

La datavisualisation, double point d'entrée du datajournalisme dans la PQR

< Marie Chagnoux ¹ >

1. CREM, Université de Lorraine
Île du Saulcy -UFR SHS, F-57045 Metz, France
Marie.Chagnoux@univ-lorraine.fr

DOI : 10.25965/interfaces-numeriques.4480

< RÉSUMÉ >

Ce travail trouve son ancrage dans une enquête de terrain sur les mutations à l'œuvre dans les rédactions de la PQR, presse quotidienne régionale. Le datajournalisme avait alors émergé comme une notion à questionner et à redéfinir tant les discours produits semblaient complexes. Entre espoirs, souvent idéalisés, d'un outil clé en main pour illustrer l'actualité ultra-locale et verrous institutionnels et individuels éprouvés lors des pratiques, il semblait y avoir un espace complexe à explorer. Nous proposons ici de présenter le rôle particulier que tient la datavisualisation dans cet espace : souvent confondue avec le datajournalisme, la datavisualisation constitue fréquemment la première étape d'une familiarisation au traitement des données, facilitée par le recours à des outils clé en main qui interviennent à deux étapes du traitement journalistique : lors de l'exploration des données par le journaliste puis dans leur mise en narration pour le lecteur.

< ABSTRACT >

This work is rooted in a survey on the changes in the editorial offices of the PQR, the regional daily press. Data-journalism had then emerged as a notion to be questioned and redefined. Between idealized hopes of a turnkey tool to illustrate ultra-local news and institutional and individual locks tested in practice, there seemed to be a complex space to explore. We propose here to present the particular role that datavisualization plays in this space. Often confused with datajournalism, datavisualization is frequently the first step in familiarizing oneself with data processing, facilitated by the use of turnkey tools that intervene at two stages of journalistic processing : during the exploration of data by the journalist and then in their narration for the reader.

< **MOTS-CLÉS** >

datajournalisme, big data, data mining, journalisme, datavisualisation, webjournalisme

< **KEYWORDS** >

datajournalism, big data, data mining, journalism, datavizualisation, webjournalism

1. Introduction

Afin de documenter les mutations des formats d'investigation et de narration induites par les techniques numériques dans la presse, nous avons mené en 2016 une enquête de terrain dans les rédactions de la presse quotidienne régionale (PQR) de la Région Grand Est, dans le cadre du projet *#LIVE*. Certains phénomènes émergents, comme le suivi *live* de l'actualité ou l'interactivité avec les lecteurs, nous avaient paru désormais durablement implantés dans le quotidien des rédactions, notamment grâce à une familiarisation de plus en plus prégnante avec les médias sociaux. À l'inverse, une forme émergente se singularisait dans le discours des journalistes interrogés : spontanément, un certain nombre d'entre eux évoquait l'écueil et la fascination que représentait le datajournalisme.

Dans les entretiens, les contours de ce « journalisme de données » se dessinaient dans la frustration d'une difficile mise en œuvre. Il semblait à la fois considéré comme une étape prometteuse et nécessaire pour donner un nouveau souffle à l'actualité ultra-locale et comme un objet aussi inaccessible que fantasmé. À l'heure de la *dataification*, cette « mise en données du monde » (Mayer-Schonberger et Cukier, 2014), il nous avait donc paru pertinent de questionner plus en avant ces représentations qu'en avaient les rédactions. En sciences de l'information et de la communication, le datajournalisme reste principalement étudié du point de vue de la reconfiguration des pratiques professionnelles (Lewis et Westlund, 2015) et du contexte anglo-saxon (Joux et Bassoni, 2018). Nous avons donc fait le choix, par une étude de terrain et une étude des productions, de placer les processus d'analyse et de mise en narration des données au cœur de notre travail pour mieux appréhender les représentations et les réalisations du datajournalisme dans les rédactions locales.

Nous verrons plus loin que formation insuffisante, temps d'enquête trop limité et écueils de traitement sont quelques-uns des obstacles auxquels se heurtent les journalistes. En revanche, il est apparu, notamment lors de l'analyse des productions, que la datavisualisation se présentait comme la première étape d'une familiarisation au traitement des données, justement parce qu'elle permettait de contourner ces obstacles. Peu coûteuse en temps, démocratisée par l'apparition de plateformes de traitement de données conviviales et accessibles, facilement déclinable, cette possibilité de transformer des données brutes et complexes en représentations visuelles intuitives et interactives séduit à l'heure où les rédactions se tournent massivement vers des formats web et où l'image joue un rôle prépondérant dans le partage d'informations sur les réseaux sociaux.

Dans un premier temps, nous montrerons que la polysémie, ou l'ambiguïté, du concept de « journalisme de données » qui émergeait des discours est entretenue par la littérature scientifique et didactique qui accorde une place prépondérante à la visualisation. Une seconde partie, basée sur nos entretiens et analyses, présentera les représentations et les imaginaires que convoque ce concept dans les rédactions de la PQR, en insistant sur les verrous logistiques et méthodologiques. Nous montrerons enfin à travers deux exemples, comment la visualisation, en tant que narration journalistique et non simple mise en image, peut créer de l'information pertinente au niveau local.

2. De la littérature au terrain : un datajournalisme polysémique et protéiforme

Si, au départ, la définition du datajournalisme, ce « journalisme de données » qui vise l'exploitation de données statistiques et leur mise à la disposition aux lecteurs, paraît non ambiguë, l'examen de la littérature comme des pratiques de terrain tend à faire émerger un concept aux frontières plus que floues. Parce qu'il implique des pratiques et des compétences hétérogènes en amont, parce qu'il se concrétise dans des réalisations protéiformes aux subtilités difficiles à appréhender pour des non-spécialistes en aval, le terme est utilisé pour renvoyer à des processus d'investigations et de restitutions différents.

2.1. La pluralité d'approches derrière le terme de « datajournalisme »

Dans leur tentative de définir le datajournalisme à partir des discours de ses acteurs, Croissant et Touboul (2013, 133) montrent que la polysémie provient principalement de la multiplicité des domaines scientifiques dans lesquels il s'inscrit : entre infographie, *fact-checking*, *visual mining*, datavisualisation, *open data*... Les auteurs éclairent la difficulté terminologique par les différences de contexte, par exemple entre États-Unis et France, et par la pluralité des approches disciplinaires convoquées pour l'appréhension des données. Elles rappellent néanmoins que la perméabilité du journalisme à d'autres pratiques sociales, qui rend impossible le consensus dans sa définition même, s'applique également au journalisme de données.

Pour Joux et Bassoni (2018), si le terme désigne toujours intrinsèquement un journalisme qui déserterait le terrain pour retrouver la culture du document, celle des données et des archives, la polysémie repose dès le départ sur les deux niveaux d'interprétation qui peuvent être faits de la notion de « donnée » dans les rédactions. Le premier niveau est celui des données du terrain permettant l'enquête du journaliste alors que le second concerne les données marketing qui permettent d'adapter la production journalistique à la perception des attentes du public. Mais la polysémie reposerait aussi sur l'acceptation plus ou moins large qui serait faite de pratiques protéiformes. Ainsi, par exemple, est cité le journalisme automatisé, comme celui de *Narrative Science*¹, société qui produit des articles générés automatiquement à partir de données et offre alors aux rédactions la « possibilité de multiplier les articles pour répondre à des demandes de niche et pour traiter de sujets que les rédactions n'avaient pas les moyens de couvrir jusqu'alors. » (*ibid*, 128 ; Desbordes, 2018).

Ainsi, le terme même de « datajournalisme », de par sa grande perméabilité aux pratiques et aux contextes, porte en lui la potentialité d'une souplesse terminologique.

¹ <https://narrativescience.com/>

2.2. Datavisualisation, le traitement de la donnée à portée de clic

Le traitement des données, tel qu'il peut être abordé par les datascientistes, suppose un travail complexe à chacune des étapes : lors de la collecte, lors du traitement et lors de la mise en cohérence des données pour élaborer de l'information ou de la connaissance. Ce travail suppose une solide formation en mathématique, notamment en statistiques et informatique, en particulier en programmation. Cet ancrage disciplinaire fait souvent défaut dans les formations initiales des journalistes et reste difficilement accessible en autoformation à des profils non scientifiques.

Ainsi, les manuels pratiques, à l'instar de la plupart des cours de datajournalisme assurés dans les formations françaises, auront une propension certaine à assimiler « datajournalisme » et « datavisualisation », faute de pouvoir proposer un apprentissage plus ambitieux. Il est alors très souvent proposé aux étudiants une initiation à base de tableurs et d'outils clé en main de datavisualisation. La visualisation des données n'est alors plus seulement un outil de mise en narration à destination des lecteurs mais également un outil d'exploration et de traitement au cœur du métier. C'est ce que professe Gray, Bounegru, Chambers et Kayser-Bril (2013, 186) dans un chapitre intitulé « La visualisation, meilleur outil du datajournaliste » :

Il paraît alors judicieux de commencer à développer des visualisations tôt dans votre travail d'investigation et régulièrement. Ne voyez pas cela comme une étape séparée, quelque chose à considérer une fois que l'article est en grande partie écrit. Laissez ce travail guider votre enquête

Ce principe méthodologique brouille ainsi encore davantage les principes définitoires qui caractériseraient le datajournalisme comme un traitement *a priori* de l'information et la datavisualisation comme un traitement *a posteriori* : la datavisualisation interviendrait à différentes étapes du travail journalistique pour guider l'exploration et la présentation des données. Arruabarrena (2015a) et Langlais (2013) évoquent tous deux la démocratisation des outils de datavisualisation. La première explique comment, initialement réservés aux experts scientifiques et professionnels, les outils se sont démocratisés avec le

lancement d'offres *Open source* et *Freemium* puis le déploiement de services web, notamment ceux de Google avec *Google Chart*, *Google Fusion Table*², ... Le second explique, à l'aide de l'exemple d'*Highcharts*, comment le développement d'un outil informatique simplifié et adapté au contexte journalistique peut jouer un « rôle crucial dans l'avènement et la légitimation du datajournalisme ».

Enfin, réduire le datajournalisme à de la datavisualisation n'est pas uniquement motivé par un souci d'accessibilité des traitements mais aussi par des impératifs économiques comme l'explique le datajournaliste Sylvain Lapoix dans Ouardi (2014, 76) :

[1] Il y a clairement un malentendu sur ce qu'on entend parfois par data journalisme, que les supports veulent parfois réduire à de la simple data visualisation, c'est-à-dire à de la mise en forme graphique de données. Ce malentendu est assis sur des enjeux économiques et politiques. Je m'explique : pour certains, faire du data journalisme ce serait simplement faire trois schémas avec de jolis dessins attractifs et faciles à comprendre pour mettre en scène quelques « chiffres clefs ». C'est une démarche marketing qui n'est pas fidèle à ce qu'est et peut être le data journalisme.

2.3. Datavisualisation, le traitement de la donnée à portée de lecteur

La tentation du « joli dessin attractif et facile à comprendre » fait écho à la définition de (Goulard, 2012, 90) : « Nous appelons ici journalisme de données la pratique journalistique qui consiste à exploiter des bases de données volumineuses pour en extraire de l'information et la présenter de façon engageante au public ».

² Il faut citer ici l'offensive que mène actuellement Google sur ce terrain de l'outillage du datajournalisme. Ainsi, le *Google News Lab* qui vise, selon le site, à « collaborer avec des journalistes et des entrepreneurs pour aider à construire l'avenir des médias », propose un cours en ligne de datajournalisme et de nombreuses ressources logiciels et méthodologiques pour la construction de datavisualisation (notamment à l'aide de DataStudio). Il a notamment publié en 2017 une très large enquête internationale sur le data journalisme (Rogers *et al*, 2017) à partir de 56 entretiens et plus de 900 réponses à un questionnaire.

Cette dimension « engageante » de la visualisation est également le parti-pris des professionnels : « Si un graphique explique mieux un sujet qu'un texte, nous choisirons le graphique. Les informations peuvent souvent être racontées en graphes, en données ou en vidéo » professe la charte des décodeurs du *Monde*³. Sur ce site de *fact-checking*, la sélection et la mise en forme graphique des données sont privilégiées, assimilées à une mise en narration, véritable outillage de ce site pour contrer les *fake-news*. La dimension pédagogique de la donnée est souvent citée comme un atout car, par-delà son accès facile, elle entretient aussi l'illusion d'une pseudo-objectivité de présentation qui permettrait au lecteur d'appréhender des données pour en tirer ses propres conclusions.

3. Représentations et expressions du datajournalisme au sein des rédactions de la PQR

Une fois ce contexte général posé, nous nous proposons de le confronter au terrain de la PQR. Deux travaux complémentaires sont ici convoqués. Le premier permet d'accéder au point de vue de journalistes non-spécialistes sur le datajournalisme. Le second au contraire sollicite des datajournalistes et leurs réalisations pour montrer comment la datavisualisation peut constituer un premier point d'entrée pour le datajournalisme.

3.1. Le datajournalisme vu par les journalistes de la PQR

Le premier travail, conduit dans le cadre du projet #LIVE, ne portait pas au départ spécifiquement sur le datajournalisme, les éléments que nous résumons ci-dessous ont donc émergé spontanément dans des rédactions et doivent à ce titre être davantage considérés comme des indicateurs plutôt que comme les constituants systématiques d'une analyse cadrée.

³<https://www.lemonde.fr/les-decodeurs/article/2014/03/10/la-charte-des-decodeurs43651064355770.html>

29 entretiens ont été menés à l'automne 2017 dans 9 rédactions de 6 titres de PQR de la région Grand Est (voir Tab. 1).

Groupe	Titre	Ville	Nbre entretiens
Ebra	<i>L'Alsace</i>	Mulhouse	4
	<i>Vosges Matin</i>	Epinal	2
	Le Républicain Lorraine	Forbach	3
		Metz	4
		Thionville	2
	L'Est républicain	Nancy	2
Rossel-La Voix	L'Est éclair	Troye	3
	L'Union L'Ardennais	Charleville Mezière	4
		Reims	5

Tableau 1 : Description du corpus des entretiens de #LIVE

Ils révèlent à la fois une familiarité des journalistes avec le concept et une difficulté à clairement énoncer comment le datajournalisme pourrait s'incarner dans leur rédaction. Trois éléments saillants ont émergé :

- la prise de conscience de l'importance du datajournalisme pour le futur de la presse locale ;
- la dimension d'émergence de l'approche en inadéquation avec la formation et le contexte ;
- l'importance du facteur individuel.

La familiarité avec le concept se traduit par la prise de conscience de la multiplication des sources de données électroniques disponibles et des politiques d'encouragement de libération des *open data*, notamment depuis la *Loi pour une République Numérique* de 2016. Les journalistes sont conscients que ces données offrent des terrains d'investigation de plus en plus accessibles. Ils y voient également un moyen de produire facilement de l'information ultra-locale en cohérence avec leur mission, confirmant ainsi l'affirmation de Joannès (2010, 147) pour lequel « plus les informations publiques deviennent accessibles et exploitables, plus

les rédactions ont intérêt à s'y investir dans une stratégie délibérée de rapprochement avec leur audience ». Mais ils confirment également le risque de modification de l'identité professionnelle du localier vers ce qu'Amiel (2017, 336) appelle le journalisme « de service », centré sur la mise à disposition d'informations pratiques.

Tous les journalistes qui ont spontanément abordé la question du datajournalisme se rejoignent sur la dimension émergente de la technologie (« *C'est quelque chose de plutôt émergent* » [E02], « *On est vraiment aux balbutiements des data* » [E12]). Pour beaucoup, ce sont les formations en interne qui ont soulevé un certain intérêt car, comme le souligne Amiel (2017, 120), l'« une des stratégies adoptées par les entreprises de presse locale pour intégrer au mieux la publication en ligne est celle de la formation ». A l'initiative des groupes de presse Ebra et Rossel, des formations sont proposées (« *la semaine dernière on avait encore des équipes en formation data. Y'a une volonté de développer cet aspect-là* » [E02]). Ces formations s'apparentent davantage à de l'initiation : une partie théorique pour « savoir ce que c'est » (extrait 1) et une partie pratique, le plus souvent de l'exploitation de données avec des tableurs en vue de générer des visualisations (extrait 2) comme nous l'exposons dans la première partie :

1. « *ça me sert évidemment culturellement, je veux pas être coupé de ça : ça me sert aussi quand mes équipes m'en parlent, qu'ils me disent « regarde j'ai récupéré telles données, comment on peut l'exploiter ? » que je leur dise pas « je sais pas ». Maintenant je peux leur dire « ben tu vas sur tel site et tu nous fais une belle infographie »*[E02]
2. « *il y a des trucs qui s'inventent partout en permanence et même quelques choses qu'on fait maintenant qu'on a été formé par un collègue des DNA, par exemple je fais des cartes interactives* » [E12]

En dehors de la génération de visualisations attractives, deux autres usages sont spontanément cités :

- la possibilité de produire des signaux faibles pour explorer des sujets émergents à partir des open data, avec le constat d'une absence de compétence (extrait 3) ;

- l'automatisation du traitement de certaines informations facilement reproductibles comme la génération de scripts pour le traitement des résultats des élections (extrait 4).

3. « *Dernièrement ils ont fait un gros travail de data sur les *zadistes originaires du Bas-Rhin, qui était quelque chose qu'on ne sait pas faire* »(E12)

4. « *Une fois que tu as le programme, tu réinventes les données à l'identique, ce qui est long, c'est le code* »(E30)

Enfin, dernier élément saillant, les disparités entre les postures individuelles apparaissent très importantes dans l'usage qui est fait des techniques de datajournalisme, entre des personnes formées mais qui ne pratiquent pas et des autodidactes passionnés qui se forment au fil des productions (extrait 5).

5. « *Beaucoup de choses se font en autodidacte mais quand on atteint un certain niveau de complexité, c'est bien de prendre deux heures de formation avec un collègue, mieux que passer une semaine à chercher à s'arracher les cheveux parce que ça ne marche pas* » (E12)

La totalité des personnes interrogées expriment leurs frustrations de ne pouvoir pleinement exploiter le potentiel offert par les données, soit que leurs compétences méthodologiques soient trop limitées, soit que les temps nécessaires au travail d'investigation et de mise en forme ne soient pas compatibles avec la temporalité des rédactions. Ces deux éléments apparaissent comme les principaux inhibiteurs à la réalisation d'articles, le second étant le plus souvent rédhitoire comme l'explique Sylvain Lapoix dans l'entretien cité plus haut (Ouardi, 2014, 78) :

le problème de la data-visualisation est qu'on demande au journaliste de consacrer un temps démesuré, mais alors DÉ-ME-SU-RÉ à des gestes qui ne relèvent pas en fait d'une réflexion sur la forme, mais à sa réalisation... au détriment parfois du travail journalistique en tant que tel. Quand j'étais chez Owni, je me suis intéressé à la cartographie, qui est une forme, pour le coup, extrêmement signifiante et efficace rhétoriquement. J'ai voulu faire des cartes interactives (notamment sur le gaz de schiste), et pour ça je me suis tourné vers les outils disponibles comme Google Map, sauf que c'est un outil... Concrètement, il m'a fallu des semaines pour réaliser une carte très moyenne, ces semaines de travail étaient retranchées de

La datavisualisation, double point d'entrée du datajournalisme dans la PQR < 11 >

mon travail de recherche et d'enquête. Un développeur fait ça en une journée...

3.2. Le datajournalisme vu par les datajournalistes de la PQR

A la suite de ce travail d'enquêtes, nous avons commencé à analyser des productions fléchées « datajournalisme » de la PQR en tant que « discours numériques natifs » (Paveau, 2017) pour appréhender le lien entre des données initiales et des mises en récit, qu'elles soient graphiques ou rédigées, dans le cadre du projet MADONA⁴. A travers ces différents exemples, nous souhaitons montrer comment ces formes embryonnaires de narrativisation des données offrent à la presse locale un nouveau rôle d'explicitation des politiques territoriales et des événements ultra-locaux. Nous avons également commencé une collaboration avec des informaticiens du LIFAT⁵ et des datajournalistes recrutés lors d'un journocamp du collectif DATA + LOCAL⁶, un collectif qui rassemble des datajournalistes d'une quinzaine de rédactions et des journalistes indépendants. La première étape de cette collaboration consistait à documenter le métier en répondant à un questionnaire en ligne.

Le questionnaire a permis de confirmer les éléments présentés en première partie :

- La formation initiale en lettres ou en sciences humaines et sociales incite à se limiter à des traitements simples de données. Seuls 5 datajournalistes sur les 18 répondants utilisent des méthodes automatisées de fouille de textes ou de *machine learning*. La moitié d'entre eux utilisent des scripts (Python, Matlab, script shell, ...) et un tiers utilisent des applications dédiées ou des logiciels clé en main (Jupyter, PwerBi, ...). Le tableur reste le principal outil de

⁴ MADONA (*Maîtriser l'Analyse interactive de DONnées pour la NArration journalistique*) est une action du GDR MADICS, Masses de Données, Informations et Connaissances en Sciences

⁵ Laboratoire d'Informatique Fondamentale et Appliquée de Tours

⁶ <https://collectif-datalocal.github.io/> Les journocamps, sont des événements qui rassemblent des journalistes sous forme de barcamp.

travail puisqu'il est systématiquement utilisé par 90 % des répondants.

- La datavisualisation est pour 75 % des répondants un outil qui permet d'explorer les données. Elle est également citée pour ses capacités d'illustration attractive d'enquêtes et comme un moyen efficace d'améliorer l'impact d'un article en proposant des illustrations qui pourront être reprises sur les médias sociaux.
- Alors que le travail d'enquêtes sur les données s'inscrit dans un temps long, le travail de mise en narration visuelle peut être extrêmement rapide, quelques minutes pour certaines visualisations comme celles liées aux résultats des élections par exemple.

Parmi les réalisations de ces journalistes de la PQR, nous avons retenu deux exemples qui s'inscrivent dans deux dynamiques complémentaires. Le premier porte sur l'exploitation de données externes, instanciées localement dans des visualisations qui permettent justement cet accès « engageant » cité plus haut. Le second montre comment la construction et l'exploitation de données internes comme le recommande Briggs (2014, 221) permet de proposer un contenu informationnel original où la donnée journalistique est valorisée. Dans les deux cas, un encart méthodologique est proposé au lecteur pour expliciter les processus d'analyse et de construction.

Le premier exemple repose sur l'exploitation pour *Rue89 Strasbourg* d'un jeu de données sur l'attribution des marchés publics de la Communauté urbaine de Strasbourg en 2014. Ce travail a généré au moins trois articles, témoignant ainsi du besoin de rentabiliser le travail d'investissement évoqué précédemment.

La datavisualisation, double point d'entrée du datajournalisme dans la PQR < 13 >

Commune	Date d'exécution	Objet 1	Objet 2	Entreprise	Montant (€)
Quatzenheim (87)	02/04/2014	MAITRISE OEUVRE POUR ETUDES ET SUIVI TRAVAUX REFECTION	GLOBALE PONT EHLMANN A PLOBSHEIM	BET FAVIER VERNE	18 125
Quatzenheim (87)	06/08/2014	MAITRISE OEUVRE REFECTION TOTALE PONT SAINT GUILLAUME		BET FAVIER VERNE	30 750
Wickersheim (87)	08/04/2014	ACQUISITION DE 5 TONDEUSES TRACTEES		MDN	4 466

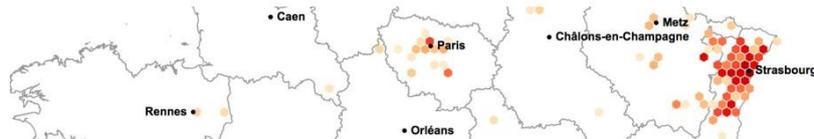


Figure 1 : visualisation interactive de la localisation des entreprises ayant remporté des marchés publics

- Le premier article, « Marchés publics : d'où sont les entreprises attributaires en 2014 ? »⁷, détaille la localisation des entreprises qui ont remporté les appels d'offre (Fig. 1). Il offre une carte interactive qui permet d'afficher la localisation conjointement à l'affichage d'un tableau qui détaille les données.

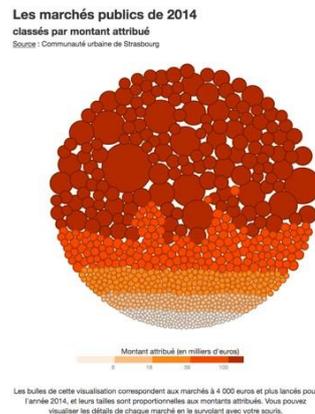


Figure 2 : visualisation dynamique et interactive des marchés publics



Figure 3 : graphique de l'évolution des marchés publics

⁷ <https://www.rue89strasbourg.com/marches-publics-dou-sont-les-entreprises-attributaires-en-2014-84199>

- Le seconde article, « Marchés publics : ce que la CUS a dépensé en 2014 » détaille à l'aide d'une infographie interactive ce que la collectivité a acheté⁸ sous forme de bulles dont la taille est proportionnelle aux montants attribués, (Fig. 2) ;
- Le troisième article, « Marchés publics : les respirations d'une mandature »⁹ détaille l'évolution des marchés publics, en nombre comme en coût total, grâce à deux graphiques statiques (Fig. 3).

L'encart « méthodologie » (Fig. 4) n'explicite la provenance des données que pour le second article.

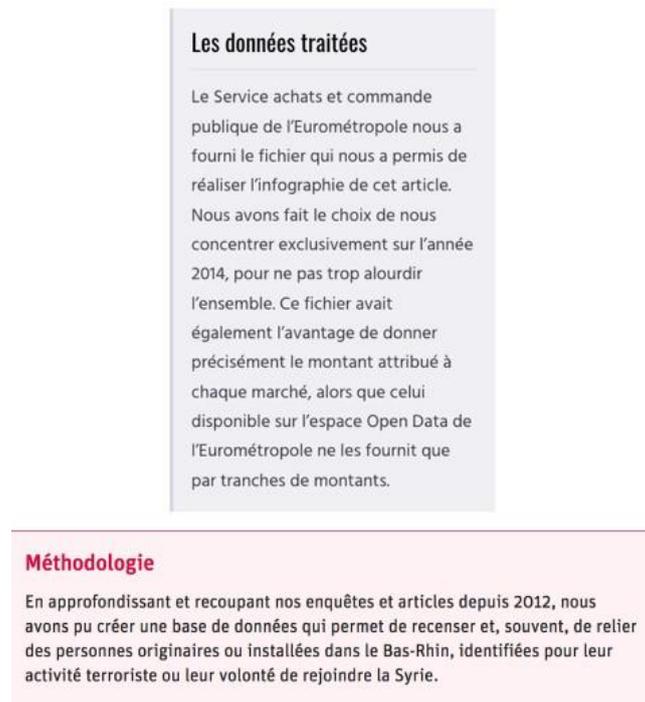


Figure 4 : encart méthodologique des articles des DNA

⁸ <https://www.rue89strasbourg.com/marches-publics-ce-que-la-cus-a-depense-en-2014-81296>

⁹ <https://www.rue89strasbourg.com/marches-publics-mandature-91311>

Le second exemple est un dossier multimédia paru sur le site des *DNA* (*Dernières Nouvelles d'Alsace*) et consacré aux terroristes djihadistes en Alsace¹⁰.

En s'appuyant sur les enquêtes et articles parus dans ce titre depuis 2012, le journal a constitué sa propre base de données sur laquelle un travail de ré-investigation a été mené. Il s'agit donc d'un processus original où l'information est créée à partir d'informations déjà générées en amont et en interne. Un autre fait intéressant réside dans la mise à jour constante des infographies. Proposé initialement en 2017, ce dossier a été par la suite modifié en 2019 après deux événements, une attaque perpétrée à Paris par un Strasbourgeois en mai et l'attentat meurtrier commis lors du marché de Noël de Strasbourg.

Le travail d'exploitation des données permet une contextualisation des événements liés au terrorisme, illustrée par différentes visualisations :

- Une première visualisation sous forme de frise chronologique interactive (Fig. 5) fait apparaître certains événements comme précurseurs des drames à venir et permet l'affichage des articles pertinents.
- Ensuite, un recensement de 38 membres de la "galaxie djihadiste" permet leur localisation dans une carte interactive (Fig. 6) dont les liens sont mis en image dans un graphe social interactif (Fig.7). Ce dernier illustre l'influence de chacun d'entre eux dans le réseau (taille des points) et permet d'accéder à de courtes notices biographiques et aux articles détaillant le parcours de chacun en cliquant sur le nœud du graphe qui lui correspond.

¹⁰ <https://www.dna.fr/dossiers/2017/11/10/bas-rhin-dans-la-toile-du-djihad-terrorisme-galaxie-alsace-bataclan-wissembourg-strasbourg-reseau>



Figure 5 : Dossier DNA, frise chronologique des événements



Figure 6 : dossier DNA, carte interactive de l'origine géographique de la "galaxie djihadiste"

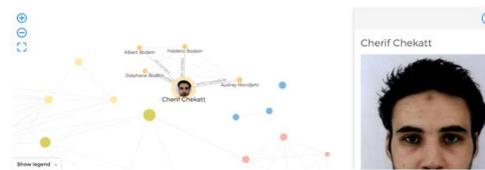


Figure 7 : Dossier DNA, graphe social interactif et accès aux bibliographies des djihadistes

L'examen de ces productions labélisées « datajournalisme » confirme :

- l'omniprésence des datavisualisations, interactives ou non, dans les articles avec un souci de varier les types de représentation ;
- le recyclage des données ou des principes méthodologiques au sein de plusieurs productions.

L'interactivité proposée dans ces représentations permet à l'internaute-lecteur de naviguer à l'intérieur des visualisations. Elle offre ainsi une personnalisation des données, intéressante au niveau de l'ultra-locale. La mise en carte fait émerger des disparités ou des similitudes entre les territoires, incitant à la glose et à l'analyse.

4. Conclusion

En croisant plusieurs travaux scientifiques, nous avons tenté de définir quelques représentations du datajournalisme dans les rédactions locales et d'identifier certains verrous techniques de sa mise en œuvre. Nous avons ainsi montré que les contraintes scientifiques et logistiques, notamment le haut niveau d'expertise technique et le temps long du traitement des données sont peu compatibles avec les réalités du travail dans les rédactions. La datavisualisation, par le biais d'outils clé en main,

offre une alternative qui permet de contourner ces difficultés. Quand la plupart des journalistes ne peuvent s'identifier à cette image du « journaliste hacker » décrite par Dagiral et Parasie (2011), ces outils vont viser la compréhension des données par le journaliste comme par le lecteur.

Le journaliste les utilisera dans un premier temps pour accéder à de larges sets de données complexes, tester des hypothèses, identifier des patterns. La datavisualisation devient alors en amont un « outil heuristique qui permet de rentrer graduellement au cœur du phénomène » (Arruabarrena, 2015a, 158). Puis dans un second temps, le lecteur y trouvera en aval une littératie visuelle qui, d'une part activera une composante productive qui utilise le visuel pour faciliter l'élaboration de connaissances, et d'autre part une composante réflexive qui permet rétroactivement de construire du sens (*ibid*, 160). La datavisualisation apparaît ainsi comme une double entrée à la mise en oeuvre de certains concepts du datajournalisme dans la PQR : lors de l'analyse et lors de la mise en narration.

Bibliographie

- Amiel Pauline (2017). *L'identité professionnelle des localiers à l'heure des mutations économiques et de la dématérialisation de la presse locale*, Thèse de Doctorat de l'Université de Toulouse-Paul Sabatier, Toulouse.
- Arruabarrena Béa (2015a). Datavisualisation : principes, enjeux et perspectives pour des utilisateurs non experts. In Broudoux Évelyne (éd.), *Big Data - Open Data : Quelles valeurs ? Quels enjeux : Actes du colloque « Document numérique et société »*, Rabat, De Boeck Supérieur, Louvain-la-Neuve, Belgique, pp. 151-163.
- Arruabarrena Béa (2015b). Datavisualisation : des données à la connaissance. *I2D - Information, données & documents* 2015/2, vol. 52, pp. 28-29.
- Briggs Mark (2014). *Manuel de journalisme web. Blogs, réseaux sociaux, multimédia, info mobile*, trad. et adapté de l'anglais par C. Robert, Eyrolles, Paris.
- Croissant Valérie, Touboul Annelise (2013). Le datajournalisme par ses acteurs, autodéfinition d'une pratique émergente en France. *Recherches en communication*, n° 40, pp. 133-149.

- Dagiral Eric, Parasie Sylvain (2011). Portrait du journaliste en programmeur : l'émergence d'une figure du journaliste "hacker", *Les Cahiers du journalisme*, n° 22/23, pp. 144-154.
- Desbordes Damien (2018). *Les robots vont-ils remplacer les journalistes ?* Plein Jour, Paris.
- Joux Alexandre, Bassoni Marc (2018). Le journalisme saisi par les Big Data ? Résistances épistémologiques, ruptures économiques et adaptations professionnelles. *Les Enjeux de l'Information et de la Communication*, n° 19/2, pp. 125-134.
- Gray Jonathan, Bounegru Liliana, Chambers Lucy, Kayser-Bril Nicoals (dirs.) (2013). *Le Guide du Datajournalisme*, Eyrolles, Paris.
- Goulard Caroline (2012). Transformation de notre paysage informationnel et journalisme de données, in Le Champion Rémy., *Journalisme 2.0 : Nouvelles formes journalistiques, nouvelles compétences*. La Documentation française. Études 5356-57, Paris, pp. 89-98.
- Joannès Alain (2010). *Data-journalisme, bases de données et visualisation de l'information*. EFE-CFPJ, Paris.
- Langlais Pierre-Carl (2013). Vers une vulgate data-visuelle ? Le cas d'Highcharts, *site web du projet Jourdain*, <https://projetjourdain.org/vers-une-vulgate-data-visuelle-le-cas-dhighcharts/>
- Lewis Seth C., Westlund Oscar (2015). Big data and journalism, Epistemology, expertise, economics and ethics. *Digital journalism*, n° 3, pp. 447-466.
- Mayer-Schonberger Viktor, Cukier Kenneth (2014). *Big data. La révolution des données est en marche*, trad. Française H. Dhifallah, Robert Laffont, Paris.
- Ouardi Samira (2014). Le data journalisme : entre retour du journalisme d'investigation et fétichisation de la donnée : Entretien avec Sylvain Lapoix. *Mouvements*, 79(3), La Découverte, Paris, pp. 74-80.
- Parasie Sylvain., Dagiral Eric (2013a). Des journalistes enfin libérés de leurs sources ? Promesses et "réalités du journalisme de données". *Sur le journalisme* [en ligne], vol. 2, n° 1, pp. 52-63.
- Parasie Sylvain., Dagiral Eric (2013b). Data-driven journalism and the public good : "Computer-assisted-reporters" and "programmer-journalists" in Chicago . *New Media & Society*, 15 (6), pp. 853-871.
- Paveau Marie-Anne (2017). *L'analyse du discours numérique. Dictionnaire des formes et des pratiques*, Hermann, Paris.
- Rogers Simon, Schwabish Jonathan, Bowers Danielle (2017) *The state of data journalism in 2017*, <https://newslab.withgoogle.com/assets/docs/data-journalism-in-2017.pdf>