



HAL
open science

Vers une mise en interface de données journalistiques

Everardo Reyes-García

► **To cite this version:**

Everardo Reyes-García. Vers une mise en interface de données journalistiques. Interfaces numériques, 2020, 9 (3), 10.25965/interfaces-numeriques.4453 . hal-03092823

HAL Id: hal-03092823

<https://hal.science/hal-03092823>

Submitted on 23 Mar 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Vers une mise en interface de données journalistiques

Visualisations journalistiques et magazines visualisés

< Everardo Reyes ¹ >

1. Laboratoire Paragraphe, Université Paris 8, Saint-Denis
2 rue de la Liberté, 93526 Saint-Denis, France
ereyes-garcia@univ-paris8.fr

DOI : 10.25965/interfaces-numeriques.4453

< RÉSUMÉ >

Les transformations constantes en termes de quantité/qualité, accès et variété de données entraînent différents changements au sein des pratiques journalistiques. Dans cet article, nous nous intéressons aux rapports entre les informations que nous caractérisons comme “données journalistiques” et leur mise en interface graphique, c’est-à-dire aux différentes manières dont les chiffres, textes et images sont organisés dans un but de visualisation et d’exploration. Après un aperçu matériel et historique des supports graphiques journalistiques, nous étudions les relations signifiantes entre trois niveaux de pertinence : thématique, visuel et interactif. Ensuite, nous discutons plusieurs cas concrets de visualisations des données publiées dans un cadre journalistique. De manière complémentaire, nous présentons nos propres visualisations d’un corpus de magazines afin de repérer des interactions entre thèmes et signes plastiques.

< ABSTRACT >

The practice of journalism is increasingly affected by continuous transformations of data in terms of quality/quantity, access, and variety of types. In this article, we discuss the relationships between information that we call “journalistic data” and their *mise-en-interface*, which encompass the different manners in which numbers, text and images are laid out with a practical goal of visualisation and exploration. First, we take a glance at journalistic graphical supports from a media archaeology perspective, we then observe meaning relations among three levels : thematic, visual, and interactive. Then, we discuss several datavisualizations published within a journalistic context. Further in our text, we present our own graphical prototypes created from a database

describing four magazines issued between 2012 and 2018. This approach allows us to see interactions between themes and colors, but also to adopt a critical and experimental perspective at a moment when numerous digital tools emerge and popularise *deep learning* techniques.

< **MOTS-CLÉS** >

mise en interface, données journalistiques, supports graphiques

< **KEYWORDS** >

interface design, journalistic data, graphical supports

1. Introduction

Dans la presse contemporaine, il est courant de constater l'existence d'une équipe dédiée au traitement et à la représentation graphique de données numériques. Tel est le cas des Décodeurs du *Monde* ou de Upshot au *New York Times*, tous les deux créés en 2014 et qui publient leurs notes sous forme de visualisations de données, d'infographies, de reportages interactifs et de récits de données. Dans cet article nous nous intéressons à ces réalisations depuis trois points de vue liés et qui contribuent de manière complémentaire à la construction d'un discours éditorial informatisé : l'aspect thématique (journalistique) met l'accent sur la manière dont les informations sont textualisées, l'aspect visuel se concentre sur les codes graphiques et l'aspect interactif fait appel au support technique d'affichage et à ses affordances.

Notre étude commence par une approche matérielle des médias pour rendre compte des relations entre les objets et leurs usages, mais aussi entre les objets et leurs caractéristiques techniques. Cette optique nous est utile afin d'identifier comment les pratiques sont déterminées par un niveau plus profond que les moyens immédiats de capture et d'affichage visuel. À son tour, l'évolution des capacités techniques répond à une configuration complexe d'exigences internes ou externes au domaine journalistique : économiques, politiques, éthiques, culturelles. Bien sûr, ces relations entre les niveaux ne sont pas exclusives au moment contemporain. De fait, un regard historique nous permet d'apprécier plusieurs traits communs que nous mettons en lumière.

Dans la deuxième partie de cet article nous menons l'analyse d'une sélection de réalisations datajournalistiques qui nous aident à identifier des tendances actuelles et qui nous fournissent des éléments pratiques pour suivre une démarche expérimentale en sciences de l'information et de la communication. Nos conclusions et travaux à venir sont aussi inspirés d'une production originale, développée à l'occasion de notre participation à la journée d'étude « Sémiotique du design d'information », organisée le 21 juin 2018 par le laboratoire MICA de l'Université Bordeaux Montaigne.

2. Aperçu historique et matériel des supports graphiques journalistiques

Du point de vue de la matérialité des médias, le support graphique fait d'abord référence à la surface sur laquelle une empreinte est tracée. On peut dire que le support graphique comprend les configurations physiques permettant qu'une inscription soit fixée et perceptible visuellement (Dondero et Reyes, 2016). La notion de *configuration* nous semble pertinente pour réunir la complexité de phénomènes qui agissent sur un affichage : les sources de lumière, la durée de la fixation, les textures du support, la luminosité.

La modification de supports graphiques pour permettre l'accueil de nouvelles formes d'inscription a toujours été un catalyseur d'évolution technique. En termes de journalisme visuel, nous pouvons situer comme point d'entrée la publication de la première photographie à demi-ton par le *Daily Graphic* à New York en 1880 (Lee, 2000, 129). Bien que le support graphique restât le papier journal, la nouveauté technique derrière cette impression consistait en l'utilisation du *halftone*. Breveté en 1852 par Henry Fox Talbot et perfectionné notamment par Max Levy en 1890, le principe du demi-ton était de placer un écran tramé entre l'objectif et le négatif pour que l'image soit formée de points de grosseur variable (Gervais 2010).

Rappelons que depuis le XV^{ème} siècle, la méthode pour préparer des images destinées à l'impression avait été essentiellement la gravure. Le bois se prête bien à la gravure en relief, utilisant des outils comme le burin pour dessiner sur un bloc de buis. Les plaques d'acier ou de cuivre étaient,

quant à elles, utilisées pour la gravure en creux. Au contraire du relief, l'encre restait dans les creux et s'imprimait par pression sur la feuille.

La gravure sur bois a été renouvelée par Thomas Bewick en 1791 et a attiré rapidement l'intérêt des journaux britanniques *The Times* et *The Observer* où des images apparaissaient sporadiquement (Jackson 1885). À partir des années 1830, les hebdomadaires illustrés ont fait leur début : *The Penny Magazine* (1832), *Le Magasin pittoresque* (1833), *The Illustrated London News* (1842), *L'illustration* (1843). Pour imprimer les illustrations avec les presses typographiques du XIX^e siècle, la priorité de la mise en page était donnée aux blocs de bois car, une fois gravés, ils ne pouvaient pas être modifiés, tandis que le texte pouvait se composer de façon plus souple autour des images (Gervais 2003). Une innovation qui mérite d'être soulevée ici est la presse rotative Ingram, conçue en 1876 par Walter Ingram spécifiquement pour *The Illustrated London News*. La solution technique était d'agrandir la plaque cylindrique sur laquelle le bloc de bois s'insérait, évitant ainsi une courbure plus prononcée du bloc (Wilson et Grey 1888).

Durant les années 1800, les techniques basées sur le demi-ton, comme la photogravure (1872), la rotogravure (1878) et la similitravure (1895), ont connu un plein essor et la mise en page des journaux comprenait systématiquement texte et image. Une plaque de zinc ou cuivre gravée servait de modèle à un matériau de carton nommé flan. Le flan, étant flexible, pouvait être courbé et servir de support pour couler une plaque en métal (en étain ou plomb) qui servait comme dispositif d'inscription pour les machines rotatives (Strulik et Kaplan, 2013).

Ce bref parcours permet d'observer l'intervention de différents supports graphiques à plusieurs étapes du processus d'impression. Pour conclure cette section, revenons sur la publication de la première photographie à demi-ton. Il est reconnu que c'est grâce à cette innovation que le photojournalisme s'est développé massivement, de pair avec une grande variété de supports techniques : plaques sèches (1864), films transparents (1884), l'appareil Kodak No. 1 (1888). En ce qui concerne son langage visuel, la photographie avait commencé à créer ses codes de représentation dans la presse (Gretton 2007), tandis que l'illustration se consacrait comme un média visuel (Drucker 2013).

3. Supports graphiques dans un contexte numérique

Le point d'entrée que nous avons choisi dans la section précédente permet d'établir quelques analogies avec notre contexte contemporain. Depuis la massification des appareils informatiques, l'écran électronique est devenu un endroit privilégié pour accéder aux contenus journalistiques. Dans un travail précédent (Reyes, 2017), nous avons décrit les contenus visuels à l'écran comme des « images-interfaces » car elles présentent la double fonctionnalité d'affichage graphique et d'interactivité. En effet, en tant que support graphique, l'écran est, comme le papier journal, la couche superficielle d'une série d'étapes d'inscription dont la principale propriété est le caractère dynamique. L'écran n'est donc pas un support fixe : à chaque seconde ses lignes se réécrivent pour rafraîchir les valeurs chromatiques des pixels (le plus courant à un taux de 60 Hz/s) et c'est ainsi que l'on perçoit le mouvement ou que l'on détecte un événement d'interaction (par exemple, pointer, cliquer, glisser-déposer).

Ainsi, l'affichage de contenus ne dépend pas exclusivement de la matérialité de l'écran. En dessous de la surface, c'est la couche de logiciels qui est sollicitée pour la gestion d'images et de textes. À considérer le Web comme le média répandu pour le datajournalisme, dans cette section nous mettons l'accent sur les capacités techniques des navigateurs tels que Mozilla Firefox ou Google Chrome.

Pour afficher des images sur les navigateurs, la balise a été introduite dans le langage HTML en 1993. Ensuite, d'autres innovations sont apparues comme <embed> (1995) et <object> (1996), et le langage de styles graphiques CSS 1.0 (1996). Ces avancements ont permis à la presse de concevoir une version en ligne et multimédia des contenus journalistiques. De fait, les premières pages Web des journaux apparaissent dans les années 1990 : *Le Monde* (1995), *The New York Times* (1996), *El País* (1996), entre autres.

La modularisation de contenus au XIX^e siècle, qui consistait à préparer les images sur des supports externes à la mise en page, était semblable aux débuts du Web. Les contenus visuels et interactifs se réalisaient avec des logiciels de publication assistée par ordinateur (PAO) et étaient ensuite exportés vers le Web dans un format compressé. Les différents

producteurs de solutions multimédia ont créé leurs propres plugins pour pouvoir afficher leurs formats spécifiques avec les balises <object> et <embed>. Par exemple, Macromedia Flash exportait vers le format SWF (utilisant le plugin Flash Player) et Director vers DCR (utilisant le plugin Shockwave).

Malgré le fait que les navigateurs contemporains ne supportent plus les plugins autrefois populaires comme Java, Shockwave et ActiveX, la recommandation HTML5, lancée en 2014, contient deux balises dédiées à l'affichage de contenus visuels : <canvas>, basée sur une grille de pixels et similaire dans sa logique à Director, et <svg> 2.0, spécifique aux images vectorielles et animations similaires à Flash.

Dans ce contexte, l'une des innovations graphiques apportées au Web par le datajournalisme a été la bibliothèque D3 (Data-Driven Documents). Inspiré du plug-in Flare pour Flash, initié par son directeur de thèse Jeffrey Heer en 2008 à Stanford, Mike Bostock avait développé le prédécesseur de D3, la bibliothèque Protovis, comme un projet d'étude. En 2011, il a présenté D3 à la conférence IEEE InfoVis et a rejoint le département graphique du *New York Times*.

D3 doit sa popularité à l'efficacité de manipulation d'éléments <svg> à l'aide de Javascript. En outre, il permet l'utilisation de divers algorithmes de visualisation de données, ces dernières structurées au format CSV et JSON. Enfin, la bibliothèque a toujours été un projet ouvert et libre. Au sens large, on peut dire que D3 a catapulté l'usage de <svg> comme support graphique pour le design d'infographies sur le Web.

Dans le panorama actuel du Web, les navigateurs fonctionnent en même temps comme support graphique et comme plateforme de production. Il existe une large gamme de logiciels sous forme de webapps dédiés au traitement de données et leur représentation graphique, entre autres : Google Maps (2005), Datawrapper (2012), RAWGraphs (2013) et Tableau Online (2013). D'un autre côté, les éléments HTML5 <canvas>, <video>, <audio> et <embed>, de la main avec des bibliothèques comme jQuery, constituent les supports graphiques *de facto* pour les reportages interactifs et les récits de données. D'autres technologies, mises à jour à partir de 2017 (CSS3, WebGL 2.0, WebRTC, Web Audio API), visent à

exploiter la visualisation et interactivité 3D, les environnements immersifs et la réalité augmentée.

4. Journalisme graphique

À partir de notre chronologie d'innovations techniques, nous pouvons aborder maintenant quelques passages de l'évolution des représentations graphiques dans la presse.

En 1848 (trente-deux ans avant l'impression de la première photographie à demi-ton), le journal *L'Illustration* avait publié les premières gravures faites à partir d'un daguerréotype (Gervais 2003). L'auteur, Charles Thibault, avait pris deux vues de la rue Saint Maur, avant et après les combats de juillet 1848. Ces gravures ont été publiées au format ovale, composées côte à côte sur une demi-page (environ 16 cm de hauteur par 14 cm de largeur). Quant à la photographie à demi-ton publiée en 1880, le photographe Stephen Horgan l'avait intitulée "A scene in Shantytown, New York" avec la légende "reproduction direct from nature". L'espace dédié sur la feuille était de 72 lignes (une hauteur d'environ d'une demi-page tabloïd), sans bords ni décorations supplémentaires.

Dans ces deux représentations, c'est la rue qui avait bénéficié du style réaliste de la photographie et sa lecture se prêtait bien à plusieurs interprétations politiques et sociales. Le fait de publier ces images réaffirmait également une posture idéologique et documentaire des journaux. Suivant cette ligne, le photojournalisme avait fini par s'instaurer comme un genre visuel vers la fin des années 1920, avec diverses sous-catégories : photo reportage, photo sensationnelle, photo de célébrités (Gernsheim 1960).

Concernant les graphiques, nous nous référons à la première carte météorologique publiée dans un journal. Elle a été préparée par le scientifique Francis Galton pour *The Times* à Londres en 1875. À partir de ses travaux en représentation climatique, Galton avait conçu une version simple et facile à comprendre par un public large. Cette publication était arrivée à un moment propice en termes de recherches en visualisation d'information. Les travaux pionniers de William Playfair sur les

représentations diagrammatiques (Playfair 1786) avaient commencé à se répandre avec un plein essor au milieu du XIX^{ème} siècle, avec les contributions de Charles Joseph Minard et d'Étienne-Jules Marey.

Le second moment des graphiques dans la presse est venu après la crise boursière de Wall Street en 1929. Le *New York Times*, par exemple, avait publié son premier diagramme de cours des actions en 1933 (Kopf 2016). Bien que d'autres graphiques aient déjà été publiés auparavant, leur parution restait épisodique, limitée à quelques-uns par an. Les thèmes des graphiques dans des journaux comme *L'Humanité*, *Le Temps* et *Le Figaro* traitaient des sujets scientifiques, industriels et économiques.

Si pendant la première moitié du XX^e siècle la presse avait exploité largement les graphiques simples (notamment les courbes ou *line chart*), nous situons un troisième moment des graphiques dans la presse à partir de l'introduction de l'informatique visuelle dans les années 1960 et du design numérique dans les années 1980. Sur cette ligne, *USA Today* a fait des contributions en termes d'accessibilité et familiarité. Créé en 1982, il inclut trois ou quatre "snapshots" par numéro : des petites vignettes qui renforcent les sections *News*, *Money*, *Sports* et *Life*. Les caricatures, les couleurs brillantes et les thèmes sont traités de manière humoristique et se présentent comme des exemples de métaphores visuelles (Bolter 1991, 52).

Avec l'émergence du Web, les grands journaux ont commencé l'intégration d'équipes spécialisées en traitement et représentation graphique de l'information. L'un des premiers a été le département graphique du *New York Times*. Établi en 1995, il a consolidé un groupe talentueux : Steve Duenes (1999), Shan Carter (2004), Amanda Cox (2005) et Mike Bostock (2012). S'il est vrai qu'au début les graphiques étaient destinés à la publication papier, à partir de 2010 les reportages

interactifs¹ et les récits de données² occupent une place prépondérante dans les versions en ligne des journaux.

5. Données journalistiques

Nous utilisons la notion de « données journalistiques » pour faire référence aux catégories pouvant servir à décrire une réalisation numérique de presse. En principe, ces catégories pourraient être structurées suivant un modèle de documentation existant. Par exemple, pour une infographie imprimée on obtiendrait : Titre, Chapeau, Légendes, Texte d'explication, Référence de la source, Crédit auteur (George-Palilonis 2006, 59). Mais la datavisualisation sur le web se refaçonne constamment et la constitution de nouvelles catégories doit être liée à une analyse exploratoire. Certes, la tâche est complexe et nécessite une optique pluridisciplinaire, mais nous esquissons dans cette section un brouillon méthodologique.

Nous commençons par distinguer trois aspects qui se manifestent au niveau de la surface : le journalistique, le visuel et l'interactif. Le premier aspect met l'accent sur la manière dont les informations sont textualisées, c'est-à-dire la forme dont les textes incarnent les valeurs et les choix hiérarchiques éditoriaux selon la méthode empirique « Qui ? Quoi ? Où ? Quand ? Comment ? Combien ? Pourquoi ? ». Le deuxième aspect se concentre sur la visualisation et ses codes graphiques. Le troisième aspect fait appel au support technique et à ses potentialités. Nous faisons allusion ici aux types de gestes corporels des utilisateurs.

Lors de la journée d'étude « Sémiotique du design d'information », nous avons sélectionné trois exemples de datavisualisation créés par *The New York Times*. Ces productions exploitent les valeurs caractéristiques de la datavisualisation traditionnelle, basée sur la « séquentialité », la « variabilité » et le « regroupement » de données quantitatives. En même temps, ces trois cas proposent une innovation graphique accompagnée

1 "Snow Fall", <http://www.nytimes.com/projects/2012/snow-fall/>

2 "Bob Beamon's Long Olympic Shadow", (nécessite Flash Player): <http://archive.nytimes.com/www.nytimes.com/interactive/2012/08/04/sports/olympics/bob-beamons-long-olympic-shadow.html>

d'une intervention technique importante. Le tableau 1 organise ces trois exemples.

Tableau 1. Visualisations journalistiques

Titre	Créateurs	Technique	Technologies
Front Row to Fashion Week (2014) ³	Mike Bostock, Shan Carter, Erik Hinton et Ruth La Ferla	Visualisation d'images	JS, CSS, HTML
A 3-D View of a Chart that Predicts the Economic Future (2015) ⁴	Gregor Aisch et Amanda Cox	Graphique 3D	WebGL, HTML
The 598 People, Places and Things Donald Trump Has Insulted on Twitter (2016) ⁵	Jasmine Lee et Kevin Quealy	Liste de tweets	D3, CSS, HTML

Au niveau journalistique, les textes des chapeaux et des légendes sont courts et précis. Les titres énoncent le sujet tandis que les légendes expriment d'abord ce que l'on doit voir et ensuite ils renvoient vers une lecture de ce qui est représenté. Au niveau visuel, ils gardent leur propre homogénéité graphique, avec une palette de couleurs peu contrastée et sans recours aux tonalités brillantes. Même dans le cas de « Fashion Week », où la variété de couleurs est précisément le point d'attractivité, l'organisation uniforme des figures est un mécanisme qui aide à renforcer la perception d'homogénéité.

Au niveau interactif, les trois réalisations diffèrent substantiellement. « Fashion Week » est un exemple typique d'image-interface car la visualisation ne sollicite pas explicitement de survoler les images pour dévoiler l'effet d'agrandissement. L'utilisateur est attiré par le cadrage, la juxtaposition des images et les textures des vêtements illustrés. C'est

3 <http://www.nytimes.com/newsgraphics/2014/02/14/fashion-week-editors-picks/index.html>

4 <https://www.nytimes.com/interactive/2015/03/19/upshot/3d-yield-curve-economic-growth.html>

5 <https://www.nytimes.com/interactive/2016/01/28/upshot/donald-trump-twitter-insults.html>

l'image elle-même qui renvoie à l'affordance d'interactivité. Ce cas est similaire aux « insultes de Trump ». Après lecture du chapeau, l'utilisateur qui clique sur le texte d'un message Twitter sera renvoyé vers sa localisation sur le réseau social et saura que tous les autres textes fonctionnent de la même manière. Les actions proposées à l'utilisateur directement via l'interface graphique sont : possibilité d'ordonner par date et de dérouler verticalement la page. Enfin, avec le « 3D Chart », il s'agit d'un autre cas d'image-interface car ce n'est qu'au moment où l'on clique sur le bouton « next » que la courbe fait une rotation sur l'axe horizontal. Pour les habitués des interfaces 3D, les règles ergonomiques sont respectées : cliquer et laisser le clic enfoncé pour faire pivoter librement le graphique.

Ces exemples rappellent l'importance de maintenir une description ouverte et non-figée des données journalistiques. Ces données peuvent être organisées dans un tableau et contenir des dimensions variées comme : quantité de texte/couleurs, fréquence de thèmes/couleurs, métaphores visuelles, figures rhétoriques, références sociales et géographiques, paradigme de programmation informatique. L'interactivité est aussi un élément à prendre en compte pour toute description car elle participe à la signification discursive. En effet, si le texte énonce les règles du jeu et les faits informatifs, les images renforcent le domaine esthétique, émotionnel et subjectif des producteurs. Pour notre part, nous dirons que l'interactivité met l'accent sur la signification technique des interfaces : c'est le canal à travers lequel les journaux transmettent les valeurs de modernité, actualité, fonctionnalité, adaptabilité et originalité.

6. Magazines visualisés

Afin d'appliquer notre notion de « données journalistiques », nous avons développé divers prototypes de datavisualisation à l'occasion de la journée d'étude « Sémiotique du design d'information ». Pour nous, l'intérêt de cette démarche expérimentale réside dans la manipulation d'informations et d'interfaces numériques elles-mêmes, autrement dit, les composants techniques des supports graphiques sont considérés comme le terrain d'enquête.

Pour faciliter nos explorations, nous avons utilisé la plateforme Google Play Newsstand car elle réunissait la plupart de journaux et magazines dans un même endroit. De plus, chaque numéro était dûment documenté avec l'image de couverture et autres métadonnées. Le corpus que nous avons créé contient 253 couvertures de magazines représentant la totalité de numéros publiés entre mars 2012 et mars 2018 par *Wired*, *Entrepreneur*, *Bazaar* et *Fashion*⁶. Ces magazines étaient listés parmi les plus populaires dans leur catégorie : « technologie », « business », « mode » et « style ».

L'aspect visuel de notre corpus a été décrit en termes d'informations visuelles quantitatives. Pour chaque image nous avons extrait les valeurs chromatiques les plus représentées à l'aide de l'outil web TinEye Extractor⁷. Ensuite, nous avons détecté automatiquement les objets, les visages et les labels avec Google Vision API, un service de vision par ordinateur du type *cloud computing*.

La corrélation quantitative entre labels et images peut donner forme à des regroupements d'images. En utilisant le programme ivpy⁸, qui fonctionne sur l'application web Jupyter, nous avons produit des clusters selon les catégories des visages détectés : *joy*, *neutral*, *headwear* et *surprise* (figure 1).



Figure 1. Labels de la reconnaissance faciale

6 Mars 2012 était la date la plus ancienne à laquelle on pouvait accéder. Par ailleurs, Google Play Newsstand a disparu peu après la journée d'étude en 2018.

7 <https://labs.tineye.com/color/>

8 Développé en 2015 par Damon Crockett, membre du Cultural Analytics Lab.

Une autre forme graphique pour exploiter des corrélations est le diagramme de réseau. Au total, nous avons obtenu une quarantaine de labels et nous les avons spatialisés avec l'outil en ligne Palladio⁹. La figure 2 montre les interconnexions entre les labels et les magazines. En vue d'ensemble, cette carte positionne les magazines selon leur degré de distanciation par les labels détectés automatiquement.

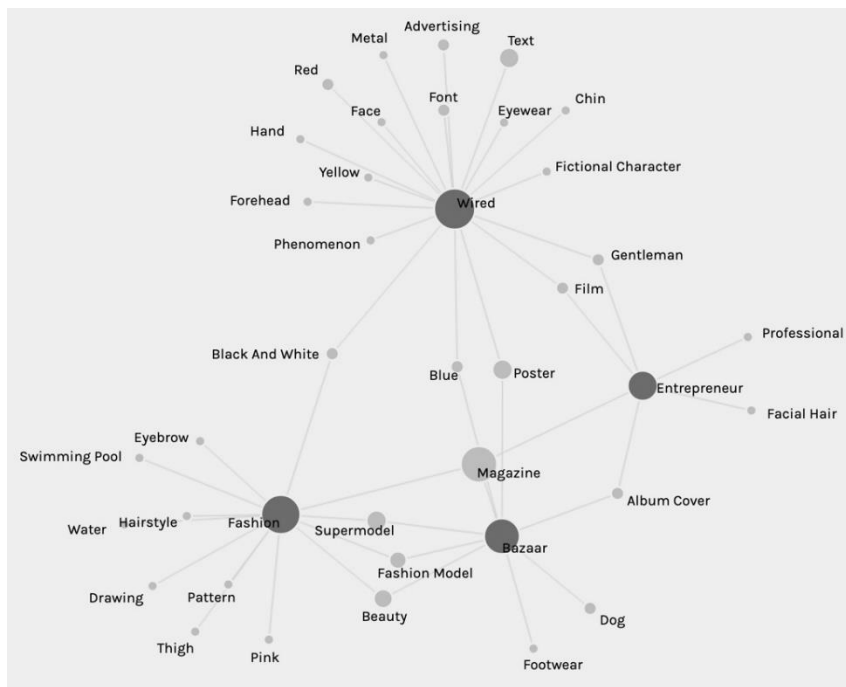


Figure 2. Diagramme de réseau de titres et labels détectés

Du point de vue thématique, ces schémas peuvent nous aider à vérifier le degré de fiabilité d'un algorithme *deep learning*. Nous avons seulement confirmé quelques intuitions de départ : quel magazine serait le plus en relation avec la notion de beauté ? Quel magazine aurait le style graphique le moins traditionnel ? Mais cet exercice laisse à découvrir les défauts et les inattendus de l'état actuel du *machine learning* où l'un des problèmes principaux est associé à l'opacité de ses mécanismes. En effet, les outils se présentent souvent comme une boîte noire dont les résultats

⁹ <http://hdlab.stanford.edu/palladio/>

du processus nous sont livrés de manière presque magique : sans savoir comment nous y sommes parvenus. Les enjeux politiques et techniques semblent s'imposer devant la visibilité et la transparence des supports graphiques.

Nous avons complété nos expériences avec la réalisation d'une visualisation dont les informations ont été obtenues par observation humaine. Dans ce but, nous avons suivi une logique d'oppositions binaires. Notre démarche consistait à réduire les thèmes (l'aspect journalistique) de chaque couverture à son degré le plus basique, c'est-à-dire au point d'avoir une transformation narrative à deux états. Ces oppositions peuvent se manifester comme deux termes énoncés explicitement dans le texte et comme deux tonalités chromatiques. À ce sujet, le travail humain consistait à choisir la tonalité qui correspond à chaque terme. Par ailleurs, cette tonalité n'était pas forcément la couleur la plus fréquente sur l'image, mais uniquement celle d'un détail de l'image. La figure 3a montre l'ensemble de thèmes de notre corpus et la figure 3b les couleurs associées aux thèmes. Dans ce cas, le thème « disaster » est lié à la tonalité #b37163 qui comprend moins de 5 % de l'espace chromatique de l'image.

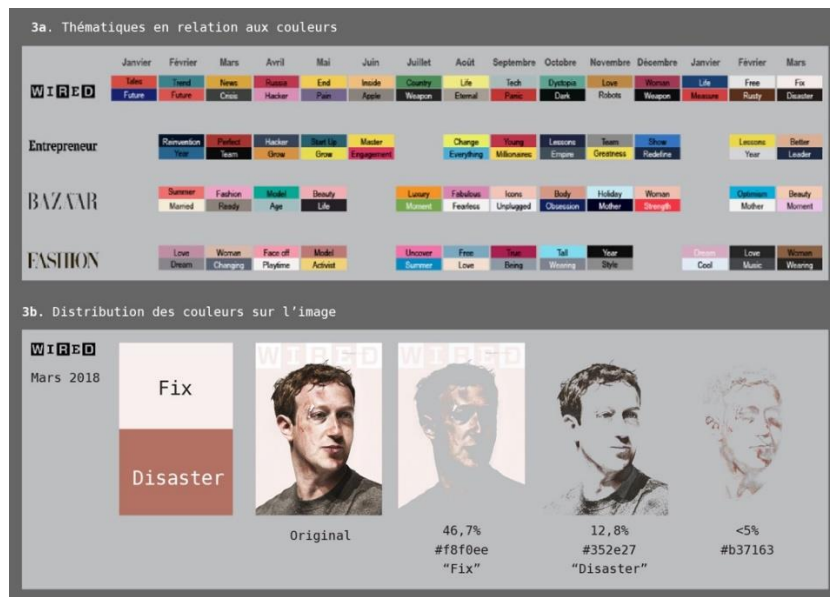


Figure 3. Thèmes et signes plastiques des couvertures

Pour aller plus loin, la figure 4 montre les tonalités de la figure précédente dans l'ensemble de possibles selon un cercle chromatique. La finalité du schéma est de visualiser l'axe paradigmatique selon le mode de couleur RVB. Bien que ce graphique mérite un développement interactif pour naviguer et filtrer les corrélations ainsi que la variation dans le temps, le mode bidimensionnel laisse voir quelques tendances des quatre magazines réunis : une accentuation de l'opposition orange-rouge avec la gamme de bleus et une absence de tonalités rose-violet et vert brillant.

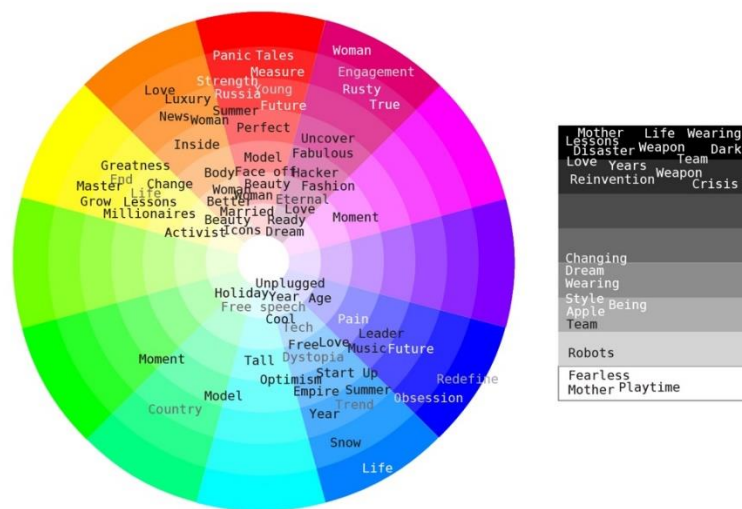


Figure 4. Thèmes des magazines dans la totalité des couleurs

7. Conclusion

Dans cet article nous avons abordé la notion de « données journalistiques » dans le contexte historique et matériel du journalisme graphique. Après avoir retracé ses principales lignes évolutives et le développement de sa représentation graphique, nous avons offert l'exemple d'une analyse appuyée sur trois aspects interdépendants : le thématique, le visuel et l'interactif.

Au moment où le datajournalisme continue d'explorer son langage, ses genres et ses méthodes, nous considérons nécessaire un regard sur les

composants techniques des supports graphiques. En effet, la croissance exponentielle d'informations disponibles sur le web pousse à l'emploi d'outils informatiques pour organiser, décrire, préserver et explorer les ressources numériques. Or, ces outils s'intègrent dans un contexte plus large où les facteurs politiques, économiques et idéologiques s'entremêlent à la technique. À l'heure où les datavisualisations mettent en scène des données en temps réel et faisant recours à l'intelligence artificielle, les données journalistiques sont une manière d'explicitier et de documenter les couches techniques et leurs implications sur la surface visible.

Nous remercions vivement Pauline Hachette pour sa relecture et commentaires.

Bibliographie

- Bolter Jay (1991). *Writing space*. Erlbaum, Londres.
- Dondero Maria Giulia, Reyes Everardo (2016). Les supports des images : de la photographie à l'image numérique, *Revue française des sciences de l'information et de la communication*, No. 9, <https://journals.openedition.org/rfsic/2124>
- Drucker Johanna, McVarish Emily (2013). *Graphic design history*, Pearson, Boston.
- George-Palilonis Jennifer (2006). *A practical guide to graphics reporting*. Elsevier, Boston.
- Gernsheim Helmut (1960). *Creative photography*. Bonanza, New York.
- Gervais Thierry (2003). D'après photographie, *Études photographiques*, No. 13, pp. 56-85.
- Gervais Thierry (2010). La similigravure, *Nouvelles de l'estampe*, No. 229, pp. 8-25.
- Gray Jonathan, et al. (2012). *The data journalism handbook*, O'Reilly, Sebastopol.
- Gretton Tom (2007). Le statut subalterne de la photographie, *Études photographiques*, No. 20, <http://journals.openedition.org/etudesphotographiques/927>
- Kopf Dan (2016). When Did Charts Become Popular?, <https://priceconomics.com/when-did-charts-become-popular/>
- Jackson Mason (1885). *The pictorial press*, Hurst and Blackett, Londres.

Lee Alfred (2000). *The daily newspaper in America vol. 1*, Routledge, Londres.

Playfair William (1786). *The Commercial and Political Atlas*, Debrett, London.

Reyes Everardo (2017). *The image-interface: graphical supports for visual information*, Hoboken, NJ, Wiley-ISTE.

Stulik Dusan, Kaplan Art (2013). *Halftone*, Los Angeles, Getty Institute.

Wilson Fred, Grey Douglas (1888). *A practical treatise upon the modern printing machinery*, Cassell & Co., Londres.