

DOSSIER 2026 A

REVUE SCIENTIFIQUE EN SCIENCES DE L'INFORMATION ET DE LA COMMUNICATION

Numéro 1/2026 (26/1) - Dossier 2026 A

Diplomatie des données

Ce supplément a été coordonné par Anne **Lehmans**

Julien **Hocine**, Julia **Bihl**, Normand **Landry**, Simon **Fitzbay**, Marie-Josée **Dupuis**, Saliha **Hadna**, Luis Roman **Arciniega Gil**, Manuel **Munier**, Chérifa **Boukacem Zeghmouri**, Candice **Fillaud**, Raphaëlle **Bats**, Mathilde **Garnier**, Sylvie **Marciset-Sognos**, Alphée Yannick Hilaire **Ngalli Ngomo**, Hubert **Boët**

Mise en page : Sonia BAZAOUI & Cédric COUSTELLIÉ
Design graphique : Laurence PAYAN
Tous droits réservés
Publication numérique : 4e trimestre 2026
© 2026 Les Enjeux de l'information et de la communication

Table des matières

Anne Lehmans	5
▶ Introduction	
Julien Hocine, Julia Bihl, Normand Landry, Simon Fitzbay et Marie-Josée Dupuis	11
▶ Une « diplomatie des données par le bas » : la défense des droits humains en contexte d'autoritarisme numérique	
Saliha Hadna	32
▶ De l'ignorance technique aux données citoyennes. Le cas des cancers pédiatriques en Pays de Retz	
Luis Roman Arciniega Gil et Manuel Munier	46
▶ Quelle gouvernance pour les données de recherche ? Diplomatie et enjeux géopolitiques	
Chérifa Boukacem Zeghmouri et Candice Fillaud	64
▶ Partage et réutilisation des données de la recherche à l'ère de la Science Ouverte. Une approche diplomatique pour penser et panser l'ouverture des données	
Raphaëlle Bats et Mathilde Garnier	79
▶ S'ouvrir à la négociation des métadonnées : ateliers expérimentaux de cartographie des projets de recherche	
Sylvie Marciset-Sognos	96
▶ Usage et ré-usage des données de la recherche : les intersections du compromis au prisme de la médiation des connaissances	
Alphée Yannick Hilaire Ngalli Ngomo	115
▶ Gouvernance et circulation des données dans les archives ouvertes en Afrique : le cas de DICAMES	
Hubert Boët	134
▶ Le « <i>Data Space Education Skills</i> » (DASES) ou la construction d'une idéologie européenne des données éducatives au regard de la domination des Big Tech sur le numérique éducatif	

Anne Lehmans

Anne Lehmans est professeure des universités en sciences de l'information et de la communication à l'université de Bordeaux et au laboratoire IMS (UMR 5218). Elle consacre ses travaux aux cultures et pratiques d'information, en particulier dans une approche politique. Elle s'intéresse aux questions d'éducation, de formation et de communication de la science.

anne.lehmans@u-bordeaux.fr

INTRODUCTION

Les données (*data*) prennent une place considérable au sein des pratiques d'information et de communication. La maîtrise des données apparaît de plus en plus comme une arme au service de l'influence et du contrôle des comportements, des discours et des pensées. Au-delà des questions d'intelligence économique, les élections sont un bon indicateur de l'importance des stratégies politiques de maîtrise de l'information en circulation à l'aide de données. Les guerres contemporaines se mènent sur les champs de bataille mais aussi sur les systèmes d'information et de circulation des données. Les sciences s'appuient sur des plateformes d'échanges de données qui imposent de nouveaux modèles de coopération, de compétition et de guerre.

Le terme de « donnée » semble pourtant tout à fait neutre. Il trouve son origine étymologique dans le verbe latin *dare* (donner) dont le substantif pluriel *data*, est utilisé en anglais. Si l'on a tendance à considérer naturellement que les données sont le reflet de la réalité, Howard Becker a montré, dès 1952, que les *data* sont en réalité des *capta*, notamment dans les sciences, le résultat de choix et d'intentions dans ce que l'on décide de prélever du réel pour construire des informations puis des connaissances ou des dispositifs de contrôle. Sylvie Leleu-Merviel (2016) ajoute à cette définition des *capta/data* les *ligata*, superstructure agrégative qui permet de passer du réel à l'information en considérant que les données ont une dimension sémiotique, matérielle et sociale. À l'inverse, au sens informatique attaché au dispositif technique, les données désignent « *la représentation d'une information sous une forme conventionnelle destinée à faciliter son traitement* » notamment informatique¹. Dans quelle mesure la diplomatie constitue-t-elle alors une catégorie d'analyse au sujet des données pour les sciences de l'information et de la communication ? Face aux mutations des pratiques d'information, cette question, au cœur des stratégies de représentation et de communication, constitue la ligne directrice de ce dossier.

Dans le monde contemporain dominé par les techniques numériques, le traitement par des machines, des algorithmes et des modèles de langage donne la possibilité de mobiliser des quantités massives de données. Celles-ci sont produites et traitées par l'intermédiaire de professionnels qui ont pour mission de les collecter, de les traiter et de les gérer, dans ce qu'il est convenu de désigner par la gouvernance des données. Celle-ci relève, selon Anke Sophia Obendiek (2023), d'une question de droits individuels, par exemple à travers les législations protégeant les données personnelles, comme le Règlement général sur la protection des données en Europe et avant cela, en France, avec la Loi Informatique et libertés de 1978. Elle concerne aussi la sécurité dans les relations entre États (et l'appel à la souveraineté des données pour les données sensibles comme les données de santé mais aussi la sécurité nationale), l'économie, parce que les données représentent des ressources centrales à travers leur exploitation par les plateformes (Isaac, 2026), enfin les valeurs collectives. Les données peuvent être considérées comme un bien commun à protéger des

hégémonies et des appropriations, particulièrement dans le domaine des sciences.

DIMENSION POLITIQUE ET STRATEGIQUE DES DONNEES

Les données ont une importance politique et stratégique fondamentale, car elles constituent la base des moyens de communication, des transactions et de potentielles manipulations. Ces dimensions politiques sont souvent masquées sous des discours portant sur l'éthique du numérique (exigence de transparence, neutralité, participation, etc.) qui permettent d'imposer des normes et révèlent des contradictions. Même dans le domaine de l'ouverture des données de la science, la question de l'industrialisation de la production scientifique vient heurter le principe de communs de la connaissance (Boukacem-Zeghmouri *et al.*, 2021). Les données sont aussi stockées et traitées suivant un cycle de vie qui, compte tenu des avancées les plus récentes de l'intelligence artificielle (IA), pose de nombreux problèmes. Celui de l'impact environnemental, notamment en matière de consommation d'énergie et d'eau, n'est pas des moindres. Ce cycle est aussi fragile du point de vue de la protection des données et des usages possibles à des fins frauduleuses, particulièrement dans certains États où des entreprises se sont spécialisées dans le piratage, dans l'exploitation et dans la manipulation des données. Les IA génératives ajoutent à ces risques la question de l'identification et de la fiabilité des données d'entraînement mobilisées par les grands modèles de langage, ainsi que celle du respect des droits des créateurs de données issues des œuvres et de l'information médiatique.

Dans le contexte de la globalisation des économies, des techniques et des systèmes politiques, la production et la gestion de données ne peuvent plus être pensées dans le cadre limité des États-nations ou des organisations. Les données circulent et leur sort ne peut être administré que dans un écosystème qui va du niveau local au niveau global. Ainsi, elles dépassent largement les frontières nationales et les conflits régionaux, et contraignent les organisations à la compétition, voire à la guerre des données (hacking, espionnage), à des négociations (Data Governance Act, AI Act, RGPD), éventuellement à des processus coopératifs (science ouverte, *data hubs*). Elles sont ressources et moyens de pouvoir. Quand une entreprise légale ou pirate capte et revend des données (dans le domaine médical avec les données de santé, dans le domaine politique avec les données électorales et les tentatives de manipulation des opinions à travers les réseaux sociaux, ou dans le domaine militaire), quand un État bloque l'accès à un réseau social sous le prétexte de protéger les données personnelles de ses citoyens, quand un mouvement citoyen « hacktiviste » lutte pour préserver l'accès aux données, ou quand des entreprises exploitent des travailleurs de la donnée dans des pays qui n'ont pas accès aux ressources produites par l'industrie informatique (Waquet, 2022), les données sont au cœur d'une nouvelle diplomatie. Celle-ci traite de ce qui constitue le matériau brut de l'information en circulation dans le monde, particulièrement à l'heure de l'explosion des usages de l'IA.

Cette diplomatie se met en place dans les instances « traditionnelles » des relations internationales, mais aussi dans des structures plus récentes qui tentent de mobiliser les organisations et de faire pression sur les politiques publiques, ou encore dans des instances de résistance à travers l'activisme des données. Le vocabulaire utilisé au sujet des données reflète d'ailleurs clairement ce rapport à la diplomatie : gouvernance et souveraineté des données, diplomatie numérique (de l'Union européenne par exemple), colonialisme des données, à propos des effets de domination associés aux rapports de production et traitement des données (Couldry, Meijas, 2023 ; Mboa Nkoudou, 2023). La diplomatie relève de la recherche d'instruments de lutte pour, ou de résistance contre le pouvoir de « datafication » et d'« archivage » des individus et des groupes. Ce pouvoir s'active à partir des traces et produit des effets de classement selon des logiques industrialisées et normalisantes. C'est le cas, par exemple, dans les algorithmes qui imposent des classements ethnicisés et invisibilisent certains groupes. Les revendications d'un droit à l'autodétermination sur les données sont centrales pour les savoirs endogènes.

L'évolution récente d'une explosion des usages de l'IA générative nécessite, en outre, des moyens industriels considérables pour son développement. Elle accroît la situation d'asymétrie en ajoutant l'exploitation du travail pour le numérique dans les pays les plus pauvres (Mosco, 2016, Casilli, 2021) et une représentation stéréotypée de l'information et des cultures (Héron, 2024). Cette situation produit de nouveaux rapports de force entre organisations et des questionnements éthiques, concernant notamment les droits humains, les données personnelles, la souveraineté des données et le respect de la création, mais aussi le contrôle social (Desveaud, 2024). L'exemple du chantier de Notre Dame de Paris illustre précisément l'importance des données dans le domaine culturel pour préserver un patrimoine commun. Le terme de « diplomatie des données » est utilisé par Livio de Luca, cité par Jacquot (2024), pour désigner le processus de gouvernance et de traitement des données produites par « *des tiers au profit des actions du Chantier scientifique. Elles impliquent la communication, l'étude, la recherche et la valorisation ainsi que la vulgarisation scientifique et excluent toute exploitation commerciale* ».

Les enjeux stratégiques et les règles du jeu de la diplomatie des données dépassent ceux de la diplomatie traditionnelle dans les relations internationales. Le constat selon lequel les dispositifs numériques ont changé les relations diplomatiques, dans l'expression « diplomatie numérique », met l'accent sur les stratégies d'influence. Celles-ci sont liées aux plateformes notamment et impliquent « l'adaptation des pratiques diplomatiques aux avancées technologiques contemporaines » (Lyubareva, Nocetti, 2024). Ce constat est déjà ancien. Mais le questionnement sur la diplomatie des données va au-delà. Il porte sur l'émergence de règles spécifiques au traitement politique et stratégique de la circulation des données. Aux *big data* est associé un changement majeur dans les relations internationales, car la circulation des données n'est pas limitée par les frontières et par l'action des États. Si la diplomatie peut être définie comme « l'art de gérer les séparations » (Badie, 2016), ou plus classiquement comme « la transmission de messages entre une communauté politique indépendante et une autre » (Bull, 1977, p.164), les diplomates apparaissent comme des « spécialistes d'une communication précise et exacte » (*Ibid.*, p.179). Dans cette perspective de diplomatie centrée sur la communication, Andy Boyd *et al.* ont proposé en 2019 de définir la diplomatie des données comme « *la mobilisation des actions et des compétences diplomatiques par un large éventail de parties prenantes pour négocier et favoriser l'accès aux données, ainsi que leur utilisation et leur compréhension généralisées* ». Dans le triptyque « *data in diplomacy* », « *diplomacy for data* » and « *data for diplomacy* », c'est le second point de la diplomatie pour les données qui importe ici.

ENJEUX DE POUVOIR

Ce dossier propose de faire de la diplomatie des données l'objet d'une réflexion sur les enjeux de pouvoir qui se dessinent autour du contrôle de la production, de la circulation et de l'exploitation des données. Trois orientations se dégagent des articles qui le constituent. L'orientation stratégique analyse les données comme des ressources de pouvoir. L'orientation de la gouvernance, particulièrement centrée sur les données de la recherche et les conséquences de leur gestion, ainsi que de leur médiation, souligne l'importance des actions de communication autour des données. Enfin, l'orientation culturelle, liée au constat des inégalités de ressources, met l'accent sur le fait que la diplomatie concerne les positions d'influence, les effets d'inertie et les perspectives de changement.

Pour la première orientation, les données apparaissent comme des enjeux de pouvoir en tant que ressources stratégiques et comme des moyens d'*empowerment* au service de la négociation dans l'arène internationale et nationale. Ces ressources, qui nécessitent une forme d'expertise, sont identifiées par Julien Hocine, Julia Bihl, Normand Landry, Simon Fitzbay et Marie-Josée Dupuis. Les auteurs ont cherché, dans une revue de littérature sur la défense des droits de l'homme, les conditions de possibilité d'une « diplomatie des

données par le bas ». **Saliha Hadna** illustre cet usage possible des données dans une perspective diplomatique qui s'appuie sur l'information issue des données épidémiologiques pour défendre les droits des victimes de cancers pédiatrique dans le pays de Retz. Car les enjeux diplomatiques des usages des données existent aussi à l'intérieur des frontières nationales.

Concernant la deuxième orientation, la gouvernance est interrogée dans le cadre des données de la recherche. La montée en puissance de la science ouverte dans un contexte de *big data* a fait émerger l'idée, introduite par Jim Grey, reprise par Tony Hey *et al.* (2009), d'un « quatrième paradigme » (Verlaet et Bachimont, 2024), après les approches empiriques, théoriques, et par simulation numérique. Ce quatrième paradigme caractériserait une épistémologie fondée sur l'exploration et l'analyse de données massives à l'aide d'algorithmes comme processus de découverte scientifique. La place des données dans la recherche scientifique a ainsi pris une ampleur considérable, dans un contexte qui appelle des arbitrages, des négociations et des conflits autour de l'accès aux bases de données. « L'écosystème » de la science ouverte est changeant, du point de vue des outils à mobiliser, du traitement scientifique et documentaire, de la normalisation pour faciliter l'interopérabilité et l'interconnexion des systèmes, la mise à disposition et le stockage des données de la recherche. Ces questions ne sont pas seulement techniques, elles posent des défis éthiques et politiques. Les données peuvent être considérées comme faisant partie des communs de la connaissance et de la culture à protéger, nécessitant la mise en place de procédures de gouvernance (Brousseau, 2022) et de règles diplomatiques qui permettent d'assurer un partage acceptable des données. **Luis Roman Arciniega Gil** et **Manuel Munier** proposent ainsi une réflexion sur les caractéristiques des « communs des données » inclusifs et durables, à partir d'un modèle de gouvernance qui entend compléter la diplomatie de l'Union européenne centrée sur la transparence et la protection des droits. Cependant, l'ouverture des données associée à la quantification dans l'évaluation de la science (Bordignon, 2021), entraîne des effets collatéraux. L'émergence de nouveaux modèles d'affaires dans la publication scientifique ainsi que la difficulté à mettre en place une véritable gestion ouverte des données sont soulignées dans l'article proposé par **Chérifa Boukacem Zeghmouri** et **Candice Fillaud** sur la réutilisation des données de la science. Cette possibilité de réutilisation crée un marché symbolique et économique, avec des dynamiques mais aussi des tensions nouvelles. **Sylvie Marciset-Sognos** confirme l'importance de la gouvernance et surtout de la médiation des données de la recherche, dans une réflexion théorique. Cette perspective théorique est complétée par l'illustration proposée par **Raphaëlle Bats** et **Mathilde Garnier** sur l'usage de la cartographie comme outil de médiation des données de la recherche auprès de publics non experts.

Concernant la troisième orientation, ce sont les enjeux culturels et les effets de domination dans la diplomatie des données qui sont analysés. En 2018, la revue *Les Enjeux de l'information et de la communication* consacrait un dossier à la production du fait social par les données, dans un processus de « datafication de la société ». La valorisation des données par les industries culturelles pose les questions de l'accès aux données, mais surtout de la capacité à les valoriser dans des perspectives d'innovation, de création et d'accès à la culture pour répondre aux défis environnementaux, sanitaires, sociaux et culturels. La capacité des organisations étatiques ou non est très inégalitaire, du fait des disproportions de moyens financiers. Cette inégalité peut être considérée comme créatrice d'injustice cognitive (Piron, 2018), voire de formes de technocolonialité (Mboa Nkoudou, 2021). Dans la redéfinition de l'ordre mondial de la connaissance scientifique, les inégalités régionales surgissent avec force. **Alphée Yannick Hilaire Ngalli Ngomo**, au sujet des bases de données bibliographiques dans les pays d'Afrique francophone, met en évidence les conséquences d'une gouvernance incomplète des données de la recherche sur la visibilité de la production scientifique et sur sa souveraineté. En Europe, les données représentent également un enjeu politique et économique majeur dans le domaine de l'éducation. **Hubert Boët** analyse l'industrialisation, la platformisation et la datafication de l'éducation pour souligner les

difficultés de la diplomatie européenne face à la domination des grandes entreprises du numérique.

Ce numéro sur la diplomatie des données ouvre un chantier dont de nombreux domaines restent à explorer. La réflexion des sciences de l'information et de la communication y est centrale, à l'interface de problématiques politiques, sociales, culturelles, économiques et techniques dans la circulation des connaissances.

NOTES

¹ Selon l'arrêté du 22 décembre 1981 sur l'enrichissement de la langue française, consultable dans le Bulletin des bibliothèques de France (BBF) : <https://bbf.enssib.fr/consulter/bbf-1982-06-0355-009>

RÉFÉRENCES BIBLIOGRAPHIQUES

Badie, Bertrand (2016), *Nous ne sommes plus seuls au monde. Un autre regard sur l'« ordre international »*, Paris : La Découverte.

Becker, Howard (1952), « Science, culture, and society », *Philosophy of Science*, vol. 19, n° 4, p. 273-287.

Bordignon, Frédérique (2021), « Comment ouvrir et mesurer la science vont de pair ? », *Arabesques*, n° 103, p. 4-5.

Boukacem-Zeghmouri, Chérifa ; Sarah Rakotoary ; Bador Pascal (2021), « La prédation dans le champ de la publication scientifique : un objet de recherche révélateur des mutations de la communication scientifique ouverte », *Natures Sciences Sociétés*, vol. 29, n° 4, p. 382-395.

Boyd, Andy ; Gatewood, Jane ; Thorson, Stuart ; de Ver Dye, Timothy (2019), « Data diplomacy », *Science & diplomacy*, vol. 8, n° 1.

Brousseau, Éric (2022), « Organiser la valorisation de l'or noir du XXI^e siècle », *Enjeux numériques*, vol. 18, n° 2, p. 15-23.

Bull Hedley (1977), *The Anarchical Society, a Study of Order in World Politics*, New York : Macmillan Publishers.

Bullich, Vincent ; Clavier, Viviane (2018), « Production des données, « Production de la société » Les Big Data et algorithmes au regard des Sciences de l'information et de la communication », *Les Enjeux de l'information et de la communication*, vol. 192, n° 2, p. 5-14.

Casilli, Antonio (2021), « Il n'y a pas d'intelligence artificielle, il n'y a que le travail du clic de quelqu'un d'autre », *Ubérisation, et après ?*, p. 33-55.

Couldry, Nick ; Mejias, Ulyses (2023), « The decolonial turn in data and technology research: what is at stake and where is it heading? », *Information, Communication & Society*, vol. 26, n° 4, p. 786-802.

Desveaud, Kathleen (2024), « Des remises en question sociétales importantes. L'intelligence artificielle décryptée », In *Comprendre les enjeux et risques éthiques de l'IA pour mieux l'appréhender*. (p. 161 -249), Paris : EMS Éditions.

Héron, Fabrice (2024), « L'intelligence artificielle générative, menace ou opportunité pour les documentalistes iconographes ? », *I2D-Information, données & documents*, vol. 242, n° 2, p. 72-78.

Hey, Antony ; Tansley, Stuart ; Tolle, Kirstin (2009), *The Fourth Paradigm: Data-Intensive*

Scientific Discovery, Redmond (USA): Microsoft Research.

Isaac, Henri (2026), *E-commerce. Nouveaux enjeux : Intelligence artificielle, durabilité, globalisation*, Paris : Dunod.

Jacquot, Kévin (2024), « Les images de la « Forêt » de Notre-Dame de Paris », *Études de communication*, n° 2, p. 113-131.

Leleu-Merviel, Sylvie (2016), « La signifiante canalisée par l'horizon de pertinence, des saisies aux agrégats via les données », *Revue des Interactions Humaines Médiatisées*, vol. 17, n° 1, p. 107-139.

Lyubareva, Inna ; Nocetti, Julien (2024), « La diplomatie numérique Évolution des stratégies diplomatiques et d'influence à l'ère (du) numérique. », *Réseaux*, vol. 245, n° 3, p. 11-35.

Mboa Nkoudou, Thomas (2021), « Stratégies de résistance à la technocolonialité : Apprendre des fablabs d'Afrique. » *Revue Possibles*, vol. 45, n° 1, p. 119-129.

Mboa Nkoudou, Thomas (2023), « We need a decolonized appropriation of AI in Africa », *Nature Human Behaviour*, vol. 7, n° 11, p. 1810-1811.

Mosco, Vincent (2016), « Après l'Internet : le Cloud, les big data et l'Internet des objets », *Les Enjeux de l'information et de la communication*, vol. 17, n° 2, p. 253-264.

Obendiek, Anke Sophia (2023), *Data governance: Value orders and jurisdictional conflicts*. Oxford University Press.

Piron, Florence (2018), « Justice et injustice cognitives : de l'épistémologie à la matérialité des savoirs humains », In *La gravité des choses*. <https://scienceetbiencommun.pressbooks.pub/gravite/chapter/justice-et-injustice-cognitives/>

Verlaet, Lise ; Bachimont, Bruno (2024), « Le 4ème paradigme en question. Appel à contribution », *Intelligibilité du numérique*, <https://www.sfsic.org/aac-publication/le-4eme-paradigme-en-question/>

Waquet, Françoise (2022), *Dans les coulisses de la science-Techniciens, petites mains et autres travailleurs invisibles*, Paris : CNRS éditions.

Une « diplomatie des données par le bas » : la défense des droits humains en contexte d'autoritarisme numérique

Article inédit, mis en ligne le 23 septembre 2026

Julien Hocine

Julien Hocine est doctorant en communication à l'Université du Québec à Montréal, assistant de recherche à la Chaire de recherche du Canada en éducation aux médias et droits humains et chargé de cours. Ses travaux de recherche portent sur les usages des technologies, les imaginaires sociotechniques, et les représentations du Nord et de l'Arctique.

Hocine.julien@courrier.uqam.ca

Julia Bihl

Julia Bihl est docteure en sciences de l'information et de la communication, collaboratrice du GReMS et chargée de cours à l'Université catholique de Louvain (Belgique). Ses travaux de recherche portent sur l'éducation aux médias, la littératie médiatique, la pensée critique et l'émancipation sociale.

Julia.bihl@uclouvain.be

Normand Landry

Normand Landry est professeur titulaire à l'Université TÉLUQ. Il est également titulaire de la Chaire de recherche du Canada en éducation aux médias et droits humains. Ses travaux se concentrent sur l'éducation aux médias, les droits de la communication, l'intimidation judiciaire, les représentations de la pauvreté, ainsi que sur la communication et les mouvements sociaux.

Normand.landry@teluq.ca

Simon Fitzbay

Simon Fitzbay est étudiant au doctorat à l'Université du Québec à Montréal et assistant de recherche à la Chaire de recherche du Canada en éducation aux médias et droits humains. Ses travaux de recherche portent sur le rôle de la technique dans le domaine de l'éducation dans une perspective d'économie politique.

fitzbay.simon@courrier.uqam.ca

Marie-Josée Dupuis

Marie-Josée Dupuis est professionnelle de recherche à la Chaire de recherche du Canada en éducation aux médias et droits humains. Ses travaux de recherche portent sur la pauvreté et les inégalités, la précarisation du marché du travail et les mouvements sociaux.

marie-josée.dupuis@teluq.ca

Plan de l'article

Introduction

Penser la gouvernance des données depuis les marges

Modalités d'analyse de la revue de littérature

Les catégories de répression numérique : surveillance, censure et propagande

Les catégories de contraintes sociotechniques

Des tactiques de résistance face à la mise en armement des données numériques

Discussion : Vers une diplomatie des données par le bas

Conclusion

Références bibliographiques

RÉSUMÉ

Dans un contexte marqué par la montée d'un autoritarisme numérique qui instrumentalise les données à des fins de surveillance, de censure et de propagande, cet article analyse les pratiques numériques des défenseur·es des droits humains. Fondé sur une revue systématique de 111 articles savants, il fait la démonstration que ces pratiques constituent une production normative implicite qui conteste les modèles étatiques de gouvernance des données. De fait, elles sont conceptualisées comme une « diplomatie des données par le bas » qui rompt avec une vision de la diplomatie comme une prérogative exclusivement étatique ou technocratique et vise à établir de nouvelles normes pour une circulation plus juste et sécurisée des données sensibles.

Mots clés

Droits humains, activisme numérique, répression numérique, mise en données, gouvernance des données, diplomatie des données.

Title

A Bottom-Up Data Diplomacy : Defending Human Rights in the Context of Digital Authoritarianism.

Abstract

In a context marked by the rise of digital authoritarianism that instrumentalizes data for surveillance, censorship, and propaganda, this article analyzes the digital practices of human rights defenders. Based on a systematic review of 111 scholarly articles, it demonstrates that these practices constitute an implicit normative production that challenges state-centric models of data governance. Indeed, they are conceptualized as a "bottom-up data diplomacy" that breaks with a vision of diplomacy as an exclusively state-based or technocratic prerogative and aims to establish new standards for a fairer and more secure circulation of sensitive data.

Keywords

Human Rights, Digital Activism, Digital Repression, Datafication, Data Governance, Data Diplomacy.

TITULO

Una diplomacia de datos desde abajo: La defensa de los derechos humanos en el contexto del autoritarismo digital.

Resumen

En un contexto marcado por el auge de un autoritarismo digital que instrumentaliza los datos para la vigilancia, la censura y la propaganda, este artículo analiza las prácticas digitales de los defensores de los derechos humanos. Basado en una revisión sistemática de 111 artículos académicos, demuestra que estas prácticas constituyen una producción normativa implícita que desafía los modelos estatales de gobernanza de datos. De hecho, se conceptualizan como una "diplomacia de datos desde abajo" que rompe con una visión de la diplomacia como una prerrogativa exclusivamente estatal o tecnocrática y busca establecer nuevas normas para una circulación más justa y segura de los datos sensibles.

Palabras clave

Derechos Humanos, activismo digital, represión digital, dataficación, gobernanza de datos, diplomacia de datos.

INTRODUCTION

Si elles ont un temps été considérées comme des vecteurs d'émancipation citoyenne, les techniques numériques se sont progressivement imposées comme des instruments centraux de contrôle politique et de répression. Les États exploitent désormais largement le potentiel des infrastructures numériques pour anticiper, surveiller et neutraliser les mobilisations collectives (Dragu et Lupu, 2021), renversant la promesse initiale d'un Internet libérateur (Diamond, 2010 ; Morozov, 2011). Cette dynamique, que *Freedom House* qualifie d'« autoritarisme numérique¹ » (Shahbaz, 2018), s'incarne dans des processus d'instrumentalisation des données à des fins de surveillance, de censure et de propagande. Parallèlement, les principales plateformes numériques privées structurent toujours plus fortement l'accès à l'information et la visibilité des discours publics, et accentuent les asymétries de pouvoir entre producteurs de données, diffuseurs et audiences (Zuboff, 2019; Gregory, 2023).

Les défenseur·es des droits humains se trouvent fréquemment positionnés en première ligne du fait de leurs activités, souvent menées dans des environnements hautement répressifs, qui visent à documenter les violations des droits fondamentaux, en imputer la responsabilité, et protéger les populations ciblées (Assemblée générale des Nations Unies, 1999 ; Land et Aronson, 2020 ; Nah et al., 2013). Face à la mise en données généralisée de leurs actions, les défenseur·es des droits humains développent des stratégies d'adaptation combinant expertise technique, prudence tactique et créativité politique (Landry *et al.*, 2025). Ces pratiques produisent, dans les faits, une contestation implicite des modèles dominants de gouvernance des données, fondés sur la centralisation, l'opacité et l'exclusion des voix dissidentes.

Cet article fait la démonstration que les pratiques numériques des défenseur·es des droits humains esquissent, en creux, une manière alternative d'envisager la régulation des données sensibles, à rebours des logiques de surveillance et de contrôle. En s'appuyant sur une revue de littérature systématique de 111 articles scientifiques en langue anglaise, nous avons dressé un état des lieux des pratiques numériques des défenseur·es des droits humains, des formes actuelles de répression numérique, et des contraintes sociotechniques² qui en découlent.

Nous mobilisons la notion de « diplomatie des données » comme cadre interprétatif pour ces pratiques, celles-ci étant comprises comme des contributions politiques à la reconfiguration des normes encadrant la régulation des données. La notion de « diplomatie des données », telle que proposée par Andrew Boyd *et al.* (2019), désigne l'intervention d'une diversité d'acteurs, au-delà des seuls représentants étatiques, dans les négociations visant à promouvoir l'accès et l'usage des données. Tout en nous inscrivant dans cette perspective qui reconnaît le rôle des acteurs non-étatiques, notre démarche la prolonge en y intégrant une dimension critique et une analyse des rapports de pouvoir. Nous concevons la diplomatie des données non seulement comme un espace de coopération, mais également comme un champ de tensions où les défenseur·es des droits humains, par leurs pratiques en contexte répressif, élaborent « par le bas » une contestation des modèles de gouvernance dominants. La discussion met ainsi en évidence trois formes de diplomatie des données susceptibles de prolonger ces pratiques : une diplomatie pour les données, une diplomatie par les données, et une diplomatie des données partagée.

PENSER LA GOUVERNANCE DES DONNÉES DEPUIS LES MARGES

L'expression « défenseur·e·s des droits humains » intègre des tensions et des ambiguïtés (voir Nah *et al.*, 2013). Formellement, elle désigne « toute personne qui, individuellement ou avec d'autres, agit à des fins de protection et de réalisation des droits humains, tant au niveau national qu'international » (Nations Unies, 1999, art. 12), sans préciser le champ d'action ni le spectre des enjeux couverts.

Trois éléments en balisent les contours. D'abord, la désignation renvoie à des formes de reconnaissance institutionnelle – politiques, juridiques et administratives – qui légitiment des mesures de protection et engagent la responsabilité des acteurs garants de la sécurité des personnes concernées. Ensuite, elle recouvre un ensemble de pratiques inscrites dans des luttes démocratiques non violentes : documentation des violations, assistance aux victimes, mise en imputabilité et promotion de réformes. Enfin, et surtout, elle suppose un cadrage – public, politique et juridique – des activités et des exactions en termes de droits humains. Discursive et performative, elle affirme une adhésion à des normes et des valeurs, et structure des formes d'identification collective. Considérée sous l'angle de la sociologie des mouvements sociaux, elle constitue le « cadre » à l'intérieur duquel se situent et prennent sens les discours et les pratiques d'intervention politique (Benford *et al.*, 2012).

À rebours d'une vision techniciste des données, cet article s'inscrit dans le champ des *Critical Data Studies*, qui conçoivent la donnée comme un objet politique, conflictuel, et situé (Deutch et Al Khatib, 2018). La mise en données (*datafication*) n'est pas un simple processus technique de conversion de l'information, mais un terrain de lutte qui redéfinit les rapports de pouvoir et soulève des enjeux de souveraineté (Nwagwu, 2025 ; Tufekci, 2017). Cette perspective rejoint les travaux francophones en sciences de l'information et de la communication qui, dans le prolongement de l'économie politique du numérique (Mosco, 2016), conçoivent la régulation des données comme un terrain contesté, indissociable des rapports de pouvoir (Bullich et Clavier, 2019). Les travaux récents sur la diplomatie des données soulignent d'ailleurs que les stratégies d'influence se déploient désormais à travers les infrastructures de données elles-mêmes (Lyubareva, Nocetti et Carey-Libbrecht, 2024). Notre démarche prolonge cette filiation en déplaçant l'orientation de l'article vers les acteurs les plus exposés à la répression.

Ce caractère politique de la donnée se manifeste dans l'ambivalence des techniques numériques : mobilisées dans des luttes prodémocratie pour documenter les violations et pour contourner la censure (Costanza-Chock, 2020 ; Lievrouw, 2023 ; Milan, 2013 ; Meikle, 2018 ; McNutt, 2018), elles servent simultanément des stratégies de répression sophistiquées portées par les États et leurs partenaires privés (Deibert, 2015 ; Michaelsen et Glasius, 2018 ; Morozov, 2011 ; Zuboff, 2019 ; Lyon, 2015). Analysés depuis la perspective des acteurs les

plus exposés, les cadres dominants de gouvernance des données – conçus du point de vue étatique ou économique et centrés sur la régulation, la sécurité nationale ou la valorisation commerciale – révèlent leur indifférence aux asymétries de pouvoir qu’ils produisent. La gouvernance doit alors être envisagée non comme un simple ensemble de règles réparties entre acteurs dominants, mais comme un dispositif attentif aux vulnérabilités générées par la mise en données.

L’hypothèse de travail est que, face à un modèle de gouvernance autoritaire qui inverse la « promesse » d’un Internet libérateur (Diamond, 2010 ; Morozov, 2011), les défenseur·es des droits humains développent non seulement de stratégies de résistance et de contournement, mais aussi, implicitement, de nouvelles formes de régulation. Menées sous la contrainte de la répression et des défaillances sociotechniques, leurs pratiques ne constituent pas de simples ajustements techniques : elles esquissent une « diplomatie des données » alternative, construite « par le bas », par et pour les acteurs les plus vulnérables. Distincte de l’activisme numérique, qui mobilise les plateformes pour amplifier des causes (Costanza-Chock, 2020), la diplomatie des données par le bas conteste les règles mêmes des infrastructures de données et désigne la participation d’acteurs non-étatiques à la reconfiguration des normes de gouvernance : qui produit, qui possède, qui légitime et qui archive les données.

Cette tension entre l’instrumentalisation autoritaire des données et leur réappropriation militante invite à formuler une question de recherche centrale : dans quelle mesure les pratiques numériques des défenseur·es des droits humains, déployées sous contrainte répressive, contribuent-elles à reconfigurer les normes de gouvernance des données au-delà des arènes diplomatiques formelles ? Nous considérons ainsi ces pratiques non comme de simples tactiques défensives, mais comme une production normative implicite qui déplace les frontières entre acteurs légitimes et acteurs périphériques de la diplomatie des données.

MODALITÉS D’ANALYSE DE LA REVUE DE LITTÉRATURE

Cet article présente les résultats d’une analyse systématique de la littérature anglophone publiée entre 2010 et 2024 sur les pratiques numériques liées à la défense et à la répression des droits humains. La collecte du corpus a été réalisée en deux phases successives :

- Phase 1 (2022-2023, période couverte 2010-2021). Une exploration préliminaire a permis d’identifier les bases de données pertinentes et de tester différentes combinaisons de mots-clés. Huit bases de données ont été retenues³. Ce choix visait une couverture à la fois des publications en sciences juridiques, en sciences humaines et sociales et dans le domaine interdisciplinaire des droits humains. La recherche par mots-clés a reposé sur des combinaisons associant des descripteurs relatifs aux acteurs (« human rights defenders », « human rights activists », « human rights activism ») et des descripteurs relatifs aux technologies (« digital technology/ies », « media technology/ies », « communication technology/ies »), articulés par les opérateurs booléens *AND* et *OR*. Les champs de recherche ont varié selon les fonctionnalités des bases (titre, résumé, mots-clés ou texte intégral). Cette phase a été complétée par une collecte par méthode boule de neige (examen des bibliographies, validation par titre/résumé, arbitrage conjoint) et a permis de retenir 52 articles.

- Phase 2 (2024-2025, période couverte 2022-2024). Les mêmes bases ont été réinterrogées avec des équations élargies, intégrant de nouveaux descripteurs identifiés dans la littérature. Quatre blocs thématiques d’équations booléennes ont été structurés, couvrant respectivement les dimensions des technologies, de la sécurité numérique, de la mobilisation en ligne et de la gouvernance informationnelle⁴. Cette seconde phase a conduit à l’ajout de 59 articles, portant le corpus total à 111 articles.

Parallèlement aux deux phases, six revues spécialisées en droits humains ont fait l’objet

d'un dépouillement intégral pour l'ensemble des numéros disponibles en ligne⁵. Les revues représentées dans le corpus relèvent de plusieurs champs disciplinaires : sciences de l'information et de la communication, sociologie, sciences politiques, relations internationales et études spécifiques sur les droits humains.

Les critères d'inclusion et d'exclusion du corpus sont les suivants :

- Critères d'inclusion : (a) publication dans une revue à comité de lecture ; (b) rédaction en langue anglaise⁶ ; (c) publication entre 2010 et 2024 ; (d) traitement explicite de pratiques numériques dans le contexte de la défense ou de la répression des droits humains (usages concrets ou stratégies délibérées de technologies numériques).

- Critères d'exclusion : (a) monographies et chapitres de livres (corpus restreint aux articles pour garantir l'homogénéité) ; (b) thèses et mémoires ; (c) littérature grise (rapports institutionnels, documents gouvernementaux, rapports d'ONG), considérée comme sources empiriques pour de futures analyses ; (d) contributions exclusivement théoriques sans référence à des pratiques numériques identifiables ; (e) articles portant sur les technologies comme objet de droits humains (ex. : bioéthique, santé numérique) sans lien avec des répertoires de défense ou de répression.

L'ensemble du corpus a été codé et analysé à l'aide du logiciel NVivo selon une démarche inductive et comparative. Le processus de codage itératif à double arbitrage (Locke *et al.*, 2022 ; Cascio *et al.*, 2019) a conduit à l'élaboration d'une grille d'analyse thématique structurée en cinq volets : (1) les catégories de discours académiques sur la contribution des techniques numériques à la défense des droits humains ; (2) les répertoires de pratiques mobilisant ces technologies ; (3) les répertoires de pratiques répressives ; (4) les facteurs intervenant dans le potentiel d'usage ; (5) les transformations de la pratique de défense des droits humains jugées liées à la numérisation du champ.

Deux concepts ont guidé la grille. Celui de « répertoire d'action collective » (Tilly, 1995 ; Tilly et Tarrow, 2015) renvoie aux routines stratégiques et tactiques partagées, disponibles pour les acteurs mobilisés (Landry *et al.*, 2015) ; la notion de « pratique », ancrée dans la sociologie des usages des TIC (Denouël et Granjon, 2011 ; Kane, 2013), désigne des activités organisées, récurrentes et orientées. Leur articulation situe les appropriations de techniques au sein des luttes sociales, comme des choix opérés parmi un éventail d'options. Le poids relatif des catégories a été établi à partir du nombre d'articles distincts les mentionnant.

Catégories	Sous-catégories
Catégories de discours sur les technologies numériques	Émancipateurs Ambivalents Critiques
Défenseur·es des droits humains : répertoires de pratiques	Pratiques communicationnelles externes Pratiques contentieuses Pratiques d'enquête et documentation Pratiques organisationnelles
Répertoires de pratiques répressives	Censure Propagande Surveillance
Catégories de facteurs	Développement des capacités Sécurité numérique Accès aux ressources matérielles et techniques Caractéristiques des technologies médiatiques Environnement politique et régulateur
Catégories de transformations	Nouvelles méthodes centrées sur les données Nouvelles pratiques d'interaction avec les publics

Notre démarche s'apparente à une revue intégrative (Torraco, 2016) : plutôt que d'agréger des résultats statistiquement comparables, elle produit une cartographie conceptuelle d'un champ fragmenté à partir d'une diversité de devis et de terrains. Les 111 articles y sont mobilisés comme des contributions dont la juxtaposition raisonnée fait émerger des régularités et des catégories transversales.

LES CATÉGORIES DE RÉPRESSION NUMÉRIQUE : SURVEILLANCE, CENSURE ET PROPAGANDE

Pour saisir la portée des pratiques de résistance, il faut d'abord caractériser le régime de contraintes dans lequel elles s'inscrivent. La revue dégage trois catégories de répression numérique – surveillance, censure et propagande – qui constituent les manifestations opératoires d'un paradigme de gouvernance des données fondé sur le contrôle et l'asymétrie de pouvoir, et révèlent une logique de « mise en armement » (*Weaponization*) des données.

La surveillance repose sur l'automatisation de l'identification et de l'extraction des données: cyber espionnage des dissidents (Anstis et LaFlèche, 2025 ; Belmonte, 2023 ; Chubb et Roberts Lyer, 2024 ; Eguridu *et al.*, 2024), captation massive et indifférenciée des flux de communication et conservation prolongée (Stevens *et al.*, 2023), reconnaissance faciale appliquée à l'identification des manifestants (Brough et Li, 2013 ; Gregory, 2010 ; Murray *et al.*, 2024), et mise sous pression des entreprises technologiques pour obtenir des renseignements sur les personnes ciblées⁷ (Anstis et Barnett, 2022 ; El Gody, 2022 ; Stevens *et al.*, 2023). Ce phénomène est soutenu par des lois et réglementations mises en place pour légitimer la surveillance de masse, la rétention des données et l'accès en temps réel aux communications numériques (Ilori, 2024 ; Yingyi et Benyera, 2024). S'y ajoutent des capacités d'infiltration destinées à repérer et anticiper les actions des opposants (Chubb et Roberts Lyer, 2024 ; Tenove, 2019).

La censure combine mesures préventives, blocage et filtrage, de la censure algorithmique (Oral, 2024 ; Yingyi et Benyera, 2024) à la coupure totale ou partielle de la connectivité (*Kill-Switch*) visant à désorganiser les mobilisations (Kasm, 2018 ; Kaun et Tréré, 2020 ; Siatitsa, 2020). Elle s'articule à des campagnes de dénigrement et de harcèlement – notamment par *doxing* – destinées à détruire la réputation des défenseur-es et à inciter à la violence (Nissen, 2023 ; Strand et Svensson, 2023 ; Svensson *et al.*, 2024).

La propagande mobilise la falsification des contenus (*Deepfake*, *Shallowfake*) reposant sur l'accès à de grands volumes de données (Gregory, 2019), le recours à l'IA générative pour produire de fausses preuves et attaquer la crédibilité des militant-es (Freeman, 2021 ; Gregory, 2022 ; Teo, 2025 ; Tripathi, 2022), ainsi que l'automatisation par botnets et « armées de trolls » qui saturent l'espace public (Tenove, 2019 ; Wilson, 2022 ; Gregory, 2023). Ces tactiques ne sont pas l'apanage des seuls régimes en place (Bradshaw et Howard, 2019) : ce qui caractérise leur usage autoritaire est moins un monopole que l'asymétrie de moyens et la capacité à conjuguer leviers étatiques et capacités computationnelles privées.

Ces modalités montrent que la donnée n'est pas un instrument de contrôle passif, mais l'objet d'une instrumentalisation offensive : le passage de la collecte (surveillance) à l'usage délibéré pour nuire, diffamer et neutraliser – campagnes de *hack-and-leak*, incitation à l'autocensure – incarne précisément la « mise en armement » des données numériques (Anstis, 2024 ; Eguridu *et al.*, 2024 ; Gregory, 2019, 2022 ; Oral, 2024).

LES CATÉGORIES DE CONTRAINTES SOCIOTECHNIQUES

Si les pratiques de surveillance, de censure et de propagande constituent les manifestations tangibles de la répression numérique, leur portée et leur efficacité sont exacerbées par les caractéristiques mêmes de l'écosystème numérique dans et à partir duquel elles se

déploient. L'agentivité des défenseur·es des droits humains est en effet doublement contrainte : non seulement par l'action coercitive directe de leurs adversaires, mais également par un ensemble de vulnérabilités qui affectent leur capacité d'action. Les analyses qui suivent soulignent que ces contraintes se déploient sur plusieurs plans : une dimension matérielle, marquée par une inégalité d'accès aux ressources techniques ; une dimension normative, où l'irrecevabilité juridique de leurs données probantes est constante ; et une dimension infrastructurelle, qui les place en situation de dépendance critique envers des infrastructures contrôlées par des acteurs étatiques ou commerciaux. Cette précarité est exacerbée par un déficit à la fois symbolique, à travers la faible reconnaissance de la légitimité de leurs données et des récits qu'ils élaborent, et politique, illustrée par la faiblesse des mécanismes de soutien diplomatique et de recours internationaux.

Les écrits recensés font état de cinq catégories interdépendantes comprenant : (1) des limitations matérielles et techniques, qui instaurent une fracture de moyens et un accès inégal aux données ; (2) des limitations liées au renforcement des capacités, qui révèlent un déficit de littératie et de ressources pour s'adapter à l'environnement numérique ; (3) des limitations liées à la sécurité numérique, qui exposent de manière disproportionnée les militants à des risques constants ; (4) des limitations imposées par l'environnement médiatique, dont la gouvernance algorithmique et la dépendance aux plateformes commerciales façonnent la visibilité des discours, et enfin : (5) des limitations de l'environnement politique et des cadres réglementaires, qui tendent à criminaliser l'action militante plutôt qu'à la protéger.

Tout d'abord, les limitations d'ordre matériel et technique concernent à la fois les inégalités d'accès aux réseaux (connexion internet déficiente ou inexistante dans certains territoires) et aux infrastructures nécessaires pour traiter des données spécifiques et stratégiques, en valider l'authenticité, les stocker et les analyser (Deutch et Al Khatib, 2018 ; Markland, 2021). La littérature souligne un accès limité à l'imagerie satellite (Herscher, 2014 ; Murray *et al.*, 2022) pour documenter des violations à grande échelle et dans des régions non accessibles (déforestation, mouvements de population, zones de conflit, etc.). Les enjeux de financement et d'équipement témoignent aussi d'une distribution inégale des ressources au sein de la communauté de défense des droits humains. Trevor A. Thrall *et al.* (2014) illustrent par exemple un cercle vicieux de concentration des ressources et de l'attention au sein du milieu des organisations non gouvernementales ; une poignée d'ONG décrites comme des *gatekeepers* dotées de budgets importants sont en mesure d'investir dans des ressources (scientifiques, sites web, vidéos) pour maximiser leur visibilité médiatique, laissant la grande majorité des autres organisations dans l'ombre.

Les limitations liées au renforcement des capacités renvoient au déficit de compétences et de moyens organisationnels pour maîtriser pleinement le cycle de vie de la donnée. La littérature identifie de faibles niveaux de littératie numérique (Strand et Svensson, 2023 ; Zanzoun, 2017) et d'expertise liée aux données et à leur visualisation (Rall *et al.*, 2016), voire des dissensions épistémiques au sein même des organisations sur l'impératif d'adopter des méthodes centrées sur les données (Freeman, 2021 ; Markland, 2021). Par ailleurs, la surabondance de données pouvant être analysées dans le cadre des enquêtes sur des violations menées par les défenseur·es des droits humains transforme un potentiel atout - l'abondance et la disponibilité de l'information - en un fardeau opérationnel, car les organisations ne disposent pas des ressources, ni des dispositions en termes de savoir-faire pour traiter et valoriser de grands volumes de données (Naimer *et al.*, 2017 ; Teo, 2025).

Les limitations liées à la sécurité numérique désignent, quant à elles, l'exposition constante et disproportionnée des défenseur·es des droits humains à des menaces qui exigent d'adopter des précautions qui sont décrites comme un fardeau (Michaelsen, 2018) qui s'étend à l'ensemble des réseaux activistes et à leurs familles. Le risque de voir la menace numérique se traduire en menace physique est jugé élevé en raison des capacités de collecte de données par les autorités et de violation de la vie privée, notamment du fait de failles

de sécurité dans les applications utilisées (Kazansky, 2015 ; Khamdamova, 2020 ; Naimer et al., 2017). En outre, la gouvernance des données est caractérisée par un oligopole numérique transnational où les États et les Gafam exercent une emprise croissante sur les processus de mise en données. Cette centralisation crée une dépendance infrastructurelle critique qui exerce une contrainte sur l'autonomie des défenseur·es des droits humains dans leurs activités. La littérature mentionne à cet égard les règles arbitraires imposées par les plateformes et leur architecture centralisée qui offre des points d'accès privilégiés pour la surveillance, en exploitant par exemple les failles de sécurité ou les mesures insuffisantes de protection de l'anonymat (Gregory, 2019 ; Kazansky, 2015).

Les limitations liées aux médias réfèrent principalement à la manière dont la structure même de la sphère publique numérique est gouvernée par des acteurs privés qui manifestent une certaine passivité à garantir un environnement exempt de violence et de discrimination à l'égard des femmes militantes (Anstis et LaFlèche, 2025 ; Faith, 2022 ; Tripathi, 2022). Les médias socio-numériques sont aussi saturés sur le plan des contenus disponibles, limitant ainsi la portée et l'impact des récits portés par les défenseur·es des droits humains, mais aussi leur capacité à retrouver, à conserver puis à analyser des preuves visuelles publiées par des témoins. En effet, le fait de produire des données, même vérifiées, ne suffit pas à garantir leur impact : la modération (Fife et al., 2023), désormais déléguée à des systèmes d'intelligence artificielle (Makhortykh, 2023) ainsi que la gouvernance algorithmique des plateformes, centrée sur l'optimisation de l'engagement des utilisateurs, contribuent à marginaliser les contenus militants (Della Valle et Simoncini, 2024 ; Gregory, 2019). Ce rapport de dépendance à l'égard des plateformes signifie que les contenus produits sont mis en concurrence avec des campagnes de désinformation et de *trolling* qui saturent l'espace médiatique et fragmentent l'opinion (Freeman, 2021 ; Kasm, 2018). À cet égard, Sam Gregory écrit : « *Actors hostile to human rights, from states to individuals, use the same technologies and platforms as human rights activists and community and grassroots journalists in order to spread misinformation, disinformation, and malinformation, to undermine civil society and democracy, and to perpetuate hate speech.* »⁸ (Gregory, 2019, p. 3-4).

Les limitations liées à l'environnement politique et aux cadres de régulation illustrent la manière dont le droit est souvent instrumentalisé pour entraver les activités des défenseur·es des droits humains et pour prévenir toute forme de contestation politique. Ces limitations concernent aussi les insuffisances du point de vue des normes d'utilisation des sources ouvertes (Murray et al., 2022) et la rétention ou la falsification de preuves par les autorités (Rall et al., 2016). Outre les règles d'utilisation contraignantes des plateformes et l'opacité de leur fonctionnement (Kazansky, 2015; Gregory, 2019; Faith, 2022), la littérature cite abondamment les lois sur la cybersécurité ou sur l'antiterrorisme qui sont fréquemment utilisées pour exiger la collaboration des entreprises privées, pour criminaliser les actes de collecte, de traitement et de diffusion de données par les défenseur·es des droits humains et, plus généralement, pour restreindre les possibilités d'action collective par le biais des technologies numériques (Della Valle et Simoncini, 2024 ; Illori, 2024 ; Nyst, 2018 ; Yingi et Benyera, 2024). La lente adoption de standards permettant l'intégration de preuves numériques devant les tribunaux internationaux dévalorise institutionnellement les fondements des activités d'enquête des défenseur·es des droits humains sur des données probantes.

Les limitations présentées précédemment ne constituent pas de simples obstacles : elles font état de vulnérabilités qui, suivant la critique des plateformes dont le modèle économique est fondé sur la monétisation de l'engagement (Kaun et Tréré, 2020), altèrent, d'une part, les capacités de mobilisation et amplifient, d'autre part, les discours haineux et la toxicité des échanges (Espindola, 2025). Le manque de capacité et de sécurité expose directement les défenseur·es des droits humains à la surveillance et au piratage, tandis que leur dépendance aux plateformes et la faiblesse des cadres juridiques contribuent davantage à leur marginalisation au sein d'un régime de gouvernance qui n'établit pas les normes de

protection et d'inclusion requises pour protéger leurs intérêts et pour reconnaître la précarité de leur situation.

Cet environnement de contraintes sociotechniques, bien que particulièrement hostile à l'action militante, agit aussi comme un catalyseur pour l'émergence de pratiques d'adaptation et d'innovation qui visent à créer des espaces d'autonomie et de sécurité. En mutualisant leurs infrastructures ou en contournant les normes officielles, les défenseur·es des droits humains ne sont pas seulement dans une posture défensive, ils expérimentent de nouvelles formes de gouvernance des données.

DES TACTIQUES DE RÉSISTANCE FACE À LA MISE EN ARMEMENT DES DONNÉES NUMÉRIQUES

La littérature scientifique met en lumière l'affirmation d'un phénomène d'« autoritarisme numérique » (Dragu et Lupu, 2021 ; Pearce et Kendzior, 2012) sur la scène internationale, positionnant *de facto* les données comme des instruments et des enjeux de pouvoir. Les adversaires des droits humains, en particulier les régimes autoritaires, exploitent les dispositifs numériques pour surveiller, pour censurer et pour intimider celles et ceux qui dénoncent leurs exactions. Les processus de mise en données participent ainsi à l'exercice et à la reproduction du pouvoir, et enjoignent les défenseur·es des droits humains à développer des stratégies défensives et de résistance au sein d'environnements numériques potentiellement hostiles où leurs pratiques les exposent à des risques accrus. Les analyses effectuées font émerger des pratiques illustrant des tactiques telles que la mutualisation des infrastructures (hébergement chiffré, *clouds* communautaires) et la formation continue pour renforcer les compétences en sécurité numérique, malgré des défis persistants soulignés par la littérature antérieure, notamment l'accès aux formations, la perception du risque et la disposition à s'engager volontairement pour développer ces compétences (Hankey et Ó Clunaigh, 2013 ; Kazansky, 2015 ; Kogan, 2013).

Sur le plan probatoire (certifier des preuves numériques et garantir la fiabilité des matériaux collectés), les défenseur·es des droits humains mettent en place des réseaux et des coalitions dans le but de mutualiser les ressources du point de vue technique, juridique et organisationnel (Acconcia *et al.*, 2024 ; Luna et Whetstone, 2024 ; Svensson *et al.*, 2024). Ils adoptent également une opacité tactique en utilisant par exemple des outils spécialisés comme *InformaCam* (van der Velden, 2015), pour sécuriser la collecte de données sensibles, mais aussi des tactiques fondées sur le *low-tech* (Hankey et Ó Clunaigh, 2013 ; MacLean, 2023) ou en fonction de la temporalité des conflits, comme l'écrivent Kirsty Fife *et al.* : « *Ownership and authorship of both records and archives need to be flexible in order to ensure security of creators and archivists, whilst also enabling the same people to later become more visible and named as central in the instrumentation of these projects* »⁹ (Fife *et al.*, 2023, p. 489). Enfin, les défenseur·es des droits humains contournent fréquemment les normes officielles en développant des méthodes d'archivage sécurisées. Des initiatives comme le *Syrian Archive*, *Mnemonic* ou le *858 Archive* mettent en œuvre cette pratique en archivant de vastes corpus documentaires de façon structurée. L'intégration de métadonnées et de protocoles de vérification à ce processus vise un triple objectif : prévenir la perte d'informations, garantir leur accessibilité et renforcer leur force probante dans le cadre de mécanismes de justice transitionnelle (Deutch et Al Khatib, 2018 ; Fife *et al.*, 2023 ; Kayyali, 2022).

Des propositions visant la création d'infrastructures d'accueil sécurisées pour les données sensibles semblables à des « cyber host States » (Marelli, 2020) et l'inclusion des défenseur·es des droits humains dans des espaces de délibération technique et réglementaire (Cath, 2021) - notamment auprès des acteurs économiques dans la conception de leurs outils et services techniques (Maréchal, 2015) - émergent par ailleurs dans les résultats d'analyse. Les stratégies de contournement adoptées par les défenseur·es des

droits humains révèlent ainsi les défaillances et les lacunes de la gouvernance des données à l'échelle locale, nationale et internationale. Elles mettent en évidence des normes internationales insuffisantes et des cadres juridiques inadaptés, tout en illustrant l'incapacité – ou le manque de volonté – des pouvoirs publics à instaurer des régimes crédibles et efficaces garantissant la responsabilité des acteurs qui entretiennent des pratiques numériques autoritaires et antidémocratiques.

VERS UNE DIPLOMATIE DES DONNÉES PAR LE BAS

L'étude présentée ici met en évidence la consolidation, dans la littérature récente, du constat d'un régime de répression numérique aux ramifications transnationales, où la mise en données devient un levier d'assujettissement des résistances sociales et politiques. À travers des usages comme la surveillance, la censure, la désinformation ou la criminalisation des activités militantes, les techniques numériques deviennent des outils qui entravent l'action des défenseur·es des droits humains. Cette répression s'appuie sur des vulnérabilités sociotechniques (inégalités d'accès, dépendance aux plateformes, opacité réglementaire, insécurité numérique, illégitimité des récits) qui restreignent l'autonomie des défenseur·es des droits humains dans l'usage des données et les exposent à des risques accrus.

Dans cet environnement hostile, les défenseur·es des droits humains développent des pratiques de contournement, de sécurisation et d'innovation qui ne relèvent pas seulement de la survie tactique. Elles dessinent les contours d'un rapport alternatif aux données, et constituent, à ce titre, des contributions politiques à part entière. Ces pratiques invitent à penser la diplomatie des données non comme une prérogative étatique ou technocratique, mais comme un champ disputé où se négocient des normes, des infrastructures et des représentations.

Pour appuyer ce constat, on peut évoquer les ressorts sur lesquels repose concrètement cette diplomatie des données construite « par le bas ». L'analyse du corpus identifie quatre dimensions constitutives qui, prises ensemble, confèrent aux pratiques des défenseur·es des droits humains une portée qui dépasse le registre de la simple résistance tactique pour s'inscrire dans celui de la production normative.

La première est celle des acteurs impliqués. Les pratiques recensées mettent en évidence une diversité de figures qui ne se réduit pas aux organisations de défense des droits humains au sens classique. Aux côtés des ONG établies et de leurs réseaux, les publications documentent le rôle croissant de collectifs d'enquête mobilisant les sources ouvertes (Deutch et Al Khatib, 2018), de coalitions transnationales spécialisées dans la collecte « forensic » comme *Physicians for Human Rights* et son outil *MediCapt* (Naimer *et al.*, 2017), de réseaux diasporiques qui prolongent l'action militante depuis l'exil (Tenove, 2019 ; Michaelsen, 2018), ou encore de communautés en ligne particulièrement exposées qui développent des pratiques de visibilité stratégique (Acconcia *et al.*, 2024 ; Svensson *et al.*, 2024). Ces acteurs partagent une caractéristique commune : ils opèrent à la marge des espaces institutionnels de délibération, et c'est précisément depuis cette position périphérique qu'ils contribuent à redéfinir les normes relatives aux données.

La deuxième dimension concerne les instruments déployés. La diplomatie des données « par le bas » ne repose pas sur des mécanismes formels de négociation, mais sur un ensemble d'outils et de protocoles qui matérialisent des choix normatifs. Le recours aux méthodes d'investigation en sources ouvertes (OSINT) dans certaines enquêtes internationales (Murray *et al.*, 2022) implique des protocoles de vérification rigoureux des données générées et partagées par des citoyens, traduisant de nouvelles normes de recevabilité probatoire. Des outils comme *InformaCam*, conçu pour intégrer des métadonnées de géolocalisation et d'horodatage au moment de la captation (van der Velden, 2015), ou encore les infrastructures d'archivage distribué développées par Syrian Archive et Mnemonic (Fife *et al.*, 2023 ; Kayyali, 2022), incarnent cette capacité à produire des standards techniques là

où les cadres officiels font défaut. La littérature souligne également l'importance des pratiques qui requièrent moins d'usages techniques (communément appelées *low-tech*), qui témoignent d'une adaptation aux contraintes matérielles des contextes d'intervention (Hankey et Ó Clunaigh, 2013 ; MacLean, 2023).

La troisième dimension est celle des espaces. Les défenseur·es des droits humains tentent d'investir des arènes de délibération hors des forums diplomatiques classiques. Les recherches mentionnent par exemple la participation aux instances de standardisation technique, à l'image des actions documentées par Corinne Cath (2021) au sein de l'*Internet Engineering Task Force* (IETF), des revendications d'inclusion dans la gouvernance des plateformes (Maréchal, 2015), ou encore des propositions concernant la création d'infrastructures d'accueil sécurisées pour les données sensibles, conçues comme des formes d'« asile numérique » (Marelli, 2020). Ces espaces se situent à l'intersection de la technique, du juridique et du politique où se négocient de manière informelle les règles qui déterminent la circulation et la valeur des données.

Enfin, la quatrième dimension est celle de la légitimité. À la différence de la diplomatie étatique, dont l'autorité découle d'un mandat politique formel, la légitimité des défenseur·es des droits humains dans le champ de la gouvernance des données repose sur d'autres fondements normatifs. Tout d'abord, la transparence méthodologique. Les protocoles de collecte, de vérification et d'authentification mis en œuvre dans les enquêtes en sources ouvertes renforcent la traçabilité qui constitue en elle-même une proposition normative puisqu'elle démontre qu'il est possible d'établir la fiabilité des données collectées hors des cadres institutionnels (Murray *et al.*, 2022 ; van der Velden, 2015). Ensuite, la vérifiabilité des données elles-mêmes. Les pratiques d'archivage structuré, d'intégration de métadonnées et de documentation des chaînes de conservation sont conçues pour favoriser la réplique et la contre-vérification par des tiers, qu'il s'agisse de chercheurs, de journalistes ou de juridictions (Deutch et Al Khatib, 2018 ; Fife *et al.*, 2023). Enfin, l'ancrage dans l'expérience située. La proximité avec les contextes de violation confère aux défenseur·es des droits humains une autorité épistémique fondée sur le témoignage direct et sur la connaissance pratique des conditions de terrain. Il s'agit d'une position que la littérature sur la gestion des visibilité en contexte répressif contribue à documenter (Albu, 2023). Ainsi, ces principes fondent la prétention des défenseur·es des droits humains à participer légitimement à la reconfiguration des normes de gouvernance des données et distinguent leur action d'un activisme numérique strictement orienté vers la mobilisation des publics et le plaidoyer pour sensibiliser les décideurs.

Ces quatre dimensions offrent la possibilité de saisir la cohérence interne de ce que nous désignons par « diplomatie des données par le bas ». De toute évidence, il ne s'agit pas d'une diplomatie formalisée dans des traités ou des protocoles internationaux, mais d'une diplomatie pratique, construite par accumulation d'initiatives convergentes. La structuration de cette diplomatie émerge de la conjonction d'acteurs qui se situent aux marges des institutions, d'instruments techniques qui matérialisent des exigences normatives, des espaces de délibération alternatifs où se négocient les règles du jeu, et une légitimité fondée non sur le mandat, mais sur la méthode. C'est dans cette perspective que la typologie de Boyd *et al.* (2019) peut être revisitée.

1. Une *diplomatie pour les données* : une diplomatie en soutien à la circulation et à la protection des données sensibles

Les pratiques documentées montrent que les défenseur·es des droits humains sont confrontés à des obstacles juridiques, techniques et politiques dans la collecte, la conservation, la transmission et la valorisation des données probantes. Une diplomatie des données centrée sur les droits humains devrait reconnaître le statut spécifique des données militantes sensibles ; garantir l'existence d'infrastructures d'asile numérique (archives sécurisées, cyberspaces protégés) ; promouvoir des standards de recevabilité internationale des preuves numériques dans les mécanismes de justice.

2. Des données pour la diplomatie : une diplomatie fondée sur la coopération par les données

Les pratiques de mutualisation, d'archivage collaboratif ou de plaidoyer fondées sur les données sont autant de formes de coopération transnationale entre défenseur·es des droits humains, journalistes, organisations non gouvernementales et chercheurs. Les données ne sont pas seulement un objet diplomatique, elles représentent aussi un médium relationnel. Elles peuvent constituer un vecteur de solidarité entre organisations exposées aux mêmes risques et servent à co-construire des narrations crédibles, à partir de multiples sources et contextes, face à la fragmentation informationnelle.

3. Une diplomatie des données partagée : une reconnaissance des défenseur·es des droits humains comme diplomates des données

Enfin, les défenseur·es des droits humains ne se contentent pas d'exiger des droits sur les données : ils produisent des récits, des savoirs et des normes. En tant que tels, ils participent activement à la fabrique diplomatique des données en rendant visibles des contextes minorés ou censurés, en questionnant les cadres dominants de légitimité, de preuve et de visibilité, et en expérimentant des formes alternatives de gouvernance informationnelle, plus décentralisées, plus sûres et plus éthiques.

Cette « diplomatie des données par le bas » ne se déploie pas de manière uniforme selon les contextes géopolitiques et sociotechniques dans lesquels elle s'inscrit. Dans les pays du « Sud global », les enjeux de dépendance infrastructurelle et de « fracture numérique » constituent des contraintes évidentes. L'accès même aux outils de collecte, de chiffrement ou d'archivage est inégalement distribué, et la gouvernance des données y est fréquemment structurée par des asymétries héritées de dynamiques coloniales (Couldry et Mejias, 2023), *a fortiori* dans les termes même des enjeux débattus dans les instances de délibération. À cet égard, Stefania Milan et Emiliano Treré (2019) défendent un renversement épistémique pour désenclaver des vécus dont la singularité est déniée. Ils mettent ainsi en garde contre un « universalisme des données » qui projette sur l'ensemble des contextes des catégories d'analyse forgées dans le « Nord global » et tend à invisibiliser les expériences situées. Dans le « Nord global », à l'inverse, la dépendance aux plateformes commerciales et la surveillance algorithmique constituent les contraintes dominantes, et la diplomatie « par le bas » prend davantage la forme de revendications d'inclusion dans les instances de standardisation et de régulation. Notre corpus, bien qu'anglophone, couvre une diversité de zones géographiques, de l'Ouganda (Strand et Svensson, 2023) à la République démocratique du Congo (Naimer *et al.*, 2017), du Zimbabwe (Yingi et Benyera, 2024) à la Syrie (Deutch et Al Khatib, 2018) ou la Tunisie (Della Valle et Simoncini, 2024), ce qui donne la possibilité d'observer cette diversité de configurations sans prétendre à l'exhaustivité.

Ces trois dimensions appellent à une redéfinition de la diplomatie des données non plus centrée sur les États et les grandes plateformes, mais ouverte aux savoirs situés, aux pratiques critiques et aux revendications normatives des acteurs vulnérables. Loin d'être de simples objets de protection, les défenseur·es des droits humains apparaissent alors comme des agents diplomatiques légitimes, porteurs d'un autre rapport à la vérité, à la mémoire, à la justice et à la circulation des savoirs.

CONCLUSION

L'autoritarisme numérique fondé sur la mise en données et sa mise en armement reconfigure profondément les rapports de pouvoir et expose les défenseur·es des droits humains à des risques accrus. En rendant visibles les régimes de surveillance, de censure et de propagande, ainsi que les vulnérabilités matérielles, infrastructurelles et juridiques qui les soutiennent, notre revue met en évidence deux contributions.

Elle documente, d'une part, un répertoire robuste et convergent de tactiques militantes

telles que la mutualisation d'infrastructures, la sécurisation des preuves, l'archivage distribué, les coopérations via des réseaux transnationaux, qui dessinent les contours d'une « diplomatie des données par le bas ». Face à cette mise en armement des données, les stratégies de résistance développées par les défenseur·es des droits humains ne peuvent être réduites à de simples ajustements techniques ; elles constituent une contestation implicite des modèles de gouvernance dominants, articulée autour de trois principes opérationnels que sont la protection (sécurité, intégrité et dignité des personnes et des données), le partage et la traçabilité (interopérabilité, préservation et authenticité des chaînes de preuve) et la participation (inclusion des défenseur·es des droits humains dans les arènes de standardisation, de régulation et de modération).

L'analyse révèle, d'autre part, les insuffisances des cadres juridiques actuels et ouvre la voie à des orientations concrètes telles que la reconnaissance d'un statut protecteur des données militantes sensibles, la garantie de la recevabilité des preuves numériques par des standards de collecte et des obligations de conservation, le renforcement des devoirs de vigilance et de transparence des plateformes, l'encadrement de la répression numérique transnationale par la clarification des responsabilités extraterritoriales, et les garanties procédurales pour les défenseur·es des droits humains.

Ce cadre interprétatif prospectif rencontre toutefois deux limites qu'il convient de formuler explicitement. La première est d'ordre méthodologique. En effet, la « diplomatie des données par le bas » émerge ici d'une revue de la littérature savante, et ne saurait capturer toute la complexité du vécu des acteurs de terrain. La manière dont les individus et les collectifs militants perçoivent leurs propres actions, négocient les risques au quotidien et donnent sens à leur engagement numérique reste un impensé de notre démarche dont l'approfondissement parviendrait à mieux inscrire les enjeux discutés dans un agenda de recherche susceptible de dépasser les limites des conceptions dominantes de la régulation de la production et des usages des données (Milan et Tréré, 2019). La seconde limite est d'ordre juridique. La fragmentation des régimes existants relatifs à la circulation transnationale des données, à la protection des informations probantes et à la responsabilité des plateformes rend délicates voire erronées les tentatives d'articuler les pratiques recensées et les instruments juridiques disponibles, qu'ils soient contraignants ou souples.

Ces limites dessinent deux chantiers prioritaires : éprouver empiriquement le cadre proposé par des entretiens et des observations auprès des défenseur·es des droits humains, afin de saisir les stratégies telles qu'elles sont vécues en contexte, et croiser l'analyse des pratiques avec une cartographie fine des régulations existantes (preuve numérique, protection des sources, cybersécurité, devoir de vigilance, export de technologies intrusives), pour identifier les leviers d'évolution et mesurer l'écart entre le droit positif et les attentes révélées.

L'essor rapide de l'intelligence artificielle générative confère une urgence supplémentaire aux chantiers mentionnés. L'IA « mise en armement » amplifie simultanément les capacités de répression par la génération d'hypertrucages destinés à discréditer des défenseur·es des droits humains (Gregory, 2022), par l'automatisation de la modération qui marginalise les contenus militants et dissidents (Makhortykh, 2023), ou par la « violence lente » qu'elle est en mesure de contribuer à exercer sur les droits fondamentaux (Teo, 2025), et les capacités de contre-surveillance, en offrant des outils d'analyse automatisée de preuves visuelles et de détection de contenus falsifiés (Freeman, 2021). La gouvernance des systèmes d'IA et de leurs données d'entraînement constitue dès lors un terrain émergent de diplomatie des données où les voix des acteurs les plus vulnérables demeurent largement absentes.

NOTES

¹ Le concept d'autoritarisme numérique renvoie à l'usage, par des États, d'un ensemble articulé de mesures, de processus et de dispositifs numériques à des fins antidémocratiques (voir Glasius et Michaelsen, 2018). Il se manifeste notamment par le contrôle et la manipulation de l'information, la surveillance (ciblée ou généralisée) des populations, les intrusions dans les systèmes, la destruction de données et d'infrastructures, ainsi que la perturbation des activités des opposants. Ces pratiques font peser des menaces directes sur la sécurité des individus et contribuent à restreindre leur capacité d'action autonome.

² Les contraintes sociotechniques désignent les limitations à l'action qui émergent de l'articulation complexe entre les architectures techniques (infrastructures numériques, plateformes et dispositifs de traitement de données) et les structures sociales, façonnées par les rapports de pouvoir, les cadres réglementaires, les normes et les inégalités d'accès et d'usage aux ressources. Elles se construisent dans l'interaction dynamique entre pratiques, dispositifs et rapports de pouvoir, modulant de manière différenciée les capacités d'agir, les marges de manœuvre et les zones de vulnérabilité des acteurs (voir Coutant, 2015, sur les approches sociotechniques).

³ Les bases de données retenues sont les suivantes : *HeinOnline - Law Journal Library* ; *Brill Online Books : Human Rights and Humanitarian Law* ; *Oxford scholarship Online - Law* ; *Scopus (Elsevier)* ; *CANLII* ; *ProQuest* ; *Google Scholar*.

⁴ Les équations détaillées de la phase 2 sont les suivantes : (a) (« human rights defenders » OR « human rights activists ») AND (« digital technologies » OR « information and communication technologies » OR « social media » OR « digital platforms ») ; (b) idem AND (« digital security » OR « surveillance » OR « censorship » OR « repression » OR « online harassment ») ; (c) idem AND (« online mobilization » OR « digital activism » OR « social media » OR « freedom of expression ») ; (d) idem AND (« data protection » OR « technology legislation » OR « internet governance » OR « hate speech » OR « disinformation »).

⁵ Les revues concernées sont les suivantes : *International Journal on Human Rights*, *Human Rights Quarterly*, *The International Human Rights Law Review*, *Asia-Pacific Journal on Human Rights and the Law*, *Journal of Human Rights Practice*, et *International Review of the Red Cross*.

⁶ Le choix d'un corpus exclusivement anglophone est délibéré, fondé sur trois constats issus du processus de recension. Premièrement, les recherches exploratoires sur les bases francophones (Cairn, Érudit, Persée) ont révélé une rareté notable de publications spécifiques à l'articulation pratiques numériques / droits humains. Deuxièmement, la variabilité terminologique en français (« droits humains », « droits de l'homme », « droits de la personne ») complexifie considérablement la systématisation et introduit un risque d'incomplétude. Troisièmement, l'objectif de l'étude était précisément de cartographier la manière dont la littérature internationale, dont l'anglais est *de facto* la langue véhiculaire dans les revues à comité de lecture, traite de cette problématique.

⁷ Les écrits mentionnent le logiciel espion Pegasus développé par l'entreprise israélienne *NSO Group* comme outil fréquemment déployé pour mettre en place cette surveillance ; qualifié de technologie posant des « risques sans précédent » par le Contrôleur européen de la protection des données (Anstis, 2024), le logiciel accorde à son opérateur un accès total au téléphone ciblé, y compris le contenu d'applications chiffrées comme *Signal* ou *WhatsApp*, ainsi qu'au micro et à la caméra du dispositif numérique.

⁸ « Les acteurs hostiles aux droits humains, qu'il s'agisse d'États ou d'individus, utilisent les mêmes techniques et plateformes que les militant·es des droits humains et les journalistes communautaires et de terrain afin de propager la mésinformation, la désinformation et les informations malveillantes, de saper la société civile et la démocratie, et de perpétuer les discours de haine. » (Gregory, 2019, p. 3-4, notre traduction).

⁹ « La propriété et la paternité des documents comme des archives doivent être suffisamment flexibles pour, d'une part, garantir la sécurité des créateurs et des archivistes, et d'autre part, permettre à ces mêmes personnes d'être ultérieurement reconnues et de gagner en visibilité en tant que figures centrales dans la conduite de ces projets. » (Fife *et al.*, 2023, p 489, notre traduction).

RÉFÉRENCES BIBLIOGRAPHIQUES

Acconcia, Giovanni ; Perego, Alessandro ; Perini, Luca (2024), « LGBTQ activism in repressive contexts: The struggle for (in)visibility in Egypt, Tunisia and Turkey », *Social Movement Studies*, vol. 23, n° 2, p. 207-225.

Albu, Oana B. (2023), « Managing visibilities: The shades and shadows of NGO work in repressive contexts », *Management Communication Quarterly*, vol. 37, n° 3, p. 659-685.

Anstis, Siena (2024), « Regulating transnational dissident cyber espionage », *International & Comparative Law Quarterly*, vol. 73, n° 1, p. 259-274.

Anstis, Siena ; Barnett, Sarah (2022), « Digital transnational repression and host states' obligation to protect against human rights abuses », *Journal of Human Rights Practice*, vol. 14, n° 2, p. 698-725.

Anstis, Siena ; LaFlèche, Émilie (2025), « Gender-based digital transnational repression as a global authoritarian practice », *Globalizations*, vol. 22, n° 4, p. 671-688.

Assemblée générale des Nations Unies (1999), *Déclaration sur le droit et la responsabilité des individus, groupes et organes de la société de promouvoir et protéger les droits de l'homme et les libertés fondamentales universellement reconnues (A/RES/53/144)*, New York : Organisation des Nations Unies.

Belmonte, Laura (2023), « Can human rights survive technology? », *Diplomatic History*, vol. 47, n° 1, p. 1-18.

Benford, Robert D. ; Snow, David A. ; Plouchard, Nicolas M. (2012), « Processus de cadrage et mouvements sociaux : présentation et bilan », *Politix*, n° 99, p. 217-255.

Boyd, Andrew ; Gatewood, Justin ; Thorson, Stuart ; Dye, Thomas D. V. (2019), « Data diplomacy », *Science & Diplomacy*, vol. 8, n° 1, [en ligne], consulté le 24 septembre 2025, <http://sciencediplomacy.org/article/2019/data-diplomacy>.

Bradshaw, Samantha ; Howard, Philip N. (2019), *The Global Disinformation Order: 2019 Global Inventory of Organised Social Media Manipulation*, Oxford : Oxford Internet Institute (Computational Propaganda Project).

Brough, Melissa ; Li, Zhen (2013), « Media systems dependency and human rights online video: The "Saffron Revolution" and WITNESS's Hub », *International Journal of Communication*, vol. 7, p. 281-304.

Bullich, Vincent ; Clavier, Viviane (2019), « Production des données, 'Production de la société'. Les Big Data et algorithmes au regard des Sciences de l'information et de la communication », *Les Enjeux de l'information et de la communication*, vol. 19, n° 2, p. 5-14.

Cascio, M. Ariel ; Lee, Eunlye ; Vaudrin, Nicole ; Freedman, Darcy A. (2019), « A team-based approach to open coding: Considerations for creating intercoder consensus », *Field Methods*, vol. 31, n° 2, p. 116-130.

Cath, Corinne (2021), « The technology we choose to create: Human rights advocacy in the Internet Engineering Task Force », *Telecommunications Policy*, vol. 45, n° 6, [en ligne], consulté le 30 avril 2026, <https://doi.org/10.1016/j.telpol.2021.102144>.

Chubb, Andrew ; Roberts Lyer, Kira (2024), « Transnational human rights violations: Addressing the evolution of globalized repression through national human rights institutions », *Journal of Human Rights Practice*, vol. 16, n° 3, p. 770-793.

Costanza-Chock, Sasha (2020), *Design Justice: Community-Led Practices to Build the Worlds We Need*, Cambridge : The MIT Press.

Couldry, Nick ; Mejias, Ulises A. (2023), « The decolonial turn in data and technology research: What is at stake and where is it heading? », *Information, Communication & Society*, vol. 26, n° 4, p. 786-802.

Coutant, Alexandre (2015), « Les approches sociotechniques dans la sociologie des usages en SIC », *Revue française des sciences de l'information et de la communication*, n° 6, [en ligne], consulté le 30 avril 2026, <https://doi.org/10.4000/rfsic.1271>.

Deibert, Ronald (2015), « Authoritarianism goes global: Cyberspace under siege », *Journal of Democracy*, vol. 26, n° 3, p. 64-78.

Della Valle, Chiara ; Simoncini, Giulia (2024), « Shades of presence in post-2011 Tunisia: Evolving political positions in feminist and queer activism », *Contemporary Politics*, p. 1-22.

Denouël, Julie ; Granjon, Fabien (2011), (dir.), *Communiquer à l'ère numérique. Regards croisés sur la sociologie des usages*, Paris : Presses des Mines.

Deutch, Jessica ; Al Khatib, Hadi (2018), « The Syrian Archive: A methodological case study of open-source investigation of state crime using video evidence from social media platforms », *State Crime Journal*, vol. 7, n° 1, [en ligne], consulté le 30 avril 2026, <https://doi.org/10.13169/statecrime.7.1.0046>.

Diamond, Larry (2010), « Liberation technology », *Journal of Democracy*, vol. 21, n° 3, p. 69-83.

Dragu, Tiberiu ; Lupu, Yonatan (2021), « Digital authoritarianism and the future of human rights », *International Organization*, vol. 75, n° 4, p. 991-1017.

Eguridu, Hope ; Tolani, Tolu ; Olusola, Opeyemi ; Sakpere, Wilson (2024), « Impact of information technology on civil rights and criminal liabilities », *African Journal of Law, Ethics and Education*, vol. 7, n° 1, [en ligne], consulté le 30 avril 2026, <https://ajleejournal.com/index.php/ajlee/article/view/168>.

El Gody, Ahmed (2022), « The revolution continues: Mapping the Egyptian Twittersphere a decade after the 2011 Revolution », *First Monday*, vol. 27, n° 8, [en ligne], consulté le 30 avril 2026, <https://doi.org/10.5210/fm.v27i8.11775>.

Espíndola, José (2025), « Attributing responsibility to Big Tech for mass atrocity: Social media and transitional justice », *Perspectives on Politics*, vol. 23, n° 3, p. 1-15.

Faith, Becky (2022), « Tackling online gender-based violence; understanding gender, development, and the power relations of digital spaces », *Gender, Technology and Development*, vol. 26, n° 3, p. 325-340.

Fife, Kirsty ; Flinn, Andrew ; Nyhan, Julianne (2023), « Documenting resistance, conflict and violence: A scoping review of the role of participatory digital platforms in the mobilisation of resistance », *Archival Science*, vol. 23, n° 3, p. 471-499.

Freeman, Lindsay (2021), « Weapons of war, tools of justice: Using artificial intelligence to investigate international crimes », *Journal of International Criminal Justice*, vol. 19, n° 1, p. 35-53.

Glasius, Marlies ; Michaelsen, Marcus (2018), « Illiberal and authoritarian practices in the digital sphere. Prologue », *International Journal of Communication*, vol. 12, p. 3795-3813.

Gregory, Sam (2010), « Cameras everywhere: Ubiquitous video documentation of human

rights, new forms of video advocacy, and considerations of safety, security, dignity and consent », *Journal of Human Rights Practice*, vol. 2, n° 2, p. 191-207.

Gregory, Sam (2019), « Cameras everywhere revisited: How digital technologies and social media aid and inhibit human rights documentation and advocacy », *Journal of Human Rights Practice*, vol. 11, n° 2, p. 373-392.

Gregory, Sam (2022), « Deepfakes, misinformation and disinformation and authenticity infrastructure responses: Impacts on frontline witnessing, distant witnessing, and civic journalism », *Journalism*, vol. 23, n° 3, p. 708-729.

Gregory, Sam (2023), « Fortify the truth: How to defend human rights in an age of deepfakes and generative AI », *Journal of Human Rights Practice*, vol. 15, n° 3, p. 702-714.

Hankey, Stephanie ; Ó Clunaigh, Daniel (2013), « Rethinking risk and security of human rights defenders in the digital age », *Journal of Human Rights Practice*, vol. 5, n° 3, p. 535-547.

Herscher, Andrew (2014), « Surveillant witnessing: Satellite imagery and the visual politics of human rights », *Public Culture*, vol. 26, n° 3, p. 469-500.

Ilori, Tunde (2024), « Protecting digital rights through soft law: Ensuring the implementation of the revised Declaration of Principles on Freedom of Expression and Access to Information in Africa », *African Human Rights Law Journal*, vol. 24, n° 1, p. 1-30.

Kane, Oumar (2013), « Les usages des TIC entre analyse sociotechnique et théories de l'appropriation : état de la littérature », *Les Enjeux de l'information et de la communication*, p. 23-42, [en ligne], consulté le 18 novembre 2025, <https://www.academia.edu/7277534>.

Kasm, Shady (2018), « Redefining publics: Mosireen, state crime and the rise of a digital public sphere », *State Crime Journal*, vol. 7, n° 1, [en ligne], consulté le 30 avril 2026, <https://doi.org/10.13169/statecrime.7.1.0100>.

Kaun, Anne ; Treré, Emiliano (2020), « Repression, resistance and lifestyle: Charting (dis)connection and activism in times of accelerated capitalism », *Social Movement Studies*, vol. 19, n° 5-6, p. 697-715.

Kayyali, Dia (2022), « Digital memory, evidence, and social media. Lessons learned from Syria », *Sociologica*, vol. 16, n° 2, p. 253-259.

Kazansky, Becky (2015), « FCJ-195 Privacy, responsibility, and human rights activism », *The Fibreculture Journal*, n° 26, p. 190-208.

Khamdamova, Feruza (2020), « The impact of digital technologies on human rights », *TSUL Legal Report International electronic scientific journal*, vol. 1, n° 1, [en ligne], consulté le 30 avril 2026, <https://legalreport.tsul.uz/index.php/journal/article/view/21>.

Kogan, Vladimir (2013), « Protecting human rights defenders in the North Caucasus: Reflections on developments from 2009 to the present », *Journal of Human Rights Practice*, vol. 5, n° 3, p. 500-511.

Land, Molly K. ; Aronson, Jay D. (2020), « Human rights and technology: New challenges for justice and accountability », *Annual Review of Law and Social Science*, vol. 16, n° 1, p. 223-240.

Landry, Normand ; Aubin, France ; Sénécal, Michel (2015), « Médias et mouvements sociaux: pratiques de mobilisations collectives », *Canadian Journal of Media Studies / Revue canadienne d'études médiatiques*, vol. 13, n° 1, p. 124-144, [en ligne], consulté le 30 avril 2026, <http://cjms.fims.uwo.ca/issues/13-01/Landry.pdf>.

Landry, Normand ; Hocine, Julien ; Fitzbay, Simon ; Dupuis, Marie-Josée (2025), « Pratiques numériques et mise en données de la défense des droits humains », *Communication*, vol. 42, n° 2, [en ligne], consulté le 30 avril 2026, <https://doi.org/10.4000/15bo5>.

Lievrouw, Leah A. (2023), *Alternative and Activist New Media*, 2e éd., Cambridge : Polity Press.

Locke, Karen ; Feldman, Martha ; Golden-Biddle, Karen (2022), « Coding practices and iterativity: Beyond templates for analyzing qualitative data », *Organizational Research Methods*, vol. 25, n° 2, p. 262-284.

Luna, Krishna C. ; Whetstone, Crystal (2024), « Women, peace and security: Digitalization and cyber feminist solidarity building in the global South », *Women's Studies International Forum*, vol. 105, [en ligne], consulté le 30 avril 2026, <https://www.sciencedirect.com/science/article/pii/S0277539524000906>.

Lyon, David (2015), *Surveillance after Snowden*, Cambridge : Polity Press.

Lyubareva, Inna ; Nocetti, Julien ; Carey-Libbrecht, Loïc (2024), « Digital diplomacy: Evolution of diplomatic and influence strategies in the digital age », *Réseaux*, n° 245, p. 11-35.

MacLean, Kate (2023), « Interactive digital platforms, human rights fact production, and the International Criminal Court », *Journal of Human Rights Practice*, vol. 15, n° 1, p. 84-99.

Makhortykh, Mykola (2023), « Unreliable narrators or untimely archivists? Challenges of using digital platforms for documenting and remembering Russia's war in Ukraine », *Georgetown Journal of International Affairs*, vol. 24, n° 2, p. 165-173.

Maréchal, Nathalie (2015), « Ranking Digital Rights: Keeping the internet safe for advocacy », *The Fibreculture Journal*, n° 26, p. 302-308.

Marelli, Massimo (2020), « Hacking humanitarians: Defining the cyber perimeter and developing a cyber security strategy for international humanitarian organizations in digital transformation », *International Review of the Red Cross*, vol. 102, n° 913, p. 367-387.

Markland, Adam (2021), « Epistemic transformation at the margins: Resistance to digitalisation and datafication within global human rights advocacy », *Global Society*, vol. 36, n° 1, p. 113-133.

McNutt, John G. (2018), (dir.), *Technology, Activism, and Social Justice in a Digital Age*, New York : Oxford University Press.

Meikle, Graham (2018), (dir.), *The Routledge Companion to Media and Activism*, London : Routledge.

Michaelsen, Marcus (2018), « Exit and voice in a digital age: Iran's exiled activists and the authoritarian state », *Globalizations*, vol. 15, n° 2, p. 248-264.

Michaelsen, Marcus ; Glasius, Marlies (2018), « Authoritarian practices in the digital age — Introduction », *International Journal of Communication*, vol. 12, p. 7.

Milan, Stefania (2013), *Social Movements and Their Technologies: Wiring Social Change*, Basingstoke : Palgrave Macmillan.

Milan, Stefania ; Treré, Emiliano (2019), « Big data from the South(s): Beyond data universalism », *Television & New Media*, vol. 20, n° 4, p. 319-335.

Morozov, Evgeny (2011), *The Net Delusion: The Dark Side of Internet Freedom*, 1re éd., New York : Public Affairs.

Mosco, Vincent (2016), « Après l'Internet : le Cloud, les big data et l'Internet des objets », *Les Enjeux de l'information et de la communication*, vol. 17, n° 2, p. 145-155.

Murray, David ; Fussey, Pete ; Hove, Kennedy ; Wakabi, Wairagala ; Kimumwe, Peter ; Saki, Obert ; Stevens, Alex (2024), « The chilling effects of surveillance and human rights: Insights from qualitative research in Uganda and Zimbabwe », *Journal of Human Rights Practice*, vol. 16, n° 1, p. 397-412.

Murray, David ; McDermott, Yvonne ; Koenig, Kristen A. (2022), « Mapping the use of open source research in UN human rights investigations », *Journal of Human Rights Practice*, vol. 14, n° 2, p. 554-581.

Nah, Alice M. ; Bennett, Karen ; Ingleton, Debbie ; Savage, James (2013), « A research agenda for the protection of human rights defenders », *Journal of Human Rights Practice*, vol. 5, n° 3, p. 401-420.

Naimer, Karen ; Brown, William ; Mishori, Ranit (2017), « MediCapt in the Democratic Republic of the Congo: The design, development, and deployment of mobile technology to document forensic evidence of sexual violence », *Genocide Studies and Prevention*, vol. 11, n° 1, p. 25-35.

Nissen, Anne (2023), « Gender-transformative remedies for women human rights defenders », *Business and Human Rights Journal*, vol. 8, n° 3, p. 369-402.

Nwagwu, Williams E. (2025), « Datafication research (1994-2023) : Three decades of evolving methodology in data science », *Topoi*, vol. 44, p. 1049-1070.

Nyst, Carly (2018), « Secrets and lies: The proliferation of state surveillance capabilities and the legislative secrecy which fortifies them — An activist's account », *State Crime Journal*, vol. 7, n° 1, p. 8-23.

Oral, Ece (2024), « The environmental rule of law and the protection of human rights defenders: Law, society, technology, and markets », *International Environmental Agreements: Politics, Law and Economics*, vol. 24, n° 2-3, p. 393-421.

Pearce, Katy E. ; Kendzior, Sarah (2012), « Networked authoritarianism and social media in Azerbaijan », *Journal of Communication*, vol. 62, n° 2, p. 283-298.

Rall, Kirsten ; Satterthwaite, Margaret L. ; Pandey, Abhishek V. ; Emerson, John ; Boy, Jeremy ; Nov, Oded ; Bertini, Enrico (2016), « Data visualization for human rights advocacy », *Journal of Human Rights Practice*, vol. 8, n° 2, p. 171-197.

Shahbaz, Adrian (2018), « Fake news, data collection, and the challenge to democracy. Freedom on the Net 2018: The rise of digital authoritarianism », *Freedom House*, [en ligne], consulté le 29 septembre 2025, <https://freedomhouse.org/report/freedom-net/2018/rise-digital-authoritarianism>.

Siatitsa, Ilia (2020), « Freedom of assembly under attack: General and indiscriminate surveillance and interference with internet communications », *International Review of the Red Cross*, vol. 102, n° 913, p. 181-198.

Stevens, Alex ; Fussey, Pete ; Murray, David ; Hove, Kennedy ; Sake, Obert (2023), « "I started seeing shadows everywhere": The diverse chilling effects of surveillance in Zimbabwe », *Big Data & Society*, vol. 10, n° 1, p. 1-14.

Strand, Cecilia ; Svensson, Jakob (2023), « Towards a situated understanding of vulnerability — An analysis of Ugandan LGBT+ exposure to hate crimes in digital spaces », *Journal of Homosexuality*, vol. 70, n° 12, p. 2806-2827.

Svensson, Jakob ; Larsson, Anders O. ; Strand, Cecilia (2024), « Who relates to whom and according to which rationale? Visibility and advocacy in the Ugandan LGBT+ Twittersphere », *New Media & Society*, vol. 27, n° 7, [en ligne], consulté le 30 avril 2026, <https://doi.org/10.1177/14614448241236853>.

Tenove, Chris (2019), « Networking justice: digitally-enabled engagement in transitional justice by the Syrian diaspora », *Ethnic and Racial Studies*, vol. 42, n° 11, p. 1950-1969.

Teo, Sherly A. (2025), « Artificial intelligence and its "slow violence" to human rights », *AI and Ethics*, vol. 5, p. 2265-2280.

Thrall, A. Trevor ; Stecula, Dominik ; Sweet, Diana (2014), « May we have your attention

please? Human-rights NGOs and the problem of global communication », *The International Journal of Press/Politics*, vol. 19, n° 2, p. 135-159.

Tilly, Charles (1995), *Popular Contention in Great Britain, 1758-1834*, Cambridge : Harvard University Press.

Tilly, Charles ; Tarrow, Sidney (2015), *Contentious Politics*, 2e éd., New York : Oxford University Press.

Torraco, Richard J. (2016), « Writing integrative literature reviews: Guidelines and examples », *Human Resource Development Review*, vol. 15, n° 4, [en ligne], consulté le 30 avril 2026, <https://doi.org/10.1177/1534484316671606>.

Tripathi, Srishti (2022), « Women and online harassment », *Feminist Dissent*, n° 6, p. 133-158.

Tufekci, Zeynep (2017), *Twitter and Tear Gas: The Power and Fragility of Networked Protest*, New Haven : Yale University Press.

van der Velden, Lonneke (2015), « Forensic devices for activism: Metadata tracking and public proof », *Big Data & Society*, vol. 2, n° 2, [en ligne], consulté le 30 avril 2026, <https://doi.org/10.1177/2053951715612823>.

Wilson, Richard A. (2022), « Digital authoritarianism and the global assault on human rights », *Human Rights Quarterly*, vol. 44, n° 4, p. 704-739.

Yingi, Emmanuel ; Benyera, Everisto (2024), « The future of democracy in the digital era: Internet shutdowns, cyber laws and online surveillance in Zimbabwe », *Alternatives: Global, Local, Political*, vol. 50, n° 2, [en ligne], consulté le 30 avril 2026, <https://doi.org/10.1177/03043754241259878>.

Zanzoun, Mustapha (2017), « The use of social media in promoting human rights among speakers of English in Morocco: The case of Facebook », *International Journal of Education and Research*, vol. 5, n° 9, p. 169-184.

Zuboff, Shoshana (2019), *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*, New York : Public Affairs.

De l'ignorance technique aux données citoyennes. Le cas des cancers pédiatriques en Pays de Retz

Article inédit, mis en ligne le 23 septembre 2026

Saliha HADNA

Saliha HADNA est maîtresse de conférences en Sciences de l'information et de la communication (SIC) à l'Université de Toulouse et au laboratoire UTOPI (UMR 5311). Ses travaux de recherche sont au croisement des SIC, de la sociologie de l'expertise et des STS pour étudier l'évolution des formes de production des connaissances et d'ignorances dans des contextes de controverses.

saliha.hadna-bremand@utoulouse.fr

Plan de l'article

Résumé et mots clés en français

Title, abstract and keywords in english

Titulo, resumen y palabras clave en espanol

Introduction

De la carence informationnelle au problème public : le rôle du collectif SCE dans l'alerte sanitaire

Les données épidémiologiques comme levier majeur dans les controverses en santé environnementale

Conclusion

Notes

Références bibliographiques

RÉSUMÉ

En s'appuyant sur une série d'entretiens et une analyse documentaire sur la controverse autour des cancers pédiatriques de Sainte-Pazanne, on montre dans un premier temps que la création du collectif Stop aux Cancers de nos Enfants (SCE) a contribué à l'institutionnalisation du mouvement citoyen. Il devient un interlocuteur privilégié des services de l'État en même temps que le révélateur de leurs limites dans la gestion du problème sanitaire. On montre dans un second temps que la mise en place de l'Institut citoyen de recherche et de prévention en santé environnementale (ICRePSE) participe à la scientification de l'argumentaire citoyen qui place au cœur de ses enjeux un concept rapidement écarté dans la démarche d'enquête des services de l'État : l'effet cocktail. La production de données épidémiologiques apparaît alors au centre des rapports de forces : tantôt évitées, tantôt recherchées. Le mouvement trouve un écho plus important à l'échelle nationale qui semble plus soucieuse de la situation locale, ce que démontre la récente adoption par l'Assemblée nationale de la proposition de loi sur la création d'un registre national des cancers.

Mots clés

Cancers pédiatriques, données épidémiologiques, effet cocktail, enquête citoyenne, rapports de force

TITLE

From technical ignorance to citizen data. The case of paediatric cancers in Pays de Retz

Abstract

Drawing on a series of interviews and a review of the literature regarding the controversy surrounding pediatric cancers in Sainte-Pazanne, we first demonstrate that the creation of the Stop aux Cancers de nos Enfants (SCE) collective contributed to the institutionalization of the citizen movement. It became a key interlocutor for government agencies while simultaneously highlighting their limitations in managing the public health issue. We then show that the establishment of the Citizen Institute for Research and Prevention in Environmental Health (ICRePSE) contributes to the scientification of the citizen argument, which places at the heart of its concerns a concept quickly dismissed in the government agencies' investigative approach: the cocktail effect. The production of epidemiological data thus appears at the center of the power struggle: sometimes avoided, sometimes sought after. The movement is finding a stronger echo at the national level, which seems more concerned with the local situation, as demonstrated by the National Assembly's recent adoption of the bill to create a national cancer registry.

Keywords

Paediatric cancers, epidemiological data, cocktail effect, citizen survey, power relations

TITULO

De la ignorancia técnica a los datos generados por los ciudadanos. El caso de los cánceres infantiles en el Pays de Retz

Resumen

A partir de una serie de entrevistas y un análisis documental sobre la controversia en torno a los cánceres infantiles de Sainte-Pazanne, se demuestra, en primer lugar, que la creación del colectivo «Stop aux Cancers de nos Enfants» (SCE) contribuyó a la institucionalización del movimiento ciudadano. Se convierte en un interlocutor privilegiado de los servicios del Estado, al tiempo que pone de manifiesto sus limitaciones en la gestión del problema sanitario. En un segundo momento, se muestra que la creación del Instituto Ciudadano de Investigación y Prevención en Salud Ambiental (ICRePSE) contribuye a la cientificación de los argumentos ciudadanos, que sitúan en el centro de sus preocupaciones un concepto rápidamente descartado en el proceso de investigación de los servicios del Estado: el efecto cóctel. La producción de datos epidemiológicos aparece entonces en el centro de las relaciones de poder: a veces evitada, a veces buscada. El movimiento encuentra un eco más importante a escala nacional, que parece más preocupada por la situación local, como demuestra la reciente aprobación por la Asamblea Nacional de la propuesta de ley sobre la creación de un registro nacional de cánceres.

Palabras clave

Cánceres infantiles, datos epidemiológicos, efecto cóctel, encuesta ciudadana, relaciones de poder

INTRODUCTION

Dans certaines controverses, les données brutes sont recherchées comme l'or dans la mesure où elles représentent la matière première pour construire une information, puis une connaissance stabilisée (Ackoff, 1989), en plus de constituer la « base du dialogue ouvert [...] entre les collectivités publiques, les entreprises et les citoyens » (Lehmans, 2018). Dans les années 1980, les principales revendications associatives concernaient l'accès à l'information auxquelles ont tenté de répondre les dispositifs administratifs, tels que la CADA (Commission d'accès aux documents administratifs) ou encore les instances d'information, puis de concertation. Mais ces revendications semblent aujourd'hui avoir évolué dans la mesure où ces associations se révèlent parfois en capacité de produire elles-mêmes les données manquantes ou considérées comme erronées.

En s'appuyant sur la controverse autour des cancers pédiatriques de Sainte-Pazanne (Loire-Atlantique), cet article met en lumière le rôle central des données comme moyen de redéfinition des rapports de force : tantôt dissimulées par certains qui les perçoivent comme une menace pour leur stratégie de maintien de l'ignorance, tantôt convoitées par ceux qui, au contraire, envisagent de lancer l'alerte (Chateauraynaud *et al.*, 2019). Qu'il s'agisse de « l'ignorance stratégique » comme outil de la gouvernance des risques (Proctor, 2011 ; Mc Goey, 2012 ; Girel, 2017 ; Hadna, 2017) ou de « l'ignorance technique » (Hadna, 2017) comme limite épistémologique, toutes deux conduisent une partie des citoyens à développer des canaux de production et de circulation de données alternatifs (Hadna, 2025).

L'article cherche à comprendre dans quelles mesures ces initiatives citoyennes sont nées d'une carence informationnelle et à quelles conditions elles sont devenues capables de produire des données en santé environnement susceptibles d'être mobilisées pour l'amélioration des politiques publiques. En outre, il s'agit d'analyser de quelle manière ce mouvement citoyen s'est peu à peu institutionnalisé, d'une part pour publiciser le problème relatif aux cancers pédiatriques, d'autre part pour contribuer plus largement à la production et à la légitimation de données fiables par un travail de collecte et de croisement de données en santé-environnement.

L'Institut citoyen de recherche et de prévention en santé environnementale (ICRePSE) a été créé en 2023 pour répondre au manque de données épidémiologiques sur le territoire du Pays de Retz où plusieurs clusters de cancers pédiatriques sont observés depuis le milieu des années 2010. En mobilisant les données issues du projet de recherche *Citizen Initiatives for Epidemiological Data* (Cied)¹, cet article cherche à comprendre comment cette initiative citoyenne parvient à s'imposer localement comme une source complémentaire de données et, *in fine*, de connaissances sur les facteurs de risques liés à la survenue des cancers.

Dans un premier temps, l'article s'attache à montrer que cette initiative se développe en réponse à une « carence informationnelle » freinant la résolution de problèmes sanitaires urgents dans un contexte global de « défiance citoyenne » (Rosanvallon, 2006) vis-à-vis des données officielles. Dans ce cadre-là, la diplomatie des données peut ici être mobilisée comme un concept offrant la possibilité de « décrire les motivations, les méthodes, les valeurs, les politiques et d'autres dimensions des négociations impliquant des données ou leur utilisation » et permettant aux parties prenantes de « susciter ou (d')empêcher le changement » (Boyd *et al.*, 2019).

Dans un second temps, l'article met en évidence que le rapport de force qui s'est établi entre le collectif « Stop aux Cancers de nos Enfants » (SCE) et les services de l'État a porté sur les données épidémiologiques, qui deviendront la pierre angulaire de l'ICRePSE. En outre, nous verrons que, face aux blocages posés à l'échelle locale, ce mouvement citoyen a ouvert sa communication à l'échelle nationale afin d'influencer les décisions politiques en faveur du développement d'un registre national de cancers.

DE LA CARENCE INFORMATIONNELLE AU PROBLÈME PUBLIC : LE RÔLE DU COLLECTIF SCE DANS L'ALERTE SANITAIRE

La carence informationnelle comme point de départ de la démarche d'enquête

La controverse autour des cancers pédiatriques de Sainte-Pazanne (Loire-Atlantique) est née de questionnements et de recherches de Marie Thibaud, une mère de famille concernée par ce drame du cancer de l'enfant. Elle sera d'ailleurs considérée comme la lanceuse d'alerte ayant permis la médiatisation et la publicisation de ce problème sanitaire majeur dans la région. En décembre 2015, Alban, son petit garçon alors âgé de 4 ans, déclare une leucémie. C'est en se rendant au service d'oncologie pédiatrique du Centre Hospitalier Universitaire (CHU) de Nantes que M. Thibaud remarque tout d'abord une petite fille qu'elle connaît.

« Je m'interroge, mais préfère certainement croire encore au hasard. Puis très rapidement j'apprends qu'un troisième, un quatrième enfant... du secteur de Sainte-Pazanne, viennent de déclarer également un cancer. Le mot hasard n'est désormais plus acceptable. Quatre enfants en 1 an sur le même secteur ! » (M. Thibaud, site internet du collectif Stop aux Cancers de nos Enfants)

En décembre 2016, la mère de famille finira par obtenir les coordonnées de l'Agence Régionale de Santé (ARS) et de la Cellule régionale de Santé publique France (SPF) via la responsable du registre des cancers pédiatriques de Loire-Atlantique et de Vendée. Elle entreprend alors un travail d'identification des familles concernées aux alentours de Sainte-Pazanne. Les coordonnées de ces familles sont transmises avec leur accord en avril 2017 à la Cellule de Veille et d'Alerte et de Gestion Sanitaire (CVAGS) de l'ARS qui saisit alors Santé Publique France (SPF) pour déclarer un « cluster ». Si cette alerte sonne comme une première victoire pour la mère de famille, le retard pris par les services de l'État dans l'évaluation de la situation et le manque d'information sur l'enquête censée être mise en place par SPF contribueront à faire émerger un sentiment de défiance. En outre, pendant cette période, de nouveaux cas de cancers sont déclarés suscitant une montée de l'inquiétude chez les parents concernés, mais aussi au sein de la population en général. Car l'émergence d'un cluster, qui correspond, dans ce cas précis, à la présence de plusieurs cas de pathologies cancéreuses sur une zone géographique donnée, contribue à accroître l'inquiétude du public résidant sur le territoire concerné, et ce, d'autant plus lorsque la source de ces pathologies reste inconnue. Ce mal invisible susceptible de frapper à la porte de toute famille se trouvant dans cette zone géographique est craint dans la mesure où il s'accompagne d'une carence informationnelle quant à ses causes. En d'autres termes, puisqu'on ne connaît pas la cause exacte de ces cancers, alors toute cause potentielle peut être source d'inquiétude. Ce sont les informations récoltées par M. Thibaud qui contribueront à informer les services de l'État de tout nouveau cas de cancer pédiatrique, le registre des cancers de Loire-Atlantique et de Vendée connaissant alors un retard de quatre années dans l'enregistrement des pathologies déclarées selon la responsable de ce registre, Jacqueline Clavel (Le Point, 11 mars 2020). Dans ce sens, « l'expertise expérientielle » de la mère de famille contribue à mettre à jour l'information de ces services de l'État par sa « compétence à mobiliser le savoir expérientiel pour apporter des réponses à des problèmes formulés en termes généraux ou spécifiques » (Akrich et Rabeharisoa, 2012, p.70). Les données issues de ce registre sont précieuses sur le plan du suivi sanitaire de la situation régionale. Mais la production de ces données reste très limitée dans la mesure où la participation à ce recensement incombe aux oncologues, bien qu'aucune obligation réglementaire ne leur impose (Thibaud et Kayser, 2020, p.30).

« [...], je les contacte tous les 2, 3 mois pour leur demander où en est leur enquête. Rien n'avance jamais. Aucune famille, nous y compris, ne seront jamais contactés par l'ARS ou SPF.

Les enfants ont continué de déclarer des cancers sur cette année 2017 et le décès d'une princesse vient se rajouter... Je les en informe à chaque nouveau cancer. Sans mes alertes,

ils n'ont rien. Je doute alors de leur engagement et de leur implication, mais je ne peux faire davantage. [...] » (M. Thibaud, site internet du collectif Stop aux Cancers de nos Enfants)

Ce que John Dewey appelle l'enquête sociale - et que nous appelons dans cet article l'enquête citoyenne - joue ici un rôle-clé dans la production d'informations puisque ce sont bien les recherches et les actions entreprises par cette mère de famille qui ont contribué au changement d'échelle de ce problème sanitaire : les cas individuels de cancers pédiatriques par leur mutualisation ont fait émerger une inquiétude mais aussi une réflexion citoyenne autour de la potentialité d'un cluster. L'enquête en faveur des données sanitaires devient alors *« un levier puissant pour mieux coupler formes de vie collective et évolutions des milieux de vie, car elle est le lieu où politique et apprentissage s'expriment dans la mise au jour de nouvelles potentialités d'actions individuelles et collectives »* (Hervé, 2023, § 10).

Selon le modèle d'Ackoff, les connaissances sont issues d'une agrégation d'informations qui sont elles-mêmes le résultat d'un ensemble de données (Ackoff, 1989). Sans données, il est ainsi impossible de créer des connaissances. En outre, comme le montrent de nombreux travaux, la diplomatie des données s'est imposée comme un véritable outil d'aide à la décision (Warin *et al.*, 2023), ce qui implique que leur présence comme leur absence peut être mobilisée stratégiquement. Les travaux de Prete ont d'ailleurs montré de quelles manières la production de données sur les effets sanitaires des pesticides était conditionnée par des facteurs politiques (Prete, 2025). Et cet enjeu de production de données va rapidement se poser, puisque l'objectif de cette mère de famille était de comprendre pour quelles raisons on recense autant de cancers pédiatriques dans la commune de Sainte-Pazanne et ses alentours. Cette question se posera d'autant plus pour M. Thibaud qu'en 2018, les services de l'État lui ont rendu des conclusions qu'elle a jugé peu satisfaisantes. A la suite de son signalement, l'ARS comme SPF n'ont en réalité pas produit d'enquête épidémiologique *stricto sensu* mais se sont appuyés sur des données existantes issues de l'*open data* pour lister d'éventuelles causes à ces cancers pédiatriques. Comme l'ont montré les travaux de Derangeon, cette réponse institutionnelle a contribué à *« une forte controverse médiatique et la colère des familles impactées »* dans la mesure où *« la transparence et la méthode d'investigation utilisée ont été remises en cause par les familles et questionnées par la presse locale et nationale »* (Derangeon, 2024, p.192). En effet, comme nous l'a confié M. Thibaud dans un entretien, aucune famille n'avait été reçue dans le cadre de cette première enquête institutionnelle. Elle la qualifiera d'ailleurs d'*« administrative »* dans la mesure où seules des bases de données publiques ont été mobilisées, comme Basol (Base de données des sites et SOLs pollués), Basias (Base de données des anciens sites industriels et activités de services) ainsi que la littérature scientifique issue du Centre International de recherche sur le cancer (CIRC) de l'Organisation Mondiale de la Santé (OMS). En outre, les conclusions de cette enquête pointent plusieurs facteurs de risques possibles sans réellement analyser ceux qui pourraient être les plus probables, laissant entrevoir les contours d'une *« stratégie de globalisation des facteurs de risques »* (Hadna, 2025). Cette stratégie consiste à regrouper ensemble plusieurs sources possibles de pollutions responsables d'une atteinte environnementale ou sanitaire, de manière à ne pas cibler l'une d'entre elles qui pourrait être susceptible de remettre en cause un enjeu jugé majeur par les décideurs, touchant par exemple à la démographie, à l'économie ou à la politique. De cette manière-là, la controverse est maintenue artificiellement, puisque le doute profite aux potentiels pollueurs. Cette situation confortera M. Thibaud dans ces premières impressions : une récolte de données efficace et répondant aux enjeux défendus par les familles concernées ne saura dépendre uniquement des services de l'État.

La création du collectif « Stop aux Cancers de nos Enfants » comme institutionnalisation des données citoyennes

Plusieurs demandes adressées par la mère de famille aux services de l'État sont restées sans réponse. Afin de mieux informer les familles résidant à Sainte-Pazanne et dans les communes situées aux alentours, une cartographie des cas de cancers avait été demandée, ainsi qu'un suivi régulier des nouveaux cas, mais ces demandes se sont heurtées à des refus catégoriques (Thibaud et Kayser, 2020, p.30). Face au refus de SPF et de l'ARS, M. Thibaud a donc monté l'association « Stop aux Cancers de nos Enfants » en mars 2019. Porté par l'inquiétude des parents confrontés à la multiplication des cas de cancers pédiatriques observés entre 2015 et 2019, ce collectif a été mis en place pour contribuer à la production de données qui restait alors encore très maigre du côté des services de l'État. L'inquiétude et l'incompréhension seront d'autant plus grandes qu'en 2019, SPF a officiellement reconnu la présence d'un cluster, sans toutefois agir en conséquence puisqu'aucune étude épidémiologique telle qu'attendue par les familles et le collectif ne sera lancée. Car l'ARS a bien saisi SPF pour mener une enquête épidémiologique. Mais, comme l'a révélé M. Thibaud dans un entretien, cette enquête a pris la forme d'un questionnaire, ce à quoi ne s'attendaient pas les parents concernés. Ce questionnaire mis en place par SPF cherchait notamment « à identifier un ou des facteurs de risque, qui seraient communs aux enfants victimes de cancer localement et que l'on ne retrouverait pas ailleurs ou de façon générique » (SPF, Questions-réponses des cancers pédiatriques à Sainte-Pazanne, note en ligne du 10 octobre 2019). Mais le format de cette enquête a soulevé de nombreuses questions, notamment concernant la méthodologie, ce dont a témoigné une des mères de famille concernée dans un journal local.

« Cette enquête épidémiologique est réalisée via un questionnaire envoyé aux familles sur leurs modes de vie. Une démarche qu'Aurélie, la maman de Juliette, devenue membre du collectif, estime faussée. »

"Ma fille, lorsqu'elle est décédée, ne vivait dans notre maison que depuis 6 mois. Il nous semblait beaucoup plus pertinent d'aller dans la maison où elle avait vécu pendant 6 ans. On nous a dit : "non l'étude porte sur votre domicile actuel, c'est ainsi. Il fallait que cela rentre dans les cases, même si ça n'avait strictement aucun sens". » (France3 Pays de la Loire, 10 mars 2021, en ligne).

Ce que regrette le collectif, c'est que l'enquête épidémiologique reposait moins sur des considérations scientifiques qu'administratives, ce que montre ce témoignage qui présente une « critérisation » discutable concernant la résidence des enfants concernés par l'enquête. Si l'on ajoute à cela le coût que représenterait pour l'État une enquête épidémiologique plus poussée dans un contexte de forte restriction budgétaire, ainsi que les rapports de force susceptibles d'émerger localement eu égard aux sources des pollutions évoquées par ces parents (agricoles et industrielles, notamment), cette démarche d'enquête épidémiologique institutionnelle rencontrait de fait un certain nombre d'obstacles. En outre, elle a finalement conclu à l'absence de cluster en 2020. Annie-Thébaud-Mony, sociologue et directrice de recherches honoraire à l'Institut national de la santé et de la recherche médicale (Inserm), Laurence Huc, toxicologue à l'Institut national de recherche pour l'agriculture, l'alimentation et l'environnement (Inrae), Patrick Dubreil, médecin généraliste à Nantes, et M. Thibaud ont conclu en consultant les rapports de l'ARS et de SPF que « la recherche des causes est orientée, moins vers des facteurs de risques collectifs que vers des comportements individuels (tabac, alcool, usage domestique des pesticides contre les moustiques et les puces) », ce qui selon eux « contribue à culpabiliser les familles » (Thébaud-Mony, et al., 2021). Les auteurs pointent aussi des limites méthodologiques dans l'enquête épidémiologique officielle, telle que l'absence de prise en compte de l'effet cocktail, ou encore le fait que la première phase de l'enquête menée en 2019 par interviews n'ait pas été menée « auprès des seuls parents concernés en visant à reconstituer de façon approfondie l'histoire personnelle de chaque enfant, en lien avec l'histoire familiale, professionnelle et résidentielle de sa famille », mais également « auprès de certains

parents d'enfants malades, de personnalités diverses, d'élus, d'enseignants, de professionnels de santé et autres » en portant « sur "le vécu" et "les attentes" » (Thébaud-Mony, et al., 2021). Très vite, les enjeux du collectif SCE vont largement dépasser la « simple » mission d'alerte pour finalement investir les membres du collectif dans une enquête citoyenne de plus grande ampleur.

« Le collectif « Stop aux Cancers de nos Enfants » dénonce cet arrêt brutal sans aucune investigation ni prélèvement de Santé Publique France pour venir chercher les causes de ces cancers pédiatriques.

Cette recherche des causes environnementales de ces cancers est inexistante en France. Ni chronicité ni effet cocktail (addition de plusieurs éléments de l'environnement) ne sont étudiés aujourd'hui.

Quand les autorités sanitaires montrent leurs limites, c'est donc aux citoyens de prendre le relais et de faire bouger les lignes » (site internet du collectif « Stop aux Cancers de nos Enfants »)

Nous distinguons ainsi trois principales missions du collectif : la diffusion de données et d'informations en santé environnementale, l'information du public et l'alerte des services de l'État. Le collectif contribue également à propager une forme de critique de l'expertise institutionnelle qui vise à souligner ses limites tout en rappelant la gravité de la situation, et les cas de pathologies, voire de décès qui continuent d'être déclarés. Cette communication a pour but de montrer l'écart existant entre la situation jugée grave et urgente, et ce qui est vécu par ces citoyens comme une forme d'« inaction publique » (Henry, 2017). Elle vise en outre à souligner que l'absence de données constitue par ricochet une absence d'actions en faveur de ces enfants malades et de prévention des risques susceptibles d'engendrer de nouveaux cas. Ainsi, dans un communiqué publié en 2019, le collectif liste plusieurs pistes évoquées par les services de l'État tout en indiquant la manière dont ces pistes ont finalement été rapidement relativisées, justifiant ainsi l'absence d'actions pour les étudier. Le communiqué est intitulé « Circulez, il n'y a rien à voir ! » (« Stop aux Cancers de nos Enfants », communiqué de presse, septembre 2019). Le gaz radon, les champs électromagnétiques ou encore des insecticides interdits depuis plusieurs années, comme la dieldrine et le lindane - et pourtant retrouvés dans l'air d'une école située à Sainte-Pazanne - sont autant d'éléments dont la présence a été reconnue par les services de l'État, mais dont la portée a été rapidement relativisée. Dans une précédente recherche portant sur les pollutions issues des anciennes mines d'uranium, nous mettions en avant cette forme d'argumentation que nous dénommons la « stratégie de relativisation du risque », et qui consiste justement à reconnaître la présence d'un risque tout en relativisant ses conséquences (Hadna, 2018). Elle permet de ne pas adopter une posture complètement sceptique en reconnaissant la présence d'un élément potentiellement risqué mais en modérant fortement l'idée qu'il puisse avoir des conséquences. En outre, cette posture permet d'éviter la production de données et d'études sur le risque concerné.

Cette montée en compétences et le processus d'institutionnalisation de ce mouvement par la création d'un collectif lui ont permis de gagner en légitimité en se positionnant comme un nouvel acteur dans la gouvernance des données à l'échelle régionale. Car, bien que cette institutionnalisation ne soit pas parvenue à offrir au collectif la capacité d'entrer dans les scènes décisionnelles, elle lui a en revanche donné la possibilité d'imposer la santé environnementale et la production des données épidémiologiques comme des enjeux centraux localement. En effet, ce mouvement citoyen a contribué à publiciser le problème des cancers pédiatriques par une démarche d'enquête citoyenne. Comme le souligne Daniel Céfaï citant John Dewey, tout part d'une « situation problématique » (Dewey, 1938) :

« L'organisme, comme être vivant (life form), doit s'engager dans une enquête pour déterminer quel est le problème et tenter de le résoudre en transformant son milieu de vie et les relations qu'il entretient avec lui. Il en va de même quand les types de réponses, routinisés et standardisés qu'une collectivité donne à son environnement s'avèrent mal

ajustés, inadéquats ou insuffisants : un trouble naît de l'indétermination de la situation que les membres d'une collectivité devront circonscrire, contenir, comprendre, contrôler » (Céfaï, 2016, p. 27).

Ainsi, la production de données intervient comme un objectif central dans la publicisation des problèmes publics marqués par l'ignorance. Dans le même temps et comme le souligne Emmanuel Henry, si les cancers et autres pathologies imputables à des facteurs environnementaux ne parviennent que rarement à devenir des problèmes publics, c'est bien parce que l'intérêt scientifique à leur égard et à celui des conséquences sanitaires de l'industrie est limité, au profit de recherches orientées vers des mécanismes d'ordre génétique (Henry, 2021). Cette institutionnalisation du mouvement citoyen sera d'autant plus efficace que SCE parvient depuis sa création à capter l'attention de parlementaires, mais aussi d'acteurs scientifiques issus du monde académique avec qui il communique régulièrement dans le cadre de manifestations scientifiques ou d'événements politiques ce qui contribue également à diffuser ses revendications. En outre, SCE constitue un siège régional central en termes de données sur les cancers pédiatriques.

Tableau Août 2021 de l'ensemble des 25 cancers pédiatriques connus par le Collectif SCE sur le secteur depuis 2015

Type de cancer	Année de déclaration du cancer	Commune au moment de la déclaration du cancer	Situation des enfants début mai 2021	Raisons de la non prise en compte des enfants par SPF		
1- Tumeur solide	3 en 2015	St Mars de Coutais Villeneuve en Retz St Hilaire de Chaleons	8 En rémission 7 Décédés 6 En traitement 3 sans information 1 en guérison	- 15 ans et plus - Communes non retenues - Déménagement hors secteur après la déclaration du cancer - Déclaration du cancer depuis fin 2018		
2- Hémopathie maligne						
3- Hémopathie maligne						
4- Tumeur solide	5 en 2016	Villeneuve en Retz Ste Pazanne St Mars de Coutais				
5- Tumeur solide						
6- Hémopathie maligne						
7- Tumeur solide						
8- Hémopathie maligne	3 en 2017	Ste Pazanne Rouans Port Saint Père				
9- Hémopathie maligne						
10- Hémopathie maligne						
11- Hémopathie maligne	6 en 2018	Cheix en Retz Machecoul Ste Pazanne Villeneuve en Retz Rouans				
12- Tumeur solide						
13- Tumeur solide						
14- Hémopathie maligne						
15- Tumeur solide						
16- Tumeur solide						
17- Hémopathie maligne	4 en 2019	Ste Pazanne St Mars de Coutais Machecoul Le Pellerin/Rouans				
18- Hémopathie maligne						
19- Tumeur solide						
20- Tumeur Solide						
21- Tumeur solide	3 en 2020	St Mars de Coutais Machecoul				
22- Tumeur solide						
23- Tumeur solide						
24-Hémopathie maligne	1 en 2021	Ste Pazanne				
25-Hémopathie maligne						
En rouge ceux ne rentrant pas dans les conclusions de l'étude de SPF de novembre 2019 , en violet ceux non comptabilisés dans le taux d'incidence de SPF (SIR). Le Collectif comptabilise 25 enfants. SPF retient 13 enfants pour l'étude du questionnaire et n'en retient ensuite que 11 sur ces 13 pour le calcul du taux d'incidence (SIR) en 2019. En septembre 2020, SPF annonce que malgré le nombre croissant de cancers pédiatriques sur ce secteur, ils ne retiennent plus le terme de cluster. Le registre départemental des cancers refuse depuis de transmettre les données des cancers pédiatriques par commune sur le département de Loire-Atlantique.						
Tumeurs solides: ostéosarcomes, tumeurs cérébrales ...						
Hémopathies malignes: leucémies, lymphomes ...						

Tableau 1. Tableau récapitulant les 25 cas de cancers pédiatriques recensés entre 2015 et 2021 (Stop aux Cancers de nos Enfants)

Le collectif a en effet entrepris un important travail de recensement des cas de cancers que l'on ne retrouve pas dans la littérature grise produite par SPF ou encore par l'ARS. Sa proximité avec les résidents des communes du Pays de Retz, et la confiance qu'il a pu acquérir auprès des familles concernées en font un acteur majeur dans la lutte contre les cancers pédiatriques et pour l'information du public. Cette montée en compétences s'est accompagnée d'une réflexion sur le rôle que pouvait jouer ce mouvement citoyen dans la production des données manquantes, plus particulièrement sur le plan épidémiologique. Car

l'un des problèmes bien identifiés dans le champ de la santé environnementale reste l'absence de données épidémiologiques qui représente souvent un frein pour l'étude de la répartition des pathologies et la compréhension des inégalités environnementales (Fosse, *et al.*, 2022, p.16). Comme nous avons pu l'observer dans notre recherche portant sur les enjeux autour de la lutte contre la pollution de l'air à Saint-Nazaire (Loire-Atlantique), cette lutte pour le développement des données épidémiologiques est susceptible de se heurter à des enjeux financiers et politiques majeurs, la pollution pouvant être liée à des activités industrielles (Hadna, 2025). En l'occurrence, cette controverse peut être vécue localement comme une menace potentielle pour l'activité agricole, l'une des plus importantes sources de l'économie locale à Sainte-Pazanne.

LES DONNÉES ÉPIDÉMIOLOGIQUES COMME LEVIER MAJEUR DANS LES CONTROVERSES EN SANTÉ ENVIRONNEMENTALE

Les données épidémiologiques comme levier dans le rapport de forces pour la reconnaissance d'un cluster

Pour le soutenir dans sa démarche de production de données épidémiologiques, le collectif SCE a mis en place l'ICRePSE. Sur le modèle de l'Institut Ecocitoyen pour la Connaissance des Pollutions (IECP) né à Fos-sur-Mer en 2010, l'ICRePSE a été mis en place pour combler le manque de données en santé environnementale, notamment causé par le désengagement de SPF. Car, en 2022, l'agence publique a mis fin au débat concernant l'existence d'un cluster : selon elle, « *la situation locale est considérée comme globalement stable d'un point de vue épidémiologique par rapport aux conclusions de l'investigation de 2019-2020* » (Santé Publique France, 17 novembre 2022). La mission de l'ICRePSE vise donc à assurer une forme de scientification de l'argumentaire du collectif SCE en tentant de produire les recherches manquantes. Porté par deux spécialistes en toxicologie, Solenn Le Bruchec, directrice, et Laurence Huc, présidente, l'institut s'appuie sur les trois principes suivants :

- Étudier les sources de pollution et mettre fin à l'effet silo dans les études épidémiologiques afin de mieux comprendre l'effet cocktail de ces pollutions ;
- Croiser les données environnementales avec les données sanitaires ;
- Impliquer les citoyens pour qu'ils continuent de contribuer aux alertes et aux recherches (source : ICRePSE).

À contre-courant de l'approche adoptée par les services de l'État, l'ICRePSE vise donc à poursuivre les investigations afin de mettre en lumière les causes qui pourraient expliquer ce qu'il reconnaît comme un cluster de cancers pédiatriques. Ces investigations prennent notamment la forme de recueils de récits de vie des familles concernées, d'analyses « *des facteurs de risque associés aux cancers pédiatriques sur les 9 communes du Sud Loire, 7 sur le territoire du Pays de Retz et 2 de Retz Sud Atlantique* », ou encore de productions de « *données d'imprégnation des populations vulnérables (enfants) du territoire par des polluants environnementaux ciblés* » (source : ICRePSE). Ces méthodes rompent avec celles mobilisées par les acteurs institutionnels dans la mesure où l'ICRePSE s'appuie notamment sur le concept d'exposome qui impose « *la nécessité d'une étude en vie réelle de la polyexposition, combinant les différentes familles de polluants (pesticides, métaux, polluants atmosphériques...) et intégrant leurs effets cumulés ou synergiques. L'enjeu est de dépasser l'approche monofacteur pour intégrer la complexité réelle des expositions tout au long des années de vie de la personne exposée* » (source : ICRePSE). Cette démarche, lorsqu'elle concerne un public mineur, engage de la part du collectif certaines démarches et accords réglementaires, puisqu'il s'agit de données très sensibles, et d'études dites de recherche impliquant la personne humaine (RIPH). Parmi les pistes investiguées, la cause possible d'un effet cocktail qui « *désigne le fait qu'associées, deux molécules (ou plus) présentent un effet toxique, là où elles sont inoffensives ou avec un effet moindre prises*

individuellement » (Gillot, 2017, p.26). Car l'une des raisons pour lesquelles SPF n'a pas souhaité poursuivre les investigations tient justement à l'absence d'un seul facteur possible pour toutes les familles enquêtées. Mais si cette observation a suffi à SPF pour abandonner l'idée d'une étude épidémiologique, pour le collectif SCE, et pour l'ICRePSE, il s'agit là d'un problème méthodologique : les outils traditionnels ne permettent pas de mesurer l'effet cocktail. En effet, les méthodes épidémiologiques actuelles ne peuvent évaluer cet effet cocktail, la toxicologie étudiant les substances une à une, et non la « synergie entre substances » (Gillot, 2017, p.27). Et si l'effet cocktail ne peut être mesuré efficacement, cela ne constitue pas un argument suffisamment robuste pour éliminer l'hypothèse des conséquences sanitaires de cet effet cocktail sur la santé des enfants.

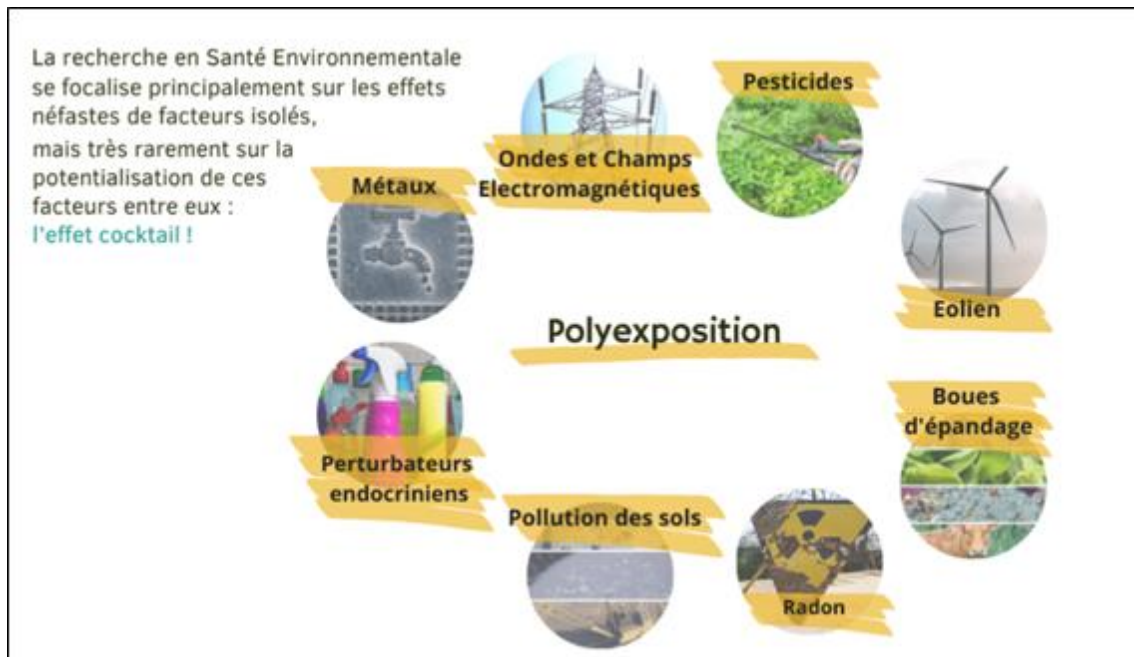


Figure 1. La représentation de l'effet cocktail mis en avant par l'ICRePSE (source : ICRePSE)

C'est donc la voie empruntée par cet institut dont une partie du travail consiste à trouver des financements destinés à la recherche en santé environnementale. Bien intégré dans le paysage académique en étant partenaire dans plusieurs projets de recherche interscientifiques, son travail se heurte toutefois à plusieurs limites. Car si sa création a permis d'engager une approche épidémiologique de la situation locale, il reste néanmoins incontestable que la gestion de la gouvernance des risques demeure sous la responsabilité exclusive des services de l'État.

La circulation des données à l'échelle nationale pour débloquer la situation à l'échelle locale

Dans ce cadre-là, l'ICRePSE a fait le choix de s'adresser à d'autres échelles d'action en diffusant ses données et ses revendications à des acteurs nationaux. Reçus le mercredi 22 mai 2024 au ministère de la Santé et de la Prévention, avec les députés de Loire-Atlantique Sandrine Josso et Yannick Haury, l'institut a notamment porté deux principaux enjeux : l'obtention d'un interlocuteur référent au sein de la Direction générale de la santé afin de faciliter les échanges avec l'échelle nationale et l'inscription à l'ordre du jour de l'Assemblée nationale de l'étude de la proposition de loi visant à mettre en place un registre national des cancers adoptée à l'unanimité au Sénat en juin 2023 (Presse Océan, jeudi 23 mai 2024). Et l'institut obtiendra gain de cause, puisque le texte a été définitivement adopté par l'Assemblée nationale le lundi 23 juin 2025 (source : Sénat.fr). Pour l'institut, la mise en place de ce registre national représente une avancée majeure, puisqu'elle vise à faciliter le développement de données épidémiologiques. Le collectif pourrait ainsi à terme bénéficier

d'une base de données équilibrée sur l'ensemble du territoire pour mieux contribuer à la surveillance de l'incidence du cancer, tout en permettant à l'échelle européenne *via* le Réseau européen des registres du cancer (ENCR) de mieux étudier la mortalité par cancer dans l'Union européenne (source : Sénat.fr). L'adoption de ce texte en faveur d'un registre national a suscité de nombreuses réactions notamment dans le domaine médical qui revendiquait également sa mise en place. Car, face aux craintes relatives au coût que peut représenter le problème de la pollution à l'échelle locale, émerge peu à peu un débat autour de l'évitabilité des pathologies, notamment cancéreuses dont on sait aujourd'hui qu'une grande part s'explique par des facteurs de risques liés à l'environnement. En outre, si la prévention des cancers de l'enfant comme de l'adulte a un coût, elle permet également d'envisager des économies importantes en termes de dépenses publiques de santé. Dans ce cadre-là, un débat à l'échelle nationale reste plus pertinent pour justement dissocier les controverses en santé environnementale de leurs enjeux économiques locaux.

CONCLUSION

À travers la montée en compétences de l'action citoyenne initialement portée par une mère de famille, ce mouvement citoyen a contribué à la médiatisation puis à la publicisation du problème sanitaire des cancers pédiatriques de Sainte-Pazanne.

Nous avons montré dans un premier temps que l'initiative individuelle de la lanceuse d'alerte M. Thibaud a permis d'engager une démarche d'enquête citoyenne dans un contexte d'ignorance a priori technique (Hadna, 2017). La carence informationnelle issue du manque d'implication des services de l'État, mais également de la fragilité du registre départemental de cancers a conduit cette mère de famille à s'engager dans une enquête citoyenne. Si cette enquête lui a permis d'acquérir un statut d'interlocuteur privilégié des services de l'État et notamment de SPF et de l'ARS, elle lui a dans le même temps donné la possibilité de révéler les limites de leur implication dans la gestion de ce problème sanitaire. Cette première étape vers l'institutionnalisation du mouvement citoyen a alors contribué à développer une expertise à travers différentes missions, telles que la production de données relatives au recensement des cas de cancers pédiatriques sur le territoire, l'information du public (et notamment des familles concernées de près ou de loin), ou encore l'alerte des acteurs publics en charge de la gouvernance des risques sanitaires.

Dans un second temps, nous avons souligné que l'enjeu principal de cette controverse locale portait sur la production de données en santé environnementale et plus particulièrement des données épidémiologiques. Les critiques qui ont émergé à la suite des études mises en place par les services de l'État s'appuyant sur des données existantes ont contribué à l'émergence d'une critique de l'expertise institutionnelle. Cette critique a souligné le fait que les facteurs de risques étaient étudiés en silo. Or, la démarche de l'ICRePSE consiste justement à renverser cette démarche d'enquête en proposant de s'intéresser à la problématique de l'effet cocktail. L'institut a alors emprunté deux voies pour tenter de faciliter ce type d'étude : son intégration dans les réseaux académiques afin de créer des collaborations et d'obtenir des financements, et la communication auprès de l'échelon national afin d'obtenir la mise en place d'un registre national des cancers.

Au-delà de l'impact qu'elle a eu sur la configuration des acteurs locaux en imposant une forme de représentation des familles touchées par ces cancers pédiatriques, cette initiative citoyenne a fait émerger une défiance constructive (Hadna, 2025) dans laquelle la critique de l'expertise des services de l'État a donné lieu à un nouveau régime de production des connaissances (Pestre, 2006) où les citoyens contribuent au développement de données épidémiologiques. Cette co-production des connaissances où les données issues de l'expérience des citoyens s'articulent aux méthodes de la toxicologie et de l'épidémiologie de l'ICRePSE vise à creuser la piste de l'effet cocktail. Ce dernier, s'il reste pour les services de l'État un obstacle majeur et un argument pour affirmer l'absence d'une cause unique à

ces cancers, reste pour l'ICRePSE le point de départ d'études et d'enquêtes citoyennes pour une meilleure connaissance des polluants et de leurs interactions.

NOTES

¹ Financé par le dispositif Toulouse Initiative for Research's Impact on Society (2024-2027).

RÉFÉRENCES BIBLIOGRAPHIQUES

Ackoff, Russell Lincoln (1989), « From Data to Wisdom », *Journal of Applied Systems Analysis*, Vol. 16, p. 3-9.

Akrich, Madeleine ; Rabeharisoa, Vololona (2012), « L'expertise profane dans les associations de patients, un outil de démocratie sanitaire », *Santé Publique*, n°24, vol°1, p. 69-74.

Boyd, Andy ; Gatewood Jane ; Thorson Stuart ; Dye Timothy De Ver (2019), "Data Diplomacy", *Science & diplomacy* vol. 8, n°1.

Chateauraynaud Francis ; Antoine Deltour ; Flore Vasseur *et ali.* (Table ronde avec) (2019), « L'alerte, un enjeu démocratique ? », *Esprit*, n°4, p. 57-70.

Cefaï, Daniel (2016), « Publics, problèmes publics, arènes publiques... Que nous apprend le pragmatisme ? », *Questions de communication*, n°30, vol°2, p. 25-64.

Derangeon, Mickaël (2024), « Eau potable et santé : une difficile évaluation des risques sanitaires » (p. 189-203), in Charvet, Ronald ; Mougin, Christian ; Rémy, Elisabeth (coord.), *Sols urbain, environnement et santé : repenser les usages*, Versailles : éditions Quæ.

Dewey John (1938), *Logique. La théorie de l'enquête*, trad. de l'américain par G. Deledalle, Paris : Presses universitaires de France.

Fosse Julien, Salesse Camille & Viennot Mathilde (2022), « Inégalités environnementales et sociales se superposent-elles ? », *France Stratégie, La note d'analyse*, n°122, p. 1-16.

Gillot, Lucie (2017), « [Perturbateurs endocriniens] Une évaluation bien empoisonnante... », *Sesame*, n°1, p. 24-28.

Girel, Mathias (2017), « Ignorance stratégique et post-vérité », *Raison présente*, n°204, vol 4, p. 83-96.

Hadna, Saliha (2017), « The Nuclear Safety Authority in France: A dogma of "Independence" and institutional fragility », *Journal of Innovation Economics & Management*, n°22, p.119-144.

Hadna, Saliha (2018), « Controverse autour des stériles uranifères : de la mise à l'agenda d'un problème public à la remise en cause de l'expertise », *Géocarrefour*, n°92 ; vol.°2.

Hadna, Saliha (2025), « Production d'ignorance en situation de crise : le cas de la catastrophe de Fukushima Daiichi », *Approches Théoriques en Information-Communication (ATIC)*, n°10, vol°1, p. 11-28.

Henry, Emmanuel (2017), *Ignorance scientifique et inaction publique. Les politiques de santé au travail*, Paris : Presses de Sciences Po, 256 p.

Henry, Emmanuel (2021), *La Fabrique des non-problèmes, ou comment éviter que la politique s'en mêle*, Paris : Presses de Sciences Po, 176 p.

Hervé, Nicolas (2023), « La théorie de l'enquête sociale de John Dewey, un cadre conceptuel pour une éducation scientifique au sociopolitique », *Recherches & éducations*, Hors-Série 2.

Lehmans, Anne (2018), « Les réinventions de la démocratie à l'aune de l'ouverture des données : du discours de la participation aux contraintes de la gouvernance », *Les Enjeux de*

l'information et de la communication, n° 19, vol. 2, p. 135-146.

McGoey, Linsey (2012), « The logic of strategic ignorance », *The British journal of sociology*, n° 63, vol° 3, p. 553-576.

Pestre, Dominique (2006), « VI. Régimes de production et de régulation des sciences en société », *Introduction aux Sciences Studies*, Paris : La Découverte, p. 94-107.

Prete, Giovanni (2025), « Pediatric diseases and pollution. Children as sentinels of health and exposure inequalities », communication présentée au Collège de France, *La production sociale des inégalités de santé*, Collège de France, Paris.

Proctor, Robert Neel (2011), *Golden Holocaust: Origins of the Cigarette Catastrophe and the Case for Abolition*, Berkeley : University of California Press.

Rosanvallon, Pierre (2006), *La contre-démocratie, la politique à l'âge de la défiance*, Paris : Seuil.

Thébaud-Mony, Annie ; Huc, Laurence ; Dubreil, Patrick ; Thibaud, Marie (2021), « Agir contre les cancers de nos enfants et leur mise en danger », *Pratiques. Cahiers de la médecine utopique*.

Thibaud, Marie ; Kayser, Marie (2020), « Stop aux cancers de nos enfants », *Pratiques, Cahiers de la médecine utopique*, n° 88.

Warin, Thierry ; De Marcellis-Warin, Nathalie ; Elimam, Sarah ; Panot, Molivann ; Schneider, Jérémy (2023), « La diplomatie à l'heure de la science des données : réflexions stratégiques et perspectives », *CIRANO Project Reports*, 2023rp-29, CIRANO.

SOURCE

Poirier, Juliette (2021), « Cancers pédiatriques du Pays de Retz en Loire-Atlantique : retour sur six années de combat », France3 Pays de la Loire, édition du 12 mars 2021. En ligne : <https://france3-regions.franceinfo.fr/pays-de-la-loire/loire-atlantique/nantes/cancers-pediatriques-du-pays-de-retz-retour-sur-six-annees-de-combat-1990363.html>

Quelle gouvernance pour les données de recherche ? Diplomatie et enjeux géopolitiques

Article inédit, mis en ligne le 23 septembre 2026

Luis Roman ARCINIEGA GIL

Docteur en droit public, chercheur associé au Centre de recherches administratives, politiques et sociales (CERAPS), UMR 8026, CNRS. Thèmes de recherche : données, vie privée et intelligence artificielle. Enseignant et consultant en droit et politiques de l'innovation. luisroman.arciniegal@univ-lille.fr

Manuel MUNIER

Maître de conférences en informatique. Thèmes de recherche : sécurité de l'information, gestion des risques. Responsable de l'équipe ISI-Know (Ingénierie et Sécurité de l'Information par la gestion des connaissances) du Laboratoire d'Informatique de l'Université de Pau et des Pays de l'Adour (LIUPPA) EA 3000. manuel.munier@univ-pau.fr

Plan de l'article

Résumé et mots clés en français

Title, abstract and keywords in English

Título, resumen y palabras clave en español

Introduction

Gouvernance des données

Dimension géopolitique : le cas de l'Union européenne

Une proposition de modèle et sa critique : la métaphore bio-inspirée ou gouverner comme un écosystème

Conclusion

Références bibliographiques

RÉSUMÉ

Dans un contexte de numérisation généralisée, les données scientifiques sont devenues une ressource stratégique se situant à la croisée de la libre circulation des savoirs, de la souveraineté et de la justice sociale. Cet article analyse les conditions d'une gouvernance équitable des *data commons* scientifiques en s'appuyant sur la théorie des communs, l'analyse comportementale du droit et la diplomatie normative. En ce sens, l'Union européenne apparaît comme une puissance régulatrice dont les instruments juridiques incarnent un modèle fondé sur la transparence et sur la protection des droits, mais qui présente également des ambivalences hégémoniques. En réponse, nous proposons un modèle de gouvernance bio-inspiré, centré sur la confiance, la réciprocité et la résilience adaptative, illustré par plusieurs exemples empiriques déjà mis en pratique dans différents contextes, et qui pourraient soutenir une application potentielle dans le domaine des données. Ce modèle ouvre la voie à une diplomatie des données plus inclusive et durable.

Mots clés

Communs numériques, données de la recherche, science ouverte, justice cognitive, gouvernance bio-inspirée.

TITLE

What form of governance for research data? Diplomacy and geopolitical issues

Abstract

In the digital age, scientific data has become a strategic resource situated at the intersection of free knowledge exchange, digital sovereignty, and social justice. This article analyses the conditions for the equitable governance of scientific data commons, drawing on commons theory, behavioural analysis of law, and normative diplomacy. While the European Union's legal instruments appear to embody a model based on transparency and the protection of rights, they also exhibit hegemonic ambivalences. In response, we propose a bio-inspired governance model centred on trust, reciprocity, and adaptive resilience, illustrated by several empirical examples already put into practice in various contexts, which could support a potential application in the field of data. This model paves the way for more inclusive and sustainable data diplomacy.

Keywords

Digital commons, research data, open science, cognitive justice, bio-inspired governance.

TITULO

¿Qué modelo de gobernanza para los datos de investigación? Diplomacia y retos geopolíticos

Resumen

En la era digital, los datos científicos se han convertido en un recurso estratégico que se sitúa en la encrucijada entre la libre circulación del conocimiento, la soberanía digital y la justicia social. Este artículo analiza las condiciones para una gobernanza equitativa de los *data commons* de investigación basándose en la teoría de los bienes comunes, el análisis conductual del derecho y la diplomacia normativa. En este sentido, la Unión Europea se perfila como una potencia reguladora cuyos instrumentos jurídicos encarnan un modelo basado en la transparencia y la protección de los derechos, aunque también presenta ambivalencias hegemónicas. Como respuesta, proponemos un modelo de gobernanza bioinspirado, centrado en la confianza, la reciprocidad y la resiliencia adaptativa, que se ilustra con varios ejemplos empíricos ya puestos en práctica en diferentes contextos, y que podrían respaldar una posible aplicación en el ámbito de los datos. Este modelo allana el camino hacia una diplomacia de datos más inclusiva y sostenible.

Palabras clave

Bienes comunes digitales, datos de investigación, ciencia abierta, justicia cognitiva, gobernanza bioinspirada.

INTRODUCTION

Dans un contexte de numérisation généralisée, les données de la recherche deviennent des ressources stratégiques suscitant des tensions croissantes entre la libre circulation des savoirs, les intérêts économiques, la souveraineté numérique et les impératifs éthiques. Au carrefour de ces dynamiques émergent de nouveaux défis pour la gouvernance de l'information scientifique, dans un contexte marqué par la multiplication des instruments juridiques européens, tels que les règlements sur la protection des données (RGPD), la gouvernance des données (RGD), les données (*Data Act*) ou encore l'intelligence artificielle (*AI Act*), qui incarnent une volonté de régulation éthique et souveraine du numérique (Veale et Zuiderveen Borgesius, 2021). Parallèlement, la montée en puissance des grandes plateformes en tant qu'acteurs quasi-régulateurs et la fragmentation des régulations internationales renforcent la portée de l'« effet Bruxelles » (Bradford, 2012 ; 2020) : l'exportation extraterritoriale des normes européennes par leur puissance de marché.

Cet article interroge les conditions d'une gouvernance équitable, collaborative et durable des données de la recherche, en les considérant comme des « communs numériques » à forte valeur sociale (Hess et Ostrom, 2005 ; Dulong de Rosnay et Stalder, 2020). Il s'inscrit dans une approche interdisciplinaire mobilisant la théorie des communs, l'analyse comportementale du droit et la diplomatie numérique, cette dernière impliquant aussi une géopolitique de la connaissance.

L'hypothèse centrale défendue est que les systèmes de gouvernance des données de la recherche pourraient être plus robustes et durables s'ils intègrent conjointement : des mécanismes de confiance, de réciprocité et d'incitation fondés sur le comportement des acteurs (Ostrom, 1990 ; Sunstein, 2014) ; une régulation transparente et adaptable, compatible avec la complexité des écosystèmes de données ; et une logique de diplomatie normative fondée sur la co-construction et la reconnaissance mutuelle (Dulong de Rosnay et Stalder, 2020).

Pour explorer cette hypothèse, l'article propose tout d'abord un cadre théorique articulant les notions de « communs de données » (*data commons*) (Citton, 2017), d'analyse comportementale du droit (Sunstein, 1997) et de diplomatie des données (Boyd *et al.*, 2019). Il se poursuit par une analyse géopolitique de la position normative de l'Union européenne en tant que puissance régulatrice globale. Enfin, à partir de ces constats, l'article formule une proposition de modèle de gouvernance bio-inspiré, fondé sur des principes systémiques, évolutifs et inclusifs, à l'image du mycélium. L'analyse critique de ce modèle sera appuyée par plusieurs exemples empiriques de gouvernance coopérative existants dans différents contextes, proposant d'ouvrir la porte à une possible applicabilité dans le domaine des données, lui donnant ainsi solidité et ancrage dans la pratique.

GOVERNANCE DES DONNÉES

Les données de la recherche comme « communs numériques »

La notion de gouvernance des données est un point d'entrée essentiel pour analyser les communs numériques scientifiques. Dans la littérature, elle est généralement définie comme l'ensemble des structures, des processus et des mécanismes par lesquels sont distribués les droits de décision, les responsabilités et les modalités de contrôle relatives aux données (Khatri et Brown, 2010 ; Abraham *et al.*, 2019). Elle ne se limite pas à une dimension technique ou juridique, mais englobe des enjeux organisationnels, politiques et épistémiques, liés notamment à la production, à la circulation et à l'interprétation des données (Leonelli, 2016, *in* Currie, 2017 ; Leonelli, 2019).

Dans ce cadre, la gouvernance des données de la recherche peut être appréhendée comme un espace de tensions entre ouverture et contrôle, coopération et compétition, intérêts publics et logiques privées. C'est précisément dans cet espace que s'inscrit notre analyse

des données comme communs numériques.

La notion de « communs numériques » désigne des ressources informationnelles partagées et gouvernées collectivement selon des règles définies par leurs communautés d'utilisateurs. Elles constituent un sous-ensemble des communs dans lequel les ressources (données, informations, culture et savoirs) sont produites et maintenues en ligne afin d'éviter leur appropriation privative et de favoriser un accès élargi, dans certaines conditions institutionnelles (Dulong de Rosnay et Stalder, 2020). Dans cette perspective, Meyer et Hudon (2018) soulignent l'importance de la participation et de la co-construction comme fondements d'une gouvernance durable.

Les données de recherche, qui deviennent, après leur publication, des biens non-rivaux en théorie économique à forte valeur sociale, peuvent ainsi être considérées comme des « communs de la connaissance » (Hess et Ostrom, 2005). Le *Knowledge Commons Framework* proposé par Hess et Ostrom (2005) montre que les technologies numériques, en permettant une duplication quasi illimitée, transforment les conditions du partage. Leur caractère non-rival et leur coût marginal quasi nul (Brousseau, 2021) militent en faveur d'un régime d'accès élargi, renforcé par le mouvement de la science ouverte et les principes FAIR - *Findable, Accessible, Interoperable, Reusable* (Mons et al., 2017), qui exigent transparence et conditions claires de réutilisation.

Ces communs incluent également des articles et des ouvrages scientifiques, dans le cadre du mouvement pour l'accès ouvert (Dulong de Rosnay et Stalder, 2020). Des outils comme le *Data Stewardship Wizard*¹ aident les chercheurs à adopter ces principes dans la pratique (Pergl, 2019). Cependant, cet idéal se heurte à des tensions, comme la conciliation entre ouverture et protection des données personnelles ou stratégiques (Leonelli, 2016, in Currie, 2017), ou encore entre coopération scientifique et compétition privative. En ce sens, la gouvernance des données est toujours marquée par des conflits normatifs et politiques (Boyd et al., 2019 ; Obendiek, 2023).

En effet, la caractérisation des données comme des biens non-rivaux ne doit pas faire oublier les dynamiques sociales et institutionnelles qui limitent leur accès et leur partage. Dans le cadre des pratiques scientifiques, les données peuvent faire l'objet de stratégies de rétention liées à des enjeux de carrière, de compétition ou de valorisation économique (Leonelli, 2016, in Currie, 2017 ; Leonelli, 2016 ; Boukacem-Zeghmouri et al., 2021). Les communs de la connaissance ne constituent donc pas des espaces naturellement coopératifs, mais des configurations sociales traversées par des rapports de pouvoir, des asymétries d'accès et des conflits d'usage.

Le droit européen illustre cette ambivalence : si le RGPD, le RGD et le *Data Act* visent à garantir un accès élargi tout en protégeant la vie privée, les secrets commerciaux et la souveraineté nationale, ces réglementations oscillent entre la promotion des communs scientifiques et la préservation d'intérêts privés. Les données de la recherche apparaissent ainsi comme des communs numériques qui recèlent un potentiel d'exploitation et de partage souvent favorisé par des cadres institutionnels et politiques, mais qui sont au cœur de tensions politiques, économiques et éthiques.

Cette ambivalence invite à envisager leur gouvernance non seulement sous l'angle juridique ou technique, mais aussi comme un enjeu de « justice cognitive » et sociale. Comme le rappelle Piron (2018), la « justice cognitive » repose sur la reconnaissance de la diversité des savoirs et d'un « universalisme inclusif » qui dépasse les normes exclusivement occidentales. La seule existence de cadres juridiques, aussi robustes soient-ils, ne suffit donc pas ; il faut intégrer des perspectives tenant compte des comportements réels des acteurs.

Comment concevoir alors des formes de gouvernance articulant ouverture, équité et soutenabilité, tout en considérant les comportements réels des acteurs avec les rapports de pouvoir qui structurent les écosystèmes scientifiques ? C'est à cette question que répond l'approche interdisciplinaire proposée dans cet article et dans laquelle s'inscrit l'analyse

comportementale du droit présentée dans la section suivante.

L'analyse comportementale du droit comme levier de gouvernance

L'analyse comportementale du droit part du constat que les individus n'agissent pas toujours de manière rationnelle, contrairement aux postulats de l'économie classique. Leurs décisions sont influencées par des biais cognitifs, des normes sociales, des heuristiques ou des incitations symboliques (Kahneman, 2011 ; Sunstein, 2014). Ces résultats, confirmés empiriquement par plusieurs études, amènent à relativiser l'efficacité des régulations prescriptives et à explorer des instruments plus souples, mieux adaptés aux comportements réels.

Les travaux fondateurs de Kahneman et Tversky (1983, 1984) ont montré que les choix humains sont davantage influencés par la perception des gains, des pertes et des cadres décisionnels qu'ils ne le sont par un calcul rationnel. Par ailleurs, Ostrom (2009) a démontré que la coopération émerge souvent dans les dilemmes sociaux lorsque la confiance et la réciprocité sont présentes. Enfin, Meyer et Hudon (2018) défendent une conception humaniste de l'action collective, tandis que Ndayisenga (2024) alerte sur les effets des technologies numériques qui tendent à marginaliser le raisonnement humain.

Les sciences comportementales identifient quatre leviers principaux, alors que de manière non exclusive, plusieurs types d'instruments de régulation peuvent être mobilisés et s'inscrivent dans des configurations hybrides de gouvernance (voir le tableau 1).

Leviers comportementaux	Instruments de régulation
Biais cognitifs : effets d'encadrement, aversion au risque et <i>dark patterns</i> , amplifiés par les algorithmes et justifiant une régulation vigilante (Kahneman et Tversky, 1983 ; 1984 ; De Filippi et Hassan, 2016 ; Veale et Borgesius, 2021).	Incitations économiques : financements conditionnés, reconnaissance académique (citations de jeux de données, valorisation institutionnelle) et autres mécanismes de récompense visant à compenser les coûts perçus du partage et à aligner les intérêts individuels sur les objectifs de la science ouverte (Timko <i>et al.</i> , 2023).
Normes sociales : déterminants comportementaux souvent plus influents que la seule information (Ostrom, 2009 ; De Filippi et Hassan, 2016).	Nudges informationnels : rappels automatisés, visualisation des bénéfices du partage, mise en avant des pratiques majoritaires, afin d'orienter les comportements sans contrainte ; efficacité dépendante du contexte et des normes professionnelles (Ioannou <i>et al.</i> , 2021).
Heuristiques : règles de décision simplifiées, sources d'erreurs systématiques (Kahneman, 2011).	Obligations réglementaires : exigences imposées par les revues ou les financeurs (dépôt des données, politiques de partage), dont l'efficacité dépend de leur mise en œuvre et des mécanismes de contrôle (European Commission, 2018 ; Woods et Pinfield, 2022).
Incitations symboliques : réputation, reconnaissance et visibilité sociale comme moteurs d'adhésion (Taylor, 2017 ; Boyd <i>et al.</i> , 2019).	—

Tableau 1. Principaux leviers comportementaux et instruments de régulation mobilisables pour la gouvernance des données.

Appliquée aux données de la recherche, l'analyse comportementale du droit permet d'éclairer les écarts persistants entre les principes de la science ouverte et les pratiques effectives des chercheurs. En effet, des études mettent en évidence que le partage des données reste limité, malgré des incitations institutionnelles croissantes, en raison de coûts perçus (temps de préparation ou risques de mauvaise utilisation), d'un manque de reconnaissance académique et de logiques concurrentielles (Woods et Pinfield, 2022 ; Rowhani-Farid *et al.*, 2017).

Ces comportements s'expliquent par des biais tels que l'aversion à la perte (crainte de perdre un avantage compétitif), le biais de *statu quo* (préférence pour les pratiques établies) ou encore des problèmes d'alignement des incitations (Woods et Pinfield, 2022 ; Rowhani-Farid *et al.*, 2017). Dès lors, la gouvernance des données ne peut seulement être envisagée en termes de règles formelles, mais doit prendre en compte les déterminants comportementaux des acteurs.

Ces constats invitent à recourir à des instruments de régulation s'inspirant des incitations douces ou « *nudges* » (Sunstein, 2014) et de la co-construction institutionnelle. Sunstein (2014) a montré que de légères incitations, préservant la liberté de choix, produisent souvent plus d'effets que des mesures coercitives. Appliqué à la gouvernance des données, cela suppose de rendre visibles les bénéfices du partage (Mons *et al.*, 2017 ; Brousseau, 2021), de simplifier les interfaces (Jaton et Vinck, 2016 ; RGD, 2022), de valoriser les bonnes pratiques par des normes (AI Act, 2024), d'intégrer les contraintes réelles des chercheurs (Boukacem-Zeghmouri *et al.*, 2021) et de créer des architectures institutionnelles favorisant la confiance et les approches polycentriques (Ostrom, 2010 ; Dulong de Rosnay et Stalder, 2020).

Toutefois, si l'efficacité de ces instruments est plus élevée, elle reste aussi conditionnelle. La littérature scientifique souligne que les *nudges* produisent des effets limités ou ambivalents, notamment lorsqu'ils ne sont pas alignés sur les incitations structurelles ou les normes professionnelles ; de plus, certaines incitations peuvent engendrer des effets pervers, comme une conformité minimale ou une instrumentalisation des dispositifs de partage (Alemanno et Spina, 2014 ; Hacker et Dimitropoulos 2017). Ces limites invitent à considérer la gouvernance des données comme un assemblage complexe d'instruments interconnectés, plutôt que comme un ensemble de solutions comportementales isolées.

La confiance et la participation des acteurs à l'élaboration des règles sont dès lors des conditions essentielles de la durabilité (Ostrom, 2010). La reconnaissance, qu'elle émane des pairs ou d'institutions légitimes, constitue un puissant levier d'adhésion (Boukacem-Zeghmouri *et al.*, 2021 ; Obendiek, 2023). Ces dynamiques se retrouvent dans les initiatives de science ouverte, les codes de conduite du RGD ou les *sandboxes*² réglementaires de l'AI Act.

Dans ce contexte, l'approche comportementale du droit s'inscrit dans une logique de *law by design* où la régulation structure les environnements institutionnels et techniques pour encourager le partage, renforcer l'équité et la justice cognitive (Taylor, 2017 ; Obendiek, 2023). Ces éléments suggèrent que la gouvernance des données repose sur des dynamiques distribuées et adaptatives, dans lesquelles les comportements individuels jouent un rôle structurant. Dans cette optique, il est possible d'envisager des modèles analytiques capables de rendre compte de ces interactions complexes, comme les approches bio-inspirées qui sont présentées dans la troisième partie de cet article. Elle invite notamment à concevoir la gouvernance des données comme une architecture de choix fondée sur la confiance, la coopération et l'engagement, offrant une vision pragmatique et inclusive des communs numériques et ouvrant la voie à la diplomatie des données.

Diplomatie des données : vers une régulation négociée des savoirs ?

La diplomatie des données désigne l'ensemble des pratiques par lesquelles des acteurs publics et privés négocient l'accès, l'utilisation et la circulation des données dans un cadre

transnational (Boyd *et al.*, 2019). Située à l'intersection de la diplomatie scientifique, des politiques numériques et des dynamiques géopolitiques, elle dépasse la coopération interétatique classique. Ressource stratégique, voire « arme » de contrôle des comportements et des représentations sociales (Boyd *et al.*, 2019), la donnée remet en question les cadres diplomatiques traditionnels et alimente une diplomatie publique élargie dans laquelle les frontières géopolitiques et organisationnelles deviennent poreuses.

Trois dimensions structurent les débats et leurs tensions s'expriment dans trois arènes concrètes (voir le tableau 2). L'UE s'affirme en effet comme un acteur central de la diplomatie normative. Grâce au RGPD, au RGD, au *Data Act* et à l'*AI Act*, elle exporte ses normes bien au-delà de ses frontières, incarnant ce que Bradford (2012 ; 2020) appelle « l'effet Bruxelles ». L'UE devient un « hégémon réglementaire » : la taille de son marché contraint les multinationales à adopter ses normes, y compris dans d'autres États (Obendiek, 2023). Cependant, cette posture ambivalente, qui promeut à la fois les droits fondamentaux et la régulation éthique, est parfois perçue comme une imposition unilatérale de normes occidentales, risquant de reproduire des logiques de domination (Taylor, 2017 ; Couldry et Mejias, 2021).

Dimension structurelle	Enjeu principal	Arène de tensions
Souveraineté numérique	Contrôle stratégique des données sensibles (santé, défense, IA) et autonomie numérique de l'UE (Obendiek, 2023)	Interopérabilité et échanges transfrontaliers : négociation des normes internationales de circulation des données (Boyd <i>et al.</i> , 2019 ; AI Act, 2023).
Hégémonies normatives	Diffusion extraterritoriale des cadres européens et américains, critiques de technocolonialisme et recomposition géopolitique avec la montée de la Chine (Bradford, 2012 ; 2020 ; Couldry et Mejias, 2021 ; Obendiek, 2023)	Extraterritorialité des régulations : extension du RGPD influençant les législations nationales (Brésil, Japon, etc.) et de l'AI Act au-delà de l'UE (considérants 21-22, 2024).
Exigences éthiques globales	Transparence, justice et neutralité face aux risques de marchandisation et d'opacité des données (Boyd <i>et al.</i> , 2019)	Localisation des infrastructures : encadrement des transferts internationaux par le RGD et le Data Act afin de protéger les droits fondamentaux et les secrets commerciaux.

Tableau 2. Dimensions structurelles et arènes de tensions de la diplomatie des données.

Face à ces tensions, la diplomatie des données offre une alternative : considérer les données comme des biens communs de la connaissance, à protéger contre les appropriations exclusives ou abusives (Dulong de Rosnay et Stalder, 2020). Leur gouvernance repose sur une participation élargie des autorités publiques, des chercheurs, des acteurs privés et des citoyens (Brousseau, 2021) ; les services d'intermédiation, y compris publics, pouvant garantir un accès équitable, notamment pour les PME et les *start-ups* (RGD, 2022). Sur le plan pratique, les licences *Creative Commons* illustrent également comment des outils contractuels privilégient les « droits d'usage » aux « droits exclusifs » et consolident des communautés de partage (Arciniega Gil, 2022 ; 2023).

Mais loin de constituer des espaces de coopération neutres, les données sont aussi marquées par des asymétries entre les acteurs, liées notamment à l'accès aux ressources techniques, aux capacités d'analyse et aux positions institutionnelles (Arciniega Gil, 2022). L'analyse des données en tant que communs numériques doit se pencher sur les rapports de pouvoir qui structurent leur production, leur circulation et leur utilisation.

Dans le champ scientifique, ces asymétries se traduisent par des inégalités entre institutions, disciplines ou régions du monde, certaines étant mieux dotées pour collecter, stocker et exploiter les données. À une échelle plus large, plusieurs travaux ont mis en évidence des dynamiques d'appropriation et de captation de la valeur des données, parfois analysées en termes de « colonialisme des données », dans lesquelles certains acteurs concentrent les bénéfices issus de ressources produites collectivement (Couldry et Mejias, 2021).

Dans ce contexte, la gouvernance des données ne peut pas être envisagée uniquement comme un problème de coordination ou d'incitation, mais elle doit être appréhendée comme un enjeu politique impliquant des arbitrages entre intérêts divergents, ainsi que des mécanismes de redistribution du pouvoir et de la valeur. Cette prise en compte des rapports de pouvoir invite à concevoir des systèmes de gouvernance qui dépassent les structures centralisées ou harmonieuses pour faire émerger des configurations dynamiques dans lesquelles la coopération et le conflit coexistent. Elle souligne ainsi l'importance de modèles capables de rendre compte de ces interactions distribuées et évolutives, comme l'approche bio-inspirée proposée ci-dessous.

Cette évolution traduit un basculement : la diplomatie ne se limite plus aux États ni aux diplomates professionnels. Tout acteur peut devenir « diplomate des données » en rejoignant des coalitions et des consortiums transnationaux pour peser sur les négociations (Boyd *et al.*, 2019). Cette diplomatie distribuée, fondée sur la transparence, la légitimité et l'adaptation continue, ouvre la voie à une régulation des savoirs plus inclusive et durable, en phase avec l'idéal de justice cognitive (Taylor, 2017 ; Piron, 2018 ; Obendiek, 2023).

DIMENSION GÉOPOLITIQUE : LE CAS DE L'UNION EUROPÉENNE

L'Union européenne comme puissance normative : régulation éthique ou hégémonie réglementaire ?

Depuis l'adoption du RGPD en 2016, puis son entrée en vigueur en 2018, l'Union européenne s'est imposée comme une puissance régulatrice mondiale. Considéré comme la législation la plus stricte en matière de protection des données, ce texte a profondément influencé les pratiques des multinationales et des agences publiques, tout en affirmant les valeurs européennes face aux scandales *Snowden* ou *Cambridge Analytica* (Obendiek, 2023). Les amendes, telles que les 746 millions d'euros infligés à Amazon et les 90 millions d'euros infligés à Google Ireland en 2021, illustrent sa force contraignante. Au-delà de ses frontières, le RGPD témoigne de la capacité de l'UE à produire des effets *de jure* et *de facto* : en 2022, 71 % des pays avaient adopté une législation s'inspirant de ce modèle (Obendiek, 2023). Ce phénomène témoigne de la capacité de l'UE à exporter ses normes par la seule force d'attraction de son marché, incitant les multinationales à harmoniser leurs pratiques à

l'échelle mondiale.

Cet arsenal s'est élargi avec le RGD de 2022, qui promeut la transparence et la confiance, le *Data Act* de 2023, qui harmonise l'accès et le partage équitable des données entre les acteurs publics et privés, et le *AI Act* de 2024, qui encadre une IA « digne de confiance », respectueuse des droits fondamentaux (Veale et Zuiderveen Borgesius, 2021). Ensemble, ces textes traduisent l'ambition de l'Europe de proposer un modèle éthique de gouvernance numérique fondé sur la responsabilité et les droits fondamentaux. Cependant, cette influence normative suscite des ambivalences. L'UE se présente comme un contre-modèle face à l'approche permissive des États-Unis ou à la stratégie centralisée de la Chine, mais elle est également perçue comme hégémonique lorsqu'elle impose ses normes grâce à la puissance de son marché (Bradford, 2012 ; 2020). Cette diplomatie normative unilatérale, qui contraint ses partenaires à adopter des règles complexes, risque d'exacerber les tensions.

Le débat est particulièrement vif dans les relations Nord-Sud. Plusieurs pays dénoncent un manque d'équité, estimant que les normes européennes renforcent les inégalités d'accès à l'innovation, aux infrastructures et à la propriété intellectuelle (Taylor, 2017). Les notions de « colonialisme des données » et de « technocolonialisme » soulignent la domination des puissances capables de valoriser les données produites ailleurs, ce qui reproduit des inégalités structurelles (Couldry et Mejias, 2021 ; Nkoudou, 2023).

Ainsi, la diplomatie des données apparaît comme un domaine marqué par des rapports de force asymétriques : les données deviennent un enjeu stratégique dans lequel l'UE oscille entre la promotion d'une régulation éthique et la tentation d'une hégémonie normative (Obendiek, 2023 ; Boyd *et al.*, 2019).

Les dynamiques de diplomatie des données peuvent alors être interprétées comme une extension, à l'échelle institutionnelle et internationale, des mécanismes de gouvernance distribuée. Elles impliquent des interactions complexes entre des acteurs publics et privés, dans un contexte marqué par des asymétries de pouvoir et des enjeux de souveraineté. Loin de se limiter à une simple coordination technique, la diplomatie des données s'inscrit dans des logiques de compétition et de négociation entre États et organisations concernant le contrôle des infrastructures, des normes et des flux d'informations. Ces dynamiques traduisent des formes contemporaines de pouvoir dans lesquelles la capacité à définir les règles d'interopérabilité et d'accès aux données constitue un levier stratégique majeur.

Dans ce contexte, les initiatives européennes en matière de gouvernance des données, comme les espaces européens de données ou les cadres réglementaires récents, peuvent être considérées comme des tentatives de concilier ouverture, protection et souveraineté numérique. Elles visent à structurer des écosystèmes de données fondés sur des valeurs spécifiques, tout en se positionnant dans un environnement international concurrentiel.

La diplomatie normative appliquée aux données scientifiques : limites, tensions et perspectives transformatrices

Appliquée à la recherche, la diplomatie des données oscille entre ambition coopérative et asymétries de pouvoir. Des initiatives telles que l'*European Open Science Cloud* (EOSC) ou la stratégie européenne de gouvernance des données visent à créer un écosystème interopérable, sécurisé et ouvert (Mons *et al.*, 2017 ; DGA, 2022). Fondés sur les principes FAIR, ces projets insistent sur la transparence et la traçabilité, mais ne signifient pas gratuité : ils exigent surtout clarté et garanties institutionnelles (Mons *et al.*, 2017). Ces avancées demeurent toutefois fragmentées et centrées sur l'Europe, avec une faible intégration nord-sud ou multilatérale (Boyd *et al.*, 2019 ; Couldry et Mejias, 2021). La gouvernance scientifique mondiale est marquée par une asymétrie : les chercheurs du Sud sont marginalisés dans les systèmes de publication dominés par le Nord (Boukacem-Zeghmouri *et al.*, 2021) et sont victimes d'aliénation épistémique, ce qui freine la production de savoirs localement pertinents (Piron, 2018). La diplomatie européenne

apparaît dès lors comme le prolongement d'un cadre technocratique occidental difficilement conciliable avec la pluralité des régimes de savoir.

À l'inverse, une diplomatie véritablement coopérative des *data commons* scientifiques devrait intégrer la justice cognitive et la diversité épistémique (Piron, 2018). La première exige la reconnaissance active de formes de savoir variées ancrées dans des traditions culturelles différentes ; la seconde invite à articuler la science ouverte, la justice des données et la diplomatie transnationale face aux logiques de « technocolonialisme » ou de « colonialisme des données » (Couldry et Mejias, 2021 ; Taylor, 2017).

Dans ce cadre, la stratégie normative de l'UE prolonge l'« effet Bruxelles », en élevant les normes mondiales, mais au prix de tensions juridiques et de conflits de souveraineté (Bradford, 2012 ; Obendiek, 2023). Trois angles morts persistent concernant l'« effet Bruxelles » appliqué à la gouvernance des données (voir le tableau 3).

Angle mort	Enjeu pour la gouvernance des données
Gouvernance distribuée	La co-construction des règles, constitutive des communs numériques, demeure insuffisamment intégrée aux approches normatives centralisées (Dulong de Rosnay et Stalder, 2020).
Innovations institutionnelles locales	Les blockchains, DAO et coopératives numériques montrent le potentiel de dispositifs favorisant transparence, participation et co-régulation (De Filippi et Hassan, 2016 ; Meyer et Hudon, 2018).
Diversité culturelle et juridique	La faible prise en compte des contextes extra-européens limite la légitimité et l'inclusivité des cadres de gouvernance (Couldry et Mejias, 2021 ; Taylor, 2017).

Tableau 3. Les principaux angles morts de l'effet Bruxelles appliqué à la gouvernance des données.

Face à ces limites, des modèles émergents ouvrent des pistes. Les organisations autonomes décentralisées (DAO, d'après son acronyme anglais) appliquées aux communs de données expérimentent une gouvernance algorithmique fondée sur la transparence et la participation (De Filippi et Merk, 2024). Grâce à leur logique adaptative et polycentrique, elles rejoignent les formes de gouvernance bio-inspirée, dans lesquelles les règles institutionnelles pourraient évoluer en fonction des contextes socio-écologiques (Ostrom, 2010). De même, une approche écosystémique fondée sur les capacités (Taylor, 2017) propose de concevoir la justice des données en fonction des libertés réelles des populations marginalisées, encourageant une participation plus large.

La diplomatie normative des données scientifiques ne saurait donc se réduire à l'exportation de normes européennes. Elle doit évoluer vers un modèle coopératif et inclusif dans lequel la reconnaissance des savoirs, l'innovation institutionnelle et la co-construction des normes deviennent des leviers d'une gouvernance durable des *data commons* scientifiques. C'est dans cette optique que le paragraphe propose un modèle bio-inspiré pensé comme une alternative adaptative aux approches descendantes, soutenu par divers exemples empiriques ayant déjà fait leurs preuves dans différents contextes et qui pourraient être envisagés pour une application potentielle dans le domaine des données.

Le modèle bio-inspiré donne la possibilité, notamment, d'éclairer les dynamiques autour des données en mettant en évidence la nature distribuée et évolutive des interactions. À l'image d'un écosystème vivant, la diplomatie des données ne reposerait pas sur un centre de décision unique, mais sur une multiplicité de nœuds interconnectés dans lesquels des règles, des normes et des flux sont négociés en permanence. Cette perspective permet d'aborder conjointement les dimensions coopératives et conflictuelles de la gouvernance internationale des données.

Ces processus révèlent une tension structurante entre, d'une part, la nécessité d'une interopérabilité accrue pour favoriser la circulation des données, d'autre part, la volonté

des acteurs de préserver leur autonomie et leur contrôle. La diplomatie des données apparaît ainsi comme un espace de négociation permanente où s'articulent coopération et rivalité. Elle constitue un terrain privilégié pour observer de quelle manière des formes de gouvernance distribuée, inspirées de logiques bio-systémiques, peuvent être mobilisées pour penser les recompositions contemporaines du pouvoir informationnel.

UNE PROPOSITION DE MODÈLE ET SA CRITIQUE : LA MÉTAPHORE BIO-INSPIRÉE DU GOUVERNER COMME UN ÉCOSYSTÈME

Le recours à la métaphore du mycélium ne relève pas ici d'un simple procédé illustratif, mais d'un outil heuristique visant à modéliser des formes de gouvernance distribuée et adaptative. Les approches bio-inspirées sont en effet mobilisées dans différents champs pour rendre compte de systèmes complexes caractérisés par des interactions locales, une absence de contrôle centralisé et une capacité d'auto-organisation.

Le mycélium, un réseau souterrain de filaments fongiques, constitue une infrastructure vivante qui relie les plantes, les arbres et les micro-organismes. Ses réseaux mycorhiziens, constitués d'hyphes ramifiés, permettent la circulation de nutriments, d'eau et de signaux chimiques entre les espèces (Simard *et al.*, 2012 ; Fricker *et al.*, 2017). Dans cet échange symbiotique, la plante fournit du carbone et reçoit en retour de l'azote, du phosphore et d'autres nutriments (Sheldrake, 2020, *in* Kerridge, 2020). Ce *Wood Wide Web*, comme certains auteurs l'appellent, agit comme une « infrastructure de résilience » capable de distribuer rapidement des ressources et de maintenir la stabilité des communautés (Simard *et al.*, 2012 ; Fricker *et al.*, 2017).

Le mycélium présente des caractéristiques particulièrement pertinentes dans le cadre de notre étude : il repose sur une organisation décentralisée, une circulation fluide des ressources et des informations, ainsi que sur des mécanismes d'adaptation aux conditions locales. Ces propriétés en font un analogue utile pour concevoir des systèmes de gouvernance dans lesquels les décisions émergent de l'interaction entre une pluralité d'acteurs plutôt que d'une autorité unique. Cette métaphore éclaire la gouvernance des communs numériques scientifiques. À l'instar des réseaux fongiques, les systèmes de données pourraient valoriser la diversité, l'interdépendance et l'adaptabilité, conditions de leur soutenabilité. Elle rompt avec les logiques centralisées ou marchandes pour envisager la gouvernance comme un processus coopératif, distribué et auto-régulé.

Les théories des communs appuient cette analogie. Hess et Ostrom (2005) estiment que la gestion durable suppose l'articulation entre communautés d'utilisateurs, systèmes de règles et droits de propriété à plusieurs niveaux. Dulong de Rosnay et Stalder (2020) décrivent les communs numériques comme des institutions sociales holistiques, au-delà du marché et de l'État. Dans ce cadre, la justice cognitive (Piron, 2018) prolonge l'analogie écologique en soulignant que la diversité des savoirs, à l'instar de la biodiversité, garantit la résilience des écosystèmes.

L'Union européenne cherche à traduire ces principes en mécanismes de transparence, de traçabilité et de confiance, notamment avec le RGPD (2018) ou les espaces européens communs de données (*data spaces*). Cependant, la gouvernance des données ne saurait se limiter à un cadre juridique descendant ; elle requiert des formes participatives et adaptatives, s'inspirant des dynamiques locales (Ostrom, 2009 ; 2010). Les innovations institutionnelles confirment cette pertinence. Les DAO, par exemple, incarnent des réseaux distribués dans lesquels la gouvernance repose sur des règles algorithmiques transparentes et participatives (De Filippi et Merk, 2024). De la même manière que les boucles du mycélium assurent robustesse et résilience (Fricker *et al.*, 2017), les mécanismes de rétroaction et de co-construction renforcent la capacité des communs numériques à faire face aux crises et

aux tensions géopolitiques.

Ainsi, le mycélium propose de dépasser la simple métaphore biologique pour inspirer des architectures distribuées, collaboratives et adaptatives. À l'image des réseaux fongiques qui nourrissent la forêt, les communs numériques scientifiques pourraient être gouvernés comme des écosystèmes, dans lesquels diversité, interdépendance et coopération sont les conditions mêmes de la durabilité.

Les cinq piliers du modèle bio-inspiré, illustrations et ancrages empiriques

À partir de la métaphore du mycélium, plusieurs correspondances peuvent être établies : les nœuds du réseau peuvent être assimilés aux acteurs de la gouvernance des données (chercheurs, institutions, plateformes et régulateurs) ; les flux au sein du mycélium renvoient aux circulations de données, d'informations et de normes ; les interactions locales traduisent les mécanismes comportementaux et institutionnels qui orientent les pratiques ; enfin, l'absence de centre unique de contrôle reflète la nature polycentrique de la gouvernance des données contemporaines.

Cette métaphore offre la possibilité de rendre compte des dynamiques émergentes dans lesquelles les comportements individuels, les incitations et les contraintes s'articulent pour produire des formes de coordination plus ou moins stables. Cinq piliers pourraient structurer un modèle de gouvernance des communs de données et traduire les dynamiques écologiques observées dans les réseaux mycorhiziens en principes opérationnels visant à renforcer la transparence, la coopération et la résilience.

Le tableau suivant, qui présente des exemples déjà mis en pratique, n'a pas pour objectif de valider empiriquement le modèle proposé, mais d'illustrer sa portée analytique dans différents contextes. Ils s'inscrivent dans une démarche exploratoire visant à montrer de quelle manière une approche bio-inspirée de la gouvernance des données peut éclairer des configurations concrètes, tout en mettant en évidence ses limites. Chacun d'entre eux illustre la possible applicabilité d'un ou plusieurs des piliers proposés, renforçant l'idée que la gouvernance des données scientifiques pourrait être envisagée comme un écosystème distribué, coopératif et adaptatif (voir le tableau 4).

Il convient toutefois de souligner que, à l'instar des écosystèmes biologiques, un modèle de gouvernance des données bio-inspiré ne saurait être exempt de tensions. Les interactions au sein du réseau donnent lieu, en effet, à des formes de compétition, de domination ou de captation des ressources, reflétant les rapports de pouvoir qui traversent la gouvernance des données. C'est là que le modèle du mycélium permet de penser la coopération et le conflit conjointement, sans réduire ces dynamiques à une logique harmonieuse.

Cela étant, en tant que métaphore, ce modèle présente aussi des limites et ne prétend pas rendre compte de l'ensemble des dimensions juridiques, politiques et économiques de la gouvernance des données. Il tend notamment à sous-représenter le rôle des institutions formelles et des rapports de pouvoir structurés, qui peuvent contraindre ou orienter les dynamiques d'auto-organisation. En ce sens, il doit être compris comme un cadre analytique complémentaire et non comme une description exhaustive.

Pilier bio-inspiré	Principe de gouvernance	Fondement théorique	Illustration empirique
Signaux chimiques	Transparence et traçabilité des flux et des règles d'usage ; prévention de la contamination des données (RGPD, RGD ; Dulong de Rosnay et Stalder, 2020).	Coordination par circulation de l'information dans les réseaux mycéliens (Sheldrake, 2020 ; Fricker <i>et al.</i> , 2017).	OpenAIRE : gouvernance fédérative fondée sur les principes FAIR, transparence, standards communs et architecture polycentrique (Mons <i>et al.</i> , 2017 ; Ostrom, 2009).
Symbiose mutualiste	Incitations douces, réciprocité et confiance pour favoriser le partage des données (Ostrom, 2010 ; Hooft <i>et al.</i> , 2023).	Coopération symbiotique et échanges réciproques de ressources (Simard <i>et al.</i> , 2012).	Système WIR : crédit mutuel, coopération volontaire, résilience et co-régulation distribuée (Stodder, 2009 ; Meyer et Hudon, 2018).
Toile sans centre	Gouvernance distribuée, polycentrisme et subsidiarité (Ostrom, 2009).	Organisation décentralisée assurant robustesse et adaptation locale (Fricker <i>et al.</i> , 2017).	Systèmes d'échange locaux et banques de temps : règles locales, confiance, reconnaissance des contributions et justice cognitive (Piron, 2018 ; Dulong de Rosnay et Stalder, 2020).
Ponts entre espèces	Interopérabilité technique et normative, diplomatie des données, standards ouverts (Boyd <i>et al.</i> , 2019 ; FAIR).	Communication entre réseaux biologiques favorisant les échanges de ressources (Simard <i>et al.</i> , 2012).	DAO et blockchain : contrats intelligents, règles transparentes, communs algorithmiques et interopérabilité (De Filippi et Hassan, 2016 ; De Filippi et Merk, 2024).
Boucles de rétroaction	Résilience, apprentissage collectif et adaptation évolutive face aux crises (Obendiek, 2023).	Reconfiguration des réseaux mycéliens après perturbation (Fricker <i>et al.</i> , 2017).	DAO : rétroaction, auto-organisation et évolution des règles de gouvernance (De Filippi et Merk, 2024).

Tableau 4. Modèle bio-inspiré de gouvernance des données : piliers, principes et illustrations empiriques.

En d'autres termes, le modèle bio-inspiré propose une grille de lecture systémique de la gouvernance des *data commons* scientifiques, dont la valeur tient moins à une validation exhaustive qu'à sa capacité à éclairer et à inspirer les pratiques existantes. Il permet d'articuler les apports de l'analyse comportementale du droit, qui met l'accent sur les déterminants individuels des pratiques, avec une approche systémique de la gouvernance des données, attentive aux interactions et aux structures émergentes. Les exemples empiriques présentés dans le tableau précédant illustrent la manière dont ces principes s'incarnent déjà, de façon partielle mais significative, dans différents contextes avec une application potentielle dans le domaine des données.

En effet, les politiques de science ouverte, portées par des financeurs et des institutions internationales, visent à promouvoir le partage des données de recherche, en même temps que leur mise en œuvre révèle des tensions entre injonctions à l'ouverture et logiques

concurrentielles propres au monde académique.

Les exemples présentés dans cette section suggèrent que le modèle bio-inspiré permet moins de prédire les comportements que de structurer leur interprétation, en mettant en évidence les interactions entre incitations, normes, rapports de pouvoir et outils techniques. Il constitue donc un outil complémentaire utile pour appréhender la complexité des systèmes de gouvernance des données. D'un point de vue comportemental, les tensions se traduisent par des stratégies de rétention ou de partage minimal, liées à l'aversion à la perte et aux incitations institutionnelles. À travers le prisme du modèle du mycélium, ces dynamiques sont à interpréter comme le résultat d'interactions locales entre les acteurs, dans lesquelles les flux de données ne circulent pas de manière uniforme, mais dépendent de conditions contextuelles et de relations de pouvoir.

Cette lecture met en évidence une gouvernance à la fois distribuée et inégalement structurée, dans laquelle certains acteurs occupent des positions centrales leur permettant d'orienter les flux et les normes. Elle souligne également des limites, notamment la difficulté à prendre en compte les contraintes juridiques formelles imposées par les institutions.

CONCLUSION

À l'heure où les données scientifiques sont devenues un levier stratégique de l'innovation, de la gouvernance et de la diplomatie mondiale, cet article propose une analyse critique des cadres normatifs existants et une exploration interdisciplinaire des alternatives possibles. En articulant la théorie des communs numériques, les dynamiques de régulation européenne et la métaphore bio-inspirée du mycélium, nous esquissons un modèle de gouvernance plus inclusif, distribué et résilient des *data commons* scientifiques.

Trois principaux résultats émergent. D'abord, l'Union européenne illustre le paradoxe d'une puissance normative dont l'influence mondiale (ou « effet Bruxelles ») élève les standards tout en générant de nouvelles asymétries, notamment envers le Sud global. Ensuite, la diplomatie des données est marquée par une fragmentation institutionnelle et un déficit d'inclusivité épistémique. Enfin, la métaphore du mycélium, inspirée du monde vivant, offre un cadre heuristique pour penser la gouvernance des données comme un écosystème fondé sur l'interdépendance, la coopération et la résilience.

Le modèle bio-inspiré proposé ne constitue pas une prescription normative, mais un cadre analytique destiné à éclairer les interactions entre les acteurs, les normes et les infrastructures. Les exemples empiriques montrent que plusieurs de ses principes trouvent déjà une traduction partielle dans des dispositifs existants, sans pour autant valider le modèle dans son ensemble.

Ces résultats invitent les décideurs publics, les institutions scientifiques et les communautés de recherche à promouvoir des formes de gouvernance plus distribuées, transparentes et coopératives, conciliant ouverture des données, protection des droits fondamentaux et diversité des contextes.

La proposition reste toutefois limitée par le caractère nécessairement partiel de la métaphore bio-inspirée, dont la transposition aux institutions humaines dépend des contextes politiques, juridiques et techniques. Les exemples mobilisés illustrent le modèle sans pour autant le valider de manière systématique.

Ces limites ouvrent autant de pistes de recherche futures et tracent un agenda pour une gouvernance bio-inspirée des communs de données scientifiques (voir tableau 5).

Axe de recherche	Objectif	Exemples de travaux
Validation empirique	Tester le modèle	Études de cas participatives sur le partage des données scientifiques
Comparaison transrégionale	Évaluer la transférabilité	Réception des normes européennes hors d'Europe (Amérique latine–Afrique–Asie)
Innovation institutionnelle	Opérationnaliser les principes bio-inspirés	DAO, plateformes coopératives, agences de données
Interdisciplinarité	Enrichir le cadre théorique	Dialogue entre droit, écologie, sciences comportementales et diplomatie

Tableau 5. Agenda de recherche pour une gouvernance bio-inspirée des communs de données scientifiques.

En définitive, considérer les données scientifiques à travers la métaphore du mycélium, c'est affirmer qu'elles ne sont pas seulement un objet de régulation technique ou juridique, mais un espace objet d'une vie collective, dans lequel l'interconnexion, la diversité et la coopération conditionnent la durabilité. La gouvernance des données devient alors une écologie politique : un art d'organiser la circulation des savoirs dans un monde interdépendant, incertain et fragmenté.

NOTES

¹ Reconnue dans le monde entier par les pionniers de la gestion des données et les instituts de recherche internationaux, cette plateforme *open source* est dédiée aux plans de gestion collaborative et évolutive des données. Il s'agit d'un outil en ligne qui permet de créer des plans de gestion des données (DMP) et qui aide les chercheurs à comprendre les exigences en matière de gestion des données selon le principe FAIR, afin qu'ils puissent élaborer leurs propres DMP. L'outil est recommandé dans le guide du programme Horizon Europe. Cf. <https://ds-wizard.org/>

² Bien qu'il n'existe pas de définition commune, les *sandboxes* ou « bacs à sable réglementaires » désignent généralement des dispositifs permettant aux entreprises de tester et d'expérimenter de nouveaux produits, services ou activités innovants sous la supervision d'une autorité de régulation, et ce, pendant une période limitée. Ils améliorent la sécurité juridique, favorisent la conformité, permettent le traitement des données personnelles et facilitent l'accès au marché pour les PME et les *start-up*. Cf. Art. 57 de l'*AI Act*.

RÉFÉRENCES BIBLIOGRAPHIQUES

Abraham, Rene ; Schneider, Johannes ; vom Brocke, Jan (2019), « Data governance : A conceptual framework, structured review, and research agenda », *International Journal of Information Management*, vol. 49, p. 424-438.

Alemanno, Alberto ; Spina, Alessandro (2014), « Nudging legally: On the checks and balances of behavioral regulation », *International Journal of Constitutional Law*, vol. 12, no. 2, p. 429-456. <https://academic.oup.com/icon/article/12/2/429/710410>

Arciniega Gil, Luis Roman (2022), *La gouvernance des données en tant que « biens communs » : étude juridico-stratégique pour une valorisation publique des données informationnelles*, thèse de doctorat en Droit public, Université de Lille, Français.

Arciniega Gil, Luis Roman (2023), « Towards a Shared Governance of Data as Digital Commons? », *University of Bristol Law School Blog* [en ligne], consulté le 15 septembre 2025. <https://legalsresearch.blogs.bris.ac.uk/2023/07/towards-a-shared-governance-of-data-as->

digital-commons/

Boukacem-Zeghmouri, Chérifa ; Rakotoary, Sarah ; Bador, Pascal (2021) « La prédation dans le champ de la publication scientifique : un objet de recherche révélateur des mutations de la communication scientifique ouverte », *Natures Sciences Sociétés*, vol. 29, n° 4, p. 382-395.

Boyd, Andy ; Gatewood, Jane ; Thorson, Stuart ; Dye, Timothy D. V. (2019), « Data Diplomacy », *Science & Diplomacy*, vol. 8, n° 1 [en ligne], consulté le 15 septembre 2025. <http://sciencediplomacy.org/article/2019/data-diplomacy>

Bradford, Anu (2012), « The Brussels Effect », *Columbia Law School: Scholarship Archive*. https://scholarship.law.columbia.edu/faculty_scholarship/271

Bradford, Anu (2022), *The Brussels Effect: How the European Union Rules the World*, Oxford University Press.

Brousseau, Éric (2022), « Organiser la valorisation de l'or noir du XXI^e siècle », *Enjeux numériques : propriété et gouvernance du numérique*, vol. 18, n° 2, p. 15-23.

Citton, Yves (2017). « Interlude. Les communs des données », *Médiarchie*, Le Seuil, p. 311-316. <https://shs.cairn.info/mediarchie--9782021349122-page-311?lang=fr>

Couldry, Nick ; Mejias, Ulises Ali (2021), « The decolonial turn in data and technology research: what is at stake and where is it heading? », *Information, Communication and Society*, vol. 26, n° 4, p. 786-802.

Currie, Adrian (2017), « Book review: Sabina Leonelli // data-centric biology: a philosophical study », *British Journal for the Philosophy of Science*, ISSN 0007-0882. <http://eprints.lse.ac.uk/id/eprint/73675>

De Filippi, Primavera ; Hassan, Samer (2016), « Blockchain technology as a regulatory technology: From code is law to law is code », *First Monday*, vol. 21, n° 12.

De Filippi, Primavera ; Merk, Tara (2024), « How to DAO: the role of trust and confidence in institutional design », *Decentralized Autonomous Organisations: Innovation and Vulnerability in the Digital Economy*, Routledge. <https://ssrn.com/abstract=5071975>

Dulong de Rosnay, Mélanie ; Stalder, Felix (2020), « Digital commons. Internet Policy Review », vol. 9, n° 4 [en ligne], consulté le 15 septembre 2025.

European Commission: Directorate-General for Research and Innovation (2018), *Turning FAIR into reality: final report and action plan from the European Commission expert group on FAIR data*. Publications Office.

Hacker, Philipp ; Dimitropoulos, Georgios (2017, « Behavioural law & economics and sustainable regulation », in Mathis, K., Huber, B. (eds), *Environmental Law and Economics. Economic Analysis of Law in European Legal Scholarship*, vol 4, Springer, Cham.

Hess, Charlotte ; Ostrom, Elinor (2005), « A Framework for Analyzing the Knowledge Commons - a chapter from Understanding Knowledge as a Commons: from Theory to Practice », *Libraries' and Librarians' Publications*. 21. <https://surface.syr.edu/sul/21/>

Hooft, Rob ; Suchánek, Marek ; Pergl, Robert (2023), « Data Stewardship Wizard. FAIR Connect », vol. 1, n° 1, p. 41-43. <https://doi.org/10.3233/FC-230501>

Ioannou, Athina ; Tussyadiah, Iis ; Miller, Graham ; Li, Shujun ; Weick, Mario (2021), « Privacy nudges for disclosure of personal information: A systematic literature review and meta-analysis », *PLoS ONE*, vol. 16, n° 8. <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0256822>

Jaton, Florian ; Vinck., Dominique (2016), « Processus frictionnels de mises en bases de données », *Revue d'anthropologie des connaissances*, vol. 10, n° 4, p. 489-504.

Kahneman, Daniel (2011), *Thinking Fast and Slow*, New York, Farrar, Straus and Giroux, traduction en français par Raymond Clarinard, *Système 1/système 2 : Les deux vitesses de la pensée*, Flammarion, 2012.

Kahneman, Daniel ; Tversky, Amos (1984), « Choices, values, and frames », *American Psychologist*, vol. 39, n° 4, p. 341-350.

Kerridge, Richard (2020), « Entangled Life by Merlin Sheldrake review - a brilliant 'door opener' book », *The Guardian*, édition du 27 août 2020 [en ligne], consulté le 15 septembre 2025. <https://www.theguardian.com/books/2020/aug/27/entangled-life-by-merlin-sheldrake-review-a-brilliant-door-opener-book>

Khatri, Vijay ; Brown, Carol V. (2010), « Designing data governance », *Communications of the ACM*, vol. 53, n° 1, p. 148-152.

Leonelli, S. (2019). Data Governance is Key to Interpretation : Reconceptualizing Data in Data Science. *Harvard Data Science Review*, 9.

Meyer, Camille ; Hudon, Marek (2019), « Money and the Commons: An Investigation of Complementary Currencies and Their Ethical Implications », *Journal for Business Ethics*, vol. 160, p. 277-292. <https://doi.org/10.1007/s10551-018-3923-1>

Mons, Barend ; Neylon, Cameron ; Velterop, Jan ; Dumontier, Michel ; da Silva Santos, Luiz Olavo Bonino ; Wilkinson Mark D. (2017); « Cloudy, increasingly FAIR; revisiting the FAIR Data guiding principles for the European Open Science Cloud », *Information Services and Use*, vol. 37, n° 1, p. 49-56.

Ndayisenga, Zenon (2024), « Decolonizing Digital Technology in Africa: A Political Thought », *Journal of Humanities & Social Sciences*, vol. 7, n° 2, p. 1-9.

Nkoudou, Thomas. H. M. (2023), « We need a decolonized appropriation of AI in Africa », *Nature Human Behaviour*, vol. 7, n° 11, p. 1810-1811.

Obendiek, Anke Sophia (2023), *Data governance: Value orders and jurisdictional conflicts (Introduction)*, Oxford University Press.

Ostrom, Elinor (1990), *Governing the commons: The evolution of institutions for collective action*, Cambridge: Cambridge University Press.

Ostrom, Elinor (2010), « Beyond Markets and States: Polycentric Governance of Complex Economic Systems », *American Economic Review*, vol. 100, n° 3, p. 641-72. DOI: 10.1257/aer.100.3.641

Ostrom, Elinor (2009), « A Polycentric Approach for Coping with Climate Change ». Policy Research Working Paper 5095, *Background Paper to the 2010 World Development Report*. <https://ssrn.com/abstract=1934353>

Pergl, Robert ; Hooft Rob ; Suchánek, Marek ; Knaisl, Vojtěch ; Slifka, Jan (2019), « Data Stewardship Wizard: A Tool Bringing Together Researchers, Data Stewards, and Data Experts around Data Management Planning », *Data Science Journal*, vol. 18, n° 59, p. 1-8.

Piron, Florence (2018), « Justice et injustice cognitives : de l'épistémologie à la matérialité des savoirs humains », in Tremblay, Émilie ; Dorcé, Ricarson (dirs), *Les Classiques des sciences sociales. 25 ans de partage des savoirs dans la francophonie*, Éditions science et bien commun p. 259-273, Québec, [en ligne], consulté le 15 septembre 2025. <https://scienceetbiencommun.pressbooks.pub/classiques25ans>

Règlement (UE) 2016/679 du Parlement européen et du Conseil du 27 avril 2016 relatif à la protection des personnes physiques à l'égard du traitement des données à caractère personnel et à la libre circulation de ces données, et abrogeant la directive 95/46/CE (règlement général sur la protection des données) (Texte présentant de l'intérêt pour l'EEE). JO L 119 du 4.5.2016, p. 1-88.

Règlement (UE) 2022/868 du Parlement européen et du Conseil du 30 mai 2022 portant sur la gouvernance européenne des données et modifiant le règlement (UE) 2018/1724 (règlement sur la gouvernance des données) (Texte présentant de l'intérêt pour l'EEE). PE/85/2021/REV/1. JO L 152, 3.6.2022, p. 1-44.

Règlement (UE) 2023/2854 du Parlement européen et du Conseil du 13 décembre 2023 concernant des règles harmonisées portant sur l'équité de l'accès aux données et de l'utilisation des données et modifiant le règlement (UE) 2017/2394 et la directive (UE) 2020/1828 (règlement sur les données) (Texte présentant de l'intérêt pour l'EEE). PE/49/2023/REV/1. JO L, 2023/2854, 22.12.2023.

Règlement (UE) 2024/1689 du Parlement européen et du Conseil du 13 juin 2024 établissant des règles harmonisées concernant l'intelligence artificielle et modifiant les règlements (CE) n° 300/2008, (UE) n° 167/2013, (UE) n° 168/2013, (UE) 2018/858, (UE) 2018/1139 et (UE) 2019/2144 et les directives 2014/90/UE, (UE) 2016/797 et (UE) 2020/1828 (règlement sur l'intelligence artificielle) (Texte présentant de l'intérêt pour l'EEE). PE/24/2024/REV/1. JO L, 2024/1689, 12.7.2024

Rowhani-Farid, Anisa ; Allen, Michelle ; Barnett, Adrian G. (2017), « What incentives increase data sharing in health and medical research? A systematic review », *Research Integrity and Peer Review*, vol. 2, n° 4. <https://link.springer.com/article/10.1186/s41073-017-0028-9>

Simard, Suzanne W. ; Beiler, Kevin J. ; Bingham, Marcus A. ; Deslippe, Julie R. ; Philip, Leanne J. ; Teste, François P. (2012), « Mycorrhizal networks: Mechanisms, ecology and modelling », *Fungal Biology Reviews*, vol. 26, n° 1, p 39-60,

Stodder, James (2009), « Complementary credit networks and macroeconomic stability: Switzerland's Wirtschaftsring », *Journal of Economic Behavior and Organization*, vol. 72, n° 1, 2009, p. 79-95, ISSN 0167-2681. <https://doi.org/10.1016/j.jebo.2009.06.002>

Sunstein, Cass R. (1997), « Behavioral Analysis of Law », *University of Chicago Law Review*, vol. 64, p. 1175.1195.

Sunstein, Cass R. (2014), « Nudging: A Very Short Guide », *Journal of Consumer Policy*, vol. 37, p. 583-588. <https://doi.org/10.1007/s10603-014-9273-1>. Harvard Public Law Working. <https://ssrn.com/abstract=2499658>

Taylor, Linnet (2017), « What is data justice? The case for connecting digital rights and freedoms globally », *Big Data & Society*, vol. 4, n° 2 [en ligne], consulté le 15 septembre 2025.

Timko, Christina ; Niederstadt, Malte ; Goel, Naman ; Faltings, Boi (2023), « Incentive Mechanism Design for Responsible Data Governance », *ACM Journal of Data and Information Quality*, vol. 15, n° 2. <https://dl.acm.org/doi/10.1145/3592617>

Tversky, Amos ; Kahneman, Daniel (1974), « Judgment under uncertainty: Heuristics and biases », *Science*, vol. 185, n° 4157, p. 1124-1131.

Veale, Michael ; Zuiderveen Borgesius, Frederik (2021), « Demystifying the Draft EU Artificial Intelligence Act — Analysing the good, the bad, and the unclear elements of the proposed approach », *Computer Law Review International*, vol. 22, n° 4, p. 97-112.

Woods, Helen Buckley ; Pinfield, Stephen (2022), « Incentivising research data sharing: a scoping review », *Wellcome Open Research*, vol. 6, 355. <https://pmc.ncbi.nlm.nih.gov/articles/PMC8808319/>

Partage et réutilisation des données de la recherche à l'ère de la Science Ouverte. Une approche diplomatique pour penser et panser l'ouverture des données

Article inédit, mis en ligne le 23 septembre 2026

Chérifa BOUKACEM ZEGHMOURI

Chérifa Boukacem-Zeghmouri est Professeure en SIC à l'Université Lyon 1, membre du laboratoire ELICO et responsable de la mention Master Information et Médiation Scientifique (IMST), spécialisé en Science Ouverte. S'appuyant au cadre théorique des industries culturelles et créatives, ses travaux analysent les transformations des formes de production, de circulation, d'évaluation et de légitimation de la recherche scientifique. Ses travaux actuels portent sur les différentes formes d'ambivalences de la Science Ouverte.

Candice FILLAUD

Candice Fillaud est doctorante en Sciences de l'Information et de la Communication à l'Université Lyon 1 et au laboratoire Elico. Elle s'est spécialisée dans les données de la recherche grâce à des études en Science Ouverte et plusieurs expériences professionnelles (CCSD, Urfist de Lyon, Dataactivist, CHIST - ERA). Elle a participé en tant que Ingénieur de recherche au projet JuDDGES financé par le programme européen CHIST- ERA, dans le cadre du WP4 « Science Ouverte ».

Plan de l'article

Résumé et mots clés en français

Title, abstract and keywords in english

Titulo, resumen y palabras clave en espanol

Introduction

Partage et réutilisation des Données vues par le discours institutionnel de la science ouverte : la panne ?

La recherche des données à réutiliser, une compétence à part entière

Recherches financées, données partagées ?

Conclusion et perspectives de recherche

Références bibliographiques

RÉSUMÉ

Alors que le partage des données de la recherche occupe une place importante des politiques de Science Ouverte, leur réutilisation qui en constitue l'oblitération reste encore peu traitée dans la littérature autour de la Science Ouverte. Or, la réutilisation permet d'aborder l'enjeu du partage au-delà des logiques normatives qui ont présidé à la définition des discours institutionnels et de l'opérationnalité des infrastructures mises à disposition. En se fondant sur les travaux de deux projets de recherche (Harbinger 1 et 2 et JuDDGES), le présent article montre que la question du partage des données traitée en relation avec celle de leur réutilisation révèle beaucoup des relations qui se tissent et se négocient entre le large prisme des acteurs impliqués dans ce que nous appellerons « l'événement de réutilisation ». L'approche diplomatique est particulièrement féconde pour prendre en

compte la complexité dans laquelle se nouent ces relations et la nature des logiques qui y interviennent. Abordée sous l'angle des SIC (médiations à l'œuvre, recherche d'information...), elle révèle les dynamiques et les tensions dans lesquelles les propriétés de la valeur des données rencontrent un marché symbolique (des savoirs) et économique (des plateformes).

Mots clés

Science Ouverte, données de la recherche, partage des données, réutilisation des données, valeur, diplomatie des données.

TITLE

On data reuse in the Open Science era. Taking a diplomatic approach to think and care about opening research data.

Abstract

Although data sharing is key in Open Science policies, data reuse, which involves its obliteration, is rarely considered in research on Open Science. However, data reuse enables us to address the issue of sharing beyond the normative logic that has governed the definition of institutional discourse and the operability of the infrastructures made available. Drawing on the findings of two research projects (Harbinger 1 and 2, and JuDDGES), this article demonstrates that considering data sharing in relation to reuse reveals much about the relationships forged and negotiated between the various stakeholders involved in the 'reuse event'. The diplomatic approach is particularly useful for considering the complexity of these relationships and the logic involved. When approached from the perspective of information and communication sciences (e.g. mediations, information retrieval), the diplomatic approach to data reveals the dynamics and tensions through which data value properties encounter a symbolic market (of knowledge) and prefigure a data reuse market.

Keywords

Open Science, data sharing, data reuse, assetization, data diplomacy.

TITULO

Compartir y reutilizar los datos de la investigación en la era de la ciencia abierta. Un enfoque diplomático para pensar y abordar la apertura de los datos.

Resumen

Si bien el intercambio de datos de investigación ocupa un lugar importante en las políticas de ciencia abierta, la reutilización de los datos, que constituye su destrucción, sigue siendo un tema poco tratado en las investigaciones sobre ciencia abierta. Sin embargo, la reutilización de los datos de investigación permite abordar la cuestión del intercambio más allá de las lógicas normativas que han presidido la definición del discurso institucional y la operatividad de las infraestructuras puestas a disposición. Basándose en los trabajos de dos proyectos de investigación (Harbinger 1 y 2 y JuDDGES), el presente artículo muestra que la cuestión del intercambio de datos, tratada en relación con la de su reutilización, revela

mucho sobre las relaciones que se tejen y se negocian entre el amplio prisma de los actores implicados en lo que llamaremos «el evento de reutilización». El enfoque diplomático es especialmente fructífero para tener en cuenta *la complejidad en la que se tejen estas relaciones y la naturaleza de las lógicas que intervienen en ellas. Abordado desde el punto de vista de las ciencias de la información y la comunicación (mediaciones en acción, búsqueda de información, etc.), el enfoque diplomático de los datos revela las dinámicas y tensiones en las que las propiedades del valor de los datos se encuentran con un mercado simbólico (el del conocimiento) y las prefiguraciones de un mercado económico (el de las plataformas).*

Palabras clave

Ciencia abierta, intercambio de datos, reutilización de datos, conversión en activos, diplomacia de datos

INTRODUCTION

Dans ses travaux de thèse, Thérèse Bru (Bru, 2017a ; Bru, 2017b), historienne, s'est intéressée à la correspondance savante des naturalistes du XVIII^e siècle pour examiner de manière originale de quelle manière elle véhiculait des « objets matériels » des sciences naturelles qui construisaient leurs méthodes d'observation. L'intérêt de son travail a été de montrer que cette correspondance a joué un rôle capital dans la circulation de données collectées dans des lieux géographiquement éloignés et difficilement accessibles pour les savants. Ces derniers ont donc entretenu des échanges épistolaires réguliers et intenses avec leurs pairs, ou avec les personnes situées sur les lieux de collecte des spécimens et impliquées dans leur collecte. Les lettres constituées des observations, des descriptions ou des mesures des spécimens collectés, ont véhiculé autant qu'elles ont partagé la matérialité des spécimens observés.

Un autre intérêt de l'approche historiographique des travaux de T. Bru est qu'elle s'est intéressée à la mise en relation de ces données (par l'information) qui a permis d'initier leur circulation, leur partage et leur réutilisation par les savants. Une des conséquences directe est qu'aux côtés des « grandes collections », comme celles qui deviendront à terme le Muséum National d'Histoire Naturelle (MNHN) à Paris qui relèvent de ce qu'elle appelle la logique verticale, on retrouve de plus petites collections, composées le plus souvent de doubles de spécimens, constituées par les savants, qu'ils pouvaient échanger entre eux, pour mener ou poursuivre leurs travaux : les doubles constituent le capital de travail des naturalistes en même temps que leur capital d'échange. Le botaniste Ambroise Marie François Joseph Palisot de Beauvois (1752-1820) offre ainsi ses derniers spécimens de plantes africaines à son confrère britannique, l'influent botaniste Robert Brown (1773-1858), contre des spécimens convoités. Dans la lettre proposant l'échange, Palisot de Beauvois note qu'il a déjà largement distribué ses spécimens africains « à plusieurs botanistes en Europe » (Bru, 2017b).

Dans cette espèce de « troc » des spécimens, l'historienne souligne tout autant le « capital recherche » que représentent ces petites collections pour les savants que la « valeur d'échange » que représente cette forme ancienne de partage des données, identifiée par V. Rebouillat (2021) dans ses travaux de thèse. Elle emploie le verbe « offrir » pour qualifier ce partage destiné à une réutilisation, dans le même sens que des travaux anglo-saxons plus récents emploient le mot « *gift* » (cadeau) pour qualifier le statut des données partagées entre chercheurs (Mattern, 2024).

Dans les modalités de partage et de réutilisation des données naturalistes du XVIII^e siècle

d'une Science Ouverte avant la lettre, les savants n'envisagent plus seulement les grandes collections, « dépôts centralisés nationaux », mis en place par les États (notamment la Grande Bretagne et France), mais ils s'appuient également sur des collections « horizontales », plus distribuées mais inventoriées, interconnectées par les relations qu'ils nouent entre eux. Dans cette approche, c'est la capacité des savants à connaître et à accéder à ces collections dispersées qui définit la taille de la collection et la portée des travaux entrepris : « la collection s'étend aussi loin que la capacité du naturaliste d'accéder aux matériaux nécessaires à son travail » (Bru, 2017b).

Cette perspective historique nous permet d'interroger différemment les résistances au partage des données documentées aujourd'hui quant aux prescriptions du discours institutionnel de la Science Ouverte. Comment expliquer que les savants du XVIIIe siècle ont su mettre en place des collections au service et au profit de leurs pratiques d'échange et de partage de spécimens ou informations sur les spécimens, alors que les chercheurs du XXIe siècle, qui disposent d'une myriade d'entrepôts de données et d'infrastructures (thématiques, nationaux, européens...), résistent autant aux prescriptions de partage des données des politiques de Science Ouverte ? Comment expliquer que la réutilisation des données avancée comme un des grands avantages du partage des données connaisse si peu d'échos auprès des chercheurs ? Car hormis quelques communautés pionnières (ex. écologie, génomique), les autres ne sont pas toujours enclines à partager leurs données (Borgman, 2012 ; Tenopir, 2015 ; Kaden, 2019 ; Fillaud, 2026) et encore moins à réutiliser celles qui sont disponibles (Pasquetto, 2019 ; Borgman, 2025).

L'approche diplomatique des données, proposée dans ce dossier, offre un cadre fécond pour prendre en charge cette question de recherche et en donner une compréhension et une analyse inédites, en dehors des logiques normatives de la définition contemporaine de la Science Ouverte. Telle que définie par Boyd (2019), la diplomatie des données a l'avantage d'envisager la dimension transactionnelle qui s'organise autour de la circulation des données, la pluralité des acteurs mobilisés autour de ces transactions (communautés scientifiques, institutions académiques, grands groupes d'éditeurs, bibliothèques académiques, intermédiaires...) et les régimes de valeurs qui se négocient explicitement ou implicitement entre producteurs et réutilisateurs des données. Même si les acteurs impliqués dans ces transactions sont mus par des motivations différentes (épistémologiques, symboliques, économiques, culturelles...) et des rapports de pouvoir, ils parviennent, par la justification de leurs actions et de leurs pratiques, à ajuster les différences qui existent entre leurs intérêts respectifs pour que les transactions puissent opérer. Cette justification rejoint donc les mécanismes d'une économie des grandeurs analysés par Luc Boltanski (1991) qui parvient à aligner intentions peu explicites et objectifs parfois divergents.

L'autre intérêt de cette approche est qu'elle envisage l'ouverture des données, et donc leur partage, de manière étroitement liée à leur réutilisation :

« Data diplomacy sits at the nexus of data, governance, and social norms, and seeks to anticipate the impact and influence of future data use. Bearing this in mind, pioneers in the field can prepare mechanisms to facilitate data transactions in ways that are acceptable to a range of stakeholders » (Boyd, 2019).

La prise en compte en même temps du partage des données et de leur réutilisation (comme modalité d'ouverture) donne donc la possibilité de revenir aux enjeux documentaires de leur production, de leur recherche (informations et métadonnées) et de leur accès en vue de leur interprétation. Elle nous conduit à proposer dans cet article ce que nous appellerons « l'événement de réutilisation » que nous définissons comme la réalisation d'un partage entre producteurs et réutilisateurs dans une relation symétrique où chacune des parties trouve son intérêt, quelle que soit sa nature, et où le partage inclut l'implicite et l'explicite qui ont présidé à la production des données. Ce que nous appelons donc « événement de réutilisation » est un partage de données qui permet une réutilisation effective dans de nouvelles recherches, dont l'agenda scientifique est défini en toute liberté et en toute

autonomie. La réutilisation est donc un événement car elle marque un infléchissement dans la réflexion du chercheur qui décide de réutiliser des données ou un jeu de données pour aborder une nouvelle question de recherche.

L'article est organisé autour de trois parties. La première constate la portée limitée des discours institutionnels de la Science Ouverte sur les pratiques de partage des données, et plus particulièrement sur leur réutilisation. Par la même occasion, cette partie interroge le bien-fondé de ces arguments incitatifs. La deuxième partie traite de la recherche des données par les chercheurs, comme compétence et enjeu crucial de la réutilisation. Elle utilise des résultats des projets Harbinger et JuDDGES pour illustrer les freins et les résistances des chercheurs à développer cette compétence et ces pratiques. La dernière partie traite de la constitution de la réutilisation des données en marché, révélé par l'approche diplomatique qui permet de décrypter les logiques de captation de la valeur qui sont en cours. L'article conclut sur des perspectives de recherches autour de la thématique encore peu explorée de la réutilisation des données.

PARTAGE ET RÉUTILISATION DES DONNÉES, VUES PAR LE DISCOURS INSTITUTIONNEL DE LA SCIENCE OUVERTE : LA PANNE ?

La Science Ouverte comme politique et discours institutionnel (MESR, 2018 ; 2021) se développe et se renforce avec l'adhésion des acteurs de la recherche, mais aussi avec l'observation d'une évolution des pratiques de la recherche. Le partage des données est au cœur de l'attention de tous les acteurs du champ de l'information et de la communication scientifiques. Il représente un défi majeur pour les régulations créatives de la Science Ouverte et les registres d'argumentation mobilisés pour justifier l'ouverture des données englobent le citoyen pour mieux souligner les enjeux sociétaux. Dès 2017, les politiques européennes (Sparc Europe, 2017), montraient que les données étaient au cœur des prescriptions et devenaient une priorité. Pourtant, l'enjeu du partage des données fédère des acteurs pour lesquels le consensus implicite est peu questionné (Tkacz, 2012 ; Mattern, 2024).

En France, c'est le second Plan National pour la Science Ouverte (PNSO 2) qui mettra une emphase particulière sur les données et leur ouverture (MESR, 2021). Les établissements académiques valorisent une définition plus large de la production scientifique de leurs chercheurs avec les ateliers de la donnée et le soutien à l'entrepôt national [RechercheDataGouv](https://www.recherche-data.gouv.fr/). Les éditeurs scientifiques développent de nouveaux produits (*Data Journals*, *Data papers*). Les agences de financement de la recherche consolident pour leur part le retour sur investissement des moyens accordés (Anger, 2022). Les plateformes d'intermédiation spécialisées telles que *FigShare*, *Zenodo* en quête de valeur, développent de nouveaux services de gestion et de partage des données. Enfin, les bibliothèques universitaires et académiques, mobilisées pour la Science Ouverte, investissent la gestion des données scientifiques et proposent des services aux chercheurs pour les accompagner (Schöpfel, 2023).

Les chercheurs objectent et voient dans certains des arguments avancés autour de l'ouverture des données une méconnaissance du fonctionnement du champ scientifique, voire une idéologie gestionnaire et néo-libérale (Mirowski, 2018 ; Fillaud, 2026). De même, tel qu'il est décrit dans les discours institutionnels de la Science Ouverte, le travail requis pour l'ouverture et pour le partage des données est appréhendé comme un levier susceptible de transformer leur métier en tâches administratives (Hostler, 2023). À titre d'exemple, les modalités de gestion des données (ex. Plans de Gestion de Données – PGD), sont perçues comme les vecteurs opératoires d'une bureaucratisation ou de nouvelles formes d'évaluation et de contrôle de la recherche.

Alors que les infrastructures techniques facilitent aujourd'hui l'ouverture et le partage des données (entrepôts, de données, DOI, format de fichiers...), que les discours politiques légitiment la nécessité de le faire et que des modèles d'affaires existent, le chercheur lui, hésite, résiste ou bien décide d'ignorer les recommandations. Ce constat se retrouve dans toutes les études réalisées, publiées en Europe, en Amérique du Nord et désormais dans les pays émergents : hormis quelques communautés pionnières (ex. astrophysique, écologie ou génomique), les autres envisagent le partage des données sous des conditions plus spécifiques à leurs cadres et cultures disciplinaires (Berberi, 2022 ; Mattern, 2024). La période de la pandémie a confirmé ce constat montrant que si les résultats de recherche ont été partagés, les données l'ont été dans une bien moindre mesure (Lucas-Dominguez, 2021).

Y a-t-il un « public » pour la réutilisation des données ?

La réutilisation des données est une des questions de la Science Ouverte les moins documentées. Il ne s'agit pas complètement d'un angle mort car nous disposons dès 2017 de travaux formalisant ses enjeux (Pasquetto, 2019). Il s'agit plutôt d'une thématique peu développée faute de données (sans mauvais jeu de mots) reposant sur des cas concrets et effectifs de réutilisation.

De même, il est difficile de s'accorder sur une définition de la réutilisation des données. D'un domaine disciplinaire à un autre, d'un contexte de recherche (fondamental ou expérimental) à un autre, la réutilisation des données prend un sens différent. En Sciences Humaines et Sociales (SHS) par exemple, la question du partage des données en vue de leur utilisation soulève des questions particulièrement intéressantes du point de vue de la reproductibilité ou de la répliquabilité de la recherche : le type de questions de recherche mobilisées, les approches méthodologiques employées et le statut des données produites, n'offrent pas la possibilité d'envisager leur réutilisation, dans les termes des recommandations ou prescriptions politiques (Kim, 2015 ; Fillaud, 2026).

Les travaux de Borgman (2015) ont montré que les données produites en contexte de recherche sont dotées d'un capital ou d'une valeur propre (*assets*), ancrée dans leur contexte de production. Ce capital est également vecteur de la crédibilité de la recherche et du collectif qui a permis de les produire. Or, cette crédibilité est un processus dynamique, constant, qui engage un ensemble de pratiques, peu formalisées, rattachées à la culture des collectifs impliqués dans l'activité de recherche et a fortiori de production de données. La crédibilité est donc, nous le rappelle Shapin (2014), un processus inachevé car en constante adaptation.

Il est d'ailleurs intéressant de constater la place des métaphores utilisées pour approcher la réutilisation des données, comme démarche, pratique ou concept. Elles agissent comme des espaces intellectuels de médiation, pour faciliter le dialogue et l'échange entre spécialistes. La métaphore de la distance (*data distance*) employée récemment par Christine Borgman (2024), se discute avec celle des voyages de données (*data journeys*) de Sabina Leonelli (2020). Cette dernière explique que pour faciliter la réutilisation des données, il est parfois nécessaire de ne pas envisager leur partage pour privilégier les narrations qui portent leurs histoires (comprises dans les deux sens du terme) (Leonelli, 2025). De plus en plus présentes dans l'univers de la communication scientifique qui se complexifie, le recours aux métaphores permet aux chercheurs de circonscrire des phénomènes en objets de recherche plus simples à appréhender et à comprendre. Les métaphores agissent tels des ponts interdisciplinaires pour des chercheurs intéressés à un même objet mais venant de cultures scientifiques différentes. Le pouvoir évocateur des métaphores joue donc un rôle important dans les formalisations théoriques proposées par ces auteurs, notamment pour agréger un ensemble hétérogène de pratiques, peu formalisées, rattachées à la culture des individus ou des collectifs impliqués dans l'activité de recherche.

L'usage des métaphores autour de la réutilisation des données indique surtout que la réutilisation des données opère d'abord dans des espaces d'indétermination et

d'interactions de relations personnelles avant qu'elle ne se fasse dans les espaces infrastructurels (techniques ou normatifs) dédiés au partage. La réutilisation s'oblitére d'abord dans les modes de travail collectifs et collaboratifs des chercheurs connectés à des lieux physiques, autant qu'à des messageries, à des ordinateurs, à des bases de données ou bien à des plateformes. L'évènement de réutilisation des données est donc l'aboutissement d'une amplification de la réflexion fondée sur des relations entre les personnes engagées et leurs modes de connaissances spécialisées, outillés par des plateformes (Pasquetto, 2019). C'est dans cet entrelacs et dans cette complexité qu'émerge l'évènement de réutilisation qui n'est pas une réponse au discours institutionnel de la Science Ouverte, mais le prolongement d'une pensée autant qu'une pratique contemporaine du travail de la recherche.

LA RECHERCHE DES DONNÉES À RÉUTILISER, UNE COMPÉTENCE À PART ENTIÈRE

La question de la recherche de données (*data discovery*) mérite d'être prise en compte dans notre analyse car elle entre directement en jeu dans le fait de favoriser l'évènement de réutilisation. La recherche de données indique une motivation à les réutiliser, ou à tout le moins à les examiner. Pour autant, comme le montre Anthony J. Million (2025), la recherche de données échappe aux théories qui ont présidé à la recherche d'information traditionnelle.

Pour la recherche des données, leurs métadonnées et les documents apportant leur contextualisation sont tout aussi importants car ils déterminent pour l'utilisateur l'examen des chances et opportunités de faire jonction avec les questions de recherche qu'il traite. Ce sont ces informations qui vont pousser le chercheur à consulter le jeu de données puis à faire basculer sa décision du côté de la réutilisation, ou pas.

Comme le soulignent les travaux de thèse de T. Friederich (2020) soutenue à l'Université d'Humboldt, bien plus que la prise en compte des principes FAIR, il s'agit ici d'accompagner les données de documents où le degré de précision des informations apportées et leur complétude sont rattachés aux connaissances sociales et scientifiques qui se nourrissent et se transmettent au cœur de l'activité de recherche, au cœur des communautés scientifiques, le plus souvent par acculturation.

Un jeu de données ouvert mais sans métadonnées (ou insuffisantes) et sans documents de contextualisation est donc un jeu de données au mieux non-réutilisé, au pire perdu. L'ouverture des données et a fortiori leur réutilisation ne dépendent pas de l'ouverture et de la qualité de leurs métadonnées.

Les travaux de Million (2025) donnent la possibilité de comprendre que les trajectoires de recherche de données sont des agencements cognitifs complexes car ils visent à questionner et à évaluer à chaque étape, tout au long de la recherche, l'adéquation du jeu de données au cadre méthodologique : les questions de recherches définies, hypothèses associées et les méthodes employées. La réutilisation des données dépend donc de cette quête d'information et d'adéquation des informations trouvées à ses propres objectifs de recherche.

Cette démarche, chronophage et risquée, représente un « coût d'entrée » dans le jeu de données qui incite le plus souvent les chercheurs à renoncer à la réutilisation pour collecter leurs propres données dans la plupart des cas. Mais c'est aussi précisément parce qu'ils sont conscients de ce coût d'entrée que, dans d'autres cas, les chercheurs ont recours à la collaboration pour obtenir le jeu de données et l'ensemble des informations et des connaissances qui lui sont associées (Pasquetto, 2019 ; Gregory, 2020).

Observer les pratiques de réutilisation des données

Le projet Harbinger a constitué un observatoire des pratiques et stratégies des jeunes chercheurs considérés comme la relève dans un système académique globalisé, car accueillant désormais les pays dits autrefois émergents (ex. Chine). Cet observatoire a rassemblé huit pays (France, Espagne, Pologne, Etats-Unis, Royaume Uni, Chine, Malaisie, Russie) sur une période allant de 2016 à 2021 (Nicholas, 2022 ; Nicholas, 2023).

Fondées sur une enquête qualitative avec des entretiens individuels semi-directifs, les analyses longitudinales du projet ont investigué 23 thématiques (dont celle des données de la recherche) relatives au travail, aux pratiques et aux représentations des jeunes chercheurs (JC) pour examiner leur potentiel de transformation du système de la communication scientifique. Ainsi, en donnant la voix à ceux qui sont situés aux avant-postes de la production des données dans les laboratoires, et les plus soumis aux pressions du système de l'évaluation de la recherche, nous avons analysé la manière dont la question du partage en lien avec la réutilisation des données était comprise, envisagée ou pratiquée. À partir de cet observatoire, ce sont donc les pratiques et les discours des 16 chercheurs et chercheuses du terrain français (13 en STM et trois en sciences sociales) qui ont été suivis pendant le projet, à mesure qu'ils cheminaient dans le monde de la recherche et de la montée des discours de Science Ouverte.

Les données de la recherche ont été abordées du point de vue de leur production, de leur gestion, de leur ouverture, de leur partage (formel et informel) et de leur éventuelle réutilisation. Hormis une chercheuse en SHS (histoire moderne) qui avait pour objet un herbier et qui a explicité la manière dont les spécimens représentaient des données qu'elle devait documenter pour qu'ils puissent être accessibles et utilisés, tous les autres chercheurs et chercheuses ont eu une compréhension contrastée de ce que représentaient ces notions.

Le partage des données a été compris et pratiqué en lien avec les publications (sans adoption des principes FAIR) et la réutilisation des données (les siennes et celles d'autrui) comme pratique, n'a pas été confirmée. L'analyse longitudinale des entretiens menés auprès des JC a identifié un ensemble de facteurs qui expliquent cette absence de pratique et le manque de motivation pour la développer.

Le premier facteur est identifié à la valeur du travail et du savoir-faire scientifique du jeune chercheur qui souhaite être reconnu et obtenir un capital symbolique dans les luttes du champ scientifique (Bourdieu, 1976). Il faut comprendre que la crédibilité du travail du JC dans sa communauté et sa reconnaissance passent par la production de nouvelles données ou par un nouveau type de données :

« J'ai passé deux années aux USA pour apprendre et maîtriser la méthode développée dans le laboratoire d'un Prix Nobel, je sais que cela représente un atout pour mon CV » (JC 3, Bio-chimie, H).

Pour les deux JC en astrophysique (domaine pionnier dans le partage et la réutilisation des données) du panel, ils ont avancé le fait d'être plus « attendus » sur leur capacité à produire de nouvelles données qu'à réutiliser celles déjà produites. Le recours à la réutilisation pourrait donc être interprété comme une créativité « en panne » ou qui aurait du mal à se développer. Se joue ainsi l'enjeu de la production de nouvelles données, preuve de la valeur du savoir-faire scientifique du JC qui endosse un rôle plus actif et plus valorisant.

Cette vision selon laquelle les recherches fondées sur des données réutilisées ont moins de valeur symbolique est partagée par les autres JC qui estiment avoir été recrutés pour produire des données et des résultats originaux valorisés dans des articles de recherche qu'ils soumettent à des revues reconnues, domine. Ceci conduit à identifier un deuxième facteur : leur compréhension de la pratique du partage des données se confond donc avec celle de la publication des jeux de données dans les articles, le plus souvent sous forme de données supplémentaires (*supplementary materials*), qui ne reposent pas sur les principes FAIR du partage des données alors même que les recherches sont financées et donc soumises aux

obligations des agences de financement. La production des données étant intrinsèquement liée aux expertises - un JC a employé le terme de « performance » - méthodologiques, elles représentent « un trésor de guerre » et sont donc valorisées via la publication. La règle implacable du *Publish or Perish* à laquelle ils sont soumis, neutralise donc toute motivation à les partager autrement que par leurs publications.

Un troisième facteur relève du travail consacré pour que des données soient « réellement partageables » est à considérer :

« Il faut y passer un sacré temps pour véritablement permettre à un jeu de données à être réutilisé clé en mains, sans que l'on contacte la personne qui les a produites, qu'on pose des tonnes de question et de précisions. C'est pour ça que quand j'ai été contacté pour mes données, j'ai préféré proposer une collaboration : tout le monde y gagne, et l'article se fait plus facilement » (JC 7, Sciences de l'environnement, H).

La figure du chercheur productif et créatif qui crée de la valeur à chacune de ses tâches, véhiculée dans les travaux critiques d'une Science Ouverte néo-libérale ne se confirme pas ici. Les JC soumis à la précarité de leur situation professionnelle sont vigilants sur le temps et sur l'effort consacrés à des tâches à fort capital symbolique, telle que celle de la production des données.

Un quatrième facteur repose sur le principe d'un partage « utile » et non moral des données. Le potentiel de la réutilisation ne peut pas être pensé qu'au bénéfice supposé des autres, mais doit aussi être pensé à son propre bénéfice. Certains JC ont remis en question l'intérêt de l'idée d'une réutilisation des données, en avançant l'argument du *small data* et donc « *trop petites données pour mériter d'être partagées* » ou bien de la grande complexité du sujet de recherche :

« Mon travail de recherche est tellement pointu qu'il ne peut pas intéresser grand monde, cela n'a aucun intérêt » (JC11, Physique, H).

Le seul cas de réutilisation des données a été présenté comme une proposition de collaboration, dont l'opportunité s'est présentée alors que le jeune chercheur évaluait un article. Il a proposé à l'auteur de collaborer en réutilisant les données dans un objectif de traitement inédit, qu'il maîtrisait, et qui leur permettait de produire de nouveaux résultats. Ce cas de figure confirme que producteurs et réutilisateurs de données se choisissent (Pasquetto, 2019), ils définissent le cadre de réutilisation qui lui permet d'être effectif et ils prennent entièrement en charge ses modalités en vue de produire l'événement de réutilisation. À l'instar des savants du XVIIIe siècle (Bru, 2017), ou des scientifiques de notre époque (Gabelica *et al.* 2022), les chercheurs s'impliquent dans le travail même d'analyse et d'interprétation qu'implique la collaboration d'une recherche commune. Cette attitude est d'autant plus importante que cela évite les risques de mauvaise interprétation (Baker & Mayernik, 2020).

RECHERCHES FINANÇÉES, DONNÉES PARTAGÉES ?

Parmi les acteurs qui œuvrent activement à faire du partage des données une pratique développée dans les communautés de recherches, figurent les agences de financement. Elles bénéficient d'une délégation d'autorité pour sélectionner et financer des programmes de recherches. Outre la publication des résultats des travaux menés par les collectifs des projets, les agences de financement prescrivent et rendent obligatoires des modalités de gestion et de partage de données. Ces politiques se sont développées à l'international, en Europe et en France, où l'ANR a suivi dès 2019 en rendant le PGD obligatoire pour tout projet de recherche financé, obligeant les chercheurs financés à décrire la manière dont leurs données sont gérées et à anticiper leur ouverture à l'issue du projet. Le célèbre *National Institute of Health* a même produit un indicateur de mesure de partage des données, le *NIH Sharing Index (S-index)*, et [organise des événements dédiés](#).

Pour autant, il est pertinent de questionner les effets de ces obligations sur les pratiques des chercheurs. Les travaux de Michael Anger (2022) se sont penchés sur les difficultés rencontrées par les agences de financement européennes à rendre leurs obligations effectives. Le grand intérêt de ces recherches inédites réside dans le fait qu'ils se sont intéressés plus particulièrement à ce qui motivait les chercheurs à développer des « bonnes » pratiques de gestion et de partage des données. Cette étude menée via des entretiens semi-directifs avec 16 agences de financement, a montré à quel point les prescriptions, peu claires et trop génériques, achoppent à susciter de l'intérêt pour les pratiques de partage de données.

Mais les agences de financement persévèrent et se coordonnent entre elles. Le programme CHIST-ERA est précisément un programme de financement, regroupant des agences européennes, dont l'ANR, qui non seulement a rendu son PGD obligatoire, mais aussi la plateforme qui permet de le produire et de l'héberger, ARGOS.

Dans le cadre du projet interdisciplinaire [JuDDGES](#), financé par le programme CHIST-ERA qui nous a donné la possibilité en tant que partenaire français d'être financés par l'ANR, nous avons déployé une enquête qui comprend un volet préparatoire, mené entre octobre 2024 et février 2025. Elle a permis de mener trois focus groupes avec des chercheurs et chercheuses de différents pays européens (France, Suisse, Espagne, Italie, Belgique) financés par le programme CHIST-ERA, avec lesquels la question de la gestion des données en vue de leur partage et de leur réutilisation a été abordée.

Un premier point de convergence dans le propos des chercheurs est leurs attentes en matière de formation, d'accompagnement, voire d'explicitation, nécessaire à la production des PGD spécifiques au programme CHIST ERA. Car les chercheurs habitués aux PGD de leur agence de financement nationale (ex. DMP Opidor pour l'ANR en France) découvraient un nouveau modèle imposé. Ceci explique le « bricolage » déployé pour en assurer la production : usage de Google Doc ou tout autre outil pour la rédaction collaborative, Word lorsque c'est une seule personne qui est assignée à la tâche, ou bien l'usage des IA génératives.

Un deuxième point de convergence relevé est le statut même du PGD produit vis-à-vis du partage des données ainsi que son devenir. Appréhendé comme un document administratif, les chercheurs lui attribuent une limite dans sa capacité à prendre en compte le « travail de gestion et d'ouverture des données », souvent implicite, réalisé pendant le projet et qui ne relève pas des catégories présentes dans les modèles du PGD. L'implicite est situé au cœur de l'activité de la recherche réalisée, invisible le plus souvent, et donc pas reconnue (Hostler, 2023).

Le troisième point qui construit la convergence des propos est le devenir même des PGD. L'expérience des chercheurs souligne l'absence de retours de la part des agences de financement. Bien précisé de nombreuses fois par les agences elles-mêmes, ce n'est pas leur rôle de répondre au questionnement actuel : « À qui sont destinées les PGD ? ». Ce silence a questionné l'utilité du PGD et plus particulièrement celle du temps et du travail consacrés à le produire. Le PGD est donc pensé comme un document à produire « pour les autres » : l'agence de financement, les gestionnaires de la recherche, les institutions..., sans valeur de preuve ou de crédibilité scientifique.

Enfin, les focus groupes indiquent qu'aucun des participants n'a mentionné le PGD comme levier de partage des données ou de leur réutilisation. Les PGD produits pour leurs projets sur la plateforme ARGOS, proposée par le programme CHIST-ERA, ne sont pas associés à des espaces de réflexions ou d'observations autour de la donnée ou des recherches qui président à sa production. En revanche, c'est le PGD lui-même qui est partagé pour aider les collègues, voire la communauté à produire un document équivalent.

Un « marché » de la réutilisation de la donnée, éclairé par l'approche diplomatique

Les travaux montrent dès les années 2010 que les données produites en contexte de recherche sont dotées d'un capital ou d'une valeur propre (*assets*), liés à leur contexte de

production (Borgman, 2012). Elles sont enrôlées dans un processus de production ou d'exploitation de la valeur (*assetization*) qui émerge à un enjeu d'économie politique de l'information et de la communication scientifique qui interroge les formes d'exploitation et de captation de valeur qui les entourent.

L'approche de la diplomatie des données adoptée dans cet article, est particulièrement pertinente pour montrer de quelle manière ce phénomène d'infiltration des plateformes privées dans les pratiques des chercheurs, fondé sur des processus d'*assetization*, introduit par la même occasion des interactions et des relations plus ou moins formalisées entre chercheurs, acteurs de la *Big Tech*, agences de financements, universités et grands groupes de l'édition scientifique. Elle permet précisément d'aborder la réutilisation des données (et donc sa gestion et son partage) en tant que marché où la multiplicité des acteurs engagés ne négocie pas seulement de la valeur économique, mais aussi de la valeur d'usage pour des enjeux de reconnaissance et de positionnement : des concours pour devenir des ambassadeurs de réutilisation des données sont organisés, des *Success Stories* sont plébiscitées et médiatisées pour montrer de quelle manière « le chercheur parvient à maximiser et à optimiser le potentiel de réutilisation des jeux de données ».

Pour les bibliothèques académiques, les organismes de recherche ou les universités, leurs événementiels autour de la Science Ouverte (ex. *Open Data Week*, Journées Science Ouverte...) d'abord tournés vers la publication scientifique ont progressivement inclus le partage des données et ils abordent depuis peu la question de leur réutilisation. Là aussi, dans les programmes des événements, des profils d'entrepreneurs de normes, chercheurs ou professionnels qui incarnent la performativité des discours institutionnels, apportent les témoignages et les retours d'expériences de partage et de réutilisation des données et en font la promotion.

L'exemple de la plateforme [Kaggle](https://www.kaggle.com/), propriété de Google, lui-même propriété du conglomérat technologique *Alphabet*, est particulièrement illustratif de ce phénomène. La plateforme est entièrement tournée et dédiée à la réutilisation des données scientifiques dans les domaines de l'informatique et plus particulièrement dans l'apprentissage automatique (*machine learning*). Elle propose la médiation entre chercheurs, voire entre communautés scientifiques, et vise à devenir le lieu de « transit » et de mise en visibilité des données partagées, à réutiliser. L'exploitation des statistiques de flux sur la plateforme (téléchargements, partages, échanges entre chercheurs, etc.) octroie un pouvoir d'observation, voire de surveillance (Pooley, 2022) des modes de production et de partage des données, complètement inédit. *Kaggle* n'est pas le seul nouvel entrant des plateformes de la réutilisation de la donnée scientifique. Un marché dédié se développe avec des stratégies d'acteurs qui se précisent et viennent consolider la numérisation et l'informationnalisation de la société (Miège, 2017 ; Miège, 2020). Ils produisent des phénomènes de désintermédiation et de captation de la valeur (*assetization*), voire d'enclosures, au fil des recompositions. Ces modalités, aux antipodes des principes de la Science Ouverte, sont pointées par des travaux qui posent la question du devenir de l'ouverture dans le monde académique, des communs et de leur gouvernance (Tkacz, 2012; Moore, 2025).

L'approche diplomatique des données est donc éclairante de la façon dont les régulations d'ouverture sont subordonnées aux *workflows* des plateformes des acteurs privés qui visent à en extraire la valeur pour un marché de la réutilisation de la donnée scientifique qu'elles cherchent à conquérir et qui demande à être étudié (Acker, 2025). Cette intrication nourrit les analyses qui identifient une promiscuité entre régulations de la Science Ouverte et celles d'une néo-libéralisation de la science (Kansa, 2014 ; Mirowski, 2018 ; Mattern, 2024). Comme le souligne A. Acker (2025) en réponse à la métaphore de la réutilisation des données « *When platforms are used in research data management and scientific enterprise, they typically render data as assets. More simply, we must distinguish research data from when it is packaged for communities by scientists from when it is packaged as commodity by*

platforms. ». Son commentaire est particulièrement important sur les enjeux de l'Intelligence Artificielle (IA) qui intervient désormais dans les nouveaux développements de ces plateformes, qui introduisent des régulations techniques, mises au service de l'*assetization*, et qui accentuent les distances entre les producteurs et les réutilisateurs de données.

CONCLUSION ET PERSPECTIVES DE RECHERCHE

Nous avons abordé dans cet article la question du partage et de la réutilisation des données de la recherche dans une approche permettant de nous écarter des cadres normatifs qui accompagnent trop souvent les travaux autour des enjeux de la Science Ouverte. En adoptant l'approche diplomatique des données, nous avons traité de la résistance des chercheurs au partage de leurs données et de leur manque de motivation à la réutilisation des données partagées, comme une thématique de recherche objectivée, qui n'appréhende pas la résistance comme un dysfonctionnement.

Cette approche a été une entrée féconde pour notre réflexion car elle a montré de quelle manière la réutilisation des données, au cœur des principes de la Science Ouverte, est révélatrice des tensions autour de la captation de la valeur des données de la recherche. Celle-ci se situe dans le rapport des chercheurs aux valeurs de la Science Ouverte, leur adhésion ou résistance à sa mise en application concrète, et les logiques managériales des plateformes de la *Big Tech*. Surtout, l'approche diplomatique des données offre la possibilité de penser les limites des propositions théoriques actuelles, qui, parfois engoncées dans des métaphores, ne tiennent pas suffisamment compte de la réalité des pratiques des chercheurs et des espaces numériques dans lesquels elles se déploient. L'approche diplomatique permet de penser autant que de « panser », au sens du *care* employé par Samuel Moore (2025), les mécanismes par lesquels l'ouverture parvient à exister entre pratiques légitimes des communautés scientifiques, plateformes numériques et prescriptions politiques (Tkacz, 2012).

Notre article conclura donc sur l'urgence d'entreprendre des travaux empiriques autour de la réutilisation des données, qui ne négligent pas la pluralité des acteurs impliqués et l'intrication de leurs stratégies déployées. De même, il souligne l'importance de penser les méthodes (quantitatives, bibliométriques ou qualitatives) à mobiliser pour traiter de l'enjeu de la réutilisation des données de la recherche, en tant que pratique à la confluence de la production du crédit scientifique, du travail de la recherche, des injonctions normatives de la Science Ouverte et des logiques de captation de la valeur développées par les propriétaires de plateformes. Enfin, il encourage à penser les méthodologies à mobiliser pour traiter de la question de la réutilisation des données, comme un axe de recherche à part entière, trop peu traité dans la littérature scientifique.

RÉFÉRENCES BIBLIOGRAPHIQUES

Acker, Amelia (2025), « How Platforms Constrain Matters of Distance and Data Reuse », *Harvard Data Science Review*, 7 (2). <https://doi.org/10.1162/99608f92.c0d03e91>.

Anger, Michael ; Wendelborn Christian ; Winkler Eva C. ; Schickhardt Christoph, (2022), « Neither Carrots nor Sticks? Challenges Surrounding Data Sharing from the Perspective of Research Funding Agencies—A Qualitative Expert Interview Study », *PLOS ONE*, 17 (9), e0273259. <https://doi.org/10.1371/journal.pone.0273259>.

Baker, Karen S. ; Mayernik Matthew S. (2020), « Disentangling Knowledge Production and Data Production », *Ecosphere*, 11 (7), e03191. <https://doi.org/10.1002/ecs2.3191>.

Berberi, Ilias ; Roche Dominique G. (2022), « No Evidence That Mandatory Open Data Policies Increase Error Correction », *Nature Ecology & Evolution*, 6 (11), p. 1630-33.

<https://doi.org/10.1038/s41559-022-01879-9>.

Boltanski, Luc ; Thévenot Laurent (1991), *De La Justification. Les Économies de La Grandeur*. Gallimard.

Borgman, Christine L. (2012), « The Conundrum of Sharing Research Data ». *Journal of the American Society for Information Science and Technology*, 63 (6), p. 1059-78. <https://doi.org/10.1002/asi.22634>.

Borgman, Christine L. (2015), *Big Data, Little Data, No Data: Scholarship in the Networked World*. The MIT Press. <https://www.jstor.org/stable/j.ctt17kk8n8>.

Borgman, Christine L.; Groth Paul (2025), « From Data Creator to Data Reuser: Distance Matters ». *Harvard Data Science Review*, 7 (2). <https://doi.org/10.1162/99608f92.35d32cfc>.

Bourdieu, Pierre (1976), « Le champ scientifique ». *Actes de la Recherche en Sciences Sociales*, 2 (2), p. 88-104. <https://doi.org/10.3406/arss.1976.3454>.

Boyd, Andy ; Gatewood Jane ; Thorson Stuart J. ; Dye Timothy (2019), *Data Diplomacy*. <https://www.sciencediplomacy.org/article/2019/data-diplomacy>.

Bru, Thérèse (2017a), « Plus vrai que nature : conversion de l'information scientifique en objets, conversion des objets en informations dans les correspondances en sciences naturelles:(xviiiè-xixè siècles, mondes britanniques et français) ». In *Matière à écrire*. Presses universitaires de Vincennes. <https://doi.org/10.3917/puv.bru.2017.01.0179>.

Bru, Thérèse (2017b), « La collection éclatée », *Artefact. Techniques, histoire et sciences humaines*, N. 6, 6. <https://doi.org/10.4000/artefact.872>.

Fillaud, Candice ; Boukacem-Zeghmouri Chérifa ; Fei Yutong ; Favel-Kapoian Valentine (2026), « The French HSS Community Speaks Out on Open Science: A Top-Down and Bottom-Up Taxonomy Approach », *The Journal of Electronic Publishing*, 29 (1). <https://doi.org/10.3998/jep.7835>.

Friedrich, Tanja (2020), *Looking for Data*. November 30. <https://doi.org/10.18452/22173>.

Gabelica, Mirko ; Bojčić Ružica ; Puljak Livia (2022), « Many Researchers Were Not Compliant with Their Published Data Sharing Statement: A Mixed-Methods Study », *Journal of Clinical Epidemiology*, 150 (October), p. 33-41. <https://doi.org/10.1016/j.jclinepi.2022.05.019>.

Gregory, Kathleen ; Groth Paul ; Scharnhorst Andrea ; Sally Wyatt Sally (2020), « Lost or Found? Discovering Data Needed for Research », *Harvard Data Science Review*, 2 (2). <https://doi.org/10.1162/99608f92.e38165eb>.

Hostler, Thomas J. (2024), « The Invisible Workload of Open Research », *Journal of Trial & Error*, 4 (1). <https://doi.org/10.36850/mr5>.

Kaden, Ben (2019), « Pourquoi les données de recherche ne sont-elles pas publiées? », *Études de communication. Langages, information, médiations*, No. 52. <https://doi.org/10.4000/edc.8783>.

Kim, Youngseek ; Adler Melissa (2015), « 'Social Scientists' Data Sharing Behaviors: Investigating the Roles of Individual Motivations, Institutional Pressures, and Data Repositories », *International Journal of Information Management* (Netherlands), 35 (4), p. 408-418. <https://doi.org/10.1016/j.ijinfomgt.2015.04.007>.

Leonelli, Sabina (2025), « What to Do About Data Distance? Responsible Alternatives to Data Sharing », *Harvard Data Science Review*, 7 (2). <https://doi.org/10.1162/99608f92.e95b5c26>.

Leonelli, Sabina ; Tempini Niccolò (eds.) (2020), *Data Journeys in the Sciences*. Springer International Publishing. <https://doi.org/10.1007/978-3-030-37177-7>.

Mattern, Janice Bially ; Kohlburn Joseph ; Heather Moulaison-Sandy Heather (2024), « Why Academics Under-Share Research Data: A Social Relational Theory », *Journal of the*

Association for Information Science and Technology, 75 (9), p. 988-1001.
<https://doi.org/10.1002/asi.24938>.

MESR (2018), « Plan national pour la Science Ouverte », *Ouvrir la Science*, November 17.
<https://www.ouvriirlascience.fr/plan-national-pour-la-science-ouverte>.

MESR (2021), « Deuxième Plan National Pour La Science Ouverte », *Ouvrir La Science*.
<https://www.ouvriirlascience.fr/deuxieme-plan-national-pour-la-science-ouverte/>.

Miège, Bernard (2017), *Les industries culturelles et créatives face à l'ordre de l'information et de la communication*. Presses universitaires de Grenoble.
<https://doi.org/10.3917/pug.miege.2017.01>.

Miège, Bernard (2020), *La numérisation en cours de la société. Points de repères et enjeux*. Presses universitaires de Grenoble. <https://doi.org/10.3917/pug.miege.2020.01>.

Million, Anthony J.; York Jeremy; Lafia Sara ; Hemphill Libby (2025), « Data, Not Documents: Moving beyond Theories of Information-Seeking Behavior to Advance Data Discovery », *Journal of the Association for Information Science and Technology*, 76 (4), p. 649-64.
<https://doi.org/10.1002/asi.24962>.

Mirowski, Philip (2018), « The Future(s) of Open Science », *Social Studies of Science*, 48 (2), p. 171-203. <https://doi.org/10.1177/0306312718772086>.

Moore, Samuel A. (2025), *Publishing Beyond the Market: Open Access, Care, and the Commons*. University of Michigan Press. <https://doi.org/10.3998/mpub.11781635>.

Nicholas, David ; Eti Herman Eti ; Boukacem-Zeghmouri Cherifa ; *et al.* (2022), « Early career researchers in the pandemic-fashioned" new scholarly normality: voices from the research frontline », *Profesional de la información*, 31(5). <https://doi.org/10.3145/epi.2022.sep.10>.

Nicholas, David ; Herman Eti ; Boukacem-Zeghmouri Cherifa ; *et al.* (2023), « Transforming Scholarly Communications: The Part Played by the Pandemic and the Contribution of Early Career Researchers », *Learned Publishing*, 36 (4), p. 492-505.
<https://doi.org/10.1002/leap.1576>.

Pasquetto, Irene V.; Borgman Christine L. ; Wofford Morgan F. (2019), « Uses and Reuses of Scientific Data: The Data Creators' Advantage », *Harvard Data Science Review*, 1 (2).
<https://doi.org/10.1162/99608f92.fc14bf2d>.

Pooley, Jeff (2022), « Surveillance Publishing », *The Journal of Electronic Publishing*, 25 (1).
<https://doi.org/10.3998/jep.1874>.

Rebouillat, Violaine (2021), « Le partage des données vu par les chercheurs : une approche par la valeur », *Les Enjeux de l'information et de la communication*, 221 (1), p. 35-53.
<https://doi.org/10.3917/enic.030.0035>.

Schöpfel, Joachim ; Rebouillat Violaine (2023), « Partage et valorisation des données de la recherche », ISTE Group. <https://www.istegroup.com/fr/produit/partage-et-valorisation-des-donnees-de-la-recherche/>.

Shapin, Steven (2014), *Une Histoire Sociale de La Vérité*. Éditions La Découverte.
https://www.editionsladecouverte.fr/une_histoire_sociale_de_la_verite-9782707173492.

Sparc Europe (2017), *A Snapshot of Open Data and Open Science Policies in Europe*.
https://sparceurope.org/wp-content/uploads/dlm_uploads/2017/12/SPARCEuropeDCC_OpenSciencePolicySnapshotOct2017.pdf.

Tenopir, Carol ; Dalton Elizabeth D. ; Allard Suzie ; *et al.* (2015), « Changes in Data Sharing and Data Reuse Practices and Perceptions among Scientists Worldwide », *PLOS ONE*, 10 (8), e0134826. <https://doi.org/10.1371/journal.pone.0134826>.

Tkacz, Nathaniel (2012), « From Open Source to Open Government : A Critique of Open Politics », *Ephemera: Theory and Politics in Organization*, Vol.12, No.4, p. 386-405. <http://www.ephemeraweb.org/journal/12-4/12-4tkacz.pdf>.

REMERCIEMENTS

Les auteurs souhaitent remercier vivement la coordinatrice du numéro pour les commentaires précieux fournis à la première version de l'article, et qui ont bénéficié à la précision des analyses proposées dans le texte.

S'ouvrir à la négociation des métadonnées : ateliers expérimentaux de cartographie des projets de recherche

Article inédit, mis en ligne le 23 septembre 2026

Raphaëlle Bats

Raphaëlle Bats est docteure en sociologie, co-responsable de l'URFIST de Bordeaux-Nouvelle Aquitaine, Université de Bordeaux, membre associée du MICA axe IDEM. Ses travaux, au croisement entre la sociologie et les sciences de l'information, sont centrés sur les sciences participatives, la médiation scientifique et les transformations des lieux de production et de diffusion des savoirs. Elle coordonne le projet ECODOC sur la circulation des savoirs experts et des savoirs expérientiels sur les territoires en transition, dans lequel elle mobilise le slam pour aborder par l'émotion la question de la lecture des données scientifiques par des non scientifiques.
raphaelle.bats@u-bordeaux.fr

Mathilde Garnier

Mathilde Garnier est chargée du projet de recherche Ecodoc à l'Université de Bordeaux et doctorante en sciences de l'information et de la communication au sein du laboratoire MICA (axe ATIIA) à l'Université Bordeaux Montaigne. Ses thèmes de recherche gravitent autour des bibliothèques, notamment autour des techno-imaginaires, des catalogues et de la circulation de l'information.
mathilde.garnier@u-bordeaux.fr

Plan de l'article

Résumé et mots clés en français

Title, abstract and keywords in english

Titulo, resumen y palabras clave en espanol

Introduction

Le récit scientifique et ses métadonnées

Le récit profane : vers l'élaboration de nouvelles métadonnées

Conclusion

Références bibliographiques

RÉSUMÉ

L'ouverture des données scientifiques, principalement organisée pour les pairs, peine à toucher les non-spécialistes alors que leur accès présente un enjeu démocratique fort. Cet article présente une expérimentation visant à rendre les données accessibles et appropriables par tou·tes à travers un travail diplomatique d'attention et d'écoute des besoins, des valeurs et des enjeux en matière de diffusion et de réception de l'information scientifique. Dans une démarche participative et lors de plusieurs types d'ateliers avec des citoyen·nes, différentes approches et lectures de documents produits dans des projets d'écologie forestière ont fait émerger des récits multiples et contextuels permettant de négocier les métadonnées nécessaires à décrire, partager et rendre réutilisables les données

scientifiques. Ce travail sensible de diplomatie permet de construire une cartographie de ces documents, outil d'analyse et de navigation, préalable à l'édition d'une base informationnelle d'un outil de médiation scientifique numérique, passerelle entre le monde académique et la société.

Mots clés

médiation scientifique, données de la recherche, science ouverte, approches sensibles, savoirs expérientiels

TITLE

Opening Up to Metadata Negotiation: Experimental Workshops on Research Project Mapping

Abstract

The opening of scientific data, primarily organised for peers, struggles to reach non-specialists, even though their access represents a strong democratic stake. This article presents an experiment aimed at making data accessible and appropriable by all through diplomatic work, attentive listening to the needs, values, and issues related to the dissemination and reception of scientific information. Through a participatory approach and various workshops with citizens, different interpretations and readings of documents produced in forest ecology research projects revealed multiple and contextual narratives. These narratives enabled the negotiation of the metadata needed to describe, share, and make scientific data reusable. This sensitive diplomatic work allows for the creation of a cartography of these documents, serving as both an analytical and navigational tool. It lays the groundwork for building an informational database for a digital science communication tool, a bridge between the academic world and society.

Keywords

science communication, research data, open science, sensitive approaches, experiential knowledge

TITULO

Abriéndose a la negociación de metadatos: talleres experimentales de cartografía de proyectos de investigación

Resumen

La apertura de los datos científicos, organizada principalmente para pares, tiene dificultades para llegar a los no especialistas, a pesar de que su acceso representa un fuerte desafío democrático. Este artículo presenta una experimentación que busca hacer que los datos sean accesibles y apropiables por todas las personas a través de un trabajo diplomático de atención y escucha de las necesidades, valores y desafíos en materia de difusión y recepción de la información científica. Mediante un enfoque participativo y varios tipos de talleres con ciudadanos, diferentes enfoques y lecturas de documentos producidos en proyectos de ecología forestal revelaron relatos múltiples y contextuales. Estos relatos permitieron negociar los metadatos necesarios para describir, compartir y hacer reutilizables los datos

científicos. Este trabajo sensible de diplomacia permite construire una cartografía de estos documentos, herramienta de análisis y navegación, que sirve como base para editar una base informacional de una herramienta digital de mediación científica, puente entre el mundo académico y la sociedad.

Palabras clave

divulgación científica, datos de investigación, ciencia abierta, enfoques sensibles, saberes experienciales

INTRODUCTION

Les milieux académiques et de la recherche sont à l'apogée de leurs appels à l'ouverture de la science en général, avec pour double objectif l'intégrité de la science et la facilitation du dialogue science et société (Mikou, 2023). De fait, les politiques en faveur de l'ouverture des données font particulièrement l'objet d'actions et d'accompagnements (Vitari et Leclercq-Vandelannoitte, 2024) : mise à disposition d'entrepôts, mise en conformité des données avec des principes et des normes, incitation à la rédaction de documents. Pourtant, ces initiatives sont davantage dirigées vers l'efficacité des organismes producteurs de sciences que vers la transparence et le transfert de connaissances envers les citoyens. Est-ce à dire que les données ne peuvent être appréhendées que selon leur utilité dans leur activité professionnelle ou comme des gages de valeur scientifique (Rebouillat, 2019) et non comme supports de médiation ? A quelles conditions faudrait-il repenser les données pour que leur ouverture soit la plus large possible ? Que faudrait-il ajouter dans un jeu de données, au-delà des données elles-mêmes, pour que les données soient appropriables par des personnes extérieures à la discipline. et appropriables par les citoyens pour leur permettre de prendre leur part dans le dialogue démocratique ?

Ces questions de médiation et d'accès aux citoyens sont l'objet central du projet de recherche présenté dans cet article qui vise à interroger le dialogue entre les savoirs scientifiques et profanes sur les territoires en transition climatique. Au cœur du projet réside une expérimentation autour d'un prototype, celui d'un outil de médiation numérique des documents et des données scientifiques produites par des projets de recherche en cours. La conception de cet outil numérique, appelé « Espace commun de documentation » ou Ecodoc, offre la possibilité de questionner les enjeux d'organisation et de découvrabilité de l'information et les choix en la matière qui manifestent l'attention portée aussi bien aux producteurs qu'aux récepteurs de l'information scientifique. Un travail diplomatique se joue alors dans ce temps de définition des axes de l'organisation de l'information où il s'agit d'anticiper les objectifs, valeurs, enjeux des différentes parties prenantes.

Ce travail de diplomatie et d'écoute a été mené sur la détermination des métadonnées, considérant que si la mise à disposition des documents et des données scientifiques auprès des pairs passe par un travail riche de description et de métadonnées, alors que l'ouverture des données aux citoyens devrait passer par un choix de métadonnées non seulement susceptibles d'informer ce public non spécialiste, mais encore susceptible de lui permettre de se réapproprier ces documents et données pour son exercice de citoyen. Dès lors, concevoir un outil numérique de médiation scientifique pour les non experts de la science rend nécessaire une forme de diplomatie de la donnée et de la métadonnée, pour assurer que les données et les documents scientifiques soient plus accessibles et favorisent la compréhension du projet, la manipulation des données pour les visualiser, une circulation dans les documents, etc.

Comme toute diplomatie, il convient dès lors de mettre en place les conditions d'écoute, la négociation entre les récits qui se font autour du projet de recherche et de la science et qui portent en eux les éléments de définition des métadonnées. Cela passe par une capacité à

écouter les récits. En d'autres termes, la diplomatie des métadonnées est affaire de relation, dans sa double dimension : qui raconte, relate et qui relie. La métadonnée, comme la donnée, « se construit autour du sens de ce qui sert de base pour élaborer un raisonnement ou un argumentaire » (Smolczewska Tona, 2021, p.141). La donnée sert de point d'appui pour construire un discours ; les métadonnées relatent par les liaisons. Partant de ce principe, pour trouver les métadonnées les plus adaptées à une médiation d'un projet scientifique à destination des citoyens, il convient de s'attacher à mettre à jour les jeux de narration, de récit et de reformulation du discours scientifique (Pétroff, 1984) afin de pouvoir faire émerger un cadre de négociation, entendu comme un support pour une diplomatie des données (Boyd, 2019), qui facilite la construction d'outils de médiation scientifique grâce à une caractérisation élargie des données produites. Cela nécessite de s'intéresser autant au récit scientifique porté par les données et par les documents produits par les chercheurs qu'au récit produit par ceux et celles qui lisent ces supports narratifs. Par ailleurs, considérant avec Brigitte Simonnot que la médiation des données appelle à prendre en compte « les conditions d'intelligibilité des formes éditoriales s'appuyant sur le traitement des données, leur interprétation en contexte situé, les difficultés d'usages, les questionnements critiques et les discours d'accompagnement, ou encore les retours des publics qu'elles suscitent » (Simonnot *et al.*, 2021, p.10), la recherche des métadonnées rend nécessaire d'attribuer une place centrale aux récepteurs, à leurs contextes de lecture et à leurs besoins et à leurs préoccupations.

La méthodologie se présente elle-même comme un temps de médiation « symétrique » (Eastes, 2021), d'exploration des données scientifiques par les non spécialistes pour en faire émerger de manière participative des descriptions attentives à leurs conditions et à leurs situations. Pour cela, la narration et la participation à la description des données scientifiques ont été étudiées en deux temps. Le premier temps de négociation a consisté à recueillir le récit de chercheurs en écologie forestière racontant leurs projets de recherche et à collecter, puis à décrire les documents et données produits. Le second temps a consisté à expérimenter des ateliers avec des citoyens pour leur permettre de se saisir des documents scientifiques et pour produire à leur tour des éléments permettant d'entrer dans la négociation sur les métadonnées. Cette diplomatie en deux temps offre la possibilité de produire des récits des projets scientifiques (relation-narration), puis, d'inférer de ces récits des métadonnées (relation-description).

La dernière étape de la méthodologie a été la constitution d'une cartographie des projets de recherche à partir des systèmes de relations mis en évidence par les métadonnées possibles pour chaque document. Cette cartographie, outil d'analyse et outil de navigation, repose sur les différentes réalités dans lesquelles le projet est saisi, et sert pour concevoir un outil de médiation scientifique qui soit un objet sensible, à partir duquel peut être travaillée une esthétique de la médiation scientifique : « un mode d'articulation entre des manières de faire, des formes de visibilité de ces manières de faire et des modes de pensabilité de leurs rapports, impliquant une certaine idée de l'effectivité de la pensée » (Rancière, 2000, p.10).

En interrogeant à la fois les conditions de la négociation des métadonnées (les ateliers) et leurs résultats (les nouvelles métadonnées), cet article présente des pistes pour étendre l'ouverture de la science et pour faciliter la saisie de celle-ci en contextes démocratiques fragiles.

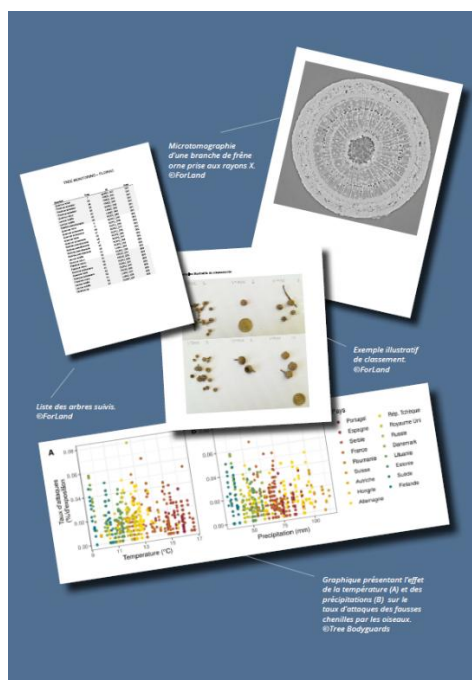
LE RÉCIT DES SCIENTIFIQUES ET SES MÉTADONNÉES

Dans la perspective de l'ouverture de la science, la première étape du projet s'est centrée sur l'ouverture entendue comme possibilité de mettre à disposition un jeu de documents permettant d'avoir accès à ce que produit et par là-même raconte la science. Or, les publications ouvertes sont habituellement liées à des projets terminés, susceptibles d'avoir

des résultats publiables. Dans le cadre du projet de recherche détaillé ici, il semblait important d'adresser la question de l'accès aux documents de la science dans le contexte de projets scientifiques en cours et encore à l'état de collecter et traiter des données, sans forcément avoir déjà des interprétations à publier, ce qui est d'autant plus nécessaire que ces recherches portent sur des questions de transition climatique et de transformation des territoires.

Ainsi, trois chercheurs en écologie forestière ont été invités à raconter leurs projets lors d'entretiens et à nous fournir au fil de leur récit les documents nécessaires à sa compréhension. Ces entretiens ont été conduits entre juillet 2023 et janvier 2024, à raison de trois entretiens pour chaque chercheur : un entretien initial en juillet 2023 pour identifier leur projet de recherche, un deuxième entretien en octobre consacré à la collecte des matériaux produits pendant leurs recherches et enfin un entretien final entre décembre 2023 et janvier 2024 pour récolter des informations et des données manquantes. Lors de ces entretiens, s'est fait entendre un récit du projet de recherche, de ses étapes de production et de ses médiations. Pour chaque document fourni pendant l'entretien, les chercheurs ont précisé le contexte de production, ses enjeux, ses auteurs, son accessibilité, ses destinataires, etc.

L'ensemble des documents réunis, qui sont compris ici comme des données, manifestent et illustrent le processus scientifique et son récit. Ainsi le projet de recherche « stress hydrique » se racontait avec une carte de la forêt étudiée, des photos des capteurs, une définition de l'embolie, des microtomographies, des graphiques, etc. La « fructification des chênes » s'ouvrait sur des photos des stades de développement des glands, des tableurs de notation des glands et de leurs stades observés sur une année, du code R utilisé pour modéliser ces données, etc. La « phénologie foliaire » avait pour particularité de montrer des photos de la forêt prises par drone mois après mois, quant au projet sur « les prédateurs des prédateurs des chênes » il présentait un extrait du protocole en BD, des photos de chenille, des listings de chants d'oiseaux, des modélisations de l'écoute des oiseaux, le résumé d'un article, etc.



Légende 1 : exemples de documents produits dans les projets de recherche étudiés. Extrait de Lectures sensibles de la science, Bats, Garnier, 2025.

Pour ouvrir l'accès à ces documents et pouvoir penser un outil de médiation de ces projets de recherche (un Ecodoc), il a été nécessaire de les décrire par un travail de définition de métadonnées, afin de pouvoir mettre en relation les documents entre eux et accompagner le récit par une architecture d'information solide. Ce travail de description est appelé ici cartographie. Elle est envisagée comme un outil qui donne la possibilité de déambuler à travers les données du récit. Elle résulte d'une stratégie classificatoire qui s'appuie sur une base de contenus, à travers l'analyse des différents éléments composants le système à décrire, plutôt que qu'une stratégie qui s'appuie sur le haut du système (Broughton, 2017). Le résultat du travail de cartographie s'apparente à une classification à facettes ou à une classification dite analytico-synthétique. La cartographie s'inscrit dans un travail de répertoriage et d'analyse des caractéristiques qui permettent de décrire précisément les documents ou les données du récit et s'inscrit dans la volonté de les ancrer épistémologiquement.

Ce travail de cartographie est déjà une négociation entre les chercheurs, écologues, qui ont raconté leur récit de recherche et les chercheuses, en sciences de l'information et de la communication (SIC), qui l'écoutent et le décrivent. Les métadonnées sont de fait comme les données elles-mêmes « construites socialement, techniquement, économiquement et situées spatialement et culturellement » (Ibekwe-Sanjuan, 2017, p.1). La lecture apposée sur le récit des écologues, en d'autres termes le travail de description et de codage de ces documents, témoigne des regards scientifiques exercés par deux chercheuses, autrices de cet article.

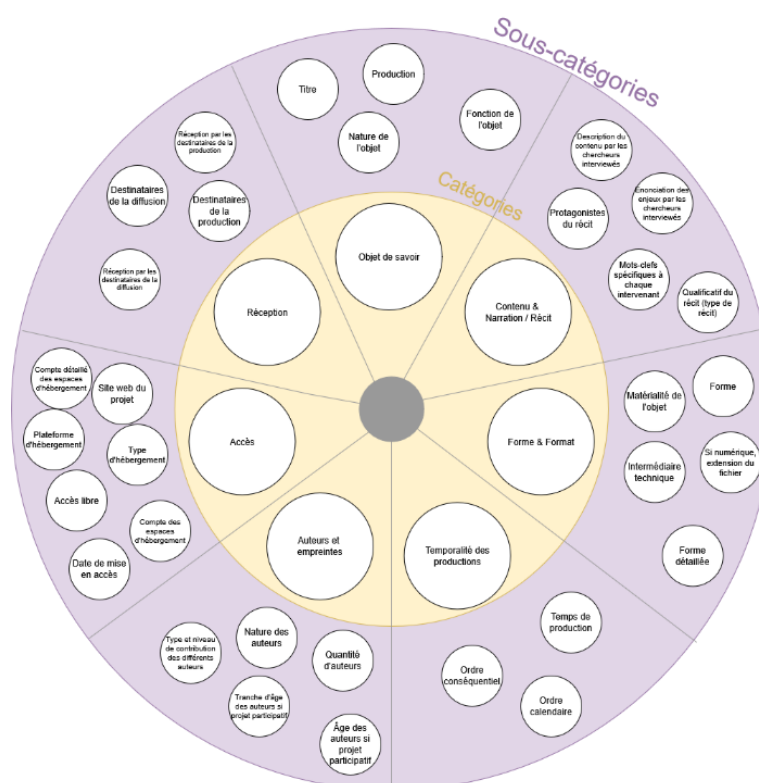
Cette lecture repose sur trois approches théoriques ; d'une part, une attention aux temps et aux étapes de la recherche en mobilisant la notion de cascades de traduction (Latour, 1996 ; Callon, 2013), pour obtenir des métadonnées pouvant témoigner du processus scientifique à l'œuvre ; d'autre part, une attention aux enjeux d'énonciation du récit scientifique (Haraway, 2020) pour obtenir des métadonnées pouvant témoigner de l'auteur du récit et de ses empreintes ; enfin, une attention aux traces (Serres, 2022) des documents pour penser leur accessibilité et leur mise à disposition. La cartographie de leurs projets de recherche, s'est manifestée *in fine* sous forme d'un tableau synoptique et de sa visualisation, organisé en catégories (7) et sous-catégories (44), qui structurent les données descriptives afin de permettre de les croiser et de les rendre potentiellement interrogeables dans un outil de navigation de l'information.

Catégorie	Sous-catégories						
Objet de savoir	Titre	Production	Nature de l'objet	Fonction de l'objet			
Contenu & Narration / Récit	Descriptif du contenu par les chercheurs interviewés	Enonciation des enjeux par les chercheurs interviewés	Qualificatif du récit (type de récit)	Protagonistes du récit	Mots-clefs spécifiques à chaque intervenant		
Forme & Format	Forme	Forme détaillée	Matérialité de l'objet	Si numérique, extension du fichier	Intermédiaire technique		
Temporalité des productions	Temps de production	Ordre calendaire	Ordre conséquentiel				
Auteurs et empreintes	Âge des auteurs si projet participatif	Âge des auteurs si projet participatif	Nature des auteurs	Quantité d'auteurs	Type et niveau de contribution des différents auteurs		
Accès	Accès libre	Site web du projet	Type d'hébergement	Plateforme d'hébergement	Compte des espaces d'hébergement	Compte détaillé des espaces d'hébergement	Date de mise en accès
Réception	Destinataires de la production	Réception par ces destinataires	Destinataire de la diffusion	Réception (destinataires de la diffusion)			

Légende 2 : tableau synoptique de structuration des données descriptives.

Ce travail de cartographie est une étape nécessaire pour construire des navigations entre les documents par la suite. Il devient alors possible de lire ce récit selon des directions différentes : le processus scientifique (temporalités diverses, nature et fonction), l'histoire et ses protagonistes (contenu, protagonistes, auteurs), le cadre de la réception (accessibilité, destinataires), etc. En appliquant ce filtre théorique de traitement pour définir les métadonnées, il s'est opéré des choix de facettes à faire apparaître et des mises en avant des éléments et des manières de raconter le projet qui échappaient à leurs producteurs ; ainsi et pour exemple la possibilité d'étudier l'accès aux documents produits dans des recherches participatives (Bats et Garnier, 2025B). Ces facettes ne font pas disparaître l'énonciation des écologues car les descripteurs des catégories et sous-catégories reposent sur les récits des chercheurs, et sont parfois gardés en texte libre.

La cartographie finale relève vraiment d'une négociation entre chercheurs, en écologie et en SIC, et permet de voir que l'apposition de métadonnées sur les données de recherche par des personnes extérieures au projet permet d'en proposer de nouvelles lectures. Les nouvelles lectures en question sont ici des lectures scientifiques pouvant donner lieu à de nouvelles recherches scientifiques dans des disciplines distinctes. Qu'il s'agisse de nourrir un entrepôt de recherche ou un outil de médiation scientifique numérique, tel que celui envisagé au terme de cette recherche appliquée, la main qui choisit les métadonnées définit les sens de lecture. L'enjeu d'ouverture de la science n'est pas uniquement l'accès au sens de mise à disposition, mais bien de penser les conditions par lesquels cet accès sera inclusif et prendra en compte d'autres lectures possibles. Cela impliquerait donc d'étendre cette négociation autour des métadonnées, cette attention diplomatique à la définition des métadonnées à un réseau de producteurs de métadonnées plus large que celui des scientifiques, et par conséquent, aussi, moins expert.



Légende 3 : visualisation des métadonnées (catégories et sous-catégories) définies suite à la première négociation des métadonnées.

LE RÉCIT PROFANE : VERS L'ÉLABORATION DE NOUVELLES METADONNÉES

L'ouverture de la science ne consiste pas seulement dans l'ouverture des documents, elle est ouverture à l'écoute des autres intelligibilités et à la participation des non experts (Unesco, 2022). Pour que la recherche d'une plus grande accessibilité des documents scientifiques soit complète, il fallait donc s'intéresser à la possibilité d'intégrer des temps de négociation des métadonnées avec des personnes éloignées des champs d'expertises des chercheur.es. Ils sont appelés ici récits profanes de la science au sens où ils sont produits par des personnes qui ne sont pas initiées à ce qui ici est le cœur du projet scientifique, et donc ne sont initiées ni à l'écologie forestière, ni à la beauté des métadonnées, ni à la démarche scientifique en général.

Enfin, en invitant des profanes à produire un récit sur la science, le pari est de voir émerger des enjeux de descriptions qui relèvent non pas de la science, mais qui relèvent de leurs manières d'observer le monde et reposent sur leurs savoirs expérientiels. Claude Levi-Strauss rappelle que « comme dans les langues de métier, la prolifération conceptuelle correspond à une attention plus soutenue envers les propriétés du réel, à un intérêt mis en éveil sur les distinctions qu'on peut y introduire » (Levi-Strauss, 1962, p.13). En d'autres termes, pour les sciences, certaines descriptions, métadonnées, permettent d'exprimer le caractère scientifique du projet. Pour les non scientifiques, ces niveaux de description ne présentent peut-être pas d'intérêt, mais d'autres en revanche s'avèrent utiles pour eux, car ils ne décrivent pas le même réel, mais marquent des « intérêts inégalement marqués et détaillés de chaque société particulière au sein de la société nationale » (Levi-Strauss, 1962, p.12).

Pour produire des récits profanes et pour toucher ce réel, des ateliers de lecture des documents collectés lors des entretiens avec les écologues ont été organisés. Ces ateliers se présentent comme une « recherche-médiation » ; à la fois outils méthodologiques, permettant d'observer ce qui se joue dans la lecture de ces documents, et outils de médiation scientifique, permettant aux participants d'aller à la rencontre de projets de recherche menés sur leurs territoires. Il s'agit de faire émerger des lectures et des réécritures du récit scientifique permettant chacun à sa manière de recomposer le réel scientifique en un réel singulier (Bats, Garnier, 2025), propice à faire émerger des caractérisations de ce réel traduisibles en métadonnées. Pour chacun des ateliers, les participants disposaient donc d'un dossier documentaire composé des impressions des différents documents fournis par les chercheurs pour raconter leurs projets, organisés en suivant la temporalité conséquentielle de production de ces objets : contexte, protocole, méthode, résultats, analyse. Ils représentaient ainsi doublement un récit scientifique du projet, au sens où il racontait le projet de recherche à travers les documents qu'il produit et où ce projet est raconté selon un angle de lecture qui est celui des chercheuses l'ayant décrit. Trois types d'ateliers ont été expérimentés pour faire émerger ces réels profanes et permettre des négociations autour des métadonnées.

« Raconter la science » : le réel organisé

Le premier atelier répondait à la nécessité de proposer une lecture parallèle à celle exercée par les deux chercheuses qui était très orientée sur l'organisation des savoirs. Aussi, l'atelier « Raconter la science » consistait à produire une organisation des documents susceptible de rendre la recherche en écologie présentée la plus intelligible possible pour le grand public. Cet atelier a fait l'objet de deux sessions, l'une en 2023 et l'autre en 2024, l'une auprès du comité de pilotage du projet (composé de chercheurs dans des disciplines variées, de médiateurs scientifiques et culturels, de bibliothécaires, d'artistes) et l'autre auprès d'un groupe de stagiaires d'une formation à la médiation scientifique en bibliothèque.

Pendant ces sessions, sept réorganisations des documents ou nouvelles lectures du récit scientifique ont été conçues. Le projet se voit alors être présenté : sous forme de calques, organisés selon une logique temporelle ; sous forme de fils à tirer, organisés en partant de

la forêt vers des arbres spécifiques ; sous forme d'un jeu de cartes, non organisé pour laisser libre cours à l'imagination ; sous forme d'un outil multimédia, organisé selon le processus scientifique depuis le protocole jusqu'aux résultats ; sous forme d'un livre accordéon, organisé autour de l'objet d'étude (le gland par exemple) et ses différentes manifestations dans le projet de recherche ; sous forme d'une frise chronologique, organisée à partir du travail hebdomadaire du chercheur ; et enfin, sous forme de fenêtre à faire coulisser (zoom), organisée de en vue d'approcher la donnée depuis sa forme la plus brute à sa forme la plus analysée.



Légende 4 : trois exemples de réorganisations des documents scientifique (de gauche à droite : calque, fil et vrac). Extrait de *Lectures sensibles de la science*, Bats, Garnier, 2025.

Ces sept nouvelles organisations des documents ont mis en évidence la demande d'explicitations, de profondeur, de compléments permettant de montrer ce que la recherche révèle étape par étape. Ces nouvelles lectures demandent à préciser le contenu de chaque document pour qu'il puisse être pensé en lien avec les autres, non pas de manière à comprendre où il se situe dans le récit dans la recherche, mais pour voir apparaître de nouvelles histoires, comme parallèles à celle-ci : celle de l'arbre, celle des saisons, etc. A travers les récits profanes d'organisation des documents ressortent ainsi de nouvelles descriptions à ajouter dans la catégorie « narration ». Ainsi, l'atelier a fait ressortir la nécessité de préciser le contenu par le lieu, la date, la population, l'environnement, les outils et instruments, etc. Il ne s'agit pas tant de nouvelles catégories de métadonnées que de nouvelles sous-catégories ou de précisions dans les sous-catégories (voir Légende 7). Il se joue ici une négociation dans la demande de précision et le contexte. Les nouveaux lecteurs de la recherche adressent des points d'attention nécessaires pour ouvrir la science et la rendre lisible, mais aussi des points de confiance.

La négociation est d'autant plus riche que les propositions d'organisation ont toutes été assorties par des propositions de visualisation, alors même que la demande était limitée au classement des documents. Chaque groupe a pris le parti de proposer aussi des manifestations plus sensibles de l'appréhension de la recherche. Chaque organisation des documents est aussi une manifestation physique de la relation. Le récit n'est plus linéaire, mais médiatisé par une structure visuelle. Les organisations proposées rendent préhensible les lectures ; elles en font des expériences sensorielles, des manifestations d'un réel organisé, scénarisé comme par la datavisualisation (Rubessi, 2021). Elles font ainsi partie « d'un horizon culturel et philosophique avec lequel nous imaginons et construisons des réalités » (Rubessi, 2021, p.176) et font de l'outil de médiation visé un outil de « réduction de la complexité qui peut s'édifier et s'accomplir alors même que le monde et la société où le phénomène a lieu sont inconnus » (Luhman, *Observation of Modernity*, 1998, cité par Rubessi, 2021, p.176). Cette manière de penser la structure et sa manifestation sensorielle est inestimable dans un projet d'ouverture de la science qui vise *in fine* à produire un outil de navigation à travers l'information scientifique dans une perspective de médiation scientifique.

« De l'arbre à la feuille » : le réel sensible

Le second atelier répondait à la nécessité de se confronter à des lectures encore plus éloignées des approches scientifiques et organisationnelles. Pour cela, des temps d'expression du récit de la science, qui reposent davantage sur des approches sensibles de la lecture des données du récit scientifique que sur des approches rationnelles, ont été expérimentés. Aussi l'atelier « De l'arbre à la feuille » consiste en un atelier de slam (poésie urbaine performée) de données scientifiques. Les participants étaient invités à écrire puis performer un texte poétique à partir des jeux de documents des recherches en écologie forestière. Pour l'étude, l'atelier a été organisé à 5 reprises, entre octobre 2023 et octobre 2024, avec 46 participants. 17 entretiens menés entre 3 et 6 mois après l'atelier. Ces entretiens ont donné la possibilité de revenir sur les textes écrits et leur signification, le rapport des participants à la science, aux forêts et au slam et les enjeux de médiation de la science. Le délai post atelier visait à leur laisser le temps d'« oublier » le contenu des dossiers documentaires, ce qui permet de mesurer la réception du récit scientifique dans le temps.

“
Bientôt plus le time
J'pourrais pas aller à Monaco
Modeler la vie c'était plus dur que prévu
J'croisais être enfin peinarde
Mais personne m'avait dit
Qu'il fallait se battre comme à Stalingrad
J'étais sûr d'être un prédateur
Il aurait pt'être fallut m'prévenir
Qu'on était tous des proies
J'pensais avoir atteint la fin
Mais ça ne s'arrêtera pas tant que le geai moqueur chantera
Réussir c'était qu'une idée zombie
La complexité des interactions en société
Ne fait que nourrir mon anxiété
Un chemin linéaire c'était c'que j'espérais
Mais j'erre entre mes choix
La théorie c'était easy
Mais la pratique ça fait chier
Seul au monde comme Chuck Noland
J'trouve pas ma place
J'grandis, mes feuilles s'épanouissent
Mais j'me fais bouffer par les chenilles
En définitive j'vis comme un fou dans un monde flou
Débordé par la haine et les changements
Je chante ma peine dans mes textes

Phylia ”

Légende 5 : un des textes de slam produits pendant les ateliers « De l'arbre à la feuille ». Extrait du Recueil de Science Slamée, Bats, Garnier, 2025.

La lecture par le biais du slam du récit scientifique fait apparaître des préoccupations intimes et singulières et chaque texte est l'expression poétique et sensible de ce que la science rencontre du réel de ses lecteurs. Les textes évoquent ainsi des inquiétudes liées à l'éducation, la société, la solidarité ou l'éco-anxiété. Or, ces regards très personnels sur le réel sont complexes à inférer en termes de métadonnées, car ils sont la manifestation d'une lecture unique et singulière (personne, temps, espace). Il semble difficile de décrire le projet sur les arbres et les prédateurs comme un projet parlant d'éco-anxiété ou de solidarité, quand bien même la lecture du projet engendre pour les lecteurs des questionnements sur l'éco-anxiété et la solidarité. Il faudrait pour cela identifier l'exhaustivité des émotions produites par la lecture d'un récit ; or, les catalogues, les dispositifs de médiation ne fonctionnent pas par l'émotion produite, mais par une rationalisation et une forme d'universalisation de la description. Les données socio-

émotionnelles nécessitent pour être définies en fonction de chaque situation d'interaction des expérimentations, des annotations, des analyses de *feedback* qui paraissent peu envisageables à mener pour chaque médiation de projet de recherche scientifique (Leder, 2024). Est-ce à dire que l'approche sensible des lectures profanes échappe totalement et représente de ce fait un aspect non négociable de la mise à disposition des données ?

Cela serait d'autant plus un problème que la transmission des savoirs n'est véritablement opérante qu'à condition de ne pas reposer entièrement sur des systèmes de conviction rationalisés, et de faire place aux sens et à l'émotion (Toomey, 2023 ; Berger et Bois, 2011). D'autant que la médiation est un « processus dynamique » susceptible de se transformer (Pinède, 2020), et devrait donc être reproduite et adaptée continuellement aux lieux, temps, usagers, outils, etc. Dans l'objectif de concevoir un outil de médiation numérique et de renouveler l'ouverture de ces documents scientifiques, cette situation interroge les interactions possibles pour la définition de nouvelles métadonnées en cours de lecture des données et des documents. On peut ainsi penser la donnée dans le temps et non pas comme le document contre l'événement (Smolczewska Tona, 2021). Les lectures qui les produisent sont liées à des réalités personnelles et donc variables. Il faut donc pouvoir à la fois entendre ces approches événementielles de la rencontre avec les données et documents scientifiques et dans le même temps proposer des métadonnées qui s'inscriront dans le temps et dans l'expérience de tous, pour avoir un outil de médiation numérique qui soit réutilisable à l'envie et en tous lieux.

Pour trouver des métadonnées qui rempliront cet enjeu de refléter l'événement de lecture tout en permettant une stabilisation de la description, une analyse textuelle des textes eux-mêmes a été menée. Il s'agissait d'identifier les mots des dossiers documentaires les plus utilisés dans les textes de slam ainsi que dans l'outil d'aide à l'écriture utilisé en atelier, à savoir le slamalauréat (Bats, StreetDefRecords, 2025).

Projet de recherche	Mots les plus utilisés dans les textes de slam ou dans le slamalauréat	Expression de description par les chercheurs	Termes communs
Fructification des chênes	Gland, avorton, reproduction, litière, chêne, fleur, restes, maturation, stade, petit, comptage	Enquête sur les capacités de reproduction de l'arbre pour comprendre l'effet du changement climatique sur celle-ci et découvrir les éléments de leur adaptation à ces transformations climatiques (phénologie). Dans ce projet, la reproduction de l'arbre s'étudie à travers la fructification des chênes (glands).	Gland, reproduction, chêne
Stress hydrique	Stress, hydrique, arbres, forêt, embolie, vaisseaux, feuilles, capteurs, chênes, canicule	Enquête sur le stress hydrique des arbres, à travers la présence d'eau dans les troncs, branches et feuilles pour étudier la capacité de résistance des arbres sur un territoire urbain en transition climatique	Stress hydrique, arbres, feuilles
Débourrement et sénescence	Feuilles, forêt, bourgeon, drone, feuillaison, foliaire, phénologie, sénescence, arbre, bourgeon, débourrement, début, espèce	Enquête sur les capacités de résistance de l'arbre au changement climatique à travers l'étude des dates d'apparition des feuilles (débourrement) et des dates de vieillissement des feuilles (sénescence). Ce domaine de recherche porte le nom de phénologie foliaire.	Arbre, phénologie, débourrement, feuilles, sénescence
Arbres et prédateurs	Oiseaux, chenilles, chênes, prédateur, diversité, prédation, modeler, chant, geai, climat, arbre	Étude des effets du changement climatique sur la diversité des espèces d'oiseaux et sur leur capacité de prédation à l'égard des insectes herbivores qui attaquent les chênes.	Oiseaux, prédation, chêne

Légende 6 : tableau des mots des dossiers documentaires les plus utilisés dans les textes de slam

Le croisement entre cette recherche de mots fréquemment utilisés et les entretiens ont permis de comprendre que ces mots ont retenu l'attention des participants, soit parce que leur sonorité ou leur sens leur ont plu, soit parce que leur présence est considérée comme clé du projet. Ces ensembles de mots peuvent servir de compléments pour proposer une recherche par mots matière du projet de recherche (voir l'image plus loin dans le texte : « Visualisation des métadonnées (catégories et sous-catégories) définies par la négation des données »). Ils complètent la description du projet qui avait été donnée par les chercheurs.

Pour exemple, la description initiale du projet sur les arbres et les prédateurs avait été décrite comme suit : « Étude des effets du changement climatique sur la diversité des espèces d'oiseaux et sur leur capacité de prédation à l'égard des insectes herbivores qui attaquent les chênes ». Les textes de slam font eux ressortir les mots : « oiseaux, chenilles, chênes, prédateur, diversité, prédation, modeler, chant, geai, climat, arbre ». La lecture par les profanes permet d'avoir un regard sur ce qui retient l'attention et qui s'éloigne potentiellement de ce qui pour les experts devrait retenir l'attention. Les mots retenus sont notamment parfois donnés dans une extension ou une précision sémantique : chêne-arbre, prédation-prédateur, oiseau-geai, etc., rendant possiblement les termes plus accessibles, non pas tant par souci de simplification que par souci d'incarnation.

Enfin, les mots choisis par les personnes ne présentent aucune « erreur » sur le projet ; au contraire, ils soulignent des aspects méthodologiques que les termes choisis par les chercheurs (écologues comme en science de l'information) ne montrent pas, allant là encore dans le sens d'une explicitation concrète du projet. Les termes retenus par les profanes donnent finalement plus de corps au projet que la description originale et originelle. Prendre soin de proposer des métadonnées qui rendent davantage compte du projet dans une dimension peut-être plus proche de réalités partagées autant par les participants que par les scientifiques que d'une réalité de recherche partagée par les scientifiques seulement. L'outil de médiation en cela se fait l'écho des démarches de médiation *Living Labs* qui permettent et rendent nécessaire à la fois « une double migration - conceptuelle et pratique - du cadran "environnements contrôlés/usagers passifs" vers le cadran "environnements réalistes/usagers actifs" » (Millet et Ducoulombier, 2021). En cela, ces nouvelles métadonnées reflètent le rapport du projet à la réalité de ceux qui la lisent et permettent de toucher au réel sensible de manière partageable dans un outil de médiation et d'accès aux documents et données scientifiques. Cela plaide dans le sens d'engager de tels ateliers poétiques pour la conception d'architecture de l'information, qu'il s'agisse de métadonnées ou non, mais aussi pour la mise à disposition des documents, dans la perspective d'une science diffusée déjà ouverte et non ouverte parce que diffusée.

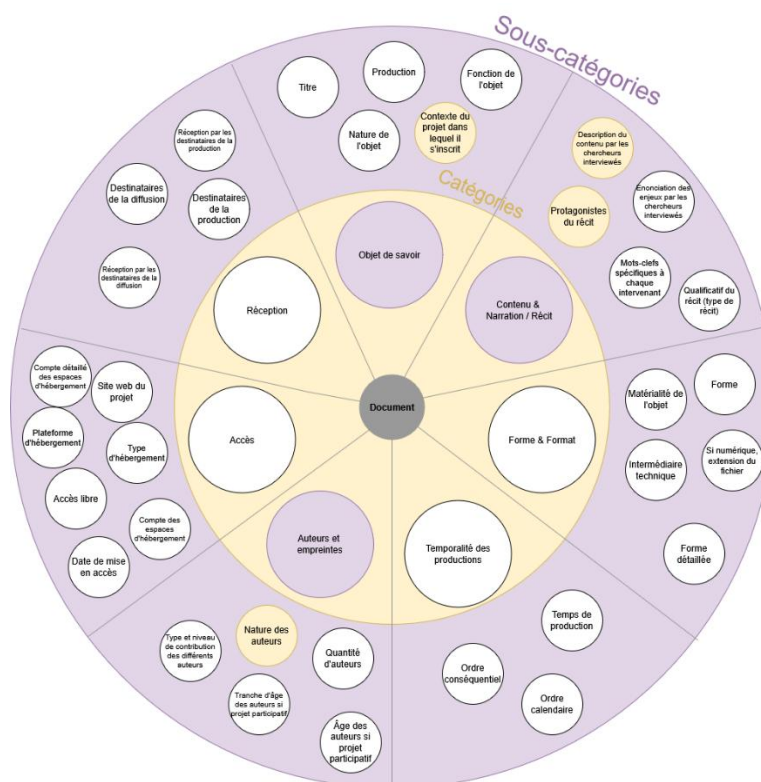
« Boussole » : le réel destiné

Le troisième atelier répondait à l'enjeu de lier la diplomatie des données avec une diplomatie des outils d'ouverture de la science. Pour cela, il fallait réinterroger l'accès aux documents et leur organisation des données au regard des préoccupations des citoyens en termes d'accès, et non plus simplement en termes de lectures. Aussi l'atelier « Boussole » avait pour objet de proposer aux participants de concevoir des outils de médiation des projets de recherche en écologie à partir des mêmes jeux de documents. Cinq sessions ont été organisées, en 2025, auprès de publics variés composés au final de 48 participants.

L'atelier « Boussole » est centré sur l'expression de préoccupations ou « concernés » (Mabi, 2014) au sujet de l'accès à l'information scientifique sur les forêts. L'atelier se présente donc comme à la fois médiation et résolution commune d'un problème qui est celui de l'accès aux projets de recherche, à leurs productions (documents et données) (Millet, 2023). À partir d'une analyse des limites des dossiers documentaires pour répondre à ces cailloux, les participants définissent des axes de mise en lecture du projet scientifique, axes susceptibles de mieux répondre à leurs préoccupations. Les ateliers « Boussole » ont permis de générer 18 propositions d'outils d'accès à l'information qui sont autant de lecture du projet scientifique. Le cadre d'analyse défini pour analyser ces propositions tient compte de la boussole en tant qu'objet (forme, support, matérialité), des publics visés, des problématiques de réception et d'accès auxquelles l'outil répond (cailloux), de ses usages possibles, de sa gestion et enfin de la scénarisation des documents qui y est proposée.

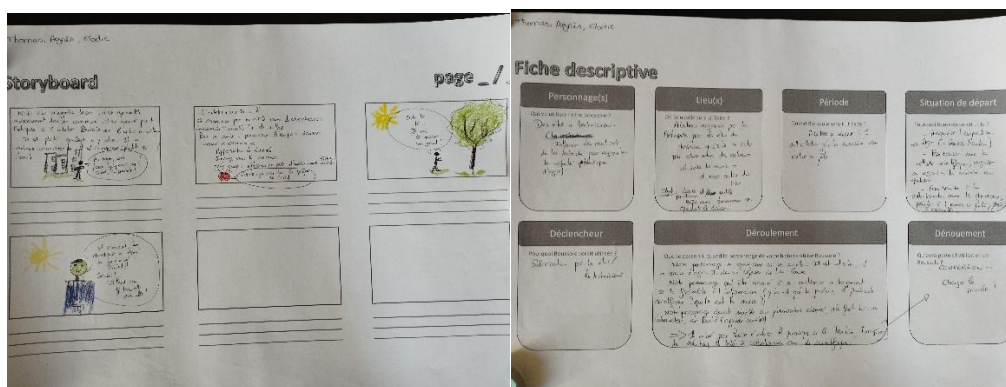
Il en ressort trois points principaux de préoccupations concernant l'accès aux documents et aux données scientifiques. Sont évoqués des enjeux d'accessibilité de l'information (adaptation aux handicaps, complexité de l'information, surabondance d'information, localisation des informations), de fiabilité des sources (pour l'information personnelle, en

cas de controverse, pour l'éducation des enfants, en cas de crise (incendie), pour pouvoir choisir ses actions) et enfin de praticité des informations (pour pouvoir agir : pour protéger la forêt, pour agir en cas d'incendie, pour protéger la biodiversité, pour anticiper les risques). En réponse à ces problématiques, les groupes proposent d'abord des réponses en termes de précision des documents par leur explicitation (graphiques, méthodologies, concepts et vocabulaire) et par l'expression plus claire du contexte de la recherche et de ses enjeux. Cela fait écho à la dimension informationnelle de la preuve, qui « *dans la tradition des pionniers de la documentation, [...] fait partie des fonctions attribuées au document dès sa définition* » (Perret, 2021, p.26). Dans le cas des ateliers « Boussole », au contraire, la preuve doit être ajoutée, le document ne suffit pas. Ce sont les métadonnées qui manquent pour assurer la fiabilité et la preuve, car celles-ci assurent la passerelle permettant de lire et de lier le document et la connaissance. Là encore des éléments doivent être ajoutés dans la cartographie initiale, notamment en termes de description des productions. L'image ci-dessous permet de visualiser les catégories et sous-catégories modifiées par les trois ateliers, dont l'atelier Boussole.



Légende 7 : Visualisation des métadonnées (catégories et sous-catégories) définies lors des deux temps de négociation des métadonnées. En couleurs inversées (en négatif), les champs modifiés à la suite des ateliers.

Par ailleurs, comme pour l'atelier « Raconter la science », la réflexion sur les métadonnées s'assortit d'une réflexion sur l'intelligibilité concrète et sensorielle de l'information. Quel jeu de langage faut-il alors ajouter au document pour que son indexabilité entre sa donnée et le connaissable qu'il rend possible soit plus évident, manifeste et intelligible (Perret, 2021, p.31-33) ? Pour approfondir cette intelligibilité, la deuxième partie de l'atelier consistait à concevoir l'outil de médiation lui-même, obligeant ainsi à décrire les axes par lesquels les données peuvent être rendues intelligibles, indexables, traduisibles.



Légende 8 : exemples de productions des groupes dans les ateliers Boussole, matériel du projet Ecodoc, atelier de mai 2025. (Ces productions ne sont qu'un aspect du travail mené et de sa restitution).

Sept catégories d'outils ont été conçus : expositions, guides, application mobile, site web, intelligence artificielle, capsules vidéo, kit pédagogique, qui toutes révèlent les attentes et les destinations envisagées pour la science et ses données du point de vue de profanes.

Boussole-type	Matérialité	Contenu	Objectif	Autre modalité de transmission
Application mobile participative	Numérique	Centré sur un ou des projets de recherche	Participatif	Forum/chat d'échange avec les chercheurs
Capsules	Numérique	Centré sur un ou des projets de recherche	Pédagogique	Scénarisation de l'information : vulgarisation
Exposition	Physique	A partir d'informations scientifiques produites dans des projets de recherche	Pédagogique	Système de QR codes pour atteindre des niveaux de complexités supérieurs au niveau présenté sur les panneaux
Guide	Numérique ou physique	Informations variées sur le thème, dont des informations scientifiques	Utilitaire (pour des guides, des naturalistes, etc.)	Utilisé dans des médiations directes avec du public
IA générative	Numérique	A partir d'informations scientifiques produites dans des projets de recherche	Informatif	Interactif
Kit pédagogique	Physique	Centré sur un ou des projets de recherche	Pédagogique	Accompagné par un kit de médiations ludiques et un carnet d'échange entre les différents utilisateurs du kit
Site web	Numérique	Informations variées sur le thème, dont des informations scientifiques	Informatif	Complété par des temps de jeux et de manipulation des données

Légende 9 : tableau des « Boussoles » catégorisées.

Ces propositions soulignent ainsi la place des projets de recherche et de leurs résultats : soit le projet est central et l'outil de médiation vise à le faire découvrir (et potentiellement à y participer), soit le projet est mobilisé comme une des sources au sein d'un ensemble d'informations susceptibles d'apporter les réponses recherchées et l'outil est centré sur les objectifs de la personne qui consulte et utilise le service. Les métadonnées de description du projet et des documents scientifiques devront donc être adaptées à la forme finale retenue pour l'outil de médiation et la place qu'elle confère aux projets de recherche. Dans les cas où le projet scientifique est central, le récit de la science tel que porté par les scientifiques s'avèrera l'axe autour duquel toutes les lectures vont se construire. Dans les cas où le projet scientifique disparaît au profit d'un ensemble d'informations fiables, mais de sources et de nature diverses, alors le récit scientifique devient un parcours possible de lecture, plutôt qu'un axe structurant toute la lecture. Dans l'expression des préoccupations et dans leur visualisation structurée, se rend visible la manifestation de réels destinés, c'est-à-dire de réalités pensées par leur utilisation et leur destination, plutôt que par l'expérience vécue. L'ouverture de la science se fait ici négociation non plus seulement dans la possibilité d'accéder à l'information, mais dans la reconnaissance de sa place et de son usage possible, *in fine* de son rôle dans la société et pour les citoyens.

CONCLUSION

Cette étude sur les récits scientifiques et profanes d'un projet de recherche avait pour objectif de montrer que la médiation de la science conduit à une multiplicité de lectures possibles d'un corpus de données scientifiques. Dès lors, un outil de médiation de l'information scientifique se manifeste comme le lieu d'une diplomatie des métadonnées, d'une valorisation par les chemins d'accès aux informations scientifiques du dialogue renouvelé entre scientifiques et citoyens. Les ateliers expérimentés ont permis d'observer les temps et les enjeux de la négociation : l'apparition de réels organisés, sensibles et destinés. En assurant un travail diplomatique d'expression et d'écoute des différents récits et des lectures qu'ils rendent possible, la médiation des données scientifiques prend une nouvelle dimension d'élaboration de jeux de relation par les choix de métadonnées. Ces ateliers sont dès lors reproductibles en tant qu'outils de médiation scientifique attentifs à la diversité des lectures de la science et en tant qu'outils de co-conception de services numériques pour faire émerger des architectures de l'information négociées.

Le travail de cartographie participatif des lectures possibles de la science qui s'appuie sur un travail sensible d'écoute et d'attention pour concevoir des métadonnées véritablement utiles aux citoyens ouvre ainsi plusieurs perspectives d'usages qui en ont en commun de consolider les passerelles entre les sciences et la société et de faciliter le rôle du citoyen en démocratie. A ce titre, prendre le temps d'une diplomatie des métadonnées revient à assumer l'entière responsabilité des enjeux de la science ouverte tels que définis par l'Unesco (2022), à savoir tout autant ouvrir la publication (ici des documents et des données), repenser les infrastructures (ici les métadonnées des documents et jeux de données partagés), faire participer les citoyens (ici dans la définition des métadonnées) que faire dialoguer le système académique avec les autres systèmes de connaissance (ici dans la diplomatie elle-même, via les ateliers expérimentés), et met en œuvre ainsi les valeurs au cœur de l'ouverture de la science, à savoir l'égalité d'accès, l'inclusion et la reconnaissance des individus, des communautés, des épistémologies et la collaboration pour produire des savoirs qui servent l'ensemble de la communauté.

RÉFÉRENCES BIBLIOGRAPHIQUES

Bats, Quentin ; StreetDef Records (2025), « Le Slamalauréat » (p.43), in Bats, Raphaëlle ; Garnier, Mathilde, *Lectures sensibles de la science, un guide de médiation de la recherche*.

Bats, Raphaëlle ; Garnier, Mathilde (2025A), *Lectures sensibles de la science : un guide de médiation de la recherche*, Bordeaux : Université de Bordeaux.

Bats, Raphaëlle ; Garnier, Mathilde (2025B), « The Traces of Participation: An Investigation into the Fate of Collectively Produced Knowledge » (p.173-191), in Pedro Sebastião, Sónia ; Cotton, Anne-Marie (dir.), *Global Public Goods Communication : Mapping Actors, Policies, and Narratives*. Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-90667-1_11

Berger, Eve ; Bois, Danis (2011), « Du Sensible au sens : un chemin d'autonomisation du sujet connaissant » (p. 117-124), in Lani-Bayle, Martine ; Bergier, Bertrand (dir.), *Quelle reconnaissance du sujet sensible en éducation ? Vol. n°16*. CERAP. <https://www.cerap.org/fr/paradigme-du-sensible/du-sensible-au-sens-un-chemin-d%E2%80%99autonomisation-du-sujet-connaissant>

Boyd, Andy ; Gatewood, Jane ; Thorson, Stuart J. ; Dye, Timothy (2019), « Data Diplomacy », *Science & Diplomacy*, [en ligne], consulté le 27 avril 2026. <https://www.sciencediplomacy.org/article/2019/data-diplomacy>

Broughton, Vanda (2017), « La classification à facettes comme théorie générale pour l'organisation des connaissances », *Les cahiers du numérique*, vol. 13, n°1, p. 25-47.

<https://doi.org/10.3166/lcn.13.1.25-47>

Callon, Michel (2013), « Quatre modèles pour décrire la dynamique de la science » (p. 201-251), in Akrich, Madeleine ; Latour, Bruno (dir.), *Sociologie de la traduction : textes fondateurs*, Paris : Presses des Mines. <https://doi.org/10.4000/books.pressesmines.1199>

Eastes, Richard-Emmanuel (2021), « Non, la médiation scientifique n'est pas politiquement neutre », in *Savoirs en société*, [en ligne], consulté le 27 avril 2026, <https://blogs.letemps.ch/richard-emmanuel-eastes/2021/06/17/non-la-mediation-scientifique-nest-pas-politiquement-neutre/>

Haraway, Donna Jeanne (2020), *Vivre avec le trouble*, Vaulx-en-Velin : Les Éditions des mondes à faire.

Ibekwe-Sanjuan, Fidelia (2017), « Vers la datafication de la société ? » (p. 31-49), in Meyer, Vincent (dir.), *Transition digitale, handicaps et travail social*, Bordeaux : LEH Editions. <https://hal.science/hal-01898457>

Latour, Bruno ; Woolgar, Steve (1996), *La vie de laboratoire : La production des faits scientifiques*, Paris : La Découverte.

Lévi-Strauss, Claude, (1962), *La pensée sauvage*, Paris : Plon.

Mabi, Clément (2014), « Comment se construit le « concernement » des publics de la démocratie dialogique ? Analyse des débats publics CNDP », *Canadian Journal of Communication*, vol. 39, n°4, p. 627-638. <https://doi.org/10.22230/cjc.2014v39n4a2737>

Mikou, Afaf (2023), « Science et Société, une relation en pleine mutation », *La Lettre de l'OCIM (Office de Coopération et d'Information Muséographique) : Musées, Patrimoine et Culture scientifiques et techniques*, vol. 206, p. 62-73. <https://journals.openedition.org/ocim/5436>

Millet, François (2023), « La donnée, objet et outil de la médiation scientifique : L'exemple du Dôme », *Cahiers de l'action*, vol. 60, n°1, p. 62-71. <https://doi.org/10.3917/cact.060.0062>

Millet, François ; Ducoulombier, Pauline (2021), « Living Lab de recherche et médiation scientifique : Une tentative d'innovation populaire », *VertigO - la revue électronique en sciences de l'environnement, Hors-série 34*. <https://doi.org/10.4000/vertigo.30249>

Perret, Arthur (2021), « Fonction documentaire de preuve et données numériques » *Information et stratégie*, (p. 25-36), in Simonnot, Brigitte ; Broudoux, Evelyne ; Chartron, Ghislaine (dir.), *Humains et données : création, médiation, décision, narration. Actes du colloque « Document numérique et société »*, Nancy, octobre 2020, De Boeck Supérieur. <https://doi.org/10.3917/dbu.simmo.2021.01>

Pétroff, André Jean (1984), « Sémiologie de la reformulation dans le discours scientifique et technique », *Langue française*, vol. 64, n°1, p. 53-67. <https://doi.org/10.3406/lfr.1984.5204>

Pinède, Nathalie (2020), « Les enjeux de la médiation des données à l'heure du numérique. L'exemple de la licence professionnelle mind », *Les cahiers de la SFSIC, 15-Varia*. <https://cahiers.sfsic.org/sfsic/index.php?id=152>

Rancière, Jacques (2000), *Le partage du sensible : Esthétique et politique*, Paris : La Fabrique.

Rebouillat, Violaine (2019), *Ouverture des données de la recherche : De la vision politique aux pratiques des chercheurs*, thèse de doctorat en Sciences de l'information et de la communication, Conservatoire national des arts et métiers, Paris. <https://theses.fr/2019CNAM1254>

Rubessi, Chiara (2021), « Visualisation de données et espace d'exposition : Enjeux et

pratiques » (p. 175-184), in Simonnot, Brigitte ; Broudoux, Evelyne ; Chartron, Ghislaine (dir.), *Humains et données : création, médiation, décision, narration. Actes du colloque « Document numérique et société »*, Nancy, octobre 2020, De Boeck Supérieur.

Serres, Alexandre (2012), « Problématiques de la trace à l'heure du numérique », *Sens-Dessous*, vol. 10, n°1, p. 84-94. <https://doi.org/10.3917/sdes.010.0084>

Simonnot, Brigitte ; Broudoux, Évelyne ; Chartron, Ghislaine (2021), Introduction (p. 7-12), in Simonnot, Brigitte ; Broudoux, Evelyne ; Chartron, Ghislaine (dir.), *Humains et données : création, médiation, décision, narration. Actes du colloque « Document numérique et société »*, Nancy, octobre 2020, De Boeck Supérieur. <https://doi.org/10.3917/dbu.simmo.2021.01>

Smolczewska Tona, Agnieszka (2021), « Enquête sur les représentations discursives des temporalités de la donnée à l'œuvre dans des articles scientifiques » (p. 137-150), in), in Simonnot, Brigitte ; Broudoux, Evelyne ; Chartron, Ghislaine (dir.), *Humains et données : création, médiation, décision, narration. Actes du colloque « Document numérique et société »*, Nancy, octobre 2020, De Boeck Supérieur. <https://doi.org/10.3917/dbu.simmo.2021.01>

Toomey, Anne Helen (2023), « Why facts don't change minds : Insights from cognitive science for the improved communication of conservation research », *Biological Conservation*, vol. 278. <https://doi.org/10.1016/j.biocon.2022.109886>

Trudel, Stéphane ; Martineau, Stéphane ; Buysse, Alexandre (2021), « Chapitre 1 Mémoire et Histoire : Des Choses Aux Mots » (p. 13-27), in *Mondes profanes*, Québec : Presses de l'Université Laval. <https://www.degruyterbrill.com/document/doi/10.1515/9782763744711-003/html>

Unesco (2022), « *Understanding open science* », *UNESCO open science toolkit*. <https://doi.org/10.54677/UTCD9302>

Vitari, Claudio ; Leclercq-Vandelannoitte, Aurélie (2024), « La science ouverte », *Systèmes d'information & management*, vol. 29, n°1, p. 3-5. <https://doi.org/10.3917/sim.241.0003>

Usage et ré-usage des données de la recherche : les intersections du compromis au prisme de la médiation des connaissances

Article inédit, mis en ligne le 23 septembre 2026

Sylvie Marciset-Sognos

Docteure en Sciences de l'information et de la communication, formatrice- documentaliste et responsable de la bibliothèque de l'ENSFEA (École Nationale Supérieure de Formation de l'Enseignement Agricole). Elle est membre du laboratoire Education, Formation, Travail, Savoirs (EFTS) de l'Université Toulouse, ENSFEA. Ses recherches portent sur les processus de médiation des savoirs à travers l'épistémologie de l'information et de la donnée. Elle développe une approche méthodologique fondée sur l'analyse qualitative de pratiques ordinaires et de recherches collaboratives orientées par la conception.
sylvie.sognos@ensfea.fr

Plan de l'article

Résumé et mots clés en français

Title, abstract and keywords in english

Titulo, resumen y palabras clave en espanol

Introduction

Les intersections du compromis au prisme de la médiation des connaissances

Les traces des compromis

Science ouverte, conflits et compromis

Conclusion

Références bibliographiques

Annexes

RÉSUMÉ

Cet article examine les enjeux scientifiques du partage et du ré-usage des données dans un contexte marqué par la « datafication » et les mutations de la communication scientifique numérique. La science ouverte, devenue priorité politique, met en tension souveraineté, gouvernance des données en soulevant des enjeux éthiques, économiques, politiques et épistémologiques. En mobilisant la notion de compromis chez Ricœur (1991) et la notion de médiation en Sciences de l'information et de la communication, l'analyse observe les intersections entre ordres de grandeur et figures de médiation. À partir de la littérature récente sur le ré-usage des données en SHS, elle expose les difficultés de leur gestion ouverte et questionne le « travail scientifique des données » (*data scholarship*) comme agencement complexe des relations entre données et recherche.

Mots clés

Données, compromis, médiation, usages sociaux de l'information, intersection

TITLE

Use and reuse of research data: intersections of compromise through the lens of knowledge mediation

Abstract

This article examines the scientific challenges of data sharing and reuse in a context characterised by ‘datafication’ and changes in digital scientific communication. Open science, now a political priority, creates tension between sovereignty and data governance by raising ethical, economic, political and epistemological issues. Drawing on Ricœur’s (1991) concept of compromise and the notion of mediation in Information and Communication Sciences, the analysis examines the intersections between orders of magnitude and forms of mediation. Drawing on recent literature on the re-use of data in the social sciences and humanities, it highlights the difficulties of open data management and examines ‘data scholarship’ as a complex arrangement of relationships between data and research.

Keywords

Data, compromise, mediation, social uses of information, intersection

TITULO

Uso y reutilización de los datos de la investigación : las intersecciones del compromiso desde la perspectiva de la mediación del conocimiento

Resumen

Este artículo examina los retos científicos que plantean el intercambio y la reutilización de datos en un contexto marcado por la «dataficación» y los cambios en la comunicación científica digital. La ciencia abierta, convertida en prioridad política, pone en tensión la soberanía y la gobernanza de los datos al plantear cuestiones éticas, económicas, políticas y epistemológicas. Recurriendo al concepto de compromiso de Ricœur (1991) y al de mediación en las Ciencias de la Información y la Comunicación, el análisis observa las intersecciones entre órdenes de magnitud y figuras de mediación. A partir de la literatura reciente sobre la reutilización de datos en las ciencias sociales y humanas, expone las dificultades de su gestión abierta y cuestiona el «trabajo científico con datos» (data scholarship) como un complejo entramado de relaciones entre datos e investigación.

Palabras clave

Datos, compromiso, mediación, usos sociales de la información, intersección

INTRODUCTION

Cette recherche exploratoire porte sur les enjeux scientifiques liés au partage des données dans un contexte de recomposition du paysage de l’information scientifique, de « datafication » (Pontille, 2017) et de profondes mutations de la communication scientifique numérique (Boukacem-Zeghmouri *et al.*, 2021). Depuis que la science ouverte est devenue

une priorité de l'État (Loi pour une République Numérique, 2016), l'objectif annoncé est l'ouverture et le ré-usage des données, soulevant une multiplicité d'enjeux en termes de souveraineté et de gouvernance de ces données. En effet, les enjeux politiques de production et de circulation des connaissances pressent le contexte d'édition scientifique entre ouverture de la science et ses déclinaisons politiques et dépôt des publications autant que possible avec leurs données, algorithmes et codes sources.

Le libre accès aux données est ainsi largement encouragé par la politique scientifique prônant, au nom de l'intérêt public, la transparence et la réutilisation des données. Pour cela, non seulement celles-ci doivent être FAIR, c'est-à-dire, faciles à trouver, accessibles, interopérables et réutilisables, mais également fiables et vérifiables. De fait, « les données de recherche ne sont précieuses, interprétables et vérifiables que dans le contexte de leur origine et avec une infrastructure suffisante pour faciliter leur réutilisation » (Borgman et Brand, 2024, p. 2). Imposée par les politiques des organismes de financement et les directives des éditeurs, ouverture et dépôt des données associées aux publications peuvent ainsi tenir un rôle important dans la démocratisation et l'accélération de la recherche.

Or, le fait d'ouvrir, c'est-à-dire de mettre à la portée de tous, et de ré-utiliser des données soulève des questions d'ordre éthique, économique, politique, mais surtout épistémologique. L'utilisation comme la réutilisation des données varie considérablement selon les disciplines universitaires, le contexte et le temps (Borgman, 2015 ; Pasqueto *et al.*, 2017). Les transcriptions d'entretiens ne se structurent pas de la même manière que des observations astronomiques par exemple, soulevant ainsi la question de la standardisation des formats de production des données, l'interopérabilité des outils logiciels de leur analyse et surtout la capacité initiale à mettre à disposition ces données en libre accès. Ces opérations ont des coûts très élevés, selon des modèles économiques encore imprécis et si les objectifs des politiques publiques en matière d'ouverture des données de recherche semblent clairs, leurs conséquences pour les acteurs et parties prenantes le semblent moins.

Cela mobilise selon Paul Ricoeur (1991) différents ordres de grandeur. La notion de grandeur renvoie aux formes du bien commun légitimes selon Luc Boltanski et Laurent Thévenot au sens où « les sciences sociales sont confrontées à des acteurs individuels dans des situations concrètes de justification face à des logiques différentes de valeur » (Boltanski et Thévenot, 2022, p. 337). Ces ordres de grandeur sont parfois en conflit et donc soulèvent des questions en termes de partage et de médiation des connaissances. Ces conflits sont observables à travers les intersections des ordres de grandeur qu'ils traversent et qu'ils sectionnent. De fait, l'intersection « est tantôt pensée comme un point d'équilibre de dernier recours qui empêche la situation de dégénérer en conflit, tantôt comme le résultat des concessions mutuelles qui permettent d'arriver à une collaboration, tantôt comme une forme de régulation conjointe » (Assayag-Gillot, 2018, p. 110).

Mais comment caractériser ce lien entre conflit et partage de connaissances ? Entre usage des données dans une recherche et leur ré-usage, comment ces données s'actualisent-elles à travers des cadres théoriques, des méthodologies, des ordres de grandeur différents ? La « fairisation » des données sert-elle effectivement au partage de connaissances fondateur de leur potentielle ouverture ? Si oui, à travers quelles formes de médiation, quels dispositifs, quels usages ? En quoi une approche diplomatique, au sens d'une observation des enjeux de pouvoirs sur le contrôle de la production, de la circulation et de l'exploitation des données scientifiques, dans le cadre de la science ouverte est-elle pertinente ?

Nous nous appuyons d'une part sur la notion de médiation en Sciences de l'information et de la communication dans ses dimensions langagière, logistique, technique et symbolique (Gardiès, 2012), sur ses figures sociales, réflexives et procédurales (Jeanneret, 2009) ; d'autre part, sur la notion de compromis (Ricoeur, 1991). Partant d'une analyse de la littérature récente sur le ré-usage des données en Sciences humaines et sociales (SHS), pointant les difficultés à mettre en place une véritable gestion ouverte des données (Boukacem-Zeghmouri, 2024), nous proposons d'observer ces intersections et ces figures du

compromis. Nous les confronterons ensuite aux figures de la médiation identifiées (Marciset-Sognos, Canizares, Gardiès, 2024), dans la perspective de questionner ce « travail scientifique des données » (*data scholarship*) dont parle Christine L. Borgman (2020) pour exprimer l'agencement complexe des relations entre données et recherche.

LES INTERSECTIONS DU COMPROMIS AU PRISME DE LA MEDIATION DES CONNAISSANCES

Donnée, usage, ré-usage

La donnée « est un fait supposé qui procède d'une différence ou d'un manque d'uniformité dans un contexte » (Floridi, 2022, cité dans Leleu-Merviel, 2017, p. 68-69). Les données de la recherche sont habituellement définies comme des enregistrements factuels (chiffres, textes, images et sons), utilisés comme sources principales pour la recherche scientifique et généralement reconnus par la science comme nécessaires pour valider des résultats de recherche (OCDE). Cette définition fait l'objet de discussions dans les SHS. Si les données y constituent des informations-sources mobilisées ou collectées, qui sous-tendent l'analyse et la production des résultats de recherche, elles sont avant tout « produites » par le regard et le questionnement scientifique du chercheur, ainsi que par les procédures d'enquête qu'il met en œuvre. Cette conception, selon laquelle les données ne sont pas « immédiatement données » mais « produites » n'enlève en rien à la valeur qu'elles prennent dans le travail des chercheurs. Bien au contraire, le fait que les données soient « produites » et puissent être considérées comme constitutives d'une recherche, leur confère une valeur intrinsèque et pose la question de leur partage et de leur conservation. Si la valeur des données fait pour l'instant l'objet d'une faible reconnaissance institutionnelle, les enquêtes menées auprès des chercheurs révèlent leur attachement aux données produites ; cet attachement se mesurant par exemple à partir de l'angoisse que suscite « l'idée de la perte ou du vol des données » (Huyghe, 2018, p. 2).

Les Sciences de l'information et de la communication (SIC) abordent les données en tenant compte de leur complexité sociale et technique, les inscrivent dans le triptyque donnée-information-connaissance, et s'appuient sur des travaux comme ceux de Marcia J. Bates (2005) qui définit la donnée comme une entité brute, n'acquérant de sens qu'au sein d'un processus de production d'information. Les données de recherche ne sont pas de simples objets techniques mais sont produites et façonnées par des outils numériques et des structures sociales. La dimension sociale et organisationnelle des données montre que leurs production et circulation sont indissociables des contextes institutionnels et technologiques. Les SIC permettent ainsi de penser la donnée comme un objet complexe, situé dans des environnements numériques et sociaux, non comme une entité neutre. Les données n'existent pas à l'état « pur », elles sont toujours affaire de mélanges, de bricolages, d'accommodements et d'agencements hybrides : objet et résultat d'un travail (Goody, 1979). Il est donc important d'analyser le travail invisible de production des données (Star, 1999) pour comprendre ce qui le constitue mais aussi les ressorts de son invisibilité (Star et Strauss, 1999).

Dans ce cadre, les SIC s'intéressent aux processus de médiation et de circulation qui transforment les données en information, puis en connaissance, en tenant compte des pratiques sociales et des outils numériques. Par ailleurs, certains auteurs comme Marie Huyghe (2018) défendent l'idée qu'il n'existe pas de données strictement qualitatives. Ce sont les démarches (interprétation, compréhension des phénomènes sociaux) qui confèrent au matériau son statut.

La question de l'usage et de la réutilisation des données de recherche occupe bien une place centrale dans les débats contemporains sur la science ouverte. Si la réutilisation des données présente de nombreux bénéfices, elle favorise la transparence des travaux scientifiques,

permet la reproductibilité des résultats et stimule l'innovation, encourageant nouvelles analyses ou croisements interdisciplinaires, ces avantages s'accompagnent de coûts importants, principalement assumés par les producteurs de données. Ceux-ci doivent investir du temps, mobiliser une expertise spécifique et mettre en place des infrastructures adaptées pour assurer qualité, documentation et mise à disposition des données (Borgman, 2024, p. 2-4, 7).

La réutilisation comporte également des risques : perte de contrôle sur les données, mésusage ou interprétations erronées, absence d'attribution correcte, voire atteintes à la propriété intellectuelle. Ces enjeux soulignent la nécessité d'un cadre de gouvernance solide. Les facteurs de succès identifiés reposent sur l'existence de standards partagés, d'une documentation riche et précise, d'infrastructures techniques fiables, ainsi que sur la reconnaissance institutionnelle et académique des efforts consentis par les producteurs.

La tendance actuelle montre que la réutilisation tend à devenir une exigence croissante dans la publication scientifique. Toutefois, ce processus demeure complexe, car il dépend de nombreux facteurs sociaux, techniques et organisationnels. C. L. Borgman (2025, p. 6-17) identifie à ce titre six dimensions de la distance qui influencent la facilité ou la difficulté de réutilisation : le domaine scientifique, les méthodes employées, les modalités de collaboration, les pratiques de conservation, les objectifs poursuivis et le facteur temps. Pour réduire ces distances, certaines conditions apparaissent essentielles : documentation rigoureuse, normalisation des formats, collaboration entre acteurs et conservation pérenne des données. Enfin, C. L. Borgman insiste sur le rôle des politiques et infrastructures qui doivent évoluer pour soutenir non seulement le partage, mais aussi la réutilisation effective des données, en tenant compte des objectifs différenciés des producteurs et des ré-utilisateurs.

Usage et réutilisation des données sont au cœur de la « diplomatie des données », qui englobe la production, le partage, l'accès, l'interprétation, la protection, la gouvernance et l'impact des données à toutes les étapes de leur cycle de vie (Boyd *et al.*, 2019). On distingue trois axes : les données dans la diplomatie - outil ou objet dans les processus diplomatiques (exemple statistiques pour la santé, droits humains, etc.), la diplomatie pour les données - négociations et cadres pour permettre l'accès, l'utilisation, la protection et la circulation des données (exemple standards, législation, consentement, gouvernance) ; et les données pour la diplomatie - usage des données pour créer ou renforcer des relations, soutenir des communautés, ou servir d'outil de *soft power* (exemple initiatives humanitaires, *empowerment* communautaire, OCAP¹).

L'utilisation des données peut donc être bénéfique ou néfaste, selon les modalités de gouvernance, de transparence, de consentement et d'équité. La participation citoyenne et la prise en compte des attentes sociales sont essentielles pour garantir une « licence sociale » à l'utilisation des données. La diplomatie des données implique ainsi une pluralité d'acteurs et nécessite des compétences, des normes et des mécanismes de confiance adaptés à la complexité croissante des usages des données.

Dimensions et figures de la médiation

La médiation en SIC « fournit des outils pour décrire avec une certaine précision les processus d'information-communication ; elle permet de requalifier socialement les dynamiques et régimes de la culture ; elle conduit le chercheur à interroger sa propre place dans la circulation sociale des savoirs » (Jeanneret, 2008, p. 37).

Le processus de médiation ne se limite pas à observer et partager des contenus ; il demande de s'intéresser aux formes de circulation, aux dispositifs, aux acteurs et aux contextes impliqués. La médiation permet également de requalifier les dynamiques culturelles pour comprendre de quelle manière les pratiques, les objets et les savoirs changent de statut lorsqu'ils circulent. En ce sens, en tant que requalification sociale, elle met en lumière les régimes de légitimité : ce qui est considéré comme savant, populaire, artistique,

scientifique... Enfin, la médiation est prise comme réflexivité du chercheur quand elle l'oblige à interroger sa propre position dans la circulation des savoirs. Le chercheur n'est pas extérieur : il participe lui-même à la médiation en produisant des discours, des analyses, des publications. Ainsi, la médiation est à la fois cadre d'analyse des processus d'information-communication, outil critique pour comprendre la transformation sociale des savoirs et des cultures et exercice réflexif qui engage le chercheur dans la dynamique qu'il étudie. La médiation a pour objectif d'assurer une articulation entre la dimension individuelle de l'utilisateur dans son rapport à l'information et la dimension sociale de la construction d'un savoir. À l'instar d'Yves Jeanneret (2009), trois dimensions principales caractérisent les processus de médiation : symbolique, langagière et logistique et technique. La dimension symbolique recouvre valeurs, normes et attentes réciproques des acteurs de la communication, leurs interactions et interrelations. Elle fait référence à un système de représentations commun à toute une société et correspond à des formes d'identifications sociales et de sociabilité ; la dimension sociale et langagière permet de cerner les contenus et fonctions des échanges et leur signification dans les processus de traduction, d'intermédiaires entre des mondes à rapprocher. Elle renvoie au processus de réécriture de l'information, créateur d'un nouveau message (Gardiès, 2012). La médiation est ainsi porteuse d'une certaine « matière informationnelle » et augmente les significations à réception ; enfin, la dimension logistique et technique, permet de repérer les objets médiateurs de la communication médiatisée et est donc liée à son organisation matérielle. C'est ce phénomène global que nous mobilisons au regard des médiations sociales qui résultent des compromis et qui peuvent s'élargir dans des médiations des savoirs, par rapport notamment, aux ré-usages des données.

Le compromis selon Paul Ricoeur

Le compromis occupe une place centrale dans la pensée de P. Ricoeur, bien qu'il ne soit jamais défini de manière systématique dans son œuvre (Assayag, 2016). Le compromis n'est pas une faiblesse mais une force fragile, indispensable à la construction d'une vie commune dans une société traversée par le conflit. Il se distingue nettement de la compromission, impliquant trahison ou abdication des principes. Le compromis est au contraire une démarche positive, fondée sur la reconnaissance mutuelle et la recherche du bien commun. P. Ricoeur (1991) conçoit le compromis comme une mise en intersection d'ordres de grandeur différents. Cette notion d'intersection est essentielle : elle désigne le point de rencontre entre valeurs ou intérêts divergents, sans qu'il y ait subordination d'un ordre sur l'autre. L'intersection permet d'éviter une hiérarchie normative et instaure une relation horizontale entre les domaines éthique, politique et économique. Plus l'intersection est large, plus le compromis est facile à construire ; plus elle est réduite, plus il devient difficile.

Dans son dialogue avec les sociologues L. Boltanski et L. Thévenot (1991), P. Ricoeur reprend l'idée que les individus évoluent dans des « ordres de grandeur » multiples (civique, marchand, domestique, industriel, ...), chacun porteur de ses propres règles et valeurs. Les conflits naissent de la coexistence de ces ordres hétérogènes. Le compromis apparaît alors comme une solution toujours précaire et révoquant, pourtant seule capable de viser le bien commun en l'absence d'un principe supérieur partagé par tous. Il peut être externe (entre différents ordres) ou interne (au sein d'un même ordre), et suppose toujours des concessions mutuelles (Assayag-Gillot, 2018).

Enfin, le compromis possède une dimension éthique et politique. Il ne s'agit pas seulement d'un outil pratique de gestion de conflits, mais d'un fondement du rapport à soi et à autrui. Sur le plan politique, il incarne l'esprit démocratique qui ne cherche pas à supprimer les conflits mais à inventer des procédures permettant expression et négociation. Il devient ainsi condition de possibilité du bien commun, dans une société pluraliste où la reconnaissance mutuelle reste à construire. Le compromis est une voie intermédiaire qui refuse fatalisme et violence, tout en acceptant pluralité et fragilité des accords humains. Il est à la fois structure (intersection des ordres), dynamique (processus renégociable) et

éthique (fondement du vivre-ensemble).

LES TRACES DES COMPROMIS

Notre corpus est constitué d'une sélection d'articles issus de la littérature récente et internationale sur les questions d'ouverture et de ré-usage des données. Nous avons procédé par recherche thématique à partir d'équations « données + usage », « données + ré-usage », « données + ouverture » sur le moteur de recherche *Google scholar*² ainsi que sur l'archive ouverte en ligne *Hal*³. Nous avons initialement retenu quarante-huit documents. Il s'agit plus ici d'une revue de littérature narrative que systématique, au sens où l'analyse documentaire « narrative » conventionnelle permet d'avoir une vue générale sur un sujet précis, pour soulever des problèmes habituellement négligés et encourager d'autres recherches (Nambiema *et al.*, 2021). Nous avons complété la collecte grâce à la lecture des bibliographies et une sélection d'ouvrages récents sur la gestion des données.

Partant du postulat que des traces de compromis étaient présentes dans cette littérature, nous avons testé cette méthode sur un premier article (Annexe 1). Nous avons effectué un relevé méthodique des éléments observables, caractéristiques du compromis : ordres de grandeur, dimensions épistémologiques, formes de médiation, dispositifs, usages, ré-exploitation des données. L'analyse a permis de distinguer les textes relevant d'une approche épistémologique sur les données, de textes rendant compte d'expériences autour du partage des données. L'analyse plus fine de ce deuxième ensemble constitué de vingt-cinq documents a mis en évidence un certain nombre de valeurs en conflit, d'intersection entre différents ordres de grandeur.

Nous avons alors procédé à une analyse qualitative (Bardin, 2013) de ce deuxième ensemble en relevant de manière systématique les traces de compromis au sens ricœurzien. La notion de compromis se noue précisément au point d'intersection des valeurs des contractants dans le cadre d'un conflit déclaré entre biens, intérêts et principes.

Lorsqu'un conflit entre les valeurs de transparence (*Open Data*) vs. protection des droits est présent, le compromis se réalise au niveau du principe d'ouverture par défaut, mais avec exceptions prévues (propriété intellectuelle, vie privée, confidentialité). « Les chercheurs ont identifié plusieurs facteurs susceptibles d'influencer les pratiques de partage et de réutilisation des données, notamment les normes culturelles, les considérations juridiques et éthiques... l'impact social (...) l'infrastructure technologique (...) et le financement ou les politiques institutionnelles » (Maurel, 2018, p. 2-3). Les intersections sont multiples entre valeurs scientifiques (collaboration), contraintes institutionnelles (juridique, financement) et valeurs sociales (impact), rendant le compromis nécessaire pour équilibrer ces dimensions.

Lors d'un conflit entre droits des administrations (bases de données) vs. libre réutilisation, le compromis se présente en termes de neutralisation du droit du producteur de bases de données au profit de l'ouverture. « La recherche a été affinée à l'aide des fonctions de recherche avancée d'Elicit [application d'intelligence artificielle assistant les chercheurs dans la recherche et l'analyse de publications scientifiques] (...) incluant les filtres pour les mots-clés, les dates de publication, les types d'études... » (Maurel, 2018, p. 4). L'intersection entre exhaustivité scientifique et contraintes méthodologiques incite au compromis méthodologique pour équilibrer rigueur et faisabilité.

Si un conflit entre droit vs éthique scientifique a lieu, la loi fixe un cadre minimal, mais les bonnes pratiques et l'éthique complètent la régulation, forme du compromis. La mise en œuvre concrète de l'ouverture repose autant sur la régulation juridique que sur des pratiques éthiques partagées. Nous avons reproduit cette méthodologie pour chacun des articles.

SCIENCE OUVERTE, CONFLITS ET COMPROMIS

Ces compromis incarnent les trois formes décrites par P. Ricœur : le point d'équilibre de dernier recours (ex. tension entre éthique et ouverture), le résultat de concessions mutuelles (ex. entre institutions et chercheurs) ou la forme de régulation conjointe (ex. entre archives quantitatives et qualitatives).

Plusieurs points d'intersection et de compromis entre valeurs, intérêts et principes parfois contradictoires sont relevés. Notre analyse montre que la science ouverte progresse par une série de compromis entre valeurs d'ouverture et de protection, entre pérennisation et contraintes financières, entre reconnaissance individuelle et partage collectif, et enfin entre normalisation et diversité.

Catégorie de valeurs en conflit	Compromis identifiés	Auteurs principaux
Confidentialité vs protection	- Documentation éthique et juridique claire dans le DMP - Anonymisation obligatoire (vie privée, données qualitatives) - Consentement éclairé (informer et anonymiser) - Gouvernance de la confidentialité (sécurité technique et malveillance)	Maurel, Jacquemin & Schöpfel, Huyghe, Ahmed, Brenel, Bonneville, Rossi
Reconnaissance vs autonomie des chercheurs	- Liberté académique et droit à l'enquête (régulation proportionnée) - Labellisation nationale et coopération inter-établissements - Proportionnalité entre ouverture et respect des droits	Rossi, Morand, Brenel, Schöpfel, Kosmopoulos
Pérennisation vs infrastructures	- Plans de gestion identifiant responsabilités et archivage - Aide au dépôt (mise en réseau, infrastructures partagées) - Certification des entrepôts (crédibilité et traçabilité)	Jacquemin, Schöpfel, Maurel, Rossi, Curty, Rebouillat, Morand, Bonneville, Late, Ochsner
Normalisation vs diversité disciplinaire	- Approche pragmatique de la normalisation - Cadres génériques mais flexibles - Services disciplinaires vs généralistes - Reconnaissance des spécificités SHS - Interopérabilité adaptable aux pratiques	Dargentas et al., Rebouillat, Brenel, Chartron, Pasquetto et al., Kosmopoulos & Schöpfel, Jiang, Prost, Barré & Billy, Rossi

Figure 1 : synthèse transversale des compromis identifiés

Conflit entre confidentialité et protection

Si l'ouverture des données scientifiques est encouragée, elle doit être conciliée avec protection des données personnelles, confidentialité et respect du consentement des participants. « Les chercheurs eux-mêmes ont tendance à considérer l'éthique comme un “frein juridique” et principal facteur empêchant une libre diffusion de leurs données. Mais la gestion des données de la recherche mobilise aussi d'autres valeurs scientifiques, comme intégrité et qualité des données, transparence du processus, solidarité, réciprocité et engagement d'excellence vis-à-vis de la société, équité et respect des personnes. [...] »

pratique scientifique, législation et politique de la recherche peuvent être en conflit en ce qui concerne la dimension éthique des données » (Jacquemin *et al.*, 2018, p. 3).

Les plans de gestion des données intègrent des sections dédiées à l'éthique, demandant de justifier toute restriction à l'ouverture par des raisons éthiques ou légales.

Le principe d'ouverture par défaut s'applique, sous exceptions et une politique « *as open as possible, as closed as necessary* » illustre ce compromis. Les données sont alors protégées par des droits de tiers et exclues de l'ouverture par défaut, avec obligation d'anonymisation stricte avant diffusion de données personnelles. En ce sens, « le cadre d'éthique de la recherche (...) indique que les considérations éthiques s'appliquent à l'archivage, au partage, à la liaison et à la réutilisation des données » (Bishop et Kuula-Luumi, 2017, p. 2). On parle alors de gouvernance de la confidentialité (sécurité technique et malveillance) (Jacquemin et Schöpfel, 2018 ; Rossi, 2023, p. 2).

Des zones d'insécurité juridique pour la recherche en SHS sont ainsi mises en lumière, en particulier autour des données issues des réseaux socio-numériques, requérant des pistes pour concilier science ouverte et protection des droits fondamentaux (Rossi, 2023, p. 4). À l'instar d'autres auteurs, Bernard Jacquemin et Joachim Schöpfel (2018) mettent l'accent sur l'intégration de l'éthique dans les plans de gestion des données. Cette dimension éthique doit être explicite, fondée sur la protection des personnes par le biais du consentement éclairé et la nécessité d'anonymiser les données sensibles. Concernant les valeurs de confidentialité, protection et éthique, nous avons donc identifié des tensions récurrentes entre ouverture des données vs protection des personnes, vie privée, droits d'auteur, RGPD. Le compromis prend la forme d'une anonymisation systématique (Huyghe, 2018 ; Dargentas⁴ *et al.*, 2007, p. 156-173 ; Rossi, 2023 ; Jacquemin et Schöpfel, 2018), d'un consentement éclairé comme garantie éthique (Curty, 2016, p. 97 ; Bonneville, 2021). L'enjeu est alors de concilier science ouverte et respect des droits fondamentaux par des dispositifs techniques, juridiques et éthiques.

Conflit entre pérennisation et infrastructures

Du point de vue de la pérennisation et des infrastructures, les tensions récurrentes se focalisent sur la durabilité des données comme sur les projets numériques souffrant de financements limités. Il est alors question de compromis en termes de certification et normalisation des entrepôts (Rebouillat 2020, Prost et Schöpfel, 2019, Jiang *et al.*, 2025), d'archivage institutionnel ou national (Morand, 2025, p. 5-47 ; Nury, 2023, avec Yareta, FORSbase/SwissUbase), de standards ouverts et interopérabilité (FAIR, TEI, EpiDoc) ou encore de mutualisation des services. Dans leurs travaux, Bishop et Kuula-Luumi (2017) montrent notamment l'évolution de l'acceptation de la réutilisation des données qualitatives, avec infrastructures et politiques de citation. Ces pratiques concrètes illustrent, dans un laboratoire pluridisciplinaire, les tensions entre données sensibles et partage, et la faible culture de réutilisation (Bonneville, 2021). Pour lutter contre ces tensions, nous assistons à une structuration nationale des plateformes Data, entre diversité locale et harmonisation via PROGEDO et le RnMSH⁵ (Morand, 2025), ou à l'essor de projets en humanités numériques (Grammateus⁶), mettant en avant les enjeux de pérennisation et de citabilité des projets numériques (Nury, 2023). L'objectif est bien de garantir la pérennité par des infrastructures certifiées, une documentation exhaustive et une mutualisation opérationnelle.

Conflit entre reconnaissance et autonomie des chercheurs

Du point de vue de la reconnaissance et de la valorisation des chercheurs, les tensions récurrentes se nouent essentiellement sur la faible reconnaissance de la réutilisation des données, en opposition avec la valorisation de la collecte originale. Le compromis se réalise grâce à un encouragement à citer les données via le DOI (Bishop et Kuula-Luumi, 2017) des indicateurs bibliométriques ouverts mesures d'impact (Bordignon *et al.*, 2021, p. 4-6), une culture du partage interne puis élargie (Huyghe, 2018) ainsi qu'un discours favorable malgré

des pratiques réticentes (Rebouillat, 2021, p. 35-47 ; Herbert *et al.* 2023, p.4-35). Le fait est que cette reconnaissance reste insuffisante, bien que des outils émergent pour valoriser la réutilisation.

Conflit entre normalisation et diversité

Enfin, bien que la normalisation soit indispensable pour assurer l'interopérabilité des données, elle doit rester souple et respectueuse des diversités disciplinaires (Dargentas *et al.*, 2007 ; Rebouillat, 2019,2020, 2021 ; Brenel, 2022 ; Pasquetto *et al.*, 2017 ; Kosmopoulos et Schöpfel, 2025).

Les tensions récurrentes sont identifiées avec les standards internationaux (FAIR, certification) par rapport à des spécificités locales et disciplinaires. Le compromis passe par une approche pragmatique (Kosmopoulos et Schöpfel, 2025), des cadres génériques mais flexibles (Jiang *et al.*, 2025), des services disciplinaires ou généralistes (Rebouillat 2019 ; Prost et Schöpfel, 2019, p. 71-98) et surtout une reconnaissance des spécificités SHS (Herbert *et al.*, 2023 ; Rossi, 2023). Ainsi, « les normes FAIR ont été adoptées par un ensemble de parties prenantes pour permettre une science ouverte, et elles intègrent toutes les parties de l'objet de recherche, du code aux données, en passant par les outils d'interprétation » (Pasquetto *et al.*, 2017, p. 3-4). Ces extraits illustrent la manière dont la collaboration et la régulation conjointe (au sens de Ricoeur) nécessitent flexibilité et concessions mutuelles sur les formats, standards et documentation, pour permettre interopérabilité et confiance. L'enjeu transversal est que cette normalisation est nécessaire pour l'interopérabilité, mais adaptable aux pratiques disciplinaires.

Nous avons ainsi identifié des compromis éthiques (entre exigences éthiques et nécessité de diffusion, protection des participants et ouverture des données), des compromis méthodologiques (entre anciens et nouveaux projets, quantitatif et qualitatif), des compromis institutionnels (entre financeurs, éditeurs, gouvernements et chercheurs), des compromis économiques (ambition de partage contre ressources limitées), des compromis académiques (intersection des finalités - recherche, enseignement, apprentissage), ou encore des compromis pratiques (abondance des collections contre usage réel sélectif).

Ces compromis dessinent un équilibre fragile mais dynamique, où la légitimité des politiques dépend de leur capacité à intégrer les spécificités des disciplines et les pratiques réelles des chercheurs. Les dimensions de la médiation présentées par Jeanneret (2008) et Gardiès (2012) permettent en ce sens de comprendre les enjeux structurels des tensions, entre valeurs, pratiques et dispositifs. Les figures de la médiation éclairent les formes d'expression de ces tensions : réflexivité du chercheur, médiatisation sociale, structuration technique. Les compromis identifiés dans le corpus sont autant d'actes de médiation : ils traduisent, réécrivent, organisent et négocient les savoirs dans des contextes contraints (Annexe 2).

La médiation est un processus d'interprétation, de transformation et de circulation des savoirs dans des contextes sociaux, culturels et techniques. Elle permet de requalifier les dynamiques culturelles, d'interroger les dispositifs de communication, et de réfléchir à la place du chercheur dans la production des savoirs. L'exemple de B. Jacquemin et J. Schöpfel (2018) montre que les auteurs mobilisent la figure sociale (représentation et médiatisation de l'expérience scientifique) en intégrant l'éthique dans les DMP. Leur approche technique (certification, sécurisation) relève de la dimension logistique et de la figure procédurale. Leur insistance sur le consentement et la protection des personnes illustre la dimension symbolique. Un autre exemple, celui de Coralie Morand (2025) fait apparaître que la mise en réseau des PUD via PROGEDO est une médiation technique (dimension logistique), tandis que l'accompagnement FAIR et la labellisation nationale sont des formes procédurales traduisant les normes en dispositifs. En outre, la diversité locale des plateformes montre une tension entre symbolique disciplinaire et normalisation technique. Autrement dit, la médiation, dans ses dimensions et figures, offre une grille d'analyse puissante pour comprendre les compromis dans la science ouverte. Elle offre la possibilité de relier les valeurs en tension aux dispositifs concrets, d'identifier les acteurs et leurs rôles dans la

circulation des savoirs et enfin de mettre en lumière les formes de légitimation, de traduction et de négociation.

CONCLUSION

Notre analyse met en lumière des tensions fondamentales : les données sont à la fois factuelles et produites, individuelles et collectives, scientifiques et sociales. Elles sont produites, interprétées et contextualisées. Leur valeur est liée à leur usage, réutilisation et circulation. La science ouverte progresse par série de compromis entre ouverture et protection, normalisation et diversité, pérennisation et contraintes financières, reconnaissance individuelle et partage collectif.

Dans ce contexte, la notion de médiation apparaît comme un outil conceptuel majeur. Ne se limitant pas à la transmission d'information, elle vise à la transformer en connaissance, en articulant dimension individuelle de l'utilisateur et dimension sociale de la construction du savoir. Les travaux en Sciences de l'information et de la communication (Jeanneret, 2009 ; Gardiès, 2012) permettent de distinguer plusieurs dimensions (symbolique, langagière, technique) et figures (réflexive, sociale, procédurale) de la médiation, éclairant les processus de circulation et d'appropriation des savoirs. Cette réflexion rejoint la pensée de P. Ricœur sur le compromis, conçu comme intersection des intérêts. Le compromis, distinct de la compromission, constitue un fondement éthique du rapport à soi et à autrui, et un projet politique démocratique. Dans le champ scientifique, il se traduit par des ajustements et des bricolages nécessaires à la structuration et au partage des données, révélant les tensions inhérentes aux pratiques collectives. Cette première exploration des liens entre médiation et compromis nous encourage à poursuivre notre réflexion constituant un axe d'étude majeur de travaux à venir. Il sera notamment pertinent de reprendre la modélisation de nos résultats (diagramme de Kiviat) pour munir les négociateurs d'un meilleur outil d'aide à la décision apte à faire émerger du compromis, notamment en faisant apparaître les différents registres (compromis éthiques, méthodologiques, institutionnels, économiques, académiques et pratiques). La médiation est ici mobilisée dans l'analyse des résultats, dans ses dimensions et figures, essentiellement pour comprendre les compromis, il serait alors intéressant de la mobiliser dans l'analyse de la réutilisation des données. Ainsi, l'étude des données de recherche en SHS invite à interroger simultanément leur construction sociale, les médiations qui en permettent la circulation, et les compromis qui en rendent possible le partage. Elle met en lumière la nécessité de penser les données non seulement comme objets techniques, mais comme réalités sociales et épistémologiques au cœur de la production du savoir.

NOTES

¹ Principes de gouvernance des données des Premières Nations : *Ownership, Control, Access, Possession* (Propriété, Contrôle, Accès, Possession). <https://rimuhc.ca/fr/-/webinaire-les-principes-ocap-des-premieres-nations>

² <https://scholar.google.com>

³ <https://hal.science>

⁴ Première expérience française de capitalisation qualitative (EDF R&D), avec charte juridique et anonymisation.

⁵ PROduction et GEstion des Données : infrastructure nationale française dédiée aux données quantitatives en SHS, chargée de leur production, gestion et diffusion.

RnMSH : Réseau national des Maisons des Sciences de l'Homme.

⁶ Programme en humanités numériques porté par l'Université de Genève.

RÉFÉRENCES BIBLIOGRAPHIQUES

Assayag, Laure (2016), “Penser la confiance avec Paul Ricœur », *Etudes Ricœuriennes / Ricœur Studies*, vol. 7, n° 2, p. 164 186, [en ligne], consulté le 28 mars 2026, <https://ricœur.pitt.edu/ojs/ricœur/article/view/318>

Assayag-Gillot, Laure (2018), « Le compromis selon Paul Ricœur », *Négociations*, n° 1, p. 103 120.

Bardin, Laurence (2013), *L'analyse de contenu*, Paris : PUF.

Bates, Marcia J. (2005), “Information and knowledge: an evolutionary framework for information science”, *Information Research*, vol. 10, n°4, [en ligne], consulté le 28 avril 2026, <https://informationr.net/ir/10-4/paper239.html>

Bishop, Libby ; Kuula-Luumi, Arja (2017), “Revisiting Qualitative Data Reuse: A Decade On”, *SAGE Open*, vol. 7, n° 1, [en ligne], consulté le 28 avril 2026, <https://doi.org/10.1177/2158244016685136>

Boltanski, Luc ; Thévenot, Laurent (2022), « Les compromis pour le bien commun » (p. 337-356), in *De la justification : les économies de la grandeur*, Paris : Gallimard.

Bonneville, A. ; Tucci, I. ; Vion, A.y ; Giglio, L. (2021). « Données de la recherche : pratiques et besoins dans un laboratoire pluridisciplinaire SHS » [rapport de recherche]. Laboratoire d'économie et sociologie du travail (LEST). <https://hal.science/hal-03265603/>

Borgman, Christine L. (2015), *Big data, little data, no data: Scholarship in the Networked World*, MIT Press.

Borgman, Christine L.; Brand, Amy (2024), “The future of data in research publishing: From nice to have to need to have?”, *Harvard Data Science Review*, [en ligne], consulté le 28 avril 2026, <https://doi.org/10.1162/99608f92.b73aae77>

Borgman, Christine L. ; Groth, Paul (2025), “From data creator to data reuser: Distance matters”, *Harvard Data Science Review*, [en ligne], consulté le 28 avril 2026, <https://hdr.mitpress.mit.edu/pub/2mvqwgmf/release/2>

Boukacem-Zeghmouri, Chérifa ; Rakotoary, Sarah ; Bador, Pascal (2021), “La prédation dans le champ de la publication scientifique : un objet de recherche révélateur des mutations de la communication scientifique ouverte », *Natures Sciences Sociétés*, vol. 29, n° 4, p. 382 395.

Bordignon, Frédérique ; Boukacem-Zeghmouri, Chérifa (2021), « Quelle place les rapports AÉRES et HCÉRES font-ils à la science ouverte ? Réponses par une analyse textométrique (2009 2021) » (p.159-172), in Mahé, Annaïg ; Mayeur, Ingrid ; Poupardin, Elsa ; Prime-Claverie, Camille (dir.), *Communication scientifique et science ouverte : opportunités, tensions et paradoxes - Actes du colloque « Document numérique et société »*, Liège, 2022, Bruxelles : De Boeck Supérieur.

Boyd, Andy ; Gatewood, Jane ; Thorson, Stuart J. ; Dye, Timothy (2019), « Data diplomacy”, *Science & Diplomacy*, [en ligne], consulté le 28 avril 2026, <https://www.sciencediplomacy.org/article/2019/data-diplomacy>

Brenel, Mireille ; Mercier, Cédric ; Suhan, Stela ; Kassas, Adib ; Ménard, Claire ; et al. (2022), *Rapport d'analyse - Enquête : Les données de la recherche à l'université Paris-Saclay, panorama et perspectives*, Université Paris-Saclay, [en ligne], consulté le 28 avril 2026, <https://universite-paris-saclay.hal.science/hal-03857804/document>

Curty, Renata Gonçalves (2016), “Factors influencing research data reuse in the social sciences: An exploratory study”, *International Journal of Digital Curation*, vol. 11, n° 1, p. 96-11.

Dargentas, Magdalini ; Le-Roux, Dominique ; Salomon, Annie-Claude ; Brugidou, Mathieu (2007), « Sur les perspectives de la recherche qualitative en France : capitalisation et ré-utilisation d'entretiens de recherche », *Recherches Qualitatives*, n° 3, p. 156-173.

Gardiès, Cécile (2012), *Dispositifs info-communicationnels de médiation des savoirs : cadre d'analyse pour l'information-documentation*, Habilitation à diriger les recherches en Sciences de l'information et de la communication - Université Toulouse 2.

Goody, Jack (1979), *La raison graphique : la domestication de la pensée sauvage*, Paris : Editions de Minuit.

Herbet, Marie-Emilia ; Montébran, Aurore ; Dugué, Alexandra ; Billy, Raphaëlle ; Maire, Mariana ; et al. (2023), *Gestion et ouverture des données de recherche en SHS : une enquête auprès des communautés de Lyon 2, Lyon 3 et la MSH Lyon St-Étienne*, [en ligne], consulté le 28 avril 2026, <https://univ-lyon3.hal.science/hal-04087568/>

Huyghe, Marie ; Cailly, Laurent ; Oppenchaim, Nicolas (2018), *Partage des données d'enquêtes en recherche qualitative : approche juridique, éthique, déontologique et épistémologique*, Rapport de recherche CITERES, [en ligne], consulté le 28 avril 2026, <https://shs.hal.science/CITERES-COST/halshs-01790983v1>

Jacquemin, Bernard ; Schöpfel, Joachim ; Chaudiron, Stéphane ; Kergosien, Eric (2018), « L'éthique des données de la recherche en sciences humaines et sociales » (p. 71-86), in Balicco, Laurence ; Broudoux, Evelyne ; Chartron, Ghislaine ; Clavier, Viviane ; Pailliar, Isabelle (dir.), *L'éthique en contexte info-communicationnel numérique : déontologie, régulation, algorithme, espace public*, Bruxelles : De Boeck Supérieur.

Jeanneret, Yves (2008), *Penser la trivialité. Vol. 1 : La vie triviale des êtres culturels*. Paris: Hermès-Lavoisier.

Jeanneret, Yves (2009), « La relation entre médiation et usage dans les recherches en information-communication en France », Reciiis, [en ligne], consulté le 28 avril 2026, <https://doi.org/10.3395/reciis.v3i3.276fr>

Jiang, Jie ; Maurici-Pollock, Danielle ; Tang, Rong (2025), « Research Data Lifecycle (RDLC): An investigation into the disciplinary focus, use cases, creator backgrounds, stages and shapes of RDLC models », *International Journal of Digital Curation*, vol. 19, n° 1.

Kosmopoulos, Christine ; Schöpfel, Joachim (2025), (dir.), *Publier, partager, réutiliser les données de la recherche : les data papers et leurs enjeux*, Villeneuve d'Ascq : Presses universitaires du Septentrion, [en ligne], consulté le 28 avril 2026, https://www.septentrion.com/html/WYSIWYGfiles/files/HNSO2_Publier-Partager-R%C3%A9utiliser.pdf

Leleu-Merviel, Sylvie (2017), *La traque informationnelle*, Londres : ISTE Editions (collection « Sciences, société et nouvelles technologies »).

Marciset-Sognos, Sylvie ; Canizares, Aurélie ; Gardiès, Cécile (2024), « Processus de médiation numérique des savoirs et dispositifs info-communicationnels : vers des figures sociales, procédurales et réflexives de la médiation », *Spirale-Revue de recherches en éducation*, vol. 73, n° 1, p. 55-68.

Maurel, Lionel (2018), « La réutilisation des données de la recherche après la loi pour une République numérique », in Ginouvès, Véronique ; Gras, Isabelle (dir.), *La diffusion numérique des données en SHS : guide de bonnes pratiques éthiques et juridiques*, Aix en Provence : Presses Universitaires de Provence.

Morand, Coralie (2025), *Etat des lieux du réseau de plateformes technologiques data*, Paris: Réseau national des Maisons des Sciences de l'Homme, [en ligne], consulté le 28 avril 2026, https://www.msh-reseau.fr/media/pages/plateformes/b82404ddd-1739888219/rapport-plateformes_data_vf.pdf

Nambiema, Aboubakari ; Fouquet, Jeanne ; Guilloteau, Juliette ; Descatha, Alexis (2021), « La revue systématique et autres types de revue de la littérature : qu'est-ce que c'est, quand, comment, pourquoi ? », *Archives des Maladies Professionnelles et de l'Environnement*, vol. 82, n° 5, p. 539-552.

Nury, Elisa (2023), « Réutiliser et pérenniser les données de la recherche : l'exemple de *grammateus* », communication présentée au *Colloque Humanistica*, juin 2023, Genève, Suisse, [en ligne], consulté le 28 avril 2026, <https://hal.science/hal-04102863v1>

Pasquetto, Irene V. ; Randles, Bernadette M. ; Borgman, Christine L. (2017), "On the reuse of scientific data", *Data Science Journal*, vol. 16, [en ligne], consulté le 28 avril 2026, <https://doi.org/10.5334/dsj-2017-008>

Pontille, David (2017), « Contributions profanes et attribution scientifique » (p. 137-152), in Leclerc, Olivier (dir.), *Savants, artistes, citoyens : tous créateurs ?*, Québec : Editions science et bien commun, [en ligne], consulté le 28 avril 2026, <https://scienceetbiencommun.pressbooks.pub/touscreateurs/>

Prost, Hélène ; Schöpfel, Joachim (2019), « Les entrepôts de données en sciences de l'information et de la communication (SIC). Une étude empirique », *Etudes de communication*, n° 52, p. 71-98.

Rebouillat, Violaine (2019), *Ouverture des données de la recherche : de la vision politique aux pratiques des chercheurs*, thèse de doctorat en Sciences de l'information et de la communication, Conservatoire national des arts et métiers - Cnam, Paris.

Rebouillat, Violaine (2020, 28 septembre), « Entrepôts de données de recherche : mesurer l'impact de l'Open Science à l'aune de la consultation des jeux de données déposés », communication présentée à la *7ème conférence Document numérique & Société - Humains et données : création, médiation, décision, narration*, 28-29 septembre 2020, Nancy, France, [en ligne], consulté le 28 avril 2026, <https://hal.science/hal-02928817v1>

Rebouillat, Violaine (2021), « Le partage des données vu par les chercheurs : une approche par la valeur », *Les Enjeux de l'Information et de la Communication*, n° 22/1, p. 35-53, [en ligne], consulté le 29 avril 2026, <https://lesenjeux.univ-grenoble-alpes.fr/2021/varia/03-le-partage-des-donnees-vu-par-les-chercheurs-une-approche-par-la-valeur/>

Ricœur, Paul (1991), « Pour une éthique du compromis. Propos recueillis par Jean-Marie Muller et François Vaillant », Fonds Ricœur [en ligne], 6p. http://www.fondsriceur.fr/uploads/medias/articles_pr/pour-une-ethique-du-compromis.pdf

Rossi, Julien (2023), *La recherche en SHS au croisement de multiples contraintes : rapport sur le cadre juridique applicable à l'usage de données issues de réseaux socio-numériques à des fins de recherche scientifique*, Centre d'expérimentation en méthodes numériques pour les recherches en Sciences Humaines et Sociales - CERES, Université Paris 8, [en ligne], consulté le 28 avril 2026, https://ceres.sorbonne-universite.fr/7ee3fa47ad31c7aa169f903a5de279ac/rapport_V6.pdf

Star, Susan Leigh (1999), "The Ethnography of Infrastructure", *American Behavioral Scientist*, vol. 43, n° 3, p. 377-391.

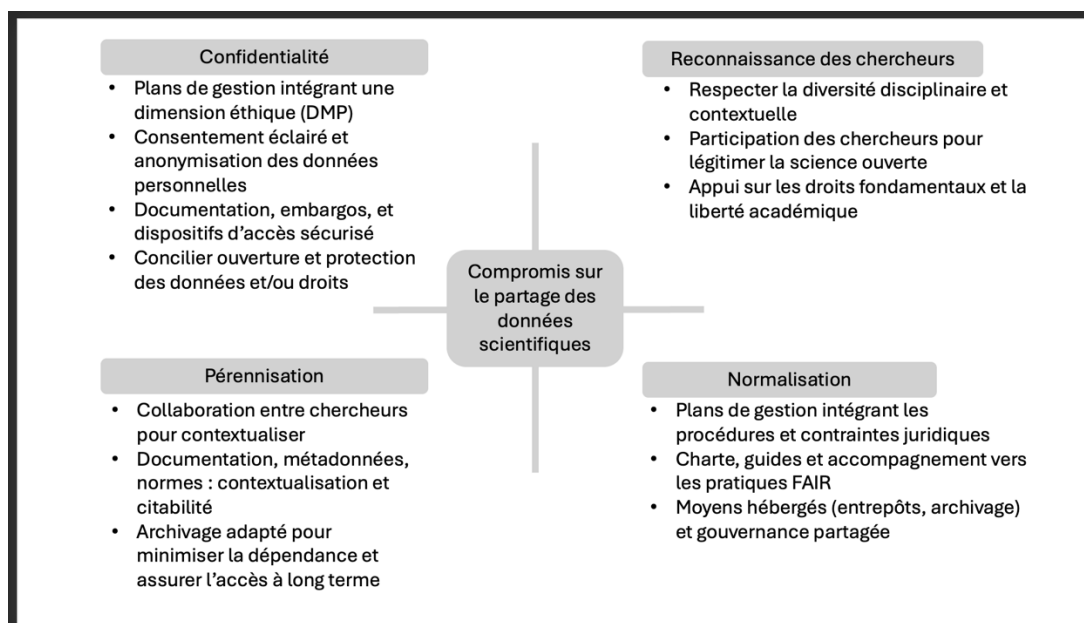
Star, Susan Leigh ; Strauss, Anselm (1999), « Layers of Silence, Arenas of Voice : The Ecology of Visible and Invisible Work », *Computer Supported Cooperative Work*, n° 8, p. 9-30.

ANNEXES

Annexe 1 : TABLEAU COMPARATIF ENTRE INTERSECTIONS DE VALEURS ET COMPROMIS ISSUS DE L'ANALYSE DU TEXTE DE L. MAUREL (2018).

Conflit / Valeurs en tension	Compromis ou intersection
Transparence (Open Data) vs protection des droits	Principe d'ouverture par défaut, mais exceptions prévues (propriété intellectuelle, vie privée, confidentialité)
Droits des administrations (bases de données) vs libre réutilisation	Neutralisation du droit du producteur de bases de données au profit de l'ouverture
Pouvoir des éditeurs privés vs liberté des chercheurs	Clauses contractuelles exclusives neutralisées, réutilisation des données garantie
Obligation légale vs liberté des projets de recherche	Licences facultatives : réutilisation possible même sans licence, mais bonnes pratiques encouragées
Ouverture des données vs propriété intellectuelle des tiers	Données protégées par des droits de tiers exclues de l'ouverture par défaut
Transparence vs protection de la vie privée	Obligation d'anonymisation stricte avant diffusion des données personnelles
Intérêt public (Open Data) vs intérêts privés (consortiums, secrets industriels)	Confidentialité contractuelle reconnue, mais tension avec ouverture par défaut
Droit vs éthique scientifique	La loi fixe un cadre minimal, mais les bonnes pratiques et l'éthique complètent la régulation

Annexe 2 : CARTE CONCEPTUELLE DES PRINCIPAUX COMPROMIS RELEVÉS DANS LE CORPUS



Annexe 3 : MEDIATION ET COMPROMIS : UNE GRILLE D'ANALYSE CROISEE

Dimensions/figures de la médiation	Nature	Liens avec les compromis identifiés
Symbolique + Figure sociale	Valeurs, normes, représentations sociales	Tensions entre ouverture/transparence et protection des données (Maurel, Silberman, Jacquemin et Schöpfel, Rossi) Représentations culturelles diverses selon les disciplines (Kosmopoulos & Schöpfel, Jiang)
Langagière et sociale + Figure réflexive	Traduction, réécriture, circulation des savoirs	Réécriture des données pour les rendre partageables (anonymisation, documentation) (Rossi, Huyghe, Dargentas) Réflexion sur les pratiques réelles versus discours institutionnel (Rebouillat, Barré & Billy)
Logistique et technique + Figure procédurale	Dispositifs, supports, infrastructures	Certification et archivage sécurisés (CoreTrust-Seal, Yareta, FSD) (Late & Ochsner, Nury, Morand) Plans de gestion des données, communautés de médiation technique et éthique (Jacquemin & Schöpfel, Rossi, Morand)

Annexe 4 : TABLEAU GLOBAL DES COMPROMIS AUTOUR DES DONNEES ET DE LA SCIENCE OUVERTE

Auteurs	Valeurs en conflit	Compromis trouvés / Intersections	Pages
Ahmed (2025)	Avantages du partage vs réticences individuelles	Identification de facteurs culturels, juridiques, sociaux, technologiques	1-2
Barré, Billy et al. (2023)	- Injonctions nationales (plans de gestion, FAIR, science ouverte) vs pratiques locales en SHS (stockage personnel, faible diffusion) - Données sensibles et personnelles vs exigences de transparence et de répliquabilité - Héritage normatif des sciences exactes vs spécificités des SHS - Besoin d'accompagnement vs lourdeur/inadéquation des obligations	- État des lieux des pratiques et besoins pour adapter les politiques aux réalités SHS - Coopération inter-établissements pour mutualiser compétences et infrastructures - Recommandations pour sensibilisation, formation et accompagnement - Reconnaissance de la diversité disciplinaire et des limites de l'ouverture intégrale	4-35
Bishop & Kuula-Luumi (2017)	- Réutilisation des données qualitatives vs limites méthodologiques - Politiques de partage vs pratiques réelles - Mesure de l'impact vs difficultés de citation	- Développement d'infrastructures (UK Data Service, FSD) - Acceptation progressive de l'analyse secondaire - Encouragement à citer les données via DOI	1-15
Bordignon (2021)	- Besoin d'indicateurs bibliométriques vs risques de réductionnisme - Bases commerciales biaisées vs exigence d'inclusion	- Initiatives ouvertes (I4OC, I4OA, OpenAlex) - Usage des archives ouvertes et métadonnées FAIR - Baromètres de science ouverte	4-6

Auteurs	Valeurs en conflit	Compromis trouvés / Intersections	Pages
	interdisciplinaire - Protection des données de citation par les éditeurs vs ouverture des métadonnées	contextualisés	
Bonneville (2021)	- Données sensibles vs pratiques de partage - Documentation scientifique vs faible formalisation technique - Réutilisation rare vs besoin de mutualisation	- Production de documentation scientifique majoritaire - Anonymisation et consentement oral privilégiés - Réflexion interne pour politique de la donnée et mutualisation	5-36
Brenel (2022)	Obligation institutionnelle vs autonomie des chercheurs	Publication FAIR imposée mais accompagnée par coopération inter-établissements	4
Curty (2016)	Commodité vs rigueur	Réutilisation = gain de temps mais effort méthodologique supplémentaire	97
Dargentas et al. (2007)	- Capitalisation des entretiens vs confidentialité - Diversité des pratiques vs homogénéisation - Réutilisation secondaire vs limites méthodologiques	- Base Verbatim (EDF R&D) avec charte juridique - Collaboration entre chercheurs pour contextualiser - Documentation exhaustive des corpus	156-173
Nury (2023)	- Pérennisation des projets numériques vs absence de financement - Données externes réutilisées vs production originale - Dynamisme des données vs stabilité	- Archivage sur Yareta avec TEI/EpiDoc - Encodage XML pour interopérabilité - Citabilité et accessibilité assurées	1
Huyghe (2018)	Intérêt scientifique vs confidentialité des enquêtes	Anonymisation obligatoire avant diffusion	3-4
Jacquemin & Schöpfel (2018)	- Ouverture/transparence vs protection des données personnelles - Respect des personnes vs réutilisation des données - Crédibilité scientifique vs risques de données erronées - Sécurité technique vs vulnérabilité humaine - Propriété intellectuelle vs accès libre	- Plans de gestion intégrant une dimension éthique (DMP) - Consentement éclairé et anonymisation - Certification et audit des entrepôts - Sécurisation technique et organisationnelle - Reconnaissance des droits tout en favorisant l'ouverture raisonnée	71-86
Jiang et al. (2025)	Spécificité disciplinaire vs généralisation	Modèles génériques et multidisciplinaires	4, 5-6
Kosmopoulos & Schöpfel (2025)	- Science ouverte vs diversité disciplinaire - Normalisation internationale vs autonomie des chercheurs	- Approche pragmatique (subsidiarité, complémentarité) - Développement d'infrastructures partagées et certification	

Auteurs	Valeurs en conflit	Compromis trouvés / Intersections	Pages
	- Pérennisation des infrastructures vs financement limité	- Nécessité d'adhésion des chercheurs	
Late & Kekäläinen (2020)	- Données quantitatives vs qualitatives - Ouverture vs protection - Téléchargements fréquents vs citations rares	- Anonymisation et conditions d'accès différenciées - Documentation et métadonnées - Réutilisation encouragée par dépôts nationaux (FSD)	1-17
Late & Ochsner (2024)	- Ouverture des archives numériques vs faible réutilisation - Étudiants vs chercheurs confirmés - Téléchargements massifs vs faible citation	- Certification CoreTrustSeal - Données surtout utilisées pour enseignement - Développement de métriques d'usage via logs	1-21
Maurel (2018)	Transparence (open data) vs protection (PI, vie privée)	Ouverture par défaut mais exceptions légales (anonymisation, droits tiers)	2-3, 7-9
Morand (2025)	- Diversité locale des plateformes vs harmonisation nationale - Besoins des chercheurs vs contraintes institutionnelles - Pérennisation financière vs dépendance aux projets	- Mise en réseau des PUD via PROGEDO - Labellisation nationale par le RnMSH - Développement de services mutualisés et accompagnement FAIR	5-47
Pasquetto et al. (2017)	Production de nouvelles données vs réutilisation	Arbitrage méthodologique entre collecte et réutilisation	5, 6
Prost & Schöpfel (2019)	- Politiques nationales de science ouverte vs pratiques spécifiques des SIC - Définition conceptuelle incertaine des données vs approche opérationnelle centrée sur infrastructures - Normalisation technique vs reconnaissance des dimensions communautaires	- Cartographie des entrepôts disciplinaires, génériques et institutionnels pour les SIC - Recommandations pour articuler certification/normalisation avec pratiques réelles - Positionnement des SIC comme acteurs de la science ouverte	71-98
Rebouillat (2019)	Services IST vs pratiques réelles des chercheurs	Usage limité mais reconnaissance de l'utilité potentielle	6-7
Rebouillat (2020)	- Politiques d'ouverture vs pratiques limitées - Diversité des infrastructures vs besoin de normalisation	- Mesure de l'impact par consultations/téléchargements - Certification (CoreTrustSeal) comme compromis - Reconnaissance des biais méthodologiques	1-20 env.
Rebouillat (2021)	- Idéal de science ouverte vs faible reconnaissance académique des données	- Trois régimes de valeur (pour soi, usage, échange) - Partage conditionné par	35-47 env.

Auteurs	Valeurs en conflit	Compromis trouvés / Intersections	Pages
	- Valeur de l'article vs valeur d'usage/échange des données - Injonctions institutionnelles vs attachement personnel	- reconnaissance symbolique - Nécessité de réformer l'évaluation de la recherche	
Rossi (2023)	- Liberté académique vs contraintes juridiques - Science ouverte vs protection des données - Obligation de plans de gestion vs incertitudes pratiques	- Droit à l'enquête scientifique (Charte UE) - Anonymisation et consentement - Interprétation proportionnée des normes - Plans de gestion équilibrés	2-53
Schöpfel (2018)	Politiques nationales vs autonomie des chercheurs	Nécessité d'adhésion des chercheurs pour légitimer science ouverte	9
Silberman (2024)	Ouverture/transparence vs confidentialité/protection	Documentation, anonymisation, embargos, dispositifs d'accès sécurisé	4, 8, 10-12

Gouvernance et circulation des données dans les archives ouvertes en Afrique : le cas de DICAMES

Article inédit, mis en ligne le 23 septembre 2026

Alphée Yannick Hilaire NGALLI NGOMO

Enseignant-chercheur à l'Université de Douala, au sein du Département de Communication. Spécialiste des Sciences de l'Information et de la Communication (SIC), des humanités numériques et du numérique éducatif. Ses travaux portent principalement sur la gouvernance de l'information, la souveraineté numérique et l'endogénéisation des technologies éducatives en contexte africain.
yannickngalli@gmail.com

PLAN DE L'ARTICLE

Résumé et mots clés en français
Title, abstract and keywords in english
Titulo, resumen y palabras clave en español
Introduction
Cadre théorique et revue de littérature
Méthodologie
Les défis de la gouvernance des données
Conclusion

RÉSUMÉ

Cette étude analyse la collection numérique DICAMES pour comprendre sa gouvernance et la circulation des données. Mobilisant une méthodologie bibliométrique, elle révèle une répartition géographique inégale et des biais structurels dans les métadonnées. L'article préconise une normalisation des données et une double validation pour corriger ces lacunes. Au-delà des aspects techniques, il mobilise les notions de souveraineté informationnelle et de diplomatie scientifique, concluant sur la nécessité de faire de ces archives des leviers stratégiques pour renforcer la position de l'Afrique dans le paysage scientifique mondial.

Mots clés

Collection numérique, archives ouvertes, gouvernance des données, circulation des données, DICAMES, diplomatie des données

TITLE

GOVERNANCE AND CIRCULATION OF DATA IN OPEN ARCHIVES IN AFRICA: THE CASE OF DICAMES

Abstract

This study analyzes the DICAMES digital collection to understand its governance and data

circulation. Utilizing a bibliometric methodology, it reveals an uneven geographical distribution and structural biases within the metadata. The paper advocates for data standardization and double validation to address these shortcomings. Moving beyond technical aspects, it engages with the concepts of informational sovereignty and science diplomacy, concluding that these archives must be leveraged as strategic tools to strengthen Africa's position in the global scientific landscape.

Keywords

Digital collection, open archives, data governance, data circulation, DICAMES, data diplomacy

TITULO

GOBERNANZA Y CIRCULACIÓN DE DATOS EN LOS ARCHIVOS ABIERTOS EN ÁFRICA: EL CASO DE DICAMES

Resumen

Este estudio analiza la colección digital DICAMES para comprender su gobernanza y la circulación de datos. A través de una metodología bibliométrica, revela una distribución geográfica desigual y sesgos estructurales en los metadatos. El artículo aboga por una normalización de los datos y una doble validación para corregir estas deficiencias. Más allá de los aspectos técnicos, moviliza las nociones de soberanía informacional y diplomacia científica, concluyendo en la necesidad de convertir estos archivos en herramientas estratégicas para fortalecer la posición de África en el panorama científico mundial.

Palabras clave

Colección digital, archivos abiertos, gobernanza de datos, circulación de datos, DICAMES, diplomacia de datos

INTRODUCTION

Lancée en 2017 par le CAMES (Conseil Africain et Malgache pour l'Enseignement Supérieur), la plateforme DICAMES est une archive ouverte visant à collecter, préserver et diffuser la production scientifique des 19 pays membres. Elle constitue un outil essentiel pour la visibilité et la circulation des savoirs en Afrique francophone, ouvrant l'accès à des travaux qui n'auraient peut-être pas trouvé d'autre canal de diffusion. Cependant, ces avantages ne se réalisent que si la gouvernance est effective.

Pour comprendre les caractéristiques de cet outil, il est nécessaire de le situer dans son contexte historique et institutionnel. Le CAMES, en tant qu'institution, a traversé plusieurs phases de développement, et DICAMES s'inscrit dans une stratégie de large modernisation et d'adaptation aux enjeux du numérique. Son lancement en 2017 répond à un double objectif : pallier l'absence de visibilité des recherches issues des établissements membres et garantir leur conservation à long terme. Au-delà de son architecture technique, la construction d'une archive ouverte multinationale comme DICAMES soulève des questions fondamentales de gouvernance des données scientifiques. Sa structure, ses règles de contribution et ses biais intrinsèques ne sont pas neutres ; ils reflètent et influencent la circulation des connaissances entre les pays membres et avec le reste du monde. Cet article

propose donc d'analyser DICAMES non seulement comme une collection, mais aussi comme un outil de diplomatie scientifique, en interrogeant ses mécanismes de gouvernance et son potentiel à réduire les déséquilibres dans la production et la diffusion des savoirs en Afrique francophone.

Trois notions structurent notre analyse : la gouvernance des données (processus et règles définissant l'accès et l'usage des données), la souveraineté informationnelle (capacité d'un État à contrôler ses infrastructures informationnelles) et la diplomatie des données (usage stratégique des données pour établir des alliances et influencer les normes internationales, s'appuyant sur le soft power (Nye, 2004).

DICAMES vise à préserver et à promouvoir l'ensemble de la production scientifique des pays africains qui ont le français en partage. Plus précisément, cette collection est principalement conçue pour combattre l'absence de visibilité et le faible accès numérique à la recherche émanant des établissements d'enseignement supérieur membres de la région CAMES, en mettant en place une approche pour garantir sa conservation à long terme dans des conditions idéales. DICAMES rassemble une large typologie de documents : thèses, mémoires, articles (avec ou sans comité scientifique), chapitres de livre, livres, rapports, communications, et autres documents jugés utiles par les gestionnaires.

La création de cette archive s'inscrit dans un mouvement global de promotion qui encourage la diffusion sans restriction des résultats de la recherche (Piron *et al.*, 2017). Cette collection compte aujourd'hui environ 8 000 documents repartis dans le tableau ci-dessous. La prééminence du Cameroun (plus de 6 000 documents) mérite d'être nuancée : elle témoigne autant de stratégies différenciées de valorisation du capital scientifique et de capacités techniques que d'une productivité inégale. DICAMES apparaît ainsi comme un instrument de cartographie des réseaux d'influence académique.

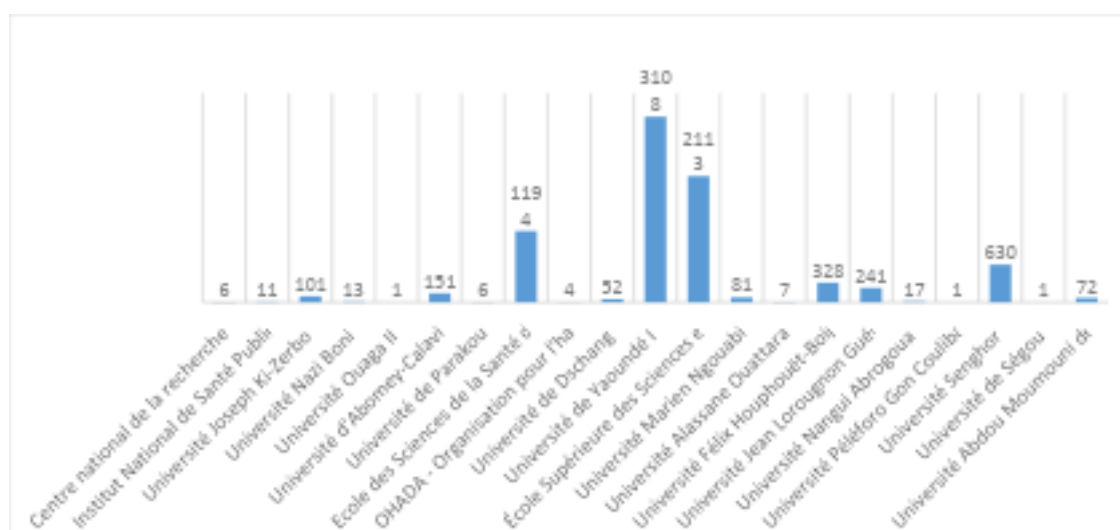


Figure 1. Contribution des établissements d'enseignement supérieur à la collection DICAMES. Données : DICAMES. Graphique : Yannick Ngalli, 2024

Du point de vue de la gouvernance de l'information, l'analyse de la collection soulève de nombreuses interrogations sur les causes d'une telle disparité. Nous pensons que la méthode choisie pour la construction de ces archives ouvertes ne facilite pas sa mise en œuvre au regard de son caractère multinational et multi-établissements. La coordination et le management d'une telle initiative constitue un enjeu majeur de gouvernance et de circulation des données à l'échelle régionale.

En plus des documents, la base dispose aussi de métadonnées associées à ces documents et c'est cet ensemble qui constitue la collection numérique. En effet, chacun des documents

présents dans la base est rattaché à un certain nombre de métadonnées descriptives. Leur rôle est de faciliter l'accès à la copie numérique dudit document. Les métadonnées choisies par DICAMES sont : titre, auteur(s), mots-clés, date de publication, éditeur résumé, pagination / nombre de pages, URI/URL, *digital object identifier* (DOI), collection(s). Ces métadonnées ne sont pas standardisées puisqu'elles ne correspondent pas aux formats normalisés comme l'ISBD (*International Standard Bibliographic Description*), ni à un encodage standard, tel que MARC (*MAchine-Readable Cataloging*). Ce manque de standardisation peut être un frein à l'interopérabilité des données de DICAMES avec d'autres systèmes et entraver la circulation internationale des données qu'elle contient.

La question de recherche est : comment l'examen des données bibliométriques éclaire-t-il le processus de constitution de DICAMES et ses répercussions sur la gouvernance des données pour l'Afrique francophone ? L'hypothèse est que l'analyse des biais dans les métadonnées révèle les mécanismes de gouvernance et identifie des leviers pour renforcer la souveraineté informationnelle.

CADRE THÉORIQUE ET REVUE DE LITTÉRATURE

Cette recherche s'inscrit dans une approche interdisciplinaire au carrefour des Sciences de l'information et de la Communication (SIC), de la sociologie des sciences et des études postcoloniales sur la circulation des savoirs. Pour analyser DICAMES comme objet de gouvernance, nous mobilisons un cadre théorique articulé autour de deux axes : la sociologie des infrastructures informationnelles et les théories des relations internationales appliquées aux sciences.

Deboin (2015) montre que les techniques et réseaux informatiques constituent le substrat permettant aux chercheurs de s'organiser en communautés, favorisant l'émergence des collections numériques.

Berthaud et al. (2021) identifient trois facteurs dans l'émergence des collections numériques (actions individuelles, mobilisation communautaire, stratégies institutionnelles). Schöpfel et al. (2022) et Schöpfel (2020) montrent, à partir du cas HAL, que les pratiques de dépôt varient selon les disciplines et les environnements institutionnels.

Trzmielewski et Gnoli (2019), quant à eux, ont choisi de réfléchir sur une évaluation de l'*Integrative Levels Classification* (ILC) en vue d'organiser des données ouvertes de la recherche en sciences humaines et sociales. Leur travail s'appuie sur une analyse des propriétés de l'ILC par rapport aux pratiques, usages et contextes de partage et de médiation des données. Les résultats qu'ils ont obtenus présentent l'ILC comme prometteuse pour la construction d'une science ouverte, dynamique et novatrice. Cependant, les travaux centrés sur les usages et les politiques nationales doivent être complétés par une perspective critique sur les enjeux de pouvoir dans la circulation des données.

S'agissant de la collection numérique que nous avons choisie pour notre étude de cas, Piron *et al* (2017) la présentent comme une plateforme ouverte qui vise la conservation et la diffusion en libre accès de toute la production scientifique des pays d'Afrique francophone subsaharienne. Leur travail décrit les principales étapes du processus de création du DICAMES en montrant les difficultés et les défis, mais aussi les solutions trouvées. Au-delà des aspects techniques et organisationnels, la littérature s'intéresse de plus en plus aux questions de gouvernance et de pouvoir dans la circulation des données de la recherche. Borgman (2015) souligne que l'ouverture des données est autant politique qu'infrastructurelle, tandis que Fabre et Gardiès (2008) montrent que l'accès à l'information scientifique est structuré par des rapports de pouvoir reproduisant les hiérarchies mondiales. Dans le contexte africain, ces enjeux sont exacerbés par la fracture numérique, mais les archives ouvertes comme DICAMES peuvent être des outils de souveraineté scientifique, à condition d'une gouvernance robuste.

Nous mobilisons le concept de réseau sociotechnique (Callon, 1991) pour appréhender DICAMES comme un assemblage d'acteurs humains, d'artefacts techniques, de règles et d'enjeux politiques. Ce cadre permet d'analyser la construction des asymétries dans la production et la circulation des savoirs. Par ailleurs, la diplomatie scientifique (Ruffini, 2017) et la diplomatie des données (Morin et al., 2021) fournissent des outils pour comprendre comment DICAMES peut contrer l'invisibilité de la recherche africaine dans les bases internationales. Notre revue de littérature met en évidence que si les aspects techniques des archives ouvertes sont bien documentés, leur dimension géopolitique dans le contexte du Sud global reste peu explorée. C'est en réponse à cette lacune que nos travaux ont été réalisés.

MÉTHODOLOGIE : UNE ANALYSE BIBLIOMÉTRIQUE COMME SONDE DE GOUVERNANCE

Notre approche méthodologique mixte est centrée sur l'analyse bibliométrique des métadonnées de DICAMES. Cette méthode, définie par Archambault et Vignola-Gagné (2004), permet l'évaluation quantitative des sciences et l'identification des biais et disparités révélateurs des mécanismes de gouvernance à l'œuvre.

Objectifs et questions de recherche

En nous référant au cadre méthodologique présenté par Robert K. Yin (2003), cette recherche se penche sur les caractéristiques essentielles des travaux effectués dans les pays membres du CAMES à travers les archives ouvertes DICAMES. Cette analyse poursuit quatre objectifs : cartographier les thématiques privilégiées, identifier les acteurs majeurs, observer l'évolution historique des tendances, et analyser la structure de gouvernance implicite à travers les disparités de contribution.

L'hypothèse principale de cette recherche est que l'examen des données bibliométriques présentes dans la collection numérique DICAMES approfondit la compréhension du processus de gouvernance et de circulation des données de cette collection numérique.

Récupération et préparation des données

La récupération et la préparation des données bibliographiques de la collection numérique DICAMES est une phase cruciale pour réaliser nos objectifs de recherche. Nous avons recours à la technique de *scrapping* en utilisant le langage de programmation Python et la bibliothèque spécialisée connue sous le nom de *BeautifulSoup*. Cette méthode nous a permis d'extraire l'intégralité des métadonnées disponibles sur le site DICAMES. Le script développé est disponible en ligne (https://github.com/YANNHIL2022/travaux_cshn) pour assurer la transparence et la reproductibilité de la recherche. L'extraction des métadonnées a été réalisée par *scrapping* avec Python et BeautifulSoup, en suivant une procédure itérative par communauté, sous-communauté et collection. Les données ont ensuite été nettoyées avec OpenRefine, enrichies via Wikidata pour les coordonnées géographiques, et visualisées avec Khartis, Tableau et Excel. Le script est disponible en ligne pour assurer la reproductibilité (https://github.com/YANNHIL2022/travaux_cshn). Les données collectées ont ensuite été nettoyées et enrichies avec *OpenRefine*, puis réconciliées avec les données de Wikidata pour ajouter des coordonnées géographiques. Les visualisations ont été produites avec Karthis, Tableau et Excel.

Analyse des données

L'analyse des données a consisté en l'examen systématique des métadonnées selon quatre axes : la répartition géographique des contributions, l'évolution temporelle des dépôts, la typologie des documents et l'analyse des auteurs. Pour chaque axe, nous avons cherché à identifier des biais ou des disparités qui pourraient indiquer des mécanismes de gouvernance spécifiques. L'analyse ne s'est pas limitée à une description des graphiques, mais a cherché

à interpréter ces résultats à la lumière du cadre théorique présenté précédemment.

Une fracture géographique dans la contribution

Sur les 19 pays membres, seuls 8 contribuent activement, avec une fragmentation interne. Cette cartographie révèle moins une simple répartition géographique qu'une analyse géopolitique des rapports de force dans la production et le partage des savoirs, distinguant pays actifs (exportateurs de capital scientifique) et pays absents (consommateurs passifs).

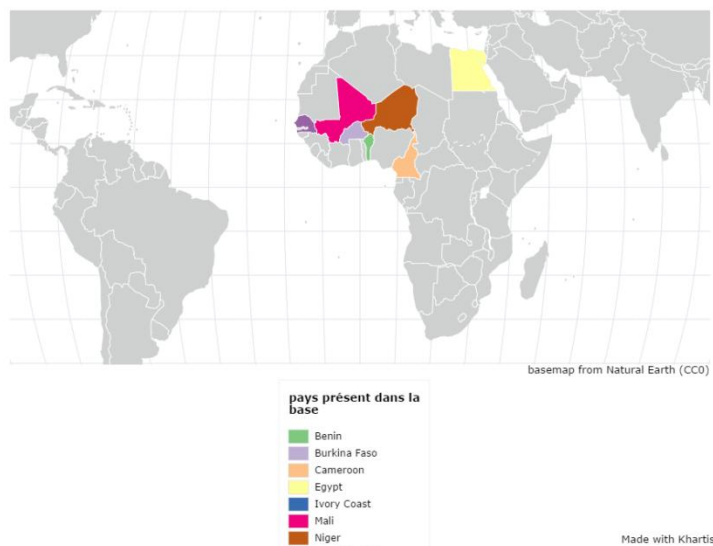


Figure 3. États contributeurs dans la collection DICAMES. Données : DICAMES. Graphique : Yannick Ngalli, 2024

La surreprésentation du Cameroun ne reflète pas nécessairement une domination académique objective, mais peut aussi résulter d'efforts de numérisation ou de politiques incitatives, tandis que les contributions modestes du Mali et du Niger peuvent s'expliquer par l'instabilité géopolitique. La concentration des contributions en Afrique Centrale et de l'Ouest, avec la présence de l'Égypte via l'Université Senghor, dessine un espace d'influence francophone structuré par la langue et les références culturelles. Cette cartographie, moins objective que politique, met en évidence une fracture dans l'accès aux données.

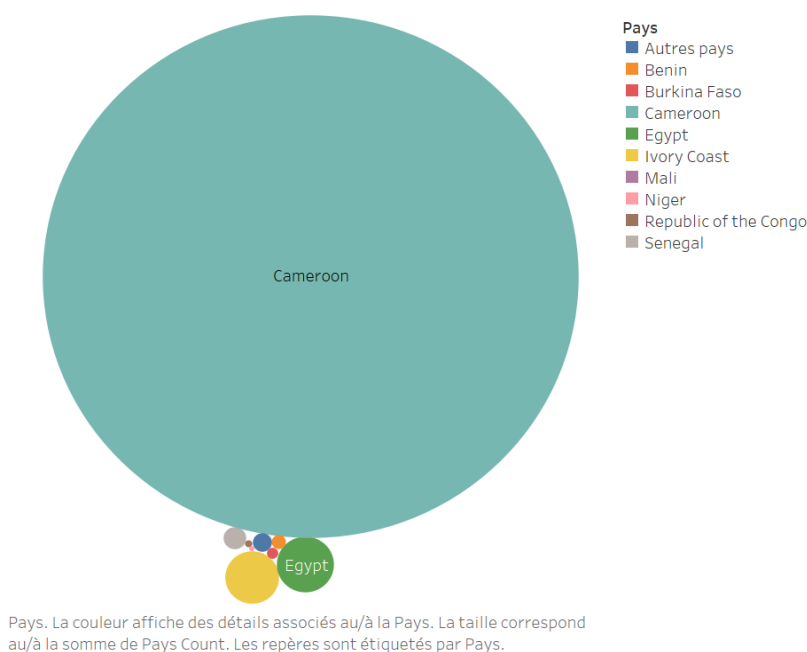


Figure 4. Proportion de la contribution des États dans la collection DICAMES. Données : DICAMES. Graphique : Yannick Ngalli, 2024

Ce déséquilibre expose les pays faiblement représentés à un risque de dépendance, réduisant leur souveraineté académique. L'inclusion des données est donc une condition essentielle pour une diplomatie des données authentiquement multilatérale.

Une temporalité de la collection marquée par la date de création de la plateforme

L'ambition de DICAMES d'intégrer rétroactivement la production antérieure se heurte à l'absence de systèmes préexistants de gestion et de numérisation, imposant une reprise manuelle coûteuse. La figure 5 montre une sur-représentation de la décennie 2010-2020, indiquant une faible intégration des travaux antérieurs.

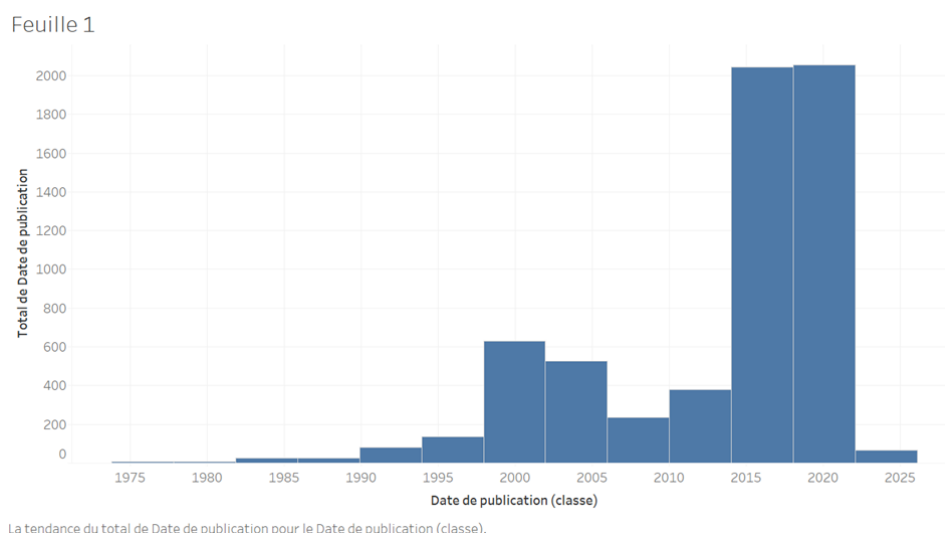


Figure 5. Evolution chronologique des données dans la collection DICAMES. Données : DICAMES. Graphique : Yannick Ngalli, 2024

Quatre stratégies de gouvernance sont envisageables : globale (numérisation exhaustive mais coûteuse), ciblée (date limite avec extension progressive), prospective (production post-adhésion, au détriment de l'héritage), ou décentralisée (auto-archivage, plus rapide mais nécessitant un contrôle qualité). Le défi commun reste de renforcer l'autonomie informationnelle et la représentativité des connaissances.

Typologie des documents conservés dans DICAMES

La base de données comprend des thèses, des mémoires, des articles (avec ou sans comité de lecture), des chapitres, des livres, des rapports, des communications, et d'autres documents académiques (ex. : projets de fin d'études professionnelles).

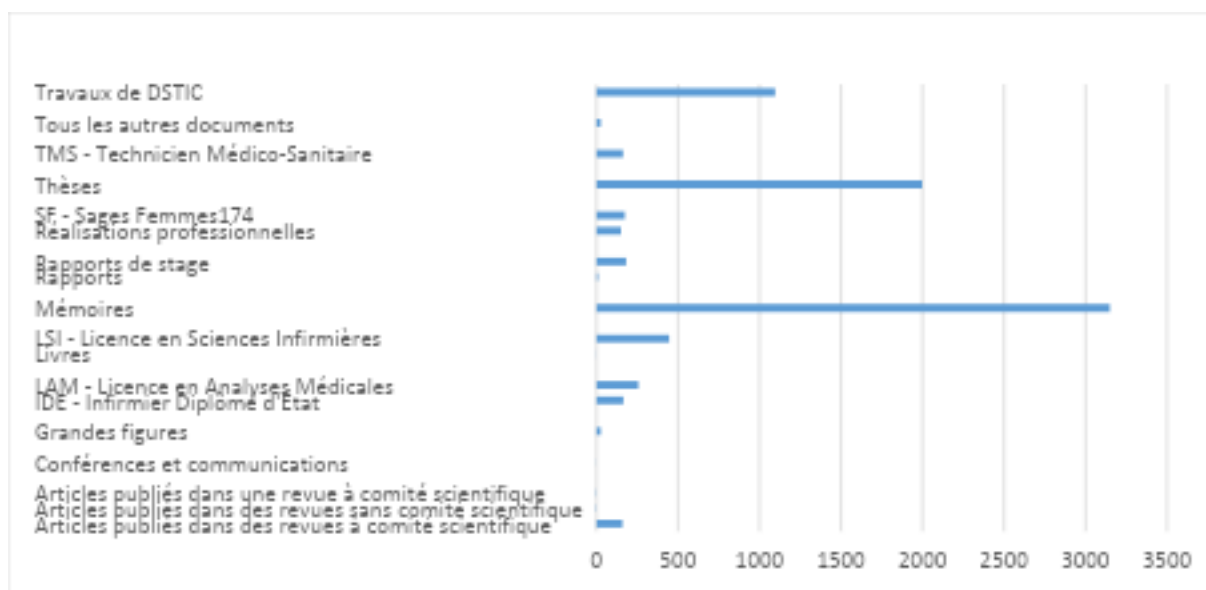


Figure 6. Nombre de documents par type de document dans la collection DICAMES. Données : DICAMES.
Graphique : Yannick Ngalli, 2024

Le graphique ci-dessus estime le volume de chaque sous collection dans la base de données. L'intention qui transparait dans le choix des documents est de regrouper tous les travaux issus de la recherche réalisée dans les établissements cibles. La base fait donc cohabiter les travaux scientifiques comme les thèses, les mémoires, les articles, les livres avec des rapports de fin d'étude de licence qui ne sont parfois que de simples comptes rendus des travaux effectués en stage. Pour bien cerner la constitution de la base en termes de typologie des documents, nous avons choisi d'analyser la catégorie la plus représentée à savoir les mémoires.

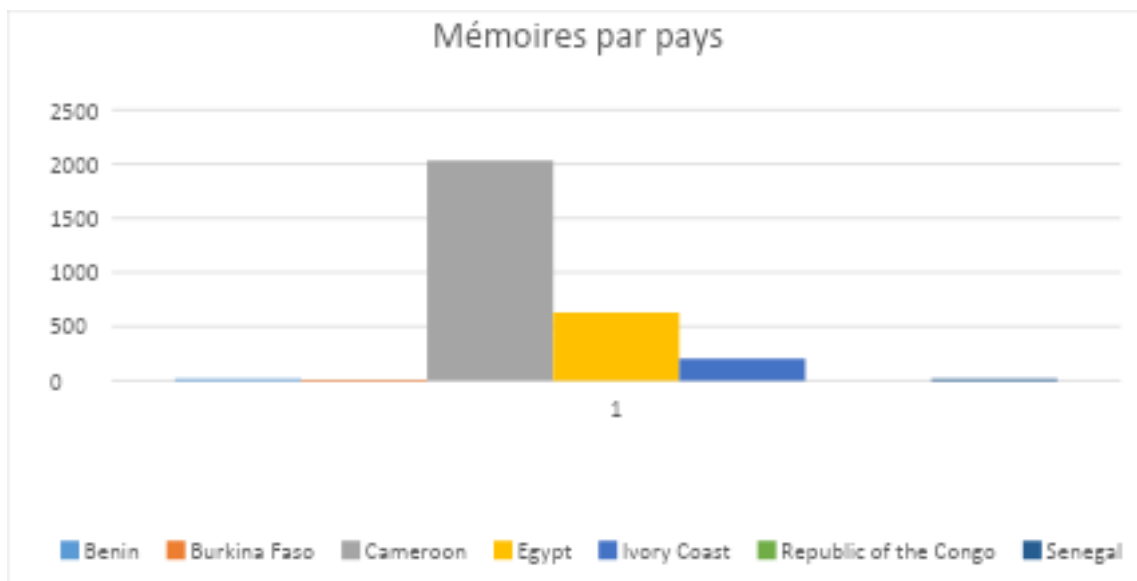


Figure 7. Nombre de mémoire recensés dans la collection DICAMES. Données : DICAMES. Graphique : Yannick Ngalli, 2024

La répartition des mémoires suit la tendance générale (hégémonie camerounaise), mais la position de l'Égypte (2e rang) révèle des contributions ciblées, reflétant les priorités et spécialisations des établissements, et les stratégies différenciées de valorisation du capital scientifique.

On observe ainsi une absence de linéarité entre le poids dataire global d'un pays et son influence dans des sous-ensembles spécialisés. Ce phénomène illustre les logiques de fragmentation et de spécialisation dans la diplomatie des données : certains acteurs peuvent compenser une faible masse critique globale par une contribution ciblée dans des niches à haute valeur académique ou symbolique. Cette situation appelle une gouvernance dataire plus fine au sein de DICAMES, capable de prendre en compte ces disparités thématiques et institutionnelles pour éviter que certains acteurs ne voient leur influence sous-évaluée en raison de biais de catégorisation ou de collecte.

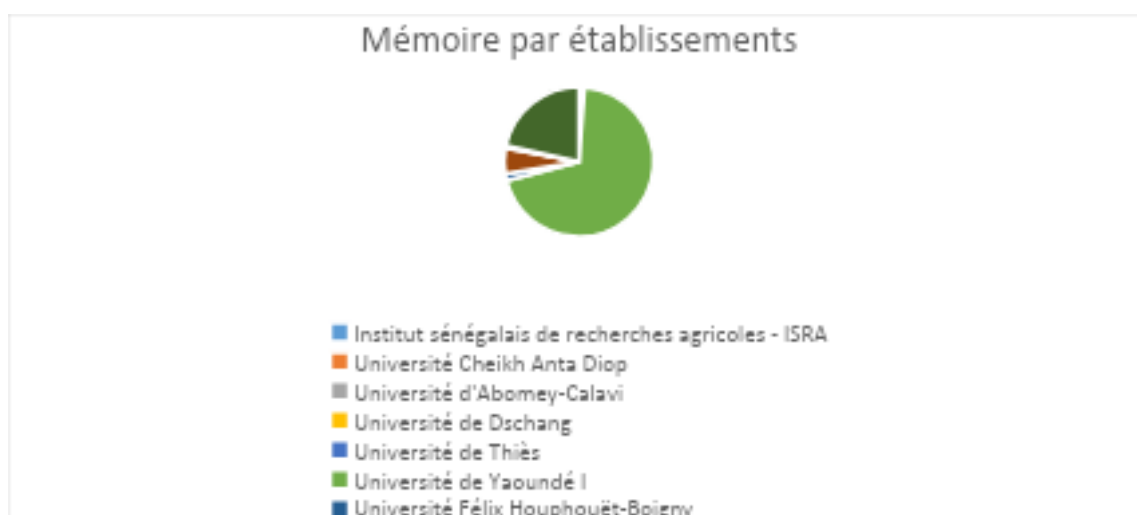


Figure 8. Nombre de mémoires par établissement dans la collection DICAMES. Données : DICAMES. Graphique : Yannick Ngalli, 2024

DICAMES : une diversité encyclopédique

La diversité encyclopédique de DICAMES résulte d'une politique délibérée de non-ségrégation disciplinaire, fondée sur le seul critère d'affiliation à un établissement membre. Si cet atout garantit une représentation équitable des savoirs, il constitue un défi en matière de découvrabilité sans mécanismes de navigation robustes.

Cette approche inclusive traduit une vision stratégique : constituer un bien commun informationnel qui reflète la diversité et la richesse des écosystèmes de connaissance nationaux, sans hiérarchie imposée, affirmant ainsi une autonomie dans la définition de ce qui constitue une production scientifique légitime.

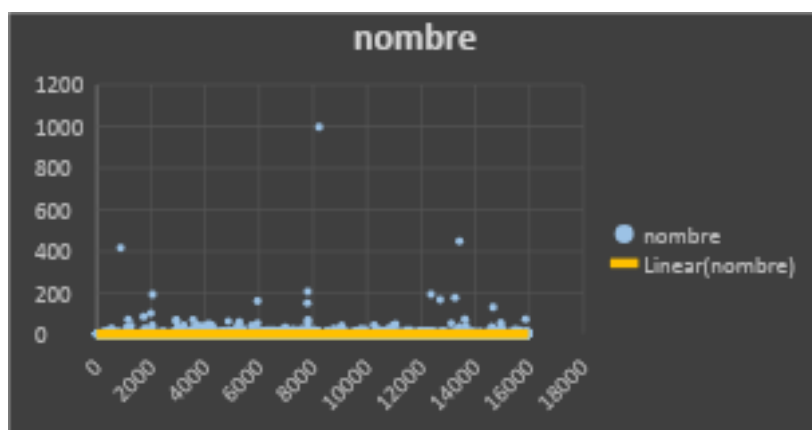


Figure 9. Thématiques dans DICAMES. Données : DICAMES. Graphique : Yannick Ngalli, 2024

Le graphique ci-dessus est une illustration du caractère encyclopédique de la collection DICAMES. En effet, sur les 16 000 thématiques qui sont dans la base, une seule revient plus de 500 fois, une dizaine reviennent plus de 200 et le reste est inférieur à 200. Il faut néanmoins noter que 80% des thématiques reviennent moins de 10 fois, ce qui démontre le caractère encyclopédique de la base qui couvre un très grand nombre de domaines de connaissance.

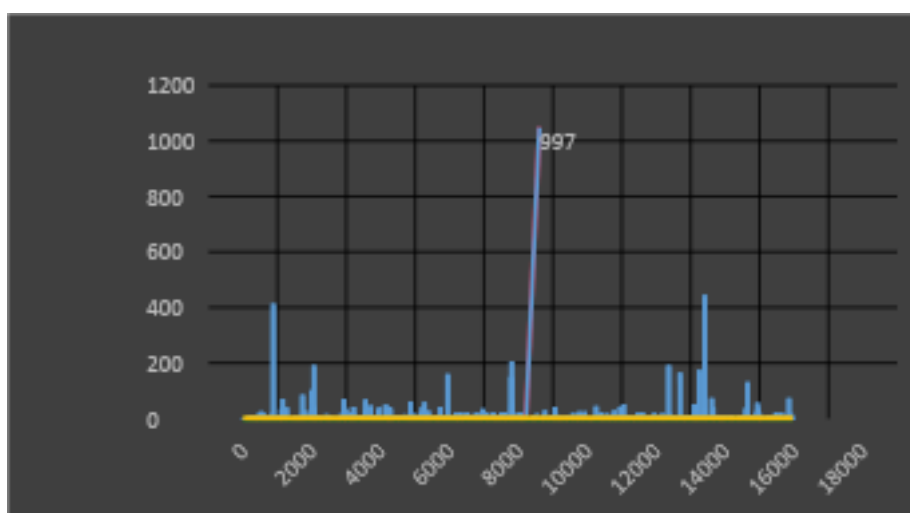


Figure 10. Thématiques ayant été étudiées dans la collection DICAMES. Données : DICAMES. Graphique : Yannick Ngalli, 2024

Sur le plan de la gouvernance des données, l'extraordinaire diversité thématique de la collection DICAMES constitue à la fois sa richesse fondamentale et un défi stratégique majeur en matière de découvrabilité et d'accès à l'information.

Enjeu de gouvernance : la politique d'inclusion exhaustive, reflet d'un engagement diplomatique en faveur de la représentation de tous les savoirs, génère un écosystème de data complexe. Sans mécanisme de navigation robuste, ce pluralisme pourrait paradoxalement nuire à l'utilité et à la valorisation du patrimoine scientifique collecté.

Solution de gouvernance : pour transformer ce défi en opportunité, la gouvernance de DICAMES a institué un cadre d'accès multidimensionnel et intuitif. Cette infrastructure technique incarne le principe selon lequel la souveraineté des données implique la capacité à les explorer efficacement.

- Recherche par métadonnées structurées (date, auteur, titre, sujet) : cette approche garantit l'intégrité et la traçabilité de l'information, favorisant une auditabilité des contenus et une recherche précise, essentielle pour la recherche académique.

- Recherche par arborescence hiérarchique (« communauté et collection ») : cette méthode offre une cartographie navigable du savoir, allant du général au particulier. Elle reflète une vision diplomatique qui organise la connaissance sans la fragmenter, respectant ainsi l'architecture naturelle des disciplines et des institutions.

- Moteur de recherche global : cet outil assure l'exhaustivité de l'exploration basée sur une interrogation transverse de l'ensemble de la collection. Il symbolise l'unité de l'espace CAMES en offrant la possibilité de découvrir des connexions inattendues entre les disciplines.

Cette pluralité de modes d'accès n'est pas une simple fonctionnalité technique ; elle est l'expression opérationnelle d'une gouvernance des données axée sur l'utilisateur. Elle assure que la richesse encyclopédique de DICAMES reste pleinement exploitable, servant ainsi l'objectif diplomatique de diffusion et de valorisation équitable de la production scientifique de l'espace CAMES.

Le problème des auteurs : une granularité sémantique insuffisante

La collection reconnaît les auteurs comme les créateurs primaires du contenu intellectuel, un principe essentiel pour assurer la traçabilité et l'attribution des savoirs. Cependant, la gouvernance actuelle des métadonnées applique une politique d'inclusion extensive en regroupant sous l'entité unique "auteur" des rôles intellectuels distincts : créateurs principaux, directeurs de recherche et membres de jury.

L'analyse des auteurs, (figures 11 à 14) révèle un biais critique dans la gouvernance des métadonnées. La collection regroupe sous l'entité unique « auteur » des rôles intellectuels distincts : créateurs principaux, directeurs de recherche et membres du jury. Cette absence de granularité sémantique a des conséquences importantes. Elle fausse les indicateurs bibliométriques (un directeur de thèse peut apparaître aussi productif que le doctorant) et nuit à la capacité de DICAMES à servir d'outil de cartographie des expertises.

Enjeu de gouvernance des métadonnées, ce choix de modélisation, bien qu'inclusif, introduit un déficit crucial de granularité sémantique. Il engendre une ambiguïté fondamentale sur la nature de la contribution de chaque acteur, portant atteinte à l'intégrité analytique de la base de données.

Conséquences sur la diplomatie des données :

- Biais d'analyse : cette absence de distinction fausse les indicateurs bibliométriques. Un professeur, contributeur indirect en tant que membre de jury, peut apparaître comme plus "productif" qu'un étudiant auteur d'une thèse, biaisant ainsi toute analyse quantitative de la production scientifique réelle.

- Obstacle à la valorisation : cette situation nuit à la capacité de DICAMES à servir d'outil de

diplomatie scientifique. Pour négocier des partenariats ou cartographier les expertises à l'échelle du CAMES, des données fiables et différenciées sur la contribution exacte de chaque chercheur sont indispensables.

- Défi de souveraineté informationnelle : une représentation aussi bruitée des compétences affaiblit la revendication de souveraineté sur la cartographie des savoirs, qui est pourtant un objectif central de ce projet.

Une refonte de la gouvernance des métadonnées visant à normaliser et à différencier les rôles (auteur, directeur, examinateur) n'est pas seulement une amélioration technique. C'est une condition sine qua non pour garantir la qualité, la crédibilité et la valeur diplomatique des données, transformant DICAMES en un instrument fiable pour l'analyse et la promotion de la recherche africaine.

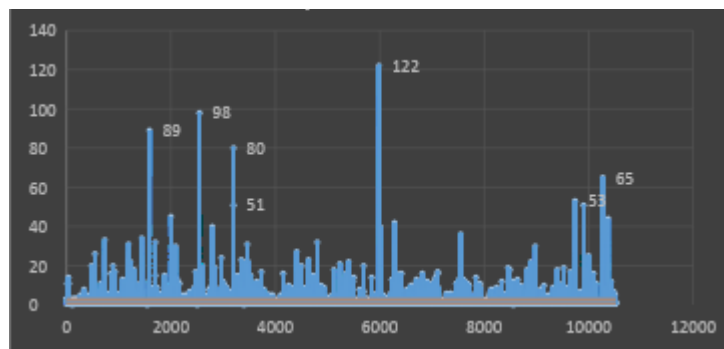


Figure 11. Publication par auteur dans la collection DICAMES. Données : DICAMES. Graphique : Yannick Ngalli, 2024

Le faible nombre de documents par auteur s'explique par le fait que c'est une collection numérique des travaux produits à l'université. À l'exception des enseignants chercheurs qui ont l'obligation d'avoir une production scientifique permanente, les autres producteurs ont la seule contrainte de produire leurs travaux de fin d'étude à savoir leurs mémoires et thèses. La plupart des auteurs étudiants qui constituent la majorité des contributeurs auront un, deux ou trois documents dans la base en fonction de leur niveau d'étude alors que les enseignants ont la possibilité d'avoir une présence égale à leur production intellectuelle en plus des travaux qu'ils dirigent et les jurys auxquels ils participent. S'agissant des auteurs par type de document, il est important de noter que les auteurs suivent la tendance générale de la base de données en termes de proportions. C'est ce que traduit le graphique ci-dessous qui présente les auteurs des mémoires par pays.

Du point de vue de la gouvernance et de la diplomatie des données, la distribution des documents par auteur dans DICAMES n'est pas une anomalie statistique, mais le reflet direct de la structure et des finalités de l'écosystème académique qu'elle documente.

Cette caractéristique est gouvernée par deux réalités distinctes :

- La nature institutionnelle des contributeurs : la majorité des auteurs sont des étudiants, pour qui la production est ponctuelle et liée à l'obtention d'un diplôme (mémoire de master, thèse de doctorat). Leur contribution se limite naturellement à un faible nombre de documents.
- Le cadre normatif de la production scientifique : seuls les enseignants-chercheurs évoluent dans un cadre incitatif exigeant une production scientifique permanente et continue. Leur présence dans la base est donc potentiellement plus régulière et diversifiée, incluant non seulement leurs propres recherches mais aussi leur rôle d'encadrement et d'évaluation.

Du point de vue de la gouvernance des données, cette disparité crée une asymétrie fondamentale dans les métadonnées. Elle soulève un défi de granularité et de qualification

des contributions (comme évoqué précédemment avec la confusion des rôles auteur/directeur/jury). Une analyse quantitative brute, sans pondération par le statut et le rôle, générerait des indicateurs biaisés et non représentatifs de la dynamique réelle de production intellectuelle.

Du point de vue de la diplomatie des données, pour servir d'outil de diplomatie scientifique, DICAMES doit distinguer et valoriser ces deux types de contribution :

- La production étudiante qui représente la cartographie de la relève académique et du potentiel futur.
- La production des enseignants-chercheurs qui reflète l'expertise active et établie.

Une gouvernance des métadonnées qui différencie clairement ces apports produirait des analyses bien plus fines et stratégiques pour les décideurs politiques et institutionnels au sein de l'espace CAMES.

Concernant la répartition des auteurs par type de document, les données confirment qu'elle épouse la structure générale de la collection. Le graphique illustrant la nationalité des auteurs de mémoires agit ainsi comme un indicateur de la répartition géographique de la formation académique au sein de l'espace CAMES, un renseignement précieux pour une diplomatie éducative et scientifique ciblée.

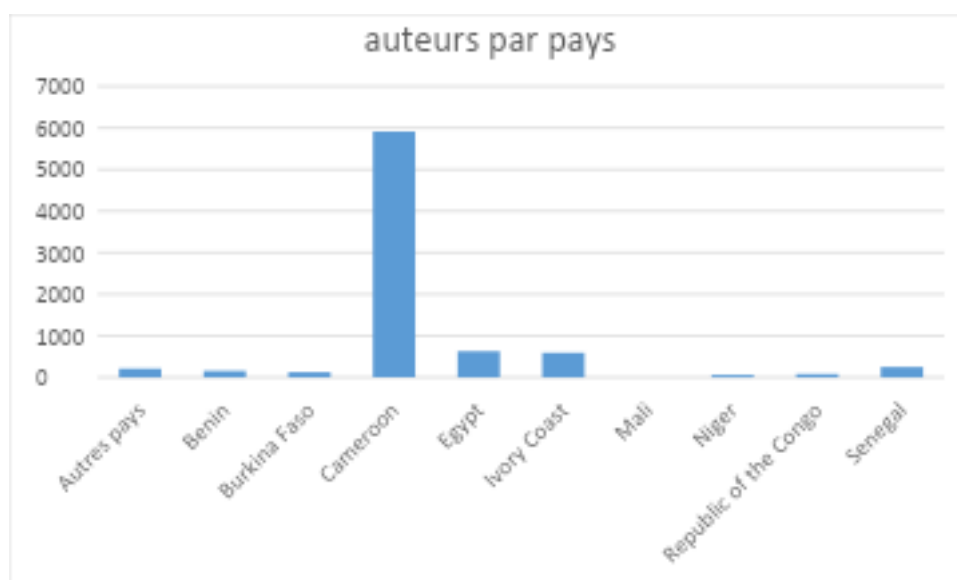


Figure 12. Représentation des auteurs par pays dans la collection DICAMES. Données : DICAMES. Graphique : Yannick Ngalli, 2024

L'histogramme ci-dessus présente les auteurs ayant contribué à la production de mémoires par pays. Le fait d'avoir 6 000 auteurs pour moins de 3 000 mémoires illustre notre propos selon lequel les directeurs et membres du jury sont cités comme auteurs. Cette situation rend les statistiques des auteurs difficiles à analyser parce qu'on y mélange des personnes ayant des fonctions très différentes (candidat, encadreur, membre de jury). Cette difficulté s'applique aussi aux auteurs par établissement comme l'illustre le graphique ci-dessous.

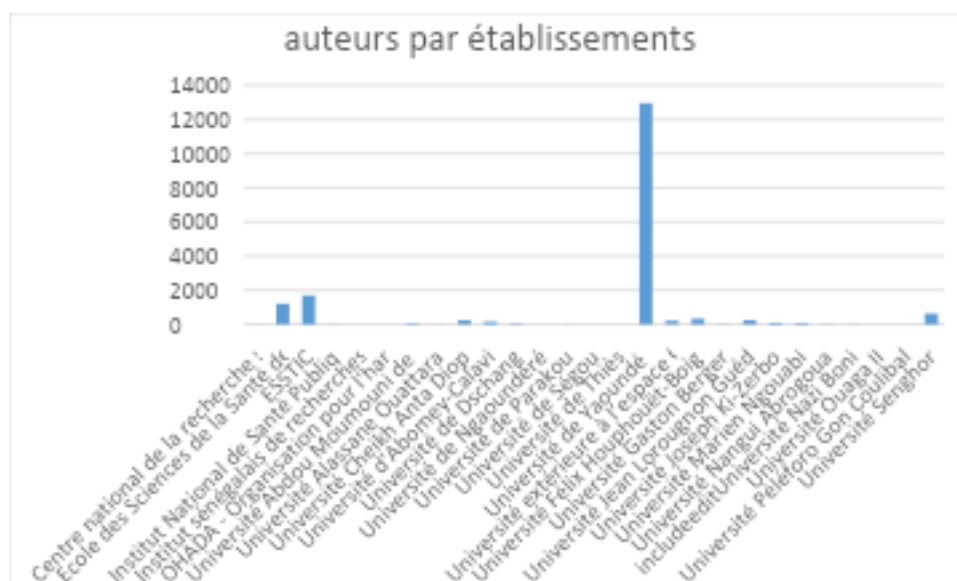


Figure 13. Représentation des auteurs par établissements dans la collection DICAMES. Données : DICAMES. Graphique : Yannick Ngalli, 2024

Le nombre d'auteurs par établissement est supérieur aux documents disponibles dans la base. Par exemple, le nombre d'auteurs de l'université de Yaoundé est l'équivalent du double de documents disponibles dans la base. Cette situation indique que le choix des métadonnées (dans ce cas des auteurs) et la façon de les représenter ont une incidence sur la construction de l'information dans une collection numérique.

Ce constat est aussi valable pour l'analyse des auteurs en fonction du type de documents.

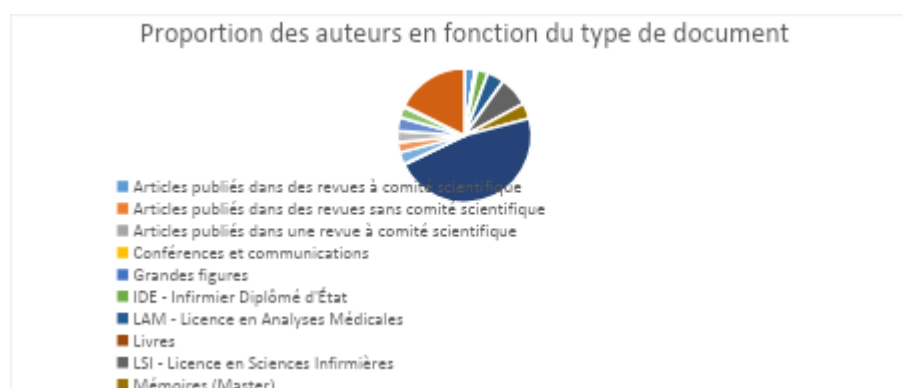


Figure 14. Proportion des auteurs en fonction du type de document dans la collection DICAMES. Données : DICAMES. Graphique : Yannick Ngalli, 2024

Malgré le fait que les auteurs aient représenté indifféremment leurs fonctions, la représentation graphique ci-dessus exprime que la proportion des auteurs par type de documents est proche de la proportion de ces documents dans la collection. La conclusion est que les mêmes critères de choix d'attribution de la qualité d'auteur ont été appliqués à tous les types de documents. Cela ne change pas les équilibres en termes de proportions dans la collection.

LES DÉFIS DE LA GOUVERNANCE DES DONNÉES

Les résultats présentés manifestent des choix de gouvernance, explicites ou implicites, structurant DICAMES. Cette discussion en tire des enseignements sur la gouvernance des données et ses implications géopolitiques.

L'analyse des métadonnées de DICAMES révèle des enjeux critiques qui compromettent son intégrité et son utilité en tant qu'outil de représentation scientifique. Les principaux biais identifiés sont les suivants: (1) la confusion des rôles sous l'entité unique "auteur" (créateur, directeur, jury), rendant les analyses bibliométriques irrecevables ; (2) les erreurs de saisie dans les titres ; (3) l'absence de vocabulaire contrôlé pour les mots-clés (plus de 14 000 occurrences non normalisées) ; (4) les incohérences typologiques (ex. : "thèses soutenues" vs "thèses").

L'absence d'alignement sur des standards internationaux tels que le format l'UNIMARC (UNiversal Machine Readable Cataloging), ou RDA (*Resource Description and Access*) limite fortement l'interopérabilité, la précision et la réutilisabilité des métadonnées. Une telle normalisation aurait permis une structuration fine des informations (données bibliographiques, descriptives, de gestion et d'accès), facilitant ainsi leur échange et leur exploitation à l'échelle internationale. Cette situation peut être interprétée comme un choix de gouvernance qui privilégie la simplicité de la collecte au détriment de la qualité des données et de leur potentiel de valorisation.

Pour garantir la qualité des données, une double validation - humaine et technique - est essentielle. Celle-ci passerait par la mise en place d'outils de contrôle automatisés couplés à une supervision par des experts métier.

L'exhaustivité limitée : une question de souveraineté informationnelle

Avec seulement 8 000 documents pour 19 pays, des centaines d'établissements et plusieurs millions d'étudiants, DICAMES représente une fraction infime de la production scientifique réelle. Ce déficit révèle une fragmentation institutionnelle et un manque de politiques incitatives. Pour devenir un outil diplomatique crédible, DICAMES doit atteindre une masse critique, ce qui suppose des circuits de collecte simplifiés, des campagnes de sensibilisation et des engagements formels avec les établissements.

DICAMES, un outil de diplomatie des données ?

La diplomatie des données (Morin et al., 2021), déclinaison technique de la diplomatie scientifique (Ruffini, 2017), recouvre les négociations et stratégies visant à contrôler et valoriser les flux de données. DICAMES pourrait devenir un instrument stratégique à trois niveaux.

Les trois dimensions diplomatiques potentielles de DICAMES

- Contrôler la narration scientifique : face aux bases internationales aux critères nord-centrés (Piron et al., 2017), DICAMES peut définir localement ce qui compte comme production scientifique (langues, types de documents, thématiques).
- Négocier des partenariats équitables : DICAMES peut servir de levier pour négocier des accords de moissonnage avec OpenAIRE, BASE, CORE, en préservant l'intégrité des métadonnées, la mention de l'origine africaine, et la réciprocité des flux.
- Renforcer la coopération Sud-Sud : DICAMES peut institutionnaliser le rôle de « hubs » informationnels (Cameroun, Côte d'Ivoire) par des accords de mutualisation, de mobilité virtuelle des données, et des communautés thématiques panafricaines.

Les obstacles actuels à une diplomatie des données effective

Quatre obstacles limitent le potentiel diplomatique de DICAMES :

- Technique : l'absence de standardisation (UNIMARC, Dublin Core) empêche l'interopérabilité.
- Politique : seuls 8 pays sur 19 contribuent, révélant priorités divergentes et méfiance envers une gouvernance centralisée.

- Épistémique : l'absence d'indicateurs stratégiques (consultations, collaborations) ne permet pas une diplomatie éclairée.
- Juridique : les législations disparates et l'absence de protocole régional sur les données fragilisent l'ambition diplomatique.

Des prérequis pour une diplomatie des données crédible

Les trois conditions préalables ci-dessous constituent une feuille de route que nous proposons :

Condition	Traduction opérationnelle
Gouvernance technique	Normalisation des métadonnées (Dublin Core, MARC21), mise en place d'une API, validation humaine et automatisée
Gouvernance institutionnelle	Signature d'une charte des données CAMES, création d'un comité de pilotage inter-états, rapports annuels sur l'état de la contribution
Gouvernance stratégique	Production d'indicateurs diplomatiques (cartes des collaborations, indice de visibilité africaine, tableau de bord des téléchargements Sud-Sud)

Tableau 1 : prérequis dans la diplomatie des données

DICAMES peut-il prétendre au soft power académique ?

Le soft power académique (Nye, 2004) désigne la capacité à influencer par l'attraction et l'exemplarité. À ce stade, DICAMES n'en génère pas : sa masse critique (8 000 documents) reste faible face à HAL (1 million), et les contributions limitées réduisent sa légitimité. Cependant, si DICAMES quadruplait ses documents, obtenait une interopérabilité et produisait des analyses bibliométriques originales, il deviendrait un outil de data power.

DICAMES possède les briques élémentaires d'un outil de diplomatie des données (assise institutionnelle, libre accès, diversité encyclopédique), mais sa transformation diplomatique est conditionnée par des choix techniques (normalisation), institutionnels (engagements étatiques) et stratégiques (indicateurs). Sans eux, il restera une archive utile sans rayonnement ; avec eux, il deviendra un modèle africain de data diplomacy.

Limites de l'étude

Cette recherche, bien qu'elle apporte des éclairages inédits sur la gouvernance des données dans les archives ouvertes africaines, comporte plusieurs limites qu'il convient d'explicitier pour situer la portée de ses conclusions. Ces limites sont de nature épistémologique, technique, méthodologique et contextuelle.

- Limites liées à la source des données

L'étude repose exclusivement sur les métadonnées, sans accès aux PDF ni aux logs d'usage, ce qui empêche de vérifier la correspondance métadonnées/contenu, d'évaluer la circulation effective des données, ou de mesurer l'impact réel. L'analyse est donc centrée sur l'offre informationnelle plutôt que sur la demande.

- Limites techniques du *scraping* et de la constitution du corpus

Le *scraping* manuel comporte des risques (omissions liées à l'architecture, absence d'API, nettoyage manuel des caractères spéciaux). La reproductibilité est assurée par le dépôt du script sur GitHub, mais une API officielle offrirait un corpus plus fiable.

- Limites de l'interprétation géopolitique

L'interprétation géopolitique des disparités (surreprésentation du Cameroun, faibles contributions sahéniennes) nécessite une enquête qualitative complémentaire. Des facteurs contextuels non mesurés (instabilité, grèves, absence de politiques nationales) ou l'effet

"institution-proximité" (hébergement initial au Cameroun) peuvent expliquer ces écarts. De plus, les productions en anglais ou langues locales échappent à DICAMES. Des entretiens seraient nécessaires pour distinguer incapacités techniques et choix politiques.

- Absence de comparaison avec d'autres archives ouvertes

L'étude est monographique et ne compare pas DICAMES à d'autres archives (AJOL, IRD, HAL), ce qui limite la généralisation. Une analyse multi-cas serait une extension naturelle. Par ailleurs, l'extraction à un instant T (octobre 2024) offre une photographie et non un film longitudinal, ne permettant pas de mesurer les dynamiques réelles ni l'impact des politiques incitatives.

- Limites temporelles et diachroniques

Les données ont été extraites à un instant T (octobre 2024). Or, DICAMES est une base évolutive. Une analyse longitudinale (plusieurs extractions à intervalles réguliers) contribuerait à mesurer la dynamique réelle des dépôts (et non un simple instantané), identifier des effets de « rattrapage » ou de « décrochage » de certains pays, évaluer l'impact des campagnes de sensibilisation ou des politiques incitatives.

Notre étude offre une photographie, non un film, sans distinction des tendances structurelles des variations conjoncturelles.

CONCLUSION

Cet article a mis en lumière le potentiel de DICAMES comme patrimoine scientifique de l'Afrique francophone, tout en identifiant des défis majeurs (biais dans les métadonnées, disparités géographiques, temporalité limitée) révélateurs d'une gouvernance implicite privilégiant la collecte à la qualité. Pour que DICAMES devienne un outil stratégique de souveraineté scientifique, deux chantiers prioritaires s'imposent : (1) la normalisation des métadonnées via des standards internationaux (UNIMARC, Dublin Core) et un contrôle qualité rigoureux ; (2) l'accroissement de l'exhaustivité par des politiques incitatives ou contraignantes auprès des établissements membres. Cette transformation exige une gouvernance active des données articulée autour de trois piliers : technique (standards et validation automatisée), institutionnelle (mécanismes incitatifs), et stratégique (diplomatie scientifique et collaborations Sud-Sud). Une stratégie de communication ambitieuse est nécessaire pour fédérer les acteurs autour de ce bien commun.

Au-delà des aspects techniques, cet article met en évidence la manière dont l'analyse bibliométrique des métadonnées révèle les mécanismes de gouvernance d'une infrastructure informationnelle, contribuant aux débats sur la souveraineté informationnelle et sur la diplomatie des données dans les pays du Sud.

RÉFÉRENCES BIBLIOGRAPHIQUES

Archambault, Éric ; Vignola-Gagné, Étienne (2004), *The use of bibliometrics in the social sciences and humanities*, Montreal : Social Sciences and Humanities Research Council of Canada (SSHRCC).

Berthaud, C.; Charnay, D.; Fargier, N. (2021). Diffuser et pérenniser le savoir scientifique : 20 ans d'histoire de HAL. Histoire de la recherche contemporaine : la revue du Comité pour l'histoire du CNRS, X(2), p. 139-149. <https://doi.org/10.4000/hrc.6330>

Borgman, Christine L. (2015), *Big Data, Little Data, No Data: Scholarship in the Networked World*, Cambridge : MIT Press.

Callon, Michel (1991), « Réseaux technico-économiques et irréversibilités », *Figures de l'irréversibilité en économie*, vol. 38, n° 1, p. 195-230.

Deboin, Marie-Claude (2015), « Déposer ses publications dans une archive ouverte, en 8 points », [en ligne], consulté le 02 mai 2026, <https://agritrop.cirad.fr/575303/>.

Fabre, Isabelle ; Gardiès, Cécile (2008), « L'accès à l'information scientifique numérique : organisation des savoirs et enjeux de pouvoir dans une communauté scientifique », *Sciences de la société*, n° 75, p. 71-84.

Morin, Jean-Frédéric ; Louafi, Selim ; Orsini, Amandine (2021), « Data Diplomacy: A New Frontier in International Relations », *Global Policy*, [en ligne], consulté le 02 mai 2026, <https://doi.org/10.1111/1758-5899.12948>.

Nye, Joseph S. (2004), *Soft Power: The Means to Success in World Politics*, New York : PublicAffairs.

Piron, Florence ; Diouf, Antonin B. ; Dibounje Madiba, Marie S. ; Mboa Nkoudou, Thomas H. ; Aubierge Ouangré, Zoé ; Tessy, Daphney R. ; ... ; Lire, Zakari (2017), « Le libre accès vu d'Afrique francophone subsaharienne », *Revue française des sciences de l'information et de la communication*, n° 11, [en ligne], consulté le 02 mai 2026, <http://journals.openedition.org/rfsic/3292>.

Ruffini, Pierre-Bruno (2017), *Science and Diplomacy: A New Dimension of International Relations*, Cham : Springer.

Schöpfel, Joachim (2020), « L'usage de la plateforme HAL par des unités de recherche : Le cas de l'Université de Lille », *I2D - Information, données & documents*, vol. 57, n° 3, p. 167-198.

Schöpfel, Joachim ; Kergosien, Éric ; Prost, Hélène ; Barrié, Julien (2022), « Pas si simple que ça... : Une enquête sur l'usage de HAL par les unités de recherche des universités IdEx », *I2D - Information, données et documents*, vol. 59, n° 2, p. 150-183.

Trzmielewski, Marcin ; Gnoli, Claudio (2019), « Une classification interdisciplinaire pour l'échange et la médiation des données ouvertes de la recherche », *Actes du 12ème Colloque international d'ISKO-France, Montpellier, 9-10 octobre 2019*.

Yin, Robert K. (2009), *Case study research: Design and methods*, Los Angeles : Sage (vol. 5).

ANNEXE

Pays contributeurs de DICAMES, les établissements associés à ces pays et le nombre de documents par établissements.

Pays	Etablissement d'enseignement supérieur	Nombre de documents
1-Burkina Faso	Centre national de la recherche scientifique et technologique - CNRST	6
	Institut National de Santé Publique	11
	Université Joseph Ki-Zerbo	101
	Université Nazi Boni	13
	Université Ouaga II	1
2- Bénin	2.1- Université d'Abomey-Calavi	151
	2.2- Université de Parakou	6
3- Cameroun	3.1- Ecole des Sciences de la Santé de l'Université Catholique d'Afrique Centrale - (ESS-UCAC)	1194
	3.2- OHADA - Organisation pour l'harmonisation en Afrique du droit des affaires	4
	3.3- Université de Dschang	52
	3.4- Université de Yaoundé I	3108
	3.5- École Supérieure des Sciences et Techniques de l'Information et de la Communication - ESSTIC	2113
4- Congo	4.1- Université Marien Ngouabi	81
5- Côte d'Ivoire	5.1- Université Alassane Ouattara	7
	5.2- Université Félix Houphouët-Boigny	328
	5.3- Université Jean Lorougnon Guédé	241
	5.4- Université Nangui Abrogoua	17
	5.5- Université Péléforo Gon Coulibaly	1
6- Egypte	6.1- Université Senghor	630
7- Mali	7.1- Université de Ségou	1
8- Niger	8.1- Université Abdou Moumouni de Niamey	72

Traitement : Yannick NGALLI, 2024.

Le « Data Space Education Skills » (DASES) ou la construction d'une idéologie européenne des données éducatives au regard de la domination des Big Tech sur le numérique éducatif

Article inédit, mis en ligne le 23 septembre 2026

Hubert Boët

*Docteur en sciences de l'information et de la communication. Il a soutenu en 2025 une thèse intitulée « Analyse contrastive de l'industrialisation des médias d'apprentissage adaptatif en France et aux États-Unis » pour laquelle il a obtenu le prix de thèse du GIS 2if. Il réalise actuellement un postdoctorat au GRIPIC/Sorbonne Université dans le cadre du programme PostGenAI@Paris, et est chercheur associé au LabSIC.
hubert.boet@yahoo.fr*

Plan de l'article

Résumé et mots clés en français

Title, abstract and keywords in english

Titulo, resumen y palabras clave en espanol

Introduction

Une éducation fondée sur les données, une utopie industrielle à l'heure de l'emprise des Big Tech sur le numérique européen

Une « idéologie infrastructurelle » fragilisée par les contraintes industrielles et diplomatiques du cloud européen

Le DASES, témoin de la puissance économique et idéologique des plateformes

Conclusion

Références bibliographiques

Notes

RÉSUMÉ

En 2020, la Commission européenne a présenté une « Stratégie européenne pour les données », visant à faire des données numériques une opportunité industrielle et politique. Cette stratégie a donné naissance au projet d'un espace de données sécurisé et partagé, le « *Data Spaces Education Skills* » (DASES), en vue de bâtir des services européens d'IA consacrés aux compétences. L'article analyse les rapports de force industriels et diplomatiques que le DASES et ses trois appels à projets sous-tendent, et évalue leur degré de rupture avec la mainmise économique et idéologique des Big Tech sur le numérique éducatif européen. Pour cela, il mêle une approche socio-économique de la filière des logiciels éducatifs aux *infrastructure* et *platform studies* et à l'analyse dynamique de l'innovation de Patrice Flichy. L'analyse des discours programmatiques révèle une utopie connexionniste « centrée sur l'humain », consistant en un projet social-libéral permis par des données et des IA « souveraines ». Cependant, la concrétisation des cas d'usage montre des rapports de force asymétriques, alors que la faiblesse des moyens engagés et la régulation erratique du cloud ne remettent pas en cause l'hégémonie des *Big Tech*. À l'aune de la théorie de l'industrialisation de la formation, ce programme témoigne en outre d'une platformisation de l'éducation, adoptant les orientations des plateformes éducatives et des offres de

Big Tech.

Mots clés

Souveraineté technologique, espaces de données, industrialisation de l'éducation, plateformes de l'éducation, diplomatie numérique, Big Tech, cloud

TITLE

The 'Data Space Education Skills' (DASES) or the construction of a European ideology of educational data in light of Big Tech's domination of digital education

Abstract

In 2020, the European Commission presented a 'European Data Strategy' aimed at making digital data an opportunity for politics and industry. This strategy gave rise to the project of a safe and shared data space, the 'Data Spaces Education Skills' (DASES), in order to develop European AI services dedicated to competencies. The article analyzes the industrial and diplomatic power relations underlying DASES and its three calls for projects and assesses the degree to which they break with the economic and ideological hold of Big Tech on European educational digital technology. To do this, it combines a socio-economic approach to the educational software sector with infrastructure and platform studies, and Patrice Flichy's dynamic analysis of innovation. The analysis of programmatic discourses reveals a 'human-centered' connectionist utopia, consisting of a social-liberal project made possible by 'sovereign' data and AI. However, the implementation of use cases reveals asymmetric power relations, while the limited resources committed and the inconsistent regulation of cloud computing fail to challenge the hegemony of Big Tech. From the perspective of the theory of the industrialization of education, this program also reflects a platformization of education, adopting the orientations of educational platforms and Big Tech offerings

Keywords

Digital sovereignty, data spaces, industrialization of education, platformization of education, digital diplomacy, Big Tech, cloud computing

TITULO

El Data Space Education Skills (DASES) o la construcción de una ideología europea de los datos educativos frente a la dominación de las Big Tech sobre lo digital educativo

Resumen

En 2020, la Comisión Europea presentó una «Estrategia Europea de Datos», cuyo objetivo era convertir los datos digitales en una oportunidad política e industrial. Esta estrategia dio origen al proyecto de un espacio de datos seguro y compartido, el «Data Spaces Education Skills» (DASES), con el fin de construir servicios europeos de IA dedicados a las competencias. El artículo analiza las relaciones de poder industriales y diplomáticas que el DASES y sus tres convocatorias de proyectos conllevan, y evalúa su grado de ruptura con el dominio económico e ideológico de las Big Tech sobre lo digital educativo europeo. Para ello combina un enfoque socioeconómico del sector de los softwares educativos con las *infrastructure* y

platform studies, así como con el análisis dinámico de la innovación de Patrice Flichy. El análisis de los discursos programáticos revela una utopía conexionista 'centrada en el ser humano', que consiste en un proyecto socioliberal posibilitado por datos e IA 'soberanos'. Sin embargo, la implementación de los casos de uso revela relaciones de poder asimétricas, mientras que la debilidad de los recursos movilizados y la regulación errática del cloud no cuestionan la hegemonía de las Big Tech. Desde la perspectiva de la teoría de la industrialización de la educación, este programa también refleja una plataformización de la educación, adoptando las orientaciones de las plataformas educativas y de las ofertas de las Big Tech.

Palabras clave

Soberanía tecnológica, espacios de datos, industrialización de la educación, plataformización de la educación, diplomacia digital, Big Tech, cloud

INTRODUCTION

Contexte social

En février 2020, la Commission européenne publie la « Stratégie européenne pour les données », soulignant leur rôle central pour le développement industriel du continent. Elle identifie toutefois trois obstacles majeurs (Commission européenne, 2020, p.3). Le premier est d'ordre commercial : des « contrôleurs d'accès » (*gatekeepers*), dominent les services en ligne via le cloud et les plateformes. Ce sont notamment les Big Tech, multinationales des technologies dont les GAFAM (Google, Amazon, Facebook devenu Meta, Apple et Microsoft) sont les représentants les plus identifiés du côté états-unien. Depuis les années 2010, ces acteurs se sont étendus « par cercles concentriques » (Thuillas et Wiart, 2023), allant de leur cœur de métier initial à l'ensemble des infrastructures et interfaces numériques (cloud, systèmes d'exploitation, e-commerce, navigateurs, etc.), imposant ainsi leurs écosystèmes et contrôlant l'accès aux données, au détriment des entreprises européennes. En découle le deuxième obstacle : l'extraterritorialité du droit états-unien, qui peut contraindre ces entreprises à transférer des données collectées en Europe vers les autorités américaines, menaçant la protection des données personnelles. Troisièmement, la Commission pointe une faible collaboration entre acteurs européens, liée à un déficit de compétences numériques, de confiance dans les logiciels, de gouvernance claire et de standards d'interopérabilité (*ibid.*, pp.7-13).

Face à ces enjeux, la Stratégie européenne vise la croissance d'entreprises continentales dans un cadre éthique, dans la continuité du Règlement Général sur la Protection des Données (RGPD). Elle s'appuie sur une initiative franco-allemande : la création en 2020 de GALA-X, association visant à établir un cadre européen sécurisé pour les échanges de données dans le cloud, avec la participation d'acteurs industriels des divers pays européens. La Stratégie promeut un « marché commun des données » et prévoit des investissements sectoriels pour favoriser l'échange de données et le développement d'une IA « éthique ». Parmi les domaines ciblés figurent les compétences, à travers le programme « *Data Space Education Skills* » (DASES), destiné à réduire l'écart entre formation et besoins du marché du travail. Ce programme est soutenu par plusieurs financements du plan « *Digital Europe : DS4Skills* (2022-2023, 1 M€), *Edge-Skills* (2022-2024, 23 M€) et *DS4Skills-GO* (2025-2027, 3 M€).

Ces initiatives s'inscrivent dans des objectifs de souveraineté technologique, de cybersécurité, de protection des valeurs européennes et de libre circulation des données. L'article propose d'analyser les rapports de pouvoir et les équilibres socio-économiques et politiques sous-jacents au DASES, ainsi que sa capacité à remettre en cause, ou non, la domination des Big Tech dans le secteur des logiciels éducatifs.

Cadre conceptuel et méthodologie

Pour répondre à cette interrogation, nous mobilisons un cadre théorique en cinq volets. D'abord, une approche socio-économique fondée sur l'économie politique de la communication, permettant de modéliser les enjeux de pouvoir traversant la production industrielle de la culture et de l'information, afin d'évaluer la possible remise en cause de la domination des Big Tech sur la filière des logiciels éducatifs.

Deuxièmement, nous examinons si le DASES renouvelle des logiques d'infrastructures et de plateformes. Les recherches en sciences de l'organisation ont défini les infrastructures comme des systèmes sociotechniques essentiels à la vie sociale, caractérisés par l'omniprésence, la fiabilité, l'invisibilité, les passerelles et les pannes (Plantin *et al.*, 2018). Les plateformes ont pour leur part été définies comme des architectures logicielles combinant des composants principaux stables et des composants complémentaires variables (Baldwin et Woodward, 2008), organisant des places de marchés dont les gestionnaires captent des revenus et des externalités positives (Bullich, 2015). Néanmoins, comme le souligne Vincent Bullich, ces recherches conduisent une perspective techniciste, qui n'envisage les stratégies industrielles que comme des effets du « design technique ». Nous empruntons pour notre part une approche info-communicationnelle qui nous permettra d'appréhender les rapports de pouvoir pluriels que ces dispositifs contribuent à naturaliser.

Troisièmement, nous appréhendons le contenu même du projet industriel du DASES sous le prisme de l'industrialisation de la formation développé par Pierre Mœglin (2016). Ainsi nous envisageons l'éducation comme obéissant à des tendances sociétales marquées par trois tendances : la *technologisation*, qui peut être définie comme le recours à des techniques qui véhiculent des manières de faire parmi les publics éducatifs ; la *rationalisation*, ou la tendance à optimiser la pédagogie selon des critères d'efficacité ; et l'*idéologisation*, autrement dit la tendance à fédérer par des « grands récits » des acteurs aux croyances et intérêts divergents. Nous éclairerons le projet éducatif et sociétal derrière le DASES à travers les filiations historiques liées au déploiement du numérique et de l'IA dans l'éducation.

Quatrièmement, la richesse de ce cas d'étude réside dans l'enchaînement d'un programme de préfiguration (*DS4Skills*) et de deux programmes de déploiement (*Edge-Skills* et *DS4Skills-GO*). Cela permet de déceler l'évolution des idéaux initiaux vers des principes pragmatiques. Nous utiliserons le modèle d'analyse de l'imaginaire technique de Patrice Flichy (2001) pour caractériser la transformation d'une *utopie* éducative contestant l'ordre établi en une *idéologie* légitimant le pouvoir d'une nouvelle organisation technique et industrielle.

Cinquièmement, une originalité de notre contribution réside dans la cohabitation entre ces questionnements industriels et les travaux portant sur la diplomatie numérique, qui met l'accent sur les stratégies d'influence liées aux plateformes et « l'adaptation des pratiques diplomatiques aux avancées technologiques » (Lyubareva et Nocetti, 2024). Nous développerons une réflexion sur les liens entre la diplomatie européenne sur les données et l'industrialisation de l'éducation à l'heure de l'IA. La première servirait-elle la seconde en contribuant à standardiser l'éducation en Europe sous couvert des idéaux de souveraineté et d'horizontalité ? Quelles règles du jeu diplomatiques entre États membres et entre l'Union et les puissances étrangères sont-elles promues à travers ce projet industriel ?

Notre méthodologie est double. D'une part, nous menons une analyse socio-économique du numérique éducatif, en croisant travaux existants, politiques européennes et stratégies des Big Tech, notamment autour du cloud souverain. D'autre part, nous réalisons une analyse de discours diachronique sur trois corpus de textes produits par des informaticiens, des acteurs politiques et des entrepreneurs impliqués dans ces projets : des textes stratégiques sur les *data spaces* (2020-2023), des documents spécifiques au DASES (appels à projets et feuille de route), et des cas d'usage récents. L'objectif est de suivre, au fil de textes contemporains émanant d'acteurs directement aux prises avec le développement de ces dispositifs, le passage d'une *utopie* à une *idéologie* autour de la donnée. Ces corpus sont également

étudiés au prisme de l'industrialisation de la formation afin d'identifier continuités et ruptures vis-à-vis de conceptions industrialistes déjà identifiées par les chercheurs. La frise suivante retrace les divers événements liés aux espaces de données et resitue, en jaune, l'enchaînement chronologique des documents constituant nos corpus de documents officiels (Figure 1).

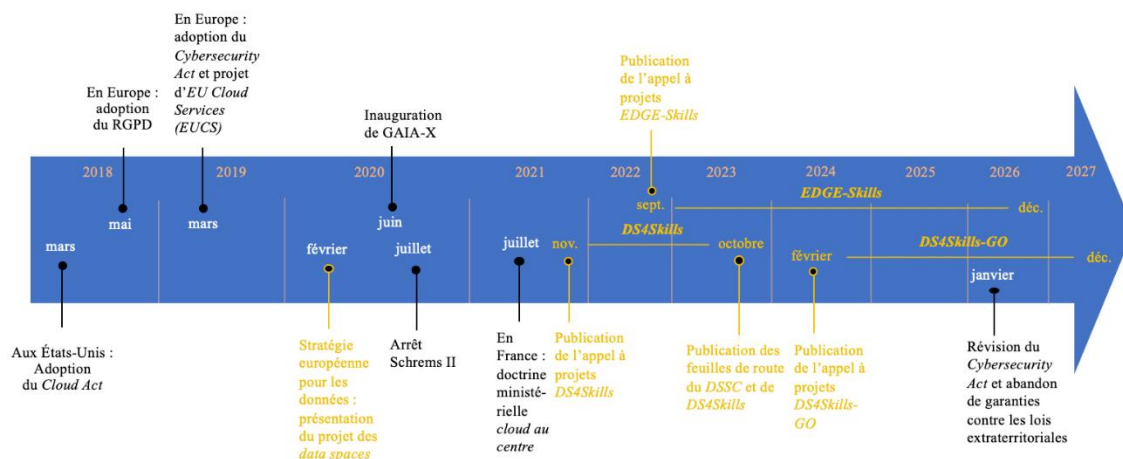


Figure 1. Chronologie retenue dans l'article pour le DASES. En jaune : les documents figurant dans nos corpus officiels.

Notre article se décompose en trois parties. Une première expose les promesses libérales, décentralisées et de souveraineté contenues dans l'utopie du DASES, en résonnance avec le contexte inégalitaire de la socio-économie européenne des logiciels éducatifs. Une deuxième partie aborde les réalisations erratiques de ces promesses, tant du point de vue des cas d'usage asymétriques que du contexte diplomatique contraire autour du cloud européen. Une troisième partie revient sur l'hypothèse d'un renouvellement du projet éducatif porté par les parties prenantes du DASES vis-à-vis des offres de Big Tech.

UNE ÉDUCATION FONDÉE SUR LES DONNÉES, UNE UTOPIE INDUSTRIELLE À L'HEURE DE L'EMPRISE DES BIG TECH SUR LE NUMÉRIQUE EUROPÉEN

Les services de *cloud computing*, pivots de l'influence des Big Tech auprès des start-ups

Pour comprendre les enjeux socio-économiques du DASES, il faut replacer ce projet dans le contexte industriel des logiciels éducatifs en Europe. Nous reprenons une modélisation en trois étapes, animées par cinq types de firmes (Boët, 2025), visible en Figure 2 : en amont, la production de contenus pédagogiques (menée par les éditeurs scolaires, les acteurs du numérique éducatif) et la conception d'algorithmes (animée par des start-ups dites « EdTech ») ; en aval, la distribution des services en mode SaaS via des plateformes d'intermédiation. Les Big Tech occupent une position centrale en intervenant à chaque étape de la chaîne.

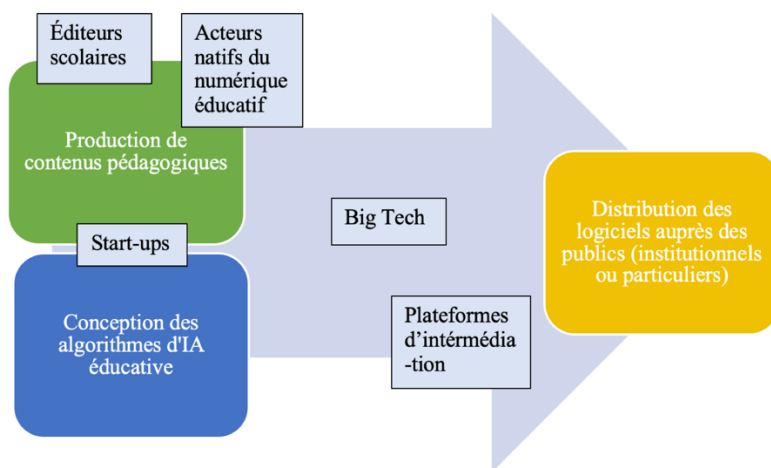


Figure 2. Modélisation de la filière des logiciels éducatifs

Un facteur clé de cette domination est leur contrôle sur le *cloud computing*, indispensable au développement des logiciels d'IA. Ces infrastructures donnent la possibilité de collecter, de stocker et de traiter les données tout en réduisant les coûts. Or le marché mondial est dominé par trois acteurs : Amazon Web Services (29 %), Microsoft Azure (20 %) et Google Cloud (13 %) au troisième trimestre 2025 (Synergy Research Group, 2025). Dans l'éducation, ces entreprises ne se limitent pas à l'hébergement : elles proposent aussi aux firmes éditrices des services algorithmiques reposant notamment sur leurs IA génératives pour leur permettre de générer des contenus pédagogiques, ainsi que des programmes d'incubation (crédits cloud, accompagnement), renforçant la dépendance des start-ups (Boët, 2025). Ce modèle favorise la centralisation des données dans des « *walled gardens* » (Plantin *et al.*, 2018), limitant leur accès à d'autres acteurs, notamment européens. Les Big Tech contrôlent ainsi un écosystème où les données d'apprentissage deviennent un levier stratégique.

Parallèlement, les Big Tech développent des services éducatifs directement auprès des établissements et des individus. Les suites comme *Google Workspace for Education* et *Microsoft 365 Education* se sont imposées, notamment lors de la pandémie, grâce à leur adaptabilité aux pics de fréquentation. Elles proposent des environnements numériques de

travail, des plateformes de ressources et participent à la formation des enseignants. Ce mouvement d'« infrastructuralisation » des plateformes - processus par lequel des plateformes tendent à devenir des infrastructures indispensables à la vie en société (Plantin *et al.*, 2018) - transforme ces plateformes en services quasi-indispensables pour les offreurs comme pour les usagers des marchés éducatifs.

Face à cette influence croissante, les pouvoirs publics européens ont engagé des régulations. L'arrêt « Schrems II » (2020) a remis en cause les transferts de données vers les États-Unis. En France, la doctrine « cloud au centre » (2021) vise à protéger les administrations, tandis que la Commission nationale de l'informatique et des libertés (CNIL) recommande de limiter l'usage de certains services états-uniens dans l'enseignement supérieur (Assemblée nationale, 2022). Ainsi, les Big Tech, devenues incontournables pour les entreprises, les établissements et les usagers, font l'objet d'un encadrement accru.

Le DASES s'inscrit dans ce contexte, avec l'ambition de proposer une alternative européenne fondée sur le partage de données et des services d'IA éthiques. L'analyse des documents de préfiguration permet d'identifier, au sens de P. Flichy, une phase d'« utopie-projet », c'est-à-dire une alternative formalisée aux dispositifs existants. Portée par la Stratégie européenne de 2020, cette utopie repose sur trois « promesses sociotechniques » (Joly, 2010) principales, devant répondre à des problèmes sociétaux.

Souveraineté, pragmatisme social-libéral et société en réseau « centrée sur l'humain » : les trois promesses du DASES

La première promesse est celle de la souveraineté, entendue d'abord comme le contrôle des individus sur leurs données, face aux risques d'utilisations contraires au respect de leur vie privée, dans la lignée du RGPD. Le *Data Spaces Support Center* (DSSC) en fait un objectif substantiel dans sa feuille de route de 2023 : « *Un espace de données est un système distribué défini par un cadre de gouvernance, qui permet des transactions de données fiables entre les participants tout en soutenant la confiance et la souveraineté des données* » (*Data Spaces Support Center*, 2023).

La souveraineté est ensuite entendue sur un plan technique : l'Union cherche à renforcer sa capacité d'action autonome dans les dispositifs numériques, face aux dépendances techniques induites par les *gatekeepers*, tout en restant ouverte à des coopérations internationales. Dans une note de juillet 2020, l'*European Parliamentary Research Service* (EPRS), service de recherche du Parlement européen, définit celle-ci comme « la capacité de l'Europe à agir de manière indépendante dans le monde numérique » et souligne qu'elle doit inclure « des mécanismes de protection et des outils offensifs pour favoriser l'innovation numérique (y compris en coopération avec des entreprises non européennes) » (Madiaga, 2020, p.1). La souveraineté constitue ainsi une réponse aux risques politiques et économiques liés à la domination des Big Tech.

La deuxième promesse formulée via le DASES est l'objectif de « réduire l'inadéquation des compétences entre le système d'éducation et de formation et les besoins du marché du travail » (Commission européenne, 2020, p.28). L'appel à projet-s de *DS4Skills* publié en novembre 2021 se réfère à une autre stratégie européenne, « *Digital Decade* », pour rappeler l'objectif de comblement d'un « fossé » des compétences numériques d'ici 2030 (Commission européenne, 2021, p.10). Ces textes programmatiques s'inscrivent dans une tradition européenne de « pragmatisme social-libéral » (Mattelart, 2001, p.90) où l'éducation est pensée comme un levier de compétitivité. Déjà en 2000, la stratégie de Lisbonne visait une « économie de la connaissance » ; en 2020, l'ambition devient celle d'une « économie fondée sur les données ». Le DASES prolonge cette logique en mobilisant les données et l'IA pour optimiser l'adéquation entre formation et emploi, dans un contexte de concurrence internationale pour les talents.

La troisième promesse est celle d'une société en réseau « centrée sur l'humain ». La feuille de route de *DS4Skills* projette un « écosystème » où individus, organisations éducatives,

entreprises et institutions publiques seraient interconnectées grâce aux données (*Data Space for Skills*, 2023a), comme le montre le visuel en Figure 3.

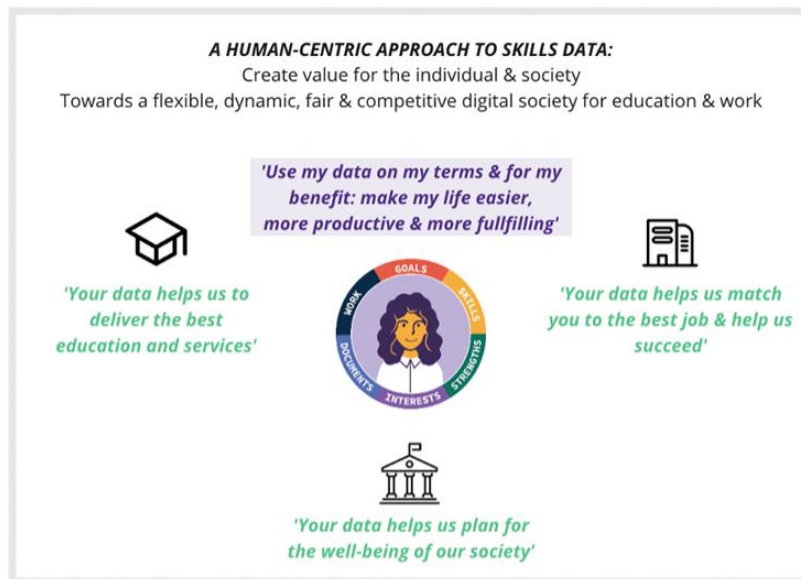


Figure 3. Schématisation de la doctrine centrée sur l'humain dans le cas du DASES.

Ce récit concorde avec le « nouvel esprit du capitalisme » théorisé par Ève Chiapello et Luc Boltanski (1999) et ses valeurs cardinales que sont l'interconnexion et l'assistance personnalisée, réactivées à travers les données et les services d'IA. Les plans détaillés du *DSSC* et de *DS4Skills*, publiés fin 2023, présentent les moyens techniques devant concrétiser cette promesse. Selon Flichy, la phase de maturation de l'utopie peut aboutir à un projet expérimental ou à une « pure fantasmagorie ». La première voie, celle de l'*utopie-projet*, est une « spécificité de l'utopie technique », où « l'auteur tente souvent de rendre concret son projet en construisant une maquette ou en réalisant une expérimentation technique » (Flichy, 2001, p.15). Le *DSSC* prend cette voie en appelant au développement de blocs de construction (*building blocks*) touchant à des aspects variés comme l'infrastructure, le mode de gouvernance ou le modèle d'affaires. Ces « blocs », destinés à être réutilisés dans des « cas d'usage » issus de différents secteurs socio-professionnels, sont présentés en Figure 4.

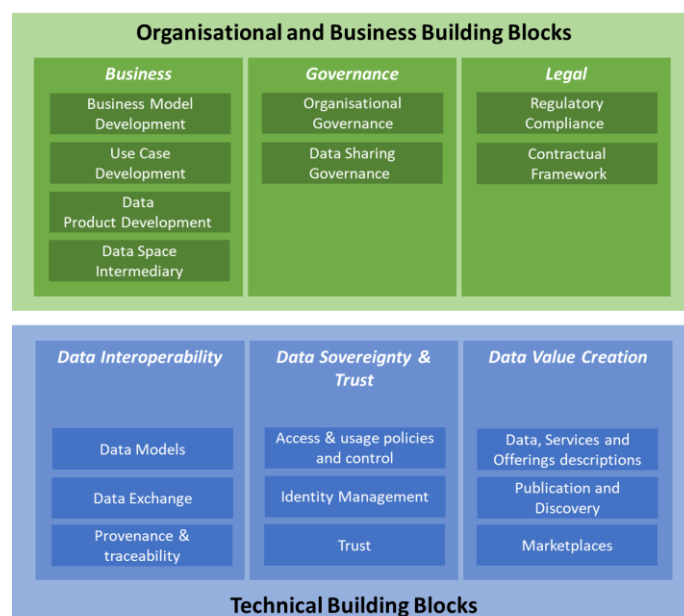


Figure 4. Schématisation des *building blocks* prônés par le *Data Space Support Center*.

Ces briques *open source* sont présentées comme des « communs numériques », infrastructures collectives permettant d'éviter les dépendances à des fournisseurs dominants (*Data Space for Skills*, 2023b). À rebours des stratégies d'accaparement des données menées par les *gatekeepers*, la référence à la théorie des « communs » d'Elinor Ostrom réactive des éléments de l'utopie séminale d'Internet décrite par P.Flichy, comme la promesse d'un partage et d'une circulation sans borne des connaissances (Flichy, 2001). Ces dispositifs traduisent le passage d'objets encore indéterminés (« objets-valises ») à des formes plus stabilisées (« objets-frontières »), selon P.Flichy (2003). Ils simplifient la complexité des secteurs pour proposer des standards techniques et organisationnels communs, favorisant l'interopérabilité et l'innovation. La gouvernance envisagée est multi-niveaux : une autorité de régulation (la *Data Space Gouvernance Authority* ou DSGA), des porteurs de cas d'usage et des participants disposant d'une certaine autonomie (*Data Space for Skills*, 2023c). Sur le plan économique, l'utopie promet des modèles d'affaires équilibrés, combinant ouverture et concurrence. Plusieurs options sont évoquées (abonnement, *freemium*, commissions, etc.), sans hiérarchisation claire. L'objectif est de permettre une circulation fluide des données entre acteurs autorisés, afin de stimuler l'innovation tout en garantissant la confiance.

Au final, le DASES ambitionne la création d'une infrastructure distribuée, reposant sur des standards communs et connectant des systèmes hétérogènes. Les programmes ultérieurs (*Edge-Skills*, *DS4Skills-GO*) visent à expérimenter ces principes à travers des cas d'usage concrets. Dans la perspective de P.Flichy, cette phase correspond à une transformation progressive de l'*utopie* en *idéologie* : les expérimentations réussies tendent à être généralisées et présentées comme des modèles, indépendamment des contextes spécifiques qui les ont rendues possibles. Il s'agit désormais d'examiner la manière dont cette utopie se confronte aux réalités socio-économiques et diplomatiques, et dans quelle mesure elle contribue à légitimer une nouvelle organisation industrielle, marquée par des rapports de pouvoir potentiellement asymétriques.

UNE « IDÉOLOGIE INFRASTRUCTURELLE » FRAGILISÉE PAR LES CONTRAINTES INDUSTRIELLES ET DIPLOMATIQUES DU CLOUD EUROPEEN

Une « idéologie infrastructurelle » en construction

Les programmes *Edge-Skills* et *DS4Skills-GO* prolongent la phase expérimentale du DASES. *Edge-Skills*, doté de 23 millions d'euros et coordonné par Prometheus-X, rassemble 28 organisations issues de sept pays. Selon le Livre blanc du projet, il vise à déployer une infrastructure de services cloud « centrée sur l'humain, distribuée et souveraine » (Prometheus-X, 2024, pp.6-7). *DS4Skills-GO*, financé à hauteur de 3 millions d'euros, soutient quant à lui des cas d'usage inspirés des travaux préparatoires de *DS4Skills*.

Ces initiatives peuvent être interprétées comme des étapes dans la construction d'une « idéologie infrastructurelle », « théorie du pouvoir et de la contestation *dans et à travers* les infrastructures médiatiques mondiales » selon Maxigas et Niels ten Oever (2023). Dans un article ultérieur, Niels ten Oever souligne que ce concept établit un lien entre la matérialité des infrastructures et la dimension discursive de l'idéologie, en ce qu'elles contribuent toutes deux à naturaliser une certaine organisation technique de la société, l'une par son omniprésence matérielle et l'autre par ses récits mobilisateurs (Ten Oever *et al.*, 2024). Dans le cas du DASES, ces rapports de pouvoir multiples seront étudiés à trois niveaux : celui de la gouvernance, celui des firmes participantes et celui des États membres.

Au niveau de la gouvernance, la DSGA joue un rôle clé. Chargée de développer les « *building blocks* » et de certifier les intermédiaires, elle détient un pouvoir structurant. L'association Prometheus-X, qui assure cette fonction, perçoit des financements publics et tire des revenus de redevances payées par les entreprises (4 à 20 % des revenus liés aux services d'IA). Bien que non lucrative, elle contrôle ainsi l'accès aux briques techniques.

Parmi celles-ci, le *Personal Data Intermediary* (PDI) est central. Présenté comme un tiers neutre garantissant la souveraineté des données, il simplifie les échanges en centralisant la gestion du consentement (*Data Spaces for Skills*, 2023d). Toutefois, cette simplification peut aussi réduire l'autonomie des acteurs en encadrant les choix et en facilitant l'exploitation des données par les fournisseurs de services. Les « *building blocks* » apparaissent ainsi comme des « objets-frontières » qui stabilisent les usages tout en concrétisant des rapports de pouvoir.

Au niveau des cas d'usage, les organisations porteuses (« *orchestrators* ») bénéficient d'un avantage stratégique. Des entreprises EdTech comme Maskott ont recours à des partenaires pour enrichir leurs services d'IA, dans des logiques proches de la sous-traitance. Témoins de cette situation, les deux projets que la firme française porte au sein d'*Edge-Skills* (« *Learning and teaching dashboards for adaptive learning* » et « *Recommender systems for teachers* ») pour lesquels elle mobilise l'expertise du laboratoire d'informatique LORIA (Université de Lorraine) et les données d'autres entreprises (Cabrilog et Rejustify) afin d'enrichir les tableaux de bord et systèmes de recommandation de son logiciel (Prometheus-X, 2026a ; Prometheus-X, 2026b).

Parallèlement, des intermédiaires comme VisionsTrust jouent un rôle central dans la mise en relation des acteurs via des plateformes d'échange de données, comme le révèlent les annonces publiées par cet acteur en Figure 5 (VisionsTrust, 2026). Ce rôle vaut à VisionsTrust des subventions conséquentes de la part de la Commission européenne, et lui permet de se rémunérer auprès des firmes participantes, qui s'abonnent pour pouvoir accéder à cet intermédiaire privilégié.

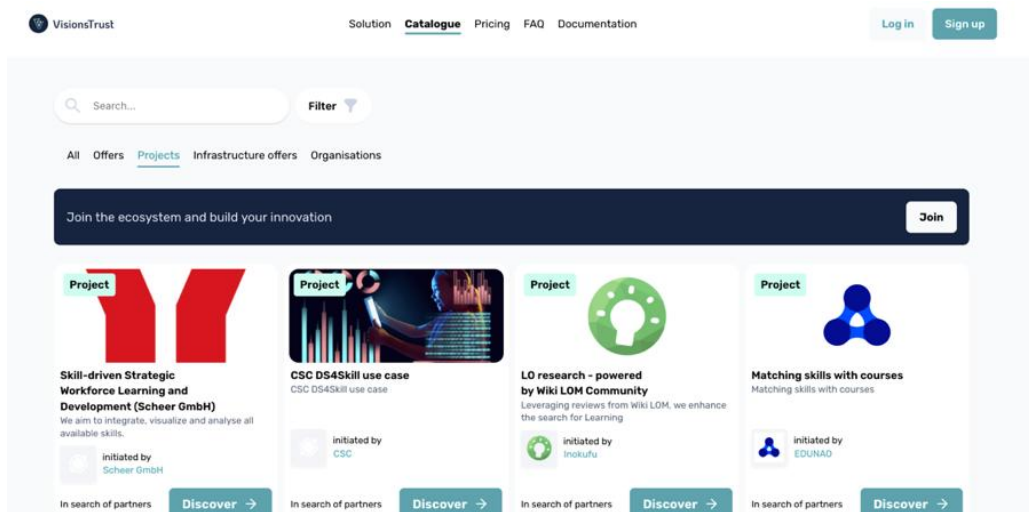


Figure 5. Aperçu de projets dans le catalogue de Visions Trust. Les deux premiers projets sont liés à DS4Skills-GO (« Mind the Gap: Skill-driven Strategic Workforce Learning and Development » et « CSC DS4Skill use case », emmenés respectivement par Scheer GmbH et CSC), quand les deux suivants (« LO research » et « Matching skills with course »), publiés par Inokufu et Edunao, émanent d'Edge-Skills.

Ces dynamiques révèlent des asymétries entre entreprises, mais aussi entre États. La France et l'Allemagne, à l'origine de GAIA-X, occupent une place dominante. La France, en particulier, concentre une grande partie des acteurs impliqués : elle fournit la majorité des « orchestrators » (11 sur 17) et des participants dans *Edge-Skills* (18 sur 28 ; EU Funding et Tenders Portal, 2026a). Dans *DS4Skills-GO*, bien que la diversité nationale soit plus large (EU Funding et Tenders Portal, 2026b), des acteurs clés issus de ces deux pays conservent des positions centrales : Prometheus-X et VisionsTrust, respectivement en tant que DSGA et place de marché pour les cas d'usage, et l'Universität Koblenz, en tant que référence académique. Ce leadership traduit une stratégie visant à structurer un marché européen des données tout en projetant des standards à l'international. La souveraineté promue n'exclut pas les acteurs étrangers, mais repose sur l'adhésion à des « valeurs européennes » codifiées dans les standards techniques. L'objectif est ainsi de concurrencer les Big Tech sans rompre avec les logiques de libre-échange.

Le DASES apparaît donc comme une « idéologie infrastructurelle » en formation, cristallisant des rapports de pouvoir à plusieurs niveaux. Toutefois, le projet reste à un stade précoce. Les cas d'usage n'ont pas encore atteint une maturité suffisante pour un déploiement massif, et le système n'a pas encore connu de « verrouillage technologique » tel que théorisé par P. Flichy (2001, p.16). Plusieurs limites interrogent en outre sa capacité future à s'imposer. Les financements restent modestes (27 millions d'euros cumulés), ce qui pourrait freiner le passage à l'échelle sans investissements privés. Surtout, le projet dépend de la question plus large du cloud européen, marquée par des tensions industrielles et diplomatiques. Ces contraintes fragilisent l'ambition de construire une infrastructure souveraine capable de rivaliser avec les plateformes dominantes.

Des obstacles industriels et diplomatiques au « cloud souverain »

L'Union européenne cherche à fonder son marché des données sur des fournisseurs de *cloud computing* respectant ses valeurs, mais le contexte industriel et diplomatique reste à ce jour incertain et favorable aux Big Tech. Deux initiatives illustrent ces tensions : GAIA-X et le projet *EU Cloud Services*.

L'association GAIA-X vise à structurer un cadre européen du cloud, mais la présence de Big Tech comme Microsoft, Amazon ou Google parmi ses plus de 300 membres suscite des critiques. Ces acteurs participent aux groupes de travail et influencent la définition des standards. Bien qu'exclus formellement de la gouvernance, ils y sont indirectement présents via des organisations comme Digital Europe, la DSGA de *Data4Skills-GO*, dont ils sont

membres fondateurs (*Digital Europe*, 2026). Certains acteurs du cloud européen dénoncent cette influence, à l'instar de Scaleway qui a quitté le projet en 2021 (Vitard, 2021), tandis que d'autres défendent une approche pragmatique. Les membres dirigeants de French Gaia-X Hub ont publié en 2021 une tribune dans *Le Monde* défendant GAIA-X en expliquant : « *Nous refusons les polémiques stériles entretenues par ceux qui se complaisent dans les chimères d'une autarcie numérique de l'Europe, ou dans l'illusion d'un souverainisme européen exclusif et absolu, dont la plupart de nos partenaires européens, d'ailleurs, ne veulent pas* » (Beylat, 2021). Ils critiquent le jusqu'au-boutisme de certains tenants de l'utopie de la souveraineté technologique, considérant plutôt la mainmise des Big Tech comme un acquis nécessitant d'inclure celles-ci dans le processus. En 2024, GAIA-X est parvenu à adopter un système de labellisation à quatre niveaux (GAIA-X, 2024).

En parallèle, le projet de certification EU Cloud Services, prévu depuis 2019, reste inabouti faute d'accord entre États membres. La France défend une approche stricte (« cloud de confiance »), reposant sur sa qualification SecNumCloud combinant exigences techniques de cybersécurité et protections contre les législations extraterritoriales, tandis que d'autres pays sont plus accommodants étant donné les besoins techniques de leurs industries. L'Allemagne, par exemple, accueille une filiale d'AWS proposant une offre cloud dite « souveraine », « *European Sovereign Cloud* », qui promet une séparation physique vis-à-vis de ses autres régions cloud avec une impossibilité d'accéder aux données hébergées (Vitard, 2023). La dernière révision du *Cybersecurity Act* en 2026 relance l'idée d'une certification, mais sans inclure de protections explicites contre les législations extraterritoriales.

La politique européenne poursuit ainsi sa stratégie ambivalente de « souveraineté numérique » (ten Oever *et al.*, 2024) : elle encadre les usages et contenus, mais laisse largement intacte l'infrastructure, afin de préserver l'ouverture de son marché numérique¹. Les Big Tech consolident quant à eux leur position en participant aux initiatives européennes et en développant des offres dites « souveraines » en partenariat avec des acteurs locaux - en témoignent les offres opérées par des acteurs européens utilisant des technologies de Big Tech comme S3NS (Thales-Google Cloud) ou Bleu (Orange-Capgemini-Microsoft) - ou de manière autonome comme Amazon. Toutefois ces services maintiennent non seulement le risque de transfert de données hors d'Europe, mais exposent également les entreprises du continent à des contraintes géopolitiques, notamment en cas de sanctions de Washington².

Ainsi les promesses de souveraineté associées aux *data spaces* reposent-elles sur des bases fragiles. Le DASES tend par ailleurs à reproduire certaines logiques de *plateformisation* caractéristiques des Big Tech, en les intégrant partiellement à son développement. Ces dynamiques éclairent les limites du projet : plutôt qu'une rupture, il pourrait contribuer à reconfigurer, sans les abolir, les rapports de domination existants dans le numérique éducatif européen.

LE DASES, TÉMOIN DE LA PUISSANCE ÉCONOMIQUE ET IDÉOLOGIQUE DES PLATEFORMES

Un développement de cas d'usage sur le mode technique et économique des plateformes

Sur le plan technique, le DASES reprend les caractéristiques des plateformes décrites par Baldwin et Woodward (2008) : des composants centraux stables (les *building blocks*), des modules variables (les services d'IA développés par les *orchestrators*) et une interface d'intermédiation (VisionsTrust). Cette dernière organise la mise en relation des acteurs et capte une part de la valeur, à la manière d'un gestionnaire de plateforme (Bullich et Guignard, 2010).

Pour la diffusion des services, la feuille de route de *DS4Skills* envisage explicitement des partenariats avec des fournisseurs de cloud et des marketplaces afin d'accélérer l'adoption :

« Pour faciliter l'intégration des blocs de construction et des intermédiaires de données de confiance avec les solutions informatiques des parties prenantes, des partenariats peuvent être établis avec les fournisseurs de services cloud et les places de marché cloud pour une intégration et une adoption plus rapides » (*Data Space for Skills*, 2023e). Sans les nommer, cela renvoie aux Big Tech et à leurs services, dont les effets de réseau sont recherchés. Ainsi, malgré l'objectif affiché de souveraineté, le DASES s'inscrit dans une logique pragmatique qui intègre ces acteurs dominants.

Cette orientation reflète une tendance plus large à la *plateformisation des infrastructures* (Plantin *et al.*, 2018), caractéristique de l'évolution du web d'un modèle ouvert et interopérable vers un système dominé par des plateformes marchandes. Le DASES, loin de rompre avec cette trajectoire, en adopte les logiques techniques et économiques.

La convergence des Big Tech et du DASES vers la plateformisation de l'éducation

Au niveau des services éducatifs, le DASES et les offres des Big Tech convergent vers le modèle de *plateformisation de l'éducation*. Comme l'a montré Vincent Bullich (2018) à propos des plateformes de MOOC, cet idéal-type repose sur une production distribuée, combinant intermédiation numérique et mise en activité des usagers, afin de maximiser la collecte de données. Deux marqueurs de l'industrialisation de l'éducation ont été caractérisés : l'idéologisation à travers des promesses de personnalisation, d'empowerment, d'inclusion, et la rationalisation via l'optimisation des pratiques pédagogiques et l'encadrement des parties prenantes au moyen de services d'accompagnement pédagogique, technique et commercial et par la contractualisation de l'activité des enseignants (*ibid.*).

Les suites éducatives des Big Tech illustrent cette logique. Elles proposent des suites éducatives comme *Google Workspace for Education* ou *Microsoft 365 Education*. Celles-ci incluent des fonctionnalités de *Learning Management Systems* (gestion de cours et de suivi des élèves via *Google Classroom* ou *Microsoft Teams*), avec l'assistance de leurs IA génératives, dans l'optique de personnaliser l'apprentissage et d'améliorer l'efficacité de l'enseignement (Google for Education, 2026 ; Microsoft Éducation, 2026). S'y ajoutent des places de marché proposant d'intégrer des applications EdTech tierces (*Google for Education App Hub*, *Microsoft Teams Store*). Ainsi ces acteurs structurent un marché multi-versant dans lequel ils se rémunèrent via des abonnements et des commissions. Ils participent à une industrialisation de l'éducation fondée sur la donnée.

Les cas d'usage du DASES, s'ils ne sont pas encore déployés à grande échelle, reprennent également les caractéristiques de la plateformisation. Sur le plan idéologique, le projet WALLY « *Learning traces from websites and browsing activities* » d'*Edge-Skills* véhicule la promesse d'une personnalisation de l'apprentissage par l'IA. Selon la page consacrée sur le site de Prometheus-X, « *en analysant l'historique complet de l'apprentissage, à la fois dans les environnements LMS et à travers la navigation sur le web, WALLY offre des expériences éducatives personnalisées qui s'adaptent aux courbes d'apprentissage et aux préférences de chacun* » (Prometheus-X, 2026c). L'autre promesse est celle d'un « *empowerment* par le consentement » : « *respectant la vie privée, WALLY permet aux apprenants de gérer leur consentement au suivi, en offrant des paramètres personnalisables qui définissent quand, comment et quels aspects de leurs activités sur le web sont surveillés* » (*ibid.*). La promesse d'inclusion dans une communauté est à trouver dans la perspective donnée aux individus d'intégrer plus facilement les marchés de la formation et du travail européens, comme l'illustre l'un des objectifs de *DS4Skills-GO*, qui est de permettre aux étudiants « *un accès facilité aux opportunités d'apprentissage des universités de l'UE* » (DS4Skills, 2026).

En termes de technologisation, une tendance dominante est l'assimilation de l'éducation à un marché optimisé par la technologie. Par exemple, le projet « *Enabling students to align skills with opportunities* » d'*Edge-Skills* vise à développer une plateforme proposant, à partir des écarts entre les compétences actuelles des étudiants et celles requises par le marché, des recommandations personnalisées en matière de formation, de stages ou d'emplois, tout en aidant les entreprises à adapter leurs programmes de recrutement et de formation

(Prometheus-X, 2026d). En creux, l'expertise de professions traditionnelles comme celle de conseiller d'orientation est invisibilisée. Une autre tendance est l'idée de pratiques éducatives multiformes et tout au long de la vie. Dans *Edge-Skills*, le projet « *Technical approach to API design for VR application and customization* » vise par exemple à faciliter l'accessibilité et l'interopérabilité des données de réalité virtuelle en développant une API utilisable sur différentes plateformes éducatives (Prometheus-X, 2026e). Une troisième forme de technologisation se focalise sur la codification du rôle de l'enseignant grâce aux visualisations et recommandations algorithmiques. Dans *DS4Skills-GO*, le cas d'usage « *Skills-driven Higher Education Institutions* » doit fournir à la communauté éducative « *des informations sur les expériences et les compétences des étudiants, évaluer les modèles d'apprentissage innovants et [la] soutenir dans l'évaluation de l'efficacité et de la satisfaction de la formation* » (DS4Skills, 2026). On retrouve là les mêmes accents que dans les offres éducatives des Big Tech.

Enfin, la rationalisation se manifeste par une recherche d'efficience fondée sur la mise en données des compétences et le suivi des parcours. L'éducation est pensée comme un système à optimiser, où les interactions entre individus et institutions sont médiées par des dispositifs numériques, dans une logique proche du « nouvel esprit du capitalisme » théorisé par Boltanski et Chiapello. Ainsi, malgré ses ambitions de souveraineté, le DASES reproduit largement les logiques de plateformes portées par les Big Tech. Bien que ces cas d'usage ne soient pas encore commercialisés, ce qui empêche de retrouver d'autres caractéristiques de la plateforme éducative telles que les productions d'accompagnement pédagogique, technique et commercial auprès des formateurs (Bullich, 2018), son projet éducatif apparaît en continuité avec leur modèle industriel, prolongeant une conception de l'éducation fondée sur la donnée, l'intermédiation et l'optimisation des parcours.

CONCLUSION

Concernant la problématique initiale, la première question concernait les rapports de force socio-économiques soulevés par le DASES. Ses textes fondateurs ont révélé une utopie connexionniste « centrée sur l'humain », consistant en un projet social-libéral permis par des données et des IA « souveraines », devant se concrétiser dans une infrastructure décentralisée empruntant à l'utopie séminale d'Internet. La phase d'expérimentation a mis au jour des rapports de force asymétriques, favorisant l'emprise technique d'une autorité, la DSGA, les visées marchandes des firmes porteuses de projet et d'intermédiaires de données, ainsi que les industries *EdTech* de certains États membres comme la France et l'Allemagne. Sur le plan diplomatique, une « idéologie infrastructurelle » a émergé, tentant de formaliser un agenda éthique face à l'*infrastructuralisation* des plateformes de Big Tech. Néanmoins la faiblesse des investissements publics et la régulation erratique des fournisseurs de cloud fragilisent le déploiement de cette infrastructure et la poursuite de ces promesses.

La seconde question, qui était de déterminer la capacité du DASES à remettre en cause la domination des Big Tech dans le secteur des logiciels éducatifs, trouve par conséquent une réponse négative. Par ailleurs, il est apparu que le DASES ne rompait pas sur le plan médiatique et productif avec ces multinationales, et véhiculait une *plateformisation des infrastructures comme de l'éducation*.

Toujours est-il que le DASES émane d'un antagonisme fort avec les Big Tech et les États-Unis sur les questions de la libre-concurrence et du respect des données personnelles. Les critiques persistantes de Big Tech comme Apple³ et de l'administration Trump envers les législations européennes indiquent que le combat politique reste vif. Pendant ce temps, les industries technologiques européennes accusent un retard compétitif à l'international. Un enjeu futur est de savoir si les totems de la libre-concurrence, du respect des données personnelles et de la souveraineté technologique, faisant la spécificité diplomatique de l'Union européenne, seront encore défendus face à ces vents industriels et

géopolitiques contraires.

NOTES

¹ Niels ten Oever et ses collègues ont déjà identifié cette ambivalence dans leur étude des sanctions visant les médias en ligne russes après l'invasion de l'Ukraine (ten Oever *et al.*, 2024) : pour contrer la désinformation russe dans l'Union, une base légale avait été trouvée pour bloquer ces médias en raison du contenu qu'ils distribuaient, via le blocage des réponses des systèmes de noms de domaine par les États membres. En revanche les fournisseurs d'accès internet russes, malgré les sanctions de l'UE, ont pu continuer à utiliser leurs numéros de réseau et adresses IP dans l'espace européen. En d'autres termes, l'UE a choisi de limiter son influence à la « couche des contenus » (« *content layer* ») présents sur Internet tout en laissant intacte l'interconnexion des infrastructures, pour éviter toute fragmentation d'Internet, et préserver ainsi ses idéaux d'ouverture et d'intégration dans la mondialisation.

² En cas de sanctions du gouvernement fédéral contre une entreprise européenne qui utiliserait les services cloud d'un géant californien, ce dernier pourrait être contraint de bloquer l'accès de cette entreprise à ses services, même si ses données et ses opérations sont basées en Europe. Cette hypothèse s'est renforcée après la suspension des services de messagerie Microsoft de Karim Khan, procureur de la Cour Criminelle Internationale basée aux Pays-Bas, en juin 2025. Celui-ci faisait alors l'objet de récriminations de l'administration fédérale pour avoir requis un mandat d'arrêt international contre le Premier Ministre israélien Benyamin Netanyahu et son Ministre de la Défense Yoav Gallant pour crimes de guerre et crime contre l'humanité commis dans la bande de Gaza (Clark, 2025).

³ Avant d'annoncer de nouvelles conditions commerciales conformes au *Digital Markets Act* (DMA) en août 2026, Apple a entretenu un bras de fer avec la Commission européenne. Condamné pour non-respect du DMA en avril 2025, le groupe californien avait riposté en septembre en demandant l'abrogation de cette réglementation : https://www.lemonde.fr/economie/article/2025/09/25/apple-appelle-l-union-europeenne-a-abroger-sa-loi-contre-les-geants-du-numerique_6642865_3234.html. Le président états-unien a multiplié les pressions sur l'Union européenne, accusant *le Digital Markets Act* et *le Digital Services Act* de cibler déloyalement les entreprises états-uniennes. Durant l'été 2025, Washington a brandi la menace de droits de douane sur les produits européens, avant de menacer, en décembre, d'imposer des restrictions à l'entrée des acteurs numériques européens sur le sol états-unien, visant notamment des entreprises comme *Mistral AI*.

RÉFÉRENCES BIBLIOGRAPHIQUES

Assemblée nationale (2022), « Question écrite n° 971 : Gratuité d'Office 365 », [en ligne], consulté le 25 avril 2026, <https://questions.assemblee-nationale.fr/q16/16-971QE.htm>

Baldwin, Carliss Y. ; Woodward, C. Jason (2008), « The Architecture of Platforms: A Unified View », Harvard Business School Finance, Working Paper No. 09-034, [en ligne], consulté le 25 avril 2026, <https://ssrn.com/abstract=1265155>. DOI : <http://dx.doi.org/10.2139/ssrn.1265155>

Beylat, Jean-Luc (2021), « Nous ne pouvons tolérer qu'une poignée d'acteurs préempte les données européennes au bénéfice exclusif de leurs modèles d'affaire », *Le Monde*, édition du 6 mai 2021, [en ligne], consulté le 25 avril 2026, https://www.lemonde.fr/idees/article/2021/05/06/nous-ne-pouvons-tolerer-qu-une-poignee-d-acteurs-preempte-les-donnees-europeennes-au-benefice-exclusif-de-leurs-modeles-d-affaire_6079350_3232.html

Boët, Hubert (2025), *Analyse contrastive de l'industrialisation des médias d'apprentissage adaptatif en France et aux États-Unis*, thèse de doctorat en Sciences de l'information et de

la communication, Université Sorbonne Paris Nord, [en ligne], consulté le 25 avril 2026, <https://theses.fr/s295672>.

Boltanski, Luc ; Chiapello, Ève. (1999), *Le nouvel esprit du capitalisme*, Paris : Gallimard.

Bullich, Vincent (2015), « Les plateformes : littérature scientifique, éléments de définition et enjeux pour les transformations des industries culturelles », *Distribution et diffusion*, journée d'études organisée par le Laboratoire d'Excellence ICCA (Industries Culturelles et Création Artistique), Maison de la Recherche de La Sorbonne Nouvelle-Paris 3, Paris, le 22 octobre 2015.

Bullich, Vincent (2018), « La « platformisation » de la formation », *Distances et médiations des savoirs*, 21 | 2018, [en ligne], consulté le 17 mai 2024, <http://journals.openedition.org/dms/2096> ; DOI : <https://doi.org/10.4000/dms.2096>

Bullich, Vincent (2021), « Plateforme, platformiser, platformisation : le péril des mots qui occultent ce qu'ils nomment », *Questions de communication*, 40(2), 47-70 [en ligne], consulté le 25 avril 2026, <https://doi.org/10.4000/questionsdecommunication.27413>.

Clark, Sam (2025), « Microsoft didn't cut services to International Criminal Court, its president says », *Politico*, édition du 4 juin 2025, [en ligne], consulté le 25 avril 2026, <https://www.politico.eu/article/microsoft-did-not-cut-services-international-criminal-court-president-american-sanctions-trump-tech-icc-amazon-google/>

Commission européenne (2020), « Une stratégie européenne pour les données. Communication de la Commission au Parlement Européen, au Conseil, au Comité Économique et Social Européen et au Comité des Régions », Bruxelles, le 19 février 2020, [en ligne], consulté le 25 avril 2026, <https://eur-lex.europa.eu/legal-content/EN/TXT/?qid=1593073685620&uri=CELEX:52020DC0066>

Commission européenne (2021), « Call: DIGITAL-2021-PREPACTS-DS-01- Preparatory actions for Data Spaces », Bruxelles, le 17 novembre 2021, [en ligne], consulté le 25 avril 2026, https://ec.europa.eu/info/funding-tenders/opportunities/docs/2021-2027/digital/wp-call/2021/call-fiche_digital-2021-prepacts-ds-01_en.pdf

Data Spaces Support Center (2023), « 2. Core concepts », [en ligne], consulté le 25 avril 2026, <https://dssc.eu/space/Glossary/176554052/2.+Core+Concepts>

Data Space for Skills (2023a), « 3.1. Principles », [en ligne], consulté le 25 avril 2026, <https://www.skillsdataspace.eu/blueprint/principles/>

Data Space for Skills (2023b), « 2.1. Blueprint Introduction: Why a data space? », [en ligne], consulté le 25 avril 2026, <https://www.skillsdataspace.eu/blueprint/test-7/>

Data Space for Skills (2023c), « 3.2. Approach », [en ligne], consulté le 25 avril 2026, <https://www.skillsdataspace.eu/blueprint/approach/>

Data Space for Skills (2023d), « 7.3. Human centricity in the technical data space architecture: the Personal Data Intermediary », [en ligne], consulté le 25 avril 2026, <https://www.skillsdataspace.eu/blueprint/human-centricity-in-the-technical-data-space-architecture-the-personal-data-intermediary/>

Data Space for Skills (2023e), « 9.5. Deployment strategy », [en ligne], consulté le 25 avril 2026, <https://www.skillsdataspace.eu/blueprint/deployment-strategy/>

Data Space for Skills (2026), « Deploying the European Skills Data Space Through Eight Sustainable Use Case », [en ligne], consulté le 25 avril 2026, <https://ds4skills.eu/use-cases/>

Digital Europe (2026), « Our corporate members », [en ligne], consulté le 25 avril 2026, <https://www.digitaleurope.org/corporate/>

EU Funding & Tenders Portal (2026a), « European Dataspace for Growth and Education - Skills (EDGE-Skills) », [en ligne], consulté le 25 avril 2026,

<https://ec.europa.eu/info/funding-tenders/opportunities/portal/screen/how-to-participate/org-details/952919756/project/101123471/program/43152860/details>

EU Funding & Tenders Portal (2026b), « Deployment of a European data space for skills (DS4SKILLS-GO) », [en ligne], consulté le 25 avril 2026, <https://ec.europa.eu/info/funding-tenders/opportunities/portal/screen/how-to-participate/org-details/999999999/project/101195165/program/43152860/details>

GAIA-X (2024), « Gaia-X Compliance Document - 24.06 Release », [en ligne], consulté le 25 avril 2026, https://gaia-x.eu/wp-content/uploads/2024/09/Compliance-Document_24.06.pdf

Madiega, Tambiama (2020), « Digital sovereignty for Europe », European Parliamentary Research Service, *EPRS Ideas Paper - Towards a more resilient EU*, [en ligne], consulté le 25 avril 2026, [https://www.europarl.europa.eu/RegData/etudes/BRIE/2020/651992/EPRS_BRI\(2020\)651992_EN.pdf](https://www.europarl.europa.eu/RegData/etudes/BRIE/2020/651992/EPRS_BRI(2020)651992_EN.pdf)

Flichy, Patrice (2001), *L'imaginaire d'Internet*, Paris : La Découverte.

Flichy, Patrice (2003), *L'innovation technique. Récents développements en sciences sociales. Vers une nouvelle théorie de l'innovation*, Paris : La Découverte.

Google for Education (2026), « GOOGLE WORKSPACE FOR EDUCATION. Renforcez l'impact des enseignants avec des outils qui transforment l'enseignement », [en ligne], consulté le 25 avril 2026, https://edu.google.com/intl/ALL_fr/workspace-for-education/editions/overview/

Joly, Pierre-Benoît (2010), « On the economics of technoscientific promises » (p. 203-222), in M. Akrich, Y. Barthe, F. Muniesa et P. Mustar (dir.), *Débordements - Mélanges offerts à Michel Callon*, Paris : Presses des Mines.

Lyubareva, Inna ; Nocetti, Julien (2024), « La diplomatie numérique. Évolution des stratégies diplomatiques et d'influence à l'ère (du) numérique », *Réseaux*, Vol. 245, n°3, p.11-35, [en ligne], consulté le 25 avril 2026, <https://shs.cairn.info/revue-reseaux-2024-3-page-11?lang=fr> DOI : <https://doi.org/10.3917/res.245.0011>

Mattelart, Armand (2001), *Histoire de la société de l'information*, Paris : La Découverte.

Maxigas ; ten Oever, Niels (2023), « Geopolitics in the infrastructural ideology of 5G ». *Global Media and China*, vol. 8, n° 3, p. 271-288. [en ligne], consulté le 25 avril 2026, <https://journals.sagepub.com/doi/10.1177/20594364231193950>. DOI : <https://doi.org/10.1177/20594364231193950>

Microsoft Éducation (2026), « Microsoft 365 Éducation. Préparez les étudiants à réussir demain grâce à Microsoft 365 Éducation » [en ligne], consulté le 25 avril 2026, <https://www.microsoft.com/fr-fr/education/products/microsoft-365#Education-plans>

Mœglin, Pierre (2016), (dir.), *Industrialiser l'éducation : anthologie commentée (1913-2012)*, Saint-Denis : Presses universitaires de Vincennes.

Plantin, Jean-Christophe ; Lagoze, Carl ; Edwards, Paul N. ; Sandvig, Christian (2016), « Infrastructure studies meet platform studies in the age of Google and Facebook », *New Media & Society*, [en ligne], consulté le 25 avril 2026, https://researchonline.lse.ac.uk/id/eprint/67571/1/plantin_Infrastructure_studies.pdf. DOI : <https://doi.org/10.1177/1461444816661553>

Prometheus-X (2024), « Edge-Skills. Creating the future of data spaces in Europe », [en ligne], consulté le 25 avril 2026, https://prometheus-x.org/wp-content/uploads/2025/02/PT-X_Whitepaper_250216.pdf

Prometheus-X (2026a), « Learning and teaching dashboards for adaptive learning », [en

ligne], consulté le 25 avril 2026, <https://prometheus-x.org/learning-and-teaching-dashboards-for-adaptive-learning/>

Prometheus-X (2026b), « Recommender systems for teachers », [en ligne], consulté le 25 avril 2026, <https://prometheus-x.org/uc-recommender-systems-for-teachers/>

Prometheus-X (2026c), « Learning traces from websites and browsing activities (WALLY) », [en ligne], consulté le 25 avril 2026, <https://prometheus-x.org/learning-traces-from-websites-and-browsing-activities-wally/>

Prometheus-X (2026d), « Enabling students to align skills with opportunities: The Schülerkarriere initiative », [en ligne], consulté le 25 avril 2026, <https://prometheus-x.org/uc-skill-gaps-analytics-for-students/>

Prometheus-X (2026e), « Technical approach to API design for VR application and customization », [en ligne], consulté le 25 avril 2026, <https://prometheus-x.org/technical-approach-to-api-design-for-vr-application-and-customization/>

Synergy Research Group (2025), « Cloud Market Share Trends - Big Three Together Hold 63% while Oracle and the Neoclouds Inch Higher », [en ligne], consulté le 25 avril 2026, <https://www.srgresearch.com/articles/cloud-market-share-trends-big-three-together-hold-63-while-oracle-and-the-neoclouds-inch-higher>

ten Oever, Niels ; Perarnaud, Clément ; Kristoff, John ; Müller, Moritz ; Resing, Max ; Filasto, Arturo ; Kanich, Chris (2024), « Sanctions and infrastructural ideologies: Assessing the material shaping of EU digital sovereignty in response to the war in Ukraine ». *Policy & Internet*, vol. 16, n° 4, p. 692-710, [en ligne], consulté le 25 avril 2026, <https://onlinelibrary.wiley.com/doi/full/10.1002/poi3.422>. DOI : <https://doi.org/10.1002/poi3.422>

VisionsTrust (2026), « Catalogue », [en ligne], consulté le 25 avril 2026, <https://visionstrust.com/catalog/projects>

Vitard, Alice (2021), « Scaleway claque la porte du projet de cloud européen Gaia-X ». *L'Usine Digitale*, édition du 19 novembre 2021, [en ligne], consulté le 25 avril 2026, <https://www.usine-digitale.fr/article/scaleway-claque-la-porte-du-projet-de-cloud-europeen-gaia-x.N1161587>

Vitard, Alice (2023), « Lancement d'un cloud "souverain" à la sauce Amazon ». *L'Usine Digitale*, édition du 25 octobre 2023, [en ligne], consulté le 25 avril 2026, <https://www.usine-digitale.fr/editorial/lancement-d-un-cloud-souverain-a-la-sauce-amazon.N2186588>

REVUE SCIENTIFIQUE EN SCIENCES DE L'INFORMATION ET DE LA COMMUNICATION

LES ENJEUX de l'information et de la communication

Revue scientifique en sciences de l'information et de la communication

Éditée par le Gresec (Groupe de recherche sur les enjeux de la communication), cette revue scientifique aborde les processus d'information-communication dans leurs développements, mutations et inscriptions sociales, politiques et économiques. Elle privilégie les travaux relevant d'approches théoriques critiques et fondés sur des études empiriques, rendant compte de recherches conduites par des auteurs confirmés, des doctorants ou de jeunes chercheurs.

Elle est constituée de 3 à 4 numéros annuels :

n°1 - Varia

n°2 - Dossier thématique

n°3 - Supplément A

n°4 - Supplément B

C'est une revue en ligne, d'accès libre, qualifiante en Sciences de l'Information et de la Communication, consultable sur :

<https://lesenjeux.univ-grenoble-alpes.fr> et sur le portail www.cairn.info.