

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.