

Agnieszka SMOLCZEWSKA TONA

**ENSSIB - École nationale supérieure des sciences de l'information et des bibliothèques,
Villeurbanne, France**

« SAUTER DANS L'INCONNU » : DE L'APPROCHE INFO-COMMUNICATIONNELLE DES TEMPORALITES DE LA DONNEE (Article)

Abstract or Résumé:

La présente communication se propose de rendre compte de la construction et de la trajectoire d'un objet de recherche à la fois original, stimulant intellectuellement et complexe. Il est question, en effet, dans les lignes qui suivent, des rapports qui peuvent exister entre deux notions - les données et le temps -, qui l'une comme l'autre sont loin d'être univoques. Nous ne traitons pas toutefois directement de tels rapports dans le cadre de cette communication. Nous choisissons plutôt de retracer la genèse de cette problématique et de sa reformulation dans nos travaux de recherche, et de présenter les différentes étapes du double cheminement, épistémologique et méthodologique, qui a balisé notre expérience de construction de cet objet de recherche. Par la même occasion, nous y aborderons quelques questions et difficultés qui se sont présentées au cours de ce processus intellectuel. Questions et difficultés qui représentent autant de sentiers possibles, sinueux et incertains, que de sauts, en avant, mais aussi en arrière ...

1. « Sauter sur l'occasion » ... ou, une invitation à la réflexion sur notre propre pratique de recherche

« *Des sauts et des trébuchements* » ... le thème même de la présente édition du congrès ACSI/CAIS apparaît comme une invitation. De surcroît, une double initiation à l'exercice de la réflexivité. Car il y a là, d'abord, un appel à penser nos problématiques et objets de recherche en termes de « sauts », à la fois : « *conceptuels, empiriques, théoriques, appliqués et/ou méthodologiques* ». Mais c'est aussi une invitation à appréhender nos conceptions et positionnements scientifiques et professionnels sous un angle, pour une fois, plus personnel. Il s'agit, en effet, d'envisager le processus de construction de nos objets scientifiques tant en termes de réussites que d'échecs, et notamment de difficultés et obstacles rencontrés, surmontés ou non, en privilégiant une réflexion sur « *des trébuchements, des chutes et des guérisons en recherche.* » (ACSI/CAIS 2024).

A ce dernier propos, il n'est pas indifférent de remarquer, que la langue française elle-même n'est pas avare d'expressions ou locutions faisant appel aux notions que l'appel à communication suscitée propose de croiser. Prenons, pour exemples, « *saut de puce* » et « *saut de géant* », deux expressions aux significations opposées¹. Les deux, employées au sens figuré, désignent métonymiquement une avancée sur un sujet ou dans une situation donnée. Cela dit, « *saut de puce* » indique une avancée si négligeable et insignifiante qu'on la compare, injustement¹, à un

¹ Qui possèdent en français un sens à la fois littéral et figuré.

saut de ce petit parasite piqueur, connu aussi, et à juste titre, pour ses sauts en hauteur et longueur. A l'opposé, « *un saut de géant* » s'emploie pour caractériser un progrès tout à fait considérable, si majeur qu'il n'est pas excessif de le rapprocher à un très grand pas, voire à un saut en avant, comparable à celui d'un géant. D'autres exemples de telles expressions idiomatiques françaises pourraient être encore citées ici, comme « *sauter aux yeux* », « *reculer pour mieux sauter* », « *trébucher dans un tapis, contre une pierre* », etc. Cela nous éloignerait toutefois du propos principal de cet article : rendre compte de la construction et de la trajectoire de l'objet de recherche annoncé dans son titre, et des défis théoriques et pratiques qui se sont présentés au cours de ce processus intellectuel. Car s'intéresser au sujet des temporalités de la donnée revient à se confronter à un nombre de questionnements épistémologiques et méthodologiques, qui représentent autant de sentiers possibles, souvent sinueux et incertains, que de sauts, en avant, mais aussi en arrière...

Cela dit, je me servirai dans ce qui suit de quelques-unes de ces expressions françaises forgées autour de l'idée de sauter et trébucher. D'abord, pour structurer mes propos, mais aussi, pour éclairer par leur biais les différentes étapes du double cheminement, épistémologique et méthodologique, qui a balisé mon expérience de construction de cet objet de recherche. Et pour commencer, il sera question de « *faire le saut* », qui signifie, au sens figuré, prendre une décision ou une résolution importante devant laquelle on a longtemps hésité.

2. « *Sauter le pas* » ... ou, de quelques considérations sur la genèse d'un objet de recherche

C'est un lieu commun de proclamer qu'aujourd'hui les données sont partout, et qu'à ce titre, elles nous concernent et affectent toutes et tous. D'innombrables exemples pourraient être invoqués pour illustrer cette assertion, à commencer par le phénomène des fuites massives de données (ou *Data Breaches*), dont la presse se fait régulièrement l'échoⁱⁱ.

Si ce type d'affaires, récurrentes depuis plusieurs années, remet les données sur le devant de la scène médiatique, il n'en demeure pas moins que ces dernières (qu'elles soient personnelles, publiques, ouvertes, massives, de recherche, etc.), sont indéniablement dans l'air du temps. Elles le sont surtout depuis les progrès accomplis au cours de ces dernières décennies qui ont conduit à l'essor des nouvelles technologies digitales, dites *data driven*, et à ce que l'on qualifie désormais de « révolution des données » (Kitchin, 2014). Le dernier avatar en date de telles technologies numériques, l'intelligence artificielle, est si intrinsèquement lié aux données, qu'il n'est pas exagéré de dire que c'est d'abord une forme d'« *intelligence des données* »ⁱⁱⁱ, en ce que ces dernières sont à la fois « *la condition pour que l'intelligence artificielle se développe et l'objet sur lequel l'intelligence artificielle intervient* » (Lorraine Maisnier-Boché, 2017, p. 25).

Face à cette omniprésence des données dans notre environnement personnel et professionnel, plusieurs questions émergent spontanément : si les données sont aujourd'hui présentes partout et à tout moment, quels sont finalement leur(s) rapport(s) au temps ? Ce temps qui passe et dont l'horloge compte des secondes, des minutes, des heures, etc. ? Ou pour le dire autrement, quelles représentations du temps et des temporalités véhiculent les données ? Comment les disent-elles, que nous apprennent-elles sur notre présent, notre passé et notre avenir ?

Ces quelques questions me taraudaient régulièrement depuis déjà quelques années, mais de façon peu engageante, sans que j'ose « *sauter le pas* » pour les penser et objectiver scientifiquement, et en faire mon objet de recherche. L'importance, voire l'urgence, de s'atteler à les explorer s'est imposée à moi, comme souvent, fortuitement. Si fortuitement que l'on pourrait même dire qu'elle

m'a littéralement « sauté aux yeux ». C'est arrivé dans un contexte très particulier : celui de la crise sanitaire liée à la pandémie de Covid-19, à travers une expérience personnelle avec un document emblématique de la gestion de cette crise sanitaire – en tout cas en France et en Europe. Je fais référence ici au *pass sanitaire européen*^{iv}, dont l'équivalent le plus proche sur le territoire canadien serait le *passport vaccinal contre la COVID-19*^v.

Il s'avère que mon premier *pass sanitaire européen* m'a été délivré uniquement à mon nom de naissance², et non à mon nom de femme mariée. Et durant de longs mois, et malgré diverses démarches administratives, il était absolument impossible de mettre à jour et de modifier cette donnée. Ce qui n'allait pas tout de même sans me poser quelques problèmes, notamment pour voyager en Europe et en dehors de ses frontières. Car étant mariée depuis longtemps, je ne possède plus aucun titre d'identité en cours de validité établi à mon nom de naissance : le seul nom qui figure sur tous mes papiers (ma carte d'identité, mon passeport et permis de conduire, etc.) est celui de mon époux. Comment alors prouver, entre autres, aux douaniers et aux polices aux frontières, exigeant de voir mon *pass sanitaire* – rappelons-le, obligatoire pour voyager – que celui-ci était bien le mien ?

Si je reviens ici sur cette petite affaire, somme toute personnelle, c'est qu'elle est, à mes yeux, révélatrice de ce que l'on pourrait appeler un « impensé »^{vi} des rapports entre temps et données. Impensé dont il est particulièrement question dans certains discours et pratiques qui définissent aujourd'hui les données. Impensé qu'il convient d'entendre, non pas tant comme un silence ou absence (volontaires ou non) des réflexions sur les rapports entre temps et données, mais comme une approche de cette question sous le prisme de « l'évidence », qu'il n'est visiblement pas utile de questionner et examiner. Impensé qui concerne, dans le cas qui nous intéresse ici, la question de l'immutabilité de nos marqueurs identitaires (tels que nos noms, prénoms, date et lieu de naissance, sexe, etc.) et des données d'identification qui leur correspondent, celles qui nous définissent au quotidien, aux yeux de nos autorités étatiques et de nos administrations, qui servent à vérifier et/ou prouver qui nous sommes. Car il est évident que pour remplir leur fonction d'identification, ces données doivent avoir une certaine consistance temporelle, c'est-à-dire être relativement stables et durables dans le temps, aussi bien à l'échelle du temps social que du temps personnel, celui qui définit notre histoire personnelle. Or, il peut arriver, et c'est ce qui s'est passé avec le « nom de jeune fille », inscrit sur mon premier *pass sanitaire*, que certaines de ces données n'aient pas cette consistance temporelle.

C'est de là qu'ont émergé ces quelques questions posées en introduction de ce papier, dont la principale est la suivante : quels sont alors le(s) rapport(s) des données au temps et notamment, leurs temporalités ? C'est dans ce contexte particulier que s'est cristallisé ce besoin impérieux de me saisir finalement de cette problématique singulière de recherche.

Cela dit, ce que suit n'apporte pas nécessairement de réponse à ces questions, un défi qu'il serait de toute façon déraisonnable de relever dans l'espace d'un seul article. Après ces quelques repères sur la genèse de la problématique au cœur de cet article, nous allons maintenant « *faire un saut* » du côté des défis théoriques et méthodologiques qu'elle soulève. Cela pour montrer comment cette problématique, surgie à partir de cet « objet concret », le *pass sanitaire*, s'est progressivement transformée en « objet de recherche » (Davallon, 2004), à la fois original, très stimulant intellectuellement, mais aussi, bien complexe.

² Dit aussi « nom de jeune fille » ou « nom patronymique », deux expressions désormais bannies du lexique administratif en France.

Et pour en parler, je reprends l'expression « *sauter par-dessus les obstacles* » : une métaphore qui exprime l'idée de ne pas se laisser décourager ou arrêter devant les difficultés ou problèmes rencontrés, mais plutôt de les affronter, voire surpasser, comme des obstacles qui peuvent se dresser sur tout chemin, connu ou inconnu, que l'on emprunte.

3. « *Sauter par-dessus les obstacles* » ... ou, de quelques considérations sur les défis de l'étude des temporalités de la donnée

Si l'étude des rapport(s) qui peuvent exister entre donnée et temps s'avère complexe, c'est parce qu'elle soulève, d'emblée, de véritables difficultés sur le plan théorique et méthodologique. Faute de place, je n'évoquerai ici que celles les plus épineuses à traiter, en les introduisant de manière un peu rapide, mais, j'espère, suffisamment éclairante pour le lecteur. La première d'entre elles, et pas la moindre, est relative à la définition de ce que l'on désigne comme « temps », et de ce que l'on appelle « donnée ».

a. « *Buter sur un obstacle* » ... celui des définitions du « temps » et de « donnée »

Commençons par le temps, dont « *la notion et la nature, qu'elle soit supposée objective ou subjective, cyclique ou continue, dissociable ou non de l'espace, lisse ou granulaire – voire même son existence, sont des questions très disputées et depuis fort longtemps, en philosophie (...), dans les sciences humaines et sociales, si ce n'est finalement dans toutes les disciplines.* » (Lebaud, 2021, p. 37). Les contradictions inhérentes à toutes ces tentatives de trouver une définition unique du temps conduisent certains auteurs, et notamment le sociologue Claude Dubard (2005), à plaider pour un changement de posture épistémologique. Cette nouvelle posture consiste à délaisser une conception uniforme, unique et unificatrice du « temps » au singulier, au profit d'une conception pluraliste et plurielle des temps. Dans cette approche, appréhender le temps, cet objet abstrait et désincarné de la réflexion philosophique, passe par le truchement des « temporalités », ce qui permet de réfléchir le temps en termes d'« *un ensemble d'objets scientifiques concrets, quantitatifs et qualitatifs, différenciés selon des échelles et des significations déterminées.* » (Dubar, 2015, p. 52).

Un constat similaire peut être fait pour les définitions de ce qu'est une donnée (ou que sont des données). De telles définitions ne manquent pas, surtout ces dernières années, propices à ce type de réflexion du fait des phénomènes, entre autres, de massification et d'ouverture des données. Et notamment dans notre discipline, où les données sont devenues aujourd'hui un objet majeur de la réflexion théorique et pratique, au même titre que l'*information*, le *document*, la *connaissance*, etc. Cela dit, qui s'emploie à examiner en détails les définitions qui en sont faites, ne peut qu'être saisi d'étonnement face à l'hétérogénéité des angles d'appréhension de l'objet qu'elles tentent de cerner. Signalons, juste à titre d'exemple, que la donnée est le plus souvent définie, dans un rapport plus ou moins étroit, à travers d'autres concepts emblématiques de notre discipline, comme : *information* (Escarpit, 1991 ; Buckland, 1991 ; Cacaly & al, 2004 ; Bates, 2005 ; Arsenault et Salaün, 2009 ; Floridi, 2005 et 2011 ; Bachimont, 2017), *connaissance* (Buckland, 1991 ; Arsenault et Salaün, 2009), *document* (Pédaque 2006 ; Leleu-Merviel, 2010), fait (Melese, 1979 ; Leleu-Merviel et Useille, 2012), *trace* (Bachimont 2016 ; Bachimont & Verlaet, 2022), etc. Et qu'il n'est pas rare que certaines de ses définitions l'inscrivent dans un rapport de (quasi) synonymie avec d'autres notions concernées, à l'instar de cette définition de l'*information*, proposée par Jean-Michel Salaün et Clément Arsenault, qui utilisent « *le mot "information" comme terme générique pour couvrir les données, l'information et les*

connaissances » (2009, p. 8). Il en va, pour aller à l'essentiel, qu'il n'existe pas encore de définition consensuelle et d'ancrage théorique universel autour du concept de donnée^{vii}.

Pour sortir de cette impasse théorique, adoptons la proposition de travail suivante : considérons que la donnée, à l'instar de tout objet informationnel ou culturel, s'ancre inévitablement dans l'axe du temps³, et permet de cette manière de l'« attraper », de le fixer, individualiser, dénommer, mettre en ordre, mesurer... en un mot, de le représenter. De ce fait, toute donnée peut être interrogée, analysée et caractérisée au prisme de ses (multiples) temporalités, comme l'entend par ailleurs la définition suivante : « *Basically, every piece of data we measure is related and often only meaningful within the context of space and time* » (Aigner & al., 2011, p. 1).

b. « Sauter dans l'inconnu » ... ou, dans les travaux portant sur les données et leurs temporalités

Mais la difficulté la plus inattendue sur laquelle j'ai buté dans ce travail, est étroitement liée à un paradoxe. Un paradoxe que l'on peut se résumer ainsi : alors que la question du temps et des temporalités des données est quasiment omniprésente dans les discours qui représentent les différentes pratiques sociales développées actuellement autour de ces objets, cette question, sans être vraiment évacuée, n'y est presque jamais abordée frontalement, en tant que telle. C'est d'autant plus vrai lorsqu'on s'intéresse à la littérature scientifique en SHS, et notamment dans notre discipline. Et c'est pour cette raison que pour en parler, je reprends l'expression ici « *sauter dans l'inconnu* » qui désigne, métaphoriquement, le fait de s'engager dans un projet sans savoir exactement où il va nous mener, sans pouvoir connaître et prévoir pleinement ses conséquences ou résultats possibles.

Je n'avancerai ici que deux exemples pour illustrer ce propos. Ainsi, il n'est pas rare que le rapport entre le concept de temps et la donnée s'exprime, de manière plus ou moins explicite, dans les définitions mêmes de cette dernière. C'est le cas de la définition citée plus haut, mais aussi de celle que nous propose l'Organisation internationale de normalisation (1993), selon laquelle : « *la donnée : représentation calendaire d'un moment dans le temps* ». Un autre exemple est celui du « cycle de vie » de la donnée (Bal, 2012), un modèle opératoire qui représente la succession des étapes par lesquelles passe la donnée au fil du temps, en commençant par sa création (p.ex. planification, collecte, acquisition, etc.), en passant par son traitement (p.ex. analyse, transformation, etc.), jusqu'à son sort final (p.ex. archivage ou éventuelle destruction). Je crois que nous serons tous d'accord pour dire que dans ce processus chaque opération effectuée sur la donnée et chaque palier inscrivent la donnée dans une nouvelle temporalité.

Et pourtant, interroger les liens entre le temps et des phénomènes et objets informationnels et du savoir est une problématique centrale en Sciences de l'information et de la communication. Plusieurs publications et manifestations scientifiques ces dernières années, témoignent de l'intérêt de ces questionnements pour notre discipline. On peut citer le Congrès annuel de l'ACSI en 2015 à Ottawa, intitulé : « Le temps comme facteur déterminant : Mémoire et participation dans la coordination des personnes et dans l'organisation des données, de l'information et des connaissances », ou le XXe Congrès de la SFSIC en 2016 à Metz, intitulé : « Temps, temporalités et information-communication » (Bonaccorsi et al., 2017 ; Domenget et al., 2017 ; Lamy et al., 2017 ; Lépine et al., 2017). On peut signaler aussi, plus récemment, ce numéro spécial de *Journal of Documentation*, paru en 2021 et intitulé « *Time and Temporality in Library and Information*

³ Et cela, quelle que soit sa nature, sa granularité, ou le contexte dans lequel on l'observe, etc.

Science » (Haider et al., 2021), entièrement consacré aux études des temporalités dans notre discipline. A quoi on peut ajouter, que cette problématique des liens entre le temps et des phénomènes et objets informationnels a été explorée aussi, notamment autour du document (Escarpit, 1976 ; Roisin et Sèdes, 2004 ; Metzger et Lallich, 2004 ; Nanard, 2004 ; Smolczewska Tona et al., 2008 ; Robert, 2010 ; Tricot et al., 2016 ; Gorichanaz, 2016 ; Bachimont, 2017 ; Smolczewska Tona et al., 2018 ; Pittalis, 2019), du Web (Schafer, 2018), de la reconnaissance (Domenget et al., 2015) ou du récit médiatique (Tétu, 2018). Mais ce n'est pas le cas – à ma connaissance en tout cas – en ce qui concerne les données.

On se trouve ainsi dans ce que l'on pourrait appeler une *terra nullius* — un territoire de recherche sans maître, où presque tout reste encore à dire, à (re)chercher, à faire... Et c'est de là aussi qu'a surgi un autre obstacle auquel je me suis aussi heurtée : comment étudier les temporalités de la donnée ?

Car si sur ce territoire de recherche « *presque tout reste encore à dire, à (re)chercher, à faire...* », il faut encore savoir comment le faire ! Je ne vous apprends rien en disant que faire de la recherche scientifique implique d'autres démarches que celle de formalisation de la question de recherche, et « *notamment celles qu'impose le verbe "faire" : trouver les moyens d'une objectivation des questions et des préoccupations, pour pouvoir les étudier (...).* » (Beillerot, 1991, p. 19). Que la démarche de choisir, élaborer et mettre en œuvre une méthode de recherche permettant d'obtenir des réponses à la question posée, est aussi importante que la question elle-même. C'est de ces quelques considérations que découle logiquement la question que je me suis posée ensuite : quel dispositif d'enquête et quelle(s) méthode(s) seraient-ils adaptés, voire le plus ou le mieux adaptés, pour « attraper » les temporalités de la donnée ?

La partie qui suit apporte l'une des réponses possibles à cette question. Elle est placée sous le titre « *Reculer pour mieux sauter* ». Une expression qui signifie que, parfois, pour surmonter une difficulté ou un obstacle, il est nécessaire de faire un pas en arrière, voire de régresser légèrement et momentanément, pour prendre le temps de réfléchir, et de retrouver ainsi l'élan nécessaire pour franchir l'obstacle en question, et ainsi, le surpasser. Dans mon cas, cette prise de recul métaphorique consistait à retourner vers les discours produits autour et sur les données. Mais cette fois-ci, il s'agissait de les appréhender avec un regard moins critique et plus avisé, constamment à la recherche de tout indice discursif, même infime, susceptible de me renseigner sur les rapports pouvant exister entre temps et données.

4. « *Reculer pour mieux sauter* » ... ou, une proposition méthodologique pour « attraper » les temporalités de la donnée

Pourquoi retourner vers ce matériau d'étude en particulier ? C'est d'abord pour une raison déjà évoquée : dans les discours qui traitent de données la question de leurs temporalités est quasiment omniprésente, même si elle l'est de manière paradoxale – presque toujours présente, et presque jamais traitée frontalement et explicitement. Mais ce choix se trouve aussi justifié par les liens épistémologiques qui existent entre les concepts de temps et de discours^{viii}. Pour les résumer à grands traits : rappelons-nous que la principale difficulté de penser le temps vient de ce qu'il n'est pas perceptible directement par nos sens, d'où la nécessité, pour s'en saisir, de le représenter (Chesneaux, 2004). Or, la réponse à cette difficulté « *s'est depuis l'Antiquité orienté[e] vers le langage : l'expérience du temps ne peut se représenter qu'à partir d'une activité langagière (...)* » (Tétu, 2018, p. 14), dont le discours^{ix} est l'une des principales manifestations.

Pour cette étude, mon choix s'est porté essentiellement sur les discours scientifiques^x. Or, ce qui m'a frappée d'emblée dans ces discours, c'est leur tendance à une sur-adjectivation du terme *donnée* ou *data*^{xi}. La plupart du temps, en effet, ce terme n'apparaît pas seul, mais accompagné d'un adjectif, voire de plusieurs, qui définissent les qualités, propriétés, valeurs, etc., de la donnée qu'ils qualifient. *Brute, primaire, secondaire, récente, observée, construite, massive, quotidienne, codée, éphémère, compilée, ouverte, agrégée, transformée, publiée, partagée...* en voilà quelques exemples, parmi tant d'autres, appartenant à ce lexique, pléthorique et hétérogène, des adjectifs qui gravitent autour du mot *donnée*, ou *data*, dans les textes scientifiques en français et anglais.

On aura par ailleurs remarqué que quelques-uns des adjectifs mentionnés ci-dessus possèdent clairement un ancrage temporel (*brute, primaire, secondaire récente, quotidienne, éphémère, actuelle*, etc.) – plus ou moins explicite. D'où l'idée m'est venue, de creuser cette observation en la transformant en piste méthodologique. Dans cet objectif, j'ai constitué un corpus de textes scientifiques, selon ces critères de sélection : des écrits scientifiques (principalement en anglais, mais aussi en français), validés par les pairs et publiés dans des circuits de l'édition scientifique (surtout des articles et chapitres d'ouvrage), écrits par des chercheurs et/ou des praticiens de la donnée. Mais le critère le plus déterminant était le suivant : que dans tous ces textes, les données⁴ soient le sujet principal de la réflexion et de ce fait, fassent l'objet de débats et tentatives de définition, description, typologie, classification, exemplification, etc. Voici, à titre d'exemple, deux documents appartenant à ce corpus, aux titres aussi évocateurs que représentatifs de ses critères de sélection : l'article de Samuelle Carlson et Ben Anderson “ *What Are Data? The Many Kinds of Data and Their Implication for Data Re-Use* ” (2007), et le chapitre de Christine L. Borgman, intitulé « *Qu'est-ce qu'une donnée ?* » (Bogman, 2015)^{xii}.

Pour récupérer dans ce corpus les adjectifs qui m'intéressaient, je me suis appuyée sur le phénomène linguistique de la cooccurrence, et plus spécifiquement de la *collocation*, résumé par ce fameux principe : « *You shall know a word by the company its keeps* » (Firth, 1957, p. 11). En pratique, ceci revient à extraire automatiquement dans les ressources textuelles exploitées, à l'aide d'un outil nommé concordancier^{xiii}, toutes les unités lexicales (dans notre cas, les *collocatifs adjectivaux*) qui apparaissent à droite et à gauche du terme pivot (dans notre cas, du terme *donnée* ou *data*). Le corpus des termes ainsi obtenu, riche de plus de cinq cent adjectifs aux significations les plus diverses, a été ensuite analysé manuellement de façon à y identifier tous les adjectifs qui expriment, plus ou moins explicitement, le(s) rapport(s) de la donnée au temps. Il en résulte un premier ensemble d'une cinquantaine de termes à connotation temporelle, tantôt synonymes (comme donnée *courante* et *active*), tantôt antonymes (comme donnée *synchrone* versus *asynchrone*), tantôt complémentaires (comme donnée *primaire, secondaire* et *tertiaire*).

Ce premier corpus a ensuite été enrichi par d'autres adjectifs, récupérés dans des ressources linguistiques qui font précisément la part belle aux cooccurrences en général, et aux collocations en particulier, à savoir les dictionnaires de cooccurrences et de collocations. A titre d'exemple, les termes comme *brute, disponible, éphémère, historique, périmée*, ont été relevés parmi une trentaine de *collocatifs adjectivaux* du terme *donnée*, répertoriés par le *Dictionnaire des cooccurrences à l'usage des écoles* de Jacques Beauchesne (2004).

Le corpus final, pas très abondant dans l'absolu (une centaine de termes environ), mais suffisamment riche et diversifié pour permettre une approche qualitative des temporalités de données, constitue aujourd'hui mon principal matériau d'étude. Dans ce vocabulaire, que je

⁴ Qu'elles soient scientifiques, personnelles, massives, ouvertes, etc.

qualifie volontairement, d'après Jean-Claude Domenget (2018), de *temporaliste*, on trouve d'abord et sans surprise, des termes plutôt attendus en matière temporalités de la donnée : ceux qui la qualifient, par exemple, de *présente, absente, existante, inexistante, préexistante, disponible, indisponible, journalière, quotidienne, mensuelle, trimestrielle, annuelle, périodique, ponctuelle, régulière, irrégulière, courante, contemporaine, récente, historique, anhistorique, intemporelle*, etc. À cela s'ajoutent de très nombreux adjectifs bien balisés dans le langage car utilisés généralement pour parler de données contemporaines, comme ceux qui les désignent comme *brutes, primaires, secondaires, tertiaires, compilées, agrégées, transformées, partagées, publiées, réutilisables, reproductibles, statiques, dynamiques, synchrones, asynchrones, actives, inactives, en temps réel*, etc. Mais le corpus en question contient aussi quelques adjectifs qui expriment les rapports de la donnée au temps de manière plus inventive et imagée. Citons parmi eux, à titre d'exemple, les termes comme *éphémère, jeune, vieillissante, survivante, morte, lente, rapide, obsolète, froide, chaude* ou *tiède*.

5. Aller vers la conclusion, sans pour autant « sauter à la conclusion » ... ou, de quelques pistes de recherche ouvertes par l'approche qualitative des adjectifs temporalistes de la donnée

En guise de conclusion, quelques éléments de réflexion autour de deux questions qui guident l'approche des temporalités de la donnée présentée ici : quelles potentialités analytiques nous offre un tel corpus d'adjectifs *temporalistes* de la donnée ? Quelles pistes de recherche peut-on concrètement explorer à partir des données recueillies dans le cadre de cette étude ?

La première réponse à ces questions est un préalable nécessaire à toute étude de ce corpus. Pour être capable de le comprendre, analyser, et dans le meilleur des cas, d'en tirer parti, il est indispensable de clarifier et fixer la signification de tous les termes qu'il contient, en dressant, par exemple, leur glossaire. Nous avons vu que certains de ses adjectifs sont métaphoriques voire anthropomorphiques (donnée *vivante*, etc.), beaucoup d'entre eux sont plus ou moins techniques ou spécialisés (donnée *en temps réel* ou *prolective* etc.), d'autres encore, ont clairement une signification ambiguë (donnée *éphémère, dynamique*, etc.). Et si l'on comprend, intuitivement, le sens attribué à la donnée, par exemple, *active* ou *inactive*⁵, que veut-on dire précisément, lorsqu'on qualifie la donnée de *passive*, de *lente*, voire de *tiède* ?

Mais comment explorer concrètement ce corpus des adjectifs *temporalistes*, avec comme visée première d'y appréhender tout ce qui a trait aux temporalités de la donnée ? Deux grandes approches, distinctes mais complémentaires, sont possibles, en tout cas à mes yeux. Chacune d'entre elles ouvre plusieurs pistes de recherche potentiellement fécondes, que j'introduis ici de façon programmatique, puisque cette recherche sur les temporalités de la donnée est toujours en cours.

La première d'entre elles peut se résumer ainsi : confronter la théorie à la réalité du terrain. Pour la problématique qui nous occupe ici, cette approche consiste à réfléchir les temporalités de la donnée à partir des différentes conceptions épistémologiques (concepts, théories, définitions, etc.) sous-jacentes aux notions du temps et de la donnée. Il s'agit alors d'identifier les éléments théoriques signifiants par rapport aux temporalités de la donnée pour les mettre en relation et les confronter aux données empiriques collectées sur le terrain. Données nous permettant, dans le pire des cas, d'approximer des éléments théoriques sélectionnées, et dans le meilleur cas, de les illustrer

⁵ C'est-à-dire, respectivement : utilisée, voire utilisée fréquemment, ou utilisée que peu, voire pas du tout.

ou exemplifier. J'illustrerai ces propos par les temporalités que la donnée semble hériter de différentes conceptions philosophiques du temps. Ces temporalités sont repérables dans notre corpus par les mêmes qualificatifs que ceux qui désignent certaines qualités du temps (ses phénomènes, propriétés, rythmes, etc.), et que l'on retrouve généralement déclinés dans les vocables qui se rapportent au temps, sous forme de couples antinomiques (Gasparini, 1994 ; cité par Grossin, 1996), tels que le temps *long* et le temps *court*, notamment. Dans notre corpus, les temporalités de la donnée directement déduites dans cette première approche, sont entre autres, celles représentées par des couples comme *statique* vs *dynamique*, *synchrone* vs *asynchrone*, *lente* vs *rapide*, *active* vs *inactive*, etc...

La deuxième approche à mobiliser dans cette étude se définit par opposition à la première. On peut la résumer ainsi : chercher à faire émerger des données empiriques rassemblées en corpus, au moins, des possibles pistes de réflexions, et au plus, des définitions d'éléments théoriques sous-jacents aux données analysées. C'est par cette approche particulière que je suis parvenue à identifier quelques temporalités de la donnée, parmi lesquelles celles associées à ce qu'on peut appeler *voyage dans le temps* par le biais des données. Plusieurs adjectifs figurant dans le corpus font penser en effet que les données flottent comme des médiateurs entre passé, présent et avenir, et de ce fait, qu'elles s'apparentent à des moyens permettant de se déplacer dans le temps... Si l'on se fie au corpus des adjectifs analysés, il apparaît que les données deviennent à la fois des points d'ancrage dans le présent (on parle bien de donnée *contemporaine*, *chaude*, *synchrone*, *courante*, etc.), des véhicules pour retourner dans le passé et l'acheminer vers le présent (lorsqu'il est question de donnée par exemple *historique*, *froide*, *asynchrone*, etc.), et même, lorsque notre passé devient le prédicteur de notre futur, des vecteurs de projection vers l'avenir, et de son anticipation (à l'instar de la donnée *prospective*, *visionnaire*, *prédictive*, etc.).

Autant de données empiriques... autant d'indices discursifs... qui nous invitent, à nouveau, à faire des sauts. Et il ne s'agit pas seulement de sauts à travers les temps : entre passé et présent, présent et futur... Mais également et plus particulièrement de sauts « *conceptuels, empiriques, théoriques, appliqués et/ou méthodologiques* » entre nos conceptions théoriques et données empiriques...

References bibliographiques

- AIGNER (W.), MIKSCH, (S.), SCHUMANN, (H.), TOMINSKI, (Ch.), *Visualization of Time-Oriented Data*. Springer, 2011.
- ARSENAULT (C.), SALAUN, (J.-M.), *Introduction. Permanence et changements*. In : Introduction aux sciences de l'information. Montréal : Presses de l'Université de Montréal, 2009. DOI : <https://doi.org/10.4000/books.pum.9713>.
- BACHIMONT, (B.), *Patrimoine et numérique - Technique et politique de la mémoire*. Ina Medias et Humanite Sciences humaines & sociales, 2017.
- BACHIMONT, (B.), VERLAET (L.) *Traces, données et preuves en contexte numérique : quelles acceptions interdisciplinaires ?*. Revue Intelligibilité du Numérique, 2, 2021.
- BALL (A.), Review of Data Management Lifecycle Models (version 1.0). REDm-MED Project Document redm1rep120110ab10. Bath, UK: University of Bath. 2012.
- BATES, (M.J.), *Information et connaissance : un cadre évolutionniste pour la Science de l'information (Information and knowledge: an evolutionary framework for information science)*,

dans *Information Research*, 10(4), July 2005. Consulté le 1/04/2024. <http://informationr.net/ir/10/4/paper239.html#goo91>

- BEAUCHESNE, (J.), Dictionnaire des cooccurrences à l'usage des écoles, Guérin, 2004.
- BEILLEROT, (J.), (1991), La « recherche », essai d'analyse. *Recherche et formation*, 9, 17-31. En ligne : <http://www.persee.fr/doc/refor_0988-1824_1991_num_9_1_1040>.
- BONACCORSI (J.), COLLET (L.), RAICHVARG (D) (dir). *Les temps des arts et des cultures*. Paris, L'Harmattan, coll. SFSIC, 2017.
- BORGMAN (C.L.), *Big Data, Little Data, No Data: Scholarship in the Networked World*. The MIT Press, 2014.
- BUCKLAND, (M. K.), *Information as Thing*. In : Journal of the American Society for Information Science. 1991.
- CACALY (S.), LE COADIC (Y.-F.), POMART (P.-D.), SUTTER (É.), *Dictionnaire de l'information*. 2e. Paris : Armand Colin, 2006.
- CHESNEAUX (J), « Cinq propositions pour appréhender le temps », *Temporalités* [En ligne], 1 | 2004, consulté le 23 octobre 2019. DOI : 10.4000/temporalites.648
- DAVALLON, (J.). *Objet concret, objet scientifique, objet de recherche*. Hermès. 2004. Vol. 38, n° 1, pp. 30.
- DOMENGET (J-C.), LARROCHE (V.) et PEYRELONG (M-F.) (dir.), *Reconnaissance et temporalités. Une approche info-communicationnelle*, Paris, L'Harmattan, coll. Communication et Civilisation, 2015.
- DOMENGET (J-C.), MIEGE (B.), PELISSIER (N.) (dir), *Le défi des temporalités : conceptualisation et approches méthodologiques en SIC*, SFSIC Congrès de Metz, vol. IV, Paris, L'Harmattan, 2017. <https://hal.archives-ouvertes.fr/hal-02418062>
- DOMENGET (J-C.), « Quelques principes pour analyser les temporalités du Web », dans SCHAFER (V.), *Temps et temporalités du Web*. Presses universitaires de Paris Nanterre, 2018.
- DUBAR, (C.), « Du temps aux temporalités : pour une conceptualisation multidisciplinaire », *Temporalités*, 20, 2014. DOI : <https://doi.org/10.4000/temporalites.2942>
- ESCARPIT (R.), *Théorie générale de l'information et de la communication*, Paris, Hachette Université, 1976 rééd., 1991.
- FLORIDI (L.), « Semantic Conceptions of Information », In *Stanford Encyclopaedia of Philosophy*, 2005. Consulté le 1/07/2020. <http://plato.stanford.edu/entries/information-semantic/>
- GORICHANAZ, (T.), “Documents and Time” dans *Proceedings from the Document Academy*, 3(1), 7, 2016.
- HAIDER (J.), JOHANSSON (V.), HAMMARFELT (B.), *Time and temporality in library and information science*. Journal of Documentation 78(1):1-17, January 2022. Disponible : <https://www.emerald.com/insight/publication/issn/0022-0418/vol/78/iss/1>
- KITCHIN (R.), “The Data Revolution. Big Data, Open Data, Data Infrastructures and Their Consequences, Londres, Sage, 2014, 240 p.

- LAMY (A.), CARRE (D.) (dir), *Temps, temporalité(s) et dispositifs de médiation*. Paris, L'Harmattan, coll. SFSIC, 2017, 166 pages.
- LAUNET, (E.), *Saut de la puce*. Libération, 2007. Disponible : https://www.liberation.fr/cahier-special/2007/08/09/saut-de-la-puce_99621/
- LEBAUD, (D.), *Notion de temps, activité de langage et linguistique*. In (ed.) Abdoul-Carime (N.) Éric Bourdonneau (E.), Mikaelian (G.), Thach (J.), *Temporalités khmères*, Edited Collection, 2021.
- LELEU-MERVIEL (S.), USEILLE (P.), « Quelques révisions du concept d'information ». Dans. PAPY (P.) (dir.), *Problématiques émergentes dans les sciences de l'information*, Paris, Lavoisier, 2008, p. 25-56. Consulté le 1/04/2024. <https://hal.archives-ouvertes.fr/hal-00695777/document>
- LELEU-MERVIEL (S.), *La traque informationnelle*. Istes éditions. Collection : Sciences, société et nouvelles technologies. 2017.
- LEPINE (V.), ALEMANNO (S.), LE MOËNNE (C.) (dir), *Communications & organisations. Accélération temporelle*. Paris, L'Harmattan, coll. SFSIC, 2017, 202 pages.
- MAISNIER-BOCHE (L.), Intelligence artificielle et données de santé, *Journal du Droit de la Santé et de l'Assurance - Maladie (JDSAM)*, vol. 17, no. 3, 2017, pp. 25-29.
- MELESE (J.), *Approches systémiques des organisations*. Editions Hommes et Techniques, Suresnes, 1979.
- METZGER (J-P.), LALLICH-BOIDIN (G.), « Temps et documents numériques », dans *Document Numérique*, 2004 – ROISIN (C.), SEDES (F.), « Temps et documents », dans *Document numérique* 2004/4 (Vol. 8)
- NANARD (J.), « Formalismes de manipulation du temps par l'auteur dans les documents multimédias », dans *Document numérique* (vol.8) 2004/4.
- PEDAUQUE (R.T.), *La redocumentarisation du monde*. Toulouse : Cepaduès éditions. 2007.
- PITTALIS (S.), « Éléments d'analyse figurative, thématique et axiologique des représentations et visualisations numériques de la temporalité », dans *Signata* (vol.10) 2019.
- SMOLCZEWSKA TONA (A.), CORDONNIER (S.), « Faire advenir le territoire dans l'enquête. Rencontre des compétences, partage des expériences et négociation des positions savantes et profanes », dans BONACCORSI (J.), CORDONNIER (S.) (dir.), *Territoires. Enquête communicationnelle*, Éditions des Archives Contemporaines, 2018.
- SMOLCZEWSKA TONA (A.), « Recontextualiser des fonds patrimoniaux numérisés de la presse régionale à travers la valorisation des dimensions temporelles » dans *Actes SFSIC'08*. Tunis, Tunisie, 2008.
- SMOLCZEWSKA TONA (A.), LANDRON (P-Y.) et LALLICH-BOIDIN (G.), « Recontextualiser des fonds patrimoniaux numérisés de la presse régionale à travers la valorisation des dimensions temporelles » dans *Actes SFSIC'08*. Tunis, Tunisie, 2008.
- SMOLCZEWSKA TONA (A.), « *Enquête sur les représentations discursives des temporalités de la donnée à l'oeuvre dans des articles scientifiques* », , Humains et données : création, médiation,

décision, narration. Actes du colloque "Document numérique et société", Nancy, octobre 2020, sous la direction de Simonnot Brigitte, Broudoux Évelyne, Chartron Ghislaine. De Boeck Supérieur, 2021, pp. 137-150.

- ROBERT (P.), *Mnémotechnologies, Une théorie générale critique des technologies intellectuelles*, Hermes, 2010.

- ROBERT (P.), *L'impensé numérique. L'impensé numérique*, Archives contemporaines, 2017.

- ROISIN (C.), SEDES (F.) (dir). « Temps et documents », dans *Document numérique* (vol.8) 2004/4.

- SCHAFER (V.) (dir). *Temps et temporalités du Web*. Presses universitaires de Paris Nanterre, 2018.

- TETU (J-F.), *Le récit médiatique et le temps*. Paris, L'Harmattan, 2018.

- TRICOT (A.), SAHUT (G.), LEMARIE (J.), *Le Document : communication et mémoire*. Louvain-la-Neuve, De Boeck Supérieur, coll. Information & stratégie, 2016.

ⁱ Tout à fait injustement, car la puce est en réalité une excellente sauteuse ! Alors qu'elle ne mesure qu'environ 5 millimètres, il paraît qu'elle est capable de bondir à peu près de 400 fois sa propre hauteur, et de sauter d'un coup une longueur d'environ 50 centimètres (Launet, 2007) – ce qui correspond à 100 fois sa propre longueur.

ⁱⁱ Un tel fait d'actualité agite la société française au moment où j'écris ces lignes (mars 2024) : la possible fuite des données personnelles suite à une cyberattaque, qui touche, potentiellement, pas moins que 43 millions de personnes inscrites, ou ayant été inscrites, à France Travail, le principal service public de l'emploi et de l'insertion professionnelle en France. Pour plus de détails, voir : Le Monde. *Cyberattaque de France Travail : les données de 43 millions de personnes ont potentiellement été exfiltrées*. [Consulté le 30/04/2024]. Disponible à l'adresse : https://www.lemonde.fr/pixels/article/2024/03/13/france-travail-victime-d-une-cyberattaque-les-donnees-de-43-millions-de-personnes-menacees_6221831_4408996.html

ⁱⁱⁱ Je reprends ici cette expression alors qu'elle recouvre aujourd'hui une réalité plus générale, à savoir, l'ensemble des méthodes et processus contribuant à la collecte et à l'analyse des données – afin de dégager et d'identifier des tendances-clés et des informations exploitables pour comprendre un marché ou un écosystème

^{iv} Je suppose qu'il est nécessaire d'expliquer au lecteur, surtout s'il n'est pas européen, ce qu'était ce *pass sanitaire européen*. En France et dans l'Union européenne, ce document, appelé aussi le *certificat COVID numérique de l'Union Européenne*, centralisait les principales données de santé des citoyens européens relatives à la Covid-19. Il attestait que son possesseur était, soit vacciné contre la Covid-19, soit négatif à la Covid-19, soit rétabli de la Covid-19. Durant la crise de la Covid-19, le *pass sanitaire européen* était obligatoire : en France, pour pouvoir se rendre, entre autres, aux cinémas, aux théâtres, aux festivals, etc., et en Europe, pour pouvoir circuler entre les pays de l'Union européenne et de l'espace Schengen.

^v Introduit également à la même époque, dans certaines provinces canadiennes si l'on se fie aux informations encore disponibles sur le site *Canada.ca – Le site officiel du gouvernement du Canada. COVID-19 : Preuve de vaccination au Canada*. [Consulté le 30/04/2024]. Disponible à l'adresse : <https://www.canada.ca/fr/sante-publique/services/maladies/maladie-coronavirus-covid-19/vaccins/preuve-vaccinale.html>

^{vi} Je reprends volontairement la notion d'« impensé » telle qu'elle est définie, entre autres, dans les travaux de Pascal Robert (2017) à propos de l'informatique et du numérique où l'impensé « *n'est pas tant un défaut de pensée qu'un ensemble de stratégies discursives et pratiques qui visent à présenter l'informatique puis le numérique sous le prisme de « l'évidence »* » (Robert, 2017).

^{vii} Ce que constate par ailleurs Christine L. Borgman, par ces lignes écrites il y a déjà quelques années : « *Now in its fifth century of use, the term data has yet to acquire a consensus definition. It is not a pure concept nor are data natural objects with an essence of their own. (...) Lists of entities that could be considered data are unsatisfactory as definitions, yet such "definitions" abound in the scholarly literature.* » (Borgman, 2015, p. 28-29).

^{viii} Liens dont j'ai eu l'occasion de rendre compte davantage dans une autre publication (Smolczewska Tona, 2021).

^{ix} Entendu ici au sens large, comme une mise en pratique du langage dans un acte de communication tant à l'écrit qu'à l'oral.

^x Cela pour travailler avec des représentations des temporalités de données construites selon une approche scientifique, qui se veut, *a priori*, rationnelle et neutre.

^{xi} Cela en fonction de la langue dans laquelle le texte était écrit bien évidemment : *donnée* pour les textes en français, et *data* pour ceux en anglais et en français, car le terme *data* est désormais fréquemment utilisé, aussi, en français.

^{xii} Paru initialement en anglais, avant sa traduction en français.

^{xiii} Dans cette étude, il s'agit du concordancier *WordSmith Tools*.