à ce qu'on observe aux États-Unis.

Pour la presse française, déjà confrontée à l'énorme enjeu du web – quel modèle économique adopter dans un univers de gratuité ? –, le journalisme de données peut sembler un épiphénomène. C'est pourtant un autre virage complexe à prendre. Cela implique des moyens et des compétences souvent difficiles à réunir, mais cela peut générer une information automatisée à fort potentiel d'audience. Dans le sport, par exemple, on sait que le public est très friand de statistiques sur les performances des sportifs, et de nombreuses applications informatives pourraient être conçues dans ce domaine.

#### **Ce nouveau type de journalisme entraîne-t-il des manières innovantes de concevoir l'information ?**

Avant d'intégrer les grands journaux américains comme le *New York Times* ou le *Chicago Tribune*, certains programmeurs-journalistes géraient des projets web dans le privé. Ils ont donc adapté à la presse traditionnelle les méthodes de gestion des start-up : définition des besoins, réunions régulières entre reporters et programmeurs, minimisation du temps passé à coder inutilement. Le journalisme de données utilise ainsi des techniques conçues hors du monde de la presse. Une façon de réinitialiser ses contenus et ses pratiques. L'utopie « *hacker* » appliquée au journalisme s'avère porteuse de promesses dans les rédactions. Avec des produits très diversifiés, il est appelé à se développer. Dans un avenir proche, de plus en plus de statistiques vont être accessibles en ligne et les informations les plus consultées du web – météo, sport, informations locales – se prêtent fort bien à ce type de traitement. ●

<sup>(1)</sup> *The data journalism handbook*, dirigé par Jonathan Gray, Liliana Bounegru et Lucy Chambers.  
<http://datajournalismhandbook.org/1.0/en/>

## **« Les journalistes hackers allient une double compétence, à la fois technique et éditoriale. »**

*sauver la démocratie !* » : c'est ainsi que Brian Boyer, qui est en charge de l'équipe des programmeurs-journalistes du *Chicago Tribune*, s'adresse à ces nouvelles recrues. Une partie d'entre elles entre dans les grands journaux américains ou lance leurs propres sites web d'information. Adrian Holovaty est l'une des figures les plus connues de ce mouvement. Diplômé en journalisme, âgé d'une trentaine d'années, ce dernier est un programmeur passionné, reconnu dans le monde du logiciel libre. Embauché par des rédactions comme celle du *Washington Post*, il développe des produits journalistiques innovants. En exploitant les données de l'armée américaine, il met ainsi au point « *Faces of the fallen* » qui recense la biographie des soldats morts en Irak ou en Afghanistan, avec photo, date de naissance et mort, âge, lieu de résidence, corps de rattachement, blessures, etc. En Europe, l'Angleterre emboîte le pas, avec notamment *The Guardian*, qui dispose d'une équipe de référence autour de Simon Rogers. En 2009, le site anglais propose un service en ligne qui per-

breuses affaires à la connaissance du public américain à partir d'une recherche informatique : des discriminations entre les citoyens, des inégalités de traitement dans les écoles ou les hôpitaux, par exemple.

#### **Pour quelles raisons la France n'est-elle pas très innovante dans ce domaine ?**

Tout d'abord, il n'existe pas en France de tradition similaire au « *computer-assisted reporting* », et les journalistes français sont peu familiers de l'utilisation des statistiques. Il est donc assez naturel qu'ils éprouvent aujourd'hui quelques difficultés à se lancer dans l'aventure. Mais cela s'explique aussi par les données auxquels les journalistes français ont aujourd'hui accès. Aux États-Unis et en Grande-Bretagne, le mouvement pour l'ouverture des « *data* » publiques a été un formidable moteur pour cette forme de journalisme. Or le mouvement est moins étendu en France. Prenons l'exemple des données criminelles. Dès 1996, la police de Chicago crée un site web où est relevé le taux de criminalité rue par rue. Cette matière a largement été exploitée par