



HAL
open science

La data literacy distribuée. Périmètres définitionnels, origines documentaire, perspectives réticulaires

Ugo Verdi, Olivier Le Deuff

► To cite this version:

Ugo Verdi, Olivier Le Deuff. La data literacy distribuée. Périmètres définitionnels, origines documentaire, perspectives réticulaires. Les Cahiers du numérique, 2020, 16 (2-4), pp.137-173. 10.3166/LCN.2020.006 . hal-04299867

HAL Id: hal-04299867

<https://hal.science/hal-04299867v1>

Submitted on 20 Apr 2024

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.

La *data literacy* distribuée

Périmètres définitionnels, origines documentaires, et perspectives réticulaires

UGO VERDI, OLIVIER LE DEUFF

Résumé :

Le but de cet article est d'interroger le concept de *data literacy* et ses fondements. La littératie des données est présentée à la fois comme un set de compétences contribuant à une formation numérique du citoyen, comme un droit inaliénable, et comme un atout majeur pour la pleine réalisation d'une société connectée. Cette résurgence de discours, que l'on retrouve dans la littérature scientifique et professionnelle, nous pousse à réinterroger les visions promues autour de la *data literacy*, les enjeux et usages actuels qui lui sont associés, ainsi que ses impacts sociaux et politiques. Pour ce faire, dans une première partie, nous resituons d'une part la *data literacy* dans son contexte actuel ainsi que dans le cadre épistémologique de la *literacy* à laquelle elle est rattachée, et d'autre part nous présentons les différents référentiels de compétences appliquées. Dans une seconde partie, nous soulignons son ancrage dans les principes de la documentation en révélant l'héritage prégnant de Paul Otlet. Dans une dernière partie, nous proposons une analyse prospective d'une *data literacy* distribuée s'appuyant sur les qualités des réseaux distribués et promouvant une réappropriation des données par les citoyens.

1. Introduction

La question « littéracique » tend à surgir lorsqu'un manque apparaît au moment où des discours prometteurs et enchantés sont parvenus à occuper l'espace public. L'*information literacy* arrive à point nommé quand émergent les discours de la « société de l'information ». La *media literacy* prend son essor dans les enjeux d'une civilisation audiovisuelle décrite et annoncée par McLuhan et le McLuhanisme (McLuhan, 1964). La *data literacy* n'échappe pas à la règle et se développe à la suite des discours sur les *Big Data* (ou Mégadonnées) et l'*Open Data* (ou Données Ouvertes). Sous-jacente parmi les spécialistes des littératies liées au numérique (Bawden, 2001), son importance est surtout soulignée quand apparaissent deux besoins essentiels : celui de disposer de travailleurs qui permettent la réalisation des promesses des mégadonnées, des villes intelligentes et ouvertes, et celui de conférer suffisamment de compétences aux usagers et citoyens pour qu'ils puissent en tirer quelques bénéfiques, au moins personnels. Si la littératie s'accompagne fréquemment du vocabulaire de crise, Enright (Enright, 2013) n'hésite pas à constater qu'elle peut s'avérer relativement violente en considérant les humains comme faisant pleinement partie du capital d'une nouvelle société industrielle et en intégrant l'information et sa gestion dans une vision économique et politique de type néolibéral. Le déplacement du prisme de l'information à la donnée semble accroître encore ce positionnement notamment dans une approche qui privilégie une mise en calcul facilitée par une raison computationnelle (Bachimont, 2016). La *data literacy* apparaît alors comme la littératie du moment au sein d'une nouvelle phase des littératies informationnelles et numériques. Pourtant, ce discours de la nouveauté, en semblant répondre aux impératifs du moment, néglige des racines historiques qui sont celles des sciences de l'information et plus particulièrement celles d'une documentation pensée de manière « étendue » par Paul Otlet qui permet de placer la *data literacy* dans une tradition documentaire. L'inscription au sein de l'histoire des réseaux apparaît également opportune pour mieux saisir que les infrastructures de l'Internet, ses protocoles et le développement du web offrent des pistes pour penser la littératie des données d'une façon originale. Nous proposons ici de rattacher la construction et la définition de la *data literacy* en dehors des seuls discours orientés mégadonnées et Intelligence Artificielle pour la rapprocher des perspectives portées par les sciences de l'information et de la communication, et plus particulièrement les perspectives des didactiques informationnelles et numériques de façon à prendre en compte les données dans un environnement plus complexe qui mêle dispositifs, processus, traitements et usages.

L'étude des littératies informationnelles, médiatiques et numériques tout comme les premiers travaux en ce qui concerne la *data literacy* (Claes et Philippette, 2020) mettent en avant deux positions : d'une part la position dite « éducative et citoyenne » qui cherche à développer les conditions d'une telle éducation et formation, et d'autre part la position *empowerment*, souvent présente dans les discours des institutions européennes, qui consiste à accompagner le volet industriel sous la forme d'un supplément éducatif, voire de kit d'auto-défense à destination du citoyen qui voit ses environnements personnels évoluer du fait de captation de données de plus en plus fréquentes. Un cadre juridique tend alors à accompagner les deux objectifs, comme c'est le cas en matière de réglementation des données personnelles, mais il aboutit à un paradoxe : celui d'accroître à nouveau les compétences en gestion de données tant pour les professionnels que pour les usagers qui se trouvent au quotidien forcés d'accepter et de gérer des conditions générales d'utilisation qui les lassent autant qu'elles les dépassent.

Alors que les conditions pour envisager des lettrés du numérique (Cormerais et Le Deuff, 2014) rien qu'au niveau de la recherche et de l'enseignement supérieur interrogent quant aux enjeux systémiques, aux protocoles à élaborer et aux infrastructures nécessaires pour parvenir au développement d'une *fair science*, les enjeux pour parvenir à envisager une *fair society of data* mettent en avant plusieurs objets et points de tension que nous souhaitons examiner dans cet article dont le but est de montrer qu'il est possible d'inscrire la *data literacy* dans des traditions documentaires et réticulaires plutôt que de céder aux impératifs de l'immédiateté et à l'urgence de besoins liés à la massification des données.

Nous observerons, dans un premier temps, les périmètres définitionnels de la *data literacy* à partir d'un corpus bibliographique constitué d'articles scientifiques et professionnels qui traitent du concept. Puis, nous aborderons dans une approche historique et épistémologique les enjeux info-documentaires, plus particulièrement ceux concernant les données, notamment en revenant sur les travaux Paul Otlet. L'objectif recherché est de resituer ces enjeux actuels, présentés comme des nouveautés, en les réexaminant à partir d'une perspective plus ancienne. Enfin, nous montrerons dans une approche

prospective, les enjeux autour d'une *data literacy* distribuée prenant appui sur les qualités des réseaux et sur l'importance d'une documentation comme technique de résistance.

2. Resituer la *data literacy*

2.1 Les sens de la *literacy*

Déceler les enjeux de la *data literacy* nécessite de revenir aux origines de la littératie dont le terme date de 1880, formé sur la notion de *literate* (Dictionnaire Merriam-Webster, 2020a). Son utilisation en 1883, au sein de l'article « The first state in the Union in literacy » du *New England Journal of Education*, servait à comparer les compétences en lecture des habitants d'un État à un autre en prenant en compte l'excellence du Massachusset en la matière (Wilson, 2008). La *literacy* s'opposait ainsi à l'*illiteracy* dont le premier usage remonte à 1660 et désigne l'incapacité de lire ou d'écrire (Dictionnaire Merriam-Webster, 2020b).

Literate est quant à lui défini comme la capacité de lire et écrire (Dictionnaire Merriam-Webster, 2020c) et est basé sur le latin « *litteratus/litteratus* », signifiant au 15^e siècle « éduqué, qui connaît les lettres » (Willinsky, 1990). Grundmann notait qu'avant le 12^e siècle, le terme « *illiteratus* » désignait celui qui ne savait ni lire, ni écrire et ne parlait pas latin, puis qu'à partir du 12^e siècle le terme a évolué pour désigner celui qui sait lire en vulgaire, mais qui ignore le latin (Grundmann, 1958).

Le concept de littératie a connu un regain de popularité grâce aux travaux de l'anthropologue Jack Goody, notamment avec la publication de l'article « The Consequences of Literacy » en 1969 où il analysait l'héritage culturel dans les sociétés sans écriture ainsi que les changements induits par l'écriture. Nous sommes « emboîtés dans les sociétés de l'écrit » : tout ce que nous faisons est basé sur l'écrit (Privat et Mohammed, 2012). L'écriture, en tant que « technologie de l'intellect », nous fait civilisés : son accès fait appel à des procédés cognitifs qui structurent la pensée (Bile, 2006). La littératie nous fournit ainsi un cadre intellectuel pour appréhender la situation de la culture de l'écrit dans la société et discuter les nécessités de relever les standards de lecture-écriture (Isnard, 2017). Elle « permet ainsi de repenser la problématique de l'alphabétisation, en tenant compte des facteurs épistémologiques, historiques, sociaux, politiques, idéologiques, linguistiques et psychologiques qui déterminent l'appropriation de l'écrit par les individus et les sociétés » (Pierre, 2003).

De plus, la littératie est chargée idéologiquement selon Brian Street : chaque version d'une pratique de littératie est construite selon des conditions sociales spécifiques et est en relation avec des structures politiques et économiques spécifiques (Street, 1985). Elle n'est pas un ensemble de compétences neutres, comme proposé au sein du courant d'une littératie dite « autonome » défendue par des auteurs comme David Olson et Angela Hildyard (Street, 1985), mais une pratique sociale en lien avec la connaissance qui comporte des valeurs relevant de la religion, du sexe, de la race, de la culture, des idéologies et du pouvoir (Street, 2001). Harvey Graff ajoute d'ailleurs que la littératie ne peut être comprise qu'au sein d'un contexte : elle ne peut être établie ni arbitrairement, ni uniformément pour l'ensemble de la population (Graff, 1979).

Sa traduction française est peu aisée. Kara soulignait que les traductions utilisées comme « lettrisme » (BO officiel de 2017) ou alphabétisation ne donnaient pas satisfaction (Kara, 2012). Le terme « littératie » que nous emploierons dans cet article nous semble le plus adéquat :

« Il nous permet en premier lieu de bien intégrer la question de l'alphabétisation. Mais, au-delà de l'alphabétisation, également la question de l'entrée dans la culture de l'écrit. » (Kara, 2012).

2.2 Les différentes littératies proches

Aujourd'hui la littératie ne s'applique plus seulement à son objet initial qu'est l'écriture : elle s'étend à de nombreux domaines parallèles tout en conservant l'idée originelle d'une maîtrise de lecture et d'écriture avec sens. Le nombre total de littératies ne peut être quantifié (Le Deuff, 2012). En rapport avec les questions numériques, nous pouvons toutefois citer l'*information literacy*, la *media literacy*, la *media and information literacy*, la *network literacy*, la *digital literacy* ou encore la *tag literacy*. L'ensemble des littératies informationnelles présentées ici ont pour objectif de faire des utilisateurs à la fois des producteurs et des

usagers au sein des domaines précisés. Les concepts de *transliteracy* (Thomas, 2005) ou de *metaliteracy* (MacKey et Jacobson, 2010), voire de littératie multimodale (Lim, 2011), définis plus tard, servent de liant entre ces différentes littératies par l'incorporation d'une variété de supports (médias, Internet, etc.).

2.3 La data literacy et ses particularités

2.3.1 La vision internationale

La *data literacy* quant à elle a été mise en lumière au cours des années 2000, tout particulièrement en 2004 avec la publication de l'article « Information Literacy, Statistical Literacy and Data Literacy » (Schield, 2004). Afin d'obtenir un aperçu global des références les plus citées de la *data literacy* à la rédaction de cet article, nous avons réalisé une exportation avec le logiciel *Publish or Perish* (CIRAD, 2020) à partir des données de *Google Scholar* et avons pu établir le tableau suivant (tableau 1).

Tableau 1. Articles les plus cités autour de la data literacy (export réalisé le 27 octobre 2020)

| Titre de l'article | Auteurs | Année de publication | Nombre de citations | Rang Google Scholar | Lieu de publication |
|--|-------------------------------------|----------------------|---------------------|---------------------|---|
| Information literacy, statistical literacy, data literacy | Schield | 2004 | 196 | 1 | International Association for Social Science Information Service and Technology |
| A systemic view of implementing data literacy in educator preparation | Mandinach, Gummer | 2013 | 385 | 2 | Educational Researcher |
| Building a Conceptual Framework for Data Literacy | Mandinach, Gummer | 2015 | 138 | 3 | Teachers College Record |
| Incorporating data literacy into information literacy programs: Core competencies and contents | Marzal, Prado | 2013 | 204 | 4 | Libri |
| Data literacy for researchers and data librarians | Koltay | 2017 | 86 | 5 | Journal of Librarianship and Information Science |
| Research data literacy | Schneider | 2013 | 75 | 6 | European Conference on Information Literacy |
| Data literacy: in search of a name and identity | Koltay | 2015 | 109 | 7 | Journal of Documentation |
| Data literacy is statistical literacy | Gould | 2017 | 44 | 8 | Statistics Education Research Journal |
| Lessons learned from a two-year experience in science data literacy education | Qin, D'Ignazio | 2010 | 91 | 9 | International Association of Scientific and Technological University Libraries |
| Data carpentry: workshops to increase data literacy for researchers | Teal, Cranston, Lapp, White, Wilson | 2015 | 71 | 10 | International Journal of data curation |

Une première lecture nous informe que les articles les plus cités sur la *data literacy* proviennent de différents domaines : les sciences des bibliothèques et de la documentation, l'*information literacy*, les sciences

de l'éducation ainsi que les domaines scientifiques qui s'intéressent à la question des données. Les contenus se scindent entre des définitions de la *data literacy* et/ou la présentation de divers projets impliquant son apprentissage et sa mise en application. On peut noter que définir la *data literacy* n'est pas une tâche achevée : les articles les plus cités continuent de proposer des définitions jusqu'en 2017. Le lien avec l'information y est central : en effet, la *data literacy* doit permettre de pouvoir extraire de l'information à partir de sets de données; la présence de l'*information literacy* (IL), la capacité de rechercher, évaluer et utiliser efficacement et de manière critique une information (ALA, 2020), souligne un lien de type méthodologique avec la *data literacy* et consolide, selon ces auteurs, le rôle central de l'information au sein de ces littératies.

2.3.2 La vision française

Sa traduction en France est multiple : on parle tantôt de culture des données (Lehmans, 2017), de data littératie (Arruabarrena, Chartron *et al.*, 2019), de littératie des données (Bodin, 2018 ; Lehmans, 2018) ou encore de Datalphabétisation (Fiaux, 2020). Nous avons renouvelé une recherche d'auteurs francophones datée du 24 décembre 2020 dans Isidore et Cairn (tableau 2) à partir des traductions françaises. Elle a permis de noter la récurrence de plusieurs auteurs francophones (Arruabarrena, Bodin, Chartron, Fiaux, Kembellec, Legoux, Lehmans, etc.) et d'autre part du terme « culture des données ».

Tableau 2. Occurrences des termes autour de la *data literacy* dans Isidore et Cairn (réalisé le 24 décembre 2020)

| Termes/Moteur de recherche | Isidore | Cairn |
|----------------------------|-------------------|-------------------|
| Data literacy | 163 occurrences | 12 occurrences |
| Culture des données | 90 occurrences | 28 occurrences |
| Littératie des données | 5 occurrences | 3 occurrences |
| Data littératie | 2 occurrences | aucune occurrence |
| Datalphabétisation | aucune occurrence | 2 occurrences |

En accord avec ce qui a été exposé plus haut vis-à-vis de la littératie, nous emploierons le terme de littératie des données pour parler de *data literacy*.

2.3.3 Les caractéristiques de la littératie des données

À l'intersection de l'*information literacy*, de la *statistical literacy* et des compétences techniques (Nayek, 2015), voire de la *numerical literacy* et de l'*PIT literacy* (Frank, Walter *et al.*, 2015), la littératie des données intègre la capacité de comprendre, d'utiliser efficacement et de manière critique la donnée (Schield, 2004), permettant à son utilisateur de la transformer en information afin de prendre des décisions éclairées (Koltay, 2017 ; Mandinach *et al.*, 2008). Dans cette même lignée, l'*Oceans of Data Institute* (ODI) en 2015 proposait de définir une personne *data literate* comme étant capable « de comprendre, expliquer et documenter l'utilité et les limites de la donnée en devenant un consommateur de donnée critique, capable de gérer ses traces numériques, de trouver du sens dans la donnée, et de prendre des décisions basées sur la donnée » (ODI, 2015). La littératie des données est ainsi nécessaire pour déterminer les données pertinentes à collecter, de percevoir les biais dans un set de données, de réaliser des analyses sensibles et de visualiser parfaitement les résultats (Kayser-Bril, 2016).

Avant de poursuivre, il semble important ici de définir le socle sur lequel s'appuie la littératie des données, à savoir la donnée. L'usage remonte au XIII^e siècle et désignait une « distribution, aumône ». Elle est définie aujourd'hui comme « représentation d'une information sous une forme conventionnelle adaptée à son exploitation ». (Académie Française, 2020). En se basant sur ce dernier point, nous retrouvons le lien évoqué ci-dessus entre la littératie des données et l'information. Les liens entre données et information ont fait l'objet de plusieurs hypothèses dont l'une des plus connues est la pyramide *Data-Information-Knowledge-Wisdom* (DIKW) qui contextualise et hiérarchise ces éléments tout en décrivant les

processus permettant la transformation de ceux-ci vers le niveau suivant (la donnée vers l'information, l'information vers la connaissance, la connaissance vers la sagesse). Si plusieurs versions de cette dernière existent (Ackoff, 1989; Rowley, 2007; Zins, 2007; Baskarada et Koronios, 2013), toutes situent la donnée à la base de la pyramide. Pour (Ackoff, 1989) elles sont des symboles qui représentent les propriétés d'un objet, d'un événement et de son environnement. Elles sont le produit de l'observation. A partir d'elles l'information est extraite, permettant de répondre aux questions primaires : qui, quoi, où et quand (Rowley, 2007). En ce sens, la littératie des données correspond à la réalisation de cette première étape et se pose comme un set de compétences permettant la maîtrise des données auquel s'ajoute une dimension politique, à savoir l'identification et la réponse critiques aux problèmes inhérents aux données, en particulier dans le cadre des mégadonnées, comme le manque de transparence, l'opacité de l'extraction des données personnelles, la complexité technologique ainsi que le contrôle sociétal (d'Ignazio et Bhargava, 2015). Etre *data literate* nécessite de plus de connaître la réalité du monde embarquée dans les données, notamment dans les formats et les langages qui les structurent et favorisent des formes et des usages spécifiques. C'est avoir conscience que la donnée brute, au sens de dénuée d'une vision ou d'une intention découlant d'un héritage culturel spécifique, n'existe pas : toute production de données provenant de processus réfléchis suivant des choix précis de logiciels (Barrowman, 2018; Iwasiński, 2020).

En cela, cela fait écho à la *critical literacy* (ou littératie critique) développée par Paulo Freire à partir de son ouvrage « *Pedagogy of the Oppressed* » en 1968 et qui se base sur la volonté de créer des lecteurs conscients que les textes positionnent les personnes de certaines façons et servent des intérêts bien précis (Lewison, 2002) : en cernant cette vision du monde véhiculée par le texte, ils se demandent surtout si cette vision est acceptable ou non et pourquoi (Gouvernement de l'Ontario, 2009). C'est, au-delà de l'esprit critique, une émancipation et un engagement du citoyen pour l'amélioration de la société. La littératie des données suit cette ligne directrice en impliquant le désir et la capacité de s'engager de manière constructive dans la société à travers ou à propos des données (Bhargava, Deahl *et al.*, 2015; Bodin, 2018). Certains auteurs avancent l'idée du concept de *critical data literacy* (Tygel et Kirsch, 2016; Špiranec, Kos *et al.*, 2019) pour intégrer cette dimension politique bien qu'à notre sens elle soit déjà présente dans la littératie des données.

C'est donc un certain « état d'esprit de la donnée » pour reprendre le mot d'(Herzog, 2015) qu'il convient d'atteindre par l'apprentissage de diverses compétences. Ainsi, outre la question de la nomination et celle des contenus qui accompagnent les discours sur les littératies, la question des compétences revient très rapidement, même si bien souvent il demeure des confusions entre compétences procédurales et des compétences plus complexes qui mêlent savoirs et savoir-faire voire savoir-être.

2.3.4 Les référentiels de compétences

La littératie des données est ainsi principalement étudiée sous l'angle des compétences que les chercheurs et les étudiants se doivent de maîtriser pour employer les données (Frank, Walter *et al.*, 2015). Dans un contexte de mégadonnées, ce besoin de maîtrise s'est considérablement accru. En effet, la multiplication exponentielle des données (33 zettaoctets en 2018) (Statista, 2019), des outils pour les manipuler ainsi que leur mise à disposition par les institutions publiques (Loi Lemaire de 2016 en France) ont durablement transformé notre rapport à celles-ci. Des services entiers en dépendent et impactent le quotidien des usagers : créer des contenus en ligne, laisser des traces digitales, user d'informations biométriques (dans le cas d'un passeport notamment), utiliser sa carte bleue ou son GPS... tout cela laisse des traces numériques qui sont ensuite réemployées par les différentes entreprises pour proposer des produits adaptés à son consommateur et/ou réaliser des statistiques. Viktor Mayer-Schönberger et Kenneth Cukier qualifiaient ce phénomène de « datafication », à savoir « la transformation de l'action sociale en données quantifiées en ligne permettant le traçage en temps réel et l'analyse prédictive » (Mayer-Schönberger et Cukier, 2013). En d'autres termes, tracer les activités des usagers et les transformer en données. Or, ces besoins étaient difficilement comblés et nécessitaient des mesures urgentes ; Carlie Idoine, analyste de l'entreprise Gartner, estimait déjà en 2018 que 80% des entreprises entreprendraient justement des actions pour former leurs personnels à la littératie des données (Pettey, 2018). Au-delà des entreprises, un intérêt particulier pour la littératie des données se remarque particulièrement dans le secteur des bibliothèques universitaires et de l'administration de la recherche (Koltay, 2015) où il s'agit de mettre en avant de nouvelles compétences notamment dans les profils de poste où les besoins font apparaître la nécessité de recruter des « data librarians » (Thiault, 2020).

C'est ainsi que, pour pallier ces différentes avaries, de nombreux métiers liés directement à la donnée se sont développés (Decideo en dénombrerait au moins 70 en 2020) (Nieuwbourg, 2020) et que plusieurs cadres de référence ont été conçus. Au Québec, le *Digital Competency Framework* dévoilé en 2019 (Roberge, 2019) a pour objectif de fournir un socle au développement de la « compétence numérique » (*digital competency*), définie comme un éventail de compétences permettant la maîtrise des technologies numériques. Divisé en douze dimensions situées sur un même continuum, il aspire à rendre le citoyen totalement autonome à mesure qu'il progresse. La compétence est ici comprise au sens de Jacques Tardif, à savoir « un savoir-agir complexe prenant appui sur la mobilisation et la combinaison efficaces d'une variété de ressources internes et externes à l'intérieur d'une famille de situations » (Tardif, 2007).

L'Union européenne, de son côté, a produit le Dig'comp, le cadre européen des compétences numériques en 2013. La compétence numérique, dès 2006, était reconnue comme l'une des huit compétences de la formation tout au long de la vie (Commission Européenne, 2006). La participation des citoyens envers la société nécessitait des compétences relatives aux technologies (Ferrari, 2013). Pour y répondre, plusieurs versions du Dig'comp ont été produites. Dans sa dernière version (2.1), son premier domaine est entièrement centré autour des littératies de l'information et des données : il comprend cinq domaines de compétences divisés en vingt-et-une compétences et huit niveaux de maîtrise. Y sont recherchés la maîtrise de l'information et des contenus digitaux (en termes de recherche, de filtre, d'évaluation, de gestion et de création), la compréhension des outils, des lois et le respect des règles en ligne (comme la Netiquette, le copyright, etc.), savoir protéger son environnement digital (ce qui inclut les supports) et enfin être capable de résoudre les problèmes techniques, digitaux et technologiques (Punie, 2017). Les stratégies européennes actuelles pour la période 2019 - 2024 doivent conduire « la transition vers une planète saine et un nouveau monde numérique » dont le premier objectif « la technologie au service des personnes » lie compétitivité économique et information du citoyen capable de prendre des décisions éclairées (Commission Européenne, 2020a). Avec la crise liée à la Covid-19 et le changement opéré dans l'utilisation de la technologie au travail du fait de la distanciation sociale, la Commission européenne a lancé au mois de juin 2020 un appel pour une modification du Dig'comp afin de l'adapter aux enjeux numériques actuels (Commission Européenne, 2020b).

En France, le Cadre de référence des compétences numériques de 2019 découle directement du Dig'comp européen. Son application concrète se traduit à travers le certificat numérique PIX (République française, 2020) qui remplace les anciens B2i et C2i dispensés au sein de l'Éducation Nationale et dans l'enseignement supérieur. Dès 2016, il avait pour objectif « de préparer la transformation digitale de l'ensemble de notre société et de notre économie » (Tice-Education, 2016). Il comporte cinq domaines divisés en seize compétences et cinq niveaux de maîtrise. Ces cinq domaines sont Information et données, Communication et collaboration, Création de contenu, Protection et sécurité, Environnement numérique.

Tableau 3. Comparaisons des différents référentiels de compétences numériques

| | Cadre de référence de la compétence numérique (Québec) | Dig'com 2.1 (Europe) | PIX (France) |
|---|---|---|--|
| Date de mise en œuvre | 2019 | 2017/2018 | 2019 |
| Domaines/dimensions | 12 | 5 | 5 |
| Nombre de compétences | 53 | 21 | 16 |
| Niveaux de maîtrise | - | 8 | 5 |
| Inclusion explicite de l' <i>Information Literacy</i> | Oui | Oui | Non |
| Objectifs visés | « Appréhender le numérique en tant qu'outil d'inclusion » | « Améliorer les compétences numériques des citoyens » | « Amener chacun d'entre nous à cultiver ses compétences numériques tout au long de sa vie. » |

Aux États-Unis, si un référentiel de compétences appliqué à l'ensemble des citoyens n'existe à pas à notre connaissance, il existe toutefois le « Federal Data Strategy Principles » (FDS, 2020) dont le plan d'action de 2020, de nature transversale, se pose en soutien aux agences gouvernementales afin de remplir un large éventail d'exigences législatives et administratives. La donnée est vue à la fois comme un levier central pour la croissance économique, rendre le gouvernement fédéral plus efficace, améliorer l'engagement civique et promouvoir la transparence. Il se compose de dix principes (à respecter en permanence) et quarante pratiques (des approfondissements des principes précédents sur une période comprise entre cinq et dix ans). Vingt étapes divisées en trois parties que sont les *Agency actions*, les *Community of Practice Action* et les *Shared Solution Actions*, devront être entreprises par les agences en 2020. Ces dernières étapes visent la maîtrise de la donnée par une amélioration des standards employés et la mise en place de projets spécifiques.

D'un point de vue entrepreneurial, la littératie des données est brandie comme un emblème dans les projets des grandes entreprises soumises à une concurrence internationale. Elles s'en approprient les définitions et les usages (Plassart, 2020 ; Soubiran, 2020). En particulier l'entreprise Qlik qui a développé en 2019 une certification Qlik en littératie des données (Qlik, 2020). Cette certification fait l'objet d'un examen divisé en sept parties : interpréter les exigences métiers, comprendre et transformer les données, concevoir et créer des visualisations, interpréter les visualisations, analyser les résultats, prendre des mesures concrètes en fonction des résultats, partager les résultats. L'outil *mydataabilities*¹ de l'entreprise australienne Data to the People propose quant à lui une auto-évaluation en ligne de ses propres compétences en littératie des données, rebaptisées *databilities* (Data to the People, 2020). Il propose quinze compétences réparties sur six niveaux et abordent les diverses dimensions de la lecture, l'écriture et la compréhension.

Ces visions de la littératie des données sont purement techniques, opérationnelles : elles ont pour objectif de former des techniciens spécialisés dans la gestion des problématiques digitales dans un contexte de concurrence autour de la donnée. Il s'agit de rendre accessible la donnée, de la manipuler et de la transmettre le plus efficacement possible. On retrouve finalement une logique présente depuis longtemps dans la documentation avec le fait de privilégier le traitement de « l'information utile ». Toutefois, le volet documentaire a toujours été accompagné par une dimension analytique.

2.4 Une tradition documentaire

Les formations universitaires se positionnent également en tant qu'acteurs de cette culture des données, en particulier le Master « Mégadonnées et Analyse sociale » (MÉDAS) créé au CNAM en 2017 qui propose de former les étudiants au métier de *data analyst*. Arruabarrena, Chartron ainsi que Kembellec soulignent que les compétences techniques, informatiques et statistiques constitutives de la « Data littératie » (leur traduction) constituent une « dimension importante des programmes de formation des data analystes » (Arruabarrena, Chartron *et al.*, 2019). Ainsi, similairement aux référentiels présentés précédemment, les compétences à acquérir ici sont centrées sur le traitement et l'analyse des données quantitatives et qualitatives (base de données, statistiques...) par, notamment, la mise en place de solutions vis-à-vis des fouilles de données, des pilotages de projets de données et de développement de stratégies numériques (CNAM, 2020). Mais le choix est délibérément celui de privilégier une position de médiation vis-à-vis des données en mettant l'accent sur l'analyse plutôt que sur le métier fantasmé de *data scientist*. Ce positionnement qui privilégie la médiation autour des données se retrouve également dans les licences professionnelles MIND (Médiations de l'information numérique et des Données) de l'IUT Bordeaux Montaigne et Data et Information de l'IUT de Tours. Ces deux licences sont également issues et portées par des formations initialement documentaires, notamment les DUT information numériques dans les organisations, historiquement créés sous l'appellation « documentation d'entreprise » dans les années 1960. Du côté du master MÉDAS, l'origine documentaire est également présente puisque la formation est portée par l'Institut national des techniques de la documentation (INTD) qui a été notamment créé par Suzanne Briet, célèbre pour son œuvre théorique sur la documentation et ses apports pédagogiques et professionnels notamment en faveur de l'analyse documentaire (Briet, 1951).

Cet intérêt et ces positionnements en ce qui concerne les aspects de formation et de médiation autour des données de la part de formations initialement documentaires ne sont en fait guère surprenants car

¹. <https://www.mydataabilities.com/>

l'intérêt pour la prise en compte des données, leur traitement et leur analyse fait pleinement partie de l'œuvre du « père de la documentation » Paul Otlet dans laquelle on peut trouver aussi bien les aspects pratiques et pragmatiques des traitements potentiels des données, que des projections avant-gardistes que certains ont pu qualifier d'utopistes.

3. Otlet et l'utopie documentaire de l'accessibilité à la plus petite unité informationnelle

Le concept de littératie des données poursuit la logique du passage d'une bibliothéconomie classique, qui demeure dans les bases de la conservation, à une documentation qui se projette vers l'avenir² (Otlet, 1934). Mark Balnaves et Michele Wilson (Balnaves et Wilson, 2011) considèrent qu'il existe deux logiques dans la recherche d'information : l'une plus traditionnelle s'inscrit dans la démarche du bibliothécaire américain Charles Cutter et vise la retrouvabilité du document au sein d'une bibliothèque. L'autre, celle de Paul Otlet, s'appuie sur le besoin d'information et le fait de privilégier l'accès à une information plutôt qu'à un document. Selon Balnaves et Wilson, le second positionnement est fondateur des sciences de l'information.

Ronald Day poursuit cette analyse en montrant que l'enjeu chez Otlet est de chercher à dépasser les contenants pour s'intéresser plus finement aux contenus et à son rapport à l'utilisateur (Day, 2014). L'accessibilité se déplace de l'accès au document à l'accès à l'information, c'est-à-dire à la donnée précise qui peut répondre à un besoin d'information spécifique. Il s'opère comme un changement de focale progressif qui consiste à se rapprocher de l'unité informationnelle la plus précise afin de pouvoir répondre à une requête. Cette accessibilité repose sur deux évolutions majeures qui ont été mises en œuvre dans les systèmes d'information informatisés et donc « post-otletiens » :

- le besoin de disposer de documents de plus en plus structurés et donc sémantisés pour saisir pleinement ce que signifient les éléments informationnels qu'ils contiennent. Cette logique est celle préconisée par les travaux de langage à balises depuis le *SGML* jusqu'aux projets de Web sémantique désormais regroupés sous le vocable de *Linked open data*. Cette entreprise de sémantisation concerne surtout des collections entretenues, telles que les catalogues de bibliothèques ou de musées, ainsi que l'information scientifique et technique ;
- le besoin d'articuler plus finement les besoins d'information et les formes documentaires interrogées qui a abouti à une logique consistant autant à comprendre l'utilisateur qu'à indexer les documents. Les stratégies de personnalisation de l'information reposent sur le développement de profils utilisateurs et la volonté des plateformes et des moteurs de recherche de faire évoluer leurs algorithmes en fonction de l'utilisateur, si bien que le principe d'une recherche universelle tend peu à peu à s'effacer au profit d'une relation individualisée entre les ressources et la personne qui effectue une requête.

Ces deux évolutions soulignent l'importance des métadonnées et le fait qu'elles ont permis, suite aux développements bibliographiques, l'apparition de nouvelles formes de catalogues, notamment informatiques, d'une part, et de déplacer les enjeux bibliothéconomiques initialement centrés sur le document vers les besoins d'information des usagers d'autre part. Cela a donc impliqué un fort développement des classifications et des langages documentaires : les métadonnées sont devenues centrales, ce que rappelle à juste titre Michael Buckland (Buckland, 2017) qui montre que cela entraîne des changements d'infrastructures. Cette évolution touche également les moteurs de recherche dans la mesure où les métadonnées vont être de plus en plus associées aux profils des usagers plutôt qu'aux seuls documents indexés.

Au projet de la documentation d'Otlet, qui avait pour but d'établir et d'assurer la qualité de l'information dans un positionnement à l'égard de la vérité qui pourrait être qualifié de positiviste, se voit opposée une accessibilité de type individuelle qui pourrait au final correspondre à un besoin d'information non seulement guidé par des besoins individuels tout à fait louables, mais également déformés par des opinions personnelles peu compatibles avec la réalité des faits.

². Otlet évoque le fait qu'« il faut des procédés nouveaux, très distincts de ceux de l'ancienne bibliothéconomie, tels qu'ils sont appliqués. » (Otlet 1934, 2).

Pour comprendre cette potentielle opposition, il faut d'abord revenir sur l'épistémologie de l'information et du document en considérant les théories de Paul Otlet (Rayward, 1975 ; Van Acker, 2011).

3.1 Une utopie à la Otlet

La littératie des données correspond finalement au désir premier de la documentation, à la volonté de penser l'accessibilité de la manière la plus poussée possible au niveau de son plus petit élément infime : le biblion.

« Le « Biblion » sera pour nous l'unité intellectuelle et abstraite, mais que l'on peut retrouver concrètement et réellement, mais revêtu de modalités diverses. Le biblion est conçu à la manière de l'atome (ion) en physique, de la cellule en biologie, de l'esprit en psychologie, de l'agrégation humaine (le socion) en sociologie. L'atome a donné lieu à une représentation de plus en plus précise et sur la base de laquelle se sont engagées toutes les recherches et discussions » (Otlet, 1934, 43).

Le concept est initialement conçu comme l'équivalent du document (Otlet, 1934, 9 et 122), mais Otlet lui confère une portée différente en s'interrogeant notamment sur son rapport à la réalité :

« Par le biblion, représentation plus ou moins adéquate de la réalité, c'est au-delà du langage que se pose le problème du nominalisme et du réalisme » (Otlet, 1934, 425).

Pour mieux saisir le biblion en tant qu'unité informationnelle, il convient de le rapporter au principe monographique qui est notamment développé au sein du projet bibliographique « Die Brücke » (Hapke, 1999) et qu'Otlet reprend d'une façon simple et explicite : une idée, une fiche.

À cette nécessité documentaire de chercher à représenter la réalité en s'appuyant sur les données, les enjeux informationnels décrits par Otlet s'articulent alors entre les différentes formes documentaires existantes : celles issues de l'audiovisuel qu'Otlet décrit, et celles du futur qu'Otlet imagine autour de son concept d'hyperdocumentation (Le Deuff et Perret, 2019). Le mot « données » apparaît d'ailleurs plus de deux cent cinquante fois dans le traité de documentation dans des sens qui diffèrent légèrement selon que le mot désigne les données du processus scientifique, ou les données disponibles dont il faut parvenir à tirer du sens avec des outils nouveaux. Au début de l'ouvrage, les données participent des éléments qui doivent être mis à jour, ajoutés et complétés au sein des documents. Chez Otlet, il n'y a pas de pensée documentaire sans les données, tout comme il n'y a pas de science des documents, et donc de documentologie, sans études et production de données (Otlet, 1934, 9).

Il s'agit donc :

- premièrement, de dépasser la pensée bibliothéconomique de base, dont l'objectif était de parvenir à resituer un document et le consulter, pour repenser la complexité informationnelle au travers des formes documentaires renouvelées, en perpétuelle évolution ;
- deuxièmement de penser les formes étendues, c'est-à-dire hyperdocumentaires, au sens des produits issus de traitements documentaires de plus en plus sophistiqués et massifs (Figure 1) ;
- troisièmement de penser les formes documentaires qui permettent de passer du document au monument et donc de la fiche à la cité mondiale (Otlet, 1934, 425), ce qui ouvre la voie à des citées « connectées »³.

Chez Otlet, l'enjeu est de pouvoir faciliter la consultation, mais également de pouvoir plus aisément décider, notamment avec une remontée de données statistiques qui s'exercent depuis la plus petite comptabilité jusqu'à celles des États et des organismes internationaux (Otlet, 1926). Pour traiter les données éparses et pour parvenir à leur donner un sens cohérent, la classification apparaît alors comme

³. « La Cité Mondiale sera un Livre colossal, dont les édifices et leurs dispositions — et non seulement leur contenu —, se liront, à la manière dont les pierres des cathédrales se « lisaient » par le peuple au moyen âge. » (Otlet 1934, 425).

une réponse adéquate pour Otlet qui imagine un langage dédié : la Classification Décimale Universelle (CDU). Il convient de nuancer la vision classificatoire d’Otlet qui serait uniquement centrée sur la CDU. Elle est en fait à comprendre de manière extensible et notamment combinatoire. Elle ne peut être réduite à une simple organisation des connaissances arborescentes tout comme elle ne peut être réduite à une simple méthode pour ranger les documents. À plusieurs reprises, Otlet évoque d’ailleurs la nécessité de disposer de nouveaux outils spécifiques pour tirer tous les bénéfices des données disponibles⁴.

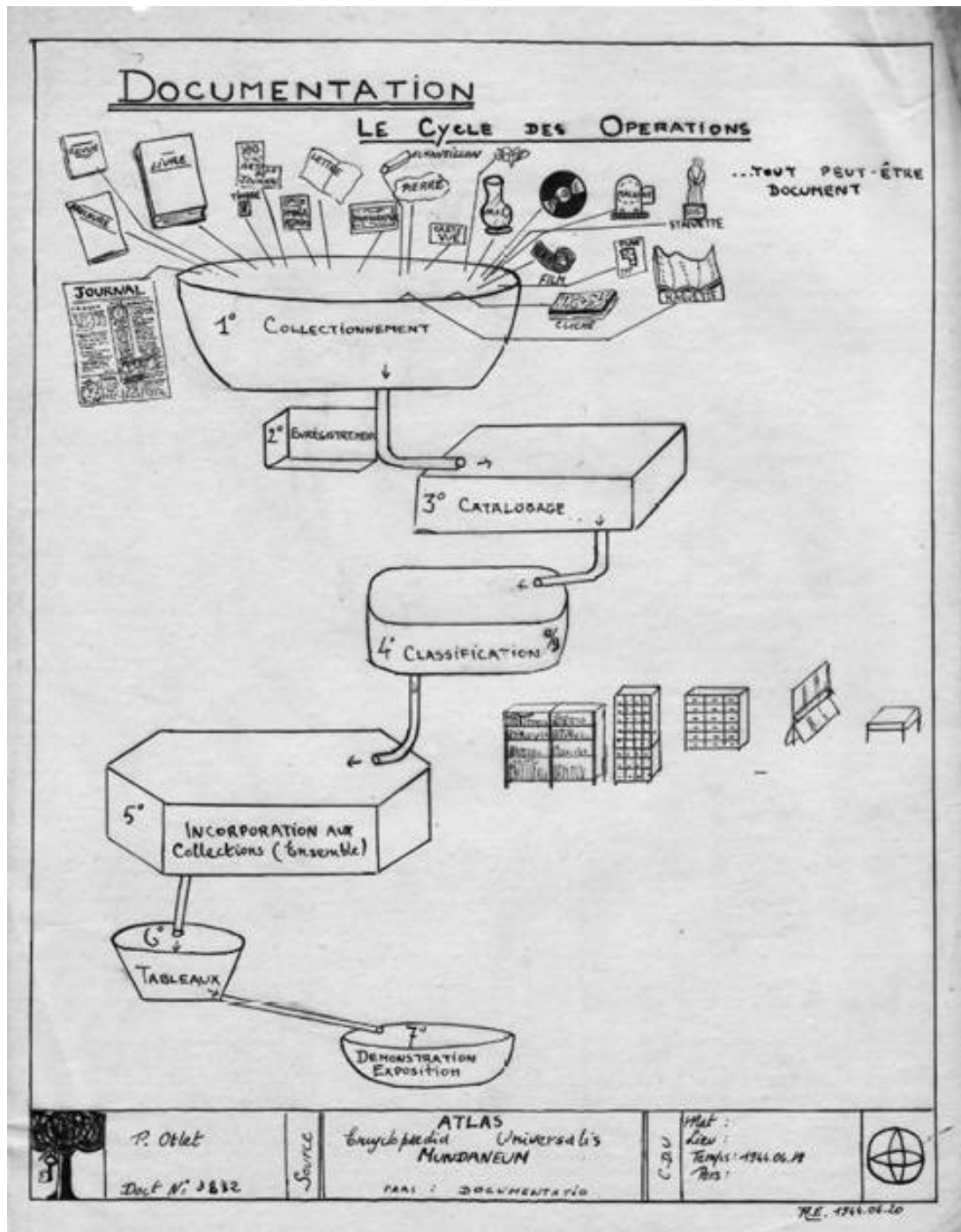


Figure 1. Le cycle des opérations selon Paul Otlet ⁵

⁴. « L’Humanité est à un tournant de son histoire. La masse des données acquises est formidable. Il faut de nouveaux instruments pour les simplifier, les condenser ou jamais l’intelligence ne saura ni surmonter les difficultés qui l’accablent, ni réaliser les progrès qu’elle entrevoit et auxquels elle aspire. » (Otlet, 1934, 430).

3.2 Une logique documentaire : formes et formats de la donnée

L'apport d'Otlet réside également dans la médiation opérée par rapport aux données. Elles ne sont pas essentialisées comme existantes en soi, mais issues d'un traitement documentaire, mieux encore elles ne se valent pas forcément :

« La forme du livre est distincte de sa substance, les données qu'il contient sont relativement indépendantes. Des données de différentes sources (différents auteurs, différents pays) peuvent être comparables au point de vue de la forme, car elles se rapportent à un même objet, au même temps, et parce qu'elles sont exprimées de la même manière. Malgré cela, ces différentes données peuvent fort bien n'être pas comparables, relativement au fond, certaines étant le fruit d'une observation consciencieuse, d'un raisonnement logique, et d'autres au contraire résultant de la fantaisie et de l'invention de toutes pièces. Ce serait commettre une erreur de les amalgamer, comparer, additionner. » (Otlet, 1934, 94)

Le rapport entre le fonds et la forme est étudié par Otlet notamment quand il s'agit de penser le partage des données et leur normalisation à l'international :

« En combinant ces différentes données, on commet de nouveau une erreur et même dans certains cas la diversité de forme est si apparente, si vivante, qu'il devient absurde de vouloir tenter un rapprochement, grouper le tout en une même colonne, un même tableau. On voit donc que les exigences de forme et de fond sont différentes et peuvent être étudiées séparément. Les exigences de forme sous un certain aspect sont même plus essentielles que les autres chaque fois qu'il s'agit de coordonner des travaux très étendus comme le sont les travaux internationaux et ceux qui portent simultanément sur les domaines de plusieurs sciences ou branches d'activité. » (Otlet, 1934, 94)

Cette logique combinatoire s'exprime également dans les possibilités de lier les données dans divers ensembles et ouvrages :

« Le troisième ordre de degré se rapporte à la complexité des données : a) manière d'incorporer dans une rédaction un fait ou une idée simple ; b) manière de combiner un nombre de données dans un ensemble : un ouvrage ; c) manière de combiner dans un ensemble divers ouvrages ; d) manière de concevoir la combinaison les uns avec les autres de l'ensemble des ouvrages. » (Otlet, 1934, 99)

Chez Otlet, ces possibilités de traitement de données participent de progrès en matière scientifique et technique et plus particulièrement en ce qui concerne le besoin de disposer de professionnels de la documentation capable d'organiser et de traiter ces masses d'informations disponibles, ce qui implique la nécessité de normalisation et de collaboration internationale pour parvenir à un travail coordonné. Le travail d'Otlet et de La Fontaine a pleinement été celui de relier les plus petites unités informationnelles souvent contenues sur des fiches à des projets de plus grande ampleur (Société des Nations, Cité Mondiale, Organisation internationale du travail intellectuel, etc.). Les expositions universelles ont constitué des lieux par excellence de cette démonstration de l'articulation entre l'infiniment petit informationnel à l'infiniment grand.

3.3 Une démonstration spectaculaire

La volonté de montrer et démontrer la puissance des données ainsi que leur potentialité n'est pas nouvelle. Les photos qui montrent Paul Otlet au sein du Mundaneum en train de manipuler des fiches issues du Répertoire Bibliographique Universel participent de cette logique spectaculaire qui consiste à montrer un monde ordonné par des techniques informationnelles avancées. Shannon Mattern (Mattern, 2020) décrit parfaitement les visions enchantées qui participent à ce monde de fiches qui précèdent nos discours actuels sur les données et les *data center* :

« Ces fiches étaient des outils permettant d'envisager de nouvelles infrastructures universelles de la connaissance, des villes efficaces, harmonieusement mises en réseau. Ces petites parties mobiles de l'information ont été transformées en de grands fantasmes de mondes futurs. Ils faisaient partie des technologies culturelles et projets techniques qui ont rendu possible le panorama des Weltprojekte, des projets mondiaux, qui ont surgi autour du tour du XXe siècle » (Mattern, 2020).

⁵. Archives du Mundaneum.

Que ce soit à partir de l'unité « fiche » ou de l'unité « donnée », on retrouve cette même projection de l'infiniment petit informationnel vers des infrastructures plus grandes. Ces dernières laissent place à des imaginaires qui vont des cités mondiales et des organisations politiques à l'échelon continental ou mondial jusqu'à des systèmes de gestion optimisés dont l'objectif est de gérer un maximum d'informations dans des temporalités restreintes avec des effets démonstratifs autant qu'admiratifs. Shannon Mattern montre d'ailleurs que durant les expositions (expositions universelles ou expositions sur les technologies et villes du futur tout au long du vingtième siècle) se trouvent toujours présentes ces unités informationnelles :

« Et si, dans chacune de ces installations, les petites pièces mobiles d'informations - cartes triées, fichiers organisés, données traitées - sont toujours présentes (...), ici ces cartes se transforment en spectacle et assument des rôles distincts en tant que technologies culturelles, en indexant leurs propres milieux historiques et culturels » (Mattern, 2020).

Cette démonstration réside d'ailleurs sur un paradoxe qui est celui d'un processus de réduction du monde en objets ou éléments interprétables et interrogeables, ce que Bruno Latour s'est attaché à montrer dans son analyse de la constitution de la science notamment autour de ses laboratoires, bibliothèques et autres milieux de savoir qu'il qualifie de « centres de calculs » (Latour, 1996). Plus encore, le côté spectaculaire des données est de désormais glorifier les contenants, les logiques d'interprétation, les visualisations... bref ce qui entoure les données plutôt que la démonstration des données elles-mêmes, comme si elles cassaient la logique du spectacle et la magie qui en résulte. Cette assimilation entre science et magie, déjà mise en avant par Latour dans *Irréductions* (Latour, 1984), rappelle finalement le besoin d'envisager des modes irréductibles comme réponses à ces processus qui dissimulent non seulement les données, mais aussi la manière dont elles sont captées, construites, « obtenues » dirait encore Bruno Latour. Cette discrétisation des données, de ses formes, de ses formats, des métadonnées englobantes, renvoie à des enjeux professionnels qui se résument trop souvent au territoire de l'informatique. Or les besoins sont souvent autres : en témoignent les propos du dirigeant de Palantir⁶, société controversée de par ses activités de traitement de données massives, qui peine parfois à recruter des professionnels pour intégrer et curer des données, car l'attractivité spectaculaire ne semble pas présente. Or, c'est justement dans cette partie, cette « boîte noire », que se jouent les plus grands défis en matière de formation et donc de littératie des données. L'enjeu d'une littératie des données s'inscrit donc autant dans ce qu'il est possible de voir que de ce qui est dissimulé dans les processus afin d'éviter le risque de l'essentialisation des données.

L'extraction ou la récupération de données personnelles s'avère un bon exemple pour comprendre les limites d'une focalisation sur les données récupérées. Nos données personnelles récupérées de comptes de type GAFAM nous donne qu'un aperçu faussé de la manière dont nous voient les plateformes : nous n'avons pas leurs instruments de mesure, leurs « lentilles ». Ainsi, au-delà des effets de démonstration, la littératie des données requiert la compréhension des réseaux et l'intégration de leurs « propriétés » en matière de distribution, ce qui rejoint la vision de Paul Otlet : « Le propre d'un réseau c'est de retrouver les éléments essentiels dans chacune de ses stations. » (Otlet, 1934, 278). Cette figure du réseau revient à plusieurs reprises dans les travaux de Paul Otlet notamment pour décliner son projet documentaire au bénéfice d'une accessibilité aux données optimisée :

« Il sera organisé un Réseau Universel mettant en rapport coopératif tous les organismes particuliers de documentation, tant publics que privés, à la fois pour la production et pour l'utilisation. Le Réseau, de quelque manière que ce soit, doit relier les uns aux autres les centres producteurs, distributeurs, utilisateurs, de toute spécialisation et de tout lieu. Il s'agit pratiquement que tout producteur ayant quelque donnée à faire connaître, quelque proposition à présenter ou à défendre — tout utilisateur ayant quelque information à recueillir pour avancer son travail théorique ou pratique — toute personne enfin puisse au moindre effort et avec un maximum de sûreté et d'abondance, entrer en possession de ce qui leur est offert. » (Otlet, 1934, 415)

⁶. Michael Steinberger. Does Palantir see too much ? *New York Times Magazine*, 21 octobre 2020. <https://www.nytimes.com/interactive/2020/10/21/magazine/palantir-alex-karp.html> L'article montre la difficulté à recruter des intégrateurs de données, car ce n'est pas jugé « sexy ». Le dirigeant de la succursale de Londres Luis Mosely le dit en des termes qui rappellent le schéma d'Otlet : « It's plumbing work, basically ».

4. Vers une littératie des données distribuée ?

Alors que la tentation est de planifier de façon arbitraire ou centralisée le développement de la littératie des données, la figure du réseau permet d'envisager le réseau de façon distribuée plutôt que de manière centralisée voire décentralisée. Les nouveaux modèles obligent à repenser la manière même dont la littératie des données peut se développer de manière réticulaire et a-centrée.

4.1 Repenser la littératie des données au sein de réseaux « distribués »

La création des réseaux Internet et du Web s'est faite dans des circonstances et temporalités qui sont celles de la remise en cause des processus centralisés et des modes de fonctionnement trop hiérarchisés comme l'exprimait Louise Merzeau :

« Dès l'origine, l'Internet s'est développé comme un hacking des espaces centraux de vérité (académiques, politiques, institutionnels...), selon un principe d'a-centrage et de dissémination des conversations. En même temps que le réseau augmentait les moyens des entreprises militaires et des laboratoires de recherche, il ouvrait un espace de partage de pair à pair non contrôlable. En actant le passage d'un graphe de documents à un graphe de profils, le Web dit « social » a radicalisé (et perversi) cette grammaire fiduciaire, pour la convertir en une industrie de l'engagement » (Merzeau, 2017).

Le principe de la distribution s'est concrétisé au travers de celui d'infrastructure technique de la transmission par paquets de par les réalisations du réseau Arpanet (Serres, 2000) ainsi qu'avec les principes de l'hypertexte appliqués au Web. Mais ce sont surtout les potentialités d'exploitation sociale et éducative des qualités du réseau qui ont été soulignées rapidement par les pionniers, notamment autour de communautés virtuelles (Rheingold, 1996), ainsi que par les travaux d'intelligence collective (Lévy, 1997). Or, ce niveau peine à pleinement se réaliser en dehors des interactions classiques sur les réseaux sociaux dont la gestion est assurée essentiellement par des plateformes tant les alternatives de type Mastodon restent encore marginales. Dès lors, il en résulte un déficit en matière de protocole d'interopérabilité ainsi qu'en matière de compétences dans la mesure où les facilités d'usage tendent à dissimuler les compétences techniques. Pourtant les enjeux de la littératie des données reviennent finalement à des principes d'analyse informationnelle avant d'en penser la stratégie d'exploitation communicationnelle. La littératie des données ne peut fonctionner sur des faux-semblants de résultats qui seraient recueillis suite à des requêtes approximatives. Elle requiert au contraire une plus grande précision dans les choix de requête, dans la sélection des sets de données potentiels et dans leur interprétation au sein de leur structuration documentaire. Elle requiert également une prise de conscience du processus d'intégration de ces données au sein même des organisations, ce qui suppose probablement de nouveaux métiers qui précèdent le rôle des *data scientists* avec des médiateurs de données comme des *data stewards* ou coordonnateur de données par exemple.

Alors que la logique des données ouvertes s'est basée sur une accessibilité technique des données, le volet médiation a été négligé et laissé finalement à des associations ou petites entreprises qui tentent de faire de la formation à ces nouvelles opportunités⁷. Le paradoxe est alors celui d'une sortie des données hors de leur cadre premier pour des réutilisations qui sont fort rares et réalisées par des personnes initiées. La valeur ajoutée et l'enjeu tendent donc à se déplacer *a priori* des besoins de rendre accessibles des données à celui de savoir comment et pourquoi les exploiter, ce qui oblige à repenser les conditions d'accessibilité technique, organisationnelle et matérielle avec les conditions d'accessibilité intellectuelle qui sont justement l'enjeu d'une littératie des données. Il convient de préciser que l'enjeu n'est pas uniquement de penser l'utilisation « à la sortie », mais d'en penser la mise en place au sein même des administrations qui ne parviennent pas à développer de médiations suffisantes pour situer et comprendre où sont les données, sous quelles formes documentaires elles circulent, notamment parce que les fonctionnaires et les employés ignorent qu'ils manipulent des données alors qu'ils ont plutôt conscience d'utiliser des formulaires et des tableaux.

⁷. Voir ce billet de Samuel Goëta qui resitue bien les enjeux de penser les usages et médiations autour de l'ouverture des données. « Ouvrir, ce n'est pas juste partager les données », Medium, Août 2018, <https://medium.com/dataactivist/ouvrir-ce-nest-pas-juste-partager-des-donn%C3%A9es-simple-basique-ec87d762f2a2>

Il en est de même pour les plateformes privées comme Google, Facebook et désormais Netflix qui ont facilité la récupération de nos données personnelles, mais dont on ne sait exactement que faire tant elles sont privées de leur contexte et peu aisées à utiliser et à comprendre pour le non-initié. L'enjeu de la distribution repose alors sur les possibilités de pouvoir accéder aux processus de classification des contenus et aux algorithmes qui privilégient la « mise en avant » de contenus qui reposent sur des logiques attentionnelles plutôt sur des logiques d'organisation des connaissances en tant que telles. Or, les littératies sont trop souvent pensées et définies dans un cadre classique de l'organisation des connaissances. Ted Striphas, en s'inscrivant dans la lignée de Raymond Williams, montre qu'il faut intégrer de nouveaux mots-clés dans cette construction culturelle qui est celle d'une culture algorithmique :

« Une liste étendue pourrait inclure analogique, application, cloud, code, contrôle, convergence, copie, données, conception, numérique, format, gratuit, ami, jeu, graphique, back, humain, identité, machine, message, mobile, réseau, bruit, pair, plateforme, protocole, recherche, sécurité, serveur, partage, social, statut, Web et bien d'autres. »⁸ (Striphas, 2015, 398)

La littératie des données apparaît ici fortement liée à la perspective d'une littératie algorithmique qui offre une meilleure compréhension des processus de sélection informationnelle ainsi que la possibilité de modifier lesdits processus qui s'organisent dans une forme de boîte noire (Rouvroy et Berns, 2013 ; Carmes et Noyer, 2014). Les travaux et méthodes en rétro-ingénierie (Diakopoulos, 2019) permettent de dévoiler les tendances fortes dans la gestion des algorithmes des plateformes mais il est difficile de pouvoir en saisir toutes les subtilités, car bien souvent la valeur ajoutée économique s'y trouve justement, ce qui incite les plateformes à rester discrètes à ce niveau. Plus encore, la mise en avant des processus entraîne alors des modifications algorithmiques, tout comme les usages jugés trop optimisés par Google en matière de référencement entraînent des modifications parfois radicales pour changer la règle du jeu. Or, le contexte dans lequel s'inscrit la littératie des données est celui d'une « donne » qui n'est pas connue d'avance et dans laquelle l'utilisateur se voit inciter à participer sans en connaître toutes les règles, encore moins les gains et pertes potentiels.

Dans ce cadre, il apparaît difficile de pleinement séparer contenu et contenant, médium et données. Si on prend l'exemple des moteurs de recherche, la question informationnelle s'est déplacée de la bonne méthode pour rechercher l'information à la nécessité de savoir évaluer l'information et donc de comprendre comment les outils, comme les moteurs de recherche, hiérarchisent et sélectionnent l'information. L'hypothèse d'une *search engine literacy* (Le Deuff, 2017) est apparue opportune en matière de formation. Ainsi, il est impensable de songer à une littératie des données qui se focalise uniquement sur les formes obtenues en négligeant les processus de traitement pour obtenir, enrichir et interpréter les données. Plus encore, les positionnements en matière de données personnelles ne peuvent s'entendre dans une relation réduite entre la plateforme et l'individu, tant cette relation ne permet pas de saisir les enjeux comparatifs qui permettent aux plateformes de tirer parti des données personnelles et donc de les monétiser. Cette transformation rejoint le constat de Bernard Stiegler qui montrait qu'il s'est produit un bouleversement dans la gestion des données, notamment en ce qui concerne les métadonnées : « *Les pouvoirs centraux sont des producteurs de métadonnées, que ce soit le Parlement, le prêtre roi, le mandarin chinois, l'Académie française de Richelieu... Un tel pouvoir est un processus top down, c'est-à-dire hiérarchique et descendant. Or, depuis 1992, depuis l'apparition du World Wide Web, la production de métadonnées – c'est-à-dire des éléments de base de la synchronisation – est devenue un processus bottom up, c'est-à-dire ascendant : la métadonnée numérique est produite par des gens de manière empirique, qui n'ont pas du tout l'ambition de prendre le pouvoir* » (Ballarini-Santonocito et Serres, 2009).

Sans prétendre à permettre une prise de pouvoir total des citoyens, la littératie des données distribuée offre au moins la piste de reprise du pouvoir sur les données personnelles ainsi que sur les opérations afférentes, ce qui suppose de donner au citoyen « pouvoir comprendre » ce qui se produit et de « pouvoir faire » différemment s'il le souhaite notamment dans ses choix d'applications.

La littératie des données implique donc autant le souci des données que des instruments qui permettent de les traiter et de les comprendre. Historiquement, chez Otlet, l'enjeu était celui d'une accessibilité fine au plus petit élément en permettant à l'utilisateur de pouvoir disposer d'un outil spécifique qui plaçait l'individu au centre du dispositif. Or, les enjeux de centralisation se sont développés et accrus au fur et à mesure que la nouvelle économie du Web a placé en situation d'oligopole de grands groupes qui cherchent à conserver

⁸. Traduction par nos soins.

leur centralité alors même qu'ils prétendent s'inscrire dans des logiques réticulaires ou de graphes. Ils veulent en être les administrateurs centraux, ce qui implique de penser la littératie des données distribuée dans une logique qui change le déséquilibre entre les gestionnaires des données et les individus fournisseurs.

Une des pistes est celle qui permet de s'emparer des outils et des processus pour les modifier en fonction de ses besoins. Si ce modèle ne correspond pas aux vellétés économiques des plateformes, il est plus fréquent dans les secteurs qui fonctionnent avec des outils *open source*. Or, dans ce cadre, l'accessibilité du code n'est pas la garantie pleine d'une utilisation potentielle tant se trouve déficitaire la question de la documentation du code qui permet à l'expert comme au néophyte de pouvoir comprendre les éléments qu'il peut facilement réutiliser et modifier selon ses propres besoins.

4.2 Documenter les processus comme moyen de résistance

Participant d'une éthique des données, une documentation des processus apparaît opportune car elle répond justement aux besoins édictés, notamment par l'*open data charter* (Benkhala, 2020). Plus encore, c'est la production documentaire qui requiert des formes de littératie des données contributives et participatives, rejoignant la perspective d'une *information literacy* citoyenne (Owens, 1976), voire d'une culture de l'information citoyenne (Le Deuff, 2009), qui prenne en compte aussi bien les enjeux de réception et d'analyse de l'information que ses potentialités de production. Ces actions de documentation reposent sur la nécessité de pouvoir collecter, organiser et traiter des données en dehors des logiques institutionnelles traditionnelles et surtout en marge des logiques managériales pour permettre au citoyen de reprendre la possession de ses milieux de vie, notamment les espaces urbains, pour bénéficier de cartographies collaboratives plus précises (*Open Street Map*) que celles proposées par les services populaires pour mesurer les effets climatiques grâce à des capteurs.

À chaque fois, la littératie des données ne peut être isolée d'enjeux matériels. C'est même d'ailleurs le paradoxe : le fait de renvoyer à des petites unités informationnelles a tendance à faire oublier les dispositifs matériels qui accompagnent le processus à chaque étape. Cela renvoie à la question controversée des réseaux Internet et de la téléphonie mobile sans qui les transmissions ne peuvent s'effectuer.

L'acquisition de données fait clairement partie de processus documentaires qui requièrent des compétences différentes à chaque étape, depuis la constitution du processus jusqu'à leur archivage et conservation. Les volontés de documenter les violences policières participent également de ces enjeux qui répondent par des stratégies de « sousveillance » (Mann, 2003) à des logiques de surveillance en tentant finalement d'utiliser des outils proches pour tenter de répondre ou de mettre en avant des phénomènes qui ne peuvent être analysés uniquement de façon isolée. Et c'est clairement l'enjeu d'une littératie des données qui dépasse le paradigme du *data literate worker* pour aller sur celui du *data literate citizen* qui cherche à produire de façon collective des infrastructures de gestion de données et de démonstration qui ne soient ni celles des infrastructures publiques, ni celles des GAFAM. Ces logiques reposent sur diverses compétences qui sont celles de professionnels (documentaliste, informaticien, archiviste, journaliste...) qui font bénéficier le collectif de leurs savoirs et savoir-faire tout en cherchant à transmettre leurs compétences et leurs bonnes pratiques au réseau constitué.

Alors que l'idéologie d'un *data literacy worker* s'inscrit autant dans sa capacité à résister aux abus publicitaires que dans sa réponse aux besoins d'un marché du travail, l'idéologie d'une littératie des données distribuée renvoie à une volonté de résister aux captations de données et aux réutilisations qui échappent en général aux citoyens, finissant même par leur imposer des décisions sans leur consultation. Entre le discours institutionnel plaçant pour les données ouvertes, celui de l'*empowerment* peinant dans sa réelle application et transformation, et les captations et centralisation réalisées par les plateformes, la littératie des données distribuée apparaît ouvrir une autre voie, ou plutôt un nouveau « milieu de savoir » (Le Deuff, 2016). Ce dernier ne peut fonctionner totalement en marge tant il a besoin de s'appuyer sur les infrastructures étatiques et la production de données ouvertes. En outre, il ne peut rester en dehors des sphères industrielles tant ces dernières n'hésitent pas à réutiliser les productions réalisées et cherchent également à séduire ce public. En effet, on va retrouver comme *civic hackers* de nombreux professionnels issus tantôt du champ institutionnel tantôt du champ des services. Le milieu apparaît alors non pas comme une alternative, telle que peuvent apparaître parfois les logiciels développés par l'association Framasoft, mais au contraire comme un champ d'influence qui transforme les vellétés illusoires d'une indépendance

du cyberspace en une constitution inter, voire hyper-dépendante, des milieux de production, de collecte et de gestion des données, quelles que soient leurs formats et leur contexte.

Cela n'est possible que si se développent des espaces de documentation spécifiques qui, d'une part, décryptent les enjeux et traitements algorithmiques des plateformes dominantes ainsi que leurs biais, et d'autre part se posent en tant qu'espaces et techniques de résistance collaborative pour partager des données et des documents (Vigneron et al., 2019) qui permettent de faire apparaître des phénomènes d'inégalités sociales, culturelles ou de violences policières. Il s'agit enfin de pouvoir bénéficier d'une documentation du code notamment en ce qui concerne les entrepôts *open source*, tant il s'agit bien souvent non seulement d'avoir accès à des logiciels ou des scripts, mais également de savoir comment les utiliser et les modifier en pouvant bénéficier pour cela d'une documentation explicative.

5. Conclusion

Les référentiels de la littératie des données actuels renvoient à une vision autonome de la littératie, à savoir un set de compétences neutres applicables en toutes sociétés et toutes circonstances (Street, 1985). Les aspects culturels semblent fréquemment évacués de ces logiques processuelles et des méthodologies, un choix qui risque de nous faire passer à côté de ce qui fait l'essentiel d'une littératie, à savoir structurer et interroger le monde qui nous entoure. Cette dimension culturelle apparaît essentielle dans la mesure où il s'agit d'intégrer plus clairement la technique à la culture (Simondon, 1989). La littératie des données distribuée, comme intermédiaire d'une vision institutionnelle et d'impératifs économiques, apparaît alors comme une condition pour envisager des réseaux ou milieux métastables, c'est-à-dire bénéficiant d'infrastructures communes partagées et entretenues collectivement, tout en garantissant une neutralité ainsi que des potentialités d'innovation. Cette manière de concevoir la littératie des données permettrait d'éviter les milieux dissociés que dénonçait Stiegler (Stiegler, 2006) qui ont le double défaut de vouloir penser et révéler le monde tout en dissimulant ses méthodes et présupposés. La constitution de la littératie des données distribuée, telle que nous avons cherché à la définir ici, permettrait aussi aux différentes communautés de pouvoir construire des savoirs et savoir-faire en partageant des valeurs communes avec d'autres communautés. Le but est notamment d'éviter le risque d'uniformité culturelle et de transposition de méthodes comme ce fut le cas en ce qui concerne l'*information literacy* dominée par des conceptions états-uniennes dupliquées mondialement sans tenir compte du contexte culturel initial de leur mise en place (Basili, 2001). Les principes d'une littératie des données distribuée apparaissent à l'inverse pouvoir se situer entre une culture occidentale à prétention universaliste et un renouveau des contre-cultures qui montrent la nécessité de réellement profiter de l'architecture des réseaux Internet et Web pour faciliter à la fois un partage des données, des dispositifs et des connaissances qui les accompagnent, ainsi que de la possibilité de prendre part à leur interprétation et aux décisions qui s'en suivent.

Bibliographie

- Académie Française. (2020). *donné, -ée*. Académie Française. <https://www.dictionnaire-academie.fr/article/A9D3038>
- Ackoff, R-L. (1989). From data to wisdom. *Journal of Applied Systems Analysis*, vol. 16, p. 3-9.
- American Library Association. (2020). *Information Literacy*. literacy.ala.org. <https://literacy.ala.org/information-literacy/>
- Arruabarrena, B., Kembellec, G., Chartron, G. (2019). Data littératie & SHS : développer des compétences pour l'analyse des données. *Colloque CODATA-France*, Université Paris-Est Marne-la-Vallée.
- Bachimont, B. (2016). *Traces, calcul et interprétation : de la mesure à la donnée*, p. 13–35. <https://doi.org/10.1400/245293>
- Ballarini-Santonocito, I., Serres, A. (2009). Entretien avec Bernard Stiegler. *Mediadoc*, n°2, p. 3–7. <http://www.apden.org/IMG/pdf/mediadoc02.pdf>
- Balnaves, M., Willson, M. (2011). *A New Theory of Information & the Internet: Public Sphere Meets Protocol* (1st New edition). Peter Lang Publishing Inc.
- Barrowan, N. (2018). Why data is never raw : On the seductive myth of information free of human judgment. *The New Atlantis*, n° 56, p. 129-135.
- Basili, C. (2011). *Information literacy : un concetto solo statunitense? Information literacy: a USA-only concept?* 2nd ed., Vol.19.

- Baskarada, S., Koronios, A. (2013). Data, Information, Knowledge, Wisdom (DIKW): A Semiotic Theoretical and Empirical Exploration of the Hierarchy and its Quality Dimension. *Australasian Journal of Information Systems*, n°18, vol. 1. <https://doi.org/10.3127/ajis.v18i1.748>
- Bawden, D. (2001). *Information and digital literacies: A review of concepts*. *Journal of Documentation*, vol.57, n° 2, p. 218–259. <http://dlist.sir.arizona.edu/895/>
- Benkalha, D. (2020). [Spotlight] Walking the walk of Data Ethics. *@opendatacharter*. <https://medium.com/@opendatacharter/spotlight-walking-the-walk-of-data-ethics-c4bdd607ac63>
- Bhargava, R., Deahl, E., et al. (2015). Beyond Data Literacy: Reinventing Community Engagement and Empowerment in the Age of Data. Data-Pop Alliance. *Data-Pop Alliance White Paper Series*. <https://www.media.mit.edu/publications/beyond-data-literacy-reinventing-community-engagement-and-empowerment-in-the-age-of-data/>
- Bile, M. (2007). Jean-Marie Privat, Mohamed Kara, « La littératie. Autour de Jack Goody », *Pratiques*, 131-132. *Questions de communication*, n°12, p. 391–394. <https://journals.openedition.org/questionsdecommunication/2520>
- Bodin, F. (2018). La littératie des données comme enjeu d'éducation, d'action et de construction d'une citoyenneté numérique [Blog]. *Medium.Com/@franckbodin*. <https://medium.com/@franckbodin/la-litt%C3%A9ratic-des-donn%C3%A9es-comme-enjeu-d%C3%A9ducation-d-action-et-de-construction-d-une-citoyennet%C3%A9-9cd48779470d>
- Briet, S. (1951). *Qu'est-ce que la documentation ?*. Paris, Éditions documentaires, industrielles et techniques, coll. « Collection de documentologie », p.48.
- Buckland, M. (2017). *Information and Society*. The MIT Press, Cambridge.
- Carmes, M., Noyer, J.-M. (2014). L'irrésistible montée de l'algorithmique. *Les Cahiers du numérique*, Vol. 10, n°4, p. 63–102. <http://www.cairn.info/revue-les-cahiers-du-numerique-2014-4-page-63.htm>
- Claes, A., Philippette, T. (2020). Defining critical data literacy for recommender systems: A media-grounded approach. *Journal of Media Literacy Education*, vol. 12, n°3, p. 17-29. <https://doi.org/10.23860/JMLE-2020-12-3-3>
- CNAM. (2020). *Master Méga Données et Analyse Sociale (MédAS)* [Universitaire]. [formation.cnam.fr. https://formation.cnam.fr/rechercher-par-discipline/master-mega-donnees-et-analyse-sociale-medas--1085595.kjsp](https://formation.cnam.fr/rechercher-par-discipline/master-mega-donnees-et-analyse-sociale-medas--1085595.kjsp)
- Commission Européenne. (2006). Recommandation du Parlement Européen et du Conseil du 18 décembre 2006 sur les compétences clés pour l'éducation et la formation tout au long de la vie. *Journal Officiel de l'Union Européenne*, L.394/10. <http://www.aede-france.org/recommandations-EU.html>
- Commission Européenne. (2020a). *Façonner l'avenir numérique de l'Europe*. [ec.europa.eu. https://ec.europa.eu/info/strategy/priorities-2019-2024/europe-fit-digital-age/shaping-europe-digital-future_fr](https://ec.europa.eu/info/strategy/priorities-2019-2024/europe-fit-digital-age/shaping-europe-digital-future_fr)
- Commission Européenne. (2020b). *Upskilling for life after the pandemic: Commission launches new digital competence guidelines*. [ec.europa.eu https://ec.europa.eu/luxembourg/news/upskilling-life-after-pandemic-commission-launches-new-digital-competence-guidelines_fr](https://ec.europa.eu/luxembourg/news/upskilling-life-after-pandemic-commission-launches-new-digital-competence-guidelines_fr)
- COPIST. (2020). *A propos de CoopIST*. [coop-ist.cirad.fr. https://coop-ist.cirad.fr/a-propos-de-coopist](https://coop-ist.cirad.fr/a-propos-de-coopist)
- Cormerais, F., Le Deuff, O. (2014). Le lettré du numérique dans la reconfiguration du savoir contemporain. *Digital Intelligence 2014*.
- Data to the People. (2020). *Databilities*. [datatothepeople.org. https://www.datatothepeople.org/databilities](https://www.datatothepeople.org/databilities)
- Day, R.-E. (2014). *Indexing It All: The Subject in the Age of Documentation, Information, and Data* (Édition : 1). The MIT Press, Cambridge.
- Diakopoulos, N. (2019). *Automating the News: How Algorithms Are Rewriting the Media* (Illustrated edition). Harvard University Press, Cambridge.
- D'Ignazio, C., Bhargava, R. (2015). Approaches to Building Big Data Literacy. In *Bloomberg Data for Good Exchange 2015*. New York, NY, USA. <https://www.media.mit.edu/publications/approaches-to-building-big-data-literacy/>
- Enright, N.-F. (2013). The violence of information literacy: Neoliberalism and the human as capital. In *Information Literacy and Social Justice: Radical Professional Praxis* (L. G. S. Higgins), p. 15–38. Library Juice Press.
- Federal Data Strategy. (2020). *Federal Data Strategy 2020 Action Plan*. Federal Data Strategy, p. 52. <https://strategy.data.gov/assets/docs/2020-federal-data-strategy-action-plan.pdf>

- Ferrari, A. (2013). *DIGCOMP: A Framework for Developing and Understanding Digital Competence in Europe*. JRC Scientific and policy reports. <https://publications.jrc.ec.europa.eu/repository/bitstream/JRC83167/lb-na-26035-enn.pdf>
- Fiaux, C. (2020). *La dataphabétisation devient un enjeu majeur au sein des entreprises* [Journal]. *Dafmag.fr*. <https://www.dafmag.fr/Thematique/data-methodologie-1238/Breves/datalphabetisation-devient-enjeu-majeur-sein-entreprises-350798.htm>
- Frank, M., Walker, J., Attard, J., Tygel, A. (2016). Data Literacy - What is it and how can we make it happen? *The Journal of Community Informatics*, vol. 12, n°3, p. 4–8.
- Freire, P. *The Pedagogy of the oppressed*. Traduction anglaise de Mira Bergman Ramos. Continuum, New York, 2005.
- Goody, J., Watt, I. (1963). The consequences of Literacy. *Comparative Studies in Society and History*, vol. 5, n°3, p. 304–345.
- Gouvernement de l'Ontario. (2009). Critical Literacy : a lens for learning. *The literacy and numeracy secretariat, capacity building series, special edition* n°9. http://www.edu.gov.on.ca/eng/literacynumeracy/inspire/research/critical_literacy.pdf
- Graff, H.-J. (1979). *The Literacy Myth: Literacy and Social Structure in the Nineteenth-Century City*. Academic Press, Routledge, Londres.
- Grundmann, H. (1958). Litteratus-Illiteratus : Der Wandel einer Bildungsnorm vom Altertum zum Mittelalter. *Archiv für Kulturgeschichte*, vol. 40, p. 1–65.
- Hapke, T. (1999). *Wilhelm Ostwald, the "Brücke" (Bridge), and connections to other bibliographic activities at the beginning of the twentieth century*. *Information Today*. <https://doi.org/10.15480/882.1>
- Herzog, D. (2015). *Data literacy : a user's guide*. SAGE Publications, p.224, Thousand Oaks, États-Unis.
- Howard, R. (1996). *Les communautés virtuelles*. Addison-Wesley France, Paris.
- Institut Universitaire de Technologie Bordeaux Montaigne. (2020). *Licence professionnelle Médiations de l'Information Numérique et des Données*. [Universitaire]. <http://www.iut.u-bordeaux-montaigne.fr/diplomes/licences-pro/mediations-de-linformation-numerique-et-des-donnees-mind/>
- Institut Universitaire de Tours. (2020). *LP Métiers de l'information spécialité Data management – Objectifs*. [Universitaire]. <https://iut.univ-tours.fr/> <https://iut.univ-tours.fr/version-francaise/departement-information-communication/lp-data-information>
- Iwasiński, Ł. (2020). *Theoretical Bases of Critical Data Studies*. Université de Varsovie. Pologne. <http://ojs.sbp.pl/index.php/zin/article/view/726/637>
- Kara, M., Privat, J.-M. (2012). *Jack Goody et l'empire de la littératie (version intégrale)*. https://www.canal-u.tv/video/canal_socio/jack_goody_et_l_empire_de_la_litteratie_version_integrale.11211
- Kayser-Bril, N. (2016). Don't ask too much from data literacy. *The Journal of Community Informatics*, 12(3), 217–222. <http://www.ci-journal.net/index.php/ciej/article/view/1297>
- Koltay, T. (2015). Data literacy: in search of a name and identity. *Journal of Documentation*, 71(2), 401–415. <https://doi.org/10.1108/JD-02-2014-0026>
- Latour, B. (1984). *Les microbes : guerre et paix suivi de Irréductions*. Editions A. M. Métailié, Paris.
- Latour, B. (1996). Ces réseaux que la raison ignore - laboratoires, bibliothèques, collections. In *Le pouvoir des bibliothèques : la mémoire dans la culture occidentale*, p. 23–46. Albin Michel.
- Le Bulletin Officiel de l'Éducation Nationale. (2017). *Bulletin officiel n°6 du 9 février 2017*. Commission d'enrichissement de la langue française - Vocabulaire de l'éducation et de l'enseignement supérieur. Ministère de l'Éducation Nationale, de la Jeunesse et des Sports. <https://www.education.gouv.fr/bo/17/Hebdo6/CTNR1636064K.htm>
- Le Deuff, O. (2009). Penser la conception citoyenne de la culture de l'information. *Les Cahiers Du Numérique*, vol. 5, n°3, p. 39–49.
- Le Deuff, O. (2012). Littératies informationnelles, médiatiques et numériques : de la concurrence à la convergence ? *Études de communication*, vol. 38, p. 131–147.

- Le Deuff O. (2016), « Humanités digitales et bibliothèques », *Revue de l'Enssib*.
<https://publications-prairial.fr/balisages/index.php?id=133>
- Le Deuff, O. (2017). Search engine literacy. *European Conference on Information Literacy*, p. 359–365.
- Le Deuff, O., Perret, A. (2019). Hyperdocumentation: origin and evolution of a concept. *Journal of Documentation*, vol. 75, n°6, p. 1463–1474. <https://doi.org/10.1108/JD-03-2019-0053>
- Lehmans, A. (2017). Données ouvertes et redéfinition de la culture de l'information dans les organisations
- Lehmans, A. (2018). Les réinventions de la démocratie à l'aune de l'ouverture des données : du discours de la participation aux contraintes de la gouvernance. *Les enjeux de l'information et de la communication*, vol.19, n°2, p. 135–146. <https://www.cairn-int.info/revue-les-enjeux-de-l-information-et-de-la-communication-2018-2-page-135.htm>
- Lewison, M. (2002). Taking on critical literacy: The journey of newcomers and novices. *Language Arts*, vol. 79, n°5. https://www.researchgate.net/publication/284055039_Taking_on_critical_literacy_The_journey_of_newcomers_and_novices
- Loi n° 2016-1321 du 7 octobre 2016 pour une République numérique, 2016–1321, ECFI1524250L 96 (2016). <https://www.legifrance.gouv.fr/jorf/id/JORFTEXT000033202746?r=Ys9sdtEJ6o>
- Levy, P. (1997). *La cyberculture. Rapport au Conseil de l'Europe*. Odile Jacob, Paris.
- Lim, F.-V., O'Halloran, K. (2011). Dimensioner af Multimodal Literacy. *Viden om Læsning*, vol. 10, p. 14–21. https://www.researchgate.net/publication/306006944_Dimensions_of_Multimodal_Literacy
- Mackey, T.-P., Jacobson, T. (2010). Reframing Information Literacy as a Metaliteracy. *College & Research Libraries*, n°72, p. 62–78. 10.5860/crl-76r1
- Mann, S., Nolan, J., Wellman, B. (2003). Sousveillance : Inventing and Using Wearable Computing Devices for Data Collection in Surveillance Environments. *Surveillance & Society*, vol. 1, n°3, p. 331-355.
- Mandinach, E.-B., Honey, M., Light, D., Brunner, C. (2008). A Conceptual Framework for Data-Driven Decision Making. *Teachers College Press*.
- Mattern, S. (2013). Methodolatry and the Art of Measure. *Places Journal*. <https://doi.org/10.22269/131105>
- Mattern, S. (2020). The Spectacle of Data: A Century of Fairs, Fiches, and Fantasies. *Theory, Culture & Society*, 0263276420958052. <https://doi.org/10.1177/0263276420958052>
- Mayer-Schönberger, V., Cukier, K. (2013). *Big Data: A Revolution That Will Transform How We Live, Work, and Think*. Eamon Dolan / Houghton Mifflin Harcourt, Boston.
- McLuhan, M. (1964). *Understanding Media : the Extensions of Man*. New American Library, New York.
- Meriam-Webster. (2020a). Illiteracy. In *Meriam-Webster*. <https://www.merriam-Webster.com/dictionary/illiteracy>
- Meriam-Webster. (2020b). Literacy. In *Meriam-Webster*. <https://www.merriam-Webster.com/dictionary/literacy>
- Meriam-Webster. (2020c). Literate. In *Meriam-Webster*. <https://www.merriam-Webster.com/dictionary/literate>
- Nayek, J.-K. (2015). Data literacy and Library: An Overview. *College libraries*, vol. 30 (I–II), p. 38–51. https://www.researchgate.net/publication/304597538_Data_Literacy_and_Library_An_Overview
- Nieuwbourg, P. (2020). *Note de recherche : glossaire des métiers de la donnée*. Decideo. P. 20. https://www.decideo.fr/Podcast-Presentation-du-Glossaire-des-Metiers-de-la-Donnee_a11950.html
- Oceans of Data Institute. (2015). *Building Global Interest in Data Literacy: A Dialogue. Workshop Report* (p. 20). Education Development Center. http://oceansofdata.org/sites/oceansofdata.org/files/ODI%20Data%20Literacy%20Report_0.pdf
- Otlet, P. (1926). *L'avenir de la comptabilité et ses rapports avec les besoins de l'organisation mondiale*. La Senne, Bruxelles.
- Otlet, P. (1934). *Traité de documentation. Le livre sur le livre*. Palais Mondial, Bruxelles.
- Owens, M.-S. (1976). *Government and Libraries*. Library Journal.
- Petty, C. (2018). *Gartner Keynote: Do You Speak Data?* [Entrepreneurial]. Gartner.com. <https://www.gartner.com/smarterwithgartner/gartner-keynote-do-you-speak-data/>
- Pierre, R. (2003). Entre alphabétisation et littératie : les enjeux didactiques. *Revue française de linguistique appliquée*, vol. VIII, n°1, p. 121–137. <https://www.cairn.info/revue-francaise-de-linguistique-appliquee-2003-1-page-121.htm>

- Plassart, P. (2020). *La Data Literacy ou l'art de donner du sens à la donnée : le véritable enjeu de nos entreprises*. [Entrepreneurial]. a5sys.com. <https://www.a5sys.com/data-literacy/>
- Punie, Y. (2017). *DigComp 2.1 : The Digital Competence Framework for Citizens* (p. 48). Joint Research Centre. <https://ec.europa.eu/jrc/en/publication/eur-scientific-and-technical-research-reports/digcomp-21-digital-competence-framework-citizens-eight-proficiency-levels-and-examples-use>
- Qlik. (2020). *Certifications et qualifications Qlik* [Entrepreneurial]. qlik.com. <https://www.qlik.com/fr-fr/services/training/certifications-and-qualifications>
- Rayward, W.-B. (1975). *The universe of information: the work of Paul Otlet for documentation and international organisation*. Published for International Federation for Documentation (FID) by All-Union Institute for Scientific and Technical Information (VINITI).
- République française. (2020). *Les compétences évaluées par Pix*. Pix.fr. <https://pix.fr/competences>
- Roberge, J.-F. (2019). *Digital competency Framework*. Ministère de l'Éducation et de l'Enseignement supérieur du Québec. P. 34. http://www.education.gouv.qc.ca/fileadmin/site_Web/documents/ministere/Cadre-referenc-competence-num-AN.pdf
- Rouvroy, A., Berns, T. (2013). Gouvernamentalité algorithmique et perspectives d'émancipation. *Réseaux*, vol. 177, n°1, p. 163–196. http://www.cairn.info/resume.php?ID_ARTICLE=RES_177_0163
- Rowley, J. (2007). The wisdom hierarchy: representations of the DIKW hierarchy. *Journal of Information Science*, vol. 33, issue 2. <https://www.semanticscholar.org/paper/The-wisdom-hierarchy%3A-representations-of-the-DIKW-Rowley/bdf94027d5410b211411157ad7aacfca05aea53d>
- Schild, M. (2005). Information Literacy, Statistical Literacy and Data Literacy. *IASSIST quarterly / International Association for Social Science Information Service and Technology*, vol. 28, n°2–3, p. 7–14. <https://doi.org/10.29173/iq790>
- Serres, A. (2000). *Aux sources d'Internet : l'émergence d'ARPANET. Exploration du processus d'émergence d'une infrastructure informationnelle. Description des trajectoires des acteurs et actants, des filières et des réseaux constitutifs de la naissance d'ARPANET. Problèmes critiques et épistémologiques posés par l'histoire des innovations*. [Phdthesis, Université Rennes 2]. <https://tel.archives-ouvertes.fr/tel-00312005>
- Simondon, G. (1989). *Du mode d'existence des objets techniques*. Editions Aubier, Paris.
- Soubiran, V. (2020). *Développer les compétences data : une urgence pour le bien-être de vos salariés*. bigdataparis.com. <https://www.bigdataparis.com/2020/developper-les-competences-data-une-urgence-pour-le-bien-etre-de-vos-salaries/>
- Špiranec, S., Kos, G., George, M. (2019). Searching for critical dimensions in data literacy. Proceedings of the Tenth International Conference on Conceptions of Library and Information Science, Ljubljana, Slovenia, June 16-19. <http://informationr.net/ir/24-4/colis/colis1922.html>
- Statista. (2019). *Digital Economy Compass 2019*. Statista. P. 230. <https://cdn.statcdn.com/download/pdf/DigitalEconomyCompass2019.pdf>
- Stiegler, B. (2006). *La télécratie contre la Démocratie*. Flammarion, Paris.
- Street, B. (1985). *Literacy theory in practice*. Cambridge University Press, Cambridge.
- Street, B. (2001). *Literacy and Development: Ethnographic Perspectives* (1st ed.). Routledge, Londres.
- Striphas, T. (2015). Algorithmic culture. *European Journal of Cultural Studies*, 18(4-5), 395-412.
- Tardif, J. (2006). *L'évaluation des compétences. Documenter le parcours de développement*. Chenelière Éducation.
- Thiault, F. (2020). *Data librarian et services aux chercheurs en bibliothèque universitaire : de nouvelles médiations en émergence*. <https://hal.archives-ouvertes.fr/hal-02972705/document>
- Tice-Education. (2016). *PIX : un projet public de plateforme en ligne d'évaluation de certification des compétences numériques*. tice-education.fr. <https://www.tice-education.fr/index.php/tous-les-articles-er-ressources/examens/1115-pix-un-projet-public-de-plateforme-en-ligne-d-evaluation-de-certification-des-competences-numeriques>
- Thomas, S., Lacceti, J., Mason, B., Joseph, C. (2007). Transliteracy: Crossing divides. *First Monday*, 12. 10.5210/fm.v12i12.2060
- Tygel, A-F., Kirsch, R. (2016). Contributions of Paulo Freire for a Critical Data Literacy: a Popular Education Approach. *The Journal of Community Informatics*, vol. 12, n°3. <https://doi.org/10.15353/joci.v12i3.3279>

- Van Acker, W. (2011). *Universalism as utopia: a historical study of the schemes and schemas of Paul Otlet (1868-1944)* [Dissertation, Ghent University]. <http://hdl.handle.net/1854/LU-1942681>
- Vigneron, N.-B., Picton, É., Beyou, S., Coadic, X. (2019). De l'hypothèse de la documentation comme technique de résistance et du wiki comme objet de ces résistances. *Sens public*. <http://sens-public.org/articles/1375/>
- Willinsky, J. (1990). *The new literacy: Redefining Reading and Writing in the schools* (1st ed.). Routledge.
- Wilson, A.-A. (2008). Learning from Past and Present Uses of the Term 'Literacy'. *American Reading Forum Yearbook*, vol. 28, n° 73. https://18970341-7a2c-4cd7-a80939ba885eda1d.filesusr.com/ugd/c10ff9_b75759189b934b04a3cc7980d253c48d.pdf
- Zins, C. (2007). Conceptual approaches for defining data, information, and knowledge. *Journal of the American Society for Information Science and Technology*, vol. 58, issue 4, p. 479 - 493. <https://doi.org/10.1002/asi.20508>