

# INTRODUCTION

## Chercher !

« Faire parler les données ! ». Je m'adresse aux chercheurs<sup>1</sup>, chargés d'études<sup>2</sup>, analystes, étudiants dont le travail est de produire des connaissances et de les communiquer. Je m'adresse aussi aux enseignants ou formateurs qui les préparent à ce travail.

Les données sont les textes, articles, livres ou documents, les nombres, comptables, économiques ou statistiques, les conversations, paroles entendues ou recueillies, les images fixes ou animées, les choses vues ou observées. Ces données nous sont extérieures, nous les abordons avec notre culture, nos connaissances, nos convictions, nos méthodes et ambitions propres. Professionnels ou apprentis du travail d'étude et de recherche, vous travaillez sur des données en mobilisant vos connaissances pour en produire de nouvelles, des connaissances pour comprendre le monde ou pour agir. Ces connaissances doivent être, nouvelles ou non, disponibles, sinon comment justifier le coût de leur production ?

« Faire parler les données », le domaine que nous allons aborder est loin d'être homogène. Il se caractérise par une grande diver-

---

1. J'appellerai chercheurs les membres du monde académique ou des grands instituts de recherche privée ou publique disposant d'une certaine liberté de recherche.

2. J'appellerai chargés d'études les salariés d'une administration ou d'une entreprise industrielle, de commerce ou de service, qui travaillent pour un client en répondant à une demande et un cahier des charges précis. On les appelle aussi chargés de projets.

sité de pratiques, milieux professionnels, approches, outils et perspectives, mais la création de connaissances fait son unité.

Chercheurs en sciences humaines et chargés d'études, les premiers s'intéressent à des connaissances générales, applicables dans des situations très diverses et enseignables, les seconds répondent à une demande située et précise. Le chercheur privilégie la compréhension, le chargé d'études l'action et la réaction. Ils ont en commun la curiosité, la volonté d'observer, mais aussi le goût pour l'analyse, la réflexion. Ils sont capables d'esprit critique autant qu'ils cherchent à être compris.

Les uns, s'inscrivent dans une tradition littéraire ou ethnologique. Ils privilégient l'observation directe, les entretiens, l'analyse de documents, la construction progressive du sens. Les autres cherchent à mesurer, à quantifier. Ils construisent des modèles, collectent systématiquement un grand nombre de données, testent des hypothèses et des modèles.

Aussi parle-t-on d'études et d'analyse qualitative ou quantitative. Une recherche<sup>3</sup> sur le Web situe l'écho de ces approches. L'évocation des recherches qualitatives est plus fréquente, mais les références aux méthodes quantitatives sont mieux documentées. Toutes les disciplines des sciences humaines sont concernées, dans tous les domaines, professionnel, de la recherche ou de l'enseignement.

Même si dans certaines disciplines académiques les clivages peuvent être très clairement établis, les frontières sont assez floues. Leur complémentarité a ainsi conduit en 2007 à la création d'une revue spécialisée sur les approches qualitatives et quantitatives : *Journal of Mixed Methods Research*, dont l'objectif est de permettre la publication de recherches ne répondant pas à une stricte orthodoxie qualitative ou quantitative et de promouvoir les approches mixtes.

Néanmoins, dans les milieux professionnels, les spécificités de chaque approche sont encore utilisées pour distinguer les offres de services « quali » et « quanti ».

---

3. Recherche avec Google analytics sur le Web français et anglais et sur les sites d'Amazon et Sage publisher : voir le site.

Ainsi, la recherche qualitative d'essence littéraire reste peu instrumentée. Le microphone et la vidéo bouleversent peu le travail de lecture, d'observation ou d'écoute. Les logiciels de traitement de texte suffisent pour un travail d'analyse qui reste essentiellement dépendant de la culture, des capacités de lecture, de réflexion et d'écriture de l'auteur. Il en va différemment pour la recherche quantitative. Les tâches de collecte ont toujours été beaucoup plus instrumentées et les logiciels indispensables pour traiter les données.

Lecteur et écrivain, ou informaticien et statisticien, les moyens mis en œuvre renforcent les différenciations. Mais tout est en train de changer.

Au cours des 40 dernières années, la révolution numérique a fait passer de la rareté à l'abondance, tout en augmentant considérablement la productivité des tâches de collecte et de traitement des données. Ce fut d'abord la démocratisation de l'accès aux puissances de calcul grâce aux ordinateurs personnels des années 1980. La numérisation des textes et la généralisation des accès numériques suivent dans les années 1990. Avec les moteurs de recherche et les enquêtes en ligne, les sources de données prolifèrent. Ce mouvement est amplifié dans les années 2000 par le Web sémantique, les réseaux sociaux et les mégadonnées<sup>4</sup>. Aujourd'hui, la mobilité et l'intelligence artificielle achèvent de bouleverser les méthodes de travail et de brouiller encore plus les paradigmes. La machine et les algorithmes, des téléphones intelligents aux nuages de données<sup>5</sup>, sont omniprésents, affectant autant la production d'information que son traitement ou sa communication. Avec cette nouvelle donne apparaissent d'autres acteurs. Ils sont spécialistes des technologies que le chercheur ou le chargé d'études doit apprendre à maîtriser pour résister à la concurrence de ces nouveaux venus. De nouveaux termes apparaissent, les « *analytics* » contre la statistique, les « *sentiment analysis* » contre l'analyse de contenu, le « *Big Data* » contre l'échantillonnage...

La figure 1 montre l'évolution, au cours des 10 dernières années, de l'usage de quelques termes emblématiques. Une nouvelle ter-

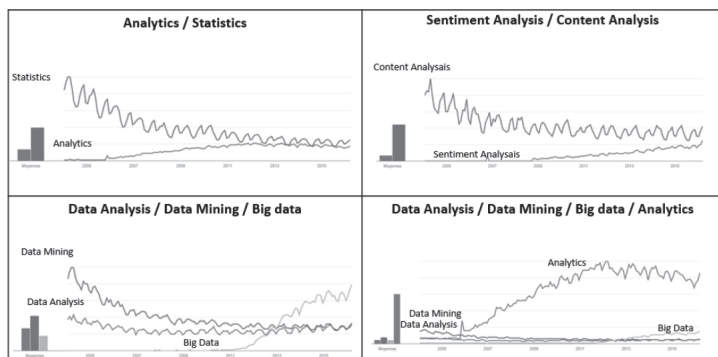
---

4. Ou données massives, termes français pour parler du jargonant *Big Data*.

5. Ou infonuagique pour les Québécois.

minologie s'impose. Elle manifeste l'émergence des nouvelles approches liées aux mutations technologiques. Elles donnent accès au traitement massif de données qualitatives ou quantitatives. Le grand nombre n'est plus le privilège des « quanti », ni l'exploration celui des « quali ». Le « Like » global et subjectif rivalise avec l'indicateur et la mesure précautionneuse.

**Figure 1 : Reflet sur le Web de quelques grandes tendances**



Même si ces évolutions sont aussi le reflet d'une communication au service des nouvelles technologies, elles manifestent l'émergence d'autres manières de faire parler les données. Ce constat correspond bien au diagnostic formulé par Klaus Krippendorff dans l'introduction de la 3<sup>e</sup> édition de son ouvrage sur l'analyse de contenu, publié pour la première fois en 1980. Il écrit : « le volume croissant des données, qui désormais dépasse la taille de ce qu'un analyste peut traiter, requiert des techniques d'une qualité différente. Ces aides informatiques participent autant à l'analyse de contenu que l'analyste humain et posent le problème de leur compréhension ».

Productivité accrue, accélération des tempos, mixité des approches, les technologies ouvrent de nouvelles perspectives. Mais les outils ne sont pas la finalité. Le rôle du chercheur est d'en examiner la logique et les usages, pour apprécier de manière critique leur contribution à la production de connaissances. Cela nécessite un retour aux fondamentaux pour une bonne

compréhension des méthodes et des spécificités du travail scientifique.

C'est l'objet de ce livre. La production de connaissances scientifiques me servira de référence : produire des connaissances nouvelles, situées par rapport aux connaissances existantes. Les fonder sur des données qui en montrent la pertinence et qui permettent de les critiquer. Se conformer aux usages de sa communauté académique ou professionnelle pour privilégier le débat.

Dans la tradition des clivages classiques, j'aborderai ensuite les méthodes quantitatives, puis les approches qualitatives, en situant ces méthodes par leur fondement informaticostatistique<sup>6</sup> et linguistico-sémantique. Je m'attacherai à décrire leurs usages en montrant que ces approches ne peuvent être réduites à la classique opposition entre positivisme et constructivisme et qu'elles sont plus complémentaires que concurrentes. Je m'appuierai enfin sur de nombreux exemples présentés plus en détail dans un site associé, où le lecteur pourra trouver données, outils et modes opératoires : site associé : <https://faireparlerlesdonnees.blog/>

J'aborde cette réflexion fort d'une longue expérience professionnelle de plus de 40 ans comme enseignant-chercheur en sciences de gestion, mais aussi comme acteur des milieux professionnels, dans lesquels je me suis trouvé engagé en tant qu'éditeur de logiciels et prestataire de services.

J'ai commencé ma vie d'universitaire à l'aube du structuralisme triomphant, elle s'achève dans l'hyperspécialisation, la globalisation des réseaux et la post-vérité. Fragmentation des disciplines, globalisation des méthodes et des outils, bouleversements technologiques. Les métiers n'ont pas (encore ?) fondamentalement évolué. Dans le monde académique, la pression à la publication est plus élevée, à l'intérieur de chaque communauté les standards s'imposent plus fortement. Dans l'entreprise, les fonctions d'étude organisées en départements spécialisés sont parfois menacées ou éclatées dans les services opération-

---

6. Ce néologisme pour rappeler que les techniques informatiques formatent la pensée statistique, comme la langue la construction du sens !

nels et confiées à des logiciels. Chez les prestataires de service, la concurrence est plus vive. Partout l'usage des technologies inexistantes il y a 30 ans change la donne en faisant apparaître de nouvelles pratiques et de nouveaux acteurs.

Cela complique l'enseignement des méthodes, la formation des chercheurs et des chargés d'étude. Si dans leur essence les fondamentaux n'ont pas beaucoup évolué, les techniques et les pratiques changent complètement la donne. Internet rend immédiatement accessible la « bibliothèque » mondiale, les puissances de calcul et la baisse des coûts accélèrent l'obsolescence de techniques statistiques développées dans une économie de pauvreté numérique, les progrès de l'intelligence artificielle et du traitement automatique des langues ouvrent de nouvelles perspectives.

Je m'appuie sur mon expérience, enrichie dans les dernières années par mes contacts avec le corps professoral et les chercheurs du Business Science Institute (BSI). Je les forme aux méthodes qui leur permettent en tant que managers de contribuer à la création de connaissances. Cela m'a aidé à formaliser cet ouvrage en sélectionnant l'essentiel des apports de la tradition scientifique, des méthodologies et des outils. Mon objectif est de décrire le champ des possibles, d'en donner une intelligence suffisante pour guider les choix et mettre en œuvre les techniques et outils appropriés. Pour cela, j'ai bénéficié de la grande variété des sujets abordés par les chercheurs du Business Science Institute et de la complicité avec la société éditrice des logiciels Sphinx qui les accompagne.

Que tous en soient remerciés !

Ce livre s'inscrit dans l'esprit des ouvrages de la collection Business Science Institute des éditions EMS<sup>7</sup>.

### Plan général du livre

I – Produire des connaissances en sciences sociales

II – Mesurer et analyser : méthodes quantitatives

---

7. Beaulieu, P., Kalika, M. (2015) ; Kalika, M. (2017), à paraître Chevalier F., Cloutier M., Mitev N., *Méthodes de recherche du DBA*, septembre 2018.

III – Observer et interpréter : méthodes qualitatives

IV – Communiquer

Le site compagnon : <https://faireparlerlesdonnees.blog/> présente de nombreux exemples et les modes opératoires nécessaires à leur mise en œuvre avec les logiciels Sphinx.

Je déroge à la tradition de débiter par l'exposé des méthodes qualitatives car les méthodes quantitatives, mieux connues et documentées, sont plus faciles à exposer. Mais ce choix exprime surtout l'idée que les méthodes qualitatives peuvent aussi être présentées comme un aboutissement plutôt qu'un préalable !