

**Jonathan ZURBACH**

**Sous la direction de Guillaume Marrel et Lise Renaud et co-encadré par Ouassim Hamzaoui**

**Laboratoire Biens Normes Contrats (LBNC) – Centre Norbert Elias (CNE)**

---

**Remarque** : Je me suis permis de rectifier votre document en précisant que M. Hamzaoui co-encadre également la thèse aux côtés de M. Marrel et de Mme Renaud, comme cela est indiqué dans les documents d'Autorisation d'inscription en doctorat ou encore dans la Convention de formation.

### **\_ Ses principaux résultats obtenus**

La thèse ayant débuté en septembre 2021, aucun résultat scientifique n'a encore été obtenu si ce n'est une clarification des notions centrales pour le traitement du sujet (p.ex. les notions de *datafication*, de formulation ou encore d'infrastructure) dans un premier document de travail et un premier compte rendu de réunion.

### **– Les apports (et limites) du travail interdisciplinaire**

La première réunion interdisciplinaire de lancement de la thèse ayant eu lieu le 24 septembre 2021, les apports et limites du travail interdisciplinaire ont pour l'instant seulement été entraperçus. Cependant, en termes d'apports, il est d'ores et déjà possible de mentionner la richesse conceptuelle provenant du cumul des patrimoines de connaissance issus de la science politique et des sciences de l'information et de la communication ; cumul permettant des hybridations dont la nature novatrice commence déjà à être appréciée. Concernant les limites du travail interdisciplinaire, le recul n'est pour l'instant pas suffisant pour les évaluer précisément. Il est cependant possible de mentionner, en guise de première analyse de l'organisation interdisciplinaire, la dominante science politique dans la configuration d'encadrement, du fait du rattachement de deux encadrants sur trois (M. Marrel et M. Hamzaoui) à cette discipline. Cette configuration pourra éventuellement influencer sur l'orientation du travail interdisciplinaire.

### **– Son intérêt par rapport à l'axe identitaire "Culture, Patrimoines, Sociétés numériques"**

L'analyse des pratiques relatives aux données en sciences humaines et sociales (SHS) s'inscrit dans trois contextes aux interfaces des notions de culture, de patrimoine et de sociétés numériques. Le premier contexte est celui du mouvement de la science ouverte dans le cadre duquel est formulée une invitation de formation de plus en plus pressante à l'égard des chercheurs pour les pousser à acquérir une « culture de la donnée » et à diffuser cette culture au sein des disciplines comme des projets de recherche qu'ils entreprennent. Le second contexte est celui des très grandes infrastructures (p.ex. Huma-Num ou Progedo) qui ont pour mission de formaliser, pour pouvoir les gérer dans la durée, les données des SHS. Dans ce second contexte, les données des SHS constituent à la fois un patrimoine scientifique pour les communautés scientifiques mais aussi, en contexte d'ouverture, un « actif stratégique » pour les entreprises ; patrimoine ou capital dont la gestion est déléguée aux ingénieurs de recherche. Enfin, le troisième contexte dans lequel s'inscrit le travail des données en SHS est celui de la *datafication* ou « mise en données du monde » propre à la « civilisation numérique ». Dans ce dernier contexte, un double mouvement a lieu qui structure et traverse le travail des données en SHS : d'une part l'accumulation des données sur les faits sociaux est considérée comme une formidable promesse de meilleure compréhension de ces derniers et, d'autre part, cette même accumulation est considérée par la société civile comme une formidable menace de surveillance exigeant une régulation stricte des traitements scientifiques opérés sur des données à caractère personnel, au moyen de l'appareil juridique du RGPD notamment.

## \_ Situation actuelle et la mission de votre travail

En préparation de la réunion de lancement de thèse du 24 septembre 2021, un premier document de travail avait été proposé. Ce document a été revu par l'équipe encadrante et a permis de cerner les éléments bibliographiques centraux pour le sujet ainsi que les concepts et méthodes permettant d'analyser ce dernier. Trois missions préliminaires ont été identifiées à l'issue de cette réunion :

1. Une analyse des formulations relatives à l'organisation du travail des données en SHS et gravitant autour de la notion de « culture de la donnée » dans les rapports institutionnels. Cette analyse adoptera une conception et une méthodologie propres au champ de l'analyse du discours. Cette conception et cette méthodologie ont été retenues car elles ont une forte coloration propre à la science politique tout en ayant une solide assise issue des sciences de l'information et de la communication.
2. Une analyse de l'émergence et du développement actuel des infrastructures de recherche fournissant le socle technologique pour la constitution des données des SHS à la fois en patrimoine scientifique pour les communautés scientifiques et en actif stratégique pour les sociétés numériques. Cette analyse permettra également de cerner les nouvelles coopérations professionnelles exigées pour la gestion des données des SHS d'une part avec les ingénieurs de recherche mais aussi, d'autre part, avec de nouveaux métiers tel que celui de *data steward* par exemple.
3. La participation à une analyse sociologique des Délégués à la Protection des Données (DPO) au service des chercheurs visant à cerner leurs profils, leurs compétences et la place qui leur est dévolue au sein des chaînes de production de données de recherche. Cette analyse, portée par M. Marrel et M. Hamzaoui, est déjà en cours et dépasse le cadre de la thèse. Dans le cadre de la thèse, cette analyse permettra de caractériser les stratégies variables de documentation des projets que proposent les DPO (notamment leur définition variable du possible et de l'impossible), qualités de stratégie juridique dont dépendent les chercheurs pour la mise en conformité de leurs traitements en matière de données et donc, en définitive, pour l'avancement de leurs recherches.

## \_ Bibliographie

- **Anderson C., 2008.** "The End of Theory: The Data Deluge Makes the Scientific Method Obsolete", Wired Magazine, <https://www.wired.com/2008/06/pb-theory/>.
- **Bonaccorsi J., 2016** "Le monde de l'Open data : les jeux sémiotiques et esthétiques de la "visualisation" comme rhétorique de la transparence". Sémir Badir; François Provenzano. *Pratiques émergentes et pensée du médium*, L'Harmattan.
- **Boullier D., 2015.** « Les sciences sociales face aux traces du big data. Société, opinion ou vibrations ? » *Revue française de science politique* 65 (5-6): 805-28.
- **Boyadjian J., 2017.** « Les conditions de scientificité des Big Data en science politique ». *Revue française de science politique* 67 (5): 919-29.
- **Broudoux É., Chartron G., (dir.) 2015.** *Open data, big data : quelles valeurs, quels enjeux ?*, Bruxelles, De Boeck.
- **Bullich V., Clavier V., (dir) 2018.** « Production des données, "production de la société". Les Big Data et algorithmes au regard des sciences de l'information et la communication », *Les Enjeux de l'information et la communication* 19 (2).
- **Chartron G., Schöpfung J., (dir.) 2017.** « Open access et Open science en débat », *Revue française des sciences de l'information et de la communication* 11, <http://journals.openedition.org/rfsic/3331>.
- **Hamzaoui O., Marrel G., 2021, à paraître.** « La datafication de (et dans) la recherche en SHS. Retour sur la construction d'un problème public », in Theviot Anaïs (dir.), *Gouverner par les données*, ENS Editions.
- **Ibekwe-Sanjuan F., 2014.** « Big Data, Big machines, Big Science : vers une société sans sujet et sans causalité ? », XIXème Congrès de la Sfsic. *Penser les techniques et les technologies : Apports des Sciences de l'Information et de la Communication et perspectives de recherches*, Toulon, France: 1-10, hal-01066202.
- **Labelle S., 2020.** *Médiations & pouvoirs de l'agir ingénieur. Mythes et infrastructures des politiques de données.* Mémoire d'HDR en Sciences de l'information et de la communication.
- **Lascoumes P., Le Galès P., 2007.** *Sociologie de l'action publique : Domaines et approches*, Paris, Armand Colin.

- **Latour B., 2011.** *Pasteur : guerre et paix des microbes. Suivi de Irréductions*, Paris, La Découverte.
- **Latour B., Woolgar S., 1988.** *La Vie de laboratoire. La production des faits scientifiques*, Paris, La Découverte.
- **Marres N., 2012.** « The redistribution of methods: on intervention in digital social research, broadly conceived », *The Sociological Review* 60: 139-165.
- **Mounier P., 2010.** « Manifeste des Digital Humanities », *Journal des anthropologues* 122-123 : 447-452.
- **Neveu É., 2015.** *Sociologie politique des problèmes publics*, Paris, Armand Colin.
- **Ollion E., Boelaert J., 2015.** « Au-delà des Big Data. Les sciences sociales et la multiplication des données numériques », *Sociologie* 6(3): 295-310.
- **Owen-Smith J., 2001.** « Managing laboratory work through skepticism : processes of evaluation and control ». *American Sociological Review* 66 (2): 427-452.
- **Paquiénéguy F., (dir.) 2016.** *Open data*, Paris, Archives contemporaines.
- **Plantin J.-R., Mabi C., Monnoyer-Smith L., (dir.) 2017.** *Ouvrir, partager, réutiliser : Regards critiques sur les données numériques*, Paris, Éditions de la Maison des sciences de l'homme, <http://books.openedition.org/editionsmsh/9026>.
- **Plantin J.-R., Russo F., 2016.** « D'abord les données, ensuite la méthode ?. Big data et déterminisme en sciences sociales ». *Socio. La nouvelle revue des sciences sociales* 6 (mai): 97-115.