Modéliser la démarche du data journaliste

Une approche nécessairement transdisciplinaire

Marie Chagnoux* — Raphaël Da Silva — Faten El Outa*** Nicolas Labroche*** — Patrick Marcel***—— Veronika Peralta*****

** Centre d'études sur les médias, les technologies et l'internationalisation (CEMTI) Université Paris 8, Maison de la Recherche, 2 rue de la Liberté, F-93526 St Denis*
Marie.Chagnoux@univ-paris8.fr

*** Journaliste de données freelance*
raphipons@gmail.com

**** Laboratoire d'Informatique Fondamentale et Appliquée de Tours (LIFAT) Université de Tours, 3 place Jean Jaurès, 41000 Blois*
{Faten.El-Outa, Nicolas.Labroche, Patrick.Marcel, Veronika.Peralta}@univ-tours.fr

RÉSUMÉ. *La mise en œuvre effective de la transdisciplinarité reste une gageure dans les projets qui rassemblent des acteurs disciplinaires alors qu'elle est indispensable pour modéliser des problèmes complexes, comme la recherche et le traitement de l'information pour produire des narrations à partir de données. Dans cet article, nous expliquons en quoi cette approche, appliquée au cas du data journalisme, a fait émerger en amont une problématique de recherche et a ébauché en aval une méthodologie et un langage communs, notamment par le recours à des modèles conceptuels. De par ses formalisations graphiques, cet outil, initialement pensé comme une étape pour la conception logicielle, a permis à tous les acteurs du projet de construire et valider le modèle au fur et à mesure.*

ABSTRACT. *Transdisciplinarity implementation remains a challenge in projects that bring together disciplinary actors while it is essential for modeling complex problems, such as data analysis to produce narratives. In this article, we explain how this approach, applied to the case of data journalism, allowed us upstream to identify a research problem and downstream to build draw a methodology and a common language, in particular through the use of conceptual models. Thanks to its graphic formalizations, this tool, initially thought of as a step for software design, allowed all the actors of the project to build and validate the mode at each stage.*

MOTS-CLÉS : *transdisciplinarité, data journalisme, modèle conceptuel*
KEYWORDS: *transdisciplinarity, data journalism, conceptual modeling*

1. Introduction

Du point de vue épistémologique, interdisciplinarité et pluridisciplinarité[1] se heurtent aujourd'hui à un paradoxe. Si elles sont reconnues et même encouragées pour résoudre des problèmes complexes, notamment lors des appels à projets, elles restent difficiles à mettre en œuvre. Formation et recherche favorisent en effet les frontières disciplinaires en isolant les communautés scientifiques dans des langages, des méthodes, des critères d'évaluations et des pratiques étanches.
Dans ce contexte, la notion de transdisciplinarité (Piaget, 1967), notamment conceptualisée par (Morin, 1994) et (Nicolescu, 1996), pourrait apparaître comme un mirage que poursuit temporairement le chercheur avant de s'en détourner finalement pour des objectifs plus productifs à court terme. La posture transdisciplinaire s'accommode mal du fonctionnement actuel de la recherche par projet qui s'inscrit dans des temporalités restreintes et vise à la production rapide de résultats aisément évaluables. Et trop souvent des projets pensés et défendus comme transdisciplinaires n'aboutissent finalement qu'à des travaux au mieux pluridisciplinaires, tant le décentrage peut décourager. « Penser la transdisciplinarité est donc une tâche instable, que le sujet abandonne pour lui préférer le retour aux routines […]. La pensée monodisciplinaire est rassurante, la pensée pluridisciplinaire et la pensée interdisciplinaire sont plus difficiles, mais elles peuvent s'appuyer sur une pensée stabilisée. La pensée transdisciplinaire est toujours en action, toujours en dépense d'énergie, toujours donc facile à abandonner, d'autant qu'elle ne s'inscrit pas bien dans les espaces que nous réserve la société » (Claverie, 2010)

Nous nous proposons ici de revenir en détails sur le travail transdisciplinaire que nous menons depuis 2017 au sein du GdR Madics[2], Groupement de Recherche dont l'objectif est justement d'encourager les activités d'animation interdisciplinaires sur les masses de données scientifiques. Au sein d'un atelier, CAJOLE (2018), puis d'une action, MADONA (2019-2022), nous avons travaillé à faire dialoguer les disciplines des sciences de l'information et de l'informatique tout en intégrant à notre collaboration des data journalistes pour « établir de véritables connexions entre concepts, outils d'analyse et modes d'interprétation de différentes disciplines » (Charaudeau, 2010). Nous aborderons donc notre cadre méthodologique, dans ses limites et ses richesses, à travers l'exemple de la modélisation du processus de mise en narration des données qu'effectuent les data journalistes. Nous montrerons en quoi associer les approches et les regards et multiplier les aller-retours entre tous les intervenants, à toutes les étapes, a permis des fertilisations croisées par-delà les résultats scientifiques. Nous montrerons enfin le rôle fondamental de la modélisation conceptuelle, point d'articulation méthodologique, qui offre un langage commun à toutes les étapes de réflexion, de négociation et finalement de validation.

---

[1] Nous reprenons comme principes définitoires la gradation usuellement adoptée, allant de la *pluridisciplinarité*, mise ensemble d'une pluralité de disciplines, à l'*interdisciplinarité*, échange collaboratif de disciplines puis à la *transdisciplinarité*, ouverture par-delà les disciplines (Létourneau, 2008).
[2] https://www.madics.fr/

Cet article se veut ainsi une réflexion épistémologique, le regard rétrospectif et réflexif que nous portons sur un cadre méthodologique en co-construction.

2. De la genèse aux résultats

Pour expliquer comment l'approche transdisciplinaire s'est imposée au fil du projet, il nous paraît pertinent de revenir sur les différentes étapes de la collaboration pour montrer comment elle a naturellement émergé pour ensuite structurer notre méthodologie.
Nos travaux trouvent leur origine dans un domaine, celui de l'exploration interactive de données, ou *interactive data exploration* (Idreos *et al*. 2015), un champ de l'informatique rattaché à l'analyse de données qui vise à concevoir des algorithmes et des outils pour faciliter l'exploitation et la compréhension des grandes masses de données accessibles aujourd'hui. C'est une problématique dans laquelle l'analyste est central puisque le but consiste à adapter le système et ses possibilités d'interaction dans le but de simplifier des tâches complexes ou répétitives d'analyse de données. Cela passe par la construction de plusieurs requêtes et traitements successifs sur différentes sources de données qui forment ce que l'on appelle une *exploration*. Comprendre le déroulement d'une analyse et à quel niveau apparaissent les découvertes, ou *insights* (Eichman et al. 2020), de nouvelles informations qui motivent les requêtes suivantes de l'exploration sont un problème majeur.
Cette tâche d'exploration autrefois dédiée à des analystes spécialistes de langages d'interrogation de données comme *SQL*, s'ouvre, en même temps que la mise à disposition croissante de données, à des publics plus variés et ne disposant généralement pas du bagage technique ou scientifique pour produire des explorations de qualité. Dans ce contexte, nous avons cherché à identifier dans l'ensemble des non spécialistes de la donnée un domaine qui aurait des besoins d'analyse et d'extraction d'information à partir de données ouvertes. Le data journalisme a alors été identifié comme un cas d'usage pertinent.

2.1. CAJOLE : Corpus pour l'Analyse de la qualité du JOurnaLisme d'Exploration des données

Des contacts ont été mis en place avec des spécialistes de ce domaine et c'est lors de cette confrontation que la nécessité d'une approche réellement transdisciplinaire est apparue. Notre première approche avait consisté à rassembler différents acteurs : journalistes, chercheurs en recherche d'information (RI), traitement automatique des langues (TAL), science de l'information et de la communication (SIC) ou gestion de connaissances, spécialistes de données ouvertes, spécialistes de l'analyse de données. Notre objectif était d'élaborer un corpus permettant d'étudier la qualité du journalisme de données en définissant des indicateurs. L'idée première était naïvement de rapprocher les méthodologies de traitement de celles faites classiquement en sciences de données pour identifier les « bonnes pratiques » du journalisme de données qui pourraient par la suite permettre d'évaluer de manière automatique des articles.

L'atelier CAJOLE, proposé à cet égard lors de la conférence CORIA-TALN 2018[3] , a alors fait émerger un certain nombre de problèmes dans l'angle adopté. Tout d'abord, en termes de compétences, comme attendu, les journalistes ne disposaient pas de l'ensemble des outils qu'un analyste spécialiste de la science des données peut mobiliser, notamment la maîtrise des langages d'interrogation de données comme SQL ou la compréhension des outils statistiques les plus avancés comme le *machine learning*. Il était donc difficile de faire des parallèles entre les activités des uns et des autres.
Plus important, l'atelier a révélé que la notion d'exploration de données n'existait pas à proprement parler en data journalisme car, au-delà des compétences nécessaires, la nature du travail d'analyse des données s'intégrait souvent difficilement avec les fonctionnements des rédactions. Les grandes enquêtes internationales très médiatisées comme celle des *Panama Papers*[4] ou d'*Openlux*[5] apparaissent "artisanales" du point de vue méthodologique et technique, elles nécessitent d'importantes équipes pluridisciplinaires de journalistes et restent marginales et peu représentatives du travail quotidien des data journalistes.
La conclusion de l'atelier fut donc qu'il n'y avait pas à ce jour de méthodologie clairement établie et communément acceptée par les journalistes, à la différence de l'informatique qui dispose de modèles, même génériques, comme CRISP-DM. Nous avons donc entrepris un travail collectif pour comprendre comment ces données étaient empiriquement travaillées pour la production de narrations. Les données, en tant que matériau technique, sont alors devenus ces « objets frontières » (Star et Griesemer, 1989) entre chercheurs et journalistes, dans un dispositif interactionniste comme a pu le décrire (Flichy, 2003 :23) : « [c]e n'est plus un acteur qui cherche à imposer sa propre vision du monde aux autres acteurs dont il a besoin, mais on assiste au contraire à l'élaboration d'un compromis ».

2.2 MADONA, Maîtriser l'Analyse interactive de DOnnées pour la NArration journalistique

Au sein de l'action MADONA, toujours dans le cadre du GdR Madics, nous avons ainsi reformulé les questions scientifiques en vue : (i) d'acquérir une meilleure compréhension des mécanismes de sélection et de traitement des données qui permettent de construire progressivement des narrations et (ii) de viser à la création et la mise à disposition d'outils permettant aux journalistes d'explorer et analyser des données, notamment ouvertes, avec une interaction simplifiée. L'objectif était, à partir d'un travail de terrain, de formuler des hypothèses, d'exprimer des primitives de haut niveau afin de sélectionner ou d'affiner des propositions de modèles.
Nous avons alors constitué une équipe composée de chercheurs en informatique, spécialistes de l'exploration interactive de données, de chercheurs en science de

[3] http://www.info.univ-tours.fr/~marcel/cajole2018/
[4] https://www.lemonde.fr/panama-papers/
[5] https://www.lemonde.fr/les-decodeurs/visuel/2021/02/08/openlux-enquete-sur-le-luxembourg-coffre-fort-de-l-europe_6069132_4355770.html

l'information et de la communication, spécialistes des pratiques de production et réception numérique et d'un journaliste de données. Nous l'avons adossée à un groupe d'une quinzaine de data journalistes volontaires de la Presse Quotidienne Régionale (*Ouest France*, *Le Parisien*, *L'Est Républicain*…) et à une formation universitaire, celle du Master *Journalisme et Médias Numérique* de l'Université de Lorraine, dans le cadre de son cours sur le data journalisme, pour recueillir les données et valider les étapes de nos travaux.

3. Les résultats scientifiques obtenus

La méthodologie suivie lors des premières années du projet s'appuie sur les méthodes agiles, populaires en ingénierie logicielle. En plaçant la communication et la collaboration au centre de la réflexion, elles permettent un développement évolutif et continu, encourageant la co-construction transdisciplinaire. La figure 1 présente un synoptique emblématique de la pratique agile (schéma de gauche) ainsi que notre déclinaison (schéma de droite).

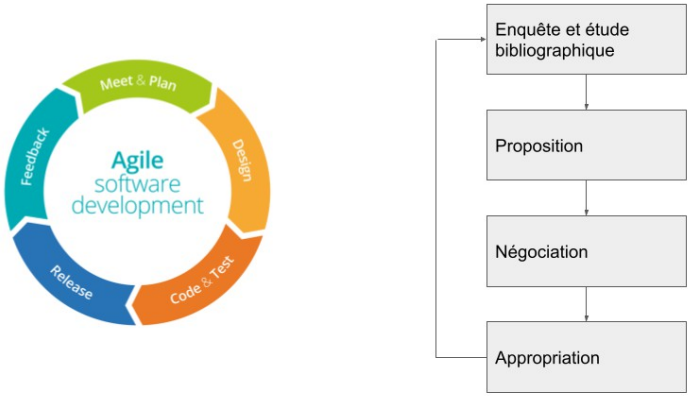


Figure 1 : *principe méthodologique*

Le schéma de la figure 2 illustre les étapes suivies et les livrables obtenus pendant les deux premières années du projet.

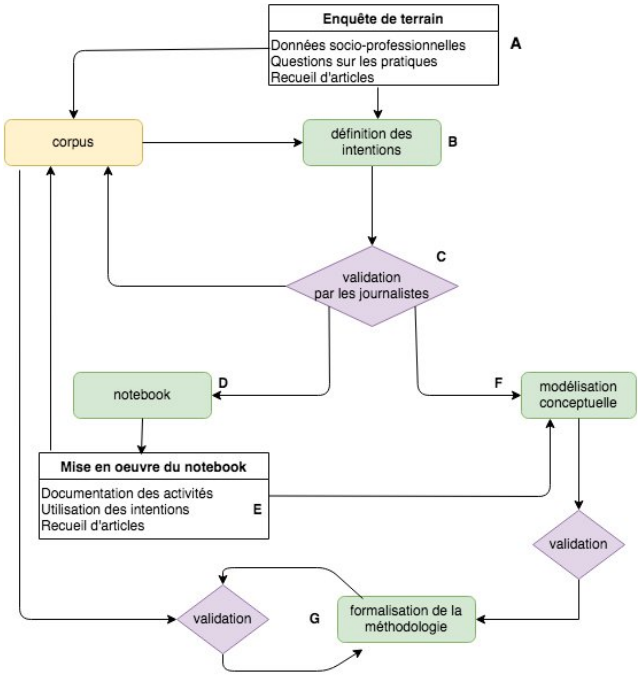


Figure 2 : *modélisation des interactions et des productions*

Un premier travail de terrain (A) a permis de documenter l'activité journalistique et de collecter des informations et des sources à partir d'un sondage en ligne sur les usages, les pratiques et les productions. Ces éléments ont servi de base à une première formalisation à l'aide d' « intentions» (B). Ces intentions, au nombre de cinq (*Décrire*, *Comparer*, *Expliquer*, *Prédire* et *Suggérer*), peuvent être assimilées à des primitives ou composantes et visent à faciliter l'expression de tâches complexes de transformation et d'analyse des données. Elles ont été soumises pour validation et application aux journalistes (C) lors d'une seconde enquête de terrain.
La figure 3 présente quelques résultats de ce travail de validation des intentions : avec des questions sur une intention précise, ici l'intention « Expliquer » et la validation de scénarisation en fonction de l'ordre de succession des intentions.

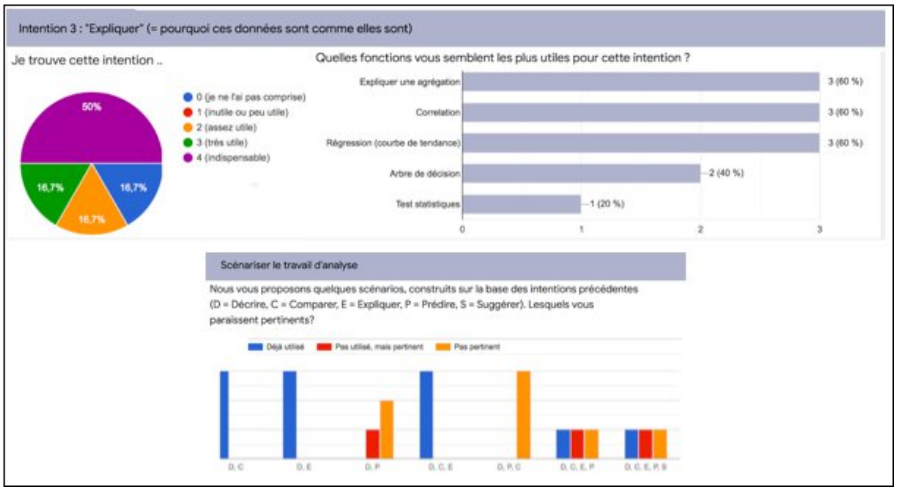


Figure 3 : *Extrait des résultats du questionnaire de validation des intentions*

Par la suite, ces intentions ont été intégrées à un notebook (D) qui a été soumis à des étudiants lors de travaux universitaires (E). Ce notebook visait à tracer le plus fidèlement possible l'enchaînement des tâches associées à chaque intention lors du travail d'enquête, les participants devaient décrire aussi exhaustivement que possible les différents scénarios explorés. L'exploitation de ce notebook et les articles produits lors de ce travail ont alors permis d'affiner la modélisation du domaine (F). A partir de cette modélisation et de la partie méthodologique contenue parfois dans les articles, nous avons proposé le modèle de narration sur lequel nous travaillons actuellement (G).¶

Comme indiqué plus haut, l'objectif de l'action MADONA était la réalisation d'une application qui viendrait en appui aux journalistes pour la création de narrations de données. Or le préalable à toute réalisation logicielle est un cahier des charges s'appuyant justement sur cette modélisation formelle du domaine. En croisant nos données et un modèle conceptuel du domaine de la narration de données inspirés par (Chatman, 1980) et (Chen et al., 2018), nous avons proposé le modèle en quatre couches illustrées dans la figure 4.

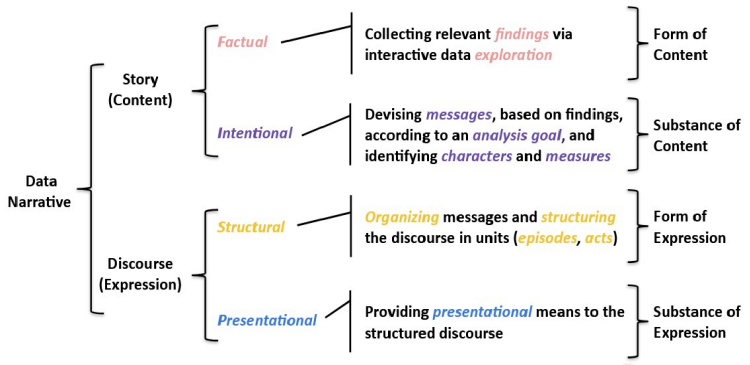


Figure 4 : *Couches du modèle de narration de données, inspirés de (Chatman, 1980)*

Les deux premières couches représentent l'histoire et les deux dernières représentent le discours. Au niveau de l'histoire, une couche « factuelle » représente la forme tandis qu'une couche « intentionnelle » représente la substance, le fond. Au niveau du discours, une couche « structurelle » représente la forme et une couche « présentationnelle » représente la substance, le rendu.

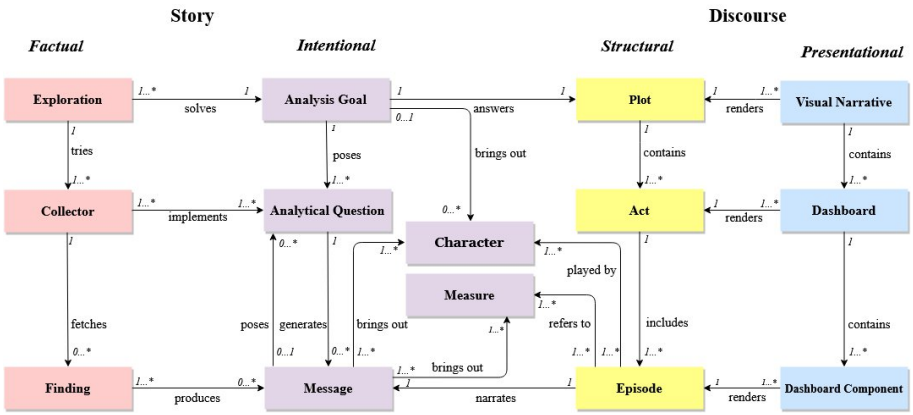


Figure 5 : *Modèle de narration de données, organisé en 4 couches (Outa et al., 2020)*

Ceci a permis de proposer dans (Outa et al. 2020) le modèle conceptuel de narration de données, exprimé comme un diagramme de classes dans le formalisme UML et illustré par la figure 5, où les quatre couches reflètent la progression, des données brutes jusqu'au rendu visuel ou discursif.

Le modèle permet de mettre en regard, comme sur la figure 6, une narration de données sous la forme d'une infographie et une partie du diagramme qui illustre sa production, de la collecte de données jusqu'à la finalisation de l'épisode.

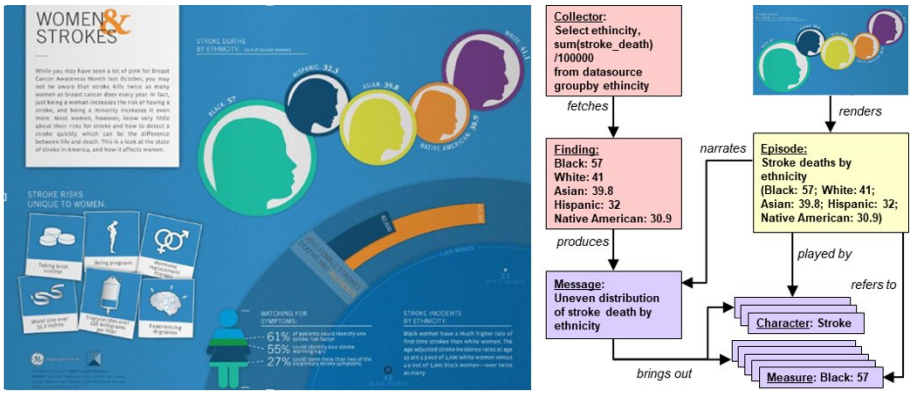


Figure 6 : *Exemple de* data narrative *(à gauche) et diagramme objet partiel pour un message particulier (à droite).*

4. De la nécessité -et de la difficulté- des interactions transdisciplinaires

Toutes les productions de la figure 2, qui représente la progression du projet, ont fait l'objet systématique d'aller-retours entre les participants, ont souvent été co-rédigées par les chercheurs et ont systématiquement été validées par tous. La formalisation du processus de mise en narration des données en utilisant un modèle conceptuel a été le vecteur d'une réelle approche transdisciplinaire sur lequel nous souhaitons insister ici.

4.1 Le modèle conceptuel au cœur des interactions

Adopter une approche transdisciplinaire impose des efforts pour appréhender de nouveaux principes méthodologiques, mais aussi, plus trivialement, pour construire des langages communs. Ainsi par exemple les termes de *dashboard* ou de *narration* ont dû faire l'objet de débats et de négociations de par les référentiels distincts qu'ils véhiculent en fonction du domaine de spécialité des participants. Nous avons également pu mesurer la persistance des barrières disciplinaires lors des communications scientifiques où nos travaux étaient souvent considérés au prisme des sujets dominants des domaines. Ainsi il nous a fallu systématiquement nous distancier de la problématique des *fake news* à laquelle nous étions assimilés par la communauté des informaticiens ou de celle de la génération automatique d'articles

par celle des chercheurs en linguistique computationnelle (TAL). Il nous a également fallu souvent redéfinir la distinction entre data journalisme et data visualisation dans la communauté des SIC.
Le fait de recourir à un modèle conceptuel pour formaliser les processus a permis de dépasser les barrières disciplinaires. Son format le rend accessible aux sciences « douces » comme aux sciences « dures » et sa réalisation intuitive a pu impliquer tous les participants et servir à établir des consensus lors des débats comme l'exprime (Guimaraes et al., 2013). A la fois outil de communication et résultat scientifique, il a été le point d'achoppement du travail collectif et, en ce sens, a pleinement rempli son rôle d'« artefact collectif » (Robinson, 1993 :190), support et instrument de la coopération. Il permet de confronter la modélisation avec la réalité du terrain en vérifiant que des étapes d'exploration des données ne sont pas omises. La puissance du modèle est qu'il permet de rendre compte du travail tel qu'il est décrit par les data journalistes et tel qu'il apparaît dans les articles du corpus.
Le modèle est toujours en cours de validation et la prochaine étape est de le soumettre aux activités de participants dans le cadre d'un hackathon. Nous souhaitons aussi tester sa pertinence dans d'autres domaines que celui du data journalisme car il nous semble que le schéma d'exploration de données et de construction de narration peut s'appliquer lors de narrations en politique ou en sciences par exemple.

4.2 Transdisciplinarité et fertilisation croisée

Par-delà les bénéfices scientifiques de cette collaboration étroite, chacun a pu mesurer ce qu'elle avait apporté dans son domaine.
L'indispensable enquête de terrain pour documenter le travail des journalistes a été l'occasion, pour les chercheurs en Sciences de l'information et de la communication, de quitter l'approche socio-professionnelle du métier pour se confronter à la dimension technique. Cela nous a ainsi permis d'affiner la perception que les journalistes avaient de la pratique du data journalisme et de faire émerger des éléments saillants, comme par exemple l'usage de la data visualisation comme outil d'enquête (Chagnoux, 2020).
Pour les journalistes, intégrer MADONA, a été l'occasion d'avoir un regard critique et réflexif sur le métier. Il nous a notamment obligé à décrire un milieu dans lequel nous évoluons trop souvent par habitude, sans remise en question. Nous avons ainsi mesuré que bien que les data journalistes soient plus nombreux qu'à la fin des années 2000, hormis de notables exceptions comme les *Décodeurs* du *Monde*[6] ou dans une moindre mesure le *Fig Data*[7], il n'y a pas eu de "grand soir" de la discipline dans le paysage national. Il y a en revanche des habitudes de travail anciennes, proches du journalisme d'investigation, qui ont été remises au goût du jour et rendues plus accessibles grâce à la récente profusion de données, mais il y a aussi de temps à autre une initiative locale qui vient donner un peu de nouveauté. Le travail de formalisation, malgré sa complexité, nous a permis de mieux appréhender nos pratiques et se révèlera fort utile lors des formations des futurs journalistes

[6] https://www.lemonde.fr/les-decodeurs/
[7] https://www.lefigaro.fr/dossier/fig-data-infographies-datavisualisation

La confrontation des différentes versions du modèle avec le terrain a permis aux informaticiens de complexifier les propositions, que ce soit par la prise en compte de faits inconnus, le recours fréquent des journalistes à des experts pour valider des étapes de leur travail par exemple, ou pour tenter de concilier la linéarité du modèle avec le foisonnement du travail d'enquête sur les données.

Conclusion

Nous avons souhaité, par cet article, revenir sur plusieurs années de collaboration entre des chercheurs et des métiers issus de disciplines différentes pour valoriser un résultat obtenu grâce à une approche résolument transdisciplinaire. Nous sommes néanmoins conscients que nos disciplines respectives sont intrinsèquement plus ouvertes à la transdisciplinarité que d'autres. De par leur histoire, leurs acteurs et leurs sujets, les Sciences de l'information et de la communication favorisent ce type d'approche et ont souvent un rôle de médiation entre des domaines éloignés. Par ailleurs, l'informatique, quand elle modélise des connaissances, automatise des processus ou se met au service des métiers, est également dans une tradition d'ouverture à des disciplines connexes ou non. Mais notre problématique comme les résultats obtenus sont tributaires de la nature intrinsèquement transdisciplinaire du projet et n'auraient pu émerger sans ce choix méthodologique.
On notera que la modélisation conceptuelle s'est révélée un efficace outil de communication et de co-construction dans un contexte transdisciplinaire. Cette notation propre au domaine informatique a facilité la validation et la distanciation par tous. Originellement pensée comme étape de la solution logicielle, elle est devenue un outil commun pour la représentation comme pour l'analyse.

Ainsi, de nos représentations négociées, de nos questionnements croisés, de la confrontation de nos cadres de référence distincts, a émergé une certaine lecture de la complexité du traitement massif des données. La validation, ou tout du moins l'évaluation, de cette compréhension passera par la pertinence des outils qui seront développés mais aussi par son adaptabilité à d'autres domaines et d'autres champs, comme ceux de la communication politique ou scientifique par exemple. Mais, par-delà ces perspectives, la réflexion épistémologique initiée lors de la rédaction de cet article se devra d'être prolongée autour du statut original des data dans une perspective de sociologie de l'innovation. Contrairement aux dispositifs traditionnels qui sont des *construits*, avec des usages pensés par la technique et en attente d'instanciation, les data seraient des *donnés*, le cadre d'usage se réalisant donc en dehors de tout cadre de fonctionnement. Il y aurait alors un terrain d'analyse fructueux à explorer et l'approche transdisciplinaire nous semble justement la plus apte à en appréhender la complexité.

Remerciements

Les auteurs tiennent à remercier tous les étudiants et les data journalistes qui ont participé aux différents travaux : Victor Alexandre, Adrien Bachy, Julien Bénéteau,

Alexandre Chanson, Ben Crulis, Arthur Diliberto, Maxime Gonzales, Ana Gressier, Lucile Jaquemart, Alexandre Léchenet, Martin Lelièvre, Audrey Margerie, Julien Rieffel Frédéric Sallet, Céline Rousseau et Lucas Ruch.

12. Bibliographie

Chagnoux M., « La datavisualisation, double point d'entrée du datajournalisme dans la PQR ». *Interfaces numériques*, vol. 9 n° 3, 2020.

Charaudeau P., « Pour une interdisciplinarité "focalisée" dans les sciences humaines et sociales », *Questions de Communication*, n°17, pp.195-222.

Chatman S., « Story and Discourse: Narrative Structure in Fiction and Film », *Cornell paperbacks*, Cornell University Press.1980.

Chen S., Li J., Andrienko G., Andrienko, N., Wang, Y., Nguyen, P.H., Turkay, C., « Supporting story synthesis: Bridging the gap between visual analytics and storytelling. » *TVCG 1-1*. 2018.

Claverie B., « Pluri-, inter-, transdisciplinarité : ou le réel décomposé en réseaux de savoir », *Projectics / Proyéctica / Projectique*, vol. 4, no. 1, 2010, pp. 5-27.

Eichmann P., Zgraggen E., Binnig C., Kraska T.,"Idebench: A benchmark for interactive data exploration ". *Proceedings of SIGMOD*, USA, 2020.¶

Flichy P., *L'innovation technique. Récents développements en sciences sociales. Vers une nouvelle théorie de l'innovation*. La Découverte, 2003

Guimarães, Maria Helena *et al.*, « Transdisciplinary Conceptual Modeling of a Social-Ecological system—A Case Study Application in Terceira Island, Azores » *Ecosystem services,* volume 3, 2013, p. e22–e31.

Idreos S., Papaemmanouil O., Chaudhuri S., « Overview of data exploration techniques. *Proceedings of SIGMOD*, Australia, 2015.

Létourneau A., « La transdisciplinarité considérée en général et en sciences de l'environnement », *VertigO - la revue électronique en sciences de l'environnement* [En ligne], Volume 8 Numéro 2 | octobre 2008.

Morin E., « Interdisciplinarité et transdisciplinarité ». *Transversales, Science, Culture*, n° 29, 1994, p. 4-8.

Nicolescu B., *La transdisciplinarité*. Paris : Éditions du Rocher, 1996

Outa F. E., Francia M., Marcel, P. Peralta V. et Vassiliadis P., « Towards a conceptual model for data narratives ». *Proceedings of ER*. 2020.

Outa F. E., Francia M., Marcel, P. Peralta V. et Vassiliadis P., « Supporting the generation of data narratives ». *Proceedings of ER Forum*, Demos and Posters. 2020.

Piaget, J., « Le système et la classification des sciences ». In J. Piaget (Ed.) *Logique et connaissance scientifique*, Paris : Gallimard, 1967. p. 1151-1224.¶

Robinson M., « Design for unanticipated use… », in Giorgio de Michelis, Carla Simone and Kjeld Schmidt, Proceedings of the Third European Conference on Computer Supported Cooperitave Work, Kluwer Academic Publishers, Dordrecht, 1993, p. 189 (traduction française in Réseaux n° 69, CNET, Paris, 1995).¶

Star S. L., Griesemer J. R., “Institutional Ecology, Translations and Boundary Objects: Amateurs and Professionals in Berkeley’s Museum of Vertebrate Zoology (1907-1939) “, *Social Studies of Sciences*, vol. 19, Sage, Londres, 1989, p. 393.