

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 ; Paillart, 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.

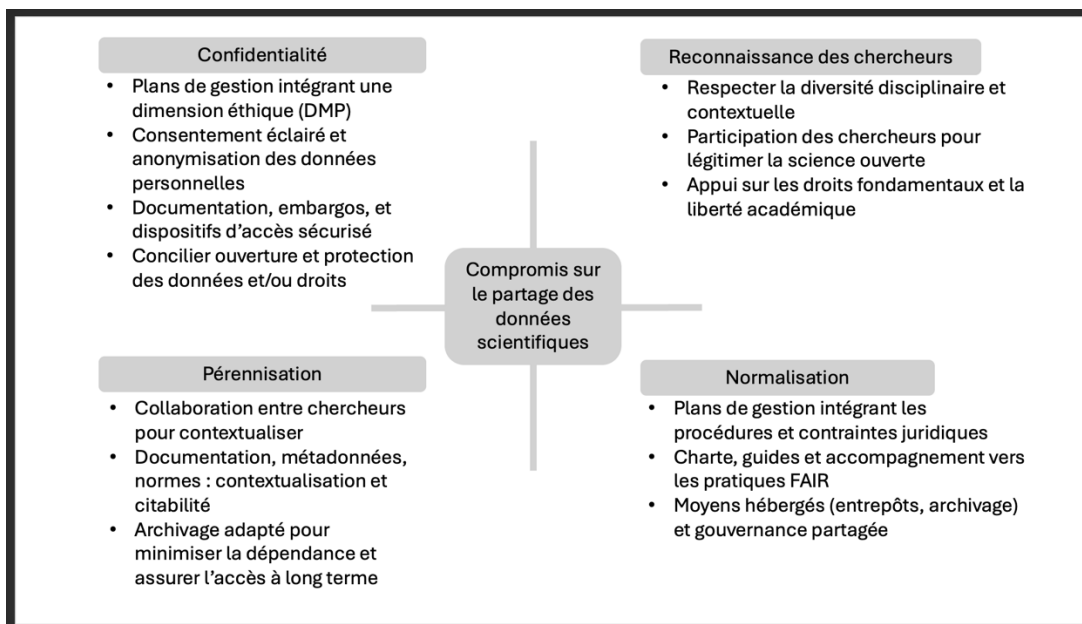
Star, Susan Leigh ; Strauss, Anselm (1999), « Layers of Silence, Arenas of Voice : The Ecology of Visible and Invisible Work », *Computer Supported Cooperative Work*, n° 8, p. 9-30.

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