

Louvain School of Management

***Comment le partage de données
entre entreprises peut représenter
une opportunité de croissance pour
les entreprises européennes ?***

Auteurs : GAUCHET Gaspard et GÉRARD Augustin
Promoteur : BELLEFLAMME Paul
Année académique 2020-2021

Remerciements

Tout d'abord, nous aimerions remercier notre promoteur M. Paul Belleflamme pour son aide, sa disponibilité et ses conseils tout au long de la rédaction de notre mémoire.

Ensuite, nous tenions à remercier EMAsphere et, plus précisément, Florence Noblet et Marine Reillon pour avoir répondu à nos nombreuses questions.

Enfin, nous souhaiterions remercier sincèrement nos amis et familles respectives pour nous avoir soutenus durant notre parcours universitaire dont ce mémoire symbolise l'aboutissement.

Tables des matières

INTRODUCTION	1
CHAPITRE 1 - DEFINITIONS	3
1.1 La donnée	3
1.1.1 <i>Le modèle DIKW</i>	3
1.2 La donnée informatique et le Big Data	6
1.2.1 <i>La donnée informatique</i>	6
1.2.2 <i>Le Big Data</i>	7
1.2.3 <i>Typologies selon la provenance des données informatiques</i>	8
1.3 Le partage de données	10
1.4 Les entreprises européennes	10
CHAPITRE 2 - LA VALEUR DE LA DONNEE	12
2.1 Évaluer des données comme des actifs intangibles	12
2.1.1 <i>L'approche par les coûts</i>	13
2.1.2 <i>L'approche par les revenus</i>	13
2.1.3 <i>L'approche par le marché</i>	14
2.2 Gérer des données comme des biens physiques	14
2.2.1 <i>La structure de coût</i>	15
2.2.2 <i>Le raffinage</i>	15
2.2.3 <i>L'architecture</i>	15
2.2.4 <i>L'obsolescence</i>	16
CHAPITRE 3 - CONTEXTUALISATION.....	17
3.1 Vue d'ensemble du marché de la donnée	17
3.1.1 <i>Structure du marché</i>	17
3.1.2 <i>Un marché en progression</i>	19
3.1.3 <i>Une répartition inégale au sein de l'UE</i>	21
3.1.4 <i>Un retard à l'international</i>	22
3.2 Le marché du partage de données entre entreprises à l'heure actuelle.....	22

3.3	Les méthodes de partage de données	24
3.3.1	<i>Les différentes méthodes de transfert</i>	24
3.3.2	<i>Complémentarité des méthodes</i>	26
CHAPITRE 4 - TYPOLOGIE DES DIFFERENTS MODELES ECONOMIQUES		27
4.1	Partage de données entre entités.....	27
4.1.1	<i>Les alliances entre firmes concurrentes</i>	27
4.1.2	<i>Relation maison-mère et filiales</i>	32
4.2	Partage de données le long de la chaîne de valeur	34
4.2.1	<i>La chaîne de valeur de Porter</i>	34
4.2.2	<i>Le Performance Planning Value Chain (PPVC)</i>	35
4.2.3	<i>Combinaison des deux modèles</i>	38
4.3	Partage transversal de données.....	40
CHAPITRE 5 - LES OBSTACLES A L'ECHANGE DE DONNEES		43
5.1	La confiance	43
5.1.1	<i>Introduction du dilemme du prisonnier</i>	43
5.1.2	<i>Dilemme 1 : Échange de données dans un contexte de non-concurrence.....</i>	43
5.1.3	<i>Dilemme 2 : Echange de données dans un contexte de concurrence directe</i>	46
5.1.4	<i>Conclusion liée aux dilemmes</i>	47
5.2	Difficultés techniques.....	47
5.2.1	<i>Manque de standardisation et interopérabilité</i>	47
5.2.2	<i>Coût lié à l'infrastructure.....</i>	49
5.2.3	<i>Qualité insuffisante des données</i>	51
5.2.4	<i>Complexité de l'architecture des plateformes</i>	52
5.2.5	<i>Indisponibilité des compétences requises.....</i>	53
5.3	Incertitude stratégique	54
5.4	Difficultés juridiques.....	54
5.4.1	<i>Perte de contrôle sur les données.....</i>	55
5.4.2	<i>RGPD</i>	55

CHAPITRE 6 - PISTES DE SOLUTIONS POUR AMELIORER LA SITUATION	59
6.1 Standardisation.....	59
6.2 Création de plateformes d'échange de données	59
6.3 Création de la marque Europe	59
6.4 Protection des données	60
6.5 Formation du personnel.....	60
CHAPITRE 7 - LES FACTEURS CLÉS DE SUCCÈS DE L'ÉCHANGE DE DONNÉES	61
7.1 Entretenir un lien de confiance entre fournisseurs de données et utilisateurs de données	61
7.2 Comprendre les demandes liées de données	62
7.3 Bâtir des partenariats.....	62
7.4 Mettre en place des mécanismes simples et faciles d'utilisation	63
7.5 Mettre en place un cadre politique et légal propice à l'échange de données.....	64
7.6 Facteurs économiques	64
CHAPITRE 8 - ÉTUDE DE CAS : EMASPHERE	66
8.1 Sélection de l'entreprise	66
8.2 Méthodologie	66
8.3 Description de l'entreprise	67
8.4 La valeur des données chez EMAsphere.....	67
8.5 Les modèles économiques utilisés chez EMAsphere.....	68
8.6 Les obstacles rencontrés et leurs solutions.....	71
8.7 Les facteurs clés de succès	73
CHAPITRE 9 - CONCLUSIONS	74
CHAPITRE 10 - LIMITES DU MEMOIRE	76
CHAPITRE 11 - FUTURES RECHERCHES	78

BIBLIOGRAPHIE	79
TABLE DES ILLUSTRATIONS.....	87
TABLE DES ANNEXES	88

Introduction

Avec l'avènement de l'informatique, les données jouent désormais un rôle central dans le quotidien des entreprises. Ces dernières génèrent, récoltent et utilisent plus de données que jamais auparavant. Ces données leur permettent d'apprendre ce qu'apprécient leurs clients, d'améliorer leurs produits, d'être plus efficace ou encore de convaincre de nouveaux clients. Certains n'hésitent pas à considérer la donnée comme « le nouvel or digital » tant elle a le potentiel de transformer le monde de l'entreprise (Deloitte, 2021).

L'émergence de l'économie de la donnée apporte également son lot de nouveautés et, parmi celles-ci : l'échange de données. Tout comme les actifs intangibles (brevet, propriété intellectuelle ...), les données peuvent être transférées d'une entreprise à une autre. Que cela soit contre une somme d'argent (une vente) ou contre d'autres données (un troc), l'échange de données entre les entreprises devient de plus en plus courant. Selon une étude de l'Union Européenne, 40% des entreprises auraient déjà échangé des données avec une autre entreprise (European Commission & Everis Benelux, 2018).

Toutefois, lorsque nous avons exploré le sujet de l'échange de données entre entreprises, nous avons constaté qu'il s'agissait d'un sujet encore peu étudié dans la littérature scientifique. Le manque d'information scientifique était encore plus criant concernant la littérature francophone.

Ainsi, nous souhaitons participer à cette littérature scientifique francophone. D'une part, notre objectif est de synthétiser au mieux la littérature internationale à ce sujet afin de faire un état de la science sur l'échange de données entre entreprises. D'autre part, nous souhaitons attirer l'attention sur cette thématique afin que d'autres chercheurs s'emparent de ce sujet d'étude.

Ce cheminement nous a conduit à notre problématique que nous développons dans ce travail : « Comment le partage de données entre entreprises peut représenter une opportunité de croissance pour les entreprises ? ». L'objectif de ce mémoire est donc de décrire de manière synthétique le phénomène d'échange de données entre entreprises. Nous avons également

choisi de focaliser notre champ de recherche aux entreprises européennes car il s'agissait de la zone économique que nous connaissions le mieux.

Dans notre premier chapitre, nous ferons un rappel théorique des éléments essentiels de notre mémoire. Nous définirons ce qu'est une donnée, le partage de données ainsi que ce que nous entendons par « entreprises européennes ».

Dans le deuxième chapitre, nous tenterons de comprendre comment nous pouvons évaluer la valeur d'une donnée. En effet, cette dernière n'est pas un bien classique et, comme nous le verrons, il peut être complexe d'en définir la valeur.

Dans le troisième chapitre, nous ferons un état des lieux de l'économie de la donnée en Europe. Nous analyserons la structure de cette économie et nous en détaillerons l'évolution historique et future.

Dans le quatrième chapitre, nous établirons des typologies permettant de classer les différents modèles économiques liés à l'échange de données. C'est notamment dans ce chapitre que nous créerons notre propre modèle sur base de deux autres modèles existants.

Dans le cinquième chapitre, nous nous intéresserons aux obstacles à l'échange de données. Autrement dit, nous tenterons de comprendre pourquoi certaines entreprises décident de ne pas échanger de données avec d'autres entreprises.

Dans le sixième chapitre, nous esquisserons quelques pistes de solutions aux obstacles précédemment cités.

Dans le septième chapitre, nous établirons une liste des facteurs clés de succès de l'échange de données. C'est-à-dire que nous tenterons de comprendre quels sont les facteurs clés qui mènent les échanges de données entre entreprises vers le succès.

Dans le huitième et dernier chapitre, nous apporterons une illustration pratique à notre mémoire. Cela a été rendu possible grâce à notre entretien avec l'entreprise EMASphere sur leurs échanges de données avec d'autres entreprises.

Chapitre 1 - Définitions

1.1 La donnée

La donnée est l'un des concepts fondamentaux de ce mémoire. Avant d'aller plus loin, il est important de prendre le temps de définir ce qu'est une donnée. En fonction de la discipline ou de l'approche, la définition que l'on attribue à la donnée peut varier. D'un point de vue technique, ce sont « *des faits ou des informations utilisés pour des buts d'analyses ou de prises de décisions* »¹ (European Commission & Everis Benelux, 2018). Cette définition est très large, c'est pourquoi nous allons définir la manière dont nous l'abordons.

Shifra Baruchson-Arbib interprète « *la donnée comme étant un stimulus sensoriel que nous percevons par nos sens* »² (Zins, 2007). Il s'agit d'une définition plus anthropologique de la donnée. Anthony Debons apporte quant à lui une définition plus technique. Il définit les « *données comme étant des symboles organisés selon des algorithmes établis* »³ (Zins, 2007).

Ces deux exemples démontrent qu'il n'y a certainement pas un consensus définitif concernant la définition de la donnée. Nous allons quant à nous baser notre mémoire sur une définition plus fonctionnelle, mieux adaptée à notre situation. En effet, la donnée peut également être vue comme la matière première permettant de créer de l'information. C'est cette dernière qui permet d'engendrer de la connaissance (Rowley, 2007). Pour comprendre son approche, nous allons détailler le modèle Data – Information – Knowledge - Wisdom (Ackoff, 1989), aussi appelé la « Wisdom Hierarchy » (Hiérarchie de la sagesse).

1.1.1 Le modèle DIKW

C'est dans un poème de 1934 qu'émerge l'idée qu'il existe une relation entre l'information et la connaissance :

¹ Traduit de l'anglais: *Facts or information used for analytical or decision-making purposes.*

² Traduit de l'anglais: *Data are sensory stimuli that we perceive through our senses.*

³ Traduit de l'anglais: *Data are symbols organized according to established algorithms.*

« *Where is the Life we have lost in living?*

(Où est la vie que nous avons perdue dans la vie ?)

Where is the wisdom we have lost in knowledge?

(Où est la sagesse que nous avons perdue dans la connaissance ?)

Where is the knowledge we have lost in information? »

Où est la connaissance que nous avons perdue dans l'information ?

T.S. Eliot, "The Rock", Faber & Faber 1934.

Il faudra néanmoins attendre 1989 pour que cette conception entre dans la littérature scientifique grâce à l'article *From data to wisdom* (Ackoff, 1989). C'est la première fois que le concept de la séquence « Data – Information – Knowledge – Wisdom » (Donnée – Information – Connaissance – Sagesse) est évoquée (Figure 1).

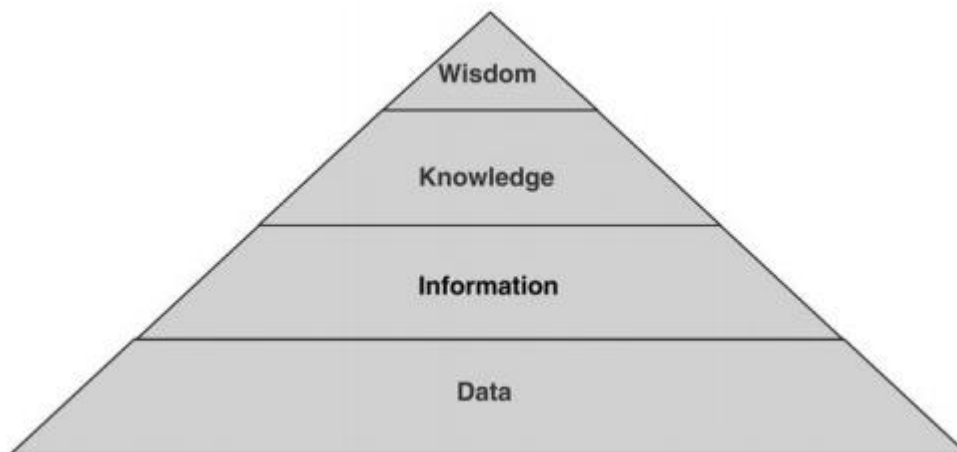


Figure 1- The DIKW Hierarchy

Selon Ackoff, les **données** sont des « *symboles qui représentent les propriétés des objets, des événements et de leur environnement. Elles sont les produits de l'observation. Mais*

elles ne sont d'aucune utilité tant qu'elles ne sont pas sous une forme utilisable »⁴ (Ackoff, 1989). Autrement dit, à ce stade-là, « *nous ne savons encore rien* »⁵ (Zeleny, 2013).

Ensuite, pour pouvoir tirer des **informations** de ces données, il faut se poser les questions « *Qui ? Quoi ? Quand ? Combien ?* ». Il s'agit en fait de remettre les données dans leur contexte et de comprendre les relations entre elles. On commence ainsi à donner du sens aux données⁶ (Bellinger *et al.*, 2004) .

Pour que les informations deviennent des **connaissances**, il va être nécessaire de se poser la question « *Comment ?* ». Autrement dit, il s'agit de rechercher des motifs récurrents entre les différentes informations. On va pouvoir alors transformer les informations en instructions. L'avantage de ces dernières est qu'il devient possible de transmettre ces instructions aux autres (Dreller, 2018).

Finalement, la **sagesse** est la « *compréhension des principes* ». C'est-à-dire répondre à la question « *Pourquoi ?* ». Cette étape permet de comprendre les phénomènes étudiés et d'en augmenter l'efficience (Bellinger *et al.*, 2004).

Il était important pour nous de revenir sur ce modèle pour comprendre l'importance de la donnée pour les entreprises. Les données sont les matériaux bruts qui permettent de comprendre des phénomènes. Comment améliorer mes ventes ? Quelles sont les caractéristiques que mes clients apprécient ? Comment promouvoir au mieux mes produits ? Toutes ces questions ont besoin de données pour pouvoir y répondre de manière empirique. Et à l'heure de la numérisation, la majorité de ces données sont sous forme informatique.

⁴ Traduit de l'anglais: *Data are defined as symbols that represent properties of objects, events and their environment. They are the products of observation. But are of no use until they are in a useable (i.e. relevant) form.*

⁵ Traduit de l'anglais: *Know-Nothing.*

⁶Traduit de l'anglais: *Information is data that has been given meaning by way of relational connection This "meaning" can be useful but does not have to be.*

1.2 *La donnée informatique et le Big Data*

1.2.1 *La donnée informatique*

Maintenant que nous avons défini les données en général, nous allons aborder plus précisément les données informatiques et le Big Data. Nous avons décidé de classer les données informatiques en deux catégories :

Les données internes - Dans cette sous-catégorie, nous plaçons toutes les informations que l'entreprise détient sur elle-même : ses employés, ses stocks, ses brevets, etc. Ce sont donc des informations qui lui sont propres.

Les données externes - Dans cette sous-catégorie, nous plaçons toutes les informations que l'entreprise détient sur son environnement : ses clients, ses fournisseurs, ses concurrents, etc. Ce sont donc des informations qu'elle a récoltées.

La particularité de ces données est qu'elles sont de nos jours entièrement informatisées. Cette informatisation permet à l'entreprise de les utiliser facilement et de les vendre/acheter en fonction de ses besoins. Néanmoins, bien qu'elles soient dématérialisées, leur stockage nécessite des installations et un personnel dédié pour maintenir l'ensemble dans un état fonctionnel.

Du point de vue académique, « *le partage de données implique généralement la collecte et l'analyse de données entre plusieurs personnes, institutions et sites de recherche. Dans ces cas-là, le partage de données va au-delà du simple échange d'ensembles de données finales, car il implique la mise en place d'infrastructures complexes d'information et de communication, ainsi que des questions d'organisation pour gérer ces collaborations* »⁷ (Birnholtz & Bietz, 2003).

⁷ Traduit de l'anglais: *Data sharing is the practice of making data used for scholarly research available to other investigators. Data sharing usually involves the collection and analysis of data across multiple individuals, institutions, and research sites. Sharing data in these cases becomes more than just the exchange of final data sets, as it involves the establishment of complex information and communication infrastructures, as well as organizational issues for managing such collaborations.*

Du point de vue concret, le partage de données est un échange d'informations entre des entreprises qui tirent toutes des avantages dans cet échange. Il s'effectue soit entre des partenaires commerciaux, entre entreprises qui partagent des produits complémentaires ou avec une entreprise spécialisée dans la récolte de données.

Ces données, bien qu'immatérielles, possèdent une grande valeur pour ceux qui savent les exploiter et sont un avantage compétitif face à des concurrents qui ne les possèderaient pas. Néanmoins, de par la manière de les obtenir et des coûts liés à leur stockage (serveurs informatiques, processeurs puissants pour traiter l'information, etc.), toutes les entreprises ne peuvent pas se permettre de le faire par elles-mêmes. C'est dans ce contexte que sont apparues des entreprises spécialisées dans ce domaine : les entreprises d'informatique en nuage (cloud computing). En 2017, Amazon Web Service était leader du marché avec 34% des parts du marché (Rama, 2017).

1.2.2 Le Big Data

La définition du Big Data a fait l'objet d'une étude complète menée par De Mauro *et al.* en 2016. Ces derniers apportent la définition suivante : « *Le Big Data est un actif informationnel caractérisé par un volume, une vélocité et une variété tellement élevés qu'il nécessite une technologie et des méthodes analytiques spécifiques pour sa transformation en valeur.* »⁸ (De Mauro *et al.*, 2016,p131). « *Il y a deux éléments qui alimentent le mouvement du Big Data : le fait qu'il y ait de plus en plus de données et l'amélioration des capacités à stocker et analyser les données* »⁹ (Marr, 2016,p2).

De par cette définition, nous pouvons déduire les défis que représentent le Big Data et plus généralement les données informatiques dans notre société qui se numérise à tout point de vue. L'accès global à Internet, l'évolution de la technologie comme des smartphones sont des

⁸ Traduit de l'anglais: *Big Data is the Information asset characterized by such a High Volume, Velocity and Variety to require specific Technology and Analytical Methods for its transformation into Value.*

⁹ Traduit de l'anglais: *There are two things that are now fuelling this Big Data movement: the fact we have more data, on anything and our improved ability to store and analyse any data.*

facteurs accélérateurs de cette dynamique (Marr, 2016). L'augmentation du télétravail à la suite de la crise du Covid-19 a encore renforcé celle-ci.

L'un des grands défis du Big Data est d'arriver à stocker toutes ces informations : bien qu'elles soient dématérialisées, elles nécessitent néanmoins des installations adaptées, notamment les serveurs. Du point de vue énergétique, « *l'ensemble des installations du stockage de données représente entre 1,1% et 1,5% de l'électricité utilisée mondialement.* »¹⁰ (Mastelic et al., 2015).

L'autre grand défi du Big Data est d'arriver à analyser l'ensemble des données. En effet, les données sont inutiles si l'entreprise ne sait pas comment en obtenir les informations qu'elle ou ses clients désirent savoir. Toute entreprise d'informatique en nuage (cloud computing) doit être apte à fournir rapidement les informations que ses clients désirent. C'est ainsi que les algorithmes d'exploration de données ont vu le jour pour accélérer le processus d'analyse.

C'est pourquoi il est important de pouvoir faire des distinctions entre les types de données récoltées, de leur provenance jusqu'à leur utilisation.

1.2.3 Typologies selon la provenance des données informatiques

Dans ce mémoire, nous ferons mention à plusieurs reprises aux données de première, deuxième ou encore troisième main. Elles sont connues plus communément dans leur dénomination anglaise : first-party, second-party et third-party data. Ces catégories permettent de classer les données informatiques selon leur provenance ; ce qui nous sera utile.

Les données de première main (first-party data)

Les données de première main font référence aux données que les entreprises ont récoltées elles-mêmes. Par exemple, les données sur leur clientèle ou la production. Ces données sont le résultat des relations et communications nouées avec leurs parties prenantes (Thar, 2019). C'est pourquoi les données de première main sont souvent considérées comme

¹⁰ Traduit de l'anglais: *Their share in power consumption generates between 1.1% and 1.5% of the total electricity use worldwide.*

les données les plus utiles. Cependant, le principal désavantage de ce type de données réside dans la quantité limitée de données accessibles. En effet, pour une entreprise, les données de première main ne peuvent provenir que de ses parties prenantes. Cela fait abstraction, par exemple, des clients potentiels.

Les données de deuxième main (second-party data)

Les données de deuxième main font allusion aux données de première main appartenant à l'origine à un autre (Schneider *et al.*, 2017). C'est le cas, par exemple, quand une entreprise tierce vend des données concernant sa clientèle à une autre entreprise. Le concept de « données de deuxième main » est donc lié de près à celui d'« échange de données ».

Les données de troisième main (third-party data)

Les données de troisième main font référence aux plateformes collectant des données de plusieurs sources. En règle générale, ces plateformes utilisent ces données pour obtenir des informations démographiques, scientifiques, techniques ou comportementales (les habitudes d'achat, les préférences, etc.). Ces informations sont ensuite revendues à des entreprises souhaitant en apprendre plus sur leur clientèle (Grether, 2016).

Utilisation dans le cadre du partage de données

Pour les données de première main, l'entreprise utilise directement les données qu'elle a recueillies pour les exploiter sauf s'il s'agit d'une entreprise spécialisée dans la collecte de données. Bien qu'elles puissent être échangées dans certains cas, il est important de rappeler qu'elles ne sont de première main que pour l'entreprise qui les a recueillies et pas pour celle qui les achète.

Pour les données de seconde main, nous sommes face à deux cas de figures :

- L'entreprise évolue au sein d'une alliance ou d'un partenariat. Ces données font donc partie des accords passés entre elles. Cette opération est effectuée pour garantir à chaque participant un maximum de données tout en diminuant les coûts supportés

par chacun. Néanmoins, elles doivent faire attention à respecter les règles de bonne concurrence et anti-cartel.

- L'entreprise veut conquérir un nouveau marché ou toucher un nouveau public cible. Pour pouvoir orienter ses actions vers la bonne cible, elle va acheter des données de première main à un acteur opérant déjà dans le milieu visé.

Pour les données de troisième main, l'entreprise va directement aller les chercher chez les entreprises spécialisées (Google, Facebook, Acxiom, etc.). L'avantage d'échanger avec ces entreprises est le grand nombre de données disponibles. En effet, puisque le business modèle de ces entreprises est la collecte et la vente de données, ces dernières doivent être les plus complètes possibles de façon à contrôler le plus de parts du marché de la vente de données.

1.3 *Le partage de données*

L'Union Européenne définit le partage de données comme étant « toutes les formes et tous les modèles possibles d'accès à des données ou de transfert de données B2B » (SWD 125, 2018). Ainsi, dans ce mémoire, nous ne parlerons pas de l'utilisation des données au sein d'une entreprise. Notre champ d'études sera focalisé sur les échanges de données entre entreprises. Autrement dit, nous porterons notre regard sur les entreprises qui génèrent des revenus externes grâce à la commercialisation de leurs données.

1.4 *Les entreprises européennes*

Dans le cadre de notre mémoire, nous avons décidé de nous focaliser sur les entreprises européennes.

Pour commencer, afin d'éviter les confusions, nous précisons que nous n'allons pas utiliser la définition de société européenne (*societas europaea*) de l'Union Européenne. **Nous définissons comme entreprise européenne les entreprises dont la majorité du chiffre d'affaires est réalisée au sein de l'Union Européenne.**

Notre choix de se focaliser sur les entreprises européennes se justifie par plusieurs facteurs. Notons tout d'abord que les auteurs de ce mémoire, tous deux européens, ont une

expertise plus avancée sur cette zone géographique. Ensuite, face à un sujet aussi large, il est essentiel de recentrer notre étude sur un champ d'application plus restreint. Enfin, notre étude a pour vocation d'être prolongée. Nous ne mettons pas de restrictions concernant la taille et le type d'entreprise.

Cependant, lors de nos comparatifs du partage de données avec d'autres échanges d'informations dans l'économie, nous ne nous restreindrons pas à étudier simplement des entreprises européennes.

Chapitre 2 - La valeur de la donnée

Afin de pouvoir échanger des données, acheteur et vendeur doivent s'accorder sur la valeur de celles-ci. Cependant, étant donné les caractéristiques intrinsèques des données, les évaluer n'est pas simple.

Il existe deux principales méthodes concernant l'évaluation d'une donnée.

D'une part, on peut considérer les données comme des actifs intangibles et utiliser des modèles existants pour les évaluer comme tels. D'autre part, il existe de nombreuses similarités entre les biens physiques et les données. Certains plaident ainsi pour qu'on évalue les deux de la même manière.

2.1 *Évaluer des données comme des actifs intangibles*

Plusieurs institutions majeures se sont penchées sur la question de l'évaluation de la donnée comme un actif intangible. Récemment, l'IVS 210 (International Valuation Standards Board) et l'IFRS 13 (International Financial Reporting Standards) ont apporté une réponse similaire pour définir la valeur d'une donnée. Ces standards listent trois grandes approches : (Moro Visconti, 2020).

- L'approche par les coûts
- L'approche par les revenus
- L'approche par le marché

Pour illustrer ces trois approches, nous utiliserons un cas d'exemple. Celui d'un département marketing disposant d'une base de données de plusieurs milliers de clients. Afin de tirer parti de ces données, l'entreprise souhaite les vendre. Nous allons donc étudier comment cette entreprise peut déterminer la valeur de cette base de données.

2.1.1 *L'approche par les coûts*

Selon cette approche, une donnée a comme valeur le coût nécessaire pour générer et entretenir celle-ci (Reilly & Schweih, 1999). Par exemple :

- Les coûts liés au travail nécessaire pour créer la donnée,
- Les prix payés lors de l'acquisition antérieure de cette donnée,
- Les coûts informatiques (logiciels, hébergement de la donnée ...)

Néanmoins, cette approche est difficile à appliquer dans la pratique. Il est difficile d'identifier avec précision les coûts liés spécifiquement à une donnée. C'est pourquoi il existe une approche alternative : l'approche du coût de remplacement. Selon celle-ci, un « *acteur du marché ne paierait pas plus pour l'actif que le coût qui serait encouru pour remplacer l'actif par un substitut d'utilité ou de fonctionnalité comparable.* » (Moro Visconti, 2020). Ainsi, on ne cherche plus à établir les coûts historiques liés à une donnée, mais bien le coût que représenterait le remplacement de celle-ci. L'un des avantages de cette approche est qu'elle prend en compte l'obsolescence des données, à savoir la perte de valeur des données au fil du temps.

2.1.2 *L'approche par les revenus*

Dans le cas de l'approche par les revenus, la valeur d'une donnée est déterminée par les bénéfices passés et futurs engendrés par celle-ci. Il s'agit essentiellement de calculer les flux de trésorerie futurs (c'est-à-dire les revenus et dépenses futurs) puis de les convertir en une valeur actualisée. Le principal défaut de cette méthode réside dans la nécessité d'évaluer ces flux de trésorerie futurs. Il peut être compliqué d'établir que tel flux de trésorerie va être lié à telle donnée. Il l'est encore plus d'évaluer quel sera le montant de ce flux de trésorerie.

Il existe néanmoins quelques méthodes permettant d'estimer ces flux :

- **L'approche des gains excédentaires** (excess earnings approach) est une méthode dérivée de l'analyse des flux de trésorerie actualisés. Cependant, au lieu de prendre en compte l'entière de l'entreprise, cette méthode isole les flux

de trésorerie qui sont associés à un ensemble de données. Il suffit ensuite d'actualiser ces flux de trésorerie pour obtenir la valeur d'un ensemble de données (Moro Visconti, 2020).

- **L'approche de l'exemption de redevance** (relief-from-royalty approach) estime la valeur d'une donnée en comptabilisant les redevances qui seraient payées à un tiers pour l'utilisation de celle-ci (Moro Visconti, 2020). Par exemple, un département marketing peut louer une base de données client pour envoyer des mails ciblés. Si ce département possédait ces données, la valeur de celles-ci serait le coût de cette location.
- **L'approche Avec-ou-sans** (with-and-without approach) compare deux scénarii. Le premier où l'entreprise ne possède pas les données. Le deuxième où elle les possède. L'entreprise calcule ensuite les flux de trésorerie actualisés afin de calculer les revenus nets générés par chacun des scénarii. L'écart entre ces deux résultats nets représente alors la valeur de ces données (Moro Visconti, 2020).
- **L'approche du terrain vierge** (greenfield approach) établit la valeur d'un ensemble de données comme étant la valeur actualisée des flux de trésorerie d'une entreprise imaginaire qui n'a aucun autre actif que cet ensemble de données (IVS 210: *Intangible Assets - Exposure Draft*, 2016).

2.1.3 L'approche par le marché

L'approche par le marché consiste à analyser le prix de vente de données semblables. La principale limite de cette approche est qu'il est parfois difficile de trouver des bases de données semblables tant chacune est unique (Moro Visconti, 2020).

2.2 Gérer des données comme des biens physiques

Il existe une tout autre école de pensée concernant l'évaluation des données : traiter ces dernières comme tout autre bien physique. Cette approche part du principe qu'il existe beaucoup de similarités entre la gestion de données et la gestion de biens physiques.

2.2.1 *La structure de coût*

Il est possible de créer un parallèle entre l'extraction de matières premières et la création de données.

D'une part, l'extraction de matières premières engendre plusieurs types de coûts : les salaires, les machines, le matériel, etc. La création de données fait intervenir des coûts similaires. Il est nécessaire d'investir dans de la main-d'œuvre et dans une infrastructure pour générer de la donnée (Fleckenstein & Fellows, 2018). On parle d'ailleurs *d'extraction de données*.

D'autre part, il y a aussi les coûts liés au stockage des matières premières. Dans le monde de la donnée, on peut interpréter ces coûts comme étant les frais d'hébergement des données. Dans un sens, les données ont une existence physique faite de zéros et de uns au sein d'un disque dur (Fleckenstein & Fellows, 2018).

Ainsi, on remarque que les structures des coûts des biens physiques et des données sont plutôt similaires ; ce qui est un premier argument en faveur de l'évaluation des données comme des biens physiques.

2.2.2 *Le raffinage*

L'une des caractéristiques intrinsèques des matières premières est leur nécessité d'être raffinées, travaillées et/ou transformées afin d'en améliorer la qualité ou l'usage. C'est également le cas avec les données. Comme nous l'avons vu avec le modèle DIKW, la donnée est la matière brute de l'information. Il est nécessaire de retravailler et nettoyer les données pour leur donner un sens. C'est ce qu'on appelle le *raffinage de données* (Morgan & Gardiner, 1990) ; ce qui démontre bien la proximité entre la gestion des données et des biens physiques.

2.2.3 *L'architecture*

Une autre similarité entre bien physique et donnée peut être trouvée dans l'architecture du stockage.

Les biens physiques sont stockés dans des entrepôts. Ceux-ci sont organisés grâce à un système de coordonnées. Par exemple, via une combinaison d'allées et de chiffres, permettant de trouver rapidement le bien. De plus, chaque entrepôt dispose d'une carte afin de s'y retrouver. Chaque objet est également identifiable grâce à un code-barre unique.

Les données quant à elles sont stockées de manière cryptée dans des serveurs. Ceux-ci sont organisés grâce à un système de bases de données et de tables permettant de trouver rapidement la donnée. De plus, chaque serveur dispose d'une carte de données (data map) afin de s'y retrouver. Chaque donnée est également identifiable grâce à un ID unique (Fleckenstein & Fellows, 2018).

On remarque une nouvelle fois une symétrie entre les deux types d'actifs.

2.2.4 *L'obsolescence*

Les données perdent de la valeur au fil du temps. En effet, avec un monde changeant rapidement, des données anciennes ne permettent plus de comprendre le monde actuel (Finger & Da Silva, 1998). Il en va de même pour les biens physiques dont la valeur diminue au fur et à mesure que le temps passe. Les biens physiques, comme les données, sont donc sujets à l'obsolescence (Fleckenstein & Fellows, 2018).

Chapitre 3 - Contextualisation

3.1 *Vue d'ensemble du marché de la donnée*

Afin de comprendre pourquoi il est intéressant pour les entreprises de s'intéresser au plus vite aux données, il est nécessaire de s'intéresser à l'état actuel et futur du marché de la donnée.

En phase avec son objectif de dynamiser ce marché, l'Union Européenne s'est équipée d'un outil d'analyse : l'« *European Data Market Monitoring Tool* » (EDM). Cet outil va nous permettre d'avoir une vue d'ensemble sur le marché de la donnée.

Avant d'aller plus loin, il est nécessaire de préciser ce que l'EDM entend par « Marché de la donnée » et « Économie de la donnée ». Par « marché de la donnée », l'EDM entend le « *marché où les données numériques sont échangées en tant que produits ou services dérivés de données brutes* ». Par « économie de la donnée », l'EDM ajoute au marché de la donnée les « *incidences économiques générées par le marché* » (The Lisbon Council, 2020).

Il nous semble également important de préciser que nous avons décidé de considérer uniquement les chiffres de l'Europe des 28 (Union Européenne + Royaume-Uni). Nous avons fait ce choix afin de garantir la continuité des données puisque certaines données historiques ne sont pas disponibles pour l'Europe des 27.

3.1.1 *Structure du marché*

L'*European Data Market Monitoring Tool* a créé un modèle permettant de synthétiser le marché de la donnée (Figure 2). Comme n'importe quel marché, celui de la donnée est défini comme un lieu de rencontre entre l'offre et la demande. D'une part, l'offre est constituée de ce qu'on appelle les « Détenteurs de données ». Ce terme regroupe les détenteurs de différents types de données, allant des informations du secteur public jusqu'aux données scientifiques, commerciales et fournies par les capteurs. En 2020, il y avait 297 350 entreprises au sein de l'UE28 considérées comme des détentrices de données. D'autre part, la demande est définie dans l'EDM comme étant des « utilisateurs de données ». Ce sont les utilisateurs finaux des

données. Autrement dit, ceux qui vont bénéficier de la valeur extraite des données. On comptait en 2020 726 110 entreprises dans l'UE28 considérées comme des utilisatrices de données (Commission européenne - European Commission, 2021a).

Tout le chemin parcouru par les données entre le détenteur de données (1) et l'utilisateur de données (3) est ce que l'EDM appelle le « Marché de la donnée » (2). On peut ainsi y retrouver les activités suivantes :

- La gestion de l'infrastructure IT (2A) – S'assurer du stockage et de l'accessibilité des données.
- L'analyse de données (2B) – Extraire de l'information grâce aux données (cf. Modèle DIKW)
- La vente / le partage de données (2C) – Le sujet sur lequel nous nous intéresserons dans ce mémoire.

Pour terminer ce tour d'horizon du marché de la donnée, reste à ajouter les autres parties prenantes (4) comme les agents régulateurs (le fisc, les douanes, etc.) ou encore les universités. Concernant les agents régulateurs, nous verrons que l'Union Européenne joue un rôle prépondérant concernant la régulation de la donnée en Europe.

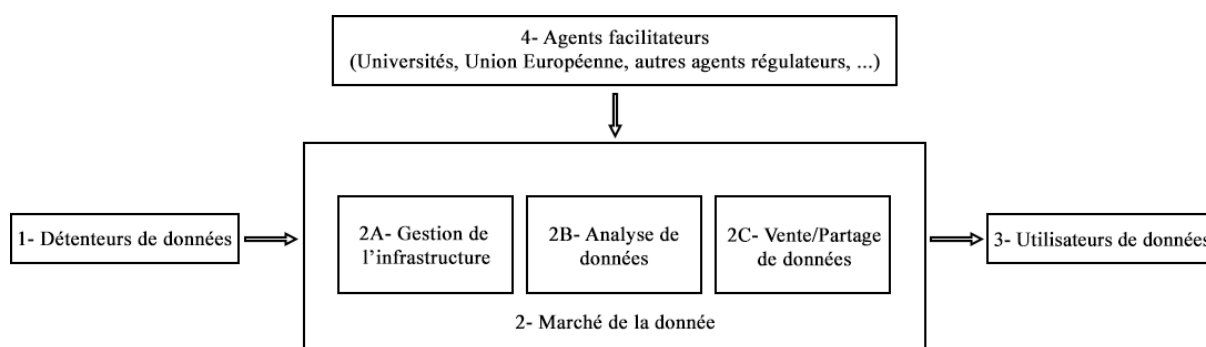


Figure 2 - La structure du marché de la donnée (Commission européenne - European Commission, 2021a)

3.1.2 Un marché en progression

Depuis près de dix ans, le marché de la donnée affiche une croissance annuelle continue entre 5% et 10% (Figure 3). Selon l'EDM, le marché a doublé de valeur durant cette même période (The Lisbon Council, 2020). Il atteint désormais plus de 80 milliards d'euros et devrait grimper jusqu'à 105 milliards d'euros d'ici 2025 (Commission européenne - European Commission, 2021a). Il est sous réserve d'un ajustement à la suite de la crise du COVID-19 (The Lisbon Council, 2020).

Cependant, il est très probable que ce secteur « *se montre résilient ces prochaines années* »¹¹ selon les experts de l'EDM. Ces derniers anticipent un rebond très rapide de l'économie dès la fin de la crise du COVID, même s'il est encore trop tôt pour chiffrer l'impact de cette crise.

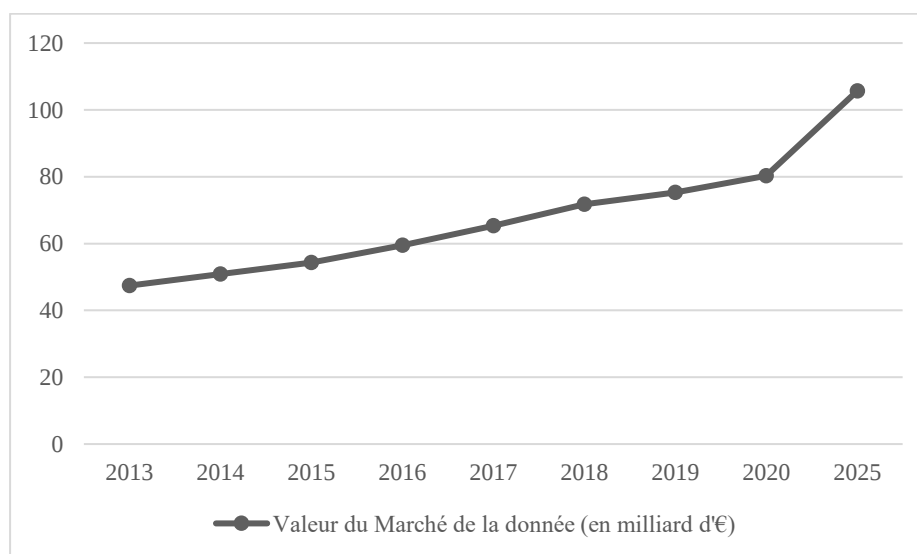


Figure 3 - Valeur du marché de la donnée (Commission européenne - European Commission, 2021a)

¹¹ Traduction de l'anglais : "In our view, the powerful negative impact of the slow-down in 2020 will be followed by a rebound and a likely return on the growth path in the next years. Many of the powerful drivers of data-driven innovation are likely to prove resilient in the next years, particularly the willingness to invest in digital technologies in order to re-launch services and create new products to stimulate demand." (The Lisbon Council, 2020)

Afin de mesurer l'ampleur du marché de la donnée, l'EDM calcule chaque année la « valeur de l'économie de la donnée ». Cet indicateur permet d'évaluer la création de valeur directe, mais aussi indirecte du marché de la donnée. En 2020, la valeur de l'économie de la donnée dans l'UE 28 (UE27 + Royaume-Uni) atteignait 444 milliards d'euros. Cela signifie que l'économie de la donnée représente à elle seule **3% du PIB de l'Union Européenne** (+ Royaume-Uni). Les prévisions sont particulièrement positives, car les experts de l'EDM estiment que ce chiffre pourrait grimper jusqu'à 675 milliards d'ici 2025 (Figure 4).

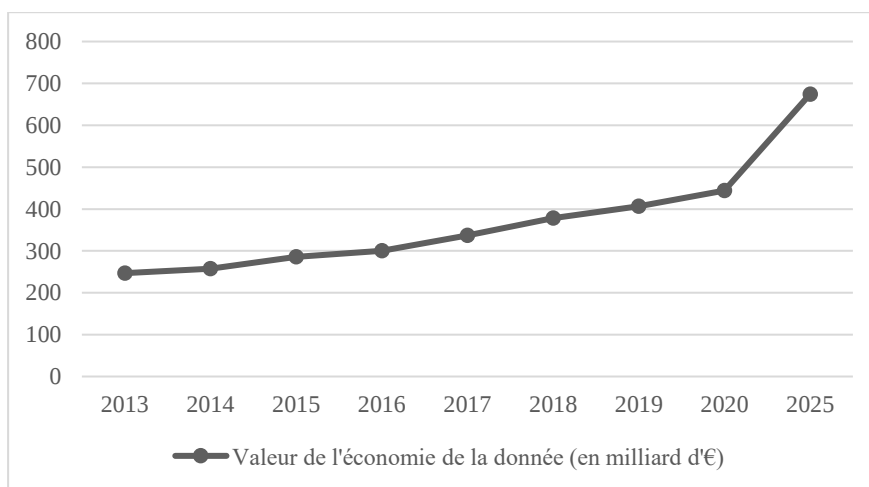


Figure 4 - Valeur de l'économie de la donnée (Commission européenne - European Commission, 2021a)

3.1.3 Une répartition inégale au sein de l'UE

Il est également intéressant de remarquer que les différents États membres de l'Union Européenne ne sont pas logés à la même enseigne (Figure 5). En termes de taille de marché, le plus large marché national au sein de l'UE est l'Allemagne. Elle représente à elle seule 28% du marché de la donnée au sein de l'UE. Avant le Brexit, le Royaume-Uni était lui aussi l'un des géants de la donnée au sein de l'Union. Ainsi, le départ du Royaume-Uni a diminué le marché de la donnée de l'Union Européenne de 23% (The Lisbon Council, 2020).

Cependant, d'autres pays -certes plus petits- possèdent également un marché de la donnée très actif. C'est principalement le cas des pays d'Europe du Nord et de quelques pays de l'Europe de l'Ouest. D'un côté, cela démontre que le marché européen est toujours assez disparate, reflétant les inégalités économiques au sein de l'UE. D'autre part, on peut y voir aussi un potentiel. Les pays de l'Est et du sud de l'Europe sont autant d'opportunités de croissance pour le marché de la donnée. Cependant, il y a encore des efforts qui doivent être fournis par l'UE pour uniformiser son marché (The Lisbon Council, 2020).

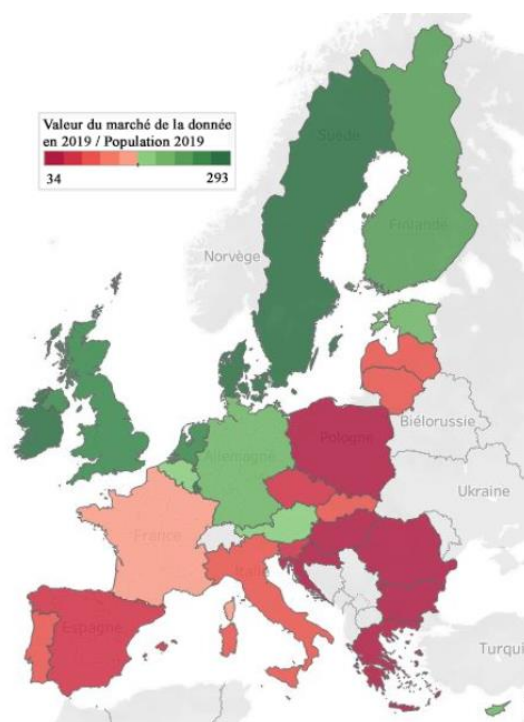


Figure 5 - Valeur du marché de la donnée en Europe¹² (The Lisbon Council, 2020)

¹² Les calculs détaillés sont accessibles dans l'annexe 1

3.1.4 Un retard à l'international

Au niveau international, l'Union Européenne accuse un retard dans l'industrie de la donnée. Comme le montre la Figure 6, le marché de la donnée est relativement deux fois plus important (à PIB égal) aux États-Unis et au Japon qu'en Europe.

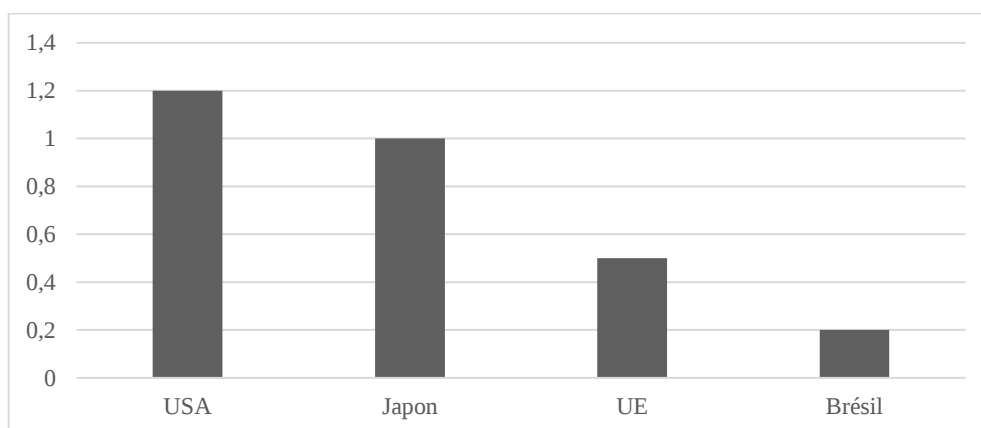


Figure 6 - Pourcentage du marché de la donnée sur le PIB (2019) (The Lisbon Council, 2020)

Même si la Commission européenne d'Ursula Von der Leyen s'est fixée comme objectif de faire de l'UE un « acteur mondial de premier plan » (Commission européenne - European Commission, 2021), l'Europe n'est pas encore en mesure de rattraper ses homologues internationaux. En effet, alors qu'en 2019 le marché de la donnée a affiché une croissance de 12,7% aux USA, de 7,2% au Brésil et de 10,5% au Japon, la croissance n'a atteint que 4,9% en Europe. Néanmoins, la Commission européenne a entamé depuis quelques années de nouveaux efforts pour pousser la croissance du secteur de la donnée en Europe (Commission européenne - European Commission, 2020). L'impact de ces mesures ne pourra être mesuré que dans quelques années.

3.2 Le marché du partage de données entre entreprises à l'heure actuelle

Le marché de la donnée est donc en pleine croissance en Europe. Mais qu'en est-il plus précisément du marché du partage de données ? Il est difficile d'apporter une réponse très précise en raison de la nature même de l'échange de données. Il s'agit essentiellement d'accords confidentiels entre deux entreprises. Il n'existe donc pas de base de données accessible pour quantifier le phénomène.

Néanmoins, il est possible d'approximer l'importance du marché du partage de données entre entreprises grâce aux différents sondages menés auprès des entreprises européennes. L'étude « *Study on data sharing between companies in Europe* », commandée par la Commission européenne, a sondé plus d'une centaine d'entreprises à travers l'Espace Économique Européen. Ce rapport (European Commission & Everis Benelux, 2018) fournit des éléments intéressants sur l'état de l'échange de données en Europe :

- Les entreprises européennes semblent partager des données entre elles. Près de 40% des répondants affirment avoir déjà partagé des données avec une autre entreprise.
- La principale raison d'acquisition des données est d'explorer la possibilité de développer de nouveaux modèles d'affaires.
- Les principales motivations des entreprises à partager leurs données sont la création de partenariats et la monétisation de leurs données.
- Le principal moyen de communication pour partager des données semble être via une API¹³.
- La principale limitation des entreprises à partager leurs données est la crainte vis-à-vis de la protection de la confidentialité de leurs données vis à vis de tiers non choisis, les règles de protection de la propriété intellectuelle et des lois contre les ententes commerciales.
- Les entreprises ne partagent donc qu'une faible proportion de leurs données. Dans le sondage, plus d'une entreprise sur deux partagent moins de 5% des données qu'elles possèdent.
- Le partage de données entre entreprises devrait suivre la même croissance que celle du marché de la donnée.

¹³ Interface de programmation applicative. Traduit de l'anglais : « *Application Programming Interface* »

3.3 Les méthodes de partage de données

Les entreprises utilisent différentes méthodes afin de partager tout type de données. Chacune avec leurs avantages et leurs inconvénients. Nous allons tenter de comprendre ces différentes méthodes.

3.3.1 Les différentes méthodes de transfert

Il existe de nombreuses méthodes permettant de transférer les données entre deux entreprises, mais nous avons décidé de nous focaliser sur les quatre développées dans l'article « *Data Exchange : Methods and Considerations* » de Charest et Rogers¹⁴.

L'API

Une API est une interface de programmation applicative¹⁵, c'est-à-dire un « Ensemble de routines standards, accessibles et documentées, qui sont destinées à faciliter au programmeur le développement d'applications » (Office Québécois de la langue française, 2005). Ainsi, une API utilise des services web pour communiquer à l'aide du protocole HTTP (Charest & Rogers, 2020). Autrement, une API est une interface conçue pour être utilisée par une autre application, et non par un humain (Figure 7). Elle permet donc d'automatiser et de faciliter l'échange de données.

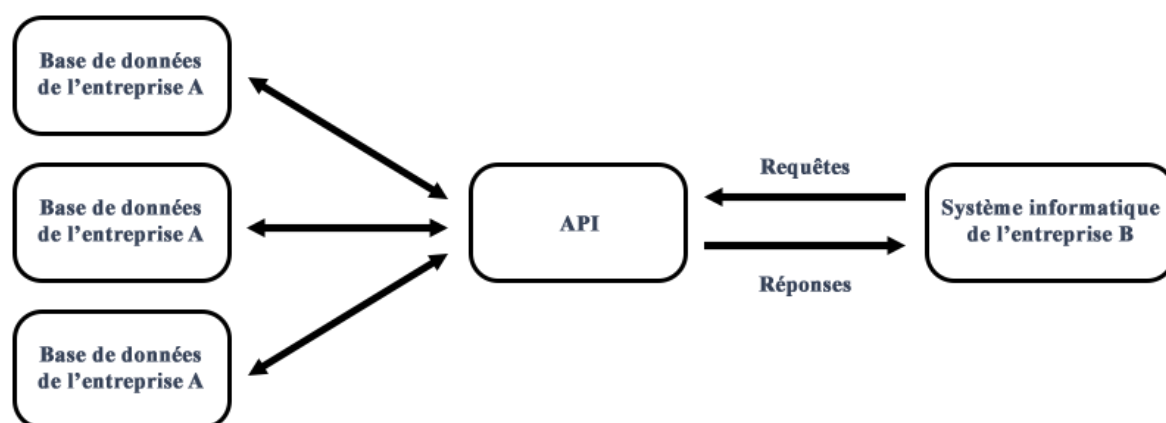


Figure 7 - Fonctionnement d'une API (IBM, 2021)

¹⁴ (Charest & Rogers, 2020)

¹⁵ Traduit de l'anglais: " Application Programming Interface "

Par exemple, une entreprise souhaite partager de la donnée avec un partenaire. Cette entreprise peut donc créer une API pour faciliter cet échange de données. Cette API servira de passerelle pour automatiser le transfert d'information. Ainsi, le système informatique du partenaire peut venir interroger lui-même l'API et cette dernière délivrera les données demandées.

La connexion directe (ETL)

Ici, la donnée est transférée via une application qui a reçu l'autorisation de se connecter directement à la base de données. Cette application va ensuite extraire, transformer et charger les données. Cette méthode est appelée « ETL », de l'anglais « *Extract, Transform and Load* » (Charest & Rogers, 2020).

Continuons avec notre exemple de l'entreprise qui souhaite partager des données avec un partenaire. Avec la méthode ETL, l'entreprise crée donc un connecteur ETL au sein de son système informatique et donne les accès de ce connecteur à son partenaire. À chaque fois que le partenaire utilisera le connecteur, il recevra une extraction complète des données.

L'envoi de fichier

Une autre méthode de transfert de données consiste à rassembler les données à partager dans un fichier organisant les données par des tableaux (généralement sous un format .JSON, .XML ou .CSV), puis à envoyer ce fichier à une autre entreprise, selon un des nombreux protocoles existants. (Charest & Rogers, 2020)

L'envoi de message

Dans ce cas-ci, on considère un message comme étant un ensemble de données envoyées à un destinataire précis. Le but de cette méthode est d'envoyer un message dès qu'un événement a lieu dans la base de données. Ces messages sont généralement gérés via un système de souscription à un type d'évènement précis. (Charest & Rogers, 2020)

Pour comprendre cette méthode, prenons l'exemple d'une plateforme de réservations de chambres d'hôtel. Cette plateforme a mis en place un système pour communiquer avec ses

hôtels partenaires sur base de messages. Ainsi, chaque hôtel partenaire recevra une notification de la part de la plateforme à chaque fois que quelqu'un réserve une de ses chambres.

3.3.2 Complémentarité des méthodes

Il est important de noter qu'il existe des avantages et inconvénients d'utiliser plus une méthode qu'une autre. D'ailleurs, il est courant que les entreprises utilisent une combinaison de ces méthodes, en fonction des usages.

Une étude comparative et technique détaillée de ces méthodes allant au-delà du cadre de ce mémoire, nous avons uniquement repris quelques critères concrets résumant les capacités des différentes méthodes :

	API	ETL	Fichier	Message
La donnée sera utilisée dans un client front-end	+	~	-	~
La donnée sera utilisée dans un système back-end	-	+	~	+
La donnée est requise dans différents formats	+	~	-	~
La donnée est fréquemment demandée	+	~	-	+
La donnée change souvent	+	~	-	+
La donnée a plusieurs flux et receveurs	~	-	-	+
La base de données est très large	-	+	+	~
La donnée doit être lisible pour un humain	-	-	+	-

Figure 8 - Résumé des avantages et inconvénients liés aux différentes méthodes de partage de données (Charest & Rogers, 2020)

Chapitre 4 - Typologie des différents modèles économiques

4.1 Partage de données entre entités

4.1.1 Les alliances entre firmes concurrentes

Avant de s'attaquer aux différents types d'alliances, il est judicieux de commencer par définir une alliance. Pour ce faire, nous nous basons sur la définition de Dominique Jolly :

Une alliance est définie comme un lien établi entre au moins deux entreprises souveraines n'appartenant pas à un même groupe qui s'entendent pour poursuivre un but conjoint dans un espace spécifié, en mettant en commun ou en échangeant des ressources, afin d'obtenir des résultats mutuellement avantageux, tout en restant indépendantes en dehors de l'alliance (Jolly, 2002, p.3).

Il est donc nécessaire d'établir de bonnes relations entre les entreprises dans le but de générer de la confiance entre celles-ci. Sur le schéma suivant (Figure 9) réalisé par Jolly, nous voyons comment sont constituées les alliances. Les deux entreprises de départ A et B apportent chacune leurs ressources à l'alliance. Il en découle qu'elles ont donc un contrôle conjoint sur cette dernière et qu'elles obtiennent toutes les deux un avantage coopératif.

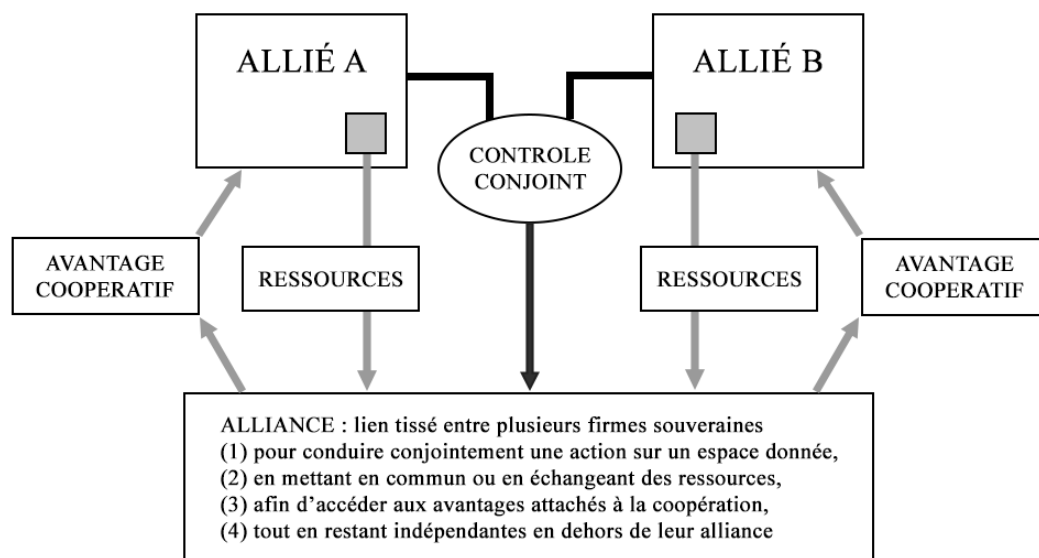


Figure 9 - La notion d'alliance en quatre points clés (Jolly, 2002)

Définitions des trois types d'alliances entre firmes

Pour définir nos différents types d'alliances entre firmes, nous nous basons sur la typologie proposée par Pierre Dussauge et Bernard Garrette dans leur livre « *Les stratégies d'alliance* ». Dans cet ouvrage, les auteurs distinguent trois grands types d'alliances entre firmes concurrentes.

Alliance dite de pseudo-concentration

Le but de ce type d'alliance est d'avoir un produit commun. Les entreprises qui fondent l'alliance s'accordent donc pour développer, produire et commercialiser ensemble un nouveau produit. Les entreprises participent à cette alliance en contribuant et en partageant leurs compétences de façon équitable. L'objectif de cette alliance est principalement de faire des économies d'échelle. Chaque membre de l'alliance participe à la hauteur des bénéfices qu'il récupère. Les alliances permettent de mutualiser des moyens de production pour assurer une meilleure couverture géographique et donner une meilleure garantie de continuité de la fourniture du produit (plusieurs moyens de productions distincts). Ces alliances sont en général métastables car le prélude à une concentration et une cessation de l'activité par l'un et l'autre membres. Cette alliance supprime la concurrence entre les entreprises qui s'allient, mais ce n'est pas pour autant que des rivalités ne subsistent pas au sein de l'alliance. Cependant, la coordination de la relation avec le client a tendance à dissimuler les tensions entre les membres. Cette alliance couvre l'ensemble de la chaîne de valeur et les positions concurrentielles sont équivalentes dans le segment visé (Dussauge & Garrette, 1995).

Les échanges de données dans ce type d'alliance sont donc généralement centrés pour réaliser un seul produit. Pour ce faire, les entreprises vont être amenées à partager leurs informations sur une partie de leur chaîne de production ainsi que sur certaines de leurs technologies. Bien que bénéfique, cette alliance peut se transformer en cheval de Troie si l'une des entreprises participantes décide d'exploiter la technologie adverse pour d'autres de ses produits.

Un exemple de ce type d'alliance est le groupe aéronautique Airbus. Ce dernier a été initié en 1970 sous la forme d'un consortium d'industriels européens dans le but d'avoir la

puissance technologique et financière nécessaire capable de produire des avions concurrentiels par rapport au géant qu'était en train de devenir Boeing, chacun des membres d'Airbus étant trop chétif pour survivre seul. Il rassemble aujourd'hui la grande majorité des constructeurs aéronautiques allemands, espagnols et français. En 2000, les plus grands acteurs européens fusionnent au sein de European Aeronautic Defence and Space company (EADS), dont Airbus devient une filiale à part entière. En 2014, le groupe prend le nom de sa filiale et s'appelle désormais Airbus Group (Airbus, 2021).

Concernant l'échange de données, Airbus, héritier de différentes alliances précédentes, possède déjà toute la technologie précédemment développée par les entreprises qui l'ont formé. Pour ce faire, elles ont toutes accepté de collaborer, partageant leurs données avec Airbus concernant l'aéronautique. Des échanges de ce type avaient déjà pu être observés sur le projet du Concorde (franco-britannique) ou sur le projet de transport Transvall (franco-allemand) (Zuliani & Jalabert, 2005).

Afin de faciliter l'échange entre les entreprises, toutes les informations sont centralisées au principal bureau d'études, situé à Toulouse dès les années 80. En plus d'effectuer ses propres recherches, le bureau d'études est en contact avec les compagnies aériennes de façon à pouvoir améliorer son offre face à leurs demandes.

Ici nous avons le cas où une alliance a été performante dans ses objectifs et, face à la fusion de ses concurrents, a adopté la même stratégie. Le partage de données est donc total dans cette nouvelle fusion.

Alliance complémentaire

L'alliance complémentaire est formée par des entreprises dont les compétences et les attributions sont complémentaires dans leur objectif commun. C'est le cas, par exemple, lorsqu'une entreprise commercialise un produit développé par une autre. En général, le produit amené par la première entreprise ne doit pas être concurrent de ceux du partenaire qui le commercialise. Cette alliance a souvent pour but d'ouvrir un produit à de nouveaux marchés stratégiques. Ce type d'alliance est très fréquent lorsqu'une entreprise veut s'internationaliser : l'entreprise entrante apporte ses produits, ses services ou ses technologies, tandis que

l'entreprise présente sur le marché/pays apporte le plus souvent son réseau de distribution déjà solidement installé (Dussauge & Garrette, 1995).

Les échanges de données de ce type d'alliance sont plus généraux. Par exemple, une entreprise peut apporter toutes les informations stratégiques là où l'autre entreprise peut apporter les informations logistiques. C'est une opération risquée qui nécessite beaucoup de confiance. En effet, aucune des deux entreprises n'est certaine que l'autre ne se retournera pas contre la première et lui dérobera ses informations et données.

Un exemple de ce type d'alliance est l'alliance entre Renault et Nissan. Un exemple de ce type d'alliance est l'alliance entre Renault et Nissan. Les deux entreprises ont créé celle-ci dans le but de développer une stratégie ainsi que des synergies communes. Dans les faits, en plus des projets développés en commun, chaque entreprise vend les produits de l'autre sous sa propre marque en fonction du marché. Par exemple, la Renault Clio est vendue sous le nom Nissan Platina au Mexique.

Cette alliance a été créée en deux étapes. Au départ, Nissan et Renault créent conjointement une joint-venture, Renault Nissan BV, détenue à parité. Le but de cette alliance est de créer des projets communs. Pour arriver à cet objectif, les deux entreprises s'étaient accordées pour partager leur expertise et leur savoir-faire ainsi que de mettre en place des stratégies communes pour produire à un coût réduit. Les deux entreprises voulaient néanmoins garder leur identité et leurs produits (Segrestin, 2005).

Le fonctionnement de l'alliance se fait au travers de la joint-venture. Les équipes n'ont pas directement fusionné lors de l'alliance. Chaque groupe d'équipes travaillait de manière indépendante. Ce fonctionnement a été mis en place pour éviter des investissements coûteux et ne pas devoir repartir de zéro au cas où l'alliance échouerait. Le partage de données ne s'est donc pas fait immédiatement. Il y a un temps de transition pour installer la confiance entre les deux entités (Segrestin, 2005).

Pour instaurer la confiance entre les deux entreprises, celles-ci ont décidé de commencer à coopérer sur des projets de taille limitée. Pour mettre en place cette dynamique, les entreprises ont créé une plateforme commune sur laquelle sont mis les composants, les besoins de chaque

équipe, les arbitrages en cas de problème, etc. Les constructeurs ont d'ailleurs continué d'approfondir leur alliance par la création de nouvelles structures communes (Segrestin, 2005). Mais ce n'est pas pour autant que toutes les tensions entre les entreprises ont disparu, l'affaire Carlos Ghosn a montré que leurs intérêts n'étaient pas toujours communs. En effet, les enquêtes et accusations sont parties de Hari Nada, vice-président de Nissan, ancien bras droit de Ghosn et administrateur de Nissan Zi-A Capital BV (« Affaire Carlos Ghosn », 2021).

Alliance de co-intégration

Pour ce troisième type d'alliance, les entreprises ont pour objectif de réaliser des économies d'échelle sur un composant bien précis ou sur un stade particulier d'un processus de production. En général, elles s'accordent de s'entraider pour de la recherche et développement ou de la production. Contrairement à l'alliance complémentaire, la concurrence peut rester frontale dans le cas de cette alliance. En effet, les produits des entreprises alliées peuvent être substituables. Par conséquent, elles sont donc sur des positions concurrentielles équivalentes. Le mode d'organisation le plus fréquent est la création d'unités ou d'ateliers spécifiques et communs aux entreprises (Dussauge & Garrette, 1995).

Les échanges de données dans ce type d'alliance sont plus limités : seules les données nécessaires au partenariat sont échangées. De par la nature de cette alliance, les entreprises prendront beaucoup de précautions concernant les échanges de données qu'elles effectuent avec leurs partenaires.

Un exemple de ce type d'alliance est l'alliance d'Apple, IBM et Motorola pour concevoir le PowerPC. En 1991, les trois entreprises se sont alliées pour concevoir et fabriquer un processeur qui puisse rivaliser avec Microsoft et Intel (« AIM Alliance », 2021).

Au sein de l'alliance, IBM et Motorola avaient créé deux filiales, Taligent and Kaleida Labs, contrairement à Apple. Les informations et données échangées ne concernaient que celles nécessaires à la conception et production du PowerPC. L'alliance prit fin en 2007 avec la fin du PowerPC (« AIM Alliance », 2021).

4.1.2 *Relation maison-mère et filiales*

Cette relation est caractérisée par un déséquilibre entre les deux entités. Par définition, une filiale est une entreprise contrôlée par une autre, la maison-mère. La maison-mère contrôlant la filiale, elle peut exiger de cette dernière l'ensemble des données que celle-ci possède alors que la réciproque n'est pas vraie. Les données que la maison-mère transmet à sa filiale sont celles qu'elle accepte de lui fournir. Dès lors, la filiale sert uniquement les intérêts de sa maison-mère et opère avec une indépendance limitée, accomplissant les objectifs fixés.

Les échanges de données dans ce type d'alliance sont diversifiés : la filiale fournit à la maison-mère toutes les données que celle-ci requiert alors que la maison-mère ne fournit à la filiale que les données dont cette dernière a besoin pour son fonctionnement.

Pour cette alliance, nous nous baserons sur l'étude *Knowledge Flows within Multinational Corporations* de Anil K. Gupta et Vijay Govindarajan. Ils ont voulu déterminer s'il y a une valeur ajoutée à la transmission de savoir entre la maison-mère et ses filiales ainsi qu'entre les filiales. Nous allons étudier leurs résultats pour décrire comment se déroule le partage de données dans ce type d'alliance.

Concernant l'échange de données entre filiales, Gupta et Govindarajan ont observé que l'échange de données était plus important lorsque les filiales étaient acquises et intégrées au groupe. En effet, l'arrivée d'une nouvelle entreprise dans un groupe pousse celle-ci à communiquer et partager ses données avec le reste du groupe (Gupta & Govindarajan, 2000). Un exemple de cette situation est lorsqu'une entreprise est achetée par un grand groupe pour un de ses brevets, ce qui est par exemple souvent le cas dans le milieu pharmaceutique (Johnson&Johnson a acheté Momenta Pharmaceuticals en août 2020 pour obtenir le nupocolimab) (« Johnson & Johnson achète Momenta pour environ 6,5 milliards de dollars », 2020).

Concernant l'échange de données entre la maison-mère et ses filiales, Gupta et Govindarajan ont observé le même phénomène : l'échange de données était plus important lorsque les filiales étaient acquises et intégrées au groupe (Gupta & Govindarajan, 2000). Face à ces résultats, il est à noter que les échanges entre maison-mère et filiales sont nombreux. Nous

assimilons cette information au fait que même si la maison-mère et ses filiales sont des entités différentes, elles évoluent néanmoins au sein d'un même groupe et doivent pour cela communiquer entre elles de façon régulière.

Les recherches de Gupta et Govindarajan se sont intéressées uniquement à l'échange entre maison-mère et filiales. Pour savoir si certains types d'entreprises ont tendance à effectuer plus d'échanges que d'autres, nous nous sommes penchés sur l'étude *Intra-organizational communication and its consequences* réalisée par Manami Suzuki, Naoki Ando et Hidehiko Nishikawa.

Leur première observation est que les entreprises qui échangent le plus de données sont celles qui communiquent le plus entre elles. Les auteurs parlent de la communication en tant qu'action et non en tant qu'exercice sémantique. Une fréquence importante ainsi que l'importance perçue de la communication sont deux facteurs qui influencent positivement l'échange entre les entreprises (Suzuki *et al.*, 2019).

Leur deuxième observation concerne la différence des échanges entre les entreprises de service et les entreprises de production. Ils ont observé que les entreprises de service obtiennent plus de bénéfices à communiquer que les entreprises de production. Les résultats impliquent que la communication avec les entreprises mères rend les connaissances des entreprises de services moins rigides et plus transférables¹⁶(Suzuki *et al.*, 2019).

Comme l'étude de Gupta et Govindarajan, l'étude de Suzuki, Ando et Nishikawa montre que la communication et les échanges sont plus importants avec les filiales acquises qu'avec celles qu'elles créent (Suzuki *et al.*, 2019).

En conclusion, le partage de données entre maison-mère et filiales est facilité par la communication et le type de filiale lors de la création de celles-ci. Les entreprises de service ont tendance à plus échanger que leurs homologues de production.

¹⁶ Traduit de l'anglais: *The findings imply that communication with parent firms makes knowledge of service firms less sticky and more transferable.*

4.2 Partage de données le long de la chaîne de valeur

4.2.1 La chaîne de valeur de Porter

La chaîne de valeur telle qu'elle est décrite par son concepteur est :

« L'idée de la chaîne de valeur est basée sur la vision des organisations en termes de processus, l'idée de voir une organisation de fabrication (ou de services) comme un système, composé de sous-systèmes ayant chacun des entrées, des processus de transformation et des sorties. Les inputs, les processus de transformation et les outputs impliquent l'acquisition et la consommation de ressources - argent, travail, matériaux, équipements, bâtiments, terrains, administration et gestion. La manière dont les activités de la chaîne de valeur sont réalisées détermine les coûts et affecte les bénéfices. »¹⁷(Porter, 1985)

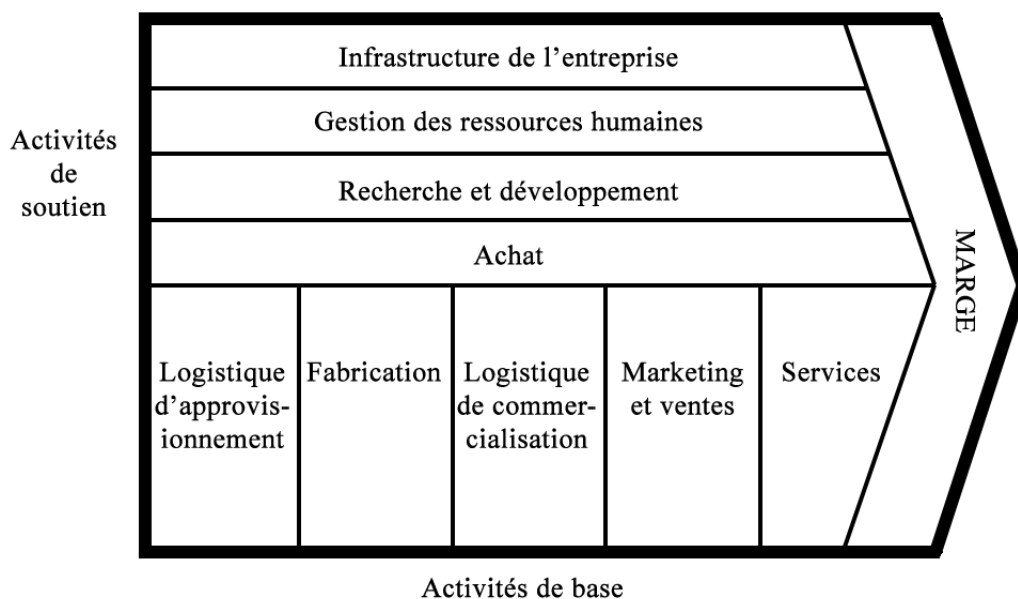


Figure 10 - La chaîne de valeur de Porter

¹⁷ Traduit de l'anglais: The idea of the value chain is based on the process view of organizations, the idea of seeing a manufacturing (or service) organization as a system, made up of subsystems each with inputs, transformation processes and outputs. Inputs, transformation processes, and outputs involve the acquisition and consumption of resources – money, labour, materials, equipment, buildings, land, administration and management. How value chain activities are carried out determines costs and affects profits.

La chaîne de valeur est divisée en deux parties : les activités de base et les activités de soutien. Les activités de base sont toutes les activités qui sont directement impliquées dans la conception du produit. Sur la chaîne de valeur, elles apparaissent dans l'ordre chronologique de la « vie » du produit de sa fabrication jusqu'à sa vente. Les activités de soutien sont toutes les activités qui forment la structure de l'entreprise et qui servent de soutien aux activités de base. Elles interviennent à toutes les étapes de la « vie » du produit au sein de l'entreprise.

Dans cette chaîne de valeur, l'échange de données se fait donc de plusieurs manières : entre les activités de base, entre les activités de soutien ainsi qu'entre les deux types d'activités. L'objectif de cette chaîne de valeur, en incluant le partage de données qui s'y déroule, est le suivant : identifier et optimiser les produits/services de l'entreprise qui lui apportent de la valeur.

4.2.2 Le Performance Planning Value Chain (PPVC)

Le Centre for Business Performance (Cranfield School of Management) a développé un processus pour faciliter l'analyse de la chaîne de valeur de Porter du point de vue des données : le Performance Planning Value Chain (PPVC). Comme le décrivent Andy Neely et Yasar Jarrar, son but est de :

- Permettre de réunir les données brutes et de les transformer de façon à obtenir des informations exploitables dans le processus de prise de décisions¹⁸ (Neely & Jarrar, 2004).
- Fournir un processus de transformation des données - souvent désorganisées ou dispersées dans leur forme originale - en informations de haute qualité et à valeur ajoutée qui permettent aux utilisateurs de prendre des décisions plus efficaces¹⁹ (Neely & Jarrar, 2004).

¹⁸ Traduit de l'anglais: *It provides a systemic process for using data to enhance decision-making, bringing together a vast array of tools to extract value from data and focus efforts on what will add real value to the organization.*

¹⁹ Traduit de l'anglais : *provide a process for the transformation of data - often disorganised or dispersed in its original form - into high-quality, value-added information that enables the users to make more effective decisions.*

- Fournir un processus qui rassemble de manière cohérente une combinaison de compétences pour analyser et interpréter des informations complexes provenant de sources diverses et la capacité de présenter des informations techniques complexes à des non-spécialistes, ainsi qu'une capacité à ajouter des idées²⁰ (Neely & Jarrar, 2004).

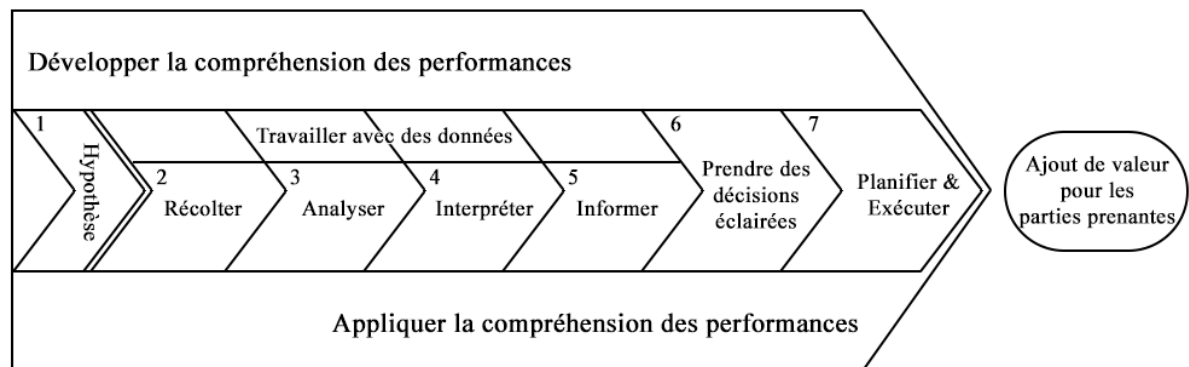


Figure 11 - Performance planning value chain²¹ (Neely & Jarrar, 2004)

Comme on peut le voir sur leur schéma (Figure 11), Neely et Jarrar ont résumé cette méthode en sept étapes :

1. **Développer une hypothèse** - Quels sont les objectifs de l'entreprise ?
2. **Recueillir les données** - En fonction des objectifs fixés, quelles sont les données que l'entreprise doit récolter ?
3. **Analyser les données** - Une fois les données récoltées, comment doivent-elles être analysées ? Que disent-elles ?
4. **Interpréter les données** - Comment utiliser ces données pour les exploiter ?
5. **Informer** - Comment transmettre les informations recueillies au sein de mon entreprise ?

²⁰ Traduit de l'anglais : *provide a process that coherently brings together a combination of skills for analysing and interpreting complex information from a variety of sources and the ability to present complex technical information to non-specialists, and an ability to add insights.*

²¹ Version originale non traduite disponible à l'annexe 3.

6. **Prendre les décisions en connaissance de cause** - En fonction de toutes les étapes précédentes, quelles sont les décisions que l'entreprise doit prendre pour maximiser ses bénéfices ?
7. **Planifier et exécuter les actions** - Quelles sont les actions à faire pour ajouter de la valeur à l'entreprise ?

Malgré le fait que ce modèle ait été développé au début du Big Data et autres mégabases de données, nous justifions notre choix avec deux arguments :

Lors de la création de ce modèle en 2004, les auteurs Neely et Jarrar avaient perçu que les managers et autres preneurs de décisions croulaient sous un nombre énorme de données. Les auteurs disaient :

À une époque où les données deviennent abondantes et accessibles à tous à des coûts plus ou moins réduits et abordables, et alors que le cycle de vie utile des données se raccourcit, la capacité à extraire rapidement et efficacement de la valeur des données, pour éclairer la prise de décision et faciliter l'action rapide, séparera les gagnants des perdants. Le PPVC est conçu pour faire passer l'organisation du stade de la manipulation des données à celui du traitement efficace de l'information et de sa transformation en connaissances à valeur ajoutée et en expérience durable pour un avantage concurrentiel²² (Neely & Jarrar, 2004, p509).

Depuis 1900 et l'apparition du modèle de la performance d'entreprises et des outils d'analyse qui l'accompagnent, de nombreux modèles ont été créés. C'est pour cela que dans les années 1980 le système de mesure des performances (performance measurement system ou PMS) a été mis en place. Il n'y a pas de consensus sur la définition du concept, Franco-Santos

²² Traduit de l'anglais: *In an age, where data are becoming abundant and available to everyone at, more or less, increasingly reduced and affordable costs, and while the useful lifecycle of data is shortening, the ability to quickly and effectively extract value from data, to inform decision-making and facilitate quick action, will separate the winners from the losers. The PPVC is designed to move the organisation from working with data to effectively handling information and turning it into value adding knowledge and sustainable experience for competitive advantage.*

et al., (2007) relèvent 17 définitions différentes. Cela s'explique par le fait que les différents auteurs n'abordent pas le PMS de la même façon.

Ainsi, l'étude *A state of the art and comparison of approaches for performance measurement systems definition and design* de Michel Stella Ravelomanantsoa, Yves Ducq et Bruno Vallespir fait l'inventaire des principaux modèles utilisés dans le cadre du PMS. Ils décrivent le PPVC comme « ...un processus systémique visant à extraire la valeur des données d'une entreprise afin d'améliorer la prise de décision et de concentrer les efforts sur ce qui apportera une réelle valeur ajoutée à l'entreprise²³. »(Ravelomanantsoa et al., 2019). De tous les modèles listés dans l'étude, le PPVC nous semble le plus pertinent à analyser dans le cadre de notre mémoire.

4.2.3 Combinaison des deux modèles

Maintenant que nous avons abordé les deux modèles, nous allons décrire comment nous allons les relier entre eux. Nous partons du fait que comme le but des deux modèles est d'augmenter les profits de l'entreprise, ils peuvent être combinés. Pour accomplir cette combinaison, nous partirons des différentes étapes du PPVC et les relierons aux activités de Porter.

Avant de commencer à étudier les différentes étapes, nous rappellerons que comme la gestion des ressources humaines est présente dans toutes les étapes. Nous prenons cette décision pour éviter de répéter pour chaque étape sa présence, ce qui allège le texte.

Pour ce qui est de l'étape de l'hypothèse, ce sont les activités de recherches et développement ainsi que le marketing qui interviennent. L'équipe en charge de la recherche et du développement va rechercher quel type de produit l'entreprise va devoir développer. Ils vont communiquer avec l'équipe marketing qui, elle, recherche la meilleure façon de vendre le produit qui sera développé. Elles vont donc rechercher le moyen d'accomplir les objectifs qui leur ont été fixés.

²³ Traduit de l'anglais : ...a systemic process to extract value from an enterprise data to enhance the decision-making to focus the efforts on what will add real value to the enterprise.

Pour ce qui est de l'étape de collecte des informations, les activités qui interviennent vont dépendre des hypothèses qui ont été formulées à l'étape précédente. Si l'entreprise veut régler un problème interne, c'est la gestion des ressources humaines et les infrastructures qui vont être les plus sollicitées. Si c'est un problème avec les clients, ce sont le marketing et les services qui le seront. Si le problème vient des fournisseurs, ce sont les achats et logistiques d'approvisionnement qui seront sollicités. Il est donc logique d'affirmer qu'ici c'est à voir au cas par cas.

Pour ce qui est de l'étape de l'analyse des données, c'est l'équipe de recherches et développement qui va travailler sur cette tâche. Il reste néanmoins possible que des analystes soient répartis dans les différentes activités de l'entreprise en fonction des besoins de cette dernière. Il n'est donc pas possible d'exclure entièrement les autres activités dans cette étape du modèle.

Pour ce qui est de l'étape de l'interprétation, la situation est en grande partie identique à l'étape du recueillement des données. Les données doivent être interprétées par les responsables des activités correspondantes. Ces derniers vont devoir étudier la façon d'exploiter de manière optimale les données pour pouvoir les utiliser dans la résolution des hypothèses formulées au départ.

Pour ce qui est de l'étape de l'information, il s'agit ici de communiquer les informations aux instances qui prennent les décisions. Ceux-ci sont représentés au sein de l'infrastructure de l'entreprise dans la chaîne de valeur de Porter. C'est un point important, car si la transmission des données analysées ne s'effectue pas correctement, toutes les étapes précédentes seront caduques et les étapes qui en découlent seront faussées. C'est pour cela que l'infrastructure de l'entreprise doit bien s'assurer du bon déroulement de cette étape.

Pour ce qui est de l'étape de la prise des décisions, c'est une fois de plus l'infrastructure de l'entreprise qui se charge de cette étape. Étant donné que c'est cette dernière qui coordonne toutes les autres, elle est donc chargée de prendre les décisions. Néanmoins, elle peut être épaulée par les autres activités si l'hypothèse formulée à la base requiert une expertise particulière ou si l'hypothèse ne concerne qu'une seule activité.

Pour ce qui est de l'étape de planifier et exécuter les actions, nous sommes dans la même situation que la récolte d'informations. En effet, une fois que l'infrastructure de l'entreprise a communiqué ses décisions au reste des activités, ces dernières doivent mettre en œuvre les actions pour réaliser les décisions prises. Les activités qui participeront à cette étape sont directement liées à l'hypothèse formulée lors de la première étape.

Pour conclure, nous avons effectué un tableau récapitulatif de cette combinaison (Annexe 2). Nous avons montré ici que les deux modèles exprimaient de manière différente la même fonction au sein d'une entreprise. Cette dernière a besoin d'énormément d'échanges de données afin de fonctionner de manière optimale.

4.3 *Partage transversal de données*

Comme son nom l'indique, ce partage de données n'est pas au sein d'une seule chaîne de valeur, mais s'étale sur plusieurs d'entre elles. Dans ce cas de figure, il n'y a pas de rivalité entre les entreprises. Ce mode d'échange de données peut être plus compliqué, car les entreprises ne sont pas naturellement en relation entre elles et doivent innover dans des secteurs qui ne sont pas les leurs (Berthier & Fulchiron, 2018).

L'exemple le plus représentatif de ce mode d'échange est la publicité en ligne. Celle-ci possède des facettes différentes et nous nous attarderons ici sur deux types: les campagnes marketing personnalisées et les campagnes marketing virales.

Campagnes marketing personnalisées

Pour les campagnes marketing personnalisées, nous utiliserons l'exemple d'une marque de vêtements qui désire passer par Facebook pour faire sa campagne de publicité. Pour cela, nous nous baserons sur le livre *Marketing Facebook : Guide stratégique pour la communication et la publicité* de Chiara Cini. Avant d'acheter des emplacements de publicité à Facebook, il faut savoir quel public on désire viser.

Pour ce faire, deux possibilités : soit rechercher soi-même par ses propres moyens (gratuit) ou avec des outils d'analyse (payant) soit de faire un appel à une société spécialisée.

Ici, nous sommes dans un premier cas d'échanges de données transversaux : une entreprise de vêtements avec une entreprise d'analyse. Une fois la population visée identifiée, l'entreprise de vêtements peut passer à l'étape suivante (Cini, 2019).

Facebook et ses deux milliards d'utilisateurs représentent une occasion intéressante pour toutes les entreprises qui veulent toucher un grand public. C'est pourquoi la véritable force de Facebook est l'énorme quantité de données qu'il possède sur eux : nom, prénom, adresse, relation amoureuse, etc. Toutes ces informations permettent de toucher de manière efficace le public souhaité (Cini, 2019). Dans son rapport annuel de 2019, Facebook déclarait : « *Nous générons la quasi-totalité de nos revenus en vendant des placements publicitaires aux spécialistes du marketing.* »²⁴.

Ainsi, en fonction de son budget, l'entreprise de vêtements pourra choisir le nombre de personnes touchées, la fréquence à laquelle ils seront touchés, etc. De plus, Facebook a mis en place le *Business Manager*, outil qui permet à une entreprise de centraliser la gestion de ses campagnes publicitaires ainsi que ses pages (Cini, 2019). Nous sommes une fois de plus dans une situation de partage de données transversal, ici entre une entreprise de vêtements et une entreprise de réseaux sociaux.

Campagne marketing virale

Le marketing viral est un marketing qui se base sur son public cible pour faire la publicité. Nous le plaçons dans le cadre du partage de données transversal dans sa façon de se diffuser. Pour cela, nous reprenons les cinq principaux paramètres développés par Mohamed Yusuf Warsame dans son mémoire *Webmarketing* et nous développons leurs liens avec le partage de données transversal.

Son premier principe, l'acte de provocation, est de susciter de l'intérêt du public. Pour cela, il faut être provocateur dans sa manière de s'exprimer afin de capter l'attention (Yusuf Warsame & Banoff, 2007). Mais c'est lors du partage et de la mise en ligne qu'un partage de données transversal peut s'effectuer. Puisque l'entreprise ne gère pas la propagation de la

²⁴ Traduit de l'anglais : *We generate substantially all of our revenue from selling advertising placements to marketers.*

publicité, celle-ci peut se retrouver sur sites, forums, etc. avec lesquels l'entreprise n'a pas de lien.

Le second principe est d'identifier la cible. Nous ne nous attarderons pas sur ce point puisqu'il a déjà été développé dans la partie précédente sur la campagne de marketing personnalisée.

Le troisième principe est la diffusion de la campagne. Dans ce principe-là, l'objectif est de choisir le lieu de départ de la campagne. L'entreprise peut choisir de collaborer avec des influenceurs, sites spécialisés, etc. Elle doit trouver le foyer qui serait le plus propice à lancer l'opération, quitte à trouver des partenaires inhabituels pour son secteur d'activité (Yusuf Warsame & Banoff, 2007).

Le quatrième principe est de mettre en place des supports pour maintenir la campagne. Ces supports sont, par exemple, la création d'un site internet dédié, des partenariats avec des influenceurs ou simplement en combinant avec une campagne marketing personnalisée. L'idée est de maintenir la campagne au centre de l'attention (Yusuf Warsame & Banoff, 2007).

Le cinquième principe est d'attirer l'attention des médias classiques. Ces derniers ont une énorme audience et apparaître à leurs antennes contribue à accentuer l'effet viral. Il est possible à ce stade que des entreprises contactent l'entreprise de départ pour effectuer des partenariats (Yusuf Warsame & Banoff, 2007).

En conclusion, ce type de campagne marketing permet un grand échange de données transversal, mais elle reste soumise à sa capacité à motiver le public et à arriver à surfer sur la vague qu'elle crée.

Chapitre 5 - Les obstacles à l'échange de données

5.1 *La confiance*

5.1.1 *Introduction du dilemme du prisonnier*

Afin de comprendre l'obstacle lié à la confiance, nous allons simuler une situation où il n'y a pas de confiance et où les entreprises agissent de manière rationnelle et égoïste. Nous allons utiliser une forme adaptée du dilemme du prisonnier (Mimbang et al., 2015).

Pour rappel, le dilemme du prisonnier est une application de la Théorie des Jeux énoncée par Albert W. Tucker en 1950. Dans ce dilemme, deux prisonniers sont retenus dans deux cellules séparées et sont interrogés sans pouvoir communiquer entre eux. Chaque prisonnier a le choix entre dénoncer son partenaire ou ne rien dire. Alors que la situation idéale pour les prisonniers semble être que les deux ne disent rien, le dilemme du prisonnier nous explique pourquoi les prisonniers pourraient décider de se trahir. (Mimbang et al., 2015).

Nous allons donc reprendre ce dilemme du prisonnier et l'adapter à l'échange de données entre entreprise. Prenons le cas de deux entreprises : Alpha et Beta. Chacune des deux entreprises possède une base de données. Alpha possède la base de données α et Beta possède la base de données β . Chacune des entreprises se trouve dans une situation où elle a le choix entre deux options : partager ses données ou ne pas les partager. De plus, elle doit faire ce choix indépendamment de celui formulé par la seconde. Dans cet exercice, nous allons établir que la valeur d'une donnée est égale au revenu qu'elle permet de générer. Il s'agit de la méthode d'évaluation utilisant l'approche par les revenus définie dans le chapitre 2.

5.1.2 *Dilemme 1 : Échange de données dans un contexte de non-concurrence*

Dans la situation initiale, les deux entreprises ne partagent pas leurs données. Arbitrairement, nous allons déterminer que chaque entreprise génère 100€ de revenus grâce à la détention de leurs données respectives. Par exemple, nous pouvons imaginer un cas où ces données sont des informations récoltées concernant les visiteurs de leur site web. Les entreprises ont ensuite remarqué que ces données leur permettent de mieux cibler leurs clients.

Ainsi, elles permettent de générer 100€ de revenus supplémentaires par rapport à la situation où elles n'avaient pas ces données.

Nous allons maintenant considérer la situation où les entreprises décident de partager ces données. Pour reprendre l'exemple précédent, accéder aux données de l'autre entreprise représente une opportunité d'améliorer le ciblage et la compréhension de sa clientèle. En connaissant mieux les besoins de vos clients, vous pouvez vous permettre d'augmenter certains prix ou d'offrir de nouveaux produits répondant mieux aux besoins des clients. Nous allons considérer ici que cet avantage représente un revenu supplémentaire de 50€.

Dernière chose, nous établissons que si une entreprise partage ses données mais que son homologue ne le fait pas, elle obtient une pénalité de 10€. Cette pénalité se justifie par le fait que cette entreprise a mis en place une infrastructure permettant de partager des données avec les autres, alors qu'elle n'en perçoit aucun bénéfice.

Vous noterez que par souci de clarté, nous avons fait l'hypothèse qu'augmenter la quantité de données implique nécessairement une augmentation du revenu. Dans les faits, cette hypothèse est erronée. La capacité des données à générer du revenu dépend également de :

- **La qualité des données** – L'obsolescence ou encore le manque de contexte peut détériorer la qualité des données, les rendant inutilisables pour générer du revenu.
- **La capacité d'analyse** – Comme nous l'avons vu avec le modèle DIKW, la donnée est la matière première de la connaissance. Néanmoins, pour pouvoir générer de la connaissance, il faut être capable d'analyser correctement les données.

Nous avons donc fait un tableau résumant la situation. Notez que les nombres ci-dessous représentent les revenus des entreprises. Le format est le suivant : (Revenu d'Alpha ; Revenu de Beta)

	Beta partage	Beta ne partage pas
Alpha partage	(150€ ; 150€)	(90€ ; 150€)
Alpha ne partage pas	(150€ ; 90€)	(100€ ; 100€)

De ce résultat, il y a deux informations majeures à en retirer :

- Nous pouvons remarquer que la situation générant au total le plus de revenus est celle où les deux entreprises partagent leurs données. C'est la situation où le maximum de revenus est généré à partir des données.
- Il existe deux équilibres de Nash dans les situations où les entreprises partagent ou ne partagent pas leurs données. C'est-à-dire que ce sont les deux situations stables où aucune des deux entreprises ne souhaite dévier de sa stratégie (Nash, 1950).

Prenant en compte ces deux informations, on peut conclure qu'une entreprise n'a aucun intérêt à partager ses données sans la certitude que son partenaire ne le fera à son tour. Vu que la situation initiale est celle où les entreprises ne partagent pas leurs données, elles n'ont pas de motivation à dévier de cette stratégie. Cela pourrait expliquer pourquoi les entreprises ne partagent pas spontanément leurs données.

Toutefois, si les entreprises se connaissent et ont confiance l'une envers l'autre, alors elles peuvent se mettre d'accord pour partager ensemble les données par le biais de partenariat. On atteindrait alors une nouvelle situation stable : celle où les deux entreprises partagent leurs données. Ce modèle pourrait expliquer pourquoi une grande partie des échanges de données

entre entreprises se fait grâce à des accords bilatéraux entre deux entreprises partenaires (European Commission & Everis Benelux, 2018).

5.1.3 Dilemme 2 : Echange de données dans un contexte de concurrence directe

Contrairement à la situation précédente, les deux entreprises Alpha et Beta sont en concurrence directe. C'est-à-dire que dès qu'une entreprise acquiert un avantage, cela se fait au dépend de ses concurrents directs qui ne possèdent pas cet avantage.

Comme la situation précédente, on considère que chaque entreprise génère 100€ de revenus grâce au fait qu'elle détient ces données. De plus, si l'entreprise possède également les données de l'autre entreprise, elle génère 50€ de revenus supplémentaires.

Il y a cependant un élément à ne pas négliger : la compétition. Si une entreprise possède plus de données que l'autre, elle possède un avantage compétitif lui permettant d'être plus efficace que son homologue. Nous allons représenter cet avantage compétitif en soustrayant 10€ à l'entreprise possédant moins de données, puis en les attribuant à l'entreprise avec le plus de données.

Nous avons donc fait un tableau résumant la situation. Notez que les nombres ci-dessous représentent les revenus des entreprises. Le format est le suivant : (Revenu d'Alpha ; Revenu de Beta)

	Beta partage	Beta ne partage pas
Alpha partage	(150€ ; 150€)	(90€ ; 160€)
Alpha ne partage pas	(160€ ; 90€)	(100€ ; 100€)

Nous pouvons remarquer que la situation générant au total le plus de revenus est celle où les deux entreprises partagent leurs données. C'est la situation où le maximum de revenus est généré à partir des données. Cependant, si on suit le raisonnement de John Nash, il serait

rationnel de ne pas partager ses données avec les autres entreprises. En effet, on remarque que pour les deux entreprises, la stratégie dominante est de ne pas partager. L'équilibre de Nash de ce jeu est donc la situation où aucune entreprise ne partage ses données (Nash, 1950).

En d'autres termes, cela revient à dire qu'une entreprise n'a aucune motivation rationnelle à partager ses données. D'une part, dans la situation où son concurrent ne partage pas ses données, l'entreprise n'a rien à gagner à partager les siennes. D'autre part, si le concurrent partage ses données, l'entreprise a tout intérêt à ne pas partager les siennes pour ne pas aider son concurrent.

5.1.4 Conclusion liée aux dilemmes

De fait, nous remarquons que le partage de données est une action altruiste dont le revenu potentiel réside dans l'espoir que ses concurrents fassent de même. Ainsi, il est nécessaire qu'un climat de confiance règne pour que le partage de données soit possible. Cette nécessité d'avoir confiance en l'autre est l'une des principales barrières empêchant les entreprises de vouloir partager leurs données.

5.2 Difficultés techniques

5.2.1 Manque de standardisation et interopérabilité

Le manque de standards concernant les données, la manière dont elles sont stockées, traitées et analysées est l'une des principales barrières à l'échange de données. Voici le problème exposé par McGinnes :

« La prolifération massive actuelle de bases de données, de pages Web et d'autres ressources d'information pose un problème d'intégration des données. [...] Il manque des mécanismes permettant une intégration facile des données de manière générale ; il est souvent difficile de combiner les ressources de données. L'une des principales raisons en est que les

*différents ensembles de données ont généralement des structures conceptuelles incompatibles.*²⁵ » (McGinnes, 2011).

Au sein de l'Union Européenne, il n'existe pas de standardisation faisant consensus. Ainsi, les méthodes de collectes, la structure ou encore les unités de mesure sont différentes d'un pays à l'autre, d'une entreprise à l'autre. Certaines bases de données peuvent donc être incompatibles « by design ». McGinnes explique très bien pourquoi cette incompatibilité pose un problème à l'échange de données entre entreprises :

*« Les ordinateurs ne peuvent pas encore comprendre les données au sens où l'entend l'homme, mais ils peuvent être programmés pour traiter judicieusement des éléments spécifiques de données, à condition que ces données soient d'un type connu ; c'est ce que nous entendons lorsque nous disons qu'un programme "comprend" un concept particulier. En pratique, cependant, la plupart des Systèmes d'Informations ne partagent ni vocabulaire ni concepts. Il serait surprenant qu'ils le fassent, étant donné la manière dont ils sont développés et la rareté avec laquelle les structures conceptuelles standards sont appliquées. Pour cette raison, il est rarement simple de relier des Systèmes d'Informations existants qui ont des schémas conceptuels hétérogènes.*²⁶ » (McGinnes, 2011).

On comprend ainsi que, si deux entreprises ont deux systèmes avec des structures différentes, le partage de données entre ces deux entreprises peut s'avérer complexe. Un travail conséquent, demandant un investissement de capitaux financier et humain important, doit être mené afin de créer un « pont » entre les deux systèmes. Si toutefois un tel pont est possible.

²⁵ Traduit de l'anglais : « The present massive proliferation of databases, web pages and other information resources presents a data integration problem. There is a need to use data in a joined-up way, but mechanisms are lacking that allow easy data integration in general case; it is often hard to combine data resources. A prime reason for this is that different datasets typically have incompatible conceptual structures. » (McGinnes, 2011)

²⁶ Traduit de l'anglais : « Computers cannot yet understand data in the sense that a human does, but they can be programmed to deal sensibly with specific items of data provided that the data is of a known type; this is what we mean when we say that a program "understands" a particular concept. In practice, however, most IS share neither vocabulary nor concepts. It would be surprising if they did, given the ways they are developed and the rarity with which standard conceptual structures are applied. For this reason, linking real-world IS that have heterogeneous conceptual schemas is rarely a simple matter. »

Nous reviendrons sur la nécessité de créer un « système passerelle » entre deux systèmes différents dans le chapitre suivant.

Afin d'illustrer cet obstacle, nous allons évoquer le cas de Green Energy, aussi appelée « *geo* » (European Commission. Directorate General for Communications Networks, Content and Technology. & everis Benelux., 2018). L'entreprise *geo* est une entreprise anglaise dont le produit historique est un appareil connecté de suivi de la consommation énergétique. Vous installez ce boîtier chez vous et vous pouvez suivre avec précision la consommation en énergie de votre foyer. Néanmoins, ils ont très vite constaté que les données récoltées avec leurs appareils avaient de la valeur. Ils ont ainsi commencé à commercialiser celles-ci. Parmi les clients intéressés, les fournisseurs d'électricité qui sont désormais capable de baser leurs prévisions sur des données en temps réel, et non plus des données historiques.

En 2015, *geo* a lancé un projet pour tenter de réduire la consommation d'énergie des foyers : mieux comprendre les liens entre les données récoltées par son thermostat et la météo extérieure. Afin d'avoir des résultats satisfaisants, ils ont mené ces analyses dans plusieurs pays différents. Or, il a été difficile pour l'entreprise de trouver des données standardisées. Une grande partie du temps lié à ce projet a donc été dédié à uniformiser les différentes bases de données ; ce qui a fatalement rendu le projet moins efficace (European Commission. Directorate General for Communications Networks, Content and Technology. & everis Benelux., 2018).

Cet exemple démontre bien l'enjeu, mais aussi les difficultés liées à l'absence de standards en Europe. Nous y reviendrons dans le chapitre suivant portant sur les solutions à ces obstacles.

5.2.2 Coût lié à l'infrastructure

Afin de pouvoir procéder à de l'échange de données avec d'autres entreprises, il est nécessaire d'accéder à une infrastructure capable de stocker, utiliser, transférer ou encore analyser ces données. Or, au plus la quantité de données est importante, au plus l'infrastructure nécessaire peut représenter un coût qui peut décourager plusieurs entreprises.

Lorsque l'on pense aux coûts liés à l'infrastructure, la première chose à laquelle on pense est le coût du stockage physique des données. Autrement dit : les serveurs. Pourtant, lorsque l'on regarde le coût lié au stockage des données, on remarque qu'il a chuté ces dernières années. On peut voir sur le graphique ci-dessous que le coût lié au stockage des données sur un serveur NAS a été divisé par 8 (Statista, 2021).

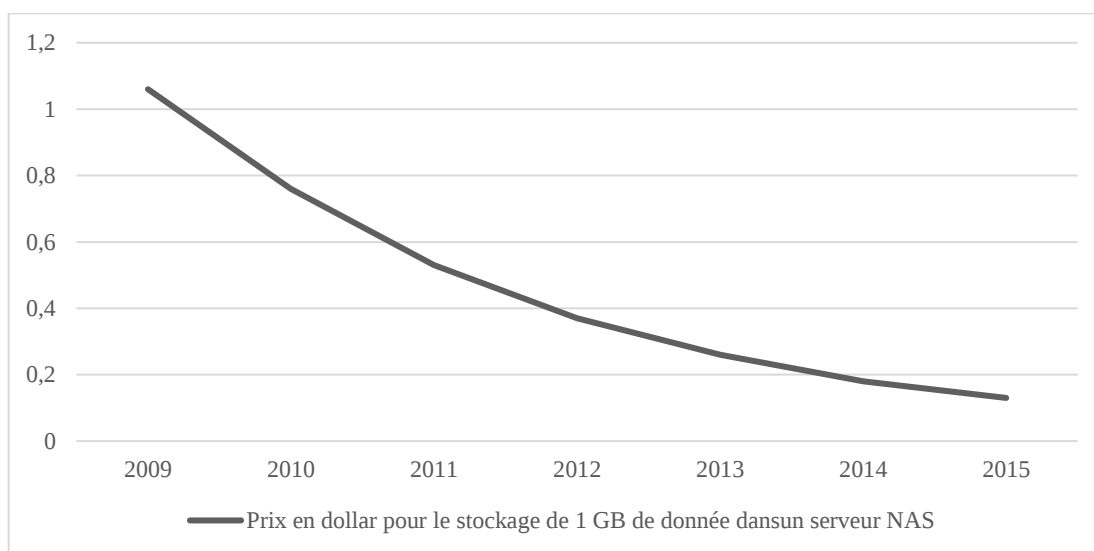


Figure 12 - Évolution du coût de stockage des données (Statista, 2021)

Selon l'étude menée par l'Union Européenne, 41% des entreprises ayant répondu au sondage ont affirmé que le coût de l'infrastructure était trop conséquent. Il y a deux raisons à cela (European Commission & Everis Benelux, 2018).

D'une part, même si le coût de stockage des données est plus faible qu'auparavant, la quantité de données que nous stockons a drastiquement augmenté. Entre 2010 et 2020, la quantité de données stockées dans le monde a été multipliée par 40 (Guo et al., 2014). On peut lier ce phénomène à celui de l'effet rebond. L'effet rebond est principalement utilisé pour analyser l'augmentation de la consommation d'énergie à la suite d'améliorations technologiques. Par exemple, la démocratisation des ampoules LED a permis de drastiquement diminuer la consommation d'une ampoule. Cependant, on constate que la consommation totale d'électricité liée à l'éclairage n'a pas diminué. La raison est simple : la démocratisation de l'ampoule LED a rendu l'éclairage peu coûteux et il est devenu peu intéressant de faire des économies. On a donc augmenté notre consommation d'éclairage, venant contrebalancer les

économies réalisées à l'échelle d'une ampoule. Concernant le stockage de données, la diminution du coût de stockage d'un Gigabyte de données n'a pas eu pour effet de diminuer la facture en hébergement des entreprises. Ces dernières consomment, comme nous l'avons vu plus haut, nettement plus de données qu'avant. Ainsi, les serveurs restent un coût important pour les entreprises (Sorrell & Dimitropoulos, 2008).

D'autre part, le coût de l'infrastructure ne se limite pas aux serveurs. Parmi les autres coûts liés à l'infrastructure, on retrouve : (Dutta & Hasan, 2013)

- Les coûts liés aux locaux (loyers, constructions)
- Les coûts énergétiques
- Les coûts en capital humain (entretien, sécurité ...)
- Les coûts d'élimination (destruction de données)
- Les coûts environnementaux

Ainsi, pour les raisons citées précédemment, l'infrastructure et ses besoins importants en investissement sont un des principaux obstacles pour certaines entreprises.

5.2.3 *Qualité insuffisante des données*

L'expression « qualité des données » renvoie au concept d'exactitude de ces données. Cette exactitude peut, par exemple, être mise à mal par des problèmes liés à l'accessibilité, la temporalité, la pertinence ou encore la crédibilité des données (Madnick et al., 2003).

Par exemple, la donnée peut être de mauvaise qualité à la suite d'erreurs de calcul lors de l'observation. Ou encore, les données sont devenues obsolètes. Prenons le cas d'une base de données de la clientèle d'une entreprise qui n'a plus été mise à jour depuis 20 ans. Cette base de données a perdu une partie importante de sa valeur. Certaines entreprises de la liste ont pu faire faillite. Les personnes de contact ont changé de travail. En résumé, il est devenu difficile d'extraire des informations utiles de cette base de données. Nous pouvons donc dire que cette dernière est de faible qualité.

Cette question de la qualité des données pose deux problèmes à l'échange de données entre entreprises :

- **Du côté des fournisseurs des données**, fournir des données de faible qualité à ses partenaires peut poser un problème. C'est par exemple le cas de l'entreprise Enedis. Enedis a signalé que la granularité des données qu'elle détient ne répond parfois pas à la demande des utilisateurs de données. Comme ces derniers exigent une granularité plus fine, Enedis ne peut pas garantir des normes de qualité suffisantes et a donc décidé de ne pas partager ces données (European Commission & Everis Benelux, 2018).
- **Du côté des utilisateurs des données**, se procurer des données représente un risque. Avant l'achat, il est difficile de mesurer la qualité des données. Il peut donc y avoir une certaine méfiance de la part des acheteurs de données.

5.2.4 Complexité de l'architecture des plateformes

Selon Baldwin et Woodard, « *La caractéristique fondamentale d'une architecture de plateforme est que certains composants restent fixes pendant toute la durée de vie de la plateforme, tandis que d'autres peuvent varier en section transversale ou changer*²⁷. » (Baldwin & Woodard, 2008,p7).

De par cette définition, chaque entreprise peut avoir une architecture de plateforme différente, même si elles étaient similaires à la base. Étant donné que c'est l'architecture qui s'adapte aux besoins de l'entreprise et non l'inverse, il est difficile d'avoir des compatibilités entre les entreprises lorsque ces dernières veulent échanger des données à cause de ces évolutions.

Nous voyons donc que la force des architectures de plateformes est leur faiblesse dans le cas de l'échange de données. L'adaptabilité des architectures rend les échanges entre ces dernières plus difficiles. Si partager des données entraîne un coût en temps et en moyen

²⁷ Traduit de l'anglais : *The fundamental feature of a platform architecture, in our view, is that certain components remain fixed over the life of the platform, while others are allowed to vary in cross-section or change over time.*

important pour permettre la compatibilité, les entreprises sont moins enclines à partager leurs données.

5.2.5 Indisponibilité des compétences requises

L'Organisation de Coopération et de Développement Economiques (OCDE) identifie plusieurs compétences nécessaires pour assurer un partage de données efficace.

Le premier point développé concerne l'importance d'avoir un personnel formé dans la gestion de données. Ce dernier doit être compétent de façon à maximiser les bénéfices tirés de l'échange de données. L'OCDE décrit ce frein de cette manière:

*Le manque d'aptitudes et de compétences liées aux données et le faible accès aux capacités de calcul et de stockage peuvent devenir des goulots d'étranglement empêchant la réutilisation efficace des données, même lorsque celles-ci sont mises à disposition grâce à un accès et un partage améliorés. La promotion des infrastructures et des compétences liées aux données ne contribuerait pas seulement à garantir que tous puissent bénéficier d'un accès et d'un partage améliorés. Elle permettrait également de maximiser la réutilisation et donc la valeur des données.*²⁸ (OECD, 2019).

Le second point développé par l'OCDE est que du personnel compétent est difficile à trouver. Comme nous venons de le voir, le personnel formé est capital pour le traitement des données. Seulement, comme le relève l'OCDE, les spécialistes en traitement de données ne représentent que 1% des emplois dans les pays de l'OCDE (OECD, 2019). Le manque de personnel formé et disponible dans le traitement de données est donc un obstacle au partage de données.

²⁸ Traduit de l'anglais : *Lack of data-related skills and competences and poor access to computation and storage capacities can become bottlenecks preventing the effective re-use of data, even where data are made available through enhanced access and sharing. Fostering data-related infrastructures and skills would not only help assure that all can benefit from enhanced access and sharing. It would also help maximise the re-use and thus the value.*

Ainsi, l'OCDE met l'accent sur l'importance du personnel qui peut être un frein important au partage de données.

5.3 *Incertitude stratégique*

Cet obstacle est décrit par Berthier et Fulchiron dans leur étude « *La deuxième vie des données* » :

Le frein le plus important au Data Sharing réside sans doute dans les incertitudes sur la valeur créée dans le futur. N'étant pas en mesure d'estimer cette valeur par avance, les entreprises détentrices de données sont réticentes à les partager, puisqu'elles n'ont pas l'assurance de récupérer ensuite une partie de cette valeur. (Berthier & Fulchiron, s. d., p7)

Nous pouvons l'interpréter en le combinant avec la situation du dilemme du prisonnier que nous avons développé dans le début de ce chapitre. En effet, comme les entreprises ne savent pas quels seront leurs bénéfices grâce aux partages de données, elles ne veulent pas prendre le risque de donner à leurs concurrents un avantage grâce à leurs données.

5.4 *Difficultés juridiques*

Le principal obstacle juridique relevé par la Commission européenne est le manque de clarté concernant les droits de la propriété des données. Cela inclut l'utilisation des données après leur partage à d'autres entreprises ainsi que les difficultés à pouvoir tracer les données et l'absence totale de contrôle sur ces dernières (European Commission & Everis Benelux, 2018). Nous pouvons donc voir que l'absence de cadre juridique dans le suivi des données est un obstacle majeur dans le partage de données.

Les deux autres obstacles relevés par la Commission européenne sont les exigences légales en matière de protection des données et les restrictions géographiques.

Concernant les exigences légales en matière de protection des données, le principal point mis en évidence est que les entreprises ne sont pas certaines du type de données qu'elles peuvent partager. Par conséquent, elles ne savent pas quel niveau de sécurité et de protection elles

doivent appliquer aux différents types de données et étant donné que ces derniers ont un coût, elles préfèrent ne pas prendre de risques (European Commission & Everis Benelux, 2018).

Concernant les restrictions géographiques, les entreprises doivent s'adapter aux différentes législations. Sachant que l'Union Européenne n'a pas les mêmes lois pour le traitement et la protection des données que les États-Unis par exemple, une entreprise opérant dans ces deux régions doit adapter ses installations et protocoles de façon à respecter les règles des différentes zones où elle se situe (European Commission & Everis Benelux, 2018).

Nous voyons donc ici que le manque de législations communes aux différents pays ainsi que le manque de cadre pour l'utilisation des données après partage constituent des obstacles aux partages entre entreprises.

5.4.1 Perte de contrôle sur les données

D'après l'enquête de la Commission européenne, l'une des principales raisons décourageant les entreprises à partager leurs données serait l'impossibilité de garantir la confidentialité des données (European Commission & Everis Benelux, 2018). En effet, les données pouvant être répliquées sans engendrer de coût, il est difficile de traquer et vérifier l'usage des données que vous partagez. Il existe ainsi des incertitudes légales liées aux droits de propriété des données.

5.4.2 RGPD

Le principal texte de loi impactant l'échange de données entre entreprises est le « Règlement général sur la protection des données », mieux connu sous l'appellation RGPD ou GDPR (Règlement (UE) 2016/679, 2016). Ce règlement, qui est en application dans l'Espace Économique Européen, a permis une uniformisation des lois nationales en la matière. L'article 2 du RGPD résume l'étendue et l'objectif de ce texte de loi :

*« Les principes et les règles régissant la protection des personnes physiques à l'égard du traitement des données à caractère personnel les concernant devraient, quelle que soit la nationalité ou la résidence de ces personnes physiques, respecter leurs libertés et droits fondamentaux, en particulier **leur droit à la protection des données à caractère personnel**. Le présent règlement vise à contribuer à la réalisation d'un espace de liberté, de sécurité et de justice et d'une union économique, au progrès économique et social, à la consolidation et à la convergence des économies au sein du marché intérieur, ainsi qu'au bien-être des personnes physiques. » (Règlement (UE) 2016/679, 2016).*

Par donnée à caractère personnel, le Règlement entend toute donnée ou information « se rapportant à une personne vivante identifiée ou identifiable » (Commission européenne - European Commission, 2021b). Cela comprend ainsi les données comme des noms, des géolocalisations, des identifiants, mais aussi des traits physiques, génétiques, économiques, culturels ou sociaux d'un individu. Le RGPD permet ainsi aux citoyens européens de contrôler leurs données à caractère personnel. Elles ne peuvent pas être récoltées par une entreprise sans son consentement et le citoyen a également le droit de demander la suppression de ses données personnelles.

Il est à noter que le RGPD n'impacte que l'échange de données personnelles. Son objectif étant de protéger les citoyens d'utilisation abusive de leurs données, il force donc les entreprises à prendre des mesures supplémentaires pour protéger les données. C'est pourquoi les entreprises anonymisent les données qu'elles possèdent dans leur base de données, car les données anonymisées ne sont pas concernées par le RGPD.

Maintenant que nous avons rappelé le RGPD, nous nous attarderons ici sur ses conséquences qui représentent un obstacle à l'échange de données. Pour ce faire, nous nous baserons d'abord sur l'étude *Measuring the Impact of the GDPR on Data Sharing in Ad Networks* de Urban et al. de 2020.

Une de leurs observations révèle que le nombre d'échanges d'identités virtuelles avec les plateformes de publicité en ligne a chuté d'environ 40% d'après les résultats de l'étude.

Bien qu'ils n'excluent pas d'autres facteurs d'influences, les auteurs supposent que cette chute est liée aux règles plus strictes instaurées par le RGPD (Urban et al., 2020). Cette observation semble logique dans le sens où le RGPD permet aux utilisateurs de mieux contrôler les données qu'ils fournissent aux sites internet.

Une autre de leurs observations est la baisse des échanges au moment de la mise en application du RGPD en mai 2018. Cette tendance a fini par s'inverser au bout de quelques mois. Selon les auteurs, cette chute serait due à la suspension des activités le temps que les entreprises mettent en place les installations pour respecter le RGPD (Urban et al., 2020). Nous voyons donc ici un deuxième obstacle lié au RGPD : la nécessité de respecter des lois plus contraignantes concernant la protection des données.

Ensuite, l'étude *The Impact of GDPR on Global Technology Development* de Li et al. contient également des éléments pour lesquels le RGPD représente un obstacle.

Grâce à l'étude de Urban et al., nous savons que les entreprises ont stoppé leur activité le temps de s'adapter aux nouvelles législations. Li et al. chiffrent ces changements à entre 1 million et 10 millions de dollars pour la majorité des entreprises américaines et qui peut dépasser les 10 millions pour une petite minorité d'entre elles (Li et al., 2019). Nous voyons donc ici que pour l'échange de données au sein de l'Union Européenne, le RGPD représente un obstacle financier.

De plus, Li et al. observent que le RGPD impose des normes contraignantes en matière de cybersécurité. Les entreprises doivent impérativement sécuriser les données personnelles qu'ils ont sur leurs utilisateurs. De plus, elles doivent les prévenir au cas où il y aurait une fuite de leurs données (Li et al., 2019). Ainsi, nous supposons que les entreprises opérant sur le territoire européen sont moins enclines à échanger leurs données avec des entreprises ne garantissant pas des normes de sécurité et de protection adaptées.

Le dernier point observé par Li et al. concerne les nouvelles technologies liées aux données. Par exemple, l'intelligence artificielle, les chaînes de blocs (bloc chains) et l'informatique dématérialisée (cloud computing) reposent sur des quantités de données et sur des algorithmes d'analyse. Les législations du RGPD vont augmenter les coûts de

développement de ces technologies, et donc vont logiquement augmenter la valeur des données ainsi que de leur analyse (Li et al., 2019).

Pour conclure, le principal obstacle que représente le RGPD concernant l'échange de données est l'impact de ces législations sur l'organisation des entreprises.

Chapitre 6 - Pistes de solutions pour améliorer la situation

6.1 *Standardisation*

Il est de plus en plus nécessaire de créer standards faisant consensus au sein de la communauté européenne. Que l'initiative vienne du public ou du privé, elle permettrait de faciliter l'échange de données. L'avènement de standards permettrait ainsi de lever de nombreuses barrières techniques et devrait stimuler le marché de la donnée (European Commission & Everis Benelux, 2018).

Dans la même lignée, il est nécessaire d'avoir toutes les données exprimées dans un langage commun. Un langage commun permet effectivement de faciliter le traitement des données ainsi que leur analyse. Un langage commun permet également de définir le niveau de risque et de confidentialité que les entreprises veulent accorder à leurs données (Levenstein, 2019).

6.2 *Création de plateformes d'échange de données*

Les difficultés liées aux différentes plateformes sont un énorme frein au partage de données. C'est pour cela que la création d'une plateforme commune pour uniformiser l'ensemble des systèmes. De plus, comme il a été souligné lorsqu'un rapport de la Commission européenne, il est important que la plateforme commune soit simple dans son utilisation ainsi qu'une interface utilisateur conviviale (European Commission & Everis Benelux, 2018). Par souci d'équité, cette plateforme devra être hébergée par une institution indépendante, telle que l'Union Européenne par exemple.

6.3 *Création de la marque Europe*

Pour faire face à la dominance américaine sur le marché des données, il faut créer un géant européen. Cette solution présente plusieurs avantages : les données européennes restent stockées sur le sol européen et restent donc entièrement soumises aux lois européennes en termes de protection des données. Cette solution maintient la souveraineté européenne dans le monde numérique (Berthier & Fulchiron, 2018).

6.4 *Protection des données*

Pour beaucoup d'entreprises, la protection de leurs données est un des freins principaux à l'échange. C'est pour cela qu'il faut leur apporter des garanties pour elles ainsi que pour les personnes sur qui elles détiennent des informations. C'est uniquement de cette façon que la confiance peut s'installer entre les entreprises et favoriser l'échange de données.

Les façons d'améliorer la confiance sont les suivantes. Il faut mettre en place des mécanismes de partage ayant une très haute sécurité informatique pour la protection des données. Chaque entreprise a un contrôle total de ses données et peut communiquer et collaborer avec les autres utilisateurs. Il serait également possible de mettre en place un système de licence, qui impliquerait entre autres que les données peuvent être légalement partagées et qu'il y a des conditions et des restrictions à l'utilisation des données (European Commission & Everis Benelux, 2018).

6.5 *Formation du personnel*

Comme l'a indiqué l'OCDE, le manque du personnel formé dans le traitement des données est un obstacle important aux échanges de données. Pour pallier cette pénurie, l'organisation suggère de former ses équipes de plusieurs manières. La première est de familiariser ses équipes aux logiciels de traitement de données. La seconde est d'apprendre aux équipes à stocker les données, c'est-à-dire les exploiter, les sauvegarder et les réutiliser dans de nouvelles situations. La troisième est de sensibiliser les équipes à l'importance des données ainsi qu'aux compétences qui y sont liées. (OECD, 2019).

Chapitre 7 - Les facteurs clés de succès de l'échange de données

Nous nous sommes déjà attardés sur les barrières qui découragent les entreprises de se lancer dans l'échange de données. Il est maintenant temps de regarder de plus près les entreprises qui ont décidé de partager leurs données.

Ce chapitre tentera donc comprendre quels sont les facteurs clés qui mènent les échanges de données entre entreprises vers le succès. Pour y parvenir, nous avons étudié différents cas. Nous sommes donc parvenus à identifier six éléments à maîtriser pour réussir ses échanges de données avec des entreprises tierces.

7.1 *Entretenir un lien de confiance entre fournisseurs de données et utilisateurs de données*

Plusieurs études montrent l'importance de la confiance dans les relations entre entreprises (Alves et al., 2017). Déjà dans le chapitre 6 sur les obstacles à l'échange de données, nous avons fait mention de l'impact de la confiance sur la décision d'une entreprise de se lancer dans l'échange de données ou non avec une autre entreprise.

La Commission européenne a mis en avant que la confiance se crée par le fait qu'une haute sécurité des données est assurée par les moyens de partages de données. Entre autres, il faut que les entreprises gardent un contrôle total sur les données qu'elles partagent. De plus, les communications entre les différents participants au partage de données jouent un rôle prédominant dans la construction d'une relation de confiance entre eux. (European Commission & Everis Benelux, 2018).

La Commission européenne propose de nombreuses solutions pour améliorer la confiance : collaboration étroite avec les utilisateurs de données, accords de licence, expliquer aux utilisateurs de données que les données détenues par une entreprise peuvent être légalement partagées, permettre des contacts directs, etc. (European Commission & Everis Benelux, 2018).

En conclusion, pour entretenir un lien de confiance dans le partage de données, le plus important est de maintenir une totale transparence dans l'utilisation de ces dernières.

7.2 Comprendre les demandes liées de données

Face à la grande quantité de données disponibles actuellement sur le marché, il faut savoir quelles données doivent être acquises ou partager pour apporter de la valeur ajoutée aux entreprises. Toutes les données n'ont pas la même valeur ni la même utilité et il est vital de savoir différencier les différents types qui sont à disposition.

La Commission européenne a mis en avant l'importance de la communication entre les utilisateurs potentiels des données et ceux qui les possèdent. Cela a pour but de permettre l'identification des besoins des entreprises ainsi que des défis auxquels elles font face. Combiné à cette communication, la Commission Européenne rappelle que la réaction aux demandes de données se doit d'être flexible, de façon à pouvoir suivre plus efficacement les besoins des partenaires (European Commission & Everis Benelux, 2018).

7.3 Bâtir des partenariats

Comme nous l'avons déjà évoqué auparavant, la confiance est un critère important dans le partage de données. C'est pour cela que les partenariats entre entreprises sont des facteurs favorisant le succès des échanges de données.

En effet, avoir un partenariat, exclusif ou non, avec une autre entreprise permet de définir un cadre précis pour l'échange de données. Ainsi, les entreprises sont dans une relation de confiance garantie par les accords qu'elles passent entre elles. Dans ces accords, on y retrouve les critères en termes de protection et utilisation, la standardisation des données, etc., et ils permettent de franchir les obstacles au partage de données.

Concrètement, la Commission européenne a identifié cinq avantages dans les partenariats :

- *Mettre en place les bons mécanismes techniques pour rendre leurs données disponibles de manière efficace et sécurisée, ce qui a permis de réaliser d'importantes économies en termes d'investissements technologiques.*
- *Permettre une meilleure compréhension des besoins des utilisateurs de données.*

- *Développer conjointement des produits ou des services à partir des données partagées.*
- *Développer les compétences nécessaires au sein de l'entreprise.*
- *Collecter des fonds pour la mise en place de leur activité de partage de données*²⁹*(European Commission & Everis Benelux, 2018,p84).*

Nous voyons donc que les entreprises ont tout intérêt à établir des partenariats si elles désirent réaliser des partages de données. Comme exemple de l'avantage du partenariat, nous pouvons citer Volkswagen qui a fait un partenariat avec HERE pour leur service de cartes en temps réel. Volkswagen fournit les données par le biais de ses camions et HERE optimise les trajets et la logistique de Volkswagen avec celles-ci (Castle, 2021).

7.4 Mettre en place des mécanismes simples et faciles d'utilisation

Puisque l'un des obstacles au partage de données est le manque de personnel formé au traitement de celles-ci, une façon de faciliter le partage est d'avoir des programmes faciles à utiliser. En effet, utiliser des programmes de traitements de données est plus pratique que recruter du personnel compétent, de par la pénurie de celui-ci.

Combiné à ces programmes, il faut également que les données soient faciles accessibles et retrouvables par tous les employés qui en auraient besoin. Pour cela, il faut que les programmes de traitements des données aient une interface claire avec un système de recherches ou de visuel pour une utilisation efficace.

De plus, si les entreprises qui procèdent au partage de données possèdent les mêmes programmes de traitement, elles n'auront pas de problèmes de compatibilité des données puisque ces dernières seront stockées de la même manière.

²⁹ Traduit de l'anglais :

- *Set up the right technical mechanisms for making their data available in an efficient and secure way, which resulted in making important savings in terms of technology investments*
- *Enable a better understanding of the data users' needs*
- *Jointly develop products or services using the data shared*
- *Build the necessary skills inside the company*
- *Raise funds for setting up their data sharing activity*

C'est un élément qu'a compris l'entreprise Dawex qui a développé une plateforme et des outils pour faciliter l'échange de données entre entreprises. Elle a créé un marché de données sécurisé sur le sol européen. Bien que ce ne soit pas le seul facteur de leur succès, leur plateforme facile à utiliser a été déterminante (European Commission. Directorate General for Communications Networks, Content and Technology. & everis Benelux., 2018).

7.5 Mettre en place un cadre politique et légal propice à l'échange de données

Comme nous l'avons vu, le manque de cadres pour l'échange de données est un frein pour les entreprises. C'est pourquoi les entreprises qui opèrent dans de bonnes conditions sont plus à même de faire du partage de données. Avoir des lois qui protègent les données et qui fournissent le contexte et les conditions pour leur partage incite les entreprises à échanger des données. Il est important de fournir des garanties aux entreprises pour que ces dernières n'aient pas de crainte concernant leurs données.

Orange et les services touristiques de la ville de Barcelone ont mis en place un partenariat pour analyser les comportements des touristes dans la ville. Pour faciliter l'échange de données, Orange a mis en place une plateforme qui fonctionne de la façon suivante : Orange analyse son réseau téléphonique et de Wifi pour détecter les touristes, compare ses données avec celles de Barcelone (tickets, réservations, etc.) et les centralisent pour les analyser. La ville propose ensuite aux touristes les meilleurs créneaux pour visiter les différents monuments. Les profils sont anonymes et la ville de Barcelone détient les données pour garantir qu'Orange ne les utilise pas (« IoT & Big Data - Sagrada Familia », 2021).

7.6 Facteurs économiques

En plus des autres facteurs que nous avons cités, il y a tout le côté économique à prendre en compte. En effet, pour que des entreprises fassent un partage de données, il faut qu'elles aient toutes les deux des bénéfices à en tirer. Comme nous l'avons évoqué avec le dilemme du prisonnier, les entreprises seront plus enclines à collaborer si elles en retirent un avantage. Nous pouvons citer ici l'exemple de Michelin avec Nespresso et McDonalds. Michelin permet aux entreprises d'utiliser son application ViaMichelin et en échange, ces dernières incitent leurs

propres clients à utiliser l'application de Michelin (European Commission. Directorate General for Communications Networks, Content and Technology. & everis Benelux., 2018).

Le partage de données demande des capacités matérielles et pécuniaires de par les installations ainsi que par le personnel impliqué. C'est pourquoi le partage de données ne peut s'effectuer que si les entreprises sont certaines d'en tirer un bénéfice ou par le biais d'une autre entreprise spécialisée. La plateforme Yggio, développée par Sensitive, est un bon exemple. Yggio a pour objectif est de centraliser toutes les données de consommation des bâtiments. Pour se financer, elle repose sur deux façons principales : les subsides gouvernementaux et les financements des entreprises partenaires. Elle doit donc convaincre de son utilité pour pouvoir perdurer (European Commission. Directorate General for Communications Networks, Content and Technology. & everis Benelux., 2018).

Bien qu'il ne soit pas directement économique, le dernier facteur économique à prendre en compte est la connaissance obtenue. Nous différencions ce facteur économique du premier car le résultat n'est pas le même. Dans le premier, le bénéfice est purement monétaire alors qu'ici il est intellectuel. Bien que sur le long terme il soit un but économique, échanger des données dans le but d'obtenir de nouveaux savoirs ou bénéficier de l'expérience de son partenaire de partage de données est facteur de succès. Nous pouvons citer ici l'exemple de Nordpool et geo. Ces deux entreprises ont collaboré sur les thermostats de geo pour pouvoir optimiser les coûts et la production de l'énergie. Grâce à cette collaboration, les entreprises ont pu développer des algorithmes et ont permis à leurs clients d'économiser 30% à 40% du prix initial tout en améliorer leurs relations avec ces derniers (European Commission. Directorate General for Communications Networks, Content and Technology. & everis Benelux., 2018).

Chapitre 8 - Étude de cas : EMAsphere

8.1 *Sélection de l'entreprise*

Afin de faire le lien entre théorie et pratique, il nous a semblé important de réaliser une étude de cas portant sur une entreprise active sur le marché du partage de données. Après quelques recherches, notre choix s'est porté sur l'entreprise EMAsphere.

EMAsphere est une entreprise éditant une plateforme cloud de gestion de la performance. La plateforme se rapproche d'une « solution BI, mais avec tout le côté préparamétré et prépackagé qui fait que la plateforme a toute une série d'indicateurs de tableau de bord prêt à l'emploi » (Annexe 5). Concrètement, la plateforme se connecte sur les bases de données de ses clients pour extraire les données. Ensuite, ces données sont transformées en informations grâce à des tableaux, des indicateurs ou encore des graphiques. Le but de l'application est d'informer au mieux les dirigeants d'entreprise pour qu'ils prennent les bonnes décisions.

De par la nature de son activité, EMAsphere est une entreprise particulièrement familière avec le concept d'échange de données entre entreprises. C'est pour cette raison que nous l'avons sélectionnée afin de rédiger cette étude de cas.

8.2 *Méthodologie*

Nous avons décidé de mener un entretien semi-dirigé avec deux collaboratrices d'EMAsphere (Annexe 5). Nous avons donc rédigé un guide d'entretien qui nous servira de fil rouge, tout en laissant place à l'improvisation (Annexe 4). Cette méthode d'entretien nous a semblé la plus naturelle, nous permettant ainsi de recueillir des témoignages plus sincères.

Afin d'obtenir une diversité dans nos intervenants, nous avons sélectionné deux collaboratrices de départements et d'implémentations différents. D'une part, Florence Noblet est la responsable du département marketing digital en Belgique. D'autre part, Marine Reillon s'occupe du volet offline. Autrement dit, elle est chargée de la gestion des événements, relations

presse, gestion des partenariats mais aussi création de contenu à destination des commerciaux. Madame Reillon est implantée à Paris, en France.

L'entretien a eu lieu le 21 juillet 2021 via vidéoconférence. Il a duré une heure trente. Vous pouvez retrouver l'intégralité de cette interview grâce à la retranscription disponible dans l'Annexe 5.

8.3 *Description de l'entreprise*

EMAsphere est une entreprise belge dont le siège social se situe à Louvain-la-Neuve. Elle a commencé à s'internationaliser en 2018, avec notamment une antenne parisienne bien implantée. Au total, plus de 60 collaborateurs sont répartis dans trois bureaux : Louvain-la-Neuve (Wallonie et Bruxelles), Ghent (Flandre) et Paris (France) (EMAsphere, 2021b).

Selon leur site web, EMAsphere compte à ce jour plus de 10.000 entreprises dans leur clientèle. Leurs clients proviennent de secteurs divers et variés. On y trouve des PME comme de plus grandes entreprises. Cela semble montrer que leur solution est adaptée à tout type d'entreprise.

Concernant leur produit, la plateforme EMAsphere repose sur des « connecteurs » (EMAsphere, 2021a). Ces connecteurs sont des petites applications dont le rôle est d'obtenir des données en provenance des systèmes informatiques de leurs clients. Ces connecteurs peuvent utiliser plusieurs types de transferts comme l'API, l'ETL ou encore l'envoi de fichiers, afin de s'adapter à tout type de bases de données. Une fois les données récoltées, elles sont analysées par la plateforme et il devient possible de les transformer en tableaux de bord (Annexe 5).

8.4 *La valeur des données chez EMAsphere*

EMAsphere ne possède pas de cadre général pour établir la valeur d'une donnée. Dans les faits, EMAsphere utilise une combinaison des trois approches que nous avons vues dans le chapitre 2.1.3.

- **L'approche par le marché** - La principale méthode utilisée par EMAsphere est d'analyser les prix des données sur le marché. Selon Florence Noblet, EMAsphere est « *contraint par les acteurs dans le sens où c'est eux qui fixent les prix et c'est nous qui tentons de négocier* ».
- **L'approche par les revenus** – EMAsphere est particulièrement attentive aux revenus qui seront générés grâce aux données qu'elle va acheter. Marine Reillon cite notamment un exemple concernant le département marketing. Elle explique qu'elle donne plus de valeur à des données appartenant à un prospect ayant déjà montré un intérêt pour EMAsphere plutôt qu'à un prospect n'ayant jamais entendu parler de l'entreprise. De fait, un prospect qui s'est intéressé à un produit est plus susceptible de conclure une vente qu'un prospect lambda. EMAsphere attribue plus de valeur à ces données.
- **L'approche par les coûts** – Comme le dit Marine Reillon : « *Les prix des données doivent être inférieurs au coût d'acquisition d'un SQL* ». Autrement dit, le prix d'une donnée n'a pas d'intérêt à excéder le coût nécessaire pour créer soi-même cette donnée.

On voit donc que les trois méthodes se complètent chez EMAsphere.

8.5 Les modèles économiques utilisés chez EMAsphere

Au cours de notre entretien avec EMAsphere, nous avons pu identifier deux modèles économiques utilisés par EMAsphere. En effet, nous avons pu identifier deux types très différents d'échanges de données entre entreprises. D'une part, il y a les échanges avec les clients grâce à sa plateforme. D'autre part, il y a les échanges de bases de données avec les partenaires en vue de trouver des opportunités de vente.

Echanges de données avec les clients

Comme nous l'avons vu, l'activité principale d'EMAsphere est d'aider à la gestion de la performance d'une entreprise en analysant ses données afin de les proposer de manière intelligible et intuitive à son dirigeant. Pour analyser le modèle économique utilisé par

EMAsphere, nous pouvons faire appel au modèle Performance Planning Value Chain que nous avons explicité au chapitre 4.2.2.

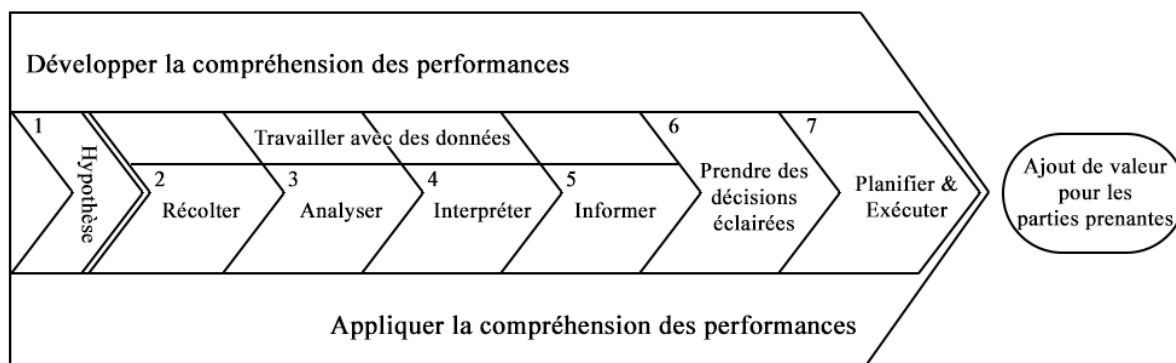


Figure 13 - Performance planning value chain³⁰ (Neely & Jarrar, 2004)

Ce modèle permet de comprendre le processus par lequel les données permettent de créer de la valeur. La plateforme d'EMAsphere peut s'intégrer dans trois maillons :

- **(3) Analyser** – Grâce à l'expertise d'EMAsphere, les données sont analysées. Pour faire écho au chapitre 1.1.1, EMAsphere transforme des données en informations utilisables.
- **(4) Interprétation** – Grâce à ses tableaux de bord et autres visuels, EMAsphere aide à l'interprétation et à la compréhension des données.
- **(5) Informer** – L'information devient plus facile à intégrer si elle est claire et intuitive.

Selon ce modèle, la plateforme EMAsphere s'apparente à de la sous-traitance d'une partie du processus de valorisation des données d'une entreprise.

Echanges de données avec les partenaires

Lors de l'entretien avec Florence Noblet et Marine Reillon, nous avons découvert qu'EMAsphere échangeait des données avec ses partenaires. Les données échangées dans ce

³⁰ Version originale non traduite disponible à l'annexe 3.

cas par EMAsphere sont des données de première main (principalement des données marketing). Il ne s'agit en aucun cas de données des clients (données de seconde main).

Cet échange part d'une constatation. Beaucoup d'entreprises possèdent les mêmes cibles marketing qu'EMAsphere, sans pour autant être concurrents. En effet, une entreprise a besoin de plusieurs logiciels : un logiciel de paie, de comptabilité, etc. Bien que les entreprises éditant des logiciels ciblent toutes les entreprises au sens large, un logiciel de paie ne fait pas concurrence à un logiciel de comptabilité.

Toutefois, dû aux obstacles que nous évoquerons plus tard, EMAsphere n'avait pas pour habitude d'échanger des données avec des entreprises similaires mais non concurrentes. Il aura fallu l'émergence de la plateforme Reveal pour qu'EMAsphere se lance dans ce genre de partenariat.

La plateforme Reveal se veut être un acteur tiers permettant à plusieurs entreprises de connecter leur application CRM³¹ (Reveal, 2021). L'instauration de cet acteur tiers a donné la confiance nécessaire à EMAsphere pour échanger des données avec des entreprises partenaires.

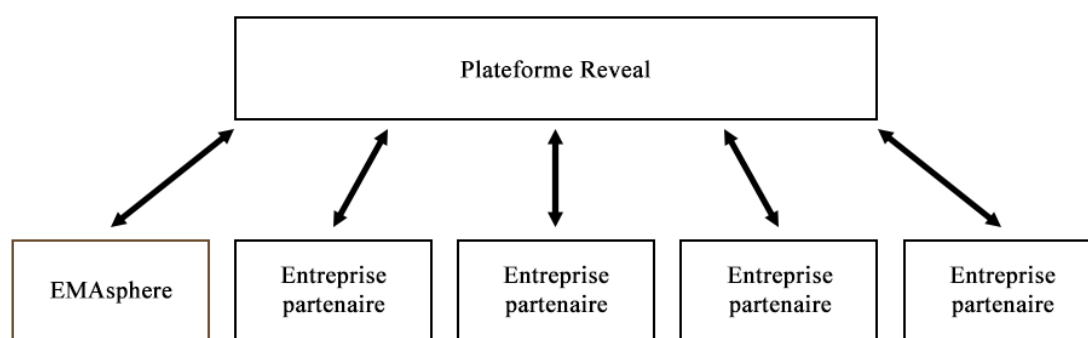


Figure 14 - Fonctionnement de la plateforme Reveal

Tout d'abord, chaque entreprise connecte son application CRM à Reveal. Ce dernier a accès à toutes les données marketing. Chaque entreprise invite ses entreprises partenaires à accéder à ses données et vice-versa. Ainsi, les données peuvent être partagées.

³¹ Customer Relation Management. En français : « Gestion de la relation client ».

La subtilité de cette plateforme, c'est qu'elle ne partage pas toutes vos données avec vos entreprises partenaires. Ces dernières n'ont accès qu'aux contacts qu'elles possèdent déjà dans son CRM. Prenons un exemple fictif : EMAsphere et l'entreprise Dupont font un partenariat sur Reveal. EMAsphere n'a pas accès à toutes les données de l'entreprise Dupont. Elle ne peut avoir que les données des contacts d'EMAsphere. L'intérêt pour EMAsphere est de pouvoir compléter sa base de données existante, pas de l'agrandir.

8.6 *Les obstacles rencontrés et leurs solutions*

Lors de notre discussion avec EMAsphere, nous avons pu identifier quelques obstacles auxquels l'entreprise a été confrontée. Nous les avons classés dans les différentes catégories énumérées dans le chapitre 5.

L'obligation d'avoir confiance

L'obstacle le plus évoqué par EMAsphere est la nécessité d'avoir confiance en un partenaire pour pouvoir échanger des données. Comme le précise Marine Reillon : « *Pour un partenariat sur le long terme, la confiance, c'est indispensable* ». Actuellement, il est hors de question d'échanger des données avec des inconnus. La principale raison est qu'EMAsphere a besoin d'une garantie avec un échange de données donnant-donnant.

En conséquence, EMAsphere partage donc des données avec peu d'acteurs. Que ça soit du côté fournisseur ou du côté utilisateur des données. Cependant, au vu de la réussite des échanges de données chez EMAsphere, cette liste ne cesse de croître.

Manque de standardisation

La plateforme EMAsphere se connecte aux systèmes informatiques de ses clients pour collecter les données dont elle a besoin. Le problème est que chaque système informatique est différent. L'équipe IT d'EMAsphere est confrontée au manque de standardisation. Cette dernière représente un coût important chez EMAsphere car cela se traduit par de nombreuses heures de travail pour les développeurs afin de parvenir à récolter les bonnes données dans le format adéquat.

L'infrastructure coûteuse

Tout d'abord, l'infrastructure nécessaire au stockage, au maintien et à l'analyse des données représente un coût important pour EMAsphere. Pour ne prendre que l'exemple du département marketing, un logiciel CRM est nécessaire pour naviguer dans la base de données clients. Ce logiciel représente déjà, à lui seul, un coût important.

De plus, il est nécessaire de garder les données constamment à jour. Pour ce faire, EMAsphere engage des étudiants pour entretenir et nettoyer la base de données.

Enfin, la mise en place de l'infrastructure et la gestion des étudiants nécessite une organisation importante. Comme Florence Noblet l'indique : « l'infrastructure a été un challenge ». Toutefois, elle précise qu'une fois tout en place, cela demande moins d'efforts.

Qualité insuffisante des données et incertitude stratégique

Il est déjà arrivé à de multiples reprises qu'EMAsphere acquière une base de données et qu'après une analyse rapide, ils s'aperçoivent que les données sont de mauvaise qualité. Acquérir une base de données, c'est toujours un risque car on ne sait pas toujours ce que l'on va obtenir.

Pour pallier ce problème, EMAsphere demande des garanties supplémentaires avant les échanges. D'une part, EMAsphere dresse une liste très précise des caractéristiques que les données doivent posséder. Par exemple : « *La base de données doit pouvoir prendre en compte les caractères spéciaux* ». D'autre part, EMAsphere exige également un échantillon de la base de données pour constater la qualité des données. Toutefois, comme le précise Florence Noblet : « *Ça ne garantit rien car les échantillons ne sont généralement pas aléatoires mais sélectionnés* ».

Les contraintes du RGPD

L'instauration du RGPD a eu un profond impact sur l'échange de données chez EMAsphere.

Lorsque le RGPD est entré en vigueur, il y a eu un temps où les entreprises n'osaient simplement plus partager des données. Il a fallu un temps d'apprentissage pour les entreprises afin de comprendre ce qu'elles pouvaient faire ou non. Durant ce temps d'apprentissage, EMAsphere s'est montré très prudente quant à l'utilisation de ses données.

Le RGPD a engendré de nouvelles contraintes pour EMAsphere, et particulièrement pour le département marketing. Un travail a été mené pour s'assurer que leur base de données respectait le nouveau règlement. EMAsphere doit également s'assurer que ses différents partenaires respectent ce règlement.

Florence Noblet et Marine Reillon ont insisté sur les bons côtés du RGPD. Le règlement a permis une conscientisation sur l'utilisation de la donnée. Comme l'illustre Florence Noblet : *« Par exemple, avant le GDPR, quand un contact téléchargeait un ebook, il entrait dans notre base de données et puis c'est tout. Maintenant, on a un encadré qui l'informe de l'intérêt légitime lié à l'utilisation de ses données. Il sait que s'il remplit le formulaire, il sera contacté par EMAsphere. Désormais, la donnée est plus propre. ».*

8.7 Les facteurs clés de succès

Selon EMAsphere, le principal facteur clé de succès de l'échange de données est la confiance et la communication avec ses partenaires. Comme le dit Marine Reillon, il s'agit de faire en sorte qu'aucune des parties prenantes ne se sente lésée. Selon elle, c'est la meilleure solution pour que l'échange de données fonctionne sur le long terme.

Florence Noblet et Marine Reillon mettent en lumière l'importance d'inclure des acteurs tiers dans l'échange de données. Ces derniers permettent d'aligner au mieux les objectifs de chacun. Ainsi, les échanges peuvent avoir lieu de manière pérenne.

Elles s'accordent également à dire que la simplicité et la sécurité de la méthode de transfert ont, elles aussi, leur rôle à jouer dans le succès de l'échange de données. Depuis qu'EMAsphere utilise la plateforme Reveal, l'entreprise est nettement plus encline à partager ses données.

Chapitre 9 - Conclusions

Ce mémoire a pour objectif de montrer les différentes manières par lesquelles le partage de données peut représenter un avantage pour les entreprises européennes en étudiant la situation actuelle ainsi que la littérature existante.

Pour cela, nous avons, dans un premier temps, fait un récapitulatif de la littérature existante sur les données et le domaine du partage de données en général. Nous avons défini notre approche de la question de recherche ainsi ce que nous entendions par entreprise européenne.

Pour bien comprendre la problématique du partage de données, nous devons comprendre comment les données sont traitées en entreprise ainsi que la façon dont leur valeur est estimée. Nous avons dû faire face à un premier obstacle : il n'existe pas de cadre faisant consensus pour calculer la valeur, il n'existe que des méthodes pour les estimer.

Un autre point important de notre étude est l'état de la situation actuelle du partage de données. Nous avons commencé par aborder l'état général du marché de la donnée et avons constaté qu'il est en pleine expansion avec de nombreuses opportunités. Cependant, la situation européenne affiche un retard par rapport aux zones asiatiques et américaines. De plus, il n'y a pas de situation uniforme entre les différents pays. Pour faire face à ce retard et ces inégalités, la Commission européenne a lancé de nouveaux plans pour accélérer la croissance du secteur tout en maintenant la protection des données du point de vue du droit.

Nous avons ensuite adapté des modèles existants pour le partage de données. Pour ce faire, nous nous sommes basés sur trois axes de types de modèle de partage de données : entre entités, sur la chaîne de valeur et transversal. Bien que le partage de données n'était pas envisagé pour chaque modèle que nous avons repris, nous avons su l'intégrer à chacun d'eux.

Nous avons abordé les différents obstacles que rencontrent les entreprises lorsqu'elles veulent faire du partage de données. Ces obstacles trouvent trois origines différentes. La première est les difficultés qu'ont les entreprises entre elles (confiance, manque de standardisation, etc.). La deuxième est liée aux objectifs de l'entreprise (coût des

infrastructures, intérêts stratégiques, etc.). Et la dernière concerne le cadre dans lequel les partages de données s'effectuent (manque de cadres légaux, perte de contrôle des données, etc.).

Pour pallier ces éléments qui font obstacles au partage des données, plusieurs solutions ont été formulées bien que nous n'excluons pas que d'autres pistes puissent être envisagées.

Nous avons logiquement abordé ensuite les facteurs clés du succès du partage de données pour augmenter le partage de données entre entreprises européennes. En plus de rétablir la confiance aux entreprises, nous avons découvert que les partenariats avec des cadres bien définis sont l'un des facteurs clés les plus importants pour le partage de données.

Pour finir, nous avons confronté la théorie avec la pratique en faisant une étude de cas avec l'entreprise EMAsphere, spécialisée dans une plateforme cloud de gestion de la performance. Cette étude a permis d'illustrer plusieurs points de théorie que nous avons développés précédemment de même que les obstacles et facteurs clés que nous avons identifiés.

Ce mémoire avait pour but premier de centraliser et adapter la théorie existante, étant donné que l'étude du domaine des données et du partage est relativement récente dans le monde universitaire. Son second objectif était d'ouvrir la voie à de futures études et travaux, grâce notamment à nos analyses sur la situation actuelle.

Chapitre 10 - Limites du mémoire

Dans ce mémoire, nous avons tenté de mettre en lumière comment le partage de données peut représenter une opportunité de croissance pour les entreprises européennes. Cependant, ce mémoire possède des limites quant au sujet abordé.

En premier lieu, ce mémoire étant la conclusion d'un master de deux années, il a été écrit sur un laps de temps relativement court. Nos analyses et nos observations sont donc limitées car nous n'avons pas pu analyser d'évolution à grandes échelles. Nous avons analysé la situation actuelle des données uniquement sur la courte période de notre master et sur des travaux effectués avant le nôtre. Le but de notre mémoire a donc vocation à être intégré dans un ensemble de recherches et d'études dont celui-ci serait l'un des points de départ.

En lien avec le point précédent, ce mémoire est avant tout une analyse de la littérature actuelle sur les données et leurs partages. Ce mémoire est un état des lieux de la situation en 2021 et il a vocation à servir de base théorique à des études futures. Ces dernières devront principalement se concentrer sur le partage de données tel qu'il est pratiqué dans les faits. Le domaine étant en renouvellement constant, il est toujours possible qu'une nouvelle technologie vienne modifier profondément le système actuel et rende archaïque une partie de la théorie que nous avons développée dans ce mémoire.

Bien que nous ayons fait une étude de cas sur une entreprise, nous ne pouvons considérer que cette dernière est représentative de la pratique. En effet, le partage de données est un domaine vaste qui requiert énormément d'analyses de différentes entreprises. Notre étude de cas peut donc être vue comme une comparaison de la théorie avec le réel sans pourtant être considérée comme la vérité absolue. Il serait donc intéressant de comparer la théorie que nous avons développée avec d'autres entreprises qui pratiquent le partage de données.

La dernière limite que nous voulons aborder concerne le domaine des données et du partage de données en général. Nous avons rencontré deux difficultés principales. Tout d'abord, c'est un domaine qui a peu de recul. Les données et le partage de celles-ci ne sont étudiés que depuis récemment. Nous étions donc limités dans la littérature qui est disponible sur le sujet. La deuxième difficulté est que c'est un domaine qui est vital pour les entreprises. En effet,

comme les données sont sensibles, les entreprises sont réticentes à partager les leurs ainsi que leur utilisation. Nous l'avons vu par exemple avec l'étude conduite par la Commission européenne qui n'a eu que 150 entreprises volontaires à y participer. Il est donc difficile d'obtenir des cas de situations concrètes.

Chapitre 11 - Futures recherches

Comme nous l'avons évoqué précédemment dans notre mémoire, ce dernier est la base pour lancer de futures recherches sur le partage de données entre entreprises. Nous classons les idées de futures recherches en plusieurs types.

La première idée concerne le type d'entreprises : nous nous sommes centrés sur les entreprises européennes dans ce mémoire et il serait intéressant d'étudier la situation des entreprises dans d'autres régions du monde (Amérique, Asie, etc.). Suite à ces études, il serait adéquat de comparer les différentes régions pour comprendre les avantages de chaque région pour le partage de données ainsi que les facteurs de réussite.

La seconde est d'effectuer une recherche sur les facteurs clés que nous avons identifiés dans ce mémoire. Il serait intéressant d'étudier l'impact de chacun de ces facteurs, ainsi que la variation de leur importance en fonction du type d'entreprise. Le but serait de pouvoir produire des guides pour chaque entreprise.

La troisième idée est d'effectuer une étude sur le long terme. Nous avons développé des idées de solutions pour augmenter le partage de données et il serait intéressant d'en implanter dans des entreprises et de suivre l'évolution du partage de données au sein de celles-ci sur le long terme.

La quatrième concerne les études de cas. Notre mémoire n'ayant que le témoignage d'une seule entreprise, il serait intéressant de faire une étude de grande échelle avec un maximum d'entreprises de façon à recueillir le maximum d'informations et de pouvoir les comparer avec la théorie que nous avons développée.

La dernière idée est de faire une étude qui aurait pour but de développer un modèle économique pour déterminer le prix des données. Comme nous l'avons vu dans ce mémoire, il n'existe actuellement aucun cadre précis pour déterminer le prix, juste des méthodes pour l'estimer. Il serait intéressant de créer ce cadre pour uniformiser la méthode du calcul du prix des données entre entreprises et faciliter l'échange.

Bibliographie

- Ackoff, R. (1989). From Data to Wisdow. *Journal of Applied Systems Analysis*, 16, 3-9.
- Affaire Carlos Ghosn. (2021). In *Wikipédia*.
https://fr.wikipedia.org/w/index.php?title=Affaire_Carlos_Ghosn&oldid=185364778
- AIM alliance. (2021). In *Wikipedia*.
https://en.wikipedia.org/w/index.php?title=AIM_alliance&oldid=1033483982
- Alves, V., Campos, P., & Felício, M. (2017). Analysis of Trust in B2B Relationships : The case of automatic storage and retrieval systems. *International Journal of Marketing, Communication and New Media*, 5(8), Article 8.
<http://u3isjournal.isvouga.pt/index.php/ijmcnm/article/view/220>
- Baldwin, C. Y., & Woodard, C. J. (2008). The Architecture of Platforms : A Unified View. *SSRN Electronic Journal*, 31. <https://doi.org/10.2139/ssrn.1265155>
- Bellinger, G., Castro, D., & Mills, A. (2004). *Data, Information, Knowledge, and Wisdom*.
<http://www.systems-thinking.org/dikw/dikw.htm>
- Berthier, A., & Fulchiron, T. (2018). *La deuxième vie des données* [Document non publié]. Mines ParisTech.
- Birnholtz, J. P., & Bietz, M. J. (2003). Data at work : Supporting sharing in science and engineering. *Proceedings of the 2003 International ACM SIGGROUP Conference on Supporting Group Work - GROUP '03*, 339. <https://doi.org/10.1145/958160.958215>
- Castle, L. (2021, août 2). Volkswagen Truck & Bus launching transportation platform for all with location by HERE. <https://360.here.com/2016/12/14/volkswagen-truck-bus-launching-transportation-platform-for-all-with-location-by-here/>
- Charest, G., & Rogers, M. (2020). *Data Exchange : Methods and Considerations*. 14.

Cini, C. (2019). *Marketing Facebook : Guide stratégique pour la communication et la publicité*. Eyrolles.

Orientations concernant le partage des données du secteur privé dans l'économie européenne des données, 125 SWD 21 (2018). <https://eur-lex.europa.eu/legal-content/FR/TXT/PDF/?uri=CELEX:52018SC0125&from=EN>

Commission européenne - European Commission. (2020). *A European strategy for data* (Rapport Final COM(2020) 66; p. 35). https://ec.europa.eu/info/sites/default/files/communication-european-strategy-data-19feb2020_en.pdf

Commission européenne - European Commission. (2021a, juin 4). *The EU Data Landscape*. Data Landscape. <https://datalandscape.eu/eu-data-landscape>

Commission européenne - European Commission. (2021, juin 7). *Façonner l'avenir numérique de l'Europe* [Text]. https://ec.europa.eu/info/strategy/priorities-2019-2024/europe-fit-digital-age/shaping-europe-digital-future_fr

Commission européenne - European Commission. (2021b, juin 9). *À quoi correspondent les données à caractère personnel?* https://ec.europa.eu/info/law/law-topic/data-protection/reform/what-personal-data_fr

De Mauro, A., Greco, M., & Grimaldi, M. (2016). A formal definition of Big Data based on its essential features. *Library Review (Glasgow)*, 65(3), 122-135. <https://doi.org/10.1108/LR-06-2015-0061>

Deloitte. (2021, août 6). *Data is the new gold*. <https://www2.deloitte.com/global/en/pages/real-estate/articles/future-real-estate-data-new-gold.html>

Dreller, A. (2018). *Creating Value from Data Sharing : Future-oriented Business Models in Theory and Practice*. Springer Fachmedien Wiesbaden. <https://doi.org/10.1007/978-3-658-23276-4>

- Dussauge, P., & Garrette, B. (1995). *Les stratégies d'alliance*. Editions d'Organisation.
- Dutta, A. K., & Hasan, R. (2013). How Much Does Storage Really Cost? Towards a Full Cost Accounting Model for Data Storage. *Economics of Grids, Clouds, Systems, and Services*, 8193, 29-43. https://doi.org/10.1007/978-3-319-02414-1_3
- EMAsphere. (2021a, mai 1). *Connecteurs—EMAsphere*. <https://www.emasphere.com/fr/solution/plateforme/connecteurs>
- EMAsphere. (2021b, août 3). *A propos—EMAsphere*. <https://www.emasphere.com/fr/a-propos>
- European Commission. Directorate General for Communications Networks, Content and Technology. & everis Benelux. (2018). *Study on data sharing between companies in Europe : Case studies*. (p. 101). Publications Office. <https://data.europa.eu/doi/10.2759/333340>
- European Commission & Everis Benelux. (2018). *Study on data sharing between companies in Europe : Final report*. (p. 158). Publications Office. <https://data.europa.eu/doi/10.2759/354943>
- Eurostat. (2021). *Population* [Database]. https://ec.europa.eu/eurostat/databrowser/view/demo_pjan/default/table?lang=en
- Finger, M., & Da Silva, F. S. (1998). Temporal data obsolescence : Modelling problems. *Proceedings. Fifth International Workshop on Temporal Representation and Reasoning (Cat. No.98EX157)*, 45-50. <https://doi.org/10.1109/TIME.1998.674130>
- Fleckenstein, M., & Fellows, L. (2018). *Modern Data Strategy*. Springer International Publishing. <https://doi.org/10.1007/978-3-319-68993-7>
- Franco-Santos, M., Kennerley, M., Micheli, P., Martinez, V., Mason, S., Marr, B., Gray, D., & Neely, A. (2007). Towards a definition of a business performance measurement system. *International journal of operations & production management*, 27(8), 784-801. <https://doi.org/10.1108/01443570710763778>

- Grether, M. (2016). Using Big Data for Online Advertising Without Wastage : Wishful Dream, Nightmare or Reality? *GfK Marketing Intelligence Review*, 8(2), 38-43. <https://doi.org/10.1515/gfkmir-2016-0014>
- Guo, H., Wang, L., Chen, F., & Liang, D. (2014). Scientific big data and Digital Earth. *Chinese Science Bulletin (Chinese Version)*, 59, 1047. <https://doi.org/10.1360/972013-1054>
- Gupta, A. K., & Govindarajan, V. (2000). Knowledge flows within multinational corporations. *Strategic management journal*, 21(4), 473-496. [https://doi.org/10.1002/\(SICI\)1097-0266\(200004\)21:4<473::AID-SMJ84>3.0.CO;2-I](https://doi.org/10.1002/(SICI)1097-0266(200004)21:4<473::AID-SMJ84>3.0.CO;2-I)
- IBM. (2021, avril 9). *What is an Application Programming Interface (API)*. <https://www.ibm.com/cloud/learn/api>
- IoT & Big Data—Sagrada Familia. (2021, août 2). *D-LAB*. <https://d-lab.tech/project-1/>
- IVS 210 : Intangible Assets—Exposure Draft*. (2016). International Valuation Standards Council. <https://www.ivsc.org/files/file/view/id/647>
- Johnson & Johnson achète Momenta pour environ 6,5 milliards de dollars. (2020, août 19). *FR24 News France*. <https://www.fr24news.com/fr/a/2020/08/johnson-johnson-achete-momenta-pour-environ-65-milliards-de-dollars.html>
- Jolly, D. (2002). Vers une typologie des alliances technologiques. *La Revue des Sciences de Gestion, Direction et Gestion*, 196-197, 121-130. <https://doi.org/10.1051/larsg:2002030>
- Levenstein, M. (2019). *The Researcher Passport : Improving Data Access and Confidentiality Protection, Global, 2017-2018. Generic*. <https://doi.org/10.3886/icpsr37454>
- Li, H., Yu, L., & He, W. (2019). The Impact of GDPR on Global Technology Development. *Journal of global information technology management : JGITM*, 22(1), 1-6. <https://doi.org/10.1080/1097198X.2019.1569186>

- Madnick, S., Wang, R., & Xian, X. (2003). The Design and Implementation of a Corporate Householding Knowledge Processor to Improve Data Quality. *Journal of Management Information Systems*, 20(3), 41-70. <https://doi.org/10.1080/07421222.2003.11045772>
- Marr, B. (2016). *Big Data in Practice : How 45 Successful Companies Used Big Data Analytics to Deliver Extraordinary Results*. John Wiley & Sons, Incorporated.
- Mastelic, T., Oleksiak, A., Claussen, H., Brandic, I., Pierson, J.-M., & Vasilakos, A. V. (2015). Cloud Computing : Survey on Energy Efficiency. *ACM Computing Surveys*, 47(2), 1-36. <https://doi.org/10.1145/2656204>
- McGinnes, S. (2011). *The Problem of Conceptual Incompatibility*. 13.
- Mimbang, J. B., Van Steenkiste, I., & 50Minutes.fr. (2015). *La théorie des jeux : Nash et le dilemme du prisonnier*. 50 Minutes.
- Morgan, C., & Gardiner, P. H. B. (1990). Data refinement by calculation. *Acta Informatica*, 27(6). <https://doi.org/10.1007/BF00277386>
- Moro Visconti, R. (2020). *The Valuation of Digital Intangibles : Technology, Marketing and Internet* (Palgrave Macmillan). Springer International Publishing. <https://doi.org/10.1007/978-3-030-36918-7>
- Nash, J. F. (1950). Equilibrium Points in N-Person Games. *Proceedings of the National Academy of Sciences of the United States of America*, 36(1), 48-49.
- Neely, A., & Jarrar, Y. (2004). Extracting value from data—The performance planning value chain. *Business process management journal*, 10(5), 506-509. <https://doi.org/10.1108/14637150410559180>
- OECD. (2019). *Enhancing Access to and Sharing of Data : Reconciling Risks and Benefits for Data Re-use across Societies*. OECD. <https://doi.org/10.1787/276aaca8-en>
- Office Québécois de la langue française. (2005). *Interface de programmation* [Dictionnaire]. http://gdt.oqlf.gouv.qc.ca/ficheOqlf.aspx?Id_Fiche=8394269

- Porter, M. E. (1985). *Competitive advantage : Creating and sustaining superior performance*. Free press.
- Rama, B. G. (2017, janvier 8). *Report : AWS Market Share Is Triple Azure's*. AWSInsider. <https://awsinsider.net/articles/2017/08/01/aws-market-share-3x-azure.aspx>
- Ravelomanantsoa, M. S., Ducq, Y., & Vallespir, B. (2019). A state of the art and comparison of approaches for performance measurement systems definition and design. *International journal of production research*, 57(15-16), 5026-5046. <https://doi.org/10.1080/00207543.2018.1506178>
- Règlement (UE) 2016/679 du Parlement européen et du Conseil du 27 avril 2016 relatif à la protection des personnes physiques à l'égard du traitement des données à caractère personnel et à la libre circulation de ces données, et abrogeant la directive 95/46/CE (règlement général sur la protection des données) (Texte présentant de l'intérêt pour l'EEE), Pub. L. No. 32016R0679, 119 OJ L 88 (2016). <http://data.europa.eu/eli/reg/2016/679/oj/fra>
- Reilly, R. F., & Schweih, R. P. (1999). *Valuing intangible assets*. McGraw-Hill.
- Reveal. (2021, août 3). *Reveal | Partnership, Growth, and Sales Ecosystem Solution*. <https://reveal.co/>
- Rowley, J. (2007). The wisdom hierarchy : Representations of the DIKW hierarchy. *Journal of Information Science*, 33(2), 163-180. <https://doi.org/10.1177/0165551506070706>
- Schneider, M. J., Jagpal, S., Gupta, S., Li, S., & Yu, Y. (2017). Protecting customer privacy when marketing with second-party data. *International Journal of Research in Marketing*, 34(3), 593-603. <https://doi.org/10.1016/j.ijresmar.2017.02.003>
- Segrestin, B. (2005). Partnering to explore : The Renault–Nissan Alliance as a forerunner of new cooperative patterns. *Research policy*, 34(5), 657-672. <https://doi.org/10.1016/j.respol.2005.02.006>

- Sorrell, S., & Dimitropoulos, J. (2008). The rebound effect : Microeconomic definitions, limitations and extensions. *Ecological Economics*, 65(3), 636-649. <https://doi.org/10.1016/j.ecolecon.2007.08.013>
- Storage : Price per capacity by technology 2009-2015*. (2021, juillet 12). Statista. <https://www.statista.com/statistics/204238/storage-price-per-capacity-by-technology/>
- Suzuki, M., Ando, N., & Nishikawa, H. (2019). Intra-organizational communication and its consequences. *Management decision*, 57(1), 71-85. <https://doi.org/10.1108/MD-02-2018-0159>
- Thar, B. D. E. (2019). How to Use Digital Marketing Data in Regulated Industries. In J. Barberis, D. W. Arner, & R. P. Buckley (Éds.), *The RegTech Book* (1^{re} éd.). Wiley. <https://doi.org/10.1002/9781119362197.ch24>
- The Lisbon Council. (2020). *The European data market monitoring tool* (Final Study Report D2.9; p. 101). European Commission. <https://digital-strategy.ec.europa.eu/en/library/european-data-market-study-update>
- Urban, T., Tatang, D., Degeling, M., Holz, T., & Pohlmann, N. (2020). Measuring the Impact of the GDPR on Data Sharing in Ad Networks. *Proceedings of the 15th ACM Asia Conference on Computer and Communications Security, ASIA CCS 2020*, 222-235. <https://doi.org/10.1145/3320269.3372194>
- Yusuf Warsame, M., & Banoff, A. (2007). *Webmarketing* [Mémoire]. Haute Ecole de Gestion de Genève.
- Zeleny, M. (2013). Integrated Knowledge Management: *International Journal of Information Systems and Social Change*, 4(4), 62-78. <https://doi.org/10.4018/jissc.2013100104>
- Zins, C. (2007). Conceptual approaches for defining data, information, and knowledge. *JASIST*, 58, 479-493. <https://doi.org/10.1002/asi.20508>

Zuliani, J.-M., & Jalabert, G. (2005). L'industrie aéronautique européenne : Organisation industrielle et fonctionnement en réseaux. *Espace géographique*, 34(2), 117-135.
<https://doi.org/10.3917/eg.342.0117>

Table des illustrations

FIGURE 1- THE DIKW HIERARCHY	4
FIGURE 2 - LA STRUCTURE DU MARCHE DE LA DONNEE (COMMISSION EUROPEENNE - EUROPEAN COMMISSION, 2021A)	18
FIGURE 3 - VALEUR DU MARCHE DE LA DONNEE (COMMISSION EUROPEENNE - EUROPEAN COMMISSION, 2021A)	19
FIGURE 4 - VALEUR DE L'ECONOMIE DE LA DONNEE (COMMISSION EUROPEENNE - EUROPEAN COMMISSION, 2021A)	20
FIGURE 5 - VALEUR DU MARCHE DE LA DONNEE EN EUROPE (THE LISBON COUNCIL, 2020)	21
FIGURE 6 - POURCENTAGE DU MARCHE DE LA DONNEE SUR LE PIB (2019) (THE LISBON COUNCIL, 2020)	22
FIGURE 7 - FONCTIONNEMENT D'UNE API (IBM, 2021)	24
FIGURE 8 - RESUME DES AVANTAGES ET INCONVENIENTS LIES AUX DIFFERENTES METHODES DE PARTAGE DE DONNEES (CHAREST & ROGERS, 2020)	26
FIGURE 9 - LA NOTION D'ALLIANCE EN QUATRE POINTS CLES (JOLLY, 2002)	27
FIGURE 10 - LA CHAINE DE VALEUR DE PORTER	34
FIGURE 11 - PERFORMANCE PLANNING VALUE CHAIN (NEELY & JARRAR, 2004)	36
FIGURE 12 - ÉVOLUTION DU COUT DE STOCKAGE DES DONNEES (STATISTA, 2021)	50
FIGURE 13 - PERFORMANCE PLANNING VALUE CHAIN (NEELY & JARRAR, 2004)	69
FIGURE 14 - FONCTIONNEMENT DE LA PLATEFORME REVEAL	70

Table des Annexes

ANNEXE 1 - VALEUR DU MARCHÉ DE LA DONNÉE EN EUROPE	89
ANNEXE 2 - RECAPITULATIF DE LA COMBINAISON DE LA CHAÎNE DE VALEUR DE PORTER ET DU PPVC	91
ANNEXE 3 - VERSION ORIGINALE DU PPCV (NEELY & JARRAR, 2004)	92
ANNEXE 4 - GUIDE D'ENTRETIEN AVEC EMASPHÈRE	93
ANNEXE 5 - RETRANSCRIPTION DE L'INTERVIEW DE FLORENCE NOBLET ET MARINE REILLON.....	97

Annexe 1 - Valeur du marché de la donnée en Europe

Pays	Market Value ³²	%UE27	%UE28	Population ³³	MV/Pop
Allemagne	16.306.713.490,00 €	28,0%	21,7%	83.019.213	196 €
Autriche	1.533.518.835,00 €	2,6%	2,0%	8.858.775	173 €
Belgique	1.903.094.122,00 €	3,3%	2,5%	11.455.519	166 €
Bulgarie	288.170.597,00 €	0,5%	0,4%	7.000.039	41 €
Chypre	177.427.910,00 €	0,3%	0,2%	875.899	203 €
Croatie	239.223.675,00 €	0,4%	0,3%	4.076.246	59 €
Danemark	1.585.107.490,00 €	2,7%	2,1%	5.806.081	273 €
Espagne	3.981.511.223,00 €	6,8%	5,3%	46.937.060	85 €
Estonie	268.438.016,00 €	0,5%	0,4%	1.324.820	203 €
Finlande	1.238.726.757,00 €	2,1%	1,6%	5.517.919	224 €
France	9.459.288.730,00 €	16,2%	12,6%	67.177.636	141 €
Grèce	515.114.082,00 €	0,9%	0,7%	10.724.599	48 €
Hongrie	525.075.752,00 €	0,9%	0,7%	9.772.756	54 €
Irlande	1.382.803.163,00 €	2,4%	1,8%	4.904.240	282 €
Italie	5.592.785.146,00 €	9,6%	7,4%	59.816.673	93 €
Lettonie	174.210.969,00 €	0,3%	0,2%	1.919.968	91 €
Lituanie	268.707.258,00 €	0,5%	0,4%	2.794.184	96 €
Luxembourg	170.182.408,00 €	0,3%	0,2%	613.894	277 €
Malte	71.063.397,00 €	0,1%	0,1%	493.559	144 €
Pays-Bas	4.362.329.672,00 €	7,5%	5,8%	17.282.163	252 €
Pologne	2.078.388.808,00 €	3,6%	2,8%	37.972.812	55 €
Portugal	1.057.664.320,00 €	1,8%	1,4%	10.276.617	103 €
République tchèque	729.104.644,00 €	1,3%	1,0%	10.649.800	68 €
Roumanie	664.243.704,00 €	1,1%	0,9%	19.414.458	34 €
Slovaquie	477.668.096,00 €	0,8%	0,6%	5.450.421	88 €

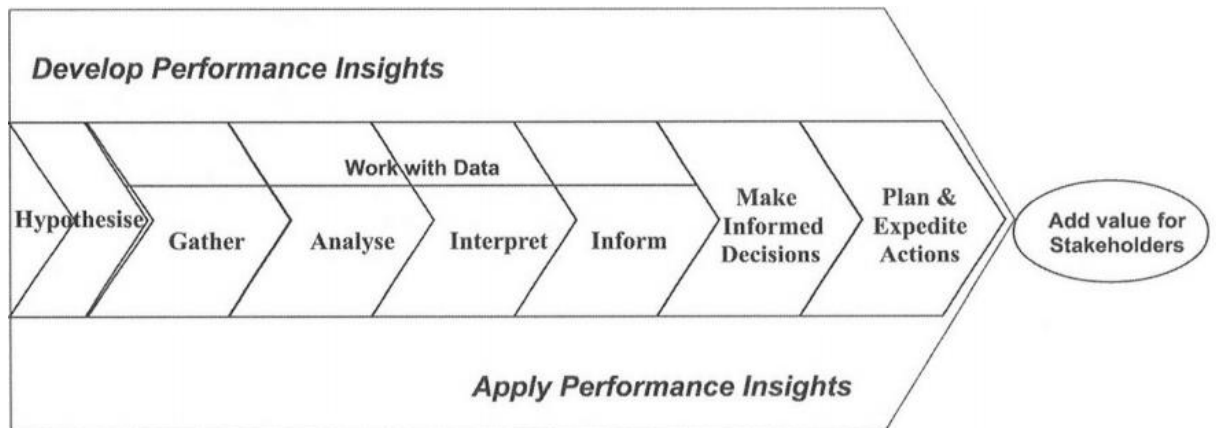
³² (The Lisbon Council, 2020)³³ (Eurostat, 2021)

Slovénie	167.034.611,00 €	0,3%	0,2%	2.080.908	80 €
Suède	2.996.074.373,00 €	5,1%	4,0%	10.230.185	293 €
TOTAL UE27	58.213.671.248,00 €	100,0%	77,3%	446.446.444	130 €
Royaume-Uni	17.060.575.778,00 €	29,3%	22,7%	66.647.112	256 €
TOTAL UE28	75.274.247.026,00 €	129,3%	100,0%	513.093.556	147 €

Annexe 2 - Récapitulatif de la combinaison de la chaîne de valeur de Porter et du PPVC

Étape du PPVC	Activité(s) de Porter
Hypothèse	- Recherches et Développement - Marketing
Récolte	- Dépend de l'hypothèse formulée, tous les activités peuvent être concernées
Analyse	- Recherches et Développement - Analystes présents dans les différentes activités
Interprétation	- Les responsables des activités concernées par l'hypothèse formulée
Information	- Infrastructure de l'entreprise
Décision	- Infrastructure de l'entreprise - Les activités directement concernées par l'hypothèse formulée à la base
Action	- Les activités directement concernées par l'hypothèse formulée à la base

Annexe 3 - Version originale du PPCV (Neely & Jarrar, 2004)



Annexe 4 - Guide d'entretien avec EMAsphere

Introduction

Bonjour, merci d'avoir accepté de m'accorder cet entretien. Il s'agit d'un entretien semi-directif. Cela veut dire que la majorité des questions seront ouvertes et que vous êtes libre dans vos réponses. Il y a toutefois un fil rouge que je tenterai de garder durant l'entretien.

Avant toute chose, permettez-moi de préciser que notre mémoire porte sur l'échange de données entre entreprises. Le but du mémoire est de comprendre ce phénomène et de l'analyser. Notez que l'on a choisi la définition la plus large possible du terme « échanger ». La vente de données est donc incluse.

1. Tout d'abord, est-ce que je peux vous demander de dire qui vous êtes ; quel est le rôle de vous avez chez EMAsphere et, ensuite, de présenter EMAsphere ?
2. Pouvez-vous nous dire quand votre entreprise a décidé de commencer à partager des données avec d'autres entreprises, ce qui a déclenché ce besoin et comment il a évolué au fil du temps ?
3. Est-ce que vous considérez l'échange de données comme étant l'activité principale ? Ou plutôt une activité secondaire (de soutien) ?

Aspect technique

4. Quel(s) type(s) de données EMAsphere partage-t-elle avec d'autres entreprises ?
5. Quels sont les mécanismes utilisés pour partager des données ? (API, envoi de fichier ...) ?
6. Quel volume de données échangez-vous environ par an ?

La valeur de la donnée

Dans notre mémoire, un chapitre est consacré à définir la valeur de la donnée.

7. Autrement dit, lorsque vous achetez une base de données, comment en fixez-vous le ?

---→ J'évoquerai ensuite les 3 méthodes inscrites dans le mémoire.

Le RGPD

En 2016, l'Union Européenne a mis en place un Règlement général sur la protection des données, mieux connu sous le nom RGPD.

8. Est-ce que ce règlement a eu un impact (positif ou négatif) sur l'échange de données entre entreprises ?
 - a. Si impacts négatifs, comment les avez-vous contourné ?

Obstacle à l'échange de données

9. Sous quelles conditions EMAsphere accepte-t-elle de partager ses données avec d'autres entreprises ?
10. Selon vous, est-ce important qu'il y ait un climat de confiance entre les entreprises qui se partagent des données ? EMAsphere partage-t-elle des données avec des entreprises qu'elle ne connaît pas ou très peu ?

Au cours du mémoire, nous avons pu constater qu'il existe une multitude de formats ou encore de structures de données différentes. Il existe encore peu de standards en la matière permettant aux entreprises d'utiliser les mêmes formats et structures de données.

11. Avez-vous déjà eu des difficultés à échanger des données suite à un manque de standardisations ?

L'échange des données nécessite une infrastructure pour stocker, utiliser, transférer ou encore analyser ces données. Autrement dit, il est nécessaire d'avoir des serveurs, du personnel IT, etc.

12. Est-ce que vous avez déjà été freiné dans l'échange de données par ces coûts liés à l'infrastructure ?

Dans le mémoire, nous définissons le concept de « qualité des données ». L'expression « qualité des données » renvoie au concept d'exactitude de ces données, à sa capacité à fournir des informations exactes.

13. Vous est-il déjà arrivé d'obtenir des données d'une autre entreprise et que, en les analysant, vous vous rendiez compte que la qualité de celles-ci est inférieure à celle que vous vous attendiez ?
14. Est-ce que l'absence d'information sur la qualité des données est quelque chose qui vous incite parfois à ne pas échanger de données ?
15. Quelles sont les connaissances techniques nécessaires pour partager des données avec d'autres entreprises ?
16. Avez-vous déjà connu des difficultés à partager des données suite à un problème lié à des connaissances techniques ?
17. Avez-vous déjà évité d'échanger des données avec une autre entreprise suite à des craintes légales (notamment suite au GDPR) ?
18. Est-ce que vous pensez que le cadre politique et légal est assez clair et propice à l'échange de données ?
19. Avez-vous des craintes concernant la perte de contrôle sur vos données en les échangeant avec d'autres entreprises ? Que mettez-vous en place pour éviter ou mitiger cette perte de contrôle ?

Dernière question concernant les obstacles. Échanger des données, c'est également un risque. Quand vous échangez des données, vous devez utiliser des moyens conséquents (en achat, en temps ...) alors que le retour sur investissement inconnu. Il est en effet difficile de connaître à l'avance quels revenus vont pouvoir être générés par ces données.

20. Est-ce que cette incertitude liée aux données a déjà représenté une barrière à l'échange de données ? Par exemple quand vous demandez l'autorisation d'acheter des données auprès de supérieurs.

Facteurs clés de succès de l'échange de données

L'un de nos chapitres porte sur les facteurs clés de succès qui permettent de faire en sorte que l'échange de données entre entreprises se passe bien et est profitable pour les deux entreprises.

Basé sur votre expérience, quels pourraient être ces facteurs qui permettent de partager des données avec succès ?

---→ J'évoquerai ensuite les 6 facteurs clés de succès inscrits dans le mémoire

Conclusion

L'interview touche à sa fin. Une nouvelle fois, merci pour le temps que vous nous avez consacré. Souhaitez-vous ajouter quelque chose ?

Annexe 5 - Retranscription de l'interview de Florence Noblet et Marine Reillon – 21 Juillet 2021

Augustin Gérard - Bonjour à vous deux. Je vais commencer du coup par vous remercier encore une fois d'avoir accepté de faire cet entretien. Il s'agit donc d'un entretien semi-directif. J'ai déjà prévu toute une série de questions qui seront le fil rouge pour cet entretien, mais vous êtes totalement libre dans les réponses. Avant toute chose, je vais vous parler un petit peu du mémoire. Donc la problématique de notre mémoire, c'est comment le partage de données entre entreprises peut représenter une opportunité de croissance pour les entreprises européennes. Globalement, on a remarqué qu'il y a un vraiment un trou dans la littérature francophone concernant le partage de données entre entreprises. Or, c'est clairement un marché qui est en plein boom. On peut voir que chaque année, c'est plus ou moins 5 à 6 % de croissance que l'année précédente. On fait donc ce mémoire pour participer aux premières pierres des travaux francophones qui vont étudier ce phénomène.

Augustin Gérard - Et on s'est rendu compte, après avoir fait toute la partie théorie que ce serait bien d'un peut montrer comment notre grille d'analyse pouvait être mise en place de manière pratique. Donc c'est là que vous intervenez.

Marine Reillon - Quand tu parles du partage de données, qu'est-ce que tu inclus ?

Augustin Gérard - On a pris la définition la plus large possible. Le terme comprend la vente, le troc ou encore le partenariat. Donc on parle d'échange de données dès qu'il y a un flux de données entre 2 entreprises.

Florence Noblet - Oh par exemple, chez EMAsphere, notre plateforme se connecte avec une API sur les logiciels comptables de nos clients. Cela rentre dans le partage de données, alors ?

Augustin Gérard - Oui, exactement.

Augustin Gérard - Tout d'abord, est-ce que je peux vous demander de dire qui vous êtes ; quel est le rôle de vous avez chez EMAsphere et, ensuite, de présenter EMAsphere ?

Florence Noblet – Je m'appelle Florence Noblet. Je m'occupe donc de toute la partie digitale au département marketing. Du coup, pour les questions que tu vas poser, je ne pourrais répondre

principalement que pour le département marketing. Et donc, je vais sans doute plus parler d'échange de données personnelles, entre guillemets.

Augustin Gérard - Oui, il n'y a pas de soucis si tu ne parles que des échanges de données dans ton département. Je crois savoir qu'EMAsphere échange beaucoup de données donc ce n'est pas grave si on n'arrive pas à parler de tous ces échanges.

Marine Reillon - Et du coup, me concernant, je suis plus sur le volet offline. Donc tout ce qui est gestion des évènements, relations presse, gestion des partenariats et on va dire ça peut être aussi création de contenu à destination des commerciaux.

Marine Reillon – Pour décrire EMAsphere, on est une société qui édite une plateforme de gestion de la performance. C'est comme ça qu'on se présente. Dans les faits, ça se rapproche d'une solution BI, mais avec tout le côté pré paramétré et pré packagé qui fait que la plateforme a tout un tas d'indicateurs de tableau de bord prêt à l'emploi. Avec, en plus, une implémentation, une facilité d'utilisation qui est plus avantageuse qu'un outil BI. On s'adresse plutôt à des fonctions dites de “décideurs” comme des C-Level.

Augustin Gérard - Du coup pour rentrer un peu globalement dans le sujet : Qu'est-ce que qu'est-ce qu'il y aurait à première vue comme le type d'échange de données chez EMAsphere ? Qu'est-ce que qu'est-ce qui vous vient à l'esprit quand vous pensez à “Echange de données chez EMAsphere”?

Florence Noblet – Concernant le département marketing, les échanges de données vont concerner principalement des données de prospects. Des données personnelles du coup. Si on veut voir plus large, on peut aussi parler des logiciels qui nous permettent de connecter une application à une autre. Par exemple, on utilise Zapier qui nous permet de connecter Livestorm -un logiciel de gestion de Webinar- à notre CRM Hubspot. Et donc il y a un échange de données entre ces différents services.

Marine Reillon - Aussi, comme disait Florence plus tôt, il y a aussi tout l'aspect lié à notre plateforme EMAsphere, où nos clients nous font confiance pour nous donner leurs données. Ensuite, elles peuvent y accéder sur notre plateforme et créer des tableaux de bord avec.

Augustin Gérard - Est-ce que je peux vous demander sous quelle forme vous partagez vos données ? Est-ce que vous utilisez des API, est ce que vous envoyez des fichiers ou peut-être vous utilisez une plateforme pour échanger des données ?

Marine Reillon - Un peu de tout. Soit tu vas avoir un simple fichier Excel qui va contenir des données. Après tu vas avoir les plateformes de transfert de données comme ShareWork. Sharework, c'est une plateforme qui permet de voir entre partenaires quelles informations on a en commun.

Augustin Gérard - Tu peux me parler de Sharework ?

Marine Reillon - En fait, ils ont changé le nom maintenant, ça s'appelle Reveal

Marine Reillon – Donc, d'un côté, chaque entreprise va connecter Reveal sur son CRM. De l'autre côté, les entreprises partenaires sont en mesure de faire une demande pour révéler les données qu'on a en commun. Ainsi, les données sont confidentielles de base mais tu peux les partager avec les entreprises que tu souhaites. Tu peux ainsi faire un matching avec le CRM des entreprises partenaires afin de voir le potentiel de prospects que tu peux avoir.

Augustin Gérard - Et avec est ce que donc tu peux savoir avec quelle entreprise tu veux partager tes informations ?

Marine Reillon - Oui, c'est ça. Imagine, toi, en tant qu'entreprise, tu "dis Ah bah j'ai ce partenaire. J'aimerais savoir combien de contacts nous avons en commun." Alors je peux l'inviter sur Reveal pour qu'il accède à certaines de nos données. S'il accepte, le partenaire va connecter aussi sa base de données et va m'inviter. Ainsi, il va devoir donner son autorisation pour qu'EMAsphere puisse voir les contacts que nous avons en commun.

Augustin Gérard - Et avec enfin, quels sont les critères pour choisir avec quelle entreprise EMAsphere partage ces données ?

Marine Reillon - Bah si je parle côté partenaire, euh, on va dire qu'on va choisir des partenaires qui éditent un logiciel SaaS, mais non concurrentiel. L'idée est qu'on a généralement la même typologie de prospect. Par contre, il faut bien sûr qu'on ne se fasse pas concurrence. Ça pourrait

être un logiciel de paie ou autres, mais pas un éditeur de logiciel de gestion de la performance. Sinon il n'y a pas trop d'intérêt.

Augustin Gérard - Alors, pour cette question, je vais plutôt parler pour le département marketing parce que pour le reste ça va être un peu difficile à avoir l'info ici. Quelle est la quantité de données que vous partagez aux autres ? Dans l'unité de mesure que vous voulez.

Florence Noblet – Difficile à dire en valeur absolue. Ce qui est sûr, c'est que nous ne partageons qu'une partie de notre CRM. Par exemple, on peut ne partager que les prospects qui ont atteint un certain niveau de maturité. Pour essayer de donner un ordre de grandeur, ça tourne autour de 50 milles – 70 milles contacts.

Marine Reillon - Non mais après tu vois si on considère qu'on prend en compte toutes les données qu'on a achetées, tu peux même dire 100000.

Augustin Gérard - Florence, tu viens de dire que du coup tu faisais une petite sélection en prenant les contacts dont le processus commercial est plus abouti, tu saurais expliquer pourquoi ?

Florence Noblet – C'était la demande d'un commercial. Je pense donc que c'était dans une démarche commerciale pour identifier les prospects qui étaient un peu plus avancés dans notre cycle de vente. C'est de mémoire, je peux me tromper, mais il me semble que c'est la demande qui m'avait fait.

Augustin Gérard - Je sais pas si quelqu'un a quelque chose à rajouter par rapport à ce qu'on a déjà discuté ?

Florence Noblet et Marine Reillon – Non !

Augustin Gérard - Du coup, l'un des premiers chapitres du mémoire, il évoque en fait la difficulté qu'il existe à attribuer une valeur à une donnée et du coup un prix. Et du coup, je me demandais si vous, chez EMASphere, vous avez une méthode ou un processus pour savoir comment vous fixez un tel prix pour tels données ?

Florence Noblet – Concernant l’achat de données, on est contraint par les acteurs dans le sens où c’est eux qui fixent les prix et c’est nous qui tentons de négocier. Ça peut très fort varier d’un pays, à l’autre. Par exemple, entre la France et la Belgique, ce sont des coûts d’acquisition qui sont vraiment complètement différents. Après, on n’a pas assez de recul pour pouvoir dire nous, on est prêts à payer autant pour un contact. Je dirais que c’est en fonction de la valeur qu’apporte un client en moyenne. C’est-à-dire de l’ACV (NDLR : Annual Contract Value – Valeur Annuelle du Contrat) moyen. Puis de comparer ça avec le taux de conversion par chaque niveau. Ainsi, après, on peut dire qu’on est prêt à payer autant pour un contact dans notre base de données.

Marine Reillon - Même si tu vois aussi on nous proposerait un nouveau contact pour 20€ dans une base de données brutes, je dirais non. Les prix des données doivent être inférieurs au coût d’acquisition d’un SQL (NDLR : Sales Qualified Leads - Prospects qualifiés pour les ventes), par exemple. Tout dépend va aussi parce qu’il faut prendre en compte la l’intention d’achat.

Marine Reillon - Par exemple, quand on achète une base de données. On va dire que ça se joue à un coût de 0,30€ jusqu’à 1,50€. Dans ce cas-là, la donnée est juste une information avec des coordonnées en fait. On n’a pas plus d’informations que ça. Par contre, quand on va faire une campagne Google Ads, on est prêt à mettre plus de budgets pour obtenir un contact qui a déjà interagi avec notre campagne. La valeur de ce contact est bien plus haute car on sait qu’il y a une intention potentielle d’achat. La donnée a plus de valeur.

Florence Noblet – Aussi, on engage des étudiants qui ont pour but de récolter et nettoyer des données. Ainsi, ça nous est déjà arrivé de nous demander “Combien ça nous coûterait en étudiant pour créer cette donnée” pour établir le prix qu’on est prêt à payer pour une donnée.

Augustin Gérard - Parfait, merci ! Je ne sais pas si l’une d’entre vous a quelque chose à ajouter ?

Florence Noblet et Marine Reillon – Non.

Augustin Gérard - Ensuite, on parle quand même dans notre mémoire du RGPD qui a changé beaucoup de choses pour les entreprises en Europe. Je pense que vous êtes familier avec le RGPD ?

Florence Noblet et Marine Reillon – Oui, bien sûr.

Augustin Gérard - Du coup la question est : “Est-ce que le RGPD a eu un impact positif ou négatif sur vos échanges de données avec les autres entreprises ?

Marine Reillon - Je pense que ça impacte, mais du fait qu'on soit une entreprise B2B, on est moins impacté. Je pense qu'il faut quand même vraiment prendre en compte que le RGPD, à la base, son but est de protéger les données personnelles en tant que toi individu. Plutôt que toi en tant qu'entreprise. On est beaucoup moins contraint qu'une entreprise qui est en C2C. Enfin, par exemple, faire d'achats de base de données comme on fait en B2C, ça serait complètement impossible. Voire mal vu. On aurait, je pense, nettement plus de désinscriptions à nos emails ou on aurait, je pense, des gens qui nous poursuivraient. En B2B, si le message est légitime et que ça ne nous entrave pas dans notre manière de faire du business, on est bon.

Florence Noblet – Oui, c'est ça. En B2B, il y a un lobbying énorme pour trouver une interprétation du RGPD qui laisse la place à l'intérêt légitime. On n'envoie pas de mails à des contacts dont on n'est pas sûr qu'il s'agit d'une adresse professionnelle. On fait attention à nos fournisseurs de données. On s'assure que les données soient GDPR compliant. Ainsi, je pense que le GDPR a aussi permis de conscientiser les entreprises sur l'utilisation des données. Cela a permis de rendre les données plus propres. Nous, chez EMAsphere, on est toujours capable de dire d'où viennent les données. Par exemple, avant le GDPR, quand un contact téléchargeait un ebook, il entrait dans notre base de données et puis c'est tout. Maintenant, on a un encadré qui l'informe de l'intérêt légitime lié à l'utilisation de ses données. Il sait que s'il remplit le formulaire, il sera contacté par EMAsphere.

Marine Reillon – Ça a aussi eu un effet bénéfique oui. On s'est rendu compte qu'on ne peut pas faire n'importe quoi, que ce sont des informations importantes. Ainsi, on sait qu'il peut y avoir des conséquences assez graves à ne pas y faire attention.

Florence Noblet – Quand ça a été mis en place, ça a été impressionnant. Le RGPD a fait peur. Je me rappelle que moi, dans ma boîte mail, je ne recevais plus aucun email marketing pendant 2-3 mois. Et puis petit à petit en fait, c'est redevenu presque comme avant.

Augustin Gérard – Parfait. Nous allons maintenant entrer dans un des gros chapitres de notre mémoire. A savoir, les obstacles à l'échange de données. Autrement dit, ce qui décourage les entreprises à faire de l'échange de données. On a parlé tout à l'heure que vous ne partagez pas de données avec les concurrents. Sous quelles autres conditions EMAsphere accepte ou refuse-t-elle de partager ses données avec d'autres entreprises ?

Marine Reillon – Une des principales conditions, c'est de savoir si on aura un retour ou pas à échanger nos données. Voilà, ça nous a coûté d'avoir ces données donc on s'attend à un retour. On partage principalement dans le cadre de partenariat gagnant-gagnant.

Florence Noblet – Il y a aussi l'aspect géographique. EMAsphere étant basée en Belgique et en France, on va principalement échanger avec des entreprises basées dans ces mêmes pays. Et, concernant le département marketing, on partage aussi plus avec les entreprises qui ont les mêmes cibles que nous.

Augustin Gérard – Dans notre mémoire, nous parlons de l'importance de la confiance envers les entreprises qui partagent des données. Est-ce que pour vous, c'est important d'avoir cette confiance d'installée quand vous échangez des données avec quelqu'un ?

Marine Reillon – Oui, c'est important la confiance pour garantir un « donnant-donnant ». L'idée c'est d'être sur un même pied d'égalité pour qu'un partenaire ne se sente pas plus lésé qu'un autre. Pour un partenariat sur le long terme, la confiance, c'est indispensable.

Florence Noblet – La confiance est aussi importante concernant les fournisseurs de bases de données. Il y en a beaucoup mais on ne travaille qu'avec quelques-uns en qui on a confiance. Il y a beaucoup d'arnaques et il faut être vigilant.

Augustin Gérard - Un autre obstacle qu'on a identifié, c'est qu'il existe une multitude de formats et de structures différentes pour les données en Europe, dans le sens, c'est très peu standardisé.

Dans le sens où, parfois, la structure d'une base de données est tellement différente qu'il est difficile de pouvoir l'intégrer à vos bases de données. Avez-vous déjà rencontré des problèmes liés à l'interopérabilité ou au manque de standards ?

Florence Noblet – Alors, en marketing, non. Cependant, notre équipe IT qui se charge de connecter EMAsphere avec les systèmes informatiques des clients, c'est un problème auquel ils sont sans cesse confrontés. Les systèmes informatiques peuvent être très différents et c'est parfois délicat de les réunir dans une seule plateforme. En marketing, les données sont plutôt standardisées. Une adresse email, ça reste une adresse email. Une chose qui me vient à l'esprit, c'est parfois la nécessité de passer par une plateforme comme Zapier pour faire la passerelle entre deux systèmes différents, mais je ne sais pas si ça rentre dans le cadre de ton analyse.

Marine Reillon – Parfois, il peut quand même y avoir des difficultés de ce genre. Par exemple, un de nos fournisseurs de données nous avait transmis une base de données avec tous les noms de famille en majuscule ; ce qui n'est pas utilisable pour nous. En marketing, on n'utilise pas ce format. Donc oui, on a déjà eu des problèmes de ce genre où il faut passer du temps à nettoyer les données.

Florence Noblet – C'est aussi pour ça qu'on travaille généralement avec les mêmes partenaires. On les connaît et on sait qu'ils travaillent bien.

Augustin Gérard – Très bien ! Un autre obstacle dont on parle dans notre mémoire, il s'agit de la nécessité d'avoir toute infrastructure. En effet, cette infrastructure est nécessaire pour stocker, utiliser, transférer ou encore analyser ces données. Autrement dit, il est nécessaire d'avoir des serveurs, du personnel IT, etc. Est-ce que vous avez déjà été freiné dans l'échange de données par ces coûts liés à l'infrastructure ?

Marine Reillon – Quand même oui. Dans le cas du département marketing, on centralise nos données dans le CRM Hubspot. Pour une PME, c'est tout de même un coût très important. Il faut convaincre le management de financer cette infrastructure. Il faut aussi ajouter à ça tous les coups liés à l'entretien des bases de données. Car, après avoir acheté une base de données, il faut la maintenir à jour. On engage donc des étudiants qui gardent à jour les données. En définitive, ce n'est pas quelque chose de compliqué mais c'est une grosse organisation à mettre en place et c'est coûteux.

Florence Noblet – Quand je suis arrivé, toute cette organisation était déjà en place. Mais je sais que mon prédécesseur, ça lui a demandé beaucoup d'efforts et de temps. Donc, par le passé, il y a eu cette phase où l'infrastructure a été un challenge mais c'est moins le cas maintenant.

Augustin Gérard – Très bien ! Dans le mémoire, nous définissons le concept de « qualité des données ». L'expression « qualité des données » renvoie au concept d'exactitude de ces données, à sa capacité à fournir des informations exactes. Vous est-il déjà arrivé d'obtenir des données d'une autre entreprise et que, en les analysant, vous vous rendiez compte que la qualité des données est inférieure à celle que vous vous attendiez ?

Florence Noblet – Oui, c'est déjà arrivé. Un exemple concret, c'est quand on a acheté une base de données qui ne gérait pas les caractères spéciaux. Par exemple, « François » devenait « Francois ». Pour un département marketing, c'est presque inutilisable. Il a fallu nettoyer manuellement toute la base de données et c'était très décevant.

Marine Reillon – Si je me souviens bien, il y avait aussi eu un achat d'une base de données au Luxembourg où les adresses mails se sont avérées être des adresses génériques comme « info@entreprise.com ». C'est très pauvre en termes de qualité pour nous.

Florence Noblet – En plus, la base de données était assez incomplète. C'est pourquoi maintenant on est très clair dans nos demandes lors d'achat de bases de données. On demande aussi un échantillon pour voir si ça convient. Mais même ça, ça ne garantit rien car les échantillons ne sont généralement pas aléatoires mais sélectionnés.

Augustin Gérard – Ce dont vous me parlez actuellement, ça me fait penser à ce qu'on a appelé l'incertitude stratégique dans notre mémoire. Simplement, c'est qu'on n'est jamais certain de la qualité qu'on achète. Est-ce que vous êtes d'accord de dire que ça représente un risque d'acheter des données ?

Marine Reillon – Oui, bien sûr.

Florence Noblet – Oui, c'est d'ailleurs pour ça qu'on essaye toujours de travailler avec les mêmes acteurs en qui on a confiance.

Augustin Gérard – Pour la prochaine question, on ne va prendre en compte que votre département. Quelles sont les connaissances techniques nécessaires pour partager des données avec d'autres entreprises ?

Florence Noblet – Des compétences en Excel. Je pense que c'est important de comprendre comment fonctionnent les structures des bases de données. Par exemple, on a des objets entreprises et des objets contacts. C'est bien de comprendre comment fonctionne la relation entre les deux. Donc il faut au moins avoir les bases en matière de gestion des bases de données.

Augustin Gérard – Et est-ce que vous avez déjà eu une situation où, pour des raisons techniques, un échange de données n'a pas pu se faire ? Par exemple à cause d'une base de données ou d'une plateforme trop complexes.

Florence Noblet – Non, ça ne me parle pas trop dans le cas du département marketing car on utilise principalement des outils qu'on maîtrise très bien. En revanche, pour l'équipe IT, ils sont souvent confrontés à ce genre de problèmes, oui.

Augustin Gérard – On revient désormais un peu sur le RGPD qui peut aussi constituer une barrière à l'échange de données. Est-ce que ce règlement a eu un impact (positif ou négatif) sur l'échange de données entre entreprises ?

Marine Reillon – Oui ! Depuis le RGPD, on va beaucoup plus loin dans la vérification de l'origine des données. On a besoin d'avoir plus confiance envers notre fournisseur de données. Donc oui, ça arrive souvent qu'on refuse des offres parce qu'on n'est pas sûr que la base de données n'est pas RGPD friendly. Par exemple, dernièrement, un fournisseur est venu nous trouver avec les bases de données clients de notre concurrent. Or, on n'y croit pas. Quelle entreprise partagerait de telles informations ? Hors cyber attaque, c'est très rare d'avoir ce genre d'information. Du coup, quand tu creuses, tu te rends compte que ce n'est pas normal. Cela montre aussi la difficulté de savoir envers qui on peut avoir confiance ou non. Au final, on doit toujours croire ses partenaires sur parole.

Florence Noblet – Au début, le RGPD avait quasi gelé tout échange de données. Cependant, de nouvelles solutions, comme Reveal, nous permettent de faire de l'échange de données légalement et de manière sécurisée.

Augustin Gérard – Dernière question concernant les obstacles. Avez-vous des craintes concernant la perte de contrôle sur vos données en les échangeant avec d'autres entreprises ? Que mettez-vous en place pour éviter ou mitiger cette perte de contrôle ?

Marine Reillon – C'est une vraie question qui se pose. Nos données pourraient facilement se résumer à quelques fichiers Excel qui pourraient rapidement fuir. Néanmoins, avec Reveal, on est passé au-dessus de cette crainte et on en tire un réel bénéfice.

Florence Noblet – Concernant les mécanismes pour contrôler ça, on a les contrats. Par exemple, quand on achète des bases de données, on n'a pas envie d'acheter un contact qu'on a déjà dans notre base de données. Du coup, on doit transmettre au fournisseur une liste de nos contacts pour qu'il les exclût de la base de données qu'il va nous vendre. Dans ce cas-là, on signe un contrat avec le fournisseur de données comme quoi il n'a pas le droit d'utiliser cette liste pour autre chose que ce qui a été prévu.

Augustin Gérard – D'accord ! Merci pour vos réponses. Ça sera tout pour les obstacles à l'échange de données. On va maintenant ouvrir un autre chapitre qui est celui qui s'intéresse aux facteurs clés de succès de l'échange de données. Ce sont ces facteurs de qui permettent de faire en sorte que l'échange de données entre entreprises se passe bien et est profitable pour les deux entreprises. Basé sur votre expérience, quels pourraient être ces facteurs qui permettent de partager des données avec succès ?

Florence Noblet – Concernant le partage de données entre partenaire, l'émergence d'acteur tiers a vraiment permis d'augmenter la confiance et donc la quantité de données échangées. Arrête-moi si je me trompe Marine.

Marine Reillon – Oui, complètement. On n'échangeait pas de données avant.

Augustin Gérard – Qu’est ce qui fait que vous avez plus confiance quand vous passez par Reveal ?

Marine Reillon – Je pense que c’est simplement le fait de passer par un tiers en qui on a confiance. C’est leur métier de faire ça et donc ils savent ce qu’ils font. En plus, l’interface qu’ils ont est clean. Tu sens que ce n’est pas une arnaque. De plus, tu sais qu’on n’est pas concurrent avec eux et donc si on perd de la donnée chez Reveal, c’est moins grave que de perdre de la donnée chez un concurrent.

Augustin Gérard – Pour reprendre un point que tu as dit, est-ce que la simplicité de la plateforme est un facteur important pour toi ?

Marine Reillon – Oui, complètement ! La plateforme est intuitive donc c’est facile. Pas besoin d’être un expert pour savoir quoi faire.

Florence Noblet – Pour répondre à la question initiale, selon moi, le plus grand facteur de succès, c’est avant tout la confiance. C’est grâce à ça que ça devient possible.

Augustin Gérard – Oui, on y consacre également une grande partie du mémoire. On a remarqué que la confiance, ce sont les fondations sur lesquelles se basent l’échange de données. Pour citer les autres facteurs qu’on a identifiés, il y a la compréhension du cadre politique et légal, la création de partenariat, la compréhension la demande des données et, un dernier, l’avantage lié au fait d’être le premier de son secteur à échanger des données.

Florence Noblet – Oui, c’est vrai.

Augustin Gérard - On arrive à la fin de mes questions. Vous souhaitez ajouter quelque chose ?

Florence Noblet – Non, on est bon !

Augustin Gérard – Du coup, je tiens à vous remercier. Cette interview va vraiment nous aider à ajouter ce côté pratique dans notre mémoire. Donc merci pour le temps que vous nous avez consacré.