



**HAL**  
open science

## Données, Valeur et Business models

Henri Isaac

► **To cite this version:**

| Henri Isaac. Données, Valeur et Business models. 2016. hal-01821836

**HAL Id: hal-01821836**

**<https://hal.science/hal-01821836>**

Submitted on 19 Oct 2018

**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.

# LES CAHIERS SCIENTIFIQUES



CAHIER N°21 : DONNÉES, VALEUR ET BUSINESS MODEL

*Henri Isaac*

-2016-







The first part of the document discusses the importance of maintaining accurate records of all transactions. It emphasizes that every sale, purchase, and payment must be properly documented to ensure the integrity of the financial statements. This includes recording the date, amount, and purpose of each transaction.

The second part of the document provides a detailed breakdown of the company's revenue. It shows the total sales for each quarter and year, along with a comparison to the budget. The analysis highlights areas where the company has exceeded expectations and identifies opportunities for improvement in other areas.

The third part of the document focuses on the company's expenses. It details the costs of goods sold, operating expenses, and other financial obligations. The goal is to identify areas where costs can be reduced without compromising the quality of the company's products or services.

The final part of the document summarizes the overall financial performance and provides recommendations for future actions. It stresses the need for continued vigilance in financial management and the importance of regular communication with stakeholders.

# Introduction

La Chaire Intelligence Économique et Stratégie des Organisations (IESO) a pour objectif d'ancrer à Dauphine et dans le monde académique la maîtrise de l'information stratégique, ou intelligence économique, comme une question centrale du management de toute organisation évoluant dans un contexte concurrentiel.

L'information stratégique renvoie aux données, aux idées ou aux faits considérés comme sensibles pour la compétitivité des organisations. La problématique de sa maîtrise est entrée dans le champ académique à la suite des travaux d'H. Simon en 1947 sur la rationalité procédurale, associés aux écrits en 1967, soit vingt ans plus tard, d'H. Wilensky sur l'intelligence organisationnelle. Si le premier met en exergue les limites cognitives des individus dans la collecte et le traitement de l'information, le second analyse les blocages structurels et culturels à la circulation de l'information nécessaire à la fonction stratégique, rompant ainsi avec le modèle de l'équilibre où l'agent décide de manière optimale en situation d'information pure et parfaite.

Parce que l'information est imparfaite, c'est-à-dire incomplète et généralement asymétrique entre les acteurs, parce que l'information est coûteuse, que ce soit pour sa collecte, pour son interprétation dans l'action collective, ou encore pour sa diffusion qui nécessite de capter l'attention des parties concernées, parce que l'information est impure, c'est-à-dire sujette à interprétations et ambiguïtés, la maîtrise de l'information sensible avant ses concurrents est au cœur des processus stratégiques et de la capacité de différenciation de toute organisation.

C'est d'autant plus le cas si certaines décisions remettent régulièrement en jeu le sort de l'entreprise dans des espaces stratégiques marqués par des relations complexes de coopération et de compétition. C'est notamment le cas dans les industries à cycle long d'investissement, sujettes à une forte concurrence internationale et des risques d'ingérence associés, et où la compétitivité est fortement liée à des choix d'innovation technologique. L'enjeu est ainsi crucial pour des entreprises comme PSA, Sanofi et Total.

L'objectif des travaux conduits dans le cadre de cette chaire est donc d'étudier et d'améliorer les dispositions individuelles et les capacités organisationnelles à maîtriser l'information sensible, avant ses concurrents, pour la compétitivité d'une organisation.

The first part of the document discusses the importance of maintaining accurate records of all transactions. It emphasizes that every entry, no matter how small, should be recorded to ensure the integrity of the financial data. This includes not only sales and purchases but also expenses and income. The document provides a detailed list of items that should be tracked, such as inventory levels, supplier payments, and customer orders. It also outlines the procedures for recording these transactions, including the use of standardized forms and the importance of double-checking entries for accuracy.

The second part of the document focuses on the analysis of the recorded data. It describes various methods for identifying trends and anomalies in the financial records. This includes comparing current performance with historical data and industry benchmarks. The document also discusses the importance of regular audits to verify the accuracy of the records and to detect any potential fraud or errors. It provides a step-by-step guide for conducting these audits, from the selection of samples to the final reporting of findings.

The final part of the document addresses the reporting and communication of the financial information. It explains how to prepare clear and concise reports that provide a comprehensive overview of the company's financial health. This includes the use of charts and graphs to visualize key data points and the inclusion of detailed explanations for any significant fluctuations. The document also discusses the importance of transparency in financial reporting and the need to communicate the results to all relevant stakeholders, including management, investors, and regulatory bodies.

Dans ce contexte, nous cherchons à la fois à étudier les pratiques mais aussi à enrichir/construire des dispositifs – conceptuels ou outillés – qui favorisent la collecte, l'interprétation et la transmission de l'information stratégique dans des organisations compétitives. L'information stratégique est entendue ici comme des données, des idées ou encore des faits sur lesquels s'est construite une conviction collective de leur caractère sensible pour la compétitivité d'une organisation.

Il s'agit ainsi de mettre au centre de la fonction stratégie l'anticipation des risques et opportunités susceptibles d'impacter la compétitivité d'une entreprise dans son marché.

Les projets de recherche qui sont développés dans le cadre de la chaire s'articulent autour de cinq axes. Ces axes se structurent en partant de l'analyse la plus micro des pratiques sociales d'intelligence économique jusqu'à la construction d'analyses macroéconomiques et géopolitiques permettant une meilleure anticipation de forces exogènes jouant un rôle pivot dans l'évolution de nos environnements.

- Axe « la circulation de l'information stratégique » dirigé par Stéphanie Dameron, Professeur de stratégie d'entreprises et Directrice de la Chaire IESO, et Lionel Garreau, Maître de conférences en stratégie d'entreprises et Adjoint à la directrice de la Chaire IESO
- Axe « Prospective et dynamique des marchés : à la recherche de signaux faibles » dirigé par Manuel Cartier, Maître de conférences en stratégie d'entreprises
- Axe « Business Model, digital et stratégie » dirigé par Henri Isaac, Maître de conférences en management des systèmes d'information
- Axe « Intelligence économique des risques et opportunités « énergie – climat » » dirigé par Patrice Geoffron, Professeur d'économie
- Axe « L'intégration de la dimension géopolitique dans l'élaboration des stratégies d'entreprises et des politiques publiques » dirigé par Bernard Guillochon, Professeur émérite d'économie





---

## *L'auteur : Henri Isaac*

---



### **Henri Isaac**

*Docteur en sciences de gestion, est maître de conférences à l'Université Paris Dauphine, chercheur au sein de Dauphine Recherches en Management (CNRS, UMR 7088). Il a été Directeur de la Recherche et Directeur Académique à Neoma Business School (2009-2012). Il a été vice-président en charge du numérique de l'université Paris Dauphine (2014-2016) à l'Université Paris-Dauphine. Co-responsable du certificat « E-commerce » CCMBenchmark-Dauphine, il est co-auteur de plusieurs ouvrages « E-commerce » (4<sup>ème</sup> édition 2014, Pearson France), « Marketing digital » (2017, Pearson), « Travail à distance & e-management », (Dunod, 2013). Spécialiste des systèmes d'information et de l'économie digitale, ses recherches, à la frontière du champ des systèmes d'information et du management, portent sur les effets des technologies sur les entreprises et leur management. Il est l'auteur de nombreuses publications académiques sur ces sujets. Il est l'auteur d'un rapport sur « l'Université Numérique » pour la Ministre de l'enseignement supérieur Mme Valérie Pécresse. Il est également Président du Think Tank Renaissance Numérique.*



# Idées clefs

La donnée est désormais considérée comme une nouvelle matière première dans l'économie digitalisée.

Cependant, le processus de création de valeur transformant la donnée en valeur pour l'entreprise et/ou le client s'avère complexe, nécessite plusieurs étapes et combine plusieurs types de ressources.

En effet, les données ont plusieurs origines et régimes juridiques. Dès lors, le potentiel de création de valeur dépend dans un premier temps de l'appréhension des données à disposition et des différents dispositifs de collecte.

La valeur des données réside davantage dans les métadonnées et les algorithmes de traitement que dans la donnée elle-même.

La restitution de la valeur créée à partir de ces traitements n'est pas encore totalement maîtrisée par les entreprises qui n'accordent pas suffisamment d'attention à cette phase essentielle dans l'appropriation de la valeur.

Un modèle général de création valeur à partir de la donnée comporte au final quatre étapes : génération de la donnée, qualification de la donnée, traitement algorithmique, médiation du traitement.

Loin de se résumer à un seul assemblage technique, la création de valeur fondée sur la donnée nécessite un design organisationnel spécifique.

Celui-ci doit intégrer une gouvernance de la donnée afin de s'assurer que celle-ci est correctement appréhendée par les différents acteurs de l'organisation.

Des compétences spécifiques sont nécessaires pour l'extraction de la valeur depuis les données collectées. Cependant, leur intégration dans les différentes composantes de leur organisation est encore plus déterminante.

Dès lors l'architecture des ressources nécessaires au traitement des données est un enjeu primordial de la performance.

Depuis de nombreuses années, des modèles d'affaires se sont construits sur l'exploitation de la donnée.

Cependant dans l'économie des plateformes, la donnée occupe une place centrale dans la création de valeur en l'exposant à tout un écosystème de partenaires. Ceci est particulièrement le cas pour l'Internet des Objets.

Dans ce cas, le processus de création de valeur à partir de la donnée en quatre étapes est très présent et au coeur de ces modèles d'affaires.

Une modélisation plus générale de la création de valeur fondée sur les données est envisageable, notamment en utilisant la méthode des options réelles comme cadre d'analyse générale.

# Sommaire

Introduction .....	13
<b>01. DE LA DONNÉE À LA VALEUR</b> .....	14
1. Problématique générale de la valeur de la donnée .....	14
2. Nature des données et valeur .....	15
3. Métadonnée et valeur .....	17
4. L'enjeu de la qualité des données.....	18
5. Algorithmes, données et valeur .....	20
6. Médiation, donnée et valeur .....	24
7. Ressources et compétences : de la gouvernance de la donnée à l'architecture de la valeur.....	26
<b>02. LES MODÈLES D'AFFAIRES FONDÉS SUR LA DONNÉE</b> .....	29
1. Une analyse typologique des modèles d'affaires fondés sur la donnée .....	29
2. Données et valeur dans les modèles d'affaires de plateforme .....	31
3. Donnée et valeur dans les modèles d'affaires de l'Internet des Objets.....	34
<b>03. CONCLUSION</b> .....	37
Bibliographie.....	38

# Introduction

La transformation digitale du monde est une nouvelle étape de sa mathématisation (Roux, 2011). Elle se traduit par une mise en donnée des interactions marchandes et sociales. Cette « datafication » conduit à l'apparition d'une nouvelle ressource, la donnée. Celle-ci est présentée comme la nouvelle unité de base d'un capital que les entreprises devraient constituer par accumulation, qu'elles devraient par la suite valoriser en organisant sa circulation, notamment dans l'industrie afin d'accélérer le passage à une logique de service et éviter ainsi la commodisation des produits (Opresnik, Taisch, 2015).

Dès lors, de plus en plus d'entreprises bâtissent donc des modèles d'affaires fondés sur la donnée (*data-driven business model*). La variété de ces modèles est importante (Hartman Et alii 2014 ; FTC 2014), traduisant de nombreux positionnements possibles le long de la chaîne de valeur de la donnée : collecte, stockage, traitement, analyse, protection, commercialisation.

Quel que soit le positionnement dans la chaîne de la donnée, la question de la valeur de la donnée est rarement discutée, comme si elle allait de soi. A cet égard, les discours produits sur les données par l'industrie des technologies de l'information ou les défenseurs de la vie privée ne portent que rarement sur cette question de la valeur de la donnée et les mécanismes par lesquels elle génère de la valeur pour les différentes parties.

L'objectif vise à décrypter les mécanismes de création de valeur fondés sur la donnée et de proposer une analyse des modèles d'affaires qui en découlent. Dans une première partie, une analyse approfondie des mécanismes d'extraction de la valeur fondés sur les données est proposée. Elle met en lumière la complexité de ces mécanismes qui rassemblent différentes ressources – données, métadonnées, algorithmes, data-visualisation – et compétences technologiques pour produire de la valeur. Ces mécanismes, loin d'être seulement un dispositif technique, combinent différentes ressources et compétences au sens stratégique du terme. Techniques, organisationnelles, humaines, ces ressources et compétences, souvent difficiles à rassembler, constituent la base d'une nouvelle architecture de la valeur. Ces assemblages technologiques et organisationnels sont à la base d'une variété de nouveaux modèles d'affaires.

Dans une seconde partie nous analysons les différents modèles d'affaires fondés sur les données. Après avoir discuté des principaux modèles qui se déploient dans l'industrie de la donnée et dans les modèles d'affaires classiques, nous proposons un approfondissement sur deux cas de modèles d'affaires dans lesquels la donnée est au cœur même de leur logique. En effet, la numérisation a fait émerger des nouvelles formes d'organisation et de production de la valeur : les plateformes. Leur succès repose en grande partie sur leur capacité à architecturer des ressources aptes à capturer et gérer l'accès aux données. Enfin, cette logique se déploiera de façon assez massive dans différentes industries, dès lors que celles-ci encastrent les logiques numériques au sein même de leurs produits. Dès lors, les modèles d'affaires de l'Internet des Objets constituent une application des nouvelles logiques d'extraction de la valeur au travers de la donnée. Un cadre d'analyse, fondé sur la maturité informationnelle, permet d'identifier plusieurs modèles d'affaires différents.

# 01. DE LA DONNÉE À LA VALEUR

## 1. Problématique générale de la valeur de la donnée

De quelle valeur parle-t-on ? Tout dépend de quel point de vue est adopté : celui du client ou de l'entreprise. Une étude menée par Orange en Europe (Orange, 2014) illustre bien le fait que les clients ont conscience de la valeur de leurs données privées sans pouvoir en évaluer réellement leur valeur monétaire, ce qui constitue un des enjeux de la notion de self-data et d'une approche *Vendor Relationship Management* (VRM)<sup>1</sup>, dans laquelle l'utilisateur gère l'accès des entreprises à ses données. Du côté des entreprises, le simple fait de collecter de la donnée ne constitue pas *per se* une source de création de valeur. En outre, créer de la valeur par l'exploitation de la donnée dépend de nombreux paramètres.

En premier lieu, la nature de la donnée (données personnelles, données techniques, données comportementales, données primaires versus secondaires) doit être prise en considération car elle détermine les traitements possibles, notamment d'un point de vue légal. Ainsi, des cadres légaux encadrent les collectes, partages et traitements possibles dans certaines activités (santé, assurance, finance), limitant les traitements possibles et les partages de données.

Cependant, la plupart des données n'ont de valeur que si elles sont complétées par des métadonnées<sup>2</sup> qui les décrivent et permettent leur exploitation. La prise de conscience de la valeur économique des métadonnées est récente car jusqu'ici, elles n'ont été perçues que pour leur rôle technique (Greenber, 2014). C'est donc la combinaison des données et des métadonnées qui conditionnera la valeur produite.

Par ailleurs, la qualité des données est un autre facteur qui conditionnera la production de valeur, et notamment à l'ère des méga-données, leur véracité (Teboul & Berthier, 2015).

En second lieu, la nature des traitements (modélisation

algorithmique, temps réel, etc.) conditionne également la valeur extraite des données. Par conséquent, la question de la création de valeur à partir de la donnée nécessite d'explicitier ces mécanismes. Ceux-ci sont désormais de plus en plus fondés sur des algorithmes auto-apprenants (*machine learning, deep learning*) (Donoho, 2015). Dès lors, le traitement algorithmique ne permet pas seulement de produire un service, il en fournit le modèle mais aussi l'améliore au fur et à mesure. La donnée a donc ici une double valeur, elle est à la base de la construction d'un modèle mais aussi de son optimisation.

En synthèse, la valeur de la donnée ne peut être que conditionnelle et non intrinsèque comme de nombreuses analyses le laissent penser.

### UN CADRE POUR L'ANALYSE DE LA VALEUR DE LA DONNÉE POUR L'ENTREPRISE

#### Legal & Privacy Framework

$$\text{Data Value} = [\sum \text{Data.Quality}].[\text{MetaData.Quality}].[\text{Algorithm.Category}]$$

Quelle que soit la valeur produite à partir des données, il demeure que leur valeur finale pour l'entreprise résidera dans la qualité de leur utilisation dans la relation de service avec le client final, comme le montre une récente étude dans l'industrie de la distribution<sup>3</sup>. Il est donc possible de représenter la valeur de la donnée pour le client à l'aide du cadre d'analyse suivant :

### UN CADRE POUR L'ANALYSE DE LA VALEUR DE LA DONNÉE POUR LE CLIENT

#### Legal & Privacy Framework

$$\text{Data Value} = [\text{MetaData.Quality}].[\text{Algorithm.Category}].[\text{Mediation.Quality}]$$

1. Voir <http://www.culturemobile.net/point-expert/self-data>

2. <https://fr.wikipedia.org/wiki/Métadonnée>

3. CapGemini Consulting, (2015), Privacy Please: Why Retailers Need to Rethink Personalization, Octobre. <http://www.slideshare.net/capgemini/privacy-please-why-retailers-need-to-rethink-personalization>

La valeur des données est donc le résultat d'un processus de production complexe<sup>4,5</sup>. En effet, la collecte de données brutes à elle seule ne suffit pas à créer une quelconque valeur. Ceci est d'autant plus vrai, qu'une partie de la valeur ne réside pas dans la donnée elle-même, mais dans les métadonnées associées aux données.

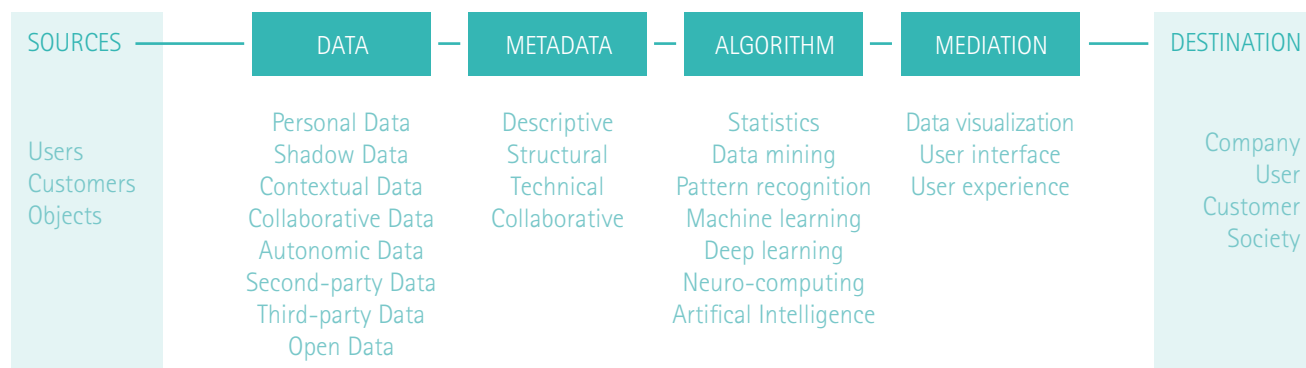
Plus encore, la valeur ne dépend pas de cette première combinaison, mais de l'extraction de la valeur par le biais de traitements algorithmiques et de processus de restitution et de valorisation aptes à créer du sens pour leurs utilisateurs (entreprises ou clients).

Enfin, il est nécessaire de comprendre quelles sont les ressources et compétences que mobilisent les organisations pour concevoir, déployer et exploiter les dispositifs à même d'extraire cette valeur et la restituer au client ou à l'utilisateur. L'architecture de ces ressources et compétences est elle-même un enjeu stratégique tant le design organisationnel nécessaire à la captation des données et leur exploitation en est le fruit, comme le modèle de plateforme l'illustre. Dans ce modèle organisationnel, la donnée devient une véritable « monnaie d'échange » entre ses partenaires (Choudary, 2015).

Ce cadre général d'analyse nécessite d'être précisé tant, chacune des étapes de la chaîne de valeur de la donnée, comporte plusieurs dimensions qui influencent l'extraction de la valeur.

Le tableau ci-dessous décrit chacune de ces étapes du processus de création de la valeur. Il identifie pour chacune d'entre elles les différentes catégories qui les composent.

TABLEAU 1 : UN CADRE D'ANALYSE DE LA CRÉATION DE VALEUR PAR LA DONNÉE



Source : Henri Isaac

4. Crié, D., Micheaux, A., (2006), « From customer data to value : What is lacking in the information chain ? », *Database Marketing & Customer Strategy Management*, vol. 13, no. 4, pp. 282-99.

5. LaValle S., Lesser E., Shockley R., Hopkin M.S., Kruschwitz N., (2011), « Big Data, Analytics and the Path From Insights to Value », *MIT Sloan Management Review*, Vol 52, n°2, Winter, pp. 21-32.

## 2. Nature des données et valeur

La nature des données utilisables dans le processus de création de valeur est large et les enjeux associés le sont également<sup>6</sup>. Le régime de production des données a considérablement évolué ces dernières années, puisqu'aux données personnelles classiques, viennent désormais s'ajouter des données produites automatiquement par différents dispositifs numériques : objets connectés, robots, capteurs autonomes. Les données sont également produites dans des contextes qui sont désormais identifiables par l'entreprise et captés en même temps que les interactions à l'origine des données. Les données peuvent aussi être le résultat d'interactions entre utilisateurs ou clients. Dans certains cas, elles peuvent être enrichies par des données étrangères à l'échange ou à la transaction. Enfin, de plus en plus de données publiques sont accessibles en mode ouvert (open data).

### a. Données personnelles & traces

Les données personnelles des utilisateurs et/ou des clients diffèrent des traces numériques liées aux usages même si dans la pratique certaines traces peuvent entrer dans le périmètre juridique des données personnelles, comme l'adresse IP. Du point de vue juridique, une donnée personnelle est définie en droit français<sup>7</sup> comme « ... une donnée à caractère personnel, ou toute information relative à une personne physique identifiée ou qui peut être identifiée, directement ou indirectement, par référence à un numéro d'identification ou à un ou plusieurs éléments qui lui sont propres. Pour déterminer si une personne est identifiable, il convient de considérer l'ensemble des moyens en vue de permettre son identification dont dispose ou auxquels peut avoir accès le responsable du traitement ou toute autre personne ».

6. Voir, Cigref, (2015), L'économie des données personnelles : les enjeux d'un business éthique, octobre, <http://www.cigref.fr/wp/wp-content/uploads/2015/11/CIGREF-Economie-donnees-perso-Enjeux-business-ethique-2015.pdf>

7. Article 2 de la Loi Informatique & Liberté



Cependant, les données issues des activités d'usage (d'un dispositif numérique en ligne ou d'un objet connecté par exemple) ne sont pas toujours des données personnelles, mais des traces liées à l'usage, ce que d'aucuns appellent l'ombre numérique<sup>8</sup>. Nous les dénommeront « *shadow data* ».

Dans cette catégorie, il faut distinguer les données explicites issues des transactions (historiques d'achats, de paiement, d'usage, etc.) des données implicites issues des dispositifs techniques eux-mêmes (données de logs, données techniques, etc.).

Ces données sont nécessaires à la bonne exécution opérationnelle de ces dispositifs numériques ou connectés. Elles permettent le pilotage de la performance de ces mêmes dispositifs. Dès lors, ces données ne peuvent être strictement considérées comme des données personnelles dans la mesure où elles ne sont produites que pour la bonne exécution du service et ne sont pas liées à l'identité d'une personne. Leur suppression dégraderait la qualité du service. Il en va de même sur les historiques d'achats des clients qui sont à la base de la gestion de la relation du client et des programmes relationnels.

### b. Données contextuelles

Les données contextuelles sont issues de l'environnement d'une transaction ou d'un échange (d'informations ou tout autre type d'échanges). Elles ne concernent ni l'utilisateur directement, ni l'échange lui-même, mais renseignent sur le contexte (localisation, temporalité, caractéristiques de l'environnement). Elles peuvent parfois être une composante des méta-données<sup>9</sup>.

Elles peuvent également provenir de dispositifs interconnectés par le biais de capteurs, comme dans le cas de la domotique où des dispositifs connectés peuvent générer des données pour produire des services contextualisés<sup>10</sup> (*Context Aware Services*<sup>11</sup>). L'utilisation des données de contexte est la base d'une

personnalisation situationnelle qui accroît la logique d'hyper-personnalisation.

### c. Données collaboratives

Une catégorie de données plus spécifiques provient des interactions sociales des individus sur des informations ou dans des échanges sur différents types d'espaces numériques (réseaux sociaux, messageries). Elles sont le fruit d'une collaboration volontaire sur les informations elles-mêmes. Il en va ainsi des activités de taggage de photos sur les réseaux sociaux où, les personnes par des actions sur ces informations, produisent au travers d'une interaction des données supplémentaires à l'information, sans qu'elles puissent strictement être considérées comme des métadonnées. Les données ajoutées enrichissent la donnée primaire grâce à une autre donnée. Les activités de partages de signets ou de favoris produisant des données mutualisées créant des catégories par agrégation rentrent également dans cette catégorie dans la mesure où la valeur provient de l'agrégation et non de la donnée prise isolément.

### d. Données automatiques

A ce type de donnée issue d'interactions entre individus, viennent s'ajouter des données produites par les objets connectés de plus en plus nombreux. S'il existe une typologie large d'objets connectés, une grande partie des données produites le seront automatiquement par le biais de capteurs. Ainsi la poubelle connectée génère elle-même des données qui mesurent son remplissage et sont utilisées pour déclencher une alerte pour procéder à son traitement. Ce sont donc des données automatiques produites par les objets eux-mêmes qui viennent enrichir la masse de données utilisables pour créer de la valeur.

### e. Données complémentaires : second party data & third party data

Toutes ces données peuvent être complétées par des données provenant d'autres acteurs économiques. Dans le cas du marketing digital, de nombreuses données viennent s'ajouter aux données de l'entreprise. Ainsi dans des campagnes de publicité en ligne (*online display*), l'annonceur peut ajouter des données de l'éditeur (*second party data*) auprès duquel se déploie sa campagne et il peut y ajouter des données de tiers (*third-party data*). Le processus de création de valeur combine alors de nombreuses données : personnelles, contextuelles, secondaires et de tierces parties.

8. Voir <https://myshadow.org>, voir également : A Coutant, T Stenger (2010), « Pratiques et temporalités des réseaux socio-numériques : logique de flux et logique d'archive », MEI-Médiation et information.

9. Comme dans le cas d'une photo pour laquelle un smartphone ajoute automatiquement dans le fichier de métadonnées EXIF les données de localisation et l'heure de la prise de vue.

10. Des services combinant des données de localisation du véhicule, de trafic routier en temps réel, météorologiques, de destination sont typiques de cette catégorie de services.

11. Voir par exemple Philip T. Moore and Hai V. Pham, (2015), « Personalization and rule strategies in data-intensive intelligent context-aware systems », *The Knowledge Engineering Review*, Volume 30, Special Issue 02, March, pp. 140-156.

## f. Les données ouvertes (Open data)

Les données ouvertes constituent depuis quelques années un type particulier de données mises à disposition gratuitement au travers d'une licence spécifique. Elles ont pour origine des données de la puissance publique ou d'établissements publics, voire d'entreprise privée. L'ouverture des données est un mouvement de fond initié aux Etats-Unis et en Europe<sup>12</sup>. La mise à disposition de données ouvertes permet aux acteurs économiques de s'appuyer sur des ressources coûteuses à produire (cartographie, base nationale d'adresses, etc.). En outre, elles enrichissent potentiellement les données de l'entreprise. Enfin, en ouvrant ses données à des tiers, dans une démarche d'open innovation, l'entreprise peut également bénéficier de traitements et d'exploitations innovants<sup>13</sup>.

En synthèse, le processus de création de valeur mobilise potentiellement une variété de données de nature différentes aux enjeux économiques et juridiques différents. Le graphique 1 résume cet ensemble de possibles.

Quel que soit le type de données et leurs combinaisons, une partie de la valeur, sinon l'essentiel dans certains services numériques, est créée non pas par les données elles-mêmes, mais par l'exploitation des métadonnées associées.

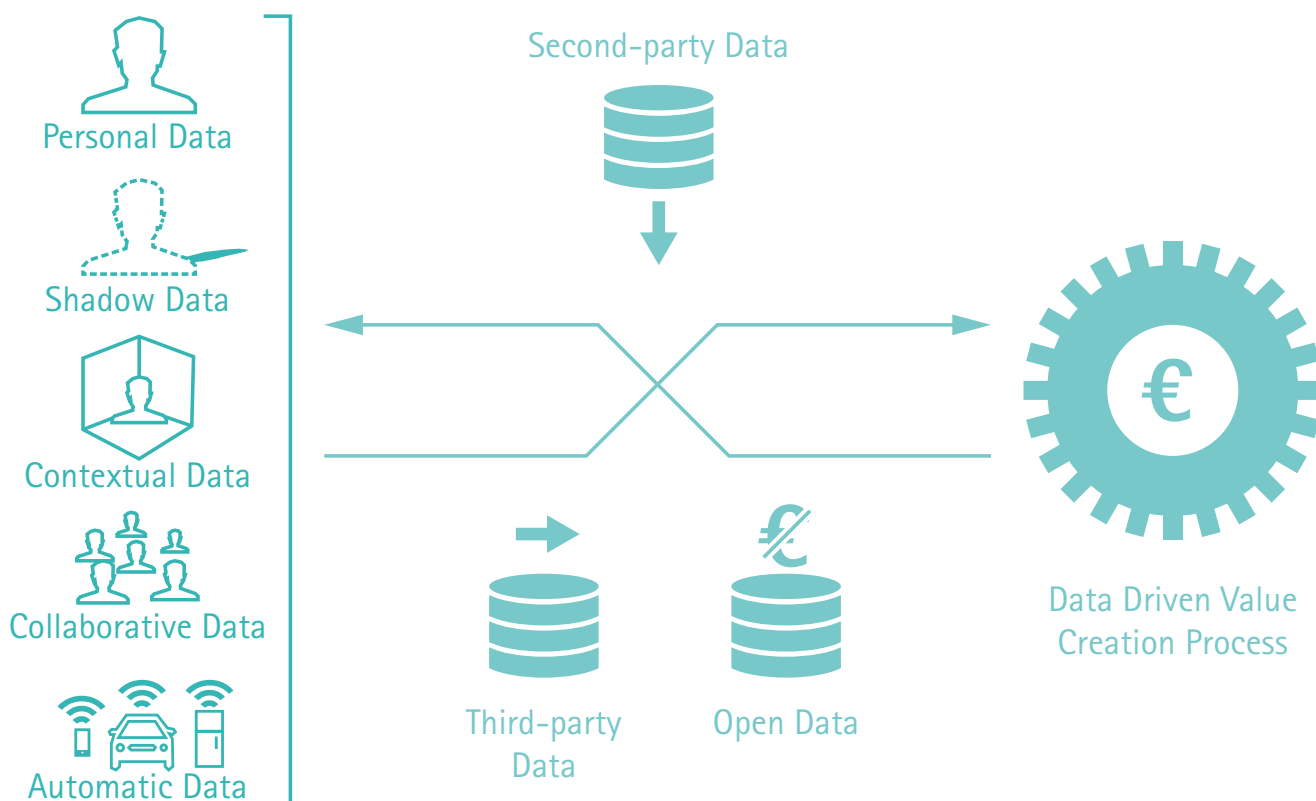
## 3. Métadonnée et valeur

### a. Définition d'une métadonnée

Une métadonnée est une donnée servant à définir ou décrire une autre donnée quel que soit son support. Par exemple, les fichiers de photos numériques captées par des terminaux électroniques comme un smartphone ou un appareil photographique intègrent des métadonnées (coordonnées GPS, exposition, taille du fichier, etc.) sur les données (pixels de la photo).

Il existe une grande variété de métadonnées, et une normalisation internationale existe sur certaines d'entre elles<sup>14</sup>.

FIGURE 1 : TYPOLOGIE DES DONNÉES ET PROCESSUS DE CRÉATION DE VALEUR



Source : Henri Isaac

12. [https://fr.wikipedia.org/wiki/Open\\_data](https://fr.wikipedia.org/wiki/Open_data)

13. Voir par exemple <https://data.sncf.com>

14. Voir par exemple : <https://fr.wikipedia.org/wiki/Métadonnée>

- Les **métadonnées descriptives** décrivent une ressource numérique à des fins telles que la découverte et l'identification. Elles peuvent inclure des éléments tels que le titre, le résumé, l'auteur et des mots-clés.
- Les **métadonnées structurelles** indiquent comment les objets sont composés et organisés, par exemple, comment les pages sont ordonnées pour former des chapitres.
- Les **métadonnées administratives** fournissent des informations pour gérer une ressource, incluant des données comme la date de création, le type de fichier et d'autres informations techniques, et les droits d'accès. On distingue les métadonnées de gestion, qui traitent des droits de propriété intellectuelle, comme les DRM (Digital Right Management), des métadonnées de conservation, qui contiennent les informations nécessaires pour archiver et préserver une ressource numérique.

## b. Donnée, métadonnée et création de valeur

Les métadonnées constituent une dimension fondamentale de la valeur des données. En effet, de très nombreux services digitaux reposent de façon quasi exclusive sur la gestion et l'exploitation des métadonnées plutôt que des données elles-mêmes. Il en va ainsi des systèmes de messagerie électronique, les moteurs de recherche, des services musicaux en ligne (les playlists proposées ne sont qu'une gestion des métadonnées), des services de catalogage de photographies, les catalogues de commerce électronique, les services de réseaux sociaux, etc.

Les données, sans leurs métadonnées, se révèlent inexploitable et leur valeur s'en trouve dès lors très fortement diminuée. Que vaut un fichier de musique digitale sans ses métadonnées ? Il ne peut être classé, ni ajouté à une liste de lecture intelligente. Aussi, une grande partie de la valeur des données réside dans la qualité de leurs métadonnées. Plus encore, c'est la combinaison de la donnée et de ses métadonnées qui est à l'origine de la valeur.

Ces combinaisons créatrices de valeur reposent sur les fonctions que les métadonnées rendent possibles. Les principales fonctions sont les suivantes :

- **La découverte** de ressources numériques. C'est une fonction essentielle qui permet, selon des critères pertinents, de trouver des ressources numériques. C'est la base de tous les moteurs de recherches inclus dans

de nombreux logiciels. Cette fonction de découverte inclut l'identification des ressources, leur tri, fournit des informations sur leur localisation.

- **L'organisation** de ressources électroniques. Les métadonnées permettent l'organisation des liens vers des ressources en fonction d'un public ou d'un sujet, la construction de pages dynamiquement à partir des métadonnées stockées dans des bases de données.
- **L'interopérabilité**. L'utilisation de métadonnées, reposant sur des protocoles de transfert normalisés, facilite la recherche entre différents systèmes techniques et rend les données interopérables, augmentant d'autant leur valeur. Dans le cadre de l'Internet des Objets, cette question de l'interopérabilité des métadonnées constituera un enjeu de taille car les données prendront de la valeur par leur intégration au sein de services combinant les données de plusieurs objets interconnectés.

Prises isolément, les données n'ont donc guère de valeur. Elles en acquièrent par leurs métadonnées. Le processus de création de valeur fondé sur les données ne saurait donc ignorer cette composante essentielle à la création de valeur.

FIGURE 2 : ÉTAPE 2 DU PROCESSUS DE CRÉATION DE VALEUR PAR LA DONNÉE



## 4. L'enjeu de la qualité des données

Les discours sur les données ignorent souvent une dimension fondamentale de leur économie, leur qualité. La plupart des données nécessitent souvent des retraitements afin de s'assurer de leur qualité, c'est même une étape souvent longue et coûteuse qui obère parfois l'extraction de la valeur elle-même<sup>15</sup>.

15. Yoon, V. Y., Aiken, P., Guimaraes, T., (2000), « Managing organizational data resources : Quality dimensions », *Information Resources Management Journal*, Jul-Sept., Vol.13, n°3.

## a. Qualité des données et ses conséquences

Selon Gartner Group, plus de 50% des programmes de CRM et des projets d'entrepôt de données (*datawarehouse*) échouent du fait de la mauvaise qualité des données<sup>16</sup>. English a identifié 122 organisations en Europe et aux États-Unis qui, ensemble, avaient perdu plus de 1.200 milliards de dollars en raison de problèmes liés à la mauvaise qualité de leurs données. Il estime qu'entre 20% et 35% du chiffre d'affaires est perdu dans l'échec des processus de gestion des données<sup>17</sup>.

Les problèmes de qualité des données ne sont pas nouveaux<sup>18</sup>. Ils ont différentes origines, depuis leur collecte (manuelle ou automatique) à leur exploitation. De nombreux travaux de recherche ont analysé cette problématique et proposent des démarches et modèles pour y remédier<sup>19</sup>.

Cependant, loin de se résumer à des enjeux purement techniques, la qualité des données ne peut être obtenue et maintenue dans le temps que grâce à une implication des différents acteurs qui gèrent des données dans l'organisation<sup>20</sup>.

C'est donc un enjeu organisationnel de gouvernance de la donnée. Une telle démarche doit rendre les données :

- **Accessibles.** Les données sont accessibles, quels que soient le format ou la source.
- **Disponibles.** Les données sont disponibles pour les utilisateurs et les applications, en temps voulu, quels que soient l'emplacement ou le mode d'utilisation.
- **Fiables.** Les données sont complètes et précises.
- **Cohérentes.** La signification des données est cohérente et il est possible de les rapprocher à travers les systèmes et les processus.
- **Sécurisées.** Les normes de sécurité et de confidentialité sont appliquées.

16. Gregory, A., (2011), « Data governance -- protecting and unleashing the value of your customer data assets », *Journal of Direct, Data and Digital Marketing Practice*, 12 (3), pp. 230-248.

17. English, L.P., (2009), *Information Quality Applied: Best Practices for Improving Business Information Processes*, Wiley Publishing Inc, Indianapolis, America.

18. Redman T. C., (1995), « Improve data quality for competitive advantage », *Sloan Management Review*, 36(2), 99

19. Voir par exemple, [https://en.wikipedia.org/wiki/Data\\_quality](https://en.wikipedia.org/wiki/Data_quality)

20. Breur T., (2009), « Data quality is everyone's business -- managing information quality -- part 2. », *Journal of Direct, Data and Digital Marketing Practice*, 11(2), 114-123.

## b. De la qualité des données à la gouvernance de la donnée

L'exploitation des données nécessite en effet une collaboration entre différents acteurs et départements et sans approche unifiée de la donnée, matérialisée par la mise en place de Master Data Management, il n'est guère envisageable de garantir la qualité de la donnée tout au long de son cycle de vie.

Cependant, si le déploiement de référentiels uniques de données vient matérialiser et supporter une démarche de gouvernance des données, le succès d'une telle démarche au niveau de l'organisation est un processus itératif qui comporte plusieurs niveaux<sup>21</sup>.

En effet, la gouvernance de la donnée n'est pas seulement une sous-catégorie de la gouvernance du système d'information. Étant donné sa finalité, elle doit être une préoccupation de gouvernance générale rattachée aux mécanismes classiques de gouvernance (Wende, 2007) qui doit définir des principes de leur gestion et de leur intégration dans les mécanismes de création de valeur<sup>22</sup>.

À ce premier niveau, s'articule un niveau managérial qui définit les processus de gestion des données (propriétaires de la donnée, mécanismes de gestion et de contrôle) qui s'articule lui-même avec un dernier niveau de gestion technique des données. Le pilotage de ces trois niveaux peut être confié à un responsable des données souvent dénommé Chief Data Officer.

Cette démarche devient essentielle alors même que la volumétrie des données ne cesse d'augmenter et que la nature des données se multiplie. De nombreuses démarches techniques actuelles, que ce soient les Data Management Platform (DMP<sup>23</sup>) pour agréger les données CRM et les données de campagnes digitales, les *data lakes*<sup>24</sup> pour colliger des données d'origines diverses, toutes ces démarches trop orientées vers la technologie négligent la question de la gouvernance de la donnée, sans laquelle le processus d'extraction de la valeur ne peut être garanti.

Une fois déployée la gouvernance des données, celles-ci nécessitent pour leur gestion et leur exploitation d'autres ressources encore, que sont les algorithmes, qui manipulent, extraient, classent, arrangent données et métadonnées.

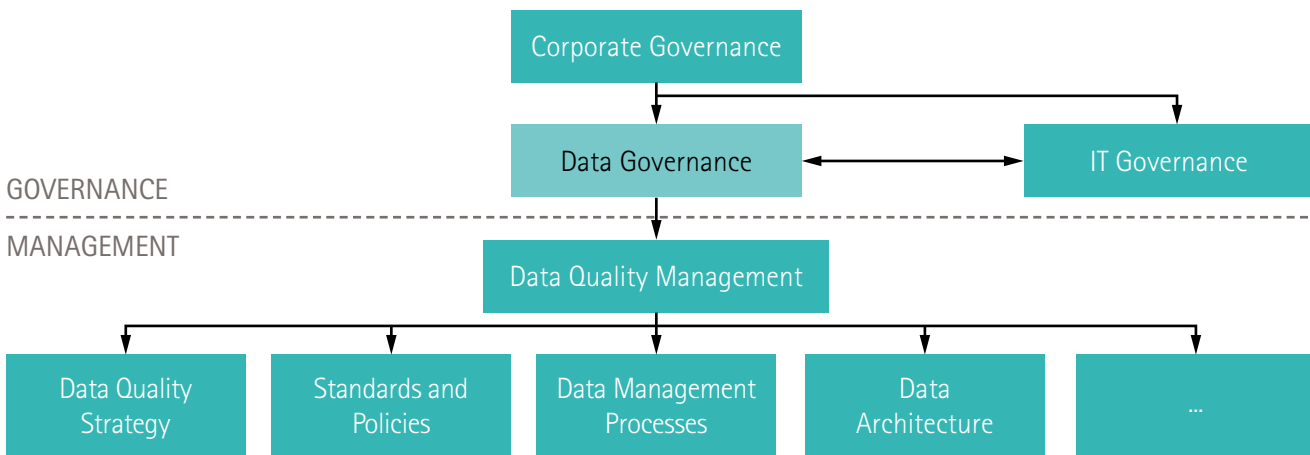
21. Baškarada S., Koronios A., (2014), « A Critical Success Factor Framework for Information Quality Management », *Information Systems Management*, Vol. 31, pp.276-295.

22. Wende K., (2007), « A Model for Data Governance – Organising Accountabilities for Data Quality Management », ACIS 2007 Proceedings

23. Voir IAB Europe, 2015, « Road to programmatic », July, White Paper, 78 p.

24. [https://en.wikipedia.org/wiki/Data\\_lake](https://en.wikipedia.org/wiki/Data_lake)

FIGURE 3 : GOUVERNANCE DE LA DONNÉE



Source (Wenke, 2007)

## 5. Algorithmes, données et valeur

Le traitement des données s'effectue au travers de nombreux algorithmes. L'ambition n'est pas de présenter une typologie exhaustive de ceux-ci, mais plutôt d'identifier les évolutions en cours dans cette étape essentielle de la création de valeur fondée sur la donnée. Si les algorithmes sont consubstantiels de l'informatique et en constituent une discipline ancienne, leurs usages toujours plus nombreux, apportent aux entreprises de nouvelles méthodes de création de valeur à partir des données.

Si l'utilisation de l'analyse de données en entreprise n'est pas récente, notamment depuis le développement des technologies d'entrepôts de données (*datawarehouse*) et des méthodes d'analyses associées (*datamining, business intelligence*), les progrès récents en algorithmie repoussent sans cesse ces possibilités. En effet, de nouvelles classes d'algorithmes comme le *machine learning* et de *deep learning* démocratisent les logiques de l'intelligence artificielle. Ce faisant ils introduisent de nouvelles possibilités pour les entreprises de créer de la valeur à partir de ces algorithmes.

### a. De la business intelligence...

L'informatisation des sociétés et le développement de leurs systèmes d'information opérationnels conduira au début des années 1990 à la création de technologies de bases de données spécialisées dans l'analyse plutôt que dans la gestion des transactions. En effet, les besoins de pilotage nécessitent un traitement des données peu compatibles avec l'efficacité opérationnelle des

systèmes de gestion. Ainsi, les technologies d'entrepôt de données (*Datawarehouse*) se sont progressivement déployées dans les entreprises pour offrir des solutions de pilotage des activités. De nombreuses techniques d'analyse de données se sont développées et sont désormais dénommées sous le terme générique de business intelligence.

Ces méthodes ont permis une première utilisation de méthodes statistiques avancées introduisant des techniques de découverte de connaissances (*knowledge discovery*) tant sur des données internes que sur les données de clients ou des données de marché. Les méthodes de *data mining* (forage de données) ont commencé à se déployer dans les organisations (Davenport, Harris, 2007). Ainsi, les méthodes de *scoring*, de détection de fraude sont typiquement le résultat de ces technologies de traitement des données. Les usages des données au travers de ces technologies de traitement ont essentiellement permis aux entreprises d'optimiser leurs processus de gestion, qu'il s'agisse du pilotage de la production ou de la relation client.

### b. ...au Big Data

Cependant, quelle que soit la nature de ces règles de gestion issues de ces analyses, elles demeurent le fruit d'une modélisation *a priori*. Les données valident ou invalident des hypothèses issues d'une modélisation théorique.

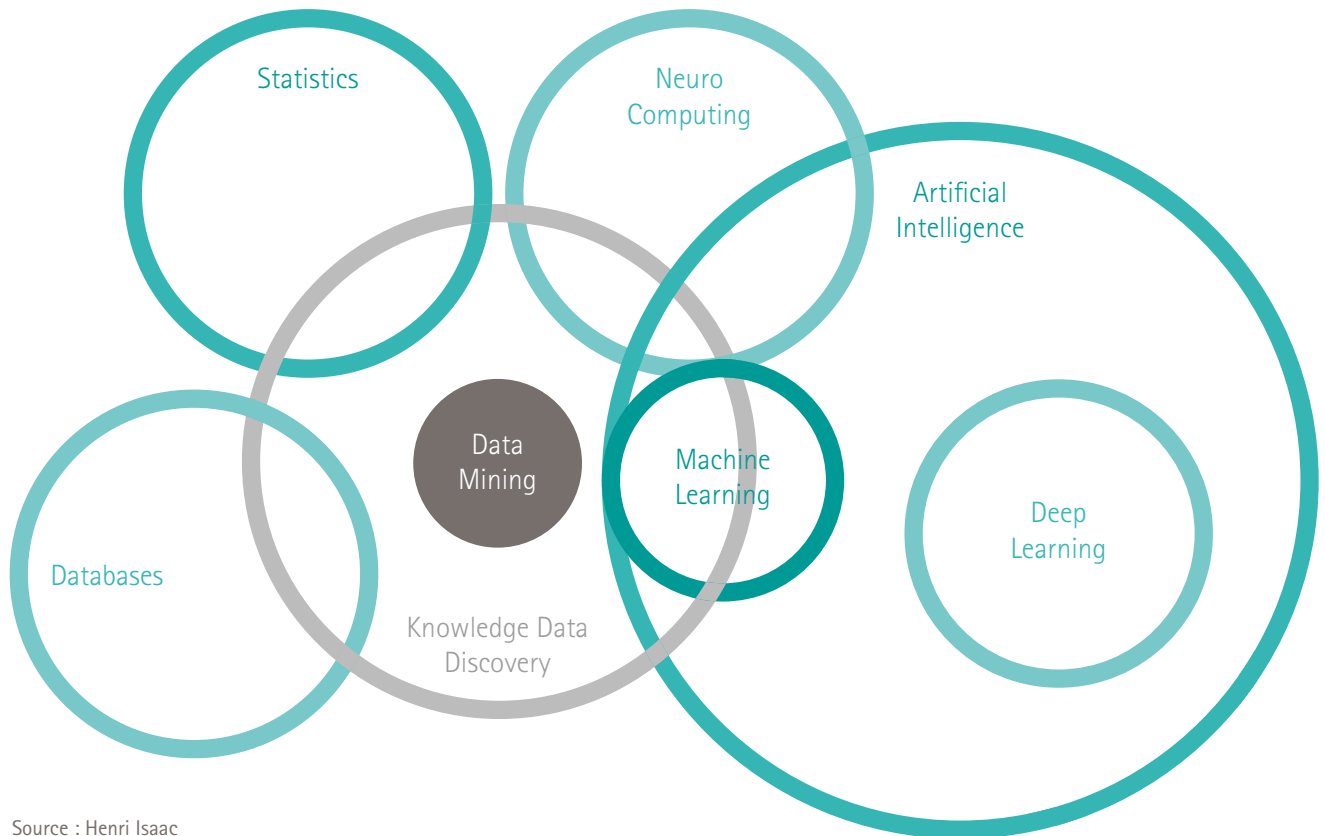
Au début des années 2000, la numérisation croissante des interactions sociales et marchandes conduira à de nouvelles méthodes de gestion des données du fait de

leur volumétrie croissante. De nouvelles méthodes de bases de données émergent (Hadoop, MapReduce), mais aussi de nouvelles approches algorithmiques sont rendues possibles par la puissance de calcul accrue des ordinateurs. Cet ensemble sera souvent résumé par le vocable « Big Data » sans qu'il ne traduise la diversité et la sophistication de ces nouvelles méthodes algorithmiques. Le Big Data rassemble en

effet, plusieurs approches aux présupposés théoriques différents et aux méthodologies bien distinctes, comme la figure 4 l'illustre.

Parmi ces méthodes, deux se distinguent tant elles contribuent à renouveler les approches dans l'utilisation des données : le *machine learning* et le *deep learning*.

FIGURE 4 : PANORAMA DES TECHNOLOGIES ALGORITHMIQUES DU BIG DATA



Source : Henri Isaac

### c. ...au Machine Learning

Le *machine learning* ou apprentissage automatique est une discipline scientifique centrée sur le développement, l'analyse et l'implémentation de méthodes automatisables qui permettent à une machine (au sens large, ici un algorithme) d'évoluer grâce à un processus d'apprentissage et ainsi de remplir des tâches qu'il est difficile ou impossible de remplir par des moyens algorithmiques plus classiques. Ainsi,

les algorithmes de *machine learning* introduisent une rupture par rapport aux modèles statistiques classiques, comme nous le résumons dans le Tableau 2. Une rupture introduite par cette catégorie d'algorithmes repose sur le fait que la modélisation est issue des données elles-mêmes. La découverte des variables est le résultat du traitement algorithmique qui ensuite ajustera le poids de celle-ci au fur et à mesure que des données nouvelles alimenteront l'algorithme.

TABLEAU 2 : DIFFÉRENCE ENTRE LE *MACHINE LEARNING* ET L'APPROCHE STATISTIQUE CLASSIQUE

Machine learning	Statistiques classiques
Réseaux, graphes	Modèle
Pondération	Paramètres
Apprentissage	Ajustement
Généralisation	Tests de performance
Apprentissage supervisé	Régression / Classification
Apprentissage non supervisé	Estimation de densité / Segmentation

Ces programmes ajustent donc un modèle qu'ils ont construit, permettant de simplifier la complexité et de l'utiliser de manière opérationnelle. Ce modèle est adaptatif, de façon à prendre en compte l'évolution de la base des informations pour lesquelles les comportements en réponse ont été validés (apprentissage).

Ceci permet d'auto-améliorer le système d'analyse ou de réponse (commande adaptative...), ce qui est une des formes que peut prendre l'intelligence artificielle. On assiste par ailleurs à une véritable démocratisation de ces méthodes algorithmiques qui sont disponibles, pour beaucoup, en open source ou sous forme d'une API<sup>25</sup> à des coûts très faibles.

Au-delà des différences de méthodes entre la business intelligence et la data science (Larson, Cheng, 2016), ce sont les usages et les finalités qui diffèrent, comme le synthétise le tableau 3.

TABLEAU 3 : DIFFÉRENCES D'USAGE ENTRE LA BI ET LE BIG DATA (LARSON, CHENG 2016)

Critères	Business Intelligence	Big Data
Type d'analyse	Descriptive, Prédictive	Prédictive, Prescriptive
Objectifs de l'analyse	Aide à la décision, Management de la performance	Gestion des opérations
Type de donnée	Structurées et définies	Non structurées, non définies
Temporalité	> 24h	Temps réel

25. voir une présentation des algorithmes de machine learning aisément accessibles ici : <http://www.kdnuggets.com/2015/12/machine-learning-data-science-apis.html>

L'objectif ici n'est pas de décrire l'ensemble des algorithmes de *machine learning* qui sont nombreux<sup>26</sup>. Ces algorithmes diffèrent tant par leurs objectifs que leurs méthodes d'apprentissage. Il n'en demeure pas moins qu'ils offrent tous la possibilité de construire un modèle à partir des données et non de spécifier le modèle *a priori*.

Dès lors, les données combinées à ces algorithmes permettent de créer des services qui apprennent des usages des clients ou d'un quelconque processus opérationnel. Le développement de telles ressources et compétences fondées sur les données et les algorithmes de *machine learning* conduit souvent à la construction d'un avantage concurrentiel difficile à imiter car il repose sur les données d'usage des utilisateurs ou clients, relevant d'une sorte d'accumulation primitive de capital qu'une entreprise entrante sur un marché aura du mal à combler sauf à recourir à une innovation de service majeure.

Deux cas illustrent cette logique, dans laquelle la création de valeur de l'entreprise est améliorée, enrichie par l'utilisation de tels algorithmes : Airbnb et Uber.

### Le machine learning chez Airbnb

Le modèle d'affaires d'Airbnb consiste non seulement à agréger une offre importante d'hébergements mais aussi à maximiser leur taux de location. En effet, le revenu d'Airbnb, qui consiste en une commission sur les transactions réalisées sur sa plateforme, dépend du taux de réservation des hébergements. Contrairement à une chaîne hôtelière classique, Airbnb ne maîtrise pas son offre de locations, les hôtes pouvant décider d'ouvrir ou fermer les réservations à leur convenance.

26. voir une présentation didactique de la plupart de ces algorithmes ici : <http://machinelearningmastery.com/a-tour-of-machine-learning-algorithms/>

Ils peuvent également refuser une demande de location. En outre, des offres nouvelles s'ajoutent à la plateforme sans qu'Airbnb ne maîtrise le volume d'offres. L'entreprise ne contrôle pas plus sa politique de prix puisque les prix sont fixés par les hôtes.

Afin de faire face à ces enjeux, Airbnb utilise plusieurs algorithmes qui optimisent la probabilité qu'une recherche effectuée par un voyageur aboutisse à une réservation. Ces algorithmes fournissent également aux loueurs des propositions de prix adaptés qui maximisent la probabilité de réservation.

Ces algorithmes de *machine learning* reposent sur un apprentissage à partir des données issues du site de l'entreprise, puis une capacité prédictive des comportements une fois la phase d'apprentissage terminée. Il continue par ailleurs à apprendre des données qui s'ajoutent au fur et à mesure des usages et du développement de l'entreprise. Ainsi, le modèle continue de s'ajuster au fur et à mesure que les données réelles permettent d'améliorer les capacités d'apprentissage des modèles utilisés.

### Le machine learning chez Uber

Tout comme Airbnb, Uber utilise des algorithmes de *machine learning* pour déployer son modèle d'affaires de mise en relation entre une demande de déplacement urbain et d'offre de transports des chauffeurs privés (VTC). Uber utilise par exemple un modèle Bayésien<sup>27</sup> qui lui permet de prédire la destination du client et ainsi de pré-positionner ses chauffeurs dans les localisations les plus susceptibles de voir une demande exister. Mais la proposition de valeur d'Uber réside dans la mise à disposition d'un chauffeur dans un temps minimum. Uber doit minimiser le temps d'attente du passager et maximiser le nombre de passagers pour un chauffeur. Il faut donc un modèle qui permette à l'offre et à la demande de répondre à ce double programme de maximisation. Ainsi dans la ville de San Francisco seulement 16% des taxis arrivent en moins de dix minutes contre 97% des chauffeurs VTC<sup>28</sup>. Cette disponibilité des véhicules Uber est au cœur de l'avantage concurrentiel<sup>29</sup> et d'autres études démontrent que plus Uber est présent depuis

longtemps, moins les clients ont de patience pour attendre un véhicule VTC<sup>30</sup>.

Ce programme s'avère complexe, car il peut différer d'une ville à l'autre, mais aussi dans la journée (heure creuse versus heure pleine). En effet, pendant les heures creuses, les chauffeurs sont prêts à parcourir de plus longues distances pour obtenir une course, ce qui n'est pas le cas aux heures d'affluence où une demande peut émerger très rapidement à proximité du chauffeur et où le chauffeur a intérêt à attendre quelques minutes plutôt que parcourir une plus longue distance pour obtenir un nouveau client. En utilisant un système multi-agents et en intégrant différents comportements du côté des passagers et des chauffeurs, Uber a pu déterminer que les conducteurs utilisant son système de dispatching central gagnent entre 25-50% de plus que les chauffeurs qui utilisent leurs seules connaissances de la ville pour la recherche d'un passager<sup>31</sup>.

Mais là où l'utilisation des algorithmes est encore plus saillante dans le modèle d'Uber est son modèle de tarification dynamique qui équilibre l'offre et la demande. Le système de tarification dynamique « *surge price* », mis en œuvre aux heures de pointes et à certaines périodes (soirées du vendredi et samedi, nuit du réveillon, etc.), qui multiplie le prix par un facteur 2 ou plus<sup>32</sup>, incite les chauffeurs à travailler à ces heures car ils augmentent leur revenu. Ainsi, le service Uber continue d'offrir des chauffeurs à des périodes où il est traditionnellement difficile de trouver un taxi. Les données montrent que l'offre augmente de 70 à 80% pendant ces périodes et que deux tiers des demandes sont absorbées<sup>33</sup>. Cet algorithme, introduit en 2011, a depuis évolué grâce aux données collectées sur toutes les courses et le modèle a pu se raffiner. Il concerne environ 10% de l'ensemble des courses sur la plateforme Uber<sup>34</sup>. Désormais, les clients sont avertis, lorsque la période est en « *surge price* », tout comme lorsque cette période se termine.

La collecte des évaluations des clients et des chauffeurs<sup>35</sup> constitue également une source

27. Voir le modèle utilisé pour prédire la destination d'un passager, ce qu'Uber arrive à détecter dans 74% des cas : <https://newsroom.uber.com/inferring-uber-rider-destinations/>

28. Voir l'étude du Transportation Center de l'Université de Californie : « App-Based, On-Demand Ride Services : Comparing Taxi and Ridesourcing Trips and User Characteristics in San Francisco », Août 2014 : <http://www.uctc.net/research/papers/UCTC-FR-2014-08.pdf>

29. Il s'agit du second motif d'utilisation d'un VTC après la facilité de paiement, op. cit, p. 15

30. Voir l'étude détaillée d'Uber sur l'acceptation du temps d'attente dans plusieurs villes et l'ancienneté de leur présence dans ces mêmes villes : <https://newsroom.uber.com/uber-expectations-as-we-grow/>

31. Voir l'algorithme utilisé par Uber pour optimiser les revenus des chauffeurs

32. Voir <http://uberestimate.com> pour les données de prix dans chaque ville

33. Voir le point de vue d'un des VC siégeant au conseil d'administration d'Uber : <http://abovethecrowd.com/2014/03/11/a-deeper-look-at-ubers-dynamic-pricing-model/>

34. Ibid.

35. <http://uk.businessinsider.com/leaked-charts-show-how-ubers-driver-rating-system-works-2015-2>



importante de données utilisées pour réguler le système et maintenir sa qualité et garantir la sécurité des passagers.

Ces deux cas mettent en évidence la capacité stratégique de ces entreprises à créer de la valeur à partir de leurs données et les capacités algorithmiques qu'elles ne cessent de faire évoluer grâce à l'utilisation d'algorithmes de *machine learning*. Les services concurrents n'utilisent pas les mêmes ressources et se voient concurrencés par des entreprises qui font reposer l'expérience du service sur des algorithmes qui ne cessent de s'ajuster à partir des données que la plateforme collecte.

De nouveaux algorithmes aux capacités encore plus développées, laissent penser que ces capacités adaptatives et prédictives augmenteront encore dans les années à venir, notamment grâce aux algorithmes de *deep learning*.

#### d. ...au Deep Learning

Les techniques de *deep learning* constituent une classe d'algorithmes d'apprentissage automatique qui diffère des algorithmes classiques de *machine learning* par l'utilisation plus massive de jeux de données pour l'entraînement de l'algorithme et l'utilisation de nombreuses couches d'apprentissage (jusqu'à 12, voire plus dans certains algorithmes) grâce à la puissance de calcul que fournissent désormais la mise en série de processeurs graphiques (GPU).

Utilisant différentes couches d'unité de traitement non-linéaire pour l'extraction et la transformation des caractéristiques, chaque couche prend en entrée la sortie de la précédente. Elles fonctionnent avec un apprentissage à plusieurs niveaux de détails ou de représentations des données. À travers les différentes couches, on passe de paramètres de bas niveau à des paramètres de plus haut niveau d'abstraction. Les algorithmes sont essentiellement supervisés et leurs applications comprennent la reconnaissance de modèles, d'images, la vision ou la reconnaissance vocale.

Ces algorithmes sont d'ores et déjà présents dans les services de reconnaissance vocale intégrés dans les smartphones (Siri, Google Voice ou Cortana), dans les réseaux sociaux pour la reconnaissance de visages dans les photos mises en ligne par les utilisateurs (Facebook aux USA), ou dans les véhicules autonomes comme la Tesla S. Ils ont vocation à s'intégrer dans des robots et d'autres types de machines autonomes.

Encore une fois, l'étendue des jeux de données de départ dans l'apprentissage supervisé importe, tout comme le déploiement à une large échelle qui améliore sans cesse l'efficacité de tels algorithmes, constituant autant de ressources et compétences difficiles à imiter pour les concurrents.

Le processus de création de valeur fondé sur la donnée combine à ce stade donc trois niveaux : les données, les métadonnées, les algorithmes dont on mesure l'importance dans leur capacité à produire des services à valeur ajoutée pour les utilisateurs et/ou les clients.

FIGURE 5 : ÉTAPE 3 DU PROCESSUS DE CRÉATION DE VALEUR PAR LA DONNÉE



Il n'en demeure pas moins que ce processus n'est pas complet tant que ce service n'est pas exposé à son destinataire et que celui-ci ne se l'est pas approprié grâce à une représentation efficace et une interface apte à la favoriser. C'est la dernière étape du processus de création de valeur fondé sur la donnée.

## 6. Médiation, donnée et valeur

### a. Interface, expérience utilisateur et valeur de la donnée

Quel que soit les processus de création de valeur fondé sur la donnée, celui-ci se traduit par un service à un utilisateur ou un client. Celui-ci valorisera ce service s'il satisfait au besoin ou le dépasse. La question de la restitution de la valeur à l'utilisateur dépend fondamentalement de chaque service : un service de listes de lecture musicales diffère fondamentalement d'un fil d'actualité personnalisé en temps réel d'un réseau social. Pourtant, l'un et l'autre mobilisent des données, des métadonnées, des algorithmes, mais le mécanisme d'appropriation est propre à chaque service et dépend fondamentalement de l'interface par laquelle la valeur est restituée à l'utilisateur.

Dès lors, la capacité à s'identifier à l'utilisateur dès la phase de conception est un enjeu majeur dans la création de service numérique. Une des méthodes

possibles, permettant une telle approche, est celle du design thinking développée à Stanford dans les années 1980 par Rolf Faste sur la base des travaux de Robert McKim. Contrairement à la pensée analytique, le « Design Thinking » est un ensemble d'espaces qui s'entrecroisent plutôt qu'un processus linéaire avec un début et une fin et se pense comme une démarche itérative dans laquelle le prototypage joue un rôle important pour trouver une solution adéquate. Mais d'autres méthodes viennent compléter cette approche générale telles que l'ergonomie d'interfaces.

Force est de reconnaître que l'interface appropriable demeure un enjeu majeur pour nombre d'entreprises qui ne perçoivent pas la part de la valeur qui se joue à cette étape du processus de création de valeur, tant il reste encore marqué, soit par la restitution de données plutôt que de services à valeur ajoutée, soit par la recherche d'une efficacité immédiate au détriment d'un service pour l'utilisateur. Prenons deux exemples dans lesquels cette compréhension est limitée voire contre-productive et destructrice de valeur.

Un premier exemple est constitué par les podomètres connectés. Ceux-ci affichent de nombreuses données qu'ils capturent au travers de leurs capteurs afin de les afficher à l'utilisateur. Cependant, quelle est la valeur d'un service qui affiche le fait d'avoir parcouru 4891 ou 12345 pas quotidiens ? Ces chiffres, en tant que tels, n'ont aucune valeur servicielle intrinsèque. Ce qui importe, c'est un service apte à inciter l'utilisateur à effectuer une activité physique qui maintient et prévient l'état de santé de l'utilisateur et affiche des messages incitatifs plutôt que des données brutes sans réelles valeurs d'usage. Dès lors, ici la restitution conduit souvent à des services de piètre qualité, conduisant souvent à un abandon rapide de ces services<sup>36</sup>.

Un second exemple, dans lequel la donnée échoue à produire de la valeur pour l'utilisateur est l'univers de la publicité en ligne, où jamais autant de données n'ont été mobilisées pour cibler un client potentiel, notamment dans la publicité « programmatique » ou le « *retargeting* ». La collecte systématique de données et traces (cookies, historiques de navigation, adresse IP, etc.) est massive et systématique dans cette industrie. Un tel usage conduit les internautes à se protéger et fuir de tels dispositifs, perçus comme trop intrusifs ou perturbateurs de leur expérience en

ligne. Dès lors, ces clients déploient des stratégies de contournement comme l'utilisation de logiciels de blocage publicitaire (comme AdBlock Plus) et des logiciels de blocage de mouchards en ligne (comme Goshtery). Les annonceurs et collecteurs de données utilisent et traitent des volumétries importantes de données, y compris en temps réel, mais par une exploitation trop intensive des données, conduisent les utilisateurs à se détourner de ces propositions commerciales. Au final, on assiste à une grande destruction de valeur et des comportements de rejet de la part des utilisateurs<sup>37</sup>.

Le sujet est d'autant plus complexe que la volumétrie des données augmentant, la question de leur utilisation mais aussi de leur représentation pour le client et/ou l'utilisateur devient un enjeu dans la création de valeur. C'est un des objectifs d'une fonction nouvelle, la datavisualisation.

## b. Datavisualisation

Cette problématique de représentation de larges volumes de données est un sujet ancien dans les sciences, qui ont depuis longtemps été confrontées à cette question<sup>38</sup>. Mais cette question s'étend désormais à de très nombreux services numériques pour lesquelles la question de la restitution et de la représentation des données et de leurs traitements devient un enjeu d'appropriation complexe et saillant<sup>39</sup>.

Une discipline tente de se constituer pour apporter des réponses à un tel enjeu, la datavisualisation qui par certains aspects rejoint le visual management qui tentait déjà de synthétiser l'information sous forme visuelle et que l'on retrouve dans le lean management.

En l'état actuel la datavisualisation se concentre presque exclusivement sur la publication ou dans des formes nouvelles de journalisme (le datajournalisme). Dans cette perspective, la datavisualisation est conçue comme l'art de représenter des données de façon visuelle. Cela se matérialise par des graphiques, des diagrammes, des cartographies, des chronologies, des infographies ou même des créations graphiques

36. Voir le cas de la société de traqueurs d'activité Fitbit : <http://www.fool.com/investing/general/2015/06/24/1-red-flag-that-fitbit-inc-investors-cant-ignore.aspx>

37. Seules 40% des annonces en ligne sont vues en Europe, voir <http://www.ad-exchange.fr/29269-29269/>

38. Voir par exemple : H. Doleisch, M. Gasser, H. Hauser, (2003), « Interactive Feature Specification for Focus+Context Visualization of Complex Simulation Data », EUROGRAPHICS - IEEE TCVG Symposium on Visualization. Voir également C. North, (2006) « Toward Measuring Visualization Insight », IEEE Computer Graphics and Applications, Mai/Juin, pp. 6-9.

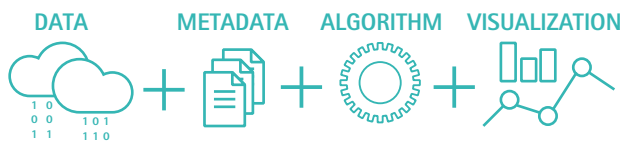
39. Berinato S., (2016), « Visualizations that really work », Harvard Business Review, June, Vol. 94 Issue 6, pp.92-100.

inédites ou encore des photos. La présentation sous une forme illustrée rend les données plus lisibles et compréhensibles. Au-delà de l'illustration, on peut même aller jusqu'à présenter des données sous la forme d'animations, voire de vidéos.

La datavisualisation reste encore trop orientée sur le seul Web et s'avère peu adaptée aux interfaces mobiles, sans parler des interfaces des objets connectés sur lesquelles la restitution des données s'orientera de plus en plus vers des interfaces homme-machine essentiellement vocales, comme Amazon Echo en est une première esquisse d'ores et déjà opérationnelle, et pour lesquels la prise de conscience d'interfaces renouvelées et dédiées est très récente<sup>40</sup>.

En conclusion, le processus de création de valeur par la donnée se représente donc comme une succession de quatre étapes fondamentales : la qualité des données, la qualité des métadonnées, l'algorithmie appliquée aux données, la restitution de la valeur au travers d'interfaces utilisateurs pertinentes, comme la figure 6 le résume.

FIGURE 6 PROCESSUS COMPLET DE CRÉATION DE VALEUR PAR LA DONNÉE



Si une ingénierie nouvelle des données se met en place avec de nouvelles fonctions dans l'entreprise pour l'orchestrer, il n'en demeure pas moins qu'à elle seule elle n'est pas en mesure de créer de la valeur s'il n'est pas au cœur des modèles d'affaires qu'elle soutient ou autorise. Dès lors c'est au niveau stratégique que la valeur de la donnée doit être appréhendée et que l'architecture des ressources nécessaires à sa création doit être pensée et architecturée dans un modèle d'affaires.

## 7. Ressources et compétences : de la gouvernance de la donnée à l'architecture de la valeur

Si le processus qui va de la donnée à la valeur s'avère complexe, il met en évidence la nécessité d'une organisation adaptée. La question de la qualité des données devrait déjà conduire les organisations à déployer une gouvernance de la donnée. Cependant, il s'agit davantage d'un prérequis que d'une réelle capacité organisationnelle. Dès lors, sans une organisation adaptée, et une capacité d'apprentissage, le déploiement du Big Data est souvent un échec (Mithas, Lee, Earley, Murugesan, 2013), ce qui était déjà le cas dans les projets de business intelligence (Peters, Wieder, Sutton, Wakefield, 2016). En outre, du fait de la nature même des méthodes du Big data, il est nécessaire d'articuler la réflexion sur les mécanismes de création de valeur et sur le modèle d'affaires avec les capacités techniques et analytiques. Ceci nécessite donc un questionnement et des réponses sur la nature des compétences à s'adjoindre et la manière de les intégrer dans l'organisation.

### a. Ressources et compétences à l'ère des données massives

Plusieurs ressources et compétences sont indispensables pour bâtir une capacité stratégique apte à extraire la valeur des données et permettre le déploiement de modèles d'affaires fondés sur la donnée.

Les compétences nécessaires procèdent de différents niveaux dans l'entreprise :

- Stratégique : capacité à définir des modèles d'affaires intégrant la donnée comme ressource centrale du processus de création de valeur ;
- Organisationnel : capacité à organiser les ressources et compétences de façon efficace, particulièrement grâce à une gouvernance de la donnée ;
- Technique : identifier les plateformes et ressources techniques les plus efficaces ;
- Managérial : capacité à gérer les différents métiers et à articuler leurs compétences dans un objectif de création de valeur.

Dans cette perspective, la capacité de collecte de données est un prérequis fondamental. Il n'est pas le seul apanage des équipes informatiques ou marketing. L'enjeu est plutôt de concevoir des produits et des services dans lesquels la collecte de données est un

40. Voir S. Lawson, « Industrial IoT still needs humans and user interfaces » : <http://www.cio.com/article/3040847/internet-of-things/industrial-iot-still-needs-humans-and-user-interfaces.html>, Mar 5, 2016

élément central dans le modèle. C'est donc dans les phases en amont de conception des produits et des services que cette question doit être envisagée et non dans les phases opérationnelles une fois le produit commercialisé. L'entreprise Shazam constitue un bon exemple. Spécialisée dans les algorithmes de reconnaissance musicale, son application mobile, collecte des données sur tous les morceaux reconnus, le lieu, l'heure, fournissant des données précieuses pour les maisons de disque. Bien plus intéressantes que les rotations des morceaux en radio, ces données sont bien sûr commercialisées par Shazam, constituant une partie non négligeable de son modèle de revenus.

Si la capacité à collecter des données est un préalable, leur exploitation nécessite des compétences spécifiques. Les seules compétences informatiques ne peuvent à elles seules garantir la capacité à créer de la valeur à partir des données. Il est nécessaire d'y adjoindre des compétences analytiques. Avec les volumétries nouvelles et l'évolution de la nature des données, celles-ci ont considérablement évoluées au cours de la dernière décennie. Les compétences analytiques sont désormais centrales dans la capacité à extraire la valeur des données. Ces compétences diffèrent des compétences historiques d'analyses présentent dans les activités de reporting et de pilotage, comme l'illustre bien la figure 7.

FIGURE 7 DIFFÉRENCES ENTRE LES ANALYSTES ET LES DATA SCIENTISTS, (HARRIS, MEHROTRA, 2014)

	ANALYSTS	DATA SCIENTISTS
Types of Data	Structured and semistructured, mostly numeric data.	All types, including unstructured, numeric and nonnumeric data (such as images, sound, text)
Preferred Tools	Statistical and modeling tools, usually contained in a data repository	Mathematical languages (such as R and Python), machine learning, natural language processing and open-source tools that access and manipulate data on multiple servers (such as Hadoop)
Nature of work	Report, predict, prescribe and optimize	Explore, discover, investigate and visualize
Typical Educational Background	Operations research, statistics, applied mathematics, predictive analytics	Computer science, data science, symbolic systems, cognitive science
Mind-set	Percentage who say they: <ul style="list-style-type: none"> <li>• are entrepreneurial: 69%</li> <li>• explore new ideas: 58%</li> <li>• gain insights outside of formal projects: 54%</li> </ul>	Percentage who say they: <ul style="list-style-type: none"> <li>• are entrepreneurial: 96%</li> <li>• explore new ideas: 85%</li> <li>• gain insights outside of formal projects: 89%</li> </ul>

Plus encore que l'identification de ces compétences, c'est la capacité à les intégrer dans les processus de décision au sein d'équipes pluridisciplinaires qui constitue le réel enjeu managérial les concernant (Harris, Mehrotra, 2014 ; Ransbotham, Kiron, Prentice, 2015, 2016).

Ce dernier point met en évidence le fait que la question de l'articulation des ressources est certainement la clé de leur efficacité. Le design organisationnel est alors l'élément clé auquel les directions doivent prêter attention.

## b. L'articulation des ressources et des compétences en données : le design organisationnel

Si l'acquisition de ressources et de compétences dans le domaine de la donnée constitue la première étape, l'essentiel proviendra de la seconde qui consiste à organiser efficacement ces ressources et compétences afin que le processus d'extraction de la valeur des données soit le plus efficace possible. Plusieurs travaux de recherche ont depuis longtemps identifié le fait que l'architecture des ressources informatiques étaient cruciale pour en retirer le meilleur bénéfice (Hitt, Brynjolfsson, 1997 ; Mendelson, 2000), rejoignant de nombreux travaux sur le design organisationnel (Galbraith JR, 2014). Le déploiement de capacités liées aux traitements des données massives (Big Data) amène à des changements organisationnels, même si ces changements peuvent s'avérer lents (Slinger, Morrison, 2014).

Les entreprises dans lesquelles les ressources et compétences sont correctement articulées ont des performances plus élevées que d'autres entreprises. Des travaux de recherche mettent en évidence que ce n'est pas la seule présence de ces compétences à l'origine de la performance mais leur répartition dans l'organisation qui est la source majeure de leur efficacité (Tambe, 2014 ; Hitt, Jin, Wu, 2015).

Dès lors, le processus d'extraction de la valeur est bien loin de se résumer à un assemblage technique de ressources et de compétences informatiques et analytiques. C'est une réelle question stratégique dans laquelle le design organisationnel des ressources et des compétences est clé dans l'obtention d'un réel avantage compétitif par la mise en oeuvre d'un modèle d'affaires fondé sur la donnée (Halaweh, El Massry, 2015).

Il en existe une grande variété. La mise en donnée du monde, liée à la digitalisation, a multiplié ceux-ci au cours de la dernière décennie. Après une présentation des différents modèles fondés sur la donnée (data-driven business models), deux d'entre eux font l'objet d'un approfondissement tant ils constituent un renouvellement profond des modèles d'affaires pour de nombreuses industries : les modèles d'affaires de l'Internet des objets, dans lesquels, la donnée sera au coeur de tous les modèles à venir ; le modèle de plateforme dans lequel la donnée est également un enjeu stratégique majeur de la création de valeur.

## 02. LES MODÈLES D'AFFAIRES FONDÉS SUR LA DONNÉE

Après avoir explicité le processus de création de valeur fondé sur les données, une analyse des modèles d'affaires fondés sur la donnée est proposée. Dans un premier temps, nous analysons les modèles d'affaires de l'industrie de la donnée, nombreux et qui souvent se combinent. Cependant, ceux-ci sont loin d'épuiser la problématique dans la mesure où, ils ne décrivent qu'une infime partie de la transformation actuelle des industries liées à la datafication du monde. Dès lors, dans un second temps, deux approfondissements sont proposés : une étude du rôle de la donnée dans les modèles de plateforme et dans les modèles de l'Internet des objets qui découlent des modèles de plateforme.

### 1. Une analyse typologique des modèles d'affaires fondés sur la donnée

Plusieurs approches des modèles d'affaires de la donnée existent. L'un centré sur l'industrie de la donnée elle-même, l'autre sur les modèles d'affaires classiques dans lesquels la donnée participe à la création de valeur.

#### a. Les modèles d'affaires dans l'industrie de la donnée

Plusieurs auteurs ont récemment proposé une analyse typologique des modèles d'affaires fondés sur les données<sup>1</sup> en analysant cent cas d'entreprises spécialisées dans la donnée. Cette approche descriptive et analytique des modèles d'affaires repose sur deux dimensions : l'origine des données et le positionnement de l'entreprise dans la chaîne de valeur de la donnée. Elle identifie principalement six modèles.

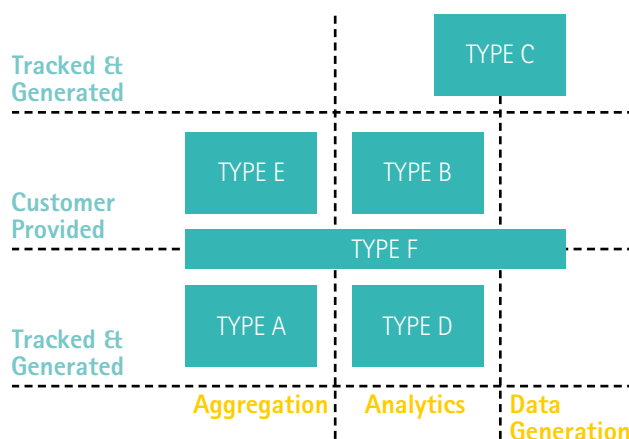
La première dimension de la typologie repose sur l'origine des données qu'utilisent les entreprises en distinguant des entreprises qui utilisent principalement des données librement et gratuitement accessibles, des entreprises qui analysent les données

fournies par leurs clients, des entreprises qui génèrent et collectent grâce à différentes méthodes des données.

La seconde dimension repose sur la nature des traitements qui s'appliquent aux données, et donc à l'activité principale sur les données, positionnant ainsi l'entreprise dans une chaîne de valeur de la donnée. Elle distingue ainsi une activité d'agrégation de la donnée d'une activité d'analyse, d'une activité de génération de nouvelles données.

La figure 8 présente la typologie obtenue par le croisement de ces deux dimensions.

FIGURE 8 : TYPOLOGIE DES MODÈLES D'AFFAIRES FONDÉS SUR LA DONNÉE (P. M. HARTMANN, M. ZAKI, N. FELDMANN, A. NEELY, (2014))



- A** Free data collector and aggregator
- B** Analytics-as-a-service
- C** Data generation and analysis
- D** Free data knowledge discovery
- E** Data-aggregation-as-a-service
- F** Multi-source data mash-up and analysis

Les entreprises du Type A sont des agrégateurs et collecteurs de données brutes accessibles gratuitement principalement sur des réseaux sociaux (Facebook, Twitter, Youtube, etc.). Leur valeur ajoutée réside dans l'agrégation et la mise à disposition de ces données souvent par le biais d'une API ou d'autre forme de visualisations.

1. P. M. Hartmann, M. Zaki, N. Feldmann, A. Neely, (2014), « Big Data for Big Business? A Taxonomy of Data-driven Business Models used by Start-up Firms », Mars, 29 p., Cambridge Service Alliance

Le type B correspond à des entreprises fournissant des services d'analyses de données de leurs clients, parfois en temps réel avec un modèle de revenus fondés principalement sur l'abonnement.

Le type C est représentatif d'entreprises qui collectent elles-mêmes les données qu'elles traitent et commercialisent auprès de leurs clients.

Le type D consiste en des entreprises qui collectent des données accessibles gratuitement et qui les analysent afin de les commercialiser auprès de leur clientèle.

Les entreprises du type E quant à elles agrègent de nombreuses sources de données appartenant ou non à leurs clients.

Enfin, les entreprises du type F, sont des sociétés qui agrègent les données fournies par leurs clients avec d'autres données externes, pour la plupart gratuites, et effectuent des analyses sur ces données. Les offres de ces entreprises se caractérisent par l'utilisation d'autres sources de données externes pour enrichir les données de leurs clients.

Pour intéressante qu'elle soit, cette typologie néglige cependant certains métiers liés à la donnée comme les *data brokers*, métier historique de la donnée, indispensables aux activités marketing de nombreuses autres industries<sup>2</sup>. Si cette typologie met en évidence la variété des positionnements possibles dans l'industrie de la donnée elle-même, elle ne permet cependant pas de comprendre comment la donnée s'intègre dans des modèles d'affaires classiques que l'on trouve dans d'autres industries que celle de la donnée.

## b. La donnée dans les modèles d'affaires « classiques »

Dans cette perspective, R. Hofman et A. van't Spijker proposent une typologie des *data-driven business models* différente, fondée sur l'utilisation du Business Models Canvas comme cadre analytique<sup>3</sup>. Ils identifient cinq modèles d'affaires génériques fondés sur la donnée.

Le premier modèle est basique : les données collectées sur une activité sont revendues à une autre entreprise sous forme anonymisées ou pas. C'est typiquement le cas d'Orange qui revend des données agrégées

de l'usage des mobiles sous la forme d'un service « Flux Vision »<sup>4</sup> à des entreprises qui ont besoin de comprendre les déplacements de leurs clients.

Le second modèle identifie l'utilisation de la donnée issue d'une première activité de l'entreprise pour innover sur un second segment de marché grâce aux données de la première activité. C'est le cas de nombreuses banques qui ont utilisé les données de ces clients bancaires pour opérer une seconde activité d'assurance auprès de ces mêmes clients.

Le troisième modèle, « *commodity swap* » repose sur la commercialisation d'un service, payant ou gratuit, qui permet la génération de données qui viennent améliorer le service de base. C'est typiquement le cas d'un compteur électrique intelligent dans lequel les données captées lors de la consommation sont potentiellement à la base d'un service complémentaire liée à l'exploitation de ces données.

Le quatrième modèle identifié repose sur l'intégration des flux de données le long d'une chaîne de valeur dans une industrie dans laquelle les entreprises échangent des données afin d'améliorer le service au client final. C'est typiquement le cas des modèles de chaîne logistique intégrée dans de nombreuses industries, comme l'automobile, ou encore la grande distribution.

Enfin, le dernier modèle, celui de la création de valeur en réseau « *Value Net Creation* », identifie le partage de données entre partenaires travaillant en réseau afin d'améliorer l'expérience globale d'un client. La différence par rapport au modèle précédent réside dans le fait que l'intégration des systèmes n'est pas nécessaire, ce que les technologies web permettent effectivement.

Encore une fois, si l'intérêt de cette typologie est indéniable, elle échoue toutefois à capturer l'essence même des modèles issus de l'économie digitalisée. En effet, la numérisation produit de nouveaux modèles d'affaires, notamment des modèles bifaces dans lesquels, la valeur est la résultante d'une interdépendance entre les deux faces d'un marché et, dans lesquels la proposition de valeur n'est pas le seul fait d'une entreprise mais d'une offre agrégée d'un écosystème.

Dès lors, une approche spécifique des modèles d'affaires des plateformes semble indispensable afin de mieux appréhender le rôle stratégique de la donnée et sa gestion dans de tels modèles. En outre, dans le

2. Voir à cet égard le rapport critique de la FTC sur l'industrie des data brokers aux USA : <https://www.ftc.gov/system/files/documents/reports/data-brokers-call-transparency-accountability-report-federal-trade-commission-may-2014/140527databrokerreport.pdf>

3. R. Hofman et A. van 't Spijker, (2013), « Patterns in data driven strategy », BlinkPaper, 11 p.

4. Voir : <https://datavenue.orange.com/flux-vision>

cadre de la digitalisation en cours, pour de nombreuses industries, avec l'encastrement de la logique réticulaire dans les objets connectés, la création de valeur fondée sur les données génère potentiellement plusieurs modèles d'affaires. Aussi, une approche spécifique de ces modèles d'affaires est nécessaire.

## 2. Données et valeur dans les modèles d'affaires de plateforme

La mise en réseau de l'information et des services produit des externalités positives au sens économique du terme<sup>5</sup>. Dans le cadre de l'économie de l'information et des réseaux électroniques, on distingue plusieurs externalités positives, les effets de réseau<sup>6</sup>. L'existence des plateformes digitales repose sur cinq types d'effets de réseau distincts<sup>7</sup>. Mais ces effets, à eux seuls, ne sont pas suffisants pour caractériser une plateforme digitale.

En effet, une plateforme relève aussi d'une logique d'écosystème qui appuie et complète les externalités positives de réseaux à l'origine des phénomènes économiques des plateformes. En outre, si une application mobile ou un site peut déclencher des effets de réseaux, cela n'implique pas pour autant qu'ils deviennent une plateforme. Dès lors, devenir une plateforme nécessite une stratégie volontariste et des décisions précises concernant l'approche technologique et le partage de la valeur entre les membres d'un écosystème.

### a. La définition d'une plateforme

Une plateforme se définit donc comme « *un service occupant une fonction d'intermédiaire dans l'accès aux informations, contenus, services ou bien édités ou fournis par des tiers. Au-delà de sa seule interface technique, elle organise et hiérarchise les contenus en vue de leur présentation et leur mise en relation aux utilisateurs finaux. À cette caractéristique commune s'ajoute parfois une dimension écosystémique caractérisée par des interrelations entre services convergents.* »<sup>8</sup>

5. L'externalité caractérise le fait qu'un agent économique crée, par son activité, un effet externe en procurant à autrui, sans contrepartie monétaire, une utilité ou un avantage de façon gratuite, ou au contraire une désutilité, un dommage sans compensation. Source : <https://fr.wikipedia.org/wiki/Externalité>

6. Voir Renaissance Numérique, (2015), « Plateformes et dynamiques concurrentielles », 34 p. [http://www.renaissancenumerique.org/images/stories/Publications/note\\_decryptage\\_plateforme\\_rn.pdf](http://www.renaissancenumerique.org/images/stories/Publications/note_decryptage_plateforme_rn.pdf)

7. Voir référence précédente

8. CNNum, (2015), Rapport "Ambition numérique: pour une politique française et européenne de la transition numérique"

La notion d'écosystèmes d'affaires forgée par James Moore en 1996<sup>9</sup> se définit comme « une communauté économique supportée par l'interaction entre des entreprises et des individus ». Un écosystème produit des biens et des services en apportant de la valeur aux clients qui font eux-mêmes partie de cet écosystème. Les membres incluent les fournisseurs, les producteurs, les concurrents et autres parties prenantes.

Les membres d'un écosystème coévoluent sur la base de leurs compétences respectives et de leurs rôles, et tendent à s'aligner eux-mêmes sur la direction d'une ou de plusieurs entreprises centrales. Le leader de l'écosystème a pour rôle d'apporter de la valeur à la communauté car il engage les membres à agir en partageant une vision pour adapter leurs investissements et trouver des rôles d'appui mutuels. L'entreprise-pivot possède une base de clients qu'elle a construite et à laquelle elle donne accès aux partenaires de l'écosystème. Elle définit les modalités de partenariats (techniques, financières), organise, anime et développe l'écosystème.

La mise en œuvre d'une telle stratégie nécessite d'adjoindre de nombreuses ressources complémentaires que ne possède pas l'entreprise-pivot et donc d'attirer des entreprises partenaires. L'interdépendance est forte et complexe : contenus/contenants, technologiques, accès aux données. On parle de coévolution des acteurs où les actions communes peuvent produire des résultats inattendus qui amènent à redéfinir les termes du partenariat.

Une plateforme se distingue d'un simple écosystème sur au moins trois éléments :

- Une plateforme exploite systématiquement les effets de réseau et plus particulièrement deux d'entre eux : indirects et croisés. Cette exploitation est le résultat d'une stratégie délibérée, ce qui n'est pas toujours le cas dans un écosystème.
- Une stratégie technologique initiale permet de bâtir les collaborations et d'ajouter des partenaires, notamment au travers d'API ou d'un SDK. La générativité de la plateforme permet de caractériser cette stratégie technologique. La notion de générativité<sup>10</sup> renvoie à la capacité plus ou moins grande d'une architecture technique à permettre le

9. Moore, James (1996), « The Death of Competition : Leadership and strategy in the age of business ecosystems », New York, Harper Business.

10. Voir Yoo Y., (2012), « The Tables Have Turned: How can the Information Systems field contribute to technology and innovation management research? », Working Paper Temple University



développement de la plateforme et donc l'étendue de celle-ci. La générativité s'apprécie au travers de plusieurs dimensions :

- Architecture des ressources informatiques (ouverte, fermée, semi-ouverte) ;
- Capacité à faciliter l'accès à celles-ci par une ou plusieurs APIs ;
- Capacité à mobiliser une communauté de développeurs ;
- Capacité à impliquer le client (*open innovation*).

• Le client est partie constitutive de la création de valeur<sup>11</sup>. La plateforme a notamment pour objectif de mobiliser le client en tant que ressource :

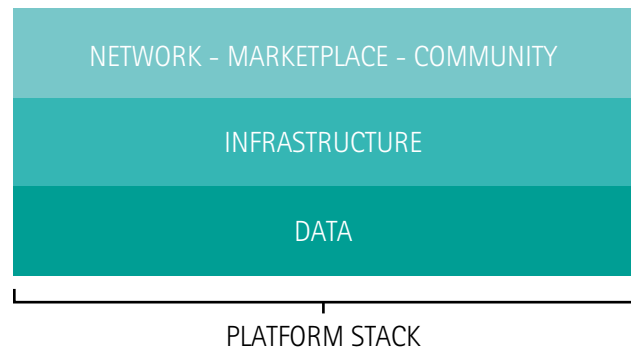
- Au travers des effets croisés, la plateforme commercialisant l'accès à cette ressource ;
- Par les données qu'il génère dans les interactions avec la plateforme (si elles font l'objet d'une exploitation par la plateforme) ;
- Par tous les mécanismes de participation du client dans la chaîne de valeur (*open innovation*, *co-conception*, *crowdsourcing*, *co-production*, *co-branding*, *marketing viral*, *communauté de clients*).

## b. La donnée dans les modèles d'affaires des plateformes

Dans une telle architecture de valeur, la question des données et du modèle d'affaires se posent d'une façon très différente de celle des modèles classiques, tels que décrits dans la section précédente.

En effet, les données constituent une ressource centrale d'une plateforme. Toute plateforme comporte trois couches interdépendantes, comme l'illustre la représentation de S. Choudary<sup>12</sup>, dont la couche basse est constituée par les données. Elle inclut tout type de données que la plateforme génère, collecte, traite, redistribue : les données des clients, les données d'usages, les données techniques, transactionnelles, etc. Selon les différents modèles de plateforme, cette couche de donnée peut-être plus ou moins importante : elle est fondamentale et centrale pour Facebook ou Google, elle est plus secondaire pour iOS d'Apple<sup>13</sup>.

FIGURE 9 : LES TROIS COMPOSANTS D'UNE PLATEFORME SELON S. CHOUDARY (2015)



Selon les plateformes, l'accès à ces données constitue une décision stratégique qui s'opère par le biais de la conception des APIs et/ou des SDKs, éléments d'infrastructure qui donneront accès à une granularité plus ou moins fine aux données, selon les décisions stratégiques de la plateforme. Ainsi, l'API OpenGraph de Facebook ouvre l'accès à une partie de ses données aux tiers, leur permettant de déployer de nouveaux services, comme le partage d'écoutes musicales (Spotify), facilitant par là même le développement de ces entreprises. Dès lors, cet accès constitue le socle de la valeur pour des entreprises tierces ; y accéder peut être stratégique pour elles. Dans le même temps, l'ouverture de la plateforme à des tiers permet de collecter d'autres données qui enrichissent celles de la plateforme.

L'accès à la plateforme (Access control dans la figure 10) et le contrôle des activités sur la plateforme (Filters dans la figure 10) sont donc deux dimensions stratégiques fondamentales d'une plateforme car elles conditionnent les contenus de la proposition de valeur de la plateforme. Ces mécanismes de pilotage de la valeur reposent fondamentalement sur une gestion des données pour maintenir et faire évoluer l'écosystème partenaire et, par là même, la proposition de valeur.

Dans cette perspective, les données sont donc une « monnaie » de la plateforme (« currency »). Des stratégies d'acquisition de données spécifiques sont déployées pour les collecter (« capture »), elles proviennent des utilisateurs de la plateforme, des partenaires, des usages, des interactions que génère la plateforme.

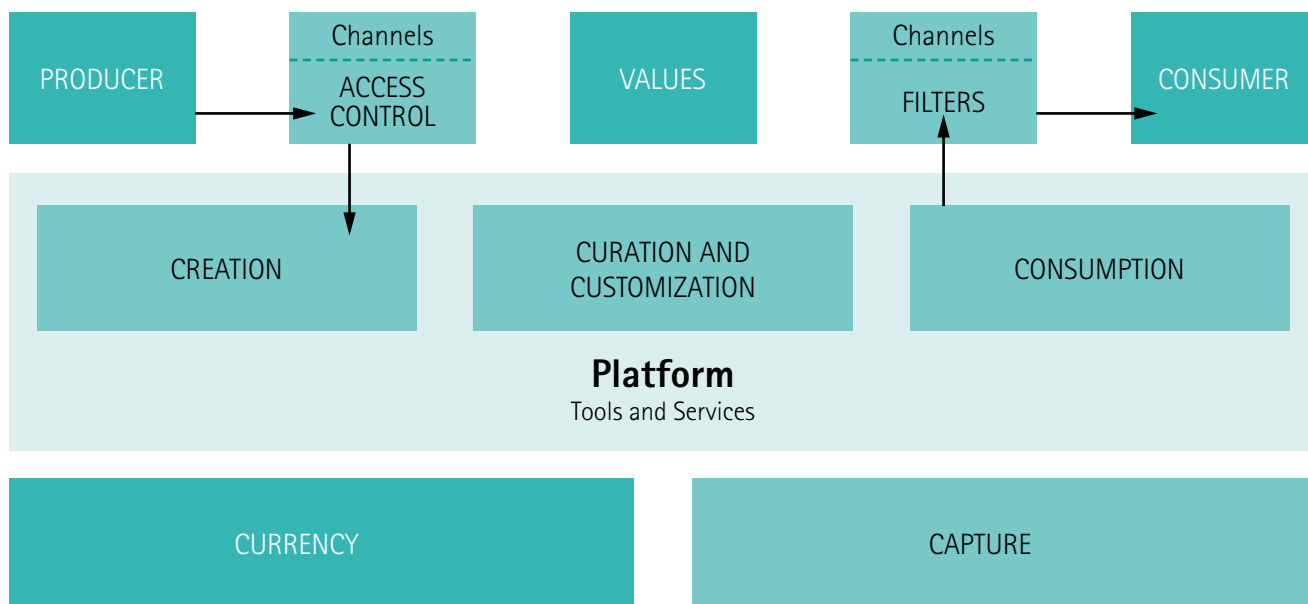
Dans cette modélisation générique des modèles d'affaires des plateformes, S. Choudary redéfinit les principes fondamentaux du management et en propose de nouveaux dans lesquels, il établit clairement deux éléments centraux identifiés dans le processus de création de la valeur : la science des données et l'algorithmique (cf. Partie I).

11. Voir « Platform Scale », Choudary S.P., qui parle de « interaction first business »

12. Sangeet Paul Choudary, (2015), « Platform scale », Platform Thinking Labs, 336 pages

13. Il existe un grand nombre de configurations possibles entre ces trois couches y compris au sein d'un même marché, voir Choudary (2015), pp. 71-77.

FIGURE 10 LE MODÈLE COMPLET DE PLATEFORME SELON S. CHOUDARY (2015)



Dans le cas d'une plateforme, la science des données remplace l'optimisation des process (BPO) de l'entreprise classique dans le sens où, les algorithmes de *machine learning* et de *deep learning* adaptent ceux-ci en temps réel par leurs mécanismes et logiques d'apprentissage intégrées. Les processus de décision sont remplacés par des logiques algorithmiques qui décident de l'allocation des ressources (prix, quantités), comme dans les modèles d'Airbnb ou d'Uber (cf. partie I).

Dans le modèle de la plateforme, la donnée est au coeur de la logique même de ces modèles qui ont désormais de très nombreuses déclinaisons dans tout type d'industrie. Ces modèles s'imposeront d'autant plus rapidement que le numérique développera la connectivité des objets, généralisant ainsi le modèle de la donnée dans les plateformes qui seront associées aux objets connectés. Dès lors, il importe de mieux comprendre le rôle de la donnée dans les mécanismes création de valeur propres aux objets connectés.

TABLEAU 4 NOUVEAUX PRINCIPES DE MANAGEMENT DES PLATEFORMES, ADAPTÉ DE S. CHOUDARY (2015)

	Company	Platform
1	Growth	Network effect
2	Human Resources Management	Community management
3	Inventory control	Liquidity management
4	Quality control	Curation & reputation
5	Sales funnels	User journeys
6	Loyalty program	Behavior design
7	Business Process Optimization	Data science
8	Sales Commission	Social feedback
9	Decision-Makers	Algorithms
10	Market Research	Real-time customization
11	Business Development	Plug and play

### 3. Donnée et valeur dans les modèles d'affaires de l'Internet des Objets

L'Internet des Objets est un champ parfois délicat à délimiter tant les modalités de connectivité sont nombreuses par l'intermédiaire de différentes technologies de réseau (Bluetooth, Wifi, 4G, Lora, etc.). Par ailleurs, l'étendue même des objets auxquels une connectivité quelconque peut être ajoutée augmente encore cette difficulté à définir le périmètre exact de ce phénomène. En effet, quelles similarités entre une poubelle connectée, une brosse à dents connectée ou une voiture connectée ?

Pour analyser le monde de l'encastrement du réseau Internet dans le monde physique des objets, il convient plutôt de raisonner en plusieurs étapes de la connectivité, comme le proposent Porter et Heppelmann<sup>14</sup>. Cette distinction permet de mieux appréhender les changements de mécanismes de création de valeur qui sont à l'œuvre dans cette transformation des objets physiques.

Elle établit également le passage du modèle d'affaires centré sur le produit à un modèle d'affaires de plateforme dans lequel le produit n'est qu'un prétexte à des prestations de service proposées par une plateforme créée autour du produit, dont la fonction réside dans le partage des données produites par l'objet connecté et enrichies par les partenaires.

#### a. Les différentes étapes de la connectivité des objets

Il est possible à partir de n'importe quel objet de raisonner sur sa connectivité et sur les modèles d'affaires qui en découlent. En partant d'un produit quelconque, il est possible de lui ajouter une fonctionnalité qui le rend intelligent grâce à des composants électroniques.

Par exemple, une chaudière peut être améliorée en lui incorporant une carte électronique qui régule la température. On passe du produit (étape 1, Figure 9) au produit intelligent doté d'une fonctionnalité supplémentaire (étape 2). En ajoutant une connectivité à ce produit on enrichit les possibilités de création de valeur. En effet, notre chaudière, devenue connectée, peut transmettre les données à un serveur qui propose des interventions en cas d'arrêt anormal de celle-ci. On assiste ici à une première migration de la valeur,

qui se déplace du produit vers un service, du fait de l'ajout d'une connectivité qui autorise un traitement des données produites par l'objet (étape 3). Le modèle d'affaires évolue donc, le produit devenant un support à des prestations de service.

Dans une étape ultérieure, les données de la chaudière pourraient être analysées en fonction des données météorologiques du lieu d'habitation et versées dans une base de données centrale qui analyse les pannes et propose des modèles de maintenance prédictive de la chaudière. La combinaison des données de la chaudière avec d'autres services permet d'accroître les services proposés et d'enrichir la proposition de valeur. La valeur continue de se déplacer vers des services, la chaudière devenant un composant dans un système de valeur (étape 4).

Dans la dernière étape, des partenaires de différentes industries accèdent à une plateforme sur laquelle ils peuvent accéder à des données produites par l'ensemble des chaudières d'un constructeur, y proposer différents services complémentaires que la plateforme autorise et contrôle par le biais de différents mécanismes : homologation, certification, contrôle d'accès aux données par l'intermédiaire d'une API. On parvient ainsi à une logique de plateforme dans laquelle la valeur est le fruit d'une plateforme et d'un écosystème (étape 5). La migration de la valeur est complète puisqu'elle n'est plus le fait du produit et de l'entreprise qui commercialise celui-ci, mais le fait d'une plateforme qui autorise et facilite la construction d'une proposition de valeur résultante d'un écosystème.

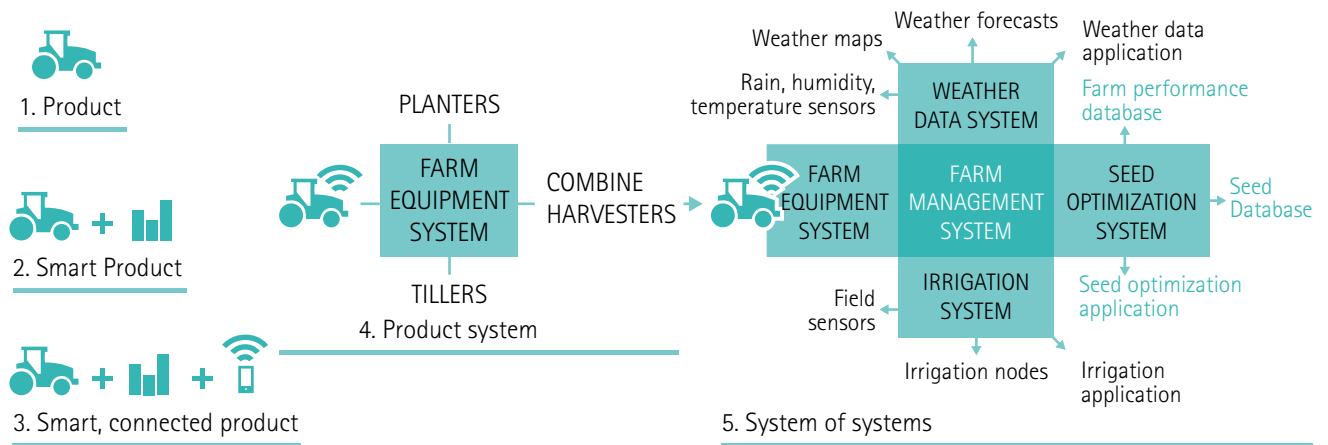
L'analyse de ces cinq étapes permet de dégager un cadre d'analyse global en raisonnant sur le processus de transformation de la donnée en valeur. En effet, dans cette logique, la création de valeur repose sur la production de données permise par la numérisation et la transmission des données issues du produit, puis à une valorisation directe ou indirecte par le biais d'une logique de plateforme.

#### b. Données, valeur et modèles d'affaires des objets connectés

La première étape du modèle de création de valeur fondé sur la donnée nécessite de doter l'objet de capteurs qui généreront les données et les transmettront. De très nombreux capteurs sont désormais possibles grâce à la miniaturisation de ceux-ci, offrant une infinité de possibilités sur un grand nombre d'objets. Cette étape consiste donc à rendre l'objet en partie « sensible » à des

14. M. E. Porter, J. E. Heppelmann, (2014), "How Smart, Connected Products Are Transforming Competition", Harvard Business review, Novembre

FIGURE 11 DU PRODUIT À LA PLATEFORME: LES DIFFÉRENTS MODÈLES D'AFFAIRES DES OBJETS CONNECTÉS, SOURCE PORTER & HEPPELMANN, 2014



usages, à son environnement, ou les deux (« *sensing* », Figure 12). Cette étape, en tant que telle ne produit guère de valeur, car les données ne font pas l'objet d'un traitement algorithmiques spécifiques et ne sont pas combinées à d'autres données, ni partagées avec des partenaires.

La seconde étape consiste à utiliser les données pour surveiller le bon fonctionnement de l'objet, apportant par exemple, par le biais d'une interface à l'utilisateur une boucle de rétroaction sur le bon ou mauvais fonctionnement de l'objet (« *monitoring* »). A ce stade la donnée ne fait l'objet que d'un traitement limité et souvent limité à l'objet, sans que celle-ci soit nécessairement transmise. On demeure dans l'objet intelligent et non dans l'objet connecté.

Les deux étapes suivantes nécessitent une transmission des données pour que les données fassent l'objet d'un

traitement qui les transforme en une information à valeur ajoutée. On peut distinguer une étape dans laquelle les données traitées permettent à l'utilisateur d'obtenir des tableaux de bord restituant les données traitées et enrichies par les traitements (« *controlling* »).

L'étape suivante repose sur une analyse plus approfondie des données notamment en les agrégeant et en leur appliquant des algorithmes plus sophistiqués qui permettent de créer de nouveaux services comme la maintenance à distance (« *diagnostics* »).

Dans l'étape suivante, l'application d'algorithmes de type machine learning permet de déployer par exemple des services de maintenance prédictive (« *prognostics* »). Les données des objets connectés sont non seulement agrégées mais alimentent des modèles d'apprentissage automatique qui en retour permettent la construction et la distribution de nouveaux services.

FIGURE 12 UN CADRE D'ANALYSE DE VALEUR DES OBJETS CONNECTÉS



La dernière étape consiste à construire une plateforme qui donne accès à des tiers et développe un écosystème de partenaires, qui à leur tour, viennent enrichir la plateforme par des services qu'ils ajoutent à la proposition de valeur initiale de l'entreprise-pivot à l'origine de l'objet connecté. Des données en provenance des tiers peuvent s'ajouter grâce à une plateforme technologique qui est accédée par des APIs ou des SDKs (« *Autonomics* »).

Dans cette perspective, la valeur de la donnée provient de son ouverture et sa réutilisation par des tiers qui peuvent l'enrichir en la combinant avec leurs données. La valeur est également conditionnée par les capacités de traitement algorithmique. Ces éléments combinés débouchent sur de nouvelles propositions de valeur qui migrent du produit au service dans une logique servicielle rendue possible par la capacité à transformer la donnée en service. Cette transformation est très bien décrite par le cadre de la théorie « *Service Dominant Logic* »<sup>15</sup> dans laquelle la valeur ne peut être que co-créée dans l'échange et ne réside pas dans l'objet lui-même. Ce cadre théorique pourrait donc bien constituer un cadre d'analyse pertinent pour l'analyse de la création de valeur de l'Internet physique qui se déploie rapidement et dans lequel, la valeur résidera fondamentalement dans la capacité à transformer des données en valeur.

---

15. Vargo, S. L. Lusch, R. F. (2004), « Evolving to a New Dominant Logic for Marketing », *Journal of Marketing* 68 (January), pp. 1-17.

## 03. CONCLUSION

A l'ère du « Big Data », nombreux sont ceux qui considèrent les données comme une nouvelle matière première et un nouvel eldorado de croissance. L'analyse développée dans ce chapitre a cherché à montrer la complexité du processus qui permet de transformer une donnée en une valeur quelconque pour l'entreprise et les clients. Cette complexité dans les mécanismes de création de valeur fondée sur les données appelle de nombreux travaux de recherche complémentaires afin de préciser les combinaisons les plus créatrices de valeur et de bâtir les modèles d'affaires pérennes. En outre, l'influence du cadre légal dans lequel se déploie cette création de valeur doit également être précisé.

Comme on l'a esquissé dans ce chapitre, la valeur d'une donnée n'est que notionnelle. Beaucoup d'étapes sont nécessaires pour en extraire sa valeur. Ces étapes nécessitent des ressources et des compétences que souvent les entreprises ne maîtrisent pas toutes simultanément. En outre, ces ressources et compétences sont de nature très différentes : techniques, analytiques, organisationnelles et stratégiques.

A cet égard, le recours à une modélisation fondée sur le modèle des options réelles est une piste possible pour élaborer un cadre général d'analyse de la valeur de la donnée. Les options réelles<sup>1</sup> ont été utilisées pour évaluer des projets réels en intégrant la valeur de flexibilité managériale et de l'incertitude dans laquelle évolue le projet. Une telle approche nous semble prometteuse tant la valeur de la donnée est dépendante de l'architecture informationnelle et technique, mais aussi stratégique de l'entreprise, notamment dans la compréhension par ses dirigeants, des enjeux d'architecture de ressources et de plateformes nécessaires pour exploiter la valeur des données de l'entreprise.

---

1. Voir une présentation simple de l'approche des options réelles : [https://fr.wikipedia.org/wiki/Analyse\\_par\\_les\\_options\\_r%C3%A9elles](https://fr.wikipedia.org/wiki/Analyse_par_les_options_r%C3%A9elles)

# Bibliographie

- Baškarada S., Koronios A., (2014), « A Critical Success Factor Framework for Information Quality Management », *Information Systems Management*, Vol. 31, pp.276–295.
- Berinato S., (2016), « Visualizations that really work », *Harvard Business Review*, June, Vol. 94 Issue 6, pp.92–100.
- Breur T., (2009), « Data quality is everyone's business -- managing information quality -- part 2. », *Journal of Direct, Data and Digital Marketing Practice*, 11(2), 114–123.
- CapGemini Consulting, (2015), *Privacy Please: Why Retailers Need to Rethink Personalization*, Octobre. <http://www.slideshare.net/capgemini/privacy-please-why-retailers-need-to-rethink-personalization>
- Choudary S.P., (2015), « Platform scale », *Platform Thinking Labs*, 336 pages
- Cigref, (2015), *L'économie des données personnelles : les enjeux d'un business éthique*, octobre, <http://www.cigref.fr/wp/wp-content/uploads/2015/11/CIGREF-Economie-donnees-perso-Enjeux-business-ethique-2015.pdf>
- Coutant A., Stenger T., (2010), « Pratiques et temporalités des réseaux socio-numériques : logique de flux et logique d'archive », *MEI-Médiation et information*
- Crié, D., Micheaux, A., (2006), « From customer data to value : What is lacking in the information chain? », *Database Marketing & Customer Strategy Management*, vol. 13, no. 4, pp. 282–99.
- Davenport, T. H., & Harris, J. G. (2007), « *Competing on analytics: the new science of winning* », Harvard Business School Press.
- Donoho D., (2015), *50 years of Data Science*, Working Paper, Sept. 18, Version 1.00
- English, L.P., (2009), *Information Quality Applied: Best Practices for Improving Business Information Processes*, Wiley Publishing Inc, Indianapolis, America.
- Erevelles S., Fukawa N., Swayne L., (2016), « Big Data consumer analytics and the transformation of marketing », *Journal of Business Research*, Vol.69, pp 897–904
- Federal Trade Commission, (2014), *Data Brokers. A call for transparency and accountability*, May, 110 p. <https://www.ftc.gov/system/files/documents/reports/data-brokers-call-transparency-accountability-report-federal-trade-commission-may-2014/140527databrokerreport.pdf>
- Galbraith JR, (2014), *Designing Organizations: Strategy, Structure, and Process at the Business Unit and Enterprise Levels*. Jossey-Bass, San Francisco, CA.
- Greenberg J., (2014), « Metadata Capital : Raising Awareness, Exploring a New Concept », *Bulletin of the Association for Information Science and Technology*, April/May, Vol. 40, N° 4, pp. 30–33
- Gregory, A., (2011), « Data governance -- protecting and unleashing the value of your customer data assets », *Journal of Direct, Data and Digital Marketing Practice*, 12 (3), pp. 230–248.
- Hartmann P. M., Zaki M., Feldmann N., Neely A., (2014), *Big Data for Big Business ? A Taxonomy of Data-driven Business Models used by Start-up Firms*, Working Paper, Cambridge service Alliance, March, 30 p.
- Hitt, L. M., and Brynjolfsson, E., (1997), "Information technology and internal firm organization: An exploratory analysis," *Journal of Management Information Systems*, pp 81–101.
- Harris, J. G., & Mehrotra, V., (2014), « Getting value from your data scientists », *MIT Sloan Management Review*, 56(1), pp. 15–18.
- R. Hofman et A. van 't Spijker, (2013), « Patterns in data driven strategy », *BlinkPaper*, 11 p.
- LaValle S., Lesser E., Shockley R., Hopkin M.S., Kruschwitz N., (2011), « Big Data, Analytics and the Path From Insights to Value », *MIT Sloan Management Review*, Vol 52, n°2, Winter, pp. 21–32
- Halaweh M., El Massry A., (2015), « Conceptual Model for Successful Implementation of Big Data in Organizations », *Journal of International Technology & Information Management*. 2015, Vol. 24 Issue 2, p21–34.
- Larson D., Chang V., (2016), « A review and future direction of agile, business intelligence, analytics and data science », *International Journal of Information Management*, Vol.36, pp. 700–710
- Mendelson, H., (2000), "Organizational architecture and success in the information technology industry," *Management science* (46:4), pp 513–529.

- Mithas, Lee, Earley, Murugesan, (2013), « Leveraging big data and business analytics », *IT Professional*, 15 (6) (2013), pp. 18-20
- Moore, James (1996), « The Death of Competition : Leadership and strategy in the age of business ecosystems », New York, Harper Business.
- Philip T. Moore and Hai V. Pham, (2015), « Personalization and rule strategies in data-intensive intelligent context-aware systems », *The Knowledge Engineering Review*, Volume 30, Special Issue 02, March, pp. 140-156.
- Opresnik D., Taisch M., (2015), « The Value of Big Data in servitization », *International Journal Production Economics*, Vol. 165, pp. 174-184.
- Orange, (2014), *The future of digital trust. A European study on the nature of consumer trust and personal data*
- Peters, Wieder, Sutton, Wakefield, (2016), « Business intelligence systems use in performance measurement capabilities: Implications for enhanced competitive advantage », *International Journal of Accounting Information Systems*, Volume 21 , Pages 1-17.
- Porter M. E., Heppelmann J. E., (2014), "How Smart, Connected Products Are Transforming Competition", *Harvard Business review*, Novembre
- Ransbotham, S., Kiron, D., & Prentice, P. K. (2016), « Beyond the hype: The hard work behind analytics success », *MIT Sloan Management Review*, 57(3)
- Ransbotham, S., Kiron, D., & Prentice, P. K. (2015), « Minding the analytics gap », *MIT Sloan Management Review*, 56(3), 63-68.
- Redman T. C., (1995), « Improve data quality for competitive advantage », *Sloan Management Review*, 36(2), 99
- Roux S., (2011), « Pour une étude des formes de la mathématisation », in *La mathématisation comme problème*, Editions des archives contemporaines, pp.3-38, <https://halshs.archives-ouvertes.fr/halshs-00813050>
- Slinger G., Morrison, R., (2014), « Will organization design be affected by big data ? », *Journal of Organization Design*, 3(3), pp. 17-26
- Tambe, P. (2014), "Big Data Investment, Skills, and Firm Value," *Management Science* (60:6), pp 1452-1469.
- Teboul B., Berthier T., (2015), « Valeur et Véracité de la donnée : Enjeux pour l'entreprise et défis pour le Data Scientist », *Actes du colloque « La donnée n'est pas donnée » cole Militaire – 23 mars 2015*, <https://hal.archives-ouvertes.fr/hal-01152219>
- Vargo, S. L. Lusch, R. F. (2004), « Evolving to a New Dominant Logic for Marketing », *Journal of Marketing* 68 (January), pp. 1-17.
- Wende K., (2007), « A Model for Data Governance – Organising Accountabilities for Data Quality Management », *ACIS 2007 Proceedings*
- Yoo Y., (2012), « The Tables Have Turned: How can the Information Systems field contribute to technology and innovation management research? », *Working Paper Temple University*
- Yoon, V. Y., Aiken, P., Guimaraes, T., (2000), « Managing organizational data resources : Quality dimensions », *Information Resources Management Journal*, Jul-Sept., Vol.13, n°3.







Place du Maréchal de Lattre de Tassigny - 75116 Paris  
Tél : +33 (0)1 44 05 43 45 - Fax : +33 (0)1 44 05 47 56  
[www.fondation.dauphine.fr](http://www.fondation.dauphine.fr)



## CHAIRE INTELLIGENCE ECONOMIQUE ET STRATÉGIE DES ORGANISATIONS

Fondation Paris-Dauphine

Si l'intelligence économique est au cœur de la stratégie de nos organisations, la prise de conscience de son rôle clef reste récente et relative en France, au-delà des cercles d'initiés. C'est pourquoi l'Université Paris-Dauphine, le groupe PSA, Sanofi, Total et la Fondation Partenariale Paris-Dauphine se sont associés pour créer la Chaire Intelligence Economique et Stratégie des Organisations en février 2011. L'objectif de cette chaire, dirigée par le Professeur Stéphanie Dameron, est de faire de l'intelligence économique une problématique clé du management. Il s'agit ainsi de développer, dans les programmes en management de l'Université Paris-Dauphine, la recherche et l'enseignement sur la veille stratégique, les opérations de soutien des avantages concurrentiels et la protection des actifs stratégiques pour toute organisation en situation de compétition économique.

