



HAL
open science

Les cultures de données : genèse d'une notion, enjeux épistémologiques et implications pratiques

Nathalie Casemajor Loustau

► To cite this version:

Nathalie Casemajor Loustau. Les cultures de données : genèse d'une notion, enjeux épistémologiques et implications pratiques. Millerand, F., Coutant, A., Latzko-Toth, G., Millette. Datafication et publics de données. Penser la mise en donnée de la société, Presses de l'Université de Montréal, A paraître. hal-04254885v1

HAL Id: hal-04254885

<https://hal.science/hal-04254885v1>

Submitted on 23 Oct 2023 (v1), last revised 22 Jan 2024 (v2)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License

Nathalie Casemajor (à paraître, 2024) « Les cultures de données : genèse d'une notion, enjeux épistémologiques et implications pratiques », in Millerand, F., Coutant, A., Latzko-Toth, G., Millette, M. (dirs.) *Datafication et publics de données. Penser la mise en donnée de la société*. Montréal : Presses de l'Université de Montréal.

Les cultures de données : genèse d'une notion, enjeux épistémologiques et implications pratiques

En 2014, Microsoft a lancé un cycle de formation à « la culture des données » (*data culture*) destiné aux analystes de données, architectes de données et autres spécialistes de la science des données. La « culture des données » y est présentée comme un champ de compétence et une mentalité à faire partager par l'ensemble des membres de l'organisation. « Aujourd'hui les organisations doivent adopter une culture axée sur les données, sans quoi elles risquent de finir à la traîne »¹ affirme Microsoft dans la promotion de sa plateforme infonuagique Azure. L'expression de « culture des données » (ou « culture de la donnée ») a ainsi connu une diffusion accélérée depuis les années 2010. Elle désigne au sens large un ensemble de dispositions permettant de s'orienter et de travailler dans un monde de plus en plus médiatisé par l'intermédiaire des données numériques.

Malgré un usage croissant, à ce jour seule une poignée de travaux se sont véritablement attelés à définir et à conceptualiser cette notion. En quoi l'ensemble des relations aux données numériques peut-il constituer une culture ? L'objectif de ce chapitre est d'interroger la notion de culture de données pour en éclairer la genèse, les enjeux épistémologiques et les implications pratiques. Dans ce cadre, analyser la relation entre données et culture permet d'éclairer comment des « publics de données » se composent à l'intérieur d'une formation discursive plus large qui est celle des cultures de données. J'adopte ici la forme doublement plurielle de *cultures de données* pour souligner la variété des cultures et des données concernées. Cette réflexion s'appuie sur une revue critique de la littérature francophone et anglophone, enrichie par des exemples tirés du secteur professionnel des arts et de la culture au Québec.

Dans un premier temps, je présenterai le contexte d'émergence de cette notion. Puis je distinguerai différents plans culturels et pôles d'activité à l'intérieur desquels situer la production de données et ses publics. J'en tirerai finalement des éléments synthétiques de définition et d'analyse. En interrogeant de manière critique la notion de culture de données et ses présuppositions, ce chapitre vise à contribuer à sa conceptualisation dans le champ des études critiques de données et des études culturelles.

¹ « Organizations today must embrace a data-driven culture or risk being left behind » (<https://azure.microsoft.com/>).

1. Enjeux : interroger les présupposés des cultures de données

La notion de cultures de données est principalement associée à deux types de données : les données ouvertes et les données massives. Ainsi, la nouvelle plateforme de données ouvertes de la Ville de Montréal a été lancée en 2020 avec l'ambition de « développ[er] une culture de la donnée à Montréal »² (voir Millerand et Meunier, le présent ouvrage). La création en 2018 à Montréal d'un *Pôle sur les données massives* en culture visait quant à elle à « développer une culture de la donnée » dans les organisations culturelles³. Situés dans différentes logiques d'exploitation des données (l'une à vocation citoyenne et l'autre davantage commerciale), ces deux discours ont en commun l'idée que « la culture des données » doit être développée.

De fait, l'acquisition d'une culture de données est présentée dans les discours qui en font la promotion comme nécessaire pour s'adapter aux nouveaux environnements numériques. L'attrait pour cette nouvelle « culture » (voire l'injonction à l'adopter) est alimentée par une croyance en l'efficacité technique et économique d'un modèle d'activité gouverné par l'exploitation des données numériques. Les données sont en effet devenues le moteur d'un nouveau paradigme de modernisation et d'innovation socio-économique. Ce paradigme est modelé par le succès phénoménal des grandes plateformes du Web dont l'activité est fondée sur une triple logique d'extraction, d'accumulation et de valorisation des données. Un tel succès rend désirable l'adoption d'une culture de données similaire qui permettrait à toute organisation de reproduire ce modèle et d'accroître ainsi sa valeur économique et sociale.

Dans le monde du travail, l'adoption d'une culture de données pose donc l'enjeu d'une transformation des pratiques professionnelles pour répondre à ce nouveau paradigme organisationnel de l'innovation par les données. Cette transformation se traduit par l'émergence de nouveaux métiers et de nouvelles compétences qui permettraient d'intégrer les données numériques dans les orientations stratégiques et dans les processus opérationnels des organisations. Poussé par la vague ascendante des données massives et de l'intelligence artificielle, le domaine de compétence qui s'est développé au croisement de l'analyse des données et de l'informatique est présenté dans ces discours comme incontournable. Pour des entreprises comme Microsoft, l'injonction à développer une culture de données revêt aussi un intérêt commercial : diffuser cette vision d'innovation et les compétences qui lui sont liées nourrit les ventes des produits et services mis en marché par les industries de la donnée (bases de données, serveurs, applications d'apprentissage machine, tableaux de bord de visualisation des données).

Or, ces discours sont chargés de présupposés normatifs rarement interrogés. Ils reposent sur les prémisses suivantes : la *tabula rasa*, autrement dit le fait que les professionnels seraient exempts de toute culture de données préalable ; l'*urgence* à acquérir « une culture » des données face à la menace de perdre sa pertinence et de disparaître ; et l'*efficacité* de cette culture qui, en optimisant la prise de décision, garantirait la croissance

² Ville de Montréal (2020) « Dévoilement du nouveau site web des données ouvertes » (<https://donnees.montreal.ca>).

³ Synapse C, « Outils » (<https://synapsec.ca/>).

et le succès des organisations. Étant donné la diffusion croissante de ce discours dans les secteurs privé, public voire associatif, il est important d'interroger les valeurs qui le sous-tendent.

À travers la notion de *cultures* de données se noue en effet l'enjeu politique du rapport entre savoir (les informations et connaissances fournies par l'analyse des données numériques) et pouvoir (la prise de décision dans les organisations, l'injonction à se conformer à un nouveau modèle d'activité, les potentielles atteintes à la vie privée des usagers). Ces enjeux concernent non seulement les professionnels sommés d'adapter leurs pratiques de travail, mais aussi leur relation avec les récepteurs des biens et services produits par cette nouvelle organisation du travail : clients, usagers ou publics appréhendés à travers leur « mise en donnée ». Pour mieux comprendre les enjeux de cette notion et ses implications pratiques dans la mise en données de la société, il est nécessaire d'en creuser la conceptualisation. Dans quelle mesure les différents contextes d'émergence de la notion influencent-ils les approches conceptuelles des cultures de données ? Et en retour, quels sont les différents effets heuristiques et pratiques de ces conceptualisations ? Je propose d'examiner les définitions existantes de cette notion en posant un double regard sur les travaux francophones et anglophones. Il s'agira d'en identifier l'origine, les points communs, les différences et les effets sur la production de savoir.

2. Problématiser les cultures de données

2.1 Origine et voisinages

L'expression anglaise « *data culture* » a émergé dès les années 1960 avec le boom de l'industrie informatique. Tombée en désuétude dans les années 1980 et 1990, elle tend alors à être supplantée par celle d'« *information culture* », comme le montre une comparaison des trajectoire des deux expressions dans Google Ngram Viewer. L'expression a suscité un nouvel intérêt dans les années 2010 en même temps que la croissance exponentielle du phénomène de « mise en données » (*datafication*) de la société, avec en corollaire l'émergence de nouveaux enjeux liés à la « littératie des données ». L'étude des cultures de données avoisine ces autres notions tout en s'en distinguant.

S'ils sont connexes, les termes de donnée et d'information ont un sens clairement distinct : une donnée est un élément basique descriptif de la réalité qui peut être codé sous forme numérique et traité par des systèmes automatisés. Seulement après avoir été interprétée et mise en contexte, une donnée prend la valeur d'information. Quant à la littératie des données, elle réfère au sens large aux connaissances et compétences nécessaires pour comprendre et utiliser les données et pour réfléchir à leurs enjeux dans l'environnement numérique. Alors que la notion de littératie des données est davantage centrée sur les apprentissages engagés au niveau individuel, la notion de cultures de données est plus large et plus englobante : par culture, on entend ici *un ensemble de répertoires de sensibilité et de signification collective partagés au sein d'un groupe*.

C'est à l'intérieur de ce cadre culturel plus large qu'on peut situer la notion de *public de données* : les manières d'imaginer les publics comme entités collectives au sein d'environnements numériques (Litt et Hargittai, 2016), les manières de les constituer en pratique (via la collecte de données, la médiation algorithmique et les mesures d'audiences; Gillespie, 2014), les manières dont ils s'assemblent effectivement à travers les environnements numériques (Marres, 2005), et les manières de se confronter à leur agentivité (Møller Hartley et al., 2021) – toutes dépendent en bonne partie de schémas d'action et d'interprétation d'ordre culturel.

Parmi les précurseurs de l'étude des cultures de données, le sociologue des sciences et technologies Geoffrey Bowker (2000) s'est intéressé dès la fin des années 1990 à l'existence de « cultures locales de données » (*local coding culture*) dans les infrastructures d'information. Il entendait par là des schémas locaux de classification de l'information propres à différentes disciplines scientifiques qui collaborent dans des projets de recherche où une multitude de données de différentes provenances (une « diversité des données » ou *datadiversity*) doivent être rassemblées dans un même système d'information.

Dans les travaux contemporains, on peut identifier deux principaux courants ayant significativement creusé le sujet des cultures de données : d'une part les études critiques des données initiées dans le champ anglophone (*critical data studies*), et d'autre part les études françaises en sciences de l'information et de la communication. S'y ajoutent de manière connexe les travaux anglophones en science de l'information (*information studies*, sous l'angle de la gestion des systèmes d'information), la science de la gestion et l'économie de la culture.

Il reste cependant plusieurs lacunes à combler dans la conceptualisation de cette notion. Premièrement, les définitions existantes tendent à se limiter à des objets restreints et à ignorer les travaux menés dans d'autres champs de recherche. Et deuxièmement, peu de croisements existent entre les conceptualisations anglophones et francophones, qui se développent en parallèle. C'est à ces manques que ce chapitre souhaite répondre, en offrant une analyse croisée de ces travaux.

2.1 Analyse croisée des définitions existantes

Si de nombreux travaux académiques utilisent la notion de culture(s) de données, très peu la problématisent réellement pour en proposer une définition explicite et argumentée. L'examen des travaux existants a permis d'identifier quatre principales définitions de la notion de cultures de données (Tableau 1). Ces résultats sont issus d'une interrogation systématique des bases de données Scopus, Isidore and Cairn en novembre 2022, couplée à une recherche exploratoire sur Google Scholar et à un processus de « boule de neige ». Ces définitions ont été sélectionnées pour la profondeur de leur élaboration épistémologique, au-delà d'un simple usage de l'expression non problématisé. Celles des auteurs anglophones (Bates, Burgess) sont ancrées dans les études critiques des données (Kitchin 2014), tandis que les auteurs francophones (Labelle, Lehmans) s'inscrivent dans le prolongement d'une perspective critique sur la culture de l'information, issue des sciences de l'information et de la communication.

Tableau 1 : Quatre définitions de la notion de cultures de données

Auteur	Objet et approche	Définition
Labelle et al. & FING (2017 : 5)	Stratégie organisationnelle. Sciences de l'information et des communications	« [...] la connaissance de gestes et de processus techniques, intellectuels et sociaux de traitement des données dans des situations professionnelles ou quotidiennes. La culture des données est composée de savoirs et savoir-faire issus des statistiques et de l'informatique. La culture des données suppose une compréhension des médias et de la circulation de l'information. »
Lehmans (2017 : 22)	Données ouvertes (organisations professionnelles). Sciences de l'information et des communications.	« La culture de la donnée est composée de savoirs, de valeurs et de compétences statistiques, informatiques, médiatiques, et informationnelles qui donnent les moyens aux professionnels de comprendre le sens de leur activité de collecte et de mise en circulation des données, de la situer dans son contexte social d'usage, et de coordonner les gestes et les procédures professionnels. Cette culture est nécessairement propre à chaque organisation qui construit ses pratiques ».
Bates (2017 : 5)	Ville intelligente. Étude critique des données.	« [...] les différentes normes culturelles, systèmes de valeurs et croyances qui informent, encadrent et justifient les pratiques des gens en matière de production, de traitement, de distribution ou d'utilisation des données (pratiques de données), ainsi que leurs efforts pour gouverner et façonner des formes particulières de pratiques de données par divers moyens sociaux et techniques » ⁴ .
Burgess dans Albury et al. (2021 : 2)	<i>Apps</i> numériques et médias sociaux. Étude critique des données et études culturelles.	« [...] les pratiques qui se forment en rapport avec et en réponse aux médias sociaux et autres données (et traces de données) que les gens génèrent dans leur vie quotidienne. Ces pratiques se forment à partir de nos divers engagements, expériences et approches de la compréhension et de la négociation de ces données » ⁵ .

⁴ « [...] the different cultural norms, value systems and beliefs, that inform, frame and justify peoples' practices of data production, processing, distribution, or use (data practices), as well as their efforts to govern and shape particular forms of data practices through a variety of social and technical means ».

⁵ « [...] the practices that form around and in response to the social media and other data (and data trails) that people generate as we go about our daily lives. These practices form from our diverse engagements with, experiences of, and approaches to understanding and negotiating these data ».

Les quatre définitions se rejoignent autour de quelques éléments centraux, mais elles se distinguent aussi par leurs objets, qui viennent influencer le cadrage de la notion. Toutes concordent sur les points suivants : il existe une *diversité* des cultures de données ; elles sont *situées* dans des contextes sociaux, matériels et culturels particuliers qui contribuent à leur donner forme ; et elles sous-tendent un engagement dans *l'action* et dans la production de *signification* par une multiplicités d'acteurs (entreprises, pouvoirs publics, société civile, travailleurs, usagers ordinaires).

Dans le champ des sciences de l'information et de la communication, les travaux qui ont abordé de front cette notion croisent les domaines de la documentation, de l'éducation et de la communication organisationnelle. Lehmans (2017, 2018) et Labelle et al. (2017) ont en commun une perspective centrée sur les transformations des modèles organisationnels et sur les nouvelles compétences professionnelles liées aux données : leur définition insiste sur le partage d'une vision commune par les membres de l'organisation et sur la compréhension critique du sens de l'action.

La définition de Labelle et al. a été élaborée dans un cahier visant la formation professionnelle, produit par la Fondation Internet Nouvelle Génération (la FING, une fondation française spécialisée dans l'accompagnement des transformations numériques des organisations) ; elle est donc davantage orientée vers le déploiement et l'appropriation d'une « nouvelle » culture des données dans les organisations. Lehmans ancre plutôt ses réflexions dans l'étude des initiatives de données ouvertes au sein des collectivités territoriales pour montrer que la transformation des pratiques organisationnelles n'est pas linéaire : elle génère des conflits de valeur, des contradictions, des bricolages et des détournements de la part des membres des organisations.

Du côté des études critiques des données, les travaux qui se sont attelés à définir la notion s'inscrivent d'une part dans l'étude des phénomènes de la ville intelligente, de la participation civique et de leurs enjeux socio-politiques. C'est le cas de Bates (2017) qui met en rapport les notions de pratique de données, de culture de données et de gouvernance des données. Sa définition insiste sur une double dimension d'analyse spatiale : elle prend en compte à la fois la spécificité locale des différents sites de pratique, et leur interconnexion dans un réseau plus large d'environnements techniques, de politiques publiques, de législations et de systèmes économiques.

D'autre part, les travaux de Burgess (2017, 2022) s'inscrivent dans l'étude du genre et de la sexualité pour étudier la transformation du rapport à l'intimité dans le contexte des médias numériques. Elle développe une perspective inspirée des études culturelles pour étudier des applications de rencontre. Sa définition propose une approche vernaculaire des cultures de données, centrée sur l'expérience ordinaire des médias socionumériques par leurs usagers. Cette conceptualisation met en lumière l'agentivité et la subjectivité des usagers ordinaires qui contribuent eux aussi à forger les cultures de données dans les activités de la vie quotidienne (Burgess *et al.*, le présent ouvrage).

3- Proposition de conceptualisation

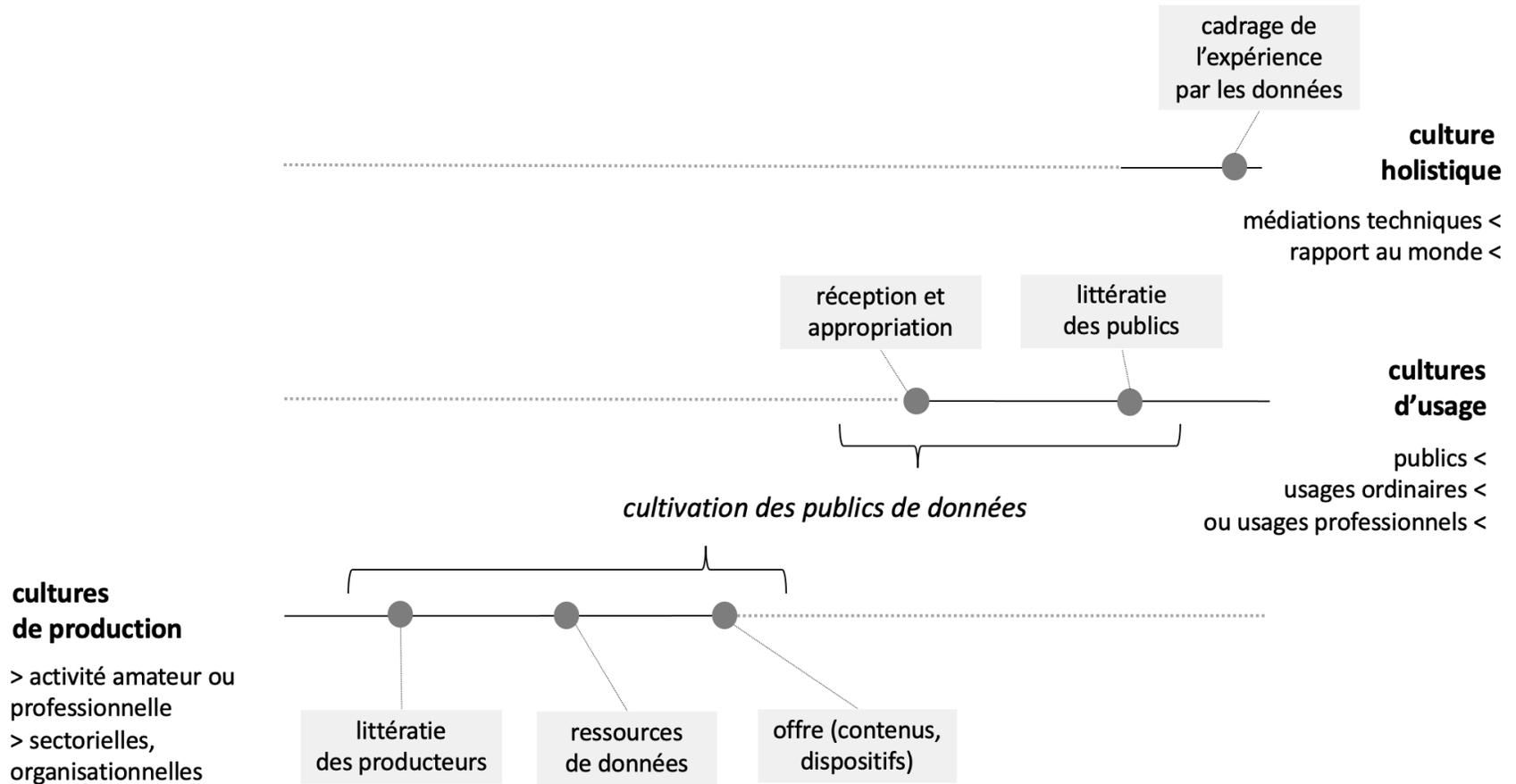
3.1 Croiser les plans culturels du rapport aux données : la cultivation des publics

Cette recension permet de faire ressortir que l'étude des cultures de données se décline sur différents plans de la notion de culture, une des plus versatiles du vocabulaire des sciences humaines et sociales. Je propose de distinguer trois plans (Figure 1) : celui de la *culture holistique*, celui de la *culture professionnelle* et celui de la *culture d'usage*. Ces trois plans se superposent et ne sont pas mutuellement exclusifs.

Le plan le plus large et le plus général est celui des cultures de données au sens holistique : il englobe les effets croisés des médiations culturelles et techniques sur les dimensions symboliques de la vie. Dans une perspective ethnoculturelle ou identitaire, la culture désigne un ensemble de traits et de dispositions socio-symboliques incorporés, largement partagés au sein d'un groupe social élargi à l'échelle d'une société, d'une nation ou d'un collectif d'appartenance, et constitutifs de son identité et de son rapport au monde. S'il peut être pertinent d'envisager les cultures de données sous l'angle de leurs spécificités nationales et locales, l'approche suggérée ici consiste plutôt à adopter la perspective d'une anthropologie culturelle des médias au sens d'une analyse des médiations techno-symboliques propres aux données numériques. Il s'agit alors d'étudier les environnements techniques de mise en données comme médiateurs d'un rapport expérientiel au monde (dans ses dimensions tant pratiques que sensibles et affectives). En effet, plusieurs dimensions fondamentales de la vie comme le rapport à la mobilité, au savoir ou à la santé sont de plus en plus expérimentés par l'intermédiaire de dispositifs numériques (applications, services) fondés sur la collecte et l'exploitation de données. Les *cadres d'expérience d'un monde mis en données* apparaissent ainsi comme un trait culturel de la modernité technique.

Le plan des cultures de production de données (et de dispositifs fondés sur des ensembles de données) est resserré sur l'activité professionnelle ou amateur visant la création ou la transformation de données. Ces cultures de production possèdent des caractéristiques, des normes et des valeurs propres à un groupe plus ou moins circonscrit (secteur socio-économique, métier, organisation). Ce plan permet d'entrer davantage en profondeur dans l'analyse de contextes situés. En prenant le secteur professionnel des arts et de la culture comme terrain de référence, on peut distinguer des catégories d'acteurs situés dans différents pôles d'activité.

Figure 1 : Les plans culturels du rapport aux données



Du côté des producteurs, les pratiques relevant des cultures de données impliquent une activité de *constitution de ressources de données* (Luka, le présent ouvrage). La nature de ces ressources est propre à chaque secteur, voire à chaque organisation ; par exemple dans le secteur du spectacle vivant, les données transactionnelles issues des plateformes de billetterie sont mises en rapport avec les métadonnées descriptives des œuvres diffusées. Ces ressources contribuent à la *production d'une offre orientée par l'analyse des données (data-driven)*, telle la recommandation personnalisée de spectacles en fonction des choix précédents de groupes de spectateurs semblables. Le développement de cette offre s'appuie sur l'acquisition d'une *littératie professionnelle* des données nécessaire pour utiliser des outils spécialisés de collecte, de gestion et d'analyse des données (bases de données ou Google Analytics par exemple). Un ensemble d'intermédiaires (plateformes, acteurs de la société civile tel Wikidata, Synapse C au Québec ou la Fondation FING en France) contribuent également à façonner les cultures de données au sens où ils participent à la médiation des données et à la formation en littératie des données.

Le plan des cultures d'usage se situe du côté des usagers et des publics⁶ des dispositifs de données. Ils développent des pratiques liées à la *réception de l'offre* et aux *usages quotidiens (ordinaires)* des dispositifs connectés. Les cultures d'usage étant influencées par une multiplicité de facteurs sociaux, culturels et techniques, leur compréhension implique l'analyse de contextes situés (les études de cas présentés dans les autres chapitres du présent ouvrage en sont l'exemple, voir Coutant et Vial, le présent ouvrage). L'activité des publics de données engage elle aussi une *littératie*, toutefois dans le cas des usagers dits ordinaires, cette littératie ne nécessite pas des compétences de collecte et d'analyse de données, mais plutôt des savoir-faire qui intègrent les dispositifs de données dans les pratiques quotidiennes.

C'est à l'intersection de ces deux derniers plans des cultures de données (celui de l'activité de production de l'offre et celui des usages) que l'on peut situer la rencontre, l'écart ou le décalage entre les représentations que les professionnels se font des publics (les publics imaginés) et leurs pratiques concrètes d'appropriation des dispositifs (les publics effectifs) (voir Harvey et Latzko-Toth, le présent ouvrage; Millerand et Meunier, le présent ouvrage). Pour creuser l'analyse de ces points de contact entre publics imaginés et effectifs, la notion de *cultivation*⁷ des publics de données (Møller Hartley et al. 2021) semble féconde car elle permet d'associer plusieurs dynamiques en jeu : les intentions et les représentations des producteurs, les affordances des dispositifs de données et des algorithmes dans leur potentiel de « calcul » et de « programmabilité » des publics (Gillespie 2014), ainsi que l'irréductible incertitude de leurs résultats en termes de formation et d'agentivité des publics.

⁶ Les termes d'usager et de public et sont à distinguer : le premier renvoie aux formes d'utilisation et d'appropriation des dispositifs, tandis que le second suppose en outre une communauté d'intérêt et l'expression d'une position (Møller Hartley et al. 2021; Millerand et al., le présent ouvrage).

⁷ Dans les études des médias, le terme de « cultivation » a été popularisé par Gerbner (1998) qui l'a mobilisé pour analyser la réception des contenus télévisuels et ses effets sur le partage de mêmes conceptions du monde chez des publics différenciés. L'usage de ce terme dans la littérature récente sur les cultures de données sont toutefois sensiblement différentes.

Le découpage proposé ici entre différents plans culturels et pôles d'activité résonne avec le découpage d'Albury, Burgess et al. (2021 : 2) qui proposent quatre acceptions du terme de cultures de donnée : « les modes de cultivation des données », « les cultures de production », les « cultures d'usage » et la « mise en données de la culture ». Le terme de *cultivation* des données que l'on retrouve ici désigne les manières dont « diverses formes de données sont créées, nettoyées, ordonnées, récoltées et entrecroisées – par des acteurs multiples et distribués mais connectés, notamment des entreprises, des gouvernements, des développeurs, des annonceurs et des utilisateurs ». Cette acception permet de garder en tête que si l'identification d'un plan de production *professionnelle* de dispositifs de données permet de creuser les logiques propres au monde du travail et aux organisations d'un secteur, elle ne doit pas obscurcir le fait que les données sont coproduites par une multiplicité d'acteurs, incluant les publics. Ainsi un « public de données » n'est pas réductible au produit d'une culture de données particulière, mais il contribue tout autant à la coproduire par ses pratiques de consommation et de contribution sur les plateformes, par ses appropriations des dispositifs de données et par ses tactiques de détournement.

3.2 Éléments de définition et dimensions d'analyse

À la lumière de ces réflexions, je propose ci-dessous une synthèse des principaux éléments de définition de la notion de cultures de données ainsi qu'une série de dimensions d'analyse. Ce cadre a été construit pour être mobilisé dans un chantier de recherche sur l'activité professionnelle du secteur des arts et de la culture. L'objectif n'est pas de proposer une définition généralisable à tous les objets de recherche, ce qui escamoterait leurs spécificités et limiterait la profondeur des analyses possibles. Il s'agit plutôt de dégager des constantes dans les définitions proposées à ce jour, et d'identifier des dimensions d'analyse qui seraient nécessairement à adapter en fonction des objets considérés.

Les éléments de définition proposés sont les suivants.

Une culture de données est un assemblage de répertoires collectifs de représentation, d'action et de justification situé dans un contexte spécifique et caractérisé par une sensibilité, un univers de sens et une rationalité construits en relation avec des ensembles de données.

Ces répertoires contribuent à façonner les données, leur statut et leurs effets sur le monde. Ils reposent sur des valeurs, des normes (explicites ou implicites), des littératies, des pratiques, des affects et des dispositifs techniques.

Ces éléments forment un système possédant une cohérence interne mais non uniforme, traversé par des lignes d'ambivalence, de tension et de contradiction. Ils s'enchevêtrent avec d'autres cultures de données, toutes animées par des dynamiques à la fois locales et globales.

Le tableau ci-dessous détaille les six principales dimensions retenues pour étudier ces répertoires : les *valeurs* associées à une culture de données spécifique, les *normes* qui lui sont propres, la *littératie* des données qu'elle présuppose, les *pratiques* par lesquelles et se

déploie, les *affects* qu'elle véhicule et les *dispositifs matériels* qui sous-tendent sa formation. Le Tableau 2 développe ces dimensions en les illustrant par des exemples tirés d'observations dans le secteur des arts et de la culture au Québec et d'un entretien collectif mené avec des professionnels des bibliothèques et archives en 2021 dans le cadre d'une étude sur la participation culturelle numérique (Sirois, Casemajor et al. 2021).

Dans l'optique d'étudier la formation et l'évolution des cultures de données d'un secteur professionnel ou d'une organisation en particulier, il est utile d'interroger la relation entre la notion de culture de données et celles de culture organisationnelle et de culture professionnelle. La première prend l'organisation comme milieu structurant pour la formation de cultures spécifiques et localisées, relativement partagées au sein d'un groupe délimité par une activité de coopération visant un but commun (dans une entreprise, une administration, une association). La seconde concerne les cultures spécifiques à chaque groupe professionnel sous l'angle des métiers. Ces cultures sont forgées par des processus de scolarisation (formation, diplôme) et de socialisation au sein d'organisations et d'associations professionnelles.

On peut tirer de la sociologie des professions d'importants éléments de conceptualisation des cultures de données. Les travaux de Montjardet en particulier proposent une critique de l'approche dite « culturaliste » de la culture professionnelle, jugée réductrice au sens où elle chercherait à identifier des attributs différentiels assignés à l'ensemble du groupe, c'est-à-dire « un catalogue de traits (ou propriétés) partagés (à des degrés divers) par tous les professionnels du même secteur » (1994 : 408). Montjardet met plutôt en évidence « à la fois un pluralisme considérable des attitudes, valeurs et pratiques et un langage commun, une identique mobilisation autour des enjeux sous-jacents aux acceptions différentes, souvent conflictuelles, des mêmes missions, du même métier » (1994 : 409). Ainsi, bien que l'étude des groupes professionnels délimités par le partage d'un même métier ne soit qu'un des aspects de la composition des cultures de données, les travaux de Montjardet invitent judicieusement à les envisager comme des objets non réductibles à une liste de propriétés. Ainsi, il serait réducteur de parler d'une culture de données propre aux bibliothécaires, qui se différencierait de manière homogène par rapport à une culture des données du secteur muséal ou musical.

Tableau 2 : Dimensions d'analyse des cultures de données

Dimension	Description	Exemples
Valeurs	Croyances fortes au sujet des données, fondées sur des principes d'action supérieurs.	<p>L'<i>objectivité</i> et la <i>complétude</i> des données comportementales (tel l'historique d'emprunt de documents recueilli via les comptes des abonnés des bibliothèques) permettraient de comprendre les pratiques des publics dans leur globalité, voire de « laisser parler les publics » en toute transparence par la mesure de leurs comportements.</p> <p>L'<i>efficacité technique</i> des dispositifs d'intelligence artificielle et de recommandation algorithmique permettraient de mettre en relation « la bonne offre » pour « le bon public ».</p>
Normes	Cadres sociaux qui régulent les comportements liés aux données ; normes d'usage ; normes professionnelles explicites (ex. code déontologique) ou implicites (normes pratiques).	<p>La <i>protection des données personnelles</i> des publics fait partie des normes professionnelles des bibliothécaires et archivistes. Cette norme est transmise durant la formation scolaire et enchâssée dans des cadres légaux (comme la <i>Loi sur les Archives</i> du Québec qui fixe les conditions de protection des renseignements personnels dans les archives privées) et dans des politiques institutionnelles (telle la <i>Politique de confidentialité</i> de Bibliothèque et Archives nationales du Québec). Ces normes peuvent modeler ou freiner les pratiques d'exploitation des données comportementales.</p>
Littératies	Savoirs et savoir-faire liées au cycle de vie des données (production/extraction, stockage, traitement, gestion, exploitation, partage et diffusion des données) et aux usages des dispositifs de données. Les connaissances et	<p>Dans le domaine des bibliothèques et archives, la littératie des données s'est d'abord développée via l'<i>informatisation des bases de données documentaires</i> (catalogage et indexation des documents) puis dans le domaine des <i>données ouvertes</i> qui concerne essentiellement les métadonnées descriptives des</p>

compétences liées aux données touchent notamment aux domaines de la statistique, de l'informatique, de l'information-documentation, du droit et des médias.

documents. La littératie des *données comportementales* est encore peu répandue chez les bibliothèques et archivistes, contrairement aux analystes de données et aux spécialistes du marketing qui travaillent dans les services de communication des institutions.

Pratiques Schémas observables d'activité, d'usage et de relations sociales stabilisés dans l'organisation de l'activité. Ces schémas se déploient dans un langage (vocabulaire), dans des processus opérationnels (procédures, protocoles, routines, modes de prises de décision, usages), dans des attitudes ou comportements et dans des modes d'interaction.

La prise de décision concernant l'acquisition de nouveaux ouvrages à intégrer dans les collections des bibliothèques est une pratique qui peut être influencée par les données comportementales : l'analyse de l'historique d'emprunt des abonnés peut orienter les acquisitions vers des types de documents populaires auprès des publics.

Affects Formes de sensibilité relatives à la mise en données et à l'expérience des dispositifs de données ; aspects émotionnels des pratiques de données ; formes symboliques et esthétiques de mise en représentation des données (ex. visualisation de données).

L'analyse des données comportementales soulève chez certains professionnels des bibliothèques l'*espoir* d'exploiter une « mine d'or sur laquelle on dort », mais aussi la *résignation* que cela ne soit qu'« un beau rêve qui ne se réalisera qu'aux États-Unis, jamais au Québec » en raison des limites technologiques des systèmes actuellement utilisés dans la province.

Dispositifs Supports matériels, outils, systèmes techniques et plateformes de données dont les affordances pratiques, normatives et axiologiques participent au cadrage des cultures de données.

Dans le secteur des bibliothèques, *les plateformes de prêt de livres numériques* sont les plus avancées en termes d'expérimentation avec les données comportementales : elles « nagent dans les données », tandis que la plupart des bibliothèques municipales opèrent avec des *systèmes au bord de l'obsolescence* qui ne leur permettent pas d'envisager de telles pratiques.

Conclusion

Problématiser la notion de cultures de données implique de faire la distinction entre ses propriétés heuristiques et les présupposés véhiculés par certains discours actuels sur l'innovation organisationnelle. L'usage croissant de cette notion dans les organisations reflète la nécessité pratique de s'adapter aux mutations numériques des environnements de travail. Mais elle s'inscrit aussi plus largement dans une formation discursive (au sens foucauldien) caractérisée par la « mise en données » de la société, autrement dit une configuration contemporaine du rapport entre dispositifs de données, savoirs et pouvoir. Dans ce contexte, l'injonction à développer de nouvelles cultures de données est alimentée par les effets idéologiques du *dataïsme*, c'est-à-dire « la croyance répandue dans la quantification objective et le suivi potentiel de tous les types de comportement humain et de la socialité grâce aux technologies numériques connectées » (van Dijck, 2014 : 198). Faire des cultures de données un objet de recherche implique donc de les historiciser de manière à saisir comment elles s'inscrivent dans des formes discursives plus larges, mais aussi d'en contextualiser les formes matérielles et les logiques d'action propres à divers milieux de pratique.

Conceptualisée de manière critique, la notion de cultures de données offre un éclairage sur les publics de données. Elle invite à prêter attention aux répertoires de sensibilité et de signification collective qui contribuent à forger le rapport entre production de dispositifs de données et formation des publics. Ce rapport engage une dynamique de coproduction entre cultures et publics de données. D'un côté, les cultures de données engagent des schémas d'action et d'interprétation (pratiques, normes, valeurs, affects, littératies) qui contribuent à forger des visions (public imaginé), des architectures algorithmiques (publics calculés) et des mesures des publics (audience). Mais en corollaire, la formation effective de publics et leur agentivité (usages, appropriations, expressions) agit sur les répertoires culturels de la mise en données. Cette agentivité des publics peut aussi bien contribuer à reproduire des répertoires en place qu'à engendrer des décalages transformatifs. Ainsi, les cultures de données ne sont pas des corpus de répertoires stables, mais plutôt des processus actifs de *cultivation*, travaillés par une multiplicité d'acteurs hétérogènes, incluant les publics.

BIBLIOGRAPHIE

- Albury, Kath, Jean Burgess et al., « Everyday Data Cultures » dans *AoIR Selected Papers of Internet Research*, 2021, p. 1-3.
- Bates, Jo, « Data cultures, power and the city », dans Rob Kitchin, Tracey Lauriault et Gavin McArdle (dirs.) *Data and the City*, Routledge, 2017, p. 189-200.
- Bowker, Geoffrey, « Biodiversity Datadiversity », *Social studies of science*, vol. 30, n°5, 2000, p. 643-683.
- Burgess, Jean, « Hookup apps' vernacular data cultures », conférence présentée à *The Social Life of Data*, RMIT University, Melbourne, 2017.
- Burgess, Jean et al., *Everyday Data Cultures*, John Wiley & Sons, 2022.

- Gerbner, George, « Cultivation analysis: An overview », *Mass Communication and Society*, vol. 1, n°3-4, 1998, p. 175-194.
- Gillespie, Tarleton, « The relevance of algorithms », dans Tarleton Gillespie, Pablo Boczkowski et Kirsten Foot (dirs.), *Media technologies: Essays on communication, materiality, and society*, MIT Press, 2014, p. 167-194.
- Kitchin, Rob, « Big Data, new epistemologies and paradigm shifts », *Big Data & Society*, vol. 1, n°1, 2014.
- Labelle, Sarah et al., *Vers une organisation data driven : mettre la culture des données au cœur de l'organisation*, Infolab – FING, 2017.
- Lehmans, Anne, « Données ouvertes et redéfinition de la culture de l'information dans les organisations. Vers une culture de la donnée », *Communication et organisation*, vol. 51, 201, p. 15-26.
- « L'horizon d'une culture de la donnée ouverte : de l'utopie aux pratiques de gouvernance des données », *Revue COSSI: communication, organisation, société du savoir et information*, n°4, 2018.
- Litt, Eden et Ezsther Hargittai, « The Imagined Audience on Social Network Sites », *Social Media+ Society*, vol.2, n°1, 2016.
- Marres Noortje, « Issues spark a public into being: A key but often forgotten point of the Lippmann–Dewey debate » dans Bruno Latour et Peter Weibel (dirs.), *Making things Public: Atmospheres of Democracy*, MIT Press, 2005, p. 208–217.
- Møller Hartley et al., « Researching publics in datafied societies: Insights from four approaches to the concept of 'publics' and a (hybrid) research agenda », *New Media & Society* (online first), 2021.
- Monjardet, Dominique, « La culture professionnelle des policiers », *Revue française de sociologie*, 1994, vol. 35, n°3, p. 393-411.
- Sirois Guillaume, Nathalie Casemajor et al., « Cultural participation through digital technology : A puzzling issue for cultural governance » dans Devin Beauregard et Jonathan Paquette (dirs.), *Canadian Cultural Policy in Transition*, Routledge, 2021, p.129-143.
- van Dijck, José, « Datafication, dataism and dataveillance: Big Data between scientific paradigm and ideology », *Surveillance & society*, vol. 12 , n°2, p. 197-208.