



HAL
open science

Fonction documentaire de preuve et données numériques

Arthur Perret

► **To cite this version:**

Arthur Perret. Fonction documentaire de preuve et données numériques. 7ème conférence Document numérique & Société, Sep 2020, Nancy, France. hal-04271477

HAL Id: hal-04271477

<https://hal.science/hal-04271477>

Submitted on 6 Nov 2023

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.

Fonction documentaire de preuve et données numériques

Arthur Perret

2021

L'une des fonctions documentaires traditionnelles est la preuve. Elle fait partie des aspects du document que le passage à l'informatique nous amène à questionner mais la littérature francophone en sciences de l'information et de la communication (SIC) ne s'en est pas encore emparée spécifiquement. En régime numérique, elle semble devenir la prérogative de systèmes informatiques comme la signature électronique, un nouvel ordre documentaire innervé par la cybersécurité. L'irruption des données en information-communication provoque une dynamique qui permet repenser cette fonction documentaire à partir d'un contexte théorique et technique actualisé. Nous proposons une formalisation de cette fonction d'après une typologie de positions et de relations qui forment un diagramme de quelques articulations conceptuelles entre information, communication et documentation. À partir d'exemples pris dans le champ scientifique (la citation et le carnet programmable ou notebook), nous montrons comment la valeur de preuve perdure et se transforme.

Preuve et document

Le document constitue un objet d'étude fondateur pour les sciences de l'information et de la communication (SIC). Autour de lui, et dans le sillage du Belge Paul Otlet (1868-1944), premier théoricien de la documentation, s'est développé tout au long du XX^e siècle une tradition de recherche appelée documentologie¹. Celle-ci s'est attachée à l'étude des racines conceptuelles, des usages sociaux et des problématiques professionnelles liées aux documents. Ce faisant, elle n'a pas manqué d'inscrire la documentation dans des problématiques travaillées plus largement par d'autres sciences humaines et sociales (SHS). Ces passerelles sont particulièrement apparentes dès qu'il est question des fonctions remplies par les documents. Ainsi, lorsque certaines écoles de pensée schématisent ces fonctions sous les termes de « notification » et de « gouvernance² », on ne peut s'empêcher de voir s'ajouter à l'enchevêtrement de l'information et de la communication, inhérent à la notion de document, celui du savoir et du pouvoir, tel que décrit par Michel Foucault³.

Un autre exemple, et que nous souhaitons développer spécifiquement dans cet article, est la **preuve**. Dans la tradition des pionniers de la documentation, elle fait partie des fonctions attribuées au do-

¹ Meyriat, « Les sciences de l'écrit », 1993.

² Sokolov, « The epistemology of documents », 2009, p. 57, 61.

³ Foucault, *Surveiller et punir*, 1993, p. 32.

cument dès sa définition. Ainsi chez Suzanne Briet : « une preuve à l'appui d'un fait ⁴ ». On retrouve cette approche dans certains travaux en bibliothéconomie et science de l'information anglo-saxonne (*Library and Information Science*) : Michael Buckland, par exemple, s'inscrit dans la lignée de Briet lorsqu'il définit les deux objectifs des systèmes d'information sous l'angle de la preuve – donner accès à ce qui a été cité comme preuve d'une assertion, et mettre en évidence ce qui pourrait être l'objet d'un tel usage ⁵. Dans d'autres champs, comme le *records management*, certains auteurs comme Marie-Anne Chabin considèrent plutôt la preuve comme le critère d'une définition distincte, qui voit le document se muer en *record*, c'est-à-dire « un objet d'information sanctionné par une instance supérieure à son auteur [...] et présentant une valeur de preuve ⁶ ».

Les ressorts de cette importance de la notion de preuve sont bien sûr millénaires. La preuve peut être définie comme ce qui est « susceptible d'établir de manière irréfutable la vérité ou la réalité de quelque chose ⁷ ». Elle s'inscrit ainsi dans la tradition philosophique et épistémologique ; le logicien Fernando Gil, par exemple, lui accorde une place significative, presque séminale :

« ses grands opérateurs – l'argumentation, le témoignage, un appel à l'expérience qui revêt des formes multiples [...] la référence à l'expérience est peut-être la dimension élémentaire de l'intelligibilité : l'expérience fait preuve ⁸ ».

Toutefois, si la preuve apparaît de nature fondamentalement rationnelle, on ne peut pas négliger l'importance de son incarnation matérielle. Le rapport entre cette notion et celle de document est quasiment constitutif de son acception moderne :

« La preuve est toujours un raisonnement : jamais la présentation d'un fait n'est en soi une preuve ; c'est simplement le moyen de l'établir. Cependant, on n'hésite pas à dire d'un fait qu'il est la preuve d'une affirmation ; et dans le langage juridique, le mot s'applique très bien à des documents ⁹ ».

Jean-Daniel Zeller rappelle que « la définition du document comme preuve par écrit est relativement récente et issue du langage juridique ¹⁰ », du fait de l'étymologie même du mot document. Le rapport entre les deux, en revanche, est bien antérieur. La preuve est étroitement liée à l'organisation des activités humaines par l'écriture, et par ce qu'on n'appellait pas encore des documents. On peut observer cela dès les sociétés anciennes, et d'abord dans leur économie, tel que le montre par exemple Jack Goody pour la Mésopotamie : « le document écrit [sert] de preuve et de garantie ¹¹ » ; il accompagne le déve-

⁴ Briet, *Qu'est-ce que la documentation ?*, 1951, p. 7.

⁵ Buckland, « What is a document? », 1997, p. 806-808.

⁶ Chabin, « Document trace et document source », 2004, p. 145.

⁷ *Trésor de la langue française*, <https://cnrtl.fr/definition/preuve>

⁸ Gil, « Preuve », 2006, p. 881.

⁹ Lalande, *Vocabulaire technique et critique de la philosophie*, 2010, p. 822-823.

¹⁰ Zeller, « Documents numériques », 2004, p. 113.

¹¹ Goody, *La logique de l'écriture*, 2018, p. 108.

loppement de processus complexes tels que le crédit ou les actions. Plus globalement, toujours selon Goody, « le passage à l'écriture est une force motrice ; elle favorise l'émergence d'un concept plus formel de la preuve, et en un certain sens, de la vérité elle-même ¹² ». L'écriture et le document viennent ainsi accroître l'importance de la preuve dans le rapport des humains au monde et à eux-mêmes. Ce faisant, elle s'affirme comme une construction culturelle et technique. C'est cet épaississement indéniable que décrit Paul Otlet lorsqu'il évoque le lien entre la documentation et le réel : « le Document offre de la Réalité une image à la sixième dérivation ¹³ ».

Le document est un objet de communication, donc de réception et d'évaluation. Cet aspect amène par exemple Ron Day à considérer la documentation parmi les « technologies du jugement ¹⁴ ». La preuve obéit à cette logique. Toutefois, le jugement n'est pas mécanique et fonctionne en tandem avec une dimension subjective liée à ce qu'on pourrait appeler la confiance :

« La preuve ne se réduit pas à la seule attestation de la vérité [...] Les méthodes de la preuve ont une contrepartie subjective que Leibniz, dans les *Nouveaux Essais*, appelle "créance" ou "degrés de l'assentiment" : assurance, confiance, ferme créance, croyance, conjecture, doute, incertitude, défiance, dans une échelle où, écrit Leibniz, "les degrés de probabilité peuvent se diversifier à l'infini" ¹⁵ ».

Cette subjectivité soulève un point crucial : une preuve n'est pas toujours vérifiée. Elle fait parfois l'objet d'un jeu rhétorique, performatif, dans lequel faire l'expérience de ce qui est présenté comme preuve est substitué à la preuve elle-même. Les logiciens du Cercle de Vienne avaient mis l'accent sur cette problématique dans le débat épistémologique au début du XX^e siècle, en considérant respectivement la « confrontation empirique d'un énoncé avec les faits » (vérification) et sa « possibilité logique » (vérifiabilité) ¹⁶. La confiance entretient donc avec la preuve un rapport qui relève de la virtualité et de l'actualisation.

C'est sur ces bases que se sont constitués depuis le XIX^e siècle des « ordres ¹⁷ » ou « régimes de pratiques ¹⁸ » documentaires successifs. Leur consolidation est indissociable de l'exercice du pouvoir, des processus de légitimation, des mécanismes d'autorité, au sein desquels la fonction et la valeur de preuve des documents représentent une constante.

¹² *Ibid.*, p. 199.

¹³ Otlet, *Traité de documentation*, 2015, p. 44.

¹⁴ Day, *Documentarity*, 2019, p. 9.

¹⁵ Gil, « Preuve », 2006, p. 882.

¹⁶ Benmakhlof, « Vérification », 2006, p. 1125-1127.

¹⁷ Salaün, *Vu, lu, su*, 2012, p. 33.

¹⁸ Müller, « À la recherche des archives de la recherche », 2006, p. 24 ; Müller, « De l'archive au document », 2011, p. 2.

Preuve et informatique

Or le devenir contemporain de ces ordres ou régimes est marqué par un changement de paradigme technique désormais largement analysé et qui a vu se généraliser l'outil informatique dans toutes les dimensions de l'information et de la communication. Ceci entraîne, sinon des ruptures, du moins des déplacements et des renouvellements. En tant qu'objet de communication, le document est toujours « cristallisation de processus, produit de métamorphoses successives ¹⁹ ». Passé au tamis numérique, il ressort encore plus processuel qu'auparavant, « pris dans une logistique dont on ne peut (pas facilement) l'extraire ²⁰ ».

Dans ce contexte, quid de la fonction documentaire de preuve ? Tous les travaux pré-cités la mentionnent, sans s'attarder spécifiquement sur elle. Il est vrai qu'il fallait d'abord établir les problématiques plus générales de la redocumentarisation ²¹ et, plus récemment, de la documentation nativement numérique ²². Dans ce contexte, la « valeur d'évidence » constitue notamment, avec la « valeur d'information », le point de départ d'une redéfinition du document par le collectif Pédauque ²³. Mais par la suite, la littérature francophone en SIC a développé d'autres pistes. La valeur et la fonction de preuve ont été subsumées sous les notions de légitimité et de légitimation ²⁴, de médium et de contrat ²⁵ ou encore d'évaluation de l'information ²⁶. L'intérêt pour la preuve n'a pas disparu, mais le concept de document est moins souvent mobilisé pour l'approcher que celui d'information et de véracité, dans le sillage de la recherche anglophone sur la falsification de l'actualité (*fake news*), des médias audiovisuels (*deep fakes*) ou des données de recherche (*fake science*).

Puisqu'on constate cette mise en retrait en SIC, et que par ailleurs le lien entre preuve et document est étroitement lié au droit, il nous paraît utile avant d'aller plus loin d'examiner ce que dit ce dernier de la preuve documentaire en régime numérique. À la question « dans quelle mesure les documents numériques obéissent à la qualification juridique d'instrument de preuve ? », Albert Dione établit que « c'est le critère de sécurité et donc de fiabilité qui est retenu pour asseoir la valeur juridique du document électronique ²⁷ ». En termes juridiques, l'attribution de la valeur de preuve est donc désormais la prérogative de systèmes informatiques comme la signature électronique. Ce basculement est confirmé par les modèles des réseaux informationnels. Celui du web sémantique, par exemple, proposé par Tim Berners-Lee ²⁸, représente la confiance (*trust*) comme reposant

¹⁹ Cotte, « Le concept de document numérique », 2004, p. 36.

²⁰ Robert et Pinède, « Le document numérique », 2012, p. 199.

²¹ Pédauque, *La redocumentarisation du monde*, 2007.

²² Epron, Pinède et Tona, « Introduction », 2020.

²³ Pédauque, *Document : forme, signe et médium*, 2003, p. 1.

²⁴ *Ibid.*, p. 22 ; Epron et Vitali-Rosati, *L'édition à l'ère numérique*, 2018, p. 7-8, 27-33.

²⁵ Salaün, *Vu, lu, su*, 2012, p. 28-30.

²⁶ Le Deuff, *La culture de l'information en reformation*, 2009, p. 362.

²⁷ Dione, *Les aspects juridiques de la dématérialisation des documents du commerce maritime*, 2018, p. 286.

²⁸ Berners-Lee, *Semantic Web on XML*, 2000, p. 10.

sur deux appuis distincts : la cryptographie (*digital signature*) et la preuve (*proof*). Le nouvel ordre documentaire est donc innervé par la cybersécurité, laquelle consiste à objectiver en partie la confiance en la mécanisant.

La fonction documentaire de preuve s'en trouve nécessairement modifiée. Jusqu'à quel point ?

Un cadre d'analyse documentaire

Nous souhaitons appliquer ici un cadre d'analyse documentaire, qu'il nous faut d'abord présenter, puis expliciter.

Éléments constitutifs

Les termes utilisés pour caractériser la communication sont ceux de Claude Shannon ²⁹ ; la typologie de l'information est celle de Michael Buckland ³⁰, complétée par Tim Gorichanaz ³¹ ; l'indexicalité est reprise dans l'acception de Ron Day ³². Nous introduisons la documentarité, qualité de ce qui fait document à nos yeux, référence à une proposition conceptuelle récente et qui procède de la même ambition syncrétique ³³. Le document occupe une position médiatrice entre émetteur et récepteur, qui met en jeu un rapport de documentarité entre récepteur et document, et de fiabilité entre document et émetteur. Il intervient également entre récepteur et connaissable, constituant une modalité d'expérience par la somme de la documentarité et de l'indexicalité.

Il en résulte un schéma (« carré documentologique », cf. figure 1) qui ne constitue pas un diagramme parfait des articulations conceptuelles entre information, communication et documentation, mais qui se veut conçu comme un « générateur de connaissance », avec une logique combinatoire ³⁴. Ainsi les cinq rectangles constituent des **positions** : émetteur, récepteur, document, donnée, connaissable. Une entité peut occuper différentes positions ; par exemple, un humain peut être considéré comme un document, ou être impliqué dans une communication en tant qu'émetteur, ou que récepteur. Certaines relations constituent la somme d'autres relations ; le graphe peut devenir dirigé en fonction des entités considérées, ou de la perspective adoptée. En somme, il s'agit d'une matrice.

Dans ce schéma, les données sont distinguées des documents. Ce choix nécessite une explication. Contrairement à ce qu'anticipait Paul Otlet, c'est l'information et non la documentation qui est devenue le concept fondateur du nouveau champ scientifique qu'il appelait de

²⁹ Shannon, « A Mathematical Theory of Communication », 1948.

³⁰ Buckland, « Information as thing », 1991.

³¹ Gorichanaz, *Understanding Self-Documentation*, 2018.

³² Day, « All that is the case: documents and Indexicality. », 2016.

³³ Perret et Le Deuff, « Documentarité et données, instrumentation d'un concept », 2019.

³⁴ Drucker, *Graphesis*, 2014, p. 105-107.

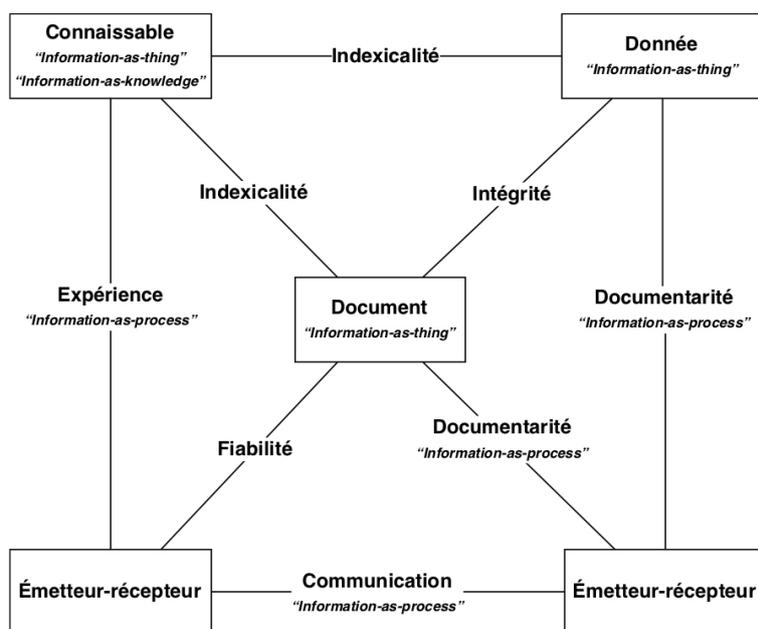


FIGURE 1: Carré documentologique.

ses vœux. Or le travail de modélisation théorique ne peut pas toujours se passer complètement d'une dimension structurale, granulaire, à laquelle l'information se prête mal. La quantité d'information de Claude Shannon par exemple, reste d'abord et avant tout le fondement d'une théorie de la communication centrée sur un contenu, dont la question du sens est évacuée. De manière générale, l'irruption des données dans les discours actuels met en lumière une certaine fragmentation théorique des sciences de l'information, avec une articulation conceptuelle compliquée par un « socle commun » qui ne l'est pas réellement³⁵.

Dans ce contexte, on peut voir dans la décomposition de l'information (ou du document) en données le retour d'une plus grande souplesse conceptuelle, non dénuée d'opérationnalité. Sylvie Leleu-Merviel, par exemple, adopte cette approche pour son modèle de la conceptualisation³⁶. Le socle de ce modèle, qui va du réel à l'information en passant par les données, est l'idée de *différence*, en référence aux travaux de Gregory Bateson et Luciano Floridi. La façon dont le modèle se déploie dessine les étapes d'une expérience instrumentée : la différence est perçue, prélevée et convertie en données, puis qualifiée et organisée en motifs reconnaissables – l'information. Cette nature construite des données constitue quasiment un consensus épistémologique trans-disciplinaire. Cela crée un paradoxe linguistique – le donné est en fait construit – mais qui se révèle

³⁵ Zins, « Conceptual approaches for defining data, information, and knowledge », 2007 ; Hjørland, « Information Science and Its Core Concepts », 2014, p. 226.

³⁶ Leleu-Merviel, « Le sens aux interstices, émergence de reliances complexes », 2010, cf. schéma p. 8.

stimulant, faisant naître des propositions terminologiques alternatives, comme obtenues plutôt que données³⁷ ou *capta* plutôt que *data*³⁸, entre autres.

Il nous semble également que la notion de données vient occuper un espace qui était autrefois celui du couple document-documentation : à la fois une catégorie, donc un indénombrable – comme « l'information », « la *data* », « le *big data* » – et une unité non seulement quantifiable mais ayant un sens structurel – « un certain nombre de données », « un jeu de données ». Ceci permet de distinguer deux plans de contenu, dans une approche documentaire classique : on désigne alors par « données » à la fois le contenu des documents, où le terme équivaut pour nous à « matériaux », et les métadonnées qui sont générées autour de ce contenu.

Fonctionnement

Afin d'illustrer ce dernier point, et le fonctionnement du diagramme de façon globale, prenons l'exemple d'une médiation documentaire classique entre une personne et un objet de connaissance potentielle : une personne lit un document dans le but de connaître quelque chose.

Ce lecteur occupe la position du récepteur dans un processus de communication. Il n'est pas en contact direct avec l'émetteur du message mais avec un document, qui joue le rôle de tiers médiateur. Face à cet objet, la personne évalue sa **documentarité** : celui-ci « fait document » ou pas, *à ses yeux*. C'est un acte interprétatif. Concrètement, la documentarité est une qualité, éventuellement quantifiable, décelée par le regard, en référence à des formes et normes documentaires que le lecteur connaît plus ou moins explicitement. Le document est ensuite lui-même le point de départ d'un autre acte interprétatif, relatif à la **fiabilité**, qui concerne l'établissement de la source et l'évaluation de son autorité. Certains auteurs approcheraient cette double question (documentarité, fiabilité) sous l'angle unique, plus générique, de la communication ou du contrat de lecture. Mais nous souhaitons tenter ici de décomposer ce type d'approche au moyen d'outils plus précis ; comme le dit Yves Jeanneret, « les SIC gagnent peut-être à déconstruire les contrats plutôt qu'à les modéliser³⁹ ».

Lorsqu'on évalue la **fiabilité**, on analyse une donnée présente dans le document : la source, l'auteur. Mais cette donnée n'est pas tout à fait sur le même plan que les autres. Dans le cas du document analo-

³⁷ Latour, « Pensée retenue, pensée distribuée », 2007, p. 609.

³⁸ Drucker, « Humanities Approaches to Graphical Display », 2011, par. 50.

³⁹ Jeanneret et Patrin-Leclère, « La métaphore du contrat », 2004, p. 139.

gique, elle fait souvent l'objet d'une disposition typographique particulière ; dans le cas du document numérique, elle n'est souvent visible que par l'intermédiaire d'une interaction spécifique. Ces modes d'énonciation se ramènent à une différence d'ordre méta (à la fois à propos et au-delà) qu'il faut franchir pour accéder à la donnée en tant que telle. Le franchissement de ce seuil relève de la **documentarité**. Son épaisseur suggère la nature particulière du « contrat » mis en œuvre. Tout ce qui est documentaire suppose cette structure particulière, organisée autour d'une propriété fondamentale que les théoriciens anglophones appellent *aboutness*⁴⁰ : il y a de l'information, et d'autre part de l'information *à propos* de cette information.

⁴⁰ Pomerantz, *Metadata*, 2015, p. 22.

Enfin, le document (et les données qu'il contient) *désignent* un objet de connaissance potentielle. La relation entre le document (ou les données) et ce **connaissable** est une relation d'**indexicalité**, c'est-à-dire un rapport métonymique entre ce qui est inscrit et ce qui est, tout court. À l'instar d'une métaphore, arriver à comprendre cette relation suppose une compétence d'écriture et de lecture, la maîtrise d'un jeu de langage particulier : il faut connaître une langue naturelle, peut-être un langage construit, une convention ou une norme, voire une bizarrerie propre à l'auteur ou aux aléas du temps.

Application

Nous déduisons du cadre exposé précédemment les composants de la fonction documentaire de preuve en régime numérique : une source (l'« émetteur » du modèle) et une information positionnée comme document, c'est-à-dire constituée de données et de métadonnées. La preuve est sollicitée au sein d'une communication. Celle-ci mobilise deux jugements : celui de la documentarité de l'information, et de la fiabilité de la source. L'actualisation de la preuve, elle, mobilise l'intégrité documentaire et l'indexicalité.

À quelles situations concrètes ces termes peuvent-ils renvoyer ? Prenons deux exemples.

Citation de références bibliographiques

La citation scientifique mobilise doublement la notion de preuve. Premièrement, il s'agit d'une preuve elle-même. La citation est enchâssée dans une rhétorique ; dans l'économie du texte scientifique, elle intervient le plus souvent pour appuyer un argument. Elle peut indiquer la réutilisation d'une idée, d'une méthodologie, ou de don-

nées ; elle peut aussi faire l'objet d'une critique, ou indiquer une étape précédente dans un cheminement.

L'informatique ouvre des possibilités de description et de mécanisation de cette rhétorique, par exemple avec la *Citation Typing Ontology*⁴¹. Ce vocabulaire contrôlé peut être associé à une technique d'annotation sémantique pour « enrichir » les citations et les listes de références bibliographiques. L'indexation de ces métadonnées permet ensuite d'en tirer parti dans la recherche et la visualisation d'information, l'éditionnalisation des textes scientifiques, et la bibliométrie⁴².

⁴¹ CiTO, <http://purl.org/spar/cito/2018-02-12>

⁴² Shotton, « CiTO, the Citation Typing Ontology », 2010.

Deuxièmement, la citation constitue elle-même une assertion à prouver. La preuve réside alors dans la présence de métadonnées d'identification (ISBN, DOI, URL, etc.) et de localisation (page, paragraphe, ancre, etc.).

La citation repose donc toujours sur la reconnaissance d'un schéma, qu'il s'agisse d'une convention rhétorique ou d'un vocabulaire contrôlé, et sur une relation présentée comme factuelle en vertu d'une indexicalité affichée. Ici, l'actualisation de la preuve nécessite bien à la fois un raisonnement qui lui est propre (l'argument présenté) et une vérifiabilité. Cette dernière dépend directement de l'intégrité entre les métadonnées référencées et l'information auxquelles elles sont associées ; cette intégrité dépend à son tour des acteurs impliqués dans la gestion de l'information (*knowledge* ou *records management*).

Carnets numériques programmables

Nous utiliserons ici le mot anglais *notebook* comme raccourci pour désigner un carnet numérique programmable, environnement d'édition qui permet d'écrire à la fois du texte et du code informatique exécutable. Il peut être publié tel quel, ou bien sous une forme éditorialisée dans laquelle il est possible d'interagir avec le code ; dans ce second cas, il est parfois appelé « essai computationnel ». De façon générale, on peut le qualifier de document-machine. Il s'agit d'un avatar de la programmation lettrée – inventée par l'informaticien Donald Knuth – dont la popularité est liée à celle des langages de programmation utilisés en statistique (comme R), en analyse de données (comme Python) ou encore en visualisation web (comme la bibliothèque Javascript D3).

Le *notebook* facilite la documentation de certaines méthodologies, en particulier quantitatives. Son pouvoir authentificateur réside dans son caractère dynamique : il permet par exemple de faire varier un pa-

ramètre dans un dispositif expérimental pour modifier les graphiques qui en découlent, voire d'éprouver tout un processus en remplaçant les données initiales par d'autres. Ce faisant, il vise une recherche reproductible, répliquable, robuste et généralisable.

En termes documentaires, la vérifiabilité de la preuve est ici conditionnée par l'opérationnalité du dispositif, qui dépend elle-même de plusieurs choses. La reproductibilité à l'identique nécessite d'être vigilant aux dépendances logicielles, c'est-à-dire l'utilisation de ressources extérieures, appelées bibliothèques de code. Ces outils doivent être identifiés dans la version exacte utilisée dans le *notebook*, au cas où un développement ultérieur introduirait des différences qui rendraient leur utilisation incompatible. Une bibliothèque peut être téléchargée et intégrée dans le document pour se prémunir en partie de ce risque. L'utilisation de nouvelles données, quant à elle, nécessite de les structurer de la même manière que les données initiales. Une telle opération sollicite plusieurs niveaux d'identification : celle des faits auxquels les données renvoient, des modes d'organisation de ces données – qui peuvent être conventionnels ou idiosyncratiques – et le format de structuration utilisé (CSV, JSON, YAML, etc.).

Synthèse

Plusieurs éléments émergent de la comparaison entre les deux exemples précédents. D'abord, les similitudes de traitement documentaire, notamment en ce qui concerne l'importance des métadonnées d'identification. La référence au numéro de version d'une bibliothèque de code utilisée dans un *notebook* n'est pas si éloignée de la référence à l'ISBN d'un ouvrage, du moins conceptuellement. Les initiatives permettant l'archivage logiciel et la citation des codes sources renforcent encore le parallèle entre ressources bibliographiques et dépendances logicielles. Nous pourrions aller jusqu'à renverser l'imitation, et parler de dépendances rhétoriques pour les documents vers lesquels pointent nos citations ; choisissons d'y voir une passerelle entre différentes pratiques de l'écriture contemporaine.

Dans les deux exemples on peut relever une nécessité de consolider la preuve, de constituer une **assurance** face à une actualisation incertaine. Cela peut être une simple stratégie de copie dans le cas de l'inclusion d'une dépendance directement dans un *notebook*. C'est impossible pour une citation : l'extraction de fragments doit être redoublée par l'annotation, et celle-ci dépend d'un écosystème informationnel, d'une bonne gestion des *connaissables*, qui permet

leur transformation en connaissances. Dans les deux domaines, ces logiques sont hyperdocumentaires et hypertextuelles, et nécessitent une immense coopération.

Ces connaissances, justement, ne nous sont pas seulement accessibles par la simple existence d'une relation d'indexicalité – encore faut-il que celle-ci soit intelligible. Quelle est la signification de `[@cites_as_authority:goody1986]`? Elle n'est pas complètement « donnée » ; pourtant, ces métadonnées constituent une donnée, un matériau potentiel de recherche. Ce qui est utile aujourd'hui pour une étude scientométrique en sciences naturelles constituera peut-être demain un outil de dialogue critique entre sciences humaines. Les futures archives numériques des chercheurs comporteront sûrement ces nouvelles médiations de la preuve, parmi d'autres à venir. Nous pouvons les anticiper en étant attentifs à notre équipement intellectuel, à l'outillage de notre expérience, démarche sans doute moins mathématique que sémiotique. Birger Hjørland nous rappelle en effet que les sciences de l'information se sont globalement éloignées du modèle mathématique de la communication pour se rapprocher des théories du langage et de la signification sociale⁴³. Entre l'actualisation de la preuve et nous, ce sont des jeux de langage qui s'interposent, qui participent à ces dérivations par rapport à la réalité dont parlait Paul Otlet. La signification langagière est importante dans un modèle documentaire, qu'il s'agisse de langue naturelle ou de langages formels (documentaires, informatiques). Nous devons donc avoir le souci de ce plurilinguisme.

Dans cette perspective, un travail sur la médiation nous attend. La preuve documentaire est un mode d'expérience de l'information. Et parler d'information comme expérience ne suppose pas l'évacuation de la médiation, mais bien la décomposition de l'expérience en médiations. Le document y conserve une place importante, car il permet d'articuler des approches structurelles et communicationnelles, dont la convergence en SIC a été soulignée notamment par Évelyne Broudoux⁴⁴. Cette évolution est concomitante d'un rejet par le droit du critère formel comme constituant de la preuve. On peut y voir le signe que cette dernière est, pour ce qui est du domaine documentaire, une fonction résolument plurivoque.

⁴³ Hjørland, « Information Science and Its Core Concepts », 2014, p. 229.

⁴⁴ Broudoux, « Contours du document numérique connecté », 2015, p. 1.

Références

- BENMAKHLOUF, Ali. « Vérification ». Dans : Lecourt, Dominique (dir.), *Dictionnaire d'histoire et philosophie des sciences*. 4e éd., rev. et augmentée. Paris : PUF, 2006, p. 1125-1128. Quadrige. Dicos poche. ISBN 978-2-13-054499-9.
- BERNERS-LEE, Tim. *Semantic Web on XML*. Washington DC, 2000. <https://www.w3.org/2000/Talks/1206-xml2k-tbl/slide10-0.html>.
- BRIET, Suzanne. *Qu'est-ce que la documentation ?* Paris : Éditions documentaires, industrielles et techniques, 1951.
- BROUDOUX, Evelyne. « Contours du document numérique connecté ». Dans : Paganelli, Céline, Chaudiron, Stéphane et Zreik, Khaldoun (dir.), *Documents et dispositifs à l'ère post-numérique*. Paris : Europa, 2015, p. 7-15. https://archivesic.ccsd.cnrs.fr/sic_01327851.
- BUCKLAND, Michael. « Information as thing ». *Journal of the American Society for Information Science*. 1991, Vol. 42, n° 5, p. 351-360. [https://doi.org/10.1002/\(SICI\)1097-4571\(199106\)42:5%3C351::AID-ASI5%3E3.0.CO;2-3](https://doi.org/10.1002/(SICI)1097-4571(199106)42:5%3C351::AID-ASI5%3E3.0.CO;2-3).
- BUCKLAND, Michael. « What is a document? » *Journal of the American Society for Information Science*. 1997, Vol. 48, n° 9, p. 804-809. [https://doi.org/10.1002/\(SICI\)1097-4571\(199709\)48:9%3C804::AID-ASI5%3E3.0.CO;2-V](https://doi.org/10.1002/(SICI)1097-4571(199709)48:9%3C804::AID-ASI5%3E3.0.CO;2-V).
- CHABIN, Marie-Anne. « Document trace et document source. La technologie numérique change-t-elle la notion de document ? » *Revue I3 - Information Interaction Intelligence*. 2004, Vol. 4, n° 1, p. 141-157. https://archivesic.ccsd.cnrs.fr/sic_00001020/.
- COTTE, Dominique. « Le concept de document numérique ». *Communication & Langages*. 2004, Vol. 140, n° 1, p. 31-41. <https://doi.org/10.3406/colan.2004.3265>.
- DAY, Ronald E. « All that is the case: documents and Indexicality. » *Scire*. 2016, Vol. 22, n° 1, p. 57-63. <http://www.ibersid.eu/ojs/index.php/scire/article/view/4304/3832>.
- DAY, Ronald E. *Documentarity: Evidence, Ontology, and Inscription*. Cambridge : MIT Press, 2019. ISBN 978-0-262-04320-5.
- DIONE, Albert Ndiack. *Les aspects juridiques de la dématérialisation des documents du commerce maritime*. Thèse de doctorat. Paris : Université Panthéon-Sorbonne - Paris I, 2018. <https://tel.archives-ouvertes.fr/tel-02060713>.
- DRUCKER, Johanna. *Graphesis: visual forms of knowledge production*. Cambridge : Harvard University Press, 2014. MetaLABprojects. ISBN 978-0-674-72493-8.
- DRUCKER, Johanna. « Humanities Approaches to Graphical Display ». *Digital Humanities Quarterly*. 2011, Vol. 5, n° 1. <http://digitalhumanities.org/dhq/vol/5/1/000091/000091.html>.
- EPRON, Benoît, PINÈDE, Nathalie et TONA, Agnieszka. « Introduction ». *Balisages*. 2020, n° 1, p. 1-5. <https://doi.org/10.35562/balisages.352>.
- EPRON, Benoît et VITALI-ROSATI, Marcello. *L'édition à l'ère numérique*. Paris : La Découverte, 2018. Collection Repères. ISBN 978-2-7071-9935-5.
- FOUCAULT, Michel. *Surveiller et punir: naissance de la prison*. Paris : Gallimard, 1993. ISBN 978-2-07-072968-5.

- GIL, Fernando. « Preuve ». Dans : Lecourt, Dominique (dir.), *Dictionnaire d'histoire et philosophie des sciences*. 4e éd., rev. et augmentée. Paris : PUF, 2006. Quadrige. Dicos poche. ISBN 978-2-13-054499-9.
- GOODY, Jack. *La logique de l'écriture : aux origines des sociétés humaines*. Paris : A. Colin, 2018. ISBN 978-2-200-62114-8.
- GORICHANAZ, Timothy J. *Understanding Self-Documentation*. Thèse de doctorat. Philadelphia : Drexel University, 2018. <http://timgorichanaz.com/pdf/gorichanaz-dissertation.pdf>.
- HJØRLAND, Birger. « Information Science and Its Core Concepts: Levels of Disagreement ». Dans : Ibekwe-SanJuan, Fidelia et Dousa, Thomas (dir.), *Theories of Information, Communication and Knowledge*. [S. l.] : Springer, 2014, p. 205-235. ISBN 978-94-007-6972-4. https://doi.org/10.1007/978-94-007-6973-1_9.
- JEANNERET, Yves et PATRIN-LECLÈRE, Valérie. « La métaphore du contrat ». *Hermès, La Revue*. 2004, n° 38, p. 133-140. <http://www.cairn.info/journal-hermes-la-revue-2004-1-page-133.htm>.
- LALANDE, André. *Vocabulaire technique et critique de la philosophie*. 18e éd. Paris : PUF, 2010. ISBN 978-2-13-058582-4.
- LATOURET, Bruno. « Pensée retenue, pensée distribuée ». Dans : Jacob, Christian (dir.), *Lieux de savoir*. Paris : Albin Michel, 2007, p. 605-615.
- LE DEUFF, Olivier. *La culture de l'information en reformation*. Thèse de doctorat. [S. l.] : Université Rennes 2, 2009. <https://tel.archives-ouvertes.fr/tel-00421928/document>.
- LELEU-MERVIEL, Sylvie. « Le sens aux interstices, émergence de reliances complexes ». Dans : *Colloque international francophone Complexité 2010*. Lille, 2010. <https://hal.archives-ouvertes.fr/hal-00526508>.
- MEYRIAT, Jean. « La bibliologie parmi les sciences de l'information et de la communication ». Dans : *Les sciences de l'écrit. Encyclopédie internationale de bibliologie*. Paris : Retz, 1993, p. 326-335. Les Encyclopédies du savoir moderne. ISBN 2-7256-1428-7.
- MÜLLER, Bertrand. « À la recherche des archives de la recherche. Problèmes de sens et enjeux scientifiques ». *Genèses*. 2006, Vol. 63, n° 2, p. 4-24. <https://doi.org/10.3917/gen.063.04>.
- MÜLLER, Bertrand. « De l'archive au document. Remarques sur l'évolution des régimes documentaires entre le XIXe et le XXIe siècle ». *Territoires contemporains*. 2011, Vol. 2011, n° 2, p. 1-10. <https://archive-ouverte.unige.ch/unige:25440>.
- OTLET, Paul. *Traité de documentation. Le livre sur le livre*. Bruxelles : Les Impressions nouvelles, 2015. ISBN 978-2-87449-299-0.
- PÉDAUQUE, Roger T. *Document : forme, signe et médium, les re-formulations du numérique* [manuscrit]. 2003. https://archivesic.ccsd.cnrs.fr/sic_00000511.
- PÉDAUQUE, Roger T. *La redocumentarisation du monde*. Toulouse : Cépaduès-Éd, 2007. ISBN 978-2-85428-728-8.
- PERRET, Arthur et LE DEUFF, Olivier. « Documentarité et données, instrumentation d'un concept ». Dans : *12ème Colloque international d'ISKO-France : Données et mégadonnées ouvertes en SHS : de nouveaux enjeux pour l'état et l'organisation des connaissances ?* 2019. <https://hal.archives-ouvertes.fr/hal-02307039>.

- POMERANTZ, Jeffrey. *Metadata*. Cambridge : MIT Press, 2015. Essential Knowledge. ISBN 978-0-262-52851-1.
- ROBERT, Pascal et PINÈDE, Nathalie. « Le document numérique : un nouvel équipement politique de la mémoire sociale ? » *Communication et organisation*. 2012, n° 42, p. 191-202. <https://doi.org/10.4000/communicationorganisation.3948>.
- SALAÜN, Jean-Michel. *Vu, lu, su: les architectes de l'information face à l'oligopole du Web*. Paris : La Découverte, 2012. Cahiers libres. ISBN 978-2-7071-7135-1.
- SHANNON, Claude E. « A Mathematical Theory of Communication ». *Bell System Technical Journal*. 1948, Vol. 27, n° 3, p. 379-423. <https://doi.org/10.1002/j.1538-7305.1948.tb01338.x>.
- SHOTTON, David. « CiTO, the Citation Typing Ontology ». *Journal of Biomedical Semantics*. 2010, Vol. 1, n° S6, p. 1-18. <https://doi.org/10.1186/2041-1480-1-S1-S6>.
- SOKOLOV, Arkady V. « The epistemology of documents (A methodological essay) ». *Automatic Documentation and Mathematical Linguistics*. 2009, Vol. 43, n° 2, p. 57-68. <https://doi.org/10.3103/S0005105509020010>.
- ZELLER, Jean-Daniel. « Documents numériques: à la recherche d'une typologie perdue... » *Document numérique*. 2004, Vol. 8, n° 2, p. 101-116. <https://doi.org/10.3166/dn.8.2.101-116>.
- ZINS, Chaim. « Conceptual approaches for defining data, information, and knowledge ». *Journal of the American Society for Information Science and Technology*. 2007, Vol. 58, n° 4, p. 479-493. <https://doi.org/10.1002/asi.20508>.