

Retour sur la Biennale du numérique de l'Enssib « La gouvernance des données : un enjeu incontournable? »

17 et 18 novembre 2025, à Villeurbanne

Judith Carbonnel

Étudiante en master 2 Sciences de l'information et des bibliothèques (SIBIST) à l'Enssib

Eva Lefebure

Étudiante en master 2 Politique des bibliothèques et de la documentation (PBD) à l'Enssib

Mathias Lenarduzzi

Étudiant en master 2 Archives numériques (ARN) à l'Enssib

Lydia Marchand

Étudiante en master 2 Archives numériques (ARN) à l'Enssib

Cindy Marechal

Étudiante en master 2 Sciences de l'information et des bibliothèques (SIBIST) à l'Enssib

Nous remercions Raphaëlle Wintrebert (master 2 SIBIST) pour son aide à la prise de notes.

*

Depuis trois ans déjà, l'Enssib [École nationale supérieure des sciences de l'information et des bibliothèques] développe une réflexion approfondie sur la question de la gouvernance des données¹. Le défi actuel ne réside pas tant dans l'apparition de nouveaux types de données que dans un changement d'échelle sans aucun précédent. En effet, la production mondiale de données a désormais atteint 180 000 milliards de gigaoctets. On parle ainsi d'un « déluge de données ». Face à cette masse considérable, la rationalisation et la gestion constituent un enjeu majeur, posant des questions sociétales, éthiques, économiques, environnementales, mais aussi et surtout professionnelles, que l'édition 2025 de la Biennale du numérique de l'Enssib a voulu explorer.

Ce compte rendu retrace les deux journées d'échanges stimulants entre professionnels et chercheurs en sciences de l'information, avec une première journée proposant une approche plutôt générale de la question, une sorte de « mise à niveau » des connaissances en la matière, et une deuxième journée centrée sur les problématiques plus spécifiquement professionnelles.

Première journée

La première matinée a été consacrée à une mise en perspective historique et contextuelle de la gouvernance des données à travers deux conférences plénières, ainsi qu'à une première entrée en matière autour d'une table ronde consacrée à la question de la « donnée » comme « bien commun ». Ainsi, le professeur Stéphane Chaudiron (Université de Lille) a analysé les conséquences de la transformation numérique et de la centralisation des données dans l'économie mondiale (*Comment en est-on arrivé là ?*²). Bien que la surcharge informationnelle soit théorisée

1 Comme aboutissement de ces réflexions, l'Enssib ouvre, en 2025, une formation permettant de délivrer un diplôme d'établissement en Gouvernance responsable des données (<https://www.enssib.fr/l-offre-de-formation/diplome-gouvernance-responsable-des-donnees-presentation>).

2 https://www.youtube.com/watch?v=wmzZwsjpzGk&list=PLEjcwQr_8kwiQ-OBiED1pSiO5OCsgWLj&index=2

depuis les années 1970, les défis actuels en matière de souveraineté, de démocratie et d'environnement imposent un changement de paradigme. L'équilibre entre ouverture, protection et durabilité des données reste précaire. Ugo Verdi, maître de conférences en sciences de l'information et de la communication (SIC) à Montpellier, a tenté, quant à lui, une définition de la « donnée » et de la « gouvernance » en montrant à la fois toute la polysémie des termes et cependant les invariants sur lesquels reposent ces notions (« *Gouvernance des données* » : *parle-t-on toujours de la même chose* ?³). Pour cela, Ugo Verdi a considéré trois aspects constitutifs de la donnée : la donnée comme concept, la donnée comme matériau et la donnée comme système.

La première table ronde (*Les données comme bien commun : promesse ou mirage* ?⁴), animée par Henri Stiller (président de l'ADBS, Association des professionnels de l'information et de la documentation), avec Valérie Peugeot (professeure affiliée à Sciences Po Paris), Jean Cattan (responsable de la démarche Café IA) et Roman Arciniega Gil (docteur en droit public), a interrogé la notion de donnée comme « bien commun ». Les intervenants ont distingué l'open data des données mises en commun. Si l'open data favorise l'innovation, elle bénéficie principalement dans les faits aux GAFAM⁵. Afin d'éviter que les communs ne deviennent une nouvelle ressource accaparée par des acteurs dominants, des solutions pour équilibrer les rapports de force ont été discutées, comme les licences restrictives, les contre-dons obligatoires et les gouvernances renforcées via des accès régulés.

Cette matinée a mis en lumière une transition fondamentale : le passage d'une ère d'administration des données vers une gouvernance rationalisée d'une masse toujours grandissante. Une gestion réfléchie des données constitue désormais un impératif éthique, économique et politique, puisqu'elles deviennent une ressource centrale pour la prise de décision organisationnelle. Cette évolution se traduit notamment par l'adoption croissante de stratégies data-driven.

En ouverture de l'après-midi de cette première journée, la table ronde n° 2, animée par Anne-Sophie Hulin (professeur adjointe à l'Université de Sherbrooke), avec Antoine Petel (juriste), Nawale Lamrini (délégué à la protection des données, Sciences Po Paris) et Anne-Laure Donzel (consultante en open data chez Dataactivist), posait la question fondamentale suivante : *Faut-il sécuriser ou libérer la gouvernance, la conservation et la réutilisation des*

données ?⁶ Cette question entraine en tension entre deux impératifs : la sécurisation juridique des données et leur libération pour permettre la réutilisation au-delà de leur finalité de base. Il a été souligné que ces deux dimensions sont indissociables. En effet, dans un paradoxe apparent, une entreprise ou une administration ne partagera ses données que si leur protection est garantie. Un bilan critique des réglementations européennes récentes a été dressé, celles-ci apparaissant légitimer une fermeture protectionniste au détriment d'une transparence dans la démarche. Face à ces tensions inhérentes au sujet, l'approche de la science ouverte a été défendue : « aussi ouvert que possible, aussi fermé que nécessaire ».

Dans le prolongement de la précédente, la table ronde n° 3, *Données et intelligence artificielle : ouvrir et protéger* !⁷, animée par Thomas Parisot (directeur général adjoint de Cairn.info, président du Gf2i, Groupement français de l'industrie de l'information), avec Philippe Masseron (directeur général du Gf2i), Bruno Marmol (responsable de production et de l'infrastructure HAL) et Christophe Drot (directeur général de DPO Consulting), s'est interrogée sur la manière dont les acteurs pouvaient réagir à la double injonction qui s'impose aux données depuis la démocratisation de l'intelligence artificielle (IA) : protéger ses données, comme l'ont fait les grands médias, tout en préparant des corpus adaptés à son entraînement, pour ne pas condamner le Web ouvert. Ici aussi l'injonction « *aussi ouvert que possible, aussi fermé que nécessaire* » s'applique. Cette tension illustre le paradoxe d'une économie de partage, où les données doivent être partagées, sans annuler les enjeux autour de la protection des données personnelles. Quelles réponses sont possibles face aux violations du droit d'auteur et à la multiplication des *bad bots* qui ne respectent pas les règles des sites collectés ? Deux approches se dessinent : l'action individuelle, comme celle de l'archive HAL qui bloque certains bots, ou l'action collective envisagée par les acteurs culturels. Toutefois, il a été souligné que ce débat arrive tardivement, puisque les principaux modèles d'IA générative sont déjà entraînés sur des données préexistantes, souvent en dépit du droit d'auteur et de la vie privée...

La dernière table ronde de cette première journée était intitulée : *Pourquoi réguler la relation entre données « publiques » et données « privées »* ?⁸ Animée par Pascal Robert (professeur des universités, Enssib), elle a réuni Paul Moussier (juriste), Guy Mélançon (professeur des universités) et Tiphaine Vacqué (direction des services et réseaux à la BnF,

3 https://www.youtube.com/watch?v=kk00veZGkrQ&list=PLEjcwQr_8kwIQ-OBiED1pSiO5OCsGWLj&index=3

4 https://www.youtube.com/watch?v=Qtl-X6nQkq8&list=PLEjcwQr_8kwIQ-OBiED1pSiO5OCsGWLj&index=4

5 GAFAM est l'acronyme des géants du Web – Google, Apple, Facebook, Amazon et Microsoft – qui sont les cinq grandes firmes américaines qui dominent le marché du numérique.

6 https://www.youtube.com/watch?v=su_p5W6xojo&list=PLEjcwQr_8kwIQ-OBiED1pSiO5OCsGWLj&index=5

7 https://www.youtube.com/watch?v=zMYrDGa8G5A&list=PLEjcwQr_8kwIQ-OBiED1pSiO5OCsGWLj&index=6

8 https://www.youtube.com/watch?v=sr1G00taqxY&list=PLEjcwQr_8kwIQ-OBiED1pSiO5OCsGWLj&index=7

Bibliothèque nationale de France). Les échanges ont porté sur la convergence entre les données publiques et privées. Cette convergence présente à la fois des avantages, comme l'apport de ressources financières pour les universités et le développement de l'open science, et des obstacles, notamment la complexité juridique liée à l'articulation des droits nationaux et européens, ainsi que les coûts importants d'accès aux données dans le public. Un constat commun a révélé l'émergence d'un marché structuré autour de quatre acteurs, à savoir les ayants droit, les fournisseurs d'IA, les fournisseurs de données et les sociétés de gestion collective, dans l'objectif d'harmoniser la concurrence.

Cette journée d'émulation intellectuelle a été clôturée par l'intervention de Samuel Goëta, sociologue et cofondateur de Dataactivist⁹, qui a posé la question : « *Est-ce que l'open data a rempli ses promesses ?* »¹⁰ Il a rappelé la promesse initiale de transparence du mouvement open data, censée permettre une meilleure connaissance et un contrôle citoyen de l'action publique, aussi bien au niveau de l'État qu'au niveau plus immédiat des collectivités territoriales, voire des établissements (comme les bibliothèques). Autant de possibles aujourd'hui encore très largement limités par le manque de transparence des données publiques en France. L'autre promesse, d'aide à l'innovation, s'est matérialisée autour de « hackathons publics ». L'analyse de Samuel Goëta décèle la trace d'un solutionnisme démocratique empreint d'une certaine naïveté, et montre également toutes les difficultés qui restent pour ouvrir les données, et il se demande si le problème, simplement, ne résiderait pas plutôt dans l'attente de résultats immédiats, au lieu de s'inscrire dans une vision sur le long terme laissant même la place au « débordement ».

Deuxième journée

Nicolas Morin (directeur de l'Abes, Agence bibliographique de l'enseignement supérieur) a initié la deuxième journée lors d'une conférence introductive sur la donnée comme infrastructure nationale (*La donnée comme infrastructure nationale*¹¹). Il s'est agi d'interroger le rôle de la puissance publique face aux innovations technologiques majeures, en soulignant les obstacles liés à la transformation numérique de l'État grevé d'une dette organisationnelle et d'une dette technique. L'analyse a rappelé que cette transformation exige du temps, car elle nécessite un véritable travail de fond sur les infrastructures pour garantir des services fiables et durables. La qualité

des données et l'interopérabilité entre systèmes centraux et outils locaux sont nécessaires pour instaurer la confiance. Le constat porte sur les systèmes informatiques coûteux remplis des données de basse qualité qui contraignent les chercheurs à devoir les nettoyer à la main. Les solutions proposent la mise en place d'un service de *master data management* (gestion des données de référence) assurant l'interopérabilité et des référentiels uniques, ainsi que des services de lignage et contrôle qualité.

Véronique Mesguich (consultante formatrice) a ensuite animé une conférence centrée sur l'apport du monde des données aux professionnels des bibliothèques (*Penser les données en bibliothèques*¹²). L'accent a été mis sur une meilleure connaissance des publics, la gestion des données d'usage et des données de recherche en bibliothèque universitaire. Les métadonnées jouent un rôle décisif dans la visibilité des ressources, avec des enjeux majeurs de qualité et d'interopérabilité. Trois catégories de données en bibliothèque ont été distinguées : les données d'usage, les données bibliographiques et les données de recherche. L'intervention a également insisté sur l'importance des IA qui imposent de repenser la qualité et le renouvellement des données brutes, tout en soulevant des questions éthiques liées à la protection des données.

Les tables rondes 5 et 6 ont permis de porter la focale sur les bibliothèques territoriales face à la gouvernance des données, et ce à l'appui de l'ouvrage intitulé *Compter pour raconter. Du bon usage des données en bibliothèque*¹³ (2025), publié sous la direction de Cécile Touitou aux Presses de l'Enssib.

Plus précisément, la table ronde n° 5 intitulée *Ce que les données font aux bibliothèques, et inversement...*¹⁴, modérée par Agnieszka Tona (maîtresse de conférences, Enssib), réunissant Cécile Touitou (responsable cellule prospective à Sciences Po Paris), Rachid Aliouat (DGDNum, Université de Lille) et Damien Day (chef du service des collections documentaires, Humathèque), a abordé l'impact réciproque entre données et bibliothèques. Les échanges ont montré que les données collectées pour établir des statistiques d'usages doivent être recontextualisées, en croisant les données issues des outils avec les besoins exprimés par les usagers. Le rôle des professionnels est défini par l'interrogation du public sur ses pratiques pour comprendre l'utilisation réelle des services proposés.

Lors de la table ronde n° 6, intitulée *cette fois-ci Quelles sont les données en bibliothèque publique, comment sont-elles produites et à quoi servent-elles*

9 <https://dataactivist.coop/fr/>

10 https://www.youtube.com/watch?v=xcAi5Ce6bhs&list=PLEjcwQr_8kwIQ-OBiED1pSiO5OCsgWlj&index=8

11 https://www.youtube.com/watch?v=BwTBSqec910&list=PLEjcwQr_8kwIQ-OBiED1pSiO5OCsgWlj&index=9

12 https://www.youtube.com/watch?v=FRlixmJ6XBE&list=PLEjcwQr_8kwIQ-OBiED1pSiO5OCsgWlj&index=10

13 <https://presses.enssib.fr/catalogue/compter-pour-raconter-du-bon-usage-des-donnees-en-bibliotheque>

14 https://www.youtube.com/watch?v=m_CCDSNejwE&list=PLEjcwQr_8kwIQ-OBiED1pSiO5OCsgWlj&index=11

concrètement ?¹⁵, animée par Cécile Touitou, les participant-es, Benjamin Luzet (chargé de mission, Bibliothèque municipale de Lyon), Hélène Bert (administratrice SIGB¹⁶, Bibliothèque des Côtes-d'Armor) et David Bédon (responsable du pôle Circulation des ressources, département du Val-d'Oise) ont discuté de la production de données en bibliothèque et de leur véritable utilité. Les statistiques bibliothéconomiques informent sur les usages mais elles manquent d'interopérabilité. Les données sont donc indispensables pour répondre aux questions sur le fonctionnement et l'utilisation des bibliothèques, et il est plus que nécessaire de développer une culture de la donnée, en sensibilisant les collègues pour encourager la collecte de données au sein des équipes.

Ghislaine Charton (professeur émérite au Cnam, Conservatoire national des arts et métiers) a proposé une mise en perspective de la formation des professionnels de l'information, un métier en « mutation constante », tout en élargissant la notion (*Comment les formations des professionnels de l'information se sont-elles adaptées aux évolutions du monde ?*¹⁷). Les vingt dernières années ont vu une spécialisation accrue des compétences avec, en parallèle, le développement d'une polyvalence des missions, au point que le tronc commun même du métier est devenu difficile à définir, tant la profession est transversale. Cette transformation s'inscrit dans une logique d'approche par blocs de compétences, formalisée notamment par la loi Avenir professionnel de 2018 et illustrée par la roue de DAMA. Ce cadre offre une meilleure visibilité et une reconnaissance accrue aux métiers de la donnée, particulièrement dans un contexte marqué par les enjeux de désinformation. Comme le souligne Ghislaine Charton : « Hier, on aidait à chercher. Aujourd'hui, on doit aider à repérer la désinformation. » La donnée est ainsi devenue une responsabilité collective qui exige préservation, nettoyage, partage et documentation.

La septième table ronde (*Métiers et activités liés aux données de la recherche : qui travaille avec qui et pour faire quoi ?*¹⁸), animée par Céline Guyon (maîtresse de conférences associée, Enssib) a réuni Claire Toussaint et Odile Nguyen (respectivement responsable pédagogique et directrice adjointe de Médiat Rhône-Alpes), Isabelle Blanc (administratrice ministérielle des données, des algorithmes et des codes sources, MESRE¹⁹) et Romain Féret (directeur de Média Normandie), et a prolongé la réflexion en

proposant une cartographie autour de sept fonctions structurantes des métiers de l'information. Ces fonctions mettent en lumière la collaboration nécessaire entre scientifiques, bibliothécaires et informaticiens. Elles sont les suivantes : collecte, traitement, analyse, diffusion, curation, gestion des métadonnées, accompagnement des chercheurs. Isabelle Blanc propose une approche par fonctions pour décroiser les rôles et valoriser la diversité des compétences. Romain Féret, lui, identifie un angle mort : les acteurs intervenant au début du cycle de vie des données, souvent situés dans les laboratoires, restent invisibles malgré leur expertise. Claire Toussaint rappelle que la reconnaissance de cette même expertise conditionne la légitimité professionnelle. Ces trois perspectives, loin de s'opposer, s'imbriquent et convergent. Pour rendre visible ce travail invisible, il est nécessaire de faciliter le dialogue entre acteurs, d'expérimenter et de construire un catalogue de services organisé par fonction plutôt que par structure. Cette approche permettrait une meilleure valorisation et reconnaissance des métiers liés aux données de recherche.

Enfin, la table ronde n° 8 intitulée *Formation des professionnels de l'information à la gouvernance des données : entre technique et humanité*²⁰, modérée par Florence Thiault (maîtresse de conférences à Rennes), avec Ambre Davat (maîtresse de conférences, Enssib), Romain Féret (directeur de Média Normandie) et Ghislaine Charton (professeur émérite au Cnam), a montré toute l'importance de former des professionnels conscients des dimensions techniques, politiques et éthiques de la donnée. La gouvernance ne relève pas seulement du management ou de la sécurité, mais suppose une culture partagée de la donnée, une compréhension globale des usages et une motivation à mettre en place des pratiques responsables. Romain Féret a notamment rappelé l'importance d'accompagner et de former en amont des projets de recherche pour clarifier les pratiques de collecte, de description et de réutilisation des données. Cela favorise des définitions communes entre communautés scientifiques. Face à la complexité des écosystèmes numériques, une acculturation mutuelle d'un métier à l'autre est indispensable, notamment pour faire dialoguer informaticiens, documentalistes et chercheurs.

Les intervenants ont également souligné le rôle central de la médiation dans ces métiers émergents. Les professionnels doivent savoir communiquer avec des acteurs aux compétences et logiques variées, plutôt que de s'en remettre à des postes prédéfinis. Le profil du *data steward*²¹ illustre cette évolution, à la croisée de la technique, de la médiation et de la compréhension des contextes scientifiques. Enfin, tous ont appelé à une montée en compétence technique

15 https://www.youtube.com/watch?v=zA2YU6XV01A&list=PLEjcwQr_8kwIQ-OBiED1pSiO5OCsgWlj&index=12

16 Système intégré de gestion de bibliothèque.

17 https://www.youtube.com/watch?v=TRPii-z3b7M&list=PLEjcwQr_8kwIQ-OBiED1pSiO5OCsgWlj&index=13

18 https://www.youtube.com/watch?v=dGu0YWAaRNc&list=PLEjcwQr_8kwIQ-OBiED1pSiO5OCsgWlj&index=14

19 Ministère de l'enseignement supérieur, de la recherche et de l'espace.

20 https://www.youtube.com/watch?v=XNpkNcnBPTI&list=PLEjcwQr_8kwIQ-OBiED1pSiO5OCsgWlj&index=15

21 <https://choisirleservicepublic.gouv.fr/metiers/numerique/data-steward/>

dans les formations, mais aussi à une meilleure reconnaissance des savoirs issus des sciences humaines et sociales, pour construire une gouvernance des données réellement partagée et incarnée.

Pour conclure cette huitième édition de la Biennale du numérique, Isabelle Blanc (administratrice ministérielle des données, des algorithmes et des codes sources, MESRE), lors de son intervention (*Conférence de clôture*²²), a proposé une réflexion sur la gouvernance des données comme enjeu stratégique et, surtout, comme catégorie d'action collective. Inscrite dans le cadre plus large des politiques publiques de la donnée et de la science ouverte, la gouvernance est envisagée non comme un simple dispositif de régulation, mais comme le passage d'unités informationnelles dispersées à un véritable commun informationnel, appropriable par des communautés professionnelles plurielles. Cette mutation suppose de dépasser une conception strictement productiviste et concurrentielle de la donnée, pour lui donner une valeur d'usage durable, souveraine et juste.

Elle a ainsi rappelé que la fonction de gouvernance s'exerce à l'intersection de plusieurs registres : identification et préservation d'un capital de données, définition de cadres de référence partagés, gestion des risques et des opportunités, pilotage des cycles de vie, arbitrages en matière d'ouverture, de confidentialité et d'interopérabilité. La gouvernance des données relève d'une fonction transversale, qui met en tension des exigences parfois contradictoires (sécurité, circulation, conformité réglementaire, innovation). Elle requiert donc une coordination fine des acteurs et des cellules au sein de l'écosystème numérique de la recherche.

Enfin, il s'est agi pour Isabelle Blanc d'élargir la perspective en identifiant plusieurs enjeux majeurs. Les données doivent être pensées comme un bien commun, dont la valeur repose sur les usages concrets qu'elles permettent, et la souveraineté numérique ne signifie pas fermeture. Elle passe par des licences adaptées qui protègent, tout en favorisant le partage. Dans cette optique, l'équité et l'inclusion sont centrales, notamment pour les pays et institutions défavorisés qui doivent aussi pouvoir accéder aux données et participer à leur valorisation. La mise en œuvre des principes FAIR²³ repose sur des infrastructures ouvertes, permettant la découvrabilité et la réutilisation. La transparence et l'éthique sont essentielles face à la désinformation, la marchandisation des savoirs et l'atteinte aux données sensibles. Quant aux technologies émergentes (IA, blockchain, informatique quantique), elles ne sont pas une fin en soi mais des accélérateurs qui transforment les pratiques et renforcent la dimension citoyenne de la culture de la donnée.

Au terme de ces réflexions, cette Biennale du numérique aura montré que la gouvernance des données est loin d'être un simple enjeu technique. Elle constitue désormais un horizon collectif dans lequel le rôle des professionnels de l'information réside peut-être d'abord dans leur contribution à la souveraineté des données et à leur capacité à transformer une masse de données hétérogènes en un véritable commun durable, interopérable et émancipateur. ☉

22 https://www.youtube.com/watch?v=qCD_Twnznj0&list=PLEjcwQr_8kwIQ-OBiED1pSiO5OCSgWLj&index=16

23 La notion de FAIR data, ou données FAIR, recouvre les manières de construire, stocker, présenter ou publier des données de manière à permettre que les données soient « faciles à trouver, accessibles, interopérables et réutilisables » (en anglais : findable, accessible, interoperable, reusable), d'où l'acronyme « FAIR ». Source : https://fr.wikipedia.org/wiki/Fair_data