

Thèse Professionnelle

Valoriser et traiter des données numérisées des micro-entreprises

MANAGER EN SYSTEME D'INFORMATION
PAR ALTERNANCE
ELEMBA ADI ERIC

**FICHE DE CONFIDENTIALITE DES RAPPORTS, MEMOIRES,
THESES ET SOUTENANCES PROFESSIONNELS**

Formation/qualification préparée : Manager en système d'information par alternance
Nom-Prénom du stagiaire : ELEMBA ADI Eric
Titre du dossier professionnel : Thèse professionnelle
Date de la soutenance : ?? juin 2021
Nom de l'entreprise : Paragon ID
Nom et qualité du représentant de l'entreprise : RUZÉ Julien, Analyste Programmeur
Noms, entreprises et fonctions des membres de jury :

Nom - Prénom	Entreprise	Fonction

Mode de diffusion autorisé

(COCHER LA CASE CORRESPONDANTE)

☐ **Diffusion libre**

Le dossier est conservé en archives au CESI, il peut être librement consulté et reproduit. Il peut être utilisé par les destinataires, les études peuvent faire l'objet de publication....

☒ **Diffusion limitée au CESI**

Les membres du jury rendent leur exemplaire au stagiaire à la fin de la soutenance. Le stagiaire est responsable de cette restitution. Un exemplaire est conservé en archives au CESI. Le dossier peut être consulté pour exemple ou illustration par les stagiaires des promotions suivantes mais il ne peut être ni sorti du CESI, ni reproduit, sauf autorisation expresse de l'auteur et de son entreprise. La mention « Diffusion limitée au CESI, reproduction interdite » doit figurer sur la page de garde.

☐ **Diffusion interdite**

Les membres du jury rendent leur exemplaire au stagiaire à la fin de la soutenance. Le stagiaire est responsable de cette restitution. Un exemplaire est conservé au CESI, à titre de preuve dans le dossier pédagogique du stagiaire. Le dossier ne peut être ni consulté, ni sorti du CESI, ni reproduit, sauf autorisation expresse de l'auteur et de son entreprise. La mention « Diffusion et reproduction interdites » doit figurer sur la page de garde.

SIGNATURES :

Pour l'entreprise :

Le stagiaire :

Le CESI :



Remerciements

En premier lieu, je tiens à remercier l'ensemble de l'équipe pédagogique du Campus d'enseignement supérieur et de formation professionnelle de Châteauroux, où se déroule la formation de manager en système d'information par l'apprentissage, pour leur appui et leurs précieux conseils durant cette troisième année de formation.

Je remercie également le responsable du service informatique, au sein de l'entreprise Paragon Id M. FEVRE Gilles, mon tuteur M. RUZE Julien, mon chef d'équipe projet M. Blon Heu et l'ensemble du service, pour leur soutien, confiance et disponibilité.

Résumé

Je suis actuellement apprenti développeur d'applications logicielles au sein de Paragon Identification, un site de production de tickets et de billets. Mon entreprise s'est spécialisée dans le développement et la fabrication de produits d'identification. Elle mène différentes études pour une gamme de produits dématérialisés.

Dans ce contexte, j'effectue donc régulièrement des veilles technologiques sur les progrès des technologies de l'information et de la communication. Parmi les nombreux thèmes, certains révèlent des tendances que je ne soupçonnais pas.

En effet, on observe dans la croissance des activités qu'une nouvelle économie numérique est apparue. Elle rassemble différents domaines, selon les participants et le public. L'activité est basée sur les technologies de l'information, les communications, la production et les services numériques.

Dans le cadre de ces activités, des organisations revendent les données qu'elles génèrent ou recueillent et détiennent à des tiers. La monétisation des données fait partie d'un processus de développement de données plus large, et pour celui-ci, à mesure que la quantité de données collectées augmente, les défis augmentent considérablement.

La forme de monétisation des données la plus connue et la plus courante est probablement liée à la monétisation des données personnelles. Il existe de nombreuses autres formes de monétisation des données à des fins de marketing.

J'ai par conséquent choisi ce sujet de thèse professionnelle en rapport avec la valorisation des données numériques collectées ou produites.

À mon avis, ce thème est très approprié pour une réflexion commerciale, des connaissances sur les nouvelles technologies et le contexte économique.

Abstract

Currently I am an apprentice software application developer at Paragon Identification, a ticket and ticket production site. My company specializes in the development and manufacture of identification products. She conducts various studies for a range of dematerialized products.

In this context, I regularly carry out technological watches on the progress of information and communication technologies. Among the many themes, some then reveal trends that I never suspected.

Indeed, we can see in the growth of activities that a new digital economy has emerged. It brings together different areas, depending on the participants and the audience. The activity is based on information technologies, communications, digital production and services.

As part of these activities, organizations resell the data generate or collect and hold to third parties. Monetizing data is part of a larger data development process, and for it, as the amount of data collected increases, the challenges increase further.

Probably the most well-known and common form of data monetization relates to personal data monetization. There are many other forms of data monetization for marketing purposes.

I therefore chose this subject for my professional thesis related to the valuation of digital data collected or generated.

In my opinion, this theme is very suitable for business thinking, knowledge about new technologies and the economic context.

Table des matières

1	INTRODUCTION.....	11
1.1	ORGANISATION DE LA THESE	11
2	LA VALORISATION DE DONNÉE NUMÉRISÉE DES MCRO-ENTRPRISES.	12
2.1	LES REVOLUTIONS DIGITALES	12
2.1.1	<i>La transformation digitale</i>	12
2.1.2	<i>Big Data</i>	15
2.2	LA DONNÉE DANS L'ECONOMIE.....	17
2.2.1	<i>Le produit commercial « Donnée »</i>	17
2.2.2	<i>La valeur d'échange de la donnée</i>	18
2.2.3	<i>Les données personnelles</i>	19
2.3	LES REGLEMENTATIONS DE LA DONNÉE.....	21
2.3.1	<i>Régime juridique de la donnée</i>	21
2.3.2	<i>Le règlement général sur la protection des données en Europe</i>	22
2.4	LES INNOVATIONS LIEES AUX DONNÉES	23
2.4.1	<i>La technologie Big Data</i>	23
2.5	LA SECURITE DES DONNÉES.....	25
2.5.1	<i>Le contexte actuel</i>	25
2.5.2	<i>La sécurité informatique</i>	26
2.5.3	<i>Les contrôles de sécurité des données numérisées</i>	26
2.6	LES ENTREPRISES CONCERNEES	29
2.6.1	<i>Les micro-entreprises</i>	29
2.6.2	<i>La Société par Actions Simplifiée</i>	29
2.6.3	<i>Le Petite ou Moyenne Entreprise</i>	29
2.6.4	<i>Leurs secteurs d'activités</i>	30
3	L'APPROCHE EMPIRIQUE DU NIVEAU DE MATURATION DES ENTREPRISES EN MATIÈRE DE VALORISATION DES DONNÉES.....	31
3.1	METHODOLOGIE.....	31
3.2	PROBLEMATIQUE.....	31
3.3	ENQUETE DE TERRAIN	32
3.3.1	<i>Les restrictions de l'enquête</i>	32
3.3.2	<i>L'échantillon ciblé</i>	32
3.3.3	<i>Choix de la méthode d'enquête</i>	32
3.3.4	<i>La méthode « Data Value Modeling – DVM »</i>	33
4	L'ANALYSE DES ENTRETIENS.....	39
4.1	DEVELOPPEMENT GLOBALE	39
4.2	LA STRATEGIE DES ENTREPRISES.....	41
4.2.1	<i>La « culture donnée »</i>	41
4.3	LES CONNAISSANCE CLIENTS	42

4.3.1	<i>Développer une vision complète des clients</i>	42
4.3.2	<i>L'Intégration en temps réel</i>	43
4.4	PARTAGE DES VALEURS	44
4.4.1	<i>Partager la valeur</i>	44
4.5	LA TRANSVERSALITE.....	45
4.5.1	<i>L'organisation et responsabiliser les participants</i>	45
4.6	L'ETHIQUE ET CONFORMITE	46
4.6.1	<i>L'éthique</i>	46
4.6.2	<i>Le cycle de vie des données pour l'innovation</i>	47
4.7	LES OUTILS ET SOLUTIONS	50
4.7.1	<i>Les outils</i>	50
4.8	CONCLUSION	53
5	PARAGON ID : LA MATURATION DES DONNÉES	55
5.1	PARAGON ID	55
5.1.1	<i>Vision</i>	55
5.1.2	<i>Mission</i>	55
5.1.3	<i>Stratégie</i>	56
5.1.4	<i>Valeurs clé</i>	56
5.1.5	<i>Les référentielle</i>	56
5.1.6	<i>Les chiffre clés</i>	57
5.2	POLITIQUE DE GOUVERNANCE DES DONNEES	58
5.2.1	<i>La donnée au centre de la transformation digitale</i>	58
5.2.2	<i>L'organisation et responsabiliser</i>	59
5.2.3	<i>Le cycle de vie des données pour l'innovation</i>	61
5.2.4	<i>L'éthique</i>	62
5.2.5	<i>Les outils</i>	62
5.3	LA CONNAISSANCE CLIENT.....	63
5.3.1	<i>La donnée autour de l'expérience client</i>	63
5.3.2	<i>Les objectifs de Paragon Id</i>	63
5.4	LA VALEUR AJOUTEE ET LE PARTAGE DES DONNEES	64
5.4.1	<i>Partager la valeur</i>	64
5.4.2	<i>Partage en interne</i>	64
5.5	CONCLUSION	65
6	CONCLUSION GÉNÉRALE	66
7	ANNEXE	67
7.1	QUESTIONNAIRE.....	67
8	BIBLIOGRAPHIE	76
9	TRAVAUX CITES	78

Table des illustrations

FIGURE 1 - PRINCIPALES DATES DE LA GENESE AMAZON (LARGERON, 2016).....	12
FIGURE 2 - DEMONSTRATION DU SCAN D'UN MEMBRE AMPUTE (DEVILLARD, 2018).....	13
FIGURE 3 – EXEMPLE DES TECHNOLOGIES UTILISEES PAR UNE ENTREPRISE SPECIALISEE DANS LA COLLECTE DE DONNEES FULLRAMA (FULLRAMA, 2020)	15
FIGURE 4 - LOGO DU LOGICIEL D'INTEGRATION DE DONNEES "TALEND" (WIKIPEDIA, 2020)	16
FIGURE 5 - VOITURE ELECTRIQUE TESLA MODEL 3 EQUIPEE DU SYSTEME AUTOPILOT (BREEZCAR, 2017)	19
FIGURE 6 - INFOGRAPHIE SUR LES CHIFFRES IMPORTANTS DES CYBERATTQUES EN 2020 (SUDRE, 2020)	25
FIGURE 7 - CREATION D'ENTREPRISES PAR ACTIVITE EN 2020 (L'INSEE, 2021)	30
FIGURE 8 - METHODOLOGIE D'INVESTIGATION.....	31
FIGURE 9 - LES SIX AXES DE RAISONNEMENT AUTOUR DE LA VALORISATION DE LA DONNEE.	33
FIGURE 10 - CYCLE DE VIE DES DONNEES.....	49

Aucune entrée de table d'illustration n'a été trouvée.

1. INTRODUCTION

1.1 Organisation de la thèse

Cette thèse aborde les méthodes qui permettent aux petites et moyennes structures de valoriser et traiter leurs données, en tenant compte de l'aspect technique et financier des technologies et services disponibles.

Celle-ci comprendra quatre chapitres retraçant la transformation digitale présentée à la section suivante via la problématique, l'étude de terrain, l'analyse et la transposition du sujet sur mon entreprise.

Le premier chapitre traite donc de la transformation numérique. Dans un premier temps, seuls les processus qui permettent aux entreprises d'intégrer toutes les technologies numériques disponibles au sein de leurs activités sont pris en compte. Ces technologies visent à améliorer de nombreuses choses au sein des entreprises et leur permettent de poursuivre leur croissance. Il est nécessaire de connaître les facteurs qui motivent cette mutation.

Le second chapitre met en lumière l'enquête de terrain, en élargissant le champ de l'étude aux micro-entreprises sensibles aux données précieuses. Tout d'abord, nous reviendrons sur le plan d'action et le cadre de l'enquête. Puis, nous examinerons les critères de d'évaluation.

Le troisième chapitre se consacre à l'analyse des résultats de la valorisation de la donnée des micro-entreprises, autour de trois axes : le type de la gouvernance à adopter, les connaissances à acquérir et la valeur ajoutée.

Enfin, cette partie fera la transposition des travaux avec la société Paragon Id autour des mêmes questions.

2. LA VALORISATION DE LA DONNÉE NUMÉRISÉE EN MICRO-ENTREPRISES.

2.1 Les révolutions digitales

2.1.1 La transformation digitale

L'émergence d'Internet a commencé à renverser la tendance dans les années 1990, principalement en termes de distribution, car la plupart des projets visaient à créer un site Web et à développer les ventes en ligne, à commencer par les industries les plus avancées, telle que le commerce.

Pour exemple, l'activité initiale d'Amazon impliquait la vente à distance de livres, puis la société a commencé à se diversifier dans la vente de produits culturels, puis de marchands. Aujourd'hui, certains aliments peuvent également être commandés via Amazon.



Figure 1 - Principales dates de la Genèse Amazon (Largeron, 2016)

De nos jours, l'arrivée des appareils mobiles, des objets connectés, de l'universalité du réseau, de la prolifération des capteurs et des téraoctets qu'ils produisent, nous fait souvent imaginer des services quasi illimités en peu de temps.

La diffusion de la technologie et son utilisation affectent au moins autant les particuliers que les professionnels. Sa démocratisation pour un usage privé a même eu lieu avant le monde du travail, ce qui a fait que cette révolution technologique affecte tous les aspects de notre vie quotidienne depuis le début.

Par contre, pour les entreprises, la nature de cette transformation doit les pousser à reconsidérer leur compréhension du changement. Étant donné que la numérisation est une industrie qui se concentre principalement sur le service à la clientèle et l'expérience du consommateur, les outils numériques ont été directement intégrés dans l'entreprise à partir du domaine privé et ont des utilisations spécifiques.

La révolution numérique signifie également le changement de modèle dans le monde du travail. Elle marque l'arrivée des méthodes de conception, de production et de collaboration de l'entreprise, qui sont aussi des modes de réflexion, de travail et d'organisation.

Dans ce cas, nous devons distinguer les résultats de cette révolution numérique, qui conduiront à la fusion de plusieurs innovations majeures :

- ◆ La numérisation des données et des communications
- ◆ L'intelligence artificielle,
- ◆ La cybernétique
- ◆ La biosciences

Ces innovations ont conduit à la conception de machines de plus en plus complexes et « intelligentes », tels que des robots, des ordinateurs, des tablettes, des smartphones, des imprimantes, des montres, des voitures et des objets connectés. Désormais, elles permettent de connecter à distance des espaces physiques grâce à des échanges automatiques (maisons, villes, usines interconnectées).

Enfin, elles conduisent à la spécification de la conception d'outils ou de fonctions, assurées par la programmation et l'automatisation du travail et de la formation (moteurs de recherche, wifi, GPS, logiciels, applications mobiles, apprentissage numérique, etc.).



Un article Science et Avenir parle d'un projet de l'ONG Handicapé international, qui teste l'impression 3D de prothèse sur mesure.

Ce projet mené à Madagascar, au Togo et en Syrie, a pour but de scanner des membres amputés sur place pour créer des modèles 3D et fabriquer à distance des fentes prothétiques.

Figure 2 - Démonstration du scan d'un membre amputé (Devillard, 2018)

La transformation numérique couvre presque tous les domaines. Elle modifie et améliore le flux de travail, par aucune série d'opérations effectuées par des individus, des départements ou des organisations entières.

La numérisation des processus permet de simplifier, de communiquer et d'optimiser la productivité. Cette transformation permet donc aux entreprises de maintenir une compétitivité et des capacités innovantes dans un marché concurrentiel en perpétuelle évolution.

En matière d'industrie, les efforts de modernisation et de numérisation des plus grandes entreprises sous le label "Industrie 4.0" témoignent de l'intention générale de mettre en œuvre une nouvelle façon d'organiser les matériaux de production.

Par exemple, Airbus Group, Bosch, Dassault Systèmes, Schneider Electric et d'autres sociétés sont toutes engagées dans le développement des usines dites intelligentes.

"Nous menons la transformation de notre appareil industriel à marche forcée". (Vincent, 2015)

Cette citation du directeur industriel du Groupe PSA Yann Vincent, dans un article de L'Usine Nouvelle, résume bien cette tendance.

En raison du développement de l'Internet des objets et des systèmes cyber physiques, ces usines numériques assureront la synchronisation automatique de divers outils et postes de travail intégrés.

Les derniers développements technologiques dans les objets connectés, la réalité augmentée, la robotique, l'imagerie, l'intelligence artificielle et même les communications affecteront nos pratiques.

Soins de santé électroniques, médecine mobile, télémédecine, CRM et e-learning ces mots doivent non seulement faire partie de notre vocabulaire, mais doivent également nous être familiers pour être apprivoiser.

2.1.2 Big Data

Il n'y a pas de définition stricte du terme big data (Schonberger, 2013). Initialement, le concept indiquait qu'en raison de l'énorme quantité d'informations, la capacité de stockage de l'ordinateur qui traitait les informations était dépassée et les ingénieurs devaient modifier les outils utilisés pour l'analyse globale.

Les objets et applications connectés, ayant considérablement augmenté, sont le canal privilégié pour la collecte de données et le suivi des activités humaines. Ils sont également devenus une véritable industrie.

Tous les domaines d'activité ont retenu l'attention de cette matière première que sont les « données ». La valeur ajoutée créée est basée sur la capacité d'analyser ces données pour les utiliser.

C'est l'origine de nouvelles technologies de traitement des données, telles que Google MapReduce et son équivalent open source Hadoop, qui ont été créées sous l'ère Yahoo. Elles permettent de gérer une plus grande quantité de données qu'auparavant, et il n'est pas nécessaire de stocker ces données classiques.



Figure 3 – Exemple des technologies utilisées par une entreprise spécialisée dans la collecte de données Fullrama (Fullrama, 2020)

Aujourd'hui, ces outils permettent un traitement des données en temps réel et à grande échelle des « Big Data ». Maintenant, nous pouvons interpréter les données pour prédire la probabilité d'un comportement futur.



Figure 4 - Logo du logiciel d'intégration de données "Talend" (Wikipedia, 2020)

Dans un avenir proche, l'intelligence artificielle développée par Google, Facebook et de nombreuses startups offrira plus de possibilités pour l'analyse et l'amélioration des données. L'utilisation intelligente de l'information à des fins de prévision est devenue le véritable problème économique et de compétitivité d'une entreprise.

L'enjeu pour les entreprises est donc d'anticiper les évolutions de leur travail afin de faire face aux risques qu'elles engendrent, d'en saisir le potentiel et de pouvoir initier d'autres changements, de manière à atteindre des performances économiques accompagnées d'une amélioration de la qualité du travail et de la vie.

2.2 La donnée dans l'économie

2.2.1 Le produit commercial « Donnée »

Avec la numérisation du monde, ses données continuent de croître dans la mesure où elles concernent désormais presque tous les aspects des activités humaines. L'explosion des données a créé une nouvelle économie dans laquelle les acteurs économiques maîtrisent ce nouveau matériau spécial, c'est-à-dire la génération de plus en plus de données pour les services numériques. (Raymond, 2020)

Par conséquent, les données semblent être une nouvelle marchandise au XXI^e siècle. Sa production, sa capture, sa possession et son exploitation sont considérées comme la source de création de valeur. De nombreuses personnes considèrent donc les données comme des produits échangeables et une source de richesse.

En outre, les entreprises offrent aux particuliers la possibilité de commercialiser leurs données personnelles, comme Datacoup¹, ou d'autres sociétés organisent des marchés de données (Dawex²), et les entreprises peuvent vendre des ensembles de données. D'autres ont tenté de monétiser les données des consommateurs directement auprès des annonceurs (Qualifo³).

Cependant, les caractéristiques des données sont loin d'être des marchandises. En effet, un examen des caractéristiques des données nous montre que la production de données à l'ère numérique est généralement un mécanisme complexe qui la sépare du concept de marchandises (Sobel, 2010) et sa valeur d'usage est plus compliquée à définir par rapport aux marchandises standards.

En plus, les données sont soumises à différentes lois et réglementations, en particulier en Europe, où plusieurs règles juridiques, notamment le règlement général sur la protection des données (CNIL, 2018) les éloignent pour l'instant du concept de biens.

¹ <https://datacoup.com>

² <https://dawex.com>

³ <https://qualifio.com>

2.2.2 La valeur d'échange de la donnée

Dans l'espace numérique, de nombreux produits sont fabriqués sans objectif de vente car ils sont produits à l'insu de la personne qui les a conçus. Par conséquent, les données obtenues par les internautes effectuant des recherches dans les moteurs de recherche de sites commerciaux ne constituent pas des données des auteurs de ces recherches.

D'autre part, le traitement secondaire de la recherche, en fait des données utiles qui peuvent être utilisées pour afficher les intentions des utilisateurs sur le site Web. Donc la valeur est générée après l'échange, et la valeur d'utilisation n'est pas définie dans l'échange ou par son producteur.

Le système de cookies sur internet est un parfait exemple de récolte et traitement de données. A l'origine, les cookies ont été développés comme un service qui fournit de la mémoire aux sites Web, dans le but de simplifier l'interaction du site avec les utilisateurs et d'améliorer l'expérience utilisateur en le rendant plus intuitif. (Cookiebot, 2020)

Les entreprises s'appuient de plus en plus sur l'échange de données pour créer de la valeur et répondre à la demande croissante de transformation numérique.

Selon le consultant IT, Laurent Lafaye nous observons deux grandes tendances dans l'échange mondial de données :

- ◆ La première observation est que l'échange de données est une activité mondiale et plus de 50% des organisations viennent de nationalités différentes. 51% de l'interaction se fait entre les pays du G7. La France, les États-Unis, l'Allemagne et le Royaume-Uni sont les plus actifs.
- ◆ Le deuxième constat est que 70% de ces échanges se font entre entreprises de différents départements. Les industries les plus actives sont les services financiers (16%), le commerce, la distribution, le commerce électronique (14%) et le marketing, la publicité et la communication (12%). L'industrie du commerce de détail est également l'industrie à la croissance la plus rapide, avec une croissance de plus de 30% en un an. (Lafaye, 2019),

Les données sont au cœur du développement économique des entreprises. La réalisation de la valeur et du potentiel des données est un défi majeur. La quantité de données augmente et crée de la valeur. Cependant, il n'est pas facile pour une entreprise de répertorier les données à réutiliser pour améliorer ses performances. Le défi est plus culturel que technique.

Pour l'avenir, les défis se concrétiseront autour de la mise en œuvre de stratégies d'échange de données, ce qui apportera à l'entreprise un avantage concurrentiel. Faciliter et assurer l'échange de données entre les acteurs de différents secteurs est essentiel pour le développement du marché de l'échange de données.

2.2.3 Les données personnelles

Bien que les données personnelles soient au cœur du règlement général européen sur la protection des données (CNIL, s.d.), de nombreuses personnes ne connaissent toujours pas sa véritable signification. Il n'y a pas de liste claire des données privées, mais nous devons noter que les données personnelles sont divisées en deux catégories :

- Les données de comportement sont des données liées à tout comportement personnel collecté en surveillant leur navigation, mettant en évidence leur comportement d'achat. Cela permet par exemple aux détaillants, d'affiner les données personnelles d'une personne, en particulier dans ces données de comportement croisées avec l'identification des données personnelles. En comprenant mieux nos intérêts personnels et nos loisirs, elles permettent d'optimiser leur processus d'achat et de leur fournir des services personnalisés.
- Les données personnelles proviennent d'individus sur Internet : données déclaratives, formulaires, inscriptions en ligne... mais l'entrée des données personnelles dépasse largement la portée déclarative.

Les gens ne sont plus les seuls producteurs de données, car les machines génèrent également des données via différents types de capteurs ou d'objets connectés. En raison de la technologie d'exploration de données, ces capteurs peuvent parfois afficher plus d'informations personnelles, qui peuvent être liées aux données d'une manière informelle et apprendre de nouvelles informations.

Par exemple, le model 3 de la Tesla est une voiture autonome dotée d'un système machine learning⁴ de conduite semi-autonome "Autopilot".



Figure 5 - Voiture électrique Tesla Model 3 équipée du système Autopilot (Breezcar, 2017)

⁴ L'apprentissage automatique (machine learning) peut être défini comme une technologie d'intelligence artificielle qui permet l'apprentissage automatique sans avoir besoin d'une programmation spéciale à l'avance.

Cependant, nous devons également souligner que la prolifération des cyberattaques a conduit à la divulgation d'informations confidentielles, ce qui entraîne des problèmes de sécurité des données, en particulier ces problèmes de stockage et d'accès.

Il existe plusieurs technologies en cours de développement qui complexifient la génération de données personnelles. Ces technologies ont considérablement enrichi l'essence des données personnelles. Ces caractéristiques ne sont plus seulement liées à l'identité officielle de l'individu, mais permettent de développer nos habitudes (navigation, achat, etc.), nos parcours d'achat en magasin ou en ligne, etc.

D'autre part, pour les auteurs, la manière dont les données sont générées n'est pas toujours consciente ou volontaire, mais en raison de leur requête sur le moteur de recherche, navigation ou activation de géolocalisation généreront autant de données que possible. En particulier, ces données peuvent être associées à des identités physiques via des adresses IP.

2.3 Les réglementations de la donnée

2.3.1 Régime juridique de la donnée

Les données sont principalement soumises à trois systèmes juridiques :

- Le système de données personnelles,
- Les données ouvertes
- Les formulaires contractuels classiques pour gérer l'échange de données personnelles et publiques.

Les deux premiers systèmes juridiques sont examinés sur une base juridique très différente. Les données ne sont pas une marchandise, les données personnelles sont une question de droits de l'homme et les données publiques ouvertes devraient être soumises à un système juridique spécifique qui ne permet pas de les considérer comme des marchandises. C'est un bien public que tout le monde peut utiliser.

Les données personnelles des individus ne sont pas considérées comme des biens mais comme des droits appartenant aux personnes. C'est ainsi que la loi sur la protection des données de 1978 (CNIL, 2019) et le règlement européen sur la protection des données de 2016 (CNIL, 2018) les ont conçus.

Les informations elles-mêmes peuvent échapper aux droits d'auteur et aux droits de propriété intellectuelle. Elles appartiennent par défaut au domaine public, même si la création de bases de données en Europe depuis les années 90 tend à leur obéir progressivement avec une logique propriétaire.

Depuis, elles n'ont cessé de se développer et ont continué à élargir le champ des bases de données commerciales, scientifiques ou économiques. C'est particulièrement vrai. Nous devons noter que les données obtenues isolément n'ont quasiment pas de valeur et ne sont pas considérées comme une marchandise.

En bref, il n'y a pas de droit à la propriété des données. Cependant, il est nécessaire d'établir un droit qui puisse contrôler efficacement ses données et déterminer son droit d'utilisation.

Quant au traitement légal des données personnelles, il repose essentiellement sur l'autorité de la Cnil, et la législation sur les données personnelles repose sur le responsable du traitement. Les citoyens européens doivent conserver le fichier pour en rester propriétaire.

2.3.2 Le règlement général sur la protection des données en Europe

Depuis le 25 Mai 2018, toutes les entreprises qui gèrent des données à caractère personnel sont soumises à des obligations réglementées. La directive européenne, dit le RGPD, oblige toutes les entreprises et les administrations, à respecter certaines règles concernant le traitement des données à caractère personnel.

L'objectif de ce règlement est de protéger les citoyens européens contre les usages malveillants de leurs données personnelles.

Par exemple :

- Un nom, un prénom, une photo, sont des données personnelles que l'on trouve sur des badges d'accès au sein de votre entreprise.
- Un numéro de téléphone via un annuaire pour retrouver les informations sur une personne.
- Les dossiers de candidature (un CV et/ou lettre de motivation) qui sont des données facilement accessibles si elles ne sont pas archivées dans les entreprises.
- Le numéro de sécurité social peut donner accès à des données médicales.
- Les images de vidéosurveillance implantées dans les entreprises sont considérées comme des données personnelles.
- Le numéro d'immatriculation permet d'obtenir le nom et adresse du titulaire de la carte grise.
- L'adresse IP enregistrée sur les serveurs pour naviguer sur les différents site web, permet de retrouver l'adresse de l'utilisateur.
- etc.

Selon cette directive, l'autorité responsable du traitement des données au sein d'une entreprise est le chef d'entreprise. Il doit appliquer une politique de sécurité informatique conforme aux règlements.

Dans le cas d'un manquement ou de négligence de la part des entreprises, le règlement général sur la protection des données prévoit, une amende administrative de 4 % de votre chiffre d'affaires annuel mondial, plus l'amende au tribunal pénal, plus les dommages et intérêts à la partie civile si une plainte a été déposée.

2.4 Les innovations liées aux données

2.4.1 La technologie Big Data

Les matériels de base actuellement les moins chers du marché permettent de développer des solutions conçues pour analyser et agréger immédiatement les données PB (1 PB⁵ = 1 000 To⁶). Bien que la plupart des entreprises soient aujourd'hui encore loin d'avoir des Pétaoctet de données, toute entreprise dont la capacité de stockage dépasse les téraoctets peut bénéficier de la technologie dit « Big Data ».

La technologie Big Data permettra également aux entreprises de réduire considérablement le temps de réponse pour le traitement des demandes ou l'exécution de tâches analytiques complexes dans des intervalles de temps plus courts.

Les nouvelles technologies pour stocker et accélérer le traitement de grandes quantités de données peuvent être divisées en trois catégories :

Accélération matérielle.

Ces périphériques matériels pour les bases de données utilisent une DRAM⁷ dynamique ou une mémoire flash au lieu de disques durs pour améliorer les performances d'accès en lecture / écriture, et peuvent également modifier le code de la base de données pour mieux utiliser les processeurs multicœurs et d'autres avancées technologiques.

Base de données d'analyse massivement parallèle (MPP).

Ces bases de données conformes à SQL sont conçues pour distribuer le traitement des données sur plusieurs machines, communément appelées « serveurs de base ». Certaines bases de données MPP utilisent une architecture de sans partage, qui distribue également le stockage de données sur plusieurs ordinateurs, éliminant ainsi les goulots d'étranglement de l'évolutivité.

⁵ Pétaoctet

⁶ Téraoctets

⁷ La mémoire vive dynamique (en anglais DRAM pour Dynamic Random Access Memory) est un type de mémoire vive compacte et peu dispendieuse.

Modèle Map-Reduce, Hadoop et autres méthodes NoSQL.

Ces méthodes permettent d'accéder aux données via des langages de programmation sans utiliser d'interfaces SQL et permettent de distribuer les données sur plusieurs ordinateurs distincts. La structure de données est un type de fichier plat ou un autre format non relationnel.

La méthode NoSQL est particulièrement importante pour l'analyse des tâches qui ne peuvent pas être efficacement codées en SQL. Cela comprend une analyse impliquant plusieurs transferts de données et des cas où les données sont non structurées ou semi-structurées (par exemple, une analyse de texte).

Map-Reduce et Hadoop sont les modèles NoSQL les plus populaires. Map-Reduce est un algorithme qui peut distribuer des tâches d'analyse de données sur plusieurs nœuds, puis "réduire" tous les résultats des nœuds dans un ensemble. Hadoop est une implémentation open source de Map-Reduce, avec d'autres fonctions d'Apache.

Grâce à ces innovations technologiques du Big Data, les entreprises peuvent désormais effectuer des analyses, ce qui n'était pas possible jusqu'à présent. En effet, trop de données doivent être traitées, l'analyse prend trop de temps ou parce que le traitement requis ne s'applique pas à SQL.

Quelle que soit la technologie de Big Data utilisée pour la tâche d'analyse, les utilisateurs professionnels doivent être capables de visualiser et de comprendre les résultats. Étant donné le nombre potentiellement élevé de résultats, les utilisateurs ont besoin d'outils de reporting et d'analyse spécialisés qui sont conçus pour trouver des données spécifiques dans une masse d'informations, y compris de puissants outils de visualisation et d'exploration, ainsi qu'un filtrage et un tri dynamique.

Ces fonctions de génération de rapports d'analyse et d'interface utilisateur doivent être capables de gérer tout type de format de données et de fonctionner en continu avec les SGBDR⁸ traditionnels, les bases de données MPP ou les fichiers distribués sur le réseau.

⁸ En [informatique](#), une **base de données relationnelle** est une [base de données](#) où l'information est organisée dans des tableaux à deux dimensions appelés des *relations* ou *tables*. (Wikipedia, 2017)

2.5 La sécurité des données

2.5.1 Le contexte actuel

Les entreprises doivent se prémunir des menaces susceptibles de ralentir leurs activités. Les ressources et les données les plus précieuses sont les cibles de ces attaques. Les infrastructures informatiques sont les points d'accès les plus favorables aux intrusions et erreurs pouvant mettre en péril une société.

Notamment, les logiciels de programme malveillant (virus et autres malwares) qui s'attaquent physiquement aux postes de travail, pour récupérer des informations sensibles. Également le piratage informatique qui consiste à s'introduire dans un système intermédiaire par le biais d'une brèche de sécurité au niveau du réseau ou d'un ou plusieurs postes de travail.

Actuellement une grande partie de la population mondiale s'est tournée vers le télétravail et les entreprises sont rapidement passées aux opérations à distance, les cyberattaques se sont aussi adaptées.

Selon une étude de l'entreprise Almeria, le travail à distance est devenu la source de 20 % des incidents de cyber sécurité, et les logiciels de malversation sont en augmentation.

LES CYBERATTQUES EN 2020 EN CHIFFRES

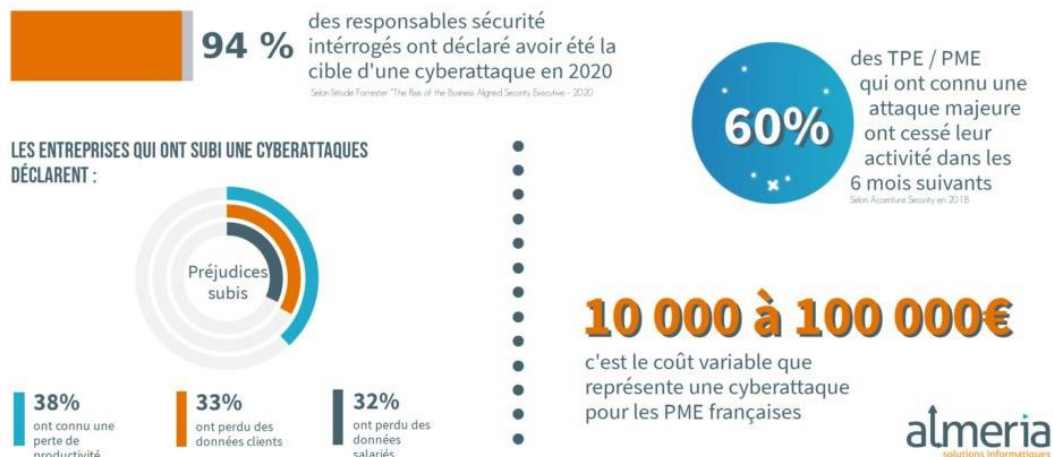


Figure 6 - Infographie sur les chiffres importants des cyberattaques en 2020 (SUDRE, 2020)

De nombreuses organisations n'ont pas encore intégré des règles suffisantes en matière de sécurité, et les vulnérabilités constituent une menace constante pour les infrastructures informatiques des entreprises.

2.5.2 La sécurité informatique

La sécurisation des traitements doit trouver un équilibre entre les droits des personnes concernées par les traitements de données et les intérêts des personnes traitant ces données.

Autrement dit, il est nécessaire de protéger les données des personnes concernées tout en tenant également compte des intérêts légitimes des responsables du traitement.

Selon le règlement du conseil du parlement européen (CNPD, 2018), les objectifs lors de la mise en œuvre d'un dispositif de sécurité informatique sont :

- Les données doivent être accessibles uniquement par les personnes autorisées.
- L'intégrité des données doit être garantie.
- Les données doivent être disponibles malgré les aléas techniques et/ou attaques du système informatique.
- Les traitements sur les données doivent être traçables et transparents.

Il faut également préciser que la sécurité des systèmes informatiques doit tenir compte des évolutions technologiques et organisationnelles d'une société afin d'assurer une sécurité optimale.

2.5.3 Les contrôles de sécurité des données numérisées

Les données créées, collectées, stockées et échangées par les entreprises sont un atout précieux. Grâce à la protection contre la corruption et l'accès non autorisé par des personnes internes ou externes, les organisations protègent leur entreprise de pertes financières, limite les risques de violation de la réputation de la marque, et de détérioration et/ou de perte de confiance des consommateurs.

Par ailleurs, la réglementation des autorités et le secteur de la sécurité des données exigent que leur entreprise respecte ces règles, quel que soit leur domaine d'activités.

Comprendre l'importance de la sécurité des données permet d'élaborer un plan pour les protéger. Il existe de nombreuses technologies et processus de sécurité des données qui peuvent contribuer à la productivité des sociétés tout en protégeant les données.

L'authentification

L'authentification est la solution la plus recommandée pour renforcer la sécurité des données et se protéger contre les violations de données. La technologie d'authentification permet de vérifier si les informations d'identification d'un utilisateur correspondent à celles stockées dans votre base de données.

Les processus d'authentification standards actuels comprennent l'utilisation d'une série de moyens pour identifier un utilisateur autorisé, tels que les mots de passe, les codes PIN, les jetons de sécurité, les cartes magnétiques ou la biométrie.

L'authentification est facilitée par la technologie unique qui, avec un seul jeton de sécurité, permet à un utilisateur identifié d'accéder à plusieurs systèmes, plateformes et applications. La technologie permettant l'autorisation détermine ce qu'un utilisateur authentifié est autorisé à faire ou à voir sur votre site Web ou votre serveur.

Contrôle d'accès

L'authentification et l'autorisation se font par un processus de contrôle d'accès. Les systèmes de contrôle d'accès peuvent notamment comprendre :

- Un contrôle d'accès facultatif qui permet d'accéder aux ressources en fonction de l'identité des utilisateurs ou des groupes,
- Un contrôle d'accès basé sur les rôles, qui attribue l'accès en fonction du rôle dans l'organisation et permet aux utilisateurs d'accéder uniquement à des informations spécifiques,
- Et un contrôle d'accès obligatoire, qui permet à l'administrateur du système de contrôler rigoureusement l'accès à toutes les informations.

Sauvegarde et récupération

La sauvegarde des données implique la copie et l'archivage des données informatiques afin de pouvoir y accéder en cas de corruption ou de perte de données. Clés USB, disques durs, stockage sur des serveurs externes, sauvegarde en ligne dans le cloud, chaque supports et systèmes de sauvegarde sont différentes et s'adapte aux besoins.

Il existe 3 types de backup :

- Full Backup (sauvegarde de l'ensemble des données) ;
- Incremental Backup (sauvegarde uniquement des données modifiées ou ajoutées depuis la dernière sauvegarde) ;
- Differential Backup (sauvegarde uniquement des données modifiées ou ajoutées depuis le dernier Full Backup ou Incremental Backup).

La récupération des données et des informations permet une reprise après un sinistre dès que possible. En fait, les données peuvent subir divers dommages pour les raisons suivantes :

- Une panne du matériel informatique ;
- Une défaillance du système d'exploitation (Windows, Linux, Mac) ;
- Une erreur de manipulation de la part d'un utilisateur ;
- Une catastrophe naturelle (tempête, inondation, incendie, ...) ;
- Une cyber-attaque ou un virus informatique.

La sauvegarde et la récupération des données consistent à anticiper ces incidents éventuels pour qu'ils n'impactent pas votre entreprise trop durement.

Chiffrement

Les systèmes de chiffrement des données renforcent efficacement la sécurité des données en utilisant un algorithme et une clé de chiffrement pour transformer un texte normal en texte chiffré. Les données chiffrées seront illisibles pour toute personne non autorisée. Ces données ne peuvent alors être déchiffrées que par un utilisateur disposant d'une clé autorisée.

Le chiffrement est utilisé pour protéger les données que vous stockez et les données échangées entre les bases de données, les appareils mobiles et le serveur distant. Les clés de chiffrement doivent être gérées de manière sûre, notamment en protégeant les systèmes de gestion essentiels, en assurant une sauvegarde sécurisée du chiffrement hors site et en limitant l'accès.

Masquage des données

Les logiciels de masquage de données cachent les données en masquant les lettres et les chiffres avec des caractères de substitution. Les données sont présentes mais derrière le masquage. Le logiciel ne restitue leur forme d'origine que lorsqu'un utilisateur autorisé reçoit ces données.

Tokénisation

La tokénisation remplace les données sensibles par des caractères aléatoires qui ne sont pas réversibles sur le plan algorithmique. Le rapport entre les données et leurs valeurs est stocké dans un tableau de consultation de base de données protégé, et non pas généré et déchiffré par un algorithme mathématique contrairement au système de chiffrement. Le jeton représentant les données réelles est utilisé dans différents systèmes pour les remplacer, tandis que les données réelles sont stockées sur une plateforme séparée et sécurisée.

Suppression

Lorsque les données numérisées ne sont plus nécessaires et doivent être définitivement effacées du système. L'effacement peut écraser ces données de manière à les rendre irrécupérables. L'effacement est différent de la suppression, processus ne consistant qu'à cacher les données de manière à en faciliter la récupération.

2.6 Les entreprises concernées

2.6.1 Les micro-entreprises

Le régime micro-entreprise est un régime fiscal est désignent les entreprises qui comptent moins de 10 associés. La responsabilité est réduite pour le gérant et ses associés.

Les entreprises doivent respecter des seuils annuels de recettes pour en bénéficier. Les recettes du régime micro-entreprise sont fixés à un seuil de :

- 176 200 euros pour les activités de vente des marchandises, d'objets, de fournitures et de denrées à emporter ou à consommer sur place, ou de fourniture de logement (hôtels, chambres d'hôtes, meublés de tourisme).
- 72 500 euros pour les autres activités de prestations de services relevant des bénéfices industriels et commerciaux⁹ et les professionnels libéraux relevant des bénéfices non commerciaux¹⁰

2.6.2 La Société par Actions Simplifiée

La Société par Actions Simplifiée (SAS) est un régime fiscal désignent les entreprises qui ont entre 2 et 9 associés. La SAS peut également ne compter qu'un seul associé, il s'agit alors d'une Société par Actions Simplifiée Unipersonnelle (SASU).

Le capital social d'une SAS peut être fixe ou variable selon les montants définis par les associés. Elle est composée d'apports d'argent et/ou d'apports de biens. Le régime SAS limite la responsabilité des associés. Concernant les dettes sociales, les associés sont responsables dans la limite de leurs apports.

2.6.3 Le Petite ou Moyenne Entreprise

La petite ou moyenne entreprise (PME) est un régime fiscal dont le nombre de salarié est compris entre 10 et 249, et le chiffre d'affaires est inférieur à 50M € ou le total du bilan ne dépasse pas 43M €.

⁹ Les Bénéfices Industriels et Commerciaux (BIC) sont une composante du revenu global soumis à l'impôt sur le revenu.

¹⁰ Les BNC, Bénéfices Non Commerciaux, correspondent aux bénéfices réalisés par les professions non commerciales et assimilées et constituent une des catégories des revenus imposables à l'impôt sur le revenu

2.6.4 Leurs secteurs d'activités

En 2020, on estime à 800 000 le nombre d'emplois supprimés, selon le Ministère de l'Économie. Cette statistique augmente le risque d'entraîner un chômage de masse mais également lancer la croissance de l'entrepreneuriat.

Malgré que la création d'entreprises ait chuté de 33% en avril 2020 selon l'article du l'INSEE (INSEE, 2020), la crise sanitaire semble stimuler les ambitions entrepreneuriales.

Créations d'entreprises par activité en 2020		
Secteur d'activité	Nombre d'entreprises créées	dont micro-entrepreneurs
Industrie	40 079	27 406
Construction	80 468	44 681
Commerce y c. réparation automobile	130 276	81 492
Transports et entreposage	101 097	88 481
Hébergement et restauration	39 083	20 003
Information et communication	42 770	27 981
Activités financières et d'assurance	21 387	3 453
Activités immobilières	35 380	15 578
Activités spécialisées, scientifiques, techniques et de soutien aux entreprises	193 757	131 641
Enseignement, santé humaine et action sociale	81 705	47 226
Autres activités de services aux ménages	82 162	60 006
Ensemble	848 164	547 948

Lecture : en 2020, 40 079 entreprises ont été créées dans le secteur de l'industrie, dont 27 406 sous le régime du micro-entrepreneur.

Figure 7 - Création d'entreprises par activité en 2020 (L'INSEE, 2021)

Une étude menée par Fondation Entreprendre (Fondation Entreprendre, 2020), démontre que 40% des sondés sont tentés par l'entrepreneuriat.

3. L'APPROCHE EMPIRIQUE DU NIVEAU DE MATURATION DES ENTREPRISES EN MATIÈRE DE VALORISATION DES DONNÉES.

3.1 Méthodologie



Figure 8 - Méthodologie d'investigation

3.2 Problématique

Les données constituent aujourd'hui un gisement de valeurs, d'opportunités et d'innovation offrant des perspectives de croissance et de compétitivité aux entreprises.

Le marché du Big data est monopolisé par les sociétés Américaines et bientôt par celles venues d'Asie. Dans ce marché déjà marqué par la présence des GAFAs (Google, Apple, Facebook et Amazon), on voit déjà l'essor des BATXI (Baidu, Alibaba, Tencent et Xiaomi), et l'Europe a toujours un train de retard.

Comment les petites et moyennes structures peuvent-elle valoriser leurs données ?

Actuellement, les projets autour de la donnée sont encore trop souvent des prouesses techniques qui ne répondent pas à des besoins métiers spécifiques ou génériques. En France, les projets de grande ampleur sur le marché restent encore un domaine réservé aux entreprises historiquement data-centric¹¹ (Chicha, 2019).

Pour initier la permutation, il faut reconsidérer la méthodologie sous l'angle de l'expérience utilisateur. Ainsi, les équipes opérationnelles pourront intégrer l'analyse et le recueil des données plus facilement et appréhenderont mieux l'intérêt du processus.

¹¹ Le data-centric est une architecture centrée sur les données. Elle comporte un composant central de type système de gestion de base de données. Des composants périphériques, appelés clients, utilisent le composant central appelé serveur de données.

La collecte et le traitement des données sont compliqués car les données sont inexploitable et nécessitent d'être restructurées pour en tirer de la valeur. Cette phase de restructuration coûte en moyenne 80% du temps et du coût d'un travail de valorisation des données

Aujourd'hui, les technologies liées à la collecte, au traitement et au stockage de données de masses ne sont plus réservées aux grandes entreprises. Désormais l'exploitation de la donnée, n'a pas besoin d'énormément de ressources.

Néanmoins une étude de terrain est nécessaire pour évaluer le niveau de conscience du capital données des entreprises.

3.3 Enquête de terrain

3.3.1 Les restrictions de l'enquête

En raison du contexte sanitaire lié à la pandémie de la COVID et les restrictions de déplacement, les entretiens directs (interviews) ont été réalisés auprès de plusieurs entreprises de la région « Centre - Val de Loire ».

3.3.2 L'échantillon ciblé

Les entreprises choisies sont des petites structures de moins de dix salariés, gérées par un ou plusieurs membres, avec un statut juridique de micro-entreprises, SAS ou PME.

3.3.3 Choix de la méthode d'enquête

Parmi les méthodes d'enquête que nous pouvons utiliser, notre choix se porte sur le questionnaire au détriment des entretiens directs. La première raison est qu'il permet la collecte de données auprès d'un grand nombre d'individus en respectant les restrictions sanitaires. La seconde, est de limiter le risque d'imprécision des sondés qui contribue à la fiabilité et à une meilleure uniformité des réponses. Les sondages semblent ainsi répondre à nos exigences

Ce sondage comprend des questions fermées afin de faciliter l'examen des réponses et de réduire le temps de réponse des personnes sondées. Les questionnaires sont réalisés selon l'approche « Data Value Modeling – DVM » (BOUNFOUR, 2018) qui définit le cadre méthodologique permettant à des structures d'établir un raisonnement autour de la question de la valorisation des données.

3.3.4 La méthode « Data Value Modeling – DVM »

Cette méthode est basée sur les travaux du Pr BOUNFOUR en rapport au capital immatériel. Elle a pour but de faire migrer les données structurées et non structurées vers la catégorisation des 3 types d'actifs (propriétaires/conjoints/non propriétaires) dans une perspective de création de valeur.

La méthode comporte 25 niveaux de pratique et 5 niveaux de maturité. Elle peut s'appliquer à tous secteurs car elle est suffisamment générique. Une des forces du modelé est son analyse comparative qu'il permet de réaliser entre entreprises pour montrer aux Directions Générales et aux directions métiers comment l'entreprise se positionne à un moment donné en matière de valorisation des données.

Les questions posées sont orientées et divisées par les 6 dimensions :



Figure 9 - Les six axes de raisonnement autour de la valorisation de la donnée.

La stratégie de l'entreprise

N° Question	Critères d'évaluation *Pratique issue du modèle DVM®
1	Des cas d'usages basés sur les données ont été déployés
2	De nouveaux business modelés ont été mis en place grâce à l'usage et l'analyse des données
3	De nouveaux des services innovants ont été mis en place grâce à l'usage et l'analyse des données
4	Les données (marketing, techniques, financières, RH ou autres) sont une composante essentielle et effective du développement des produits et services*
5	Une forme de rétribution (bitcoins par exemple) est offerte au client qui partage ses données (contrat de confiance) : contractualisation de la fidélité client
6	Des accords de partenariats sur le partage de la valeur existent (sous-traitants, co-traitants, fournisseurs...)
7	Les données sont exploitées pour éclairer les décisions

Les connaissances clients

N° Question	Critères d'évaluation *Pratique issue du modèle DVM®
8	Des actions sont mises en place pour identifier les clients/les partenaires
9	Des données sont collectées sur les attentes des clients et contribuent à personnaliser les services et l'expérience client
10	Au-delà des données clients, les données produits sont également exploitées (pour améliorer la connaissance des produits, des processus, et <i>in fine</i> la connaissance des clients)
11	Nous mobilisons les données pour simuler/prédire/anticiper les comportements des clients/écosystème*
12	Des stratégies marketing et commerciales adaptées aux besoins des Métiers sont définies à partir de la connaissance des données clients, produits et/ou des usages
13	Une surveillance des concurrents et des nouveaux entrants disrupteurs par l'analyse des données externes est en place
14	Une veille active est assurée sur les réseaux sociaux en temps réel pour anticiper, prendre des décisions...

La transversalité

N° Question	Critères d'évaluation *Pratique issue du modèle DVM®
15	La valeur de la donnée est partagée en interne
16	La connaissance des outils d'analyse des données est diffusée de façon organisée au sein de l'entreprise
17	Les données issues de capteurs et objets connectés internes sont collectées et analysées
18	Nous mobilisons les données pour offrir des combinaisons de produits/services les plus adaptés*
19	Des données sont collectées pour mesurer l'usage réel des produits et services
20	Les données sont partagées avec des concurrents/partenaires pour améliorer la productivité de l'activité
21	Les processus RH intègrent une perspective donnée, en particulier celle relative à la valorisation du capital humain de l'entreprise*

Partage des valeurs

N° Question	Critères d'évaluation *Pratique issue du modèle DVM®
22	La donnée est reconnue comme un actif clé au sein de l'entreprise et son management est organisé
23	Il existe une stratégie et des politiques de valorisation des données
24	La cohérence de la politique de gestion des données et sa traduction opérationnelle fait l'objet d'un suivi régulier au niveau de la DG*
25	L'organisation, les rôles et les responsabilités sont clairement définis en matière de gestion des données
26	Une démarche d'identification des sources de données non-structurées existe
27	Une politique d'ouverture des données est mise en œuvre à l'externe (open data/publication de données et/ou d'algorithmes), à l'écosystème
28	Des méthodes de préparation des données existent pour créer des ensembles de données cohérents

Éthique & Conformité

N° Question	Critères d'évaluation *Pratique issue du modèle DVM®
29	Des règles éthiques sur l'usage et l'exploitation des données sont définies, partagées et appliquées.
30	Un système de gestion de crise est défini et activable en cas de fuites et/ou d'usages détournés des données
31	Un processus d'anonymisation des données personnelles est en place pour permettre leur exploitation
32	Les règles de protection des données personnelles sont prises en compte dès la conception.
33	Les DPI (Droits de Propriété Intellectuelle) sont clairement identifiés et encadrés

Outils & Solutions

N° Question	Critères d'évaluation *Pratique issue du modèle DVM®
34	Des connecteurs, APIs, et autres <i>web services</i> ont été développés pour favoriser le partage des données
35	Des systèmes/outils sont mis en place pour capter les données en temps réel
36	L'interopérabilité est prise en compte, notamment avec les données de l' <i>IoT</i>
37	Une architecture et/ou des applications <i>Big Data</i> sont en place

4. L'ANALYSE DES ENTRETIENS

4.1 Développement global

Le questionnaire a été soumis à quatre-vingt-dix entreprises, une quarantaine de réponses ont été collectées. Le graphique ci-dessous est le résultat visuel de ces différentes auto-évaluations du niveau de développement en matière de valorisation des données. Il donne une première idée du niveau de maturité globale des petites ou moyennes entreprises :

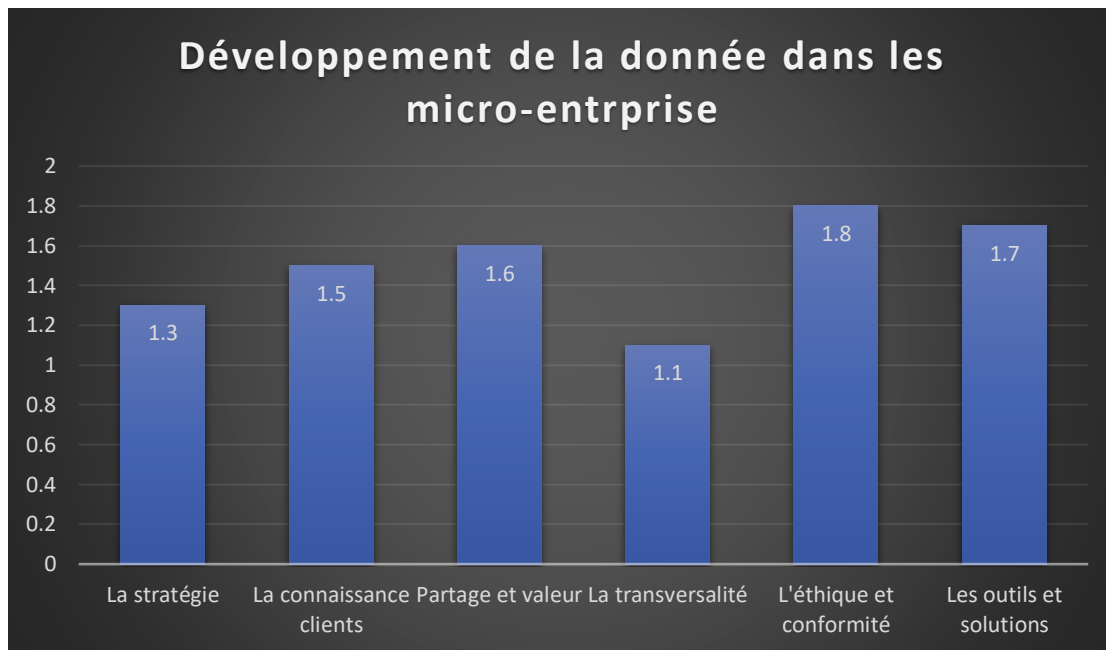


Tableau 1 - Développement de la donnée dans les micro-entreprises

Sur une moyenne sur trois, les axes « La transversalité » (1,1de moyenne) et « La stratégie d'entreprise » (1,3 de moyenne) représentent les plus faibles taux de maturité des entreprises. Les principales difficultés liées à la valorisation des données en entreprise reflètent ainsi certaines contraintes :

- Partage des données (interne et à l'externe) ;
- Accessibilité des données ;
- Décloisonnement des Métiers ;
- Optimisation de la qualité des produits et/ou des services ;
- Capacité à prendre des décisions éclairées, à partir de l'analyse de données (structurées et non structurées) en temps réel ;
- Capacité à développer des partenariats pour faire évoluer le modèle d'entreprise et à se différencier de la concurrence grâce à la valorisation des données.

Néanmoins, le radar montre que les points les forts sont les axes « L'Éthique & Conformité » (1,8 de moyenne) et « Les Outils et Solutions » (1,7 de moyenne). Il existe donc, pour une majorité d'entreprises une vision stratégique liée aux données. Il semble essentiel pour les entités sondées de faire de l'éthique un élément de différenciation, en garantissant à leurs clients finaux le respect de leur vie privée. La confiance est vue comme un argument de vente. Parallèlement, les entreprises s'équipent pour explorer et expérimenter les différents usages possibles des données, via notamment la mise en œuvre de systèmes d'architecture Big Data, mais aussi de services web.

Entre ces deux pôles extrêmes, se trouvent les axes « Le connaissance clients » (1,5 de moyenne) et « Partage et valeur » (1,7- de moyenne), qui reflètent ainsi un niveau de maturité moyen, laissant donc une certaine marge de progression. Ceci illustre bien les difficultés constatées chez les sondées autour de la gouvernance des données et de leur valorisation, encore confuses dans de nombreux micro-entreprises.

Globalement, ce graphique montre que les entreprises sondées sont bien conscientes de la valeur potentielle des données qui sont stockées dans leurs systèmes d'information et en dehors. Mais il démontre également qu'il reste encore beaucoup d'efforts à fournir avant que les petites structures ne deviennent des entreprises où la donnée est un élément structurant et essentiel à la culture de l'entreprise et où elle est gérée comme un actif clé.

4.2 La stratégie des entreprises

4.2.1 La « culture donnée »

La valorisation de données associée à la mise en œuvre de la gouvernance, conduit à supposer que la direction considère les données comme un atout stratégique de l'entreprise. Placer ce concept au cœur des micro-entreprises signifie des changements de culture et de modes de fonctionnement qui doivent être accompagnés.

Les données d'origine ont de la valeur, mais au fur et à mesure que le traitement progresse, la valeur aura tendance à augmenter, ce qui la rendra plus stratégique, et nécessitera la mise en place de mécanismes de gouvernance adaptatifs pour introduire la classification, les règles et les propriétaires au sein de l'entreprise.

Le principe du processus de vérification des données est de définir ce que sont des données précieuses. Voici les étapes à développer pour définir les exigences fonctionnelles :

- Définir les questions clés pour déterminer la valeur.
 - La validité : les critères correspondent au degré de conformité des données selon les règles ou contraintes définies.
 - La précision : délimiter le niveau d'exactitude et de netteté.
 - L'intégralité : examiner le niveau de complétude
 - La cohérence : observer sa pertinence.
 - L'uniformité : définir des normes à l'ensemble des données.
- Déterminer ce que nous voulons faire avec les données, planifier leur utilisation, organiser la transversalité pour éliminer les redondances de données, éliminer les incohérences dans le traitement du cycle de vie des données et responsabiliser le propriétaire.
- Trier en utilisant diverses méthodes. Déterminer la finalité des données : D'un point de vue macro, on peut en distinguer trois.
 - Objectif de communication,
 - Objectif commercial,
 - Objectif d'apprentissage.
- Mettre en place un système de contrôle de la qualité des données tout au long du cycle de vie.

4.3 Les connaissances clients

4.3.1 Développer une vision complète des clients

Avoir une vision globale du client est l'un des défis posés par le numérique. De nombreuses entreprises cherchent à développer cette vision. Certaines d'entre elles déploient des procédures centralisées destinées à « traduire » les données entre entreprises via des entités dédiées ou des succursales numériques.

Dans ce cas, une culture de valorisation peut devenir un atout pour l'entreprise, qui peut développer à la fois sa vision complète et sa connaissance de l'écosystème,

L'environnement du client, d'un point de vue "donnée", est très étendu : les données clients classiques sur les réseaux sociaux (données personnelles, centres d'intérêt, etc.) et les données externes (plus ou moins personnalisées) peuvent enrichir la connaissance client (données sociales, les habitudes de consommation...). Toutes ces dimensions peuvent être traitées par la technologie du Big Data pour produire des résultats pertinents.

Dans la distribution, la stratégie des sociétés est naturellement orientée vers la connaissance client, mais de nos jours, les problèmes de qualification sont principalement liés aux données produites.

D'un point de vue technique, un problème se pose pour mieux comprendre le produit, concernant tous les champs d'informations existants. Un croisement des informations avec le volume de produits et le volume de clients est nécessaire pour obtenir des éléments concrets fournissant de nouveaux services aux clients.

Dans le domaine du service, simplifier le processus de gestion des clients sur Internet est un défi, et l'utilisation des données peut devenir un levier : remplir automatiquement des questionnaires à partir des données commandées. Ces méthodes de service pragmatiques peuvent prolonger le parcours client.

Voici un exemple d'informations permettant l'enrichissement continu :

- Données démographiques sociales : identité, sexe, âge, profession, adresse, etc.
- Données de navigation : produits et pages consultés, cookies, paniers abandonnés, etc.
- Données de transaction : historique des achats et des retours ...
- Données interactives : service client, support technique, avis publiés ...

4.3.2 L'Intégration en temps réel

Toutes les entreprises ont des besoins et objectifs spécifiques (traitement des réclamations clients, optimisation des réseaux de distribution ou de transport et analyse des sinistres, etc.). Certaines d'entre elles font la distinction entre « temps d'interaction » et « temps manuel », allant de quelques secondes à quelques mois.

Les données en temps réel ne sont pas non plus nécessaires pour le système, c'est pourquoi pour effectuer une analyse de consommation, certaines entreprises préfèrent utiliser des données statiques pour le travail.

Dans le domaine des médias, la publicité en ligne aura également le même problème. En effet, la charge de travail des entreprises utilisant des outils de visualisation en temps réel est encore relativement faible.

La réelle difficulté réside dans la bonne compréhension des informations générées par chaque relais. Il est impossible d'analyser tous les phénomènes, vous devez donc savoir comment arrêter l'augmentation du débit lorsque vous n'en avez pas besoin, ou comment réduire la quantité de traitement lorsque vous en avez besoin.

Pour certaines entreprises, l'intégration des données en temps réel est un élément indispensable de la gestion de crise, notamment dans les cas de cyber sécurité, mais aussi pour analyser les dégâts des catastrophes naturelles ou gérer sa réputation sur les réseaux sociaux.

4.4 Partage des valeurs

4.4.1 Partager la valeur

L'échange de données n'est pas nouveau. Les individus, les organisations et les gouvernements ont longtemps échangé des informations avant l'existence d'ordinateurs et de réseaux.

En principe, la mise en place d'un système de partage dans un réseau signifie qu'on le réalise sous certaines conditions et surtout pour une cible spécifique. Les problèmes sont généralement liés à l'emplacement où sont stockées les informations à partager, à la qualité du flux que vous souhaitez garantir (si vous utilisez des vidéos, vous avez besoin d'un débit très rapide.), au niveau de sécurité d'accès et de stockage lui-même.

Aussi, il est important de connaître le type de fichiers pour les échanges. Dans le cas d'un fichier simple, un serveur de fichiers sera configuré. Mais si ces fichiers sont exploités par un logiciel alors un serveur de données est plus approprié.

Par exemple, dans le cas où l'information à partager est un fichier vidéo. Ce fichier peut transiter par téléchargement ou être diffusé en continu. Cela nécessite une configuration ou service spécifique pour chaque mode de diffusion.

Les actions ne sont pas simples. Les entreprises ont besoin d'une étude approfondie des besoins. Une nécessité qui pense non seulement aux données mais aussi aux programmes qui les administreront. Le problème des professionnels est lié aux droits d'accès, mais aussi à la capacité de stockage et du terminal.

Les avantages de la mise en place d'un partage de données qui répondent aux besoins spécifiques sont :

- Les organisations et les gouvernements peuvent partager davantage de données les uns avec les autres et peuvent remplir cela dans un type de données sûr, juste, légitime et respectueux concerné.
- La combinaison de données de différentes sources peut augmenter la performance et la valeur des services dans l'ordre. Il permet une meilleure recherche, un meilleur développement et fournit également de meilleurs produits.
- L'accessibilité des données permet une coopération efficace avec un minimum de risque.

4.5 La transversalité

4.5.1 L'organisation et la responsabilisation les participants

Afin d'adopter une méthode d'évaluation des données et de diffuser une 'culture donnée' dans toute l'entreprise. Il semble indispensable que dans une organisation, le manager nomme un responsable des données dans sa hiérarchie. Son rôle est :

- De coordonner les processus à tous les niveaux de l'entreprise et dans tous les domaines d'activité,
- D'être entièrement responsable de la gouvernance et de la stratégie des données.

Il est également nécessaire de se réunir régulièrement pour convoquer différents acteurs afin de vérifier et de mesurer les métriques du tableau de bord de gestion. La qualité des données, la cohérence et l'exhaustivité du catalogue, l'ouverture du référentiel, la présentation et l'expression des exigences. Une gouvernance appropriée doit également prendre en compte le rapport investissement / coût / valeur (dépenses ou pertes évitées en termes de coûts/bénéfices).

Voici un exemple de politique de la donnée à adopter par des points importants à destination des managers :

- Connaître les données générées par son entité.
- Savoir catégoriser les données utilisées par ses entités, et comprendre les classifications pertinentes et leur signification traitement.
- Reconnaître les données générées par son entité qui peuvent être partager en interne.
- Identifier les données que son entité souhaite récupérer.
- Percevoir les données générées par son entité qui peuvent être publiées en externe.
- Connaître le/s responsable/s de la qualité et de la sécurité des données utilisées par son entité.
- Discerner les responsables des publications des données générées par son entité.
- Avoir une réglementation écrite sur la durée de conservation des données.
- Rédiger des règles sur les données qui peuvent être utilisées.
- Avoir une parfaite connaissance des outils qui peuvent être utilisés pour partager des données en interne et externe.

4.6 L'éthique et conformité

4.6.1 L'éthique

La confiance est un argument de vente. En effet, dans certaines entreprises, il existe un délégué à la protection des données dont la responsabilité est de combiner l'éthique d'utilisation et le potentiel stratégique des données. Il est important de faire en sorte que cette éthique soit « une culture de la simplification de la complexité »

Cependant, l'idée d'éthique contre les entreprises est très puissante. L'idée opposée toujours "la vie privée est bonne pour les affaires et nous avons plutôt développé une "approche de conformité" plutôt qu'une approche client.

En raison de nos restrictions plus strictes sur l'Europe, nos progrès sont plus lents que les GAFAs et les objectifs de différenciation sur le marché, ne peuvent donc pas être atteints.

Le comité d'éthique, le Groupe des Nations Unies pour l'évaluation (GNUE) a défini une directive fixant un code de conduite pour les évaluateurs ainsi que des obligations d'ordre éthique relatives à la collecte et au stockage de données pour les personnes participant à une évaluation.

Et dans la plupart des cas, le comité doit approuver formellement la méthode de collecte des données, leur gestion et leur utilisation dans les évaluations. Il s'agit de garantir que ces pratiques respectent les obligations éthiques de tous ceux qui collectent des données auprès d'elles.

Les bonnes pratiques à exploiter sont :

- Développer des politiques internes pour gérer / utiliser les données dans l'entreprise et faire connaître les données aux clients finaux.
- Faire confiance aux avantages concurrentielles.
- S'assurer que les données client ne sont pas utilisées à mauvais escient.

Les limites pratiques souvent rencontrées dans la collecte de données sont les suivantes :

- Dans les cas suivants, les principaux initiés, participants et sites Web associés peuvent être indisponibles/inaccessibles :
 - Les informations sur le site prévu et/ou les participants sont incomplets,
 - Ils n'ont pas été contactés assez tôt, avant ou pendant la collecte des données,
 - L'évaluation est effectuée après la fin de la procédure ;
- Les sources de données pertinentes existantes, telles que les données d'enquête, sont limitées dans le temps et/ou de contenu (par exemple, manque d'indicateurs importants), Et/ou qualité (il y a des problèmes majeurs avec la qualité des données.);
- Dans le cas de grands programmes, les informations sur la qualité du service fourni par les différents agents d'exécution sont généralement limitées.

4.6.2 Le cycle de vie des données pour l'innovation

La définition de l'état d'assainissement des données en fonction du processus et des méthodes de vérification des données peut garantir la qualité des données contenues dans le référentiel pendant tout son cycle de vie.

Cela permet de confirmer que la collecte de données prévue couvre toutes les questions clés de l'évaluation, détermine si la triangulation entre les différentes sources de données est suffisante et aide à la conception des données.

Des outils de collecte de données (tels que des questionnaires, des questions d'entrevue, des outils d'extraction de données et des outils d'observation pour l'examen des documents) pour s'assurer que les informations nécessaires sont collectées.

Après avoir examiné les informations disponibles, il est important de créer une matrice d'évaluation, de spécifier les méthodes de collecte et d'analyse des données utilisées pour répondre à chaque question d'évaluation clé, puis d'identifier et de hiérarchiser les lacunes dans la collecte de nouvelles données qui doivent être collectées.

Question clés d'évaluation	Enquête sur les partis prenants	Entretiens avec les informateurs clés	Dossier projet	L'observation de la mise en œuvre
Question 1 Quelle a été la qualité de la mise en œuvre ?		✓	✓	✓
Question 2 Dans quelle mesure les objectifs ont-ils été atteints ?	✓	✓	✓	
Question 3 Quels ont été les autres impacts du programme ?	✓	✓		
Question 4 Comment le programme peut-il être amélioré ?		✓		✓

Tableau 2 - Matrice d'évaluation : faire correspondre la collecte de données aux questions clés d'évaluation

Il existe de nombreuses méthodes de collecte de données :

Options	Sources et méthodes
Recherche de documents et de données existants	<ul style="list-style-type: none"> ▪ Documents officiels relatifs aux politiques, plans de mise en œuvre du programme et rapports ▪ Statistiques officielles ▪ Données de suivi du programme ▪ Dossiers du programme
Collecte de données auprès de groupes ou d'individus	<ul style="list-style-type: none"> ▪ Entretiens avec des informateurs clés, individus, groupes, groupes de discussion, méthodes projectives ▪ Questionnaires ou enquêtes : par courriel, sur Internet, en face à face, données mobiles
Observation	<ul style="list-style-type: none"> ▪ Structuré ou non structuré ▪ Avec/sans participants ▪ Participatif ou non participatif ▪ Utiliser des notes, des photos ou des enregistrements vidéo
Mesures physiques	<ul style="list-style-type: none"> ▪ Mesure biophysique ▪ Information géographique

Tableau 3 - Exemple des sources de données possibles.

Ensuite, le problème qui se pose, notamment en termes de gouvernance, est celui de l'ouverture du référentiel de données afin de pouvoir l'alimenter et constituer le data lake¹² : on peut imaginer que dans la démarche, l'observatoire compétent (ou institution) sera les données sortantes prouvant aux parties prenantes.

¹² Le Data Lake est avant tout un moyen de stockage de grosses volumétries sans prétraitement des données brutes de différents formats et sources.

En résumé, la gouvernance données peut être exprimées comme ceci :

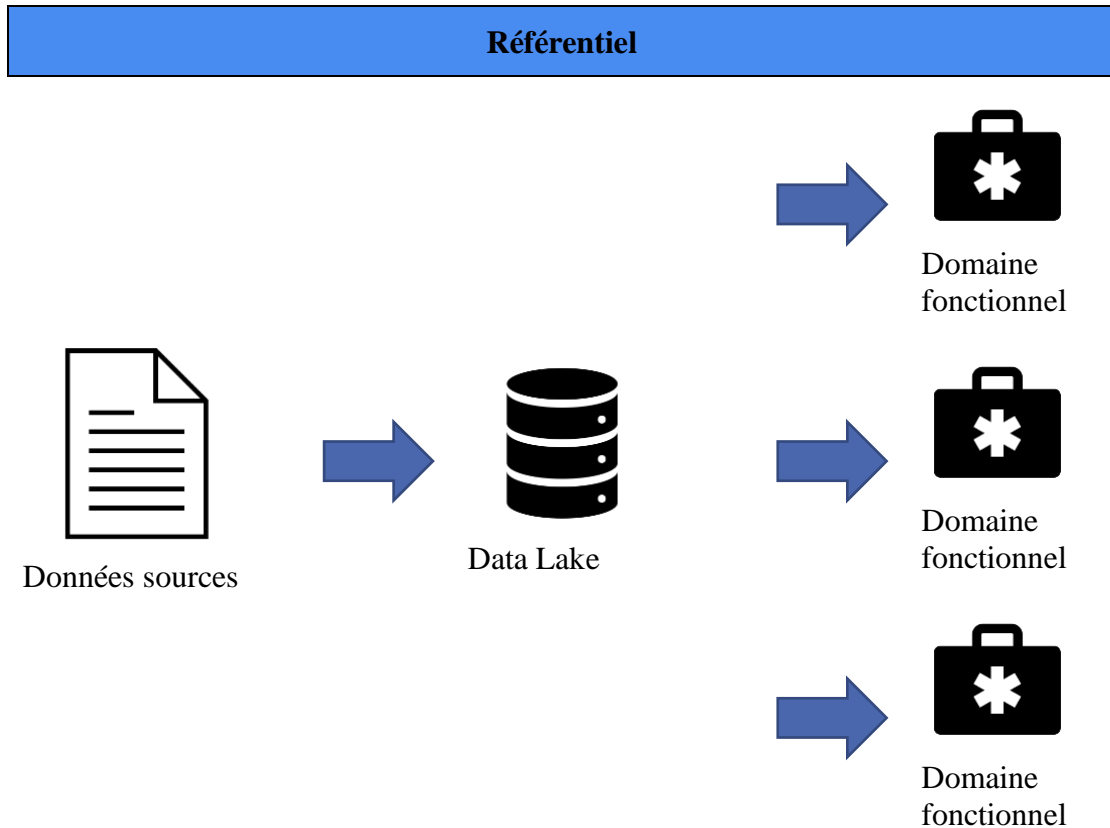


Figure 10 - Cycle de vie des données

La phase « sources données » :

Stocker toutes les données dans le « Data Lake » puis évaluer l'utilisation réelle en fonction de l'utilisation déterminée par une entité.

La phase « Data Lake » :

Des règles de gestion ont été définies et des outils ont été mis en place pour filtrer les données. Des réunions sont organisées pour communiquer sur la pertinence de ces données.

La phase « Domaine fonctionnel » :

Tout au long du processus, la qualité de toutes les données doit être garantie, même si ce n'est pas à tous les niveaux critiques. On peut imaginer la gestion minimale des données contenues dans le Data Lake, et responsabiliser chaque direction commerciale de la qualité de ces données, afin de constituer un catalogue de données de qualité, et créer un environnement concurrentiel favorable. La création de valeur peut être la communication, l'apprentissage ou l'utilisation commerciale.

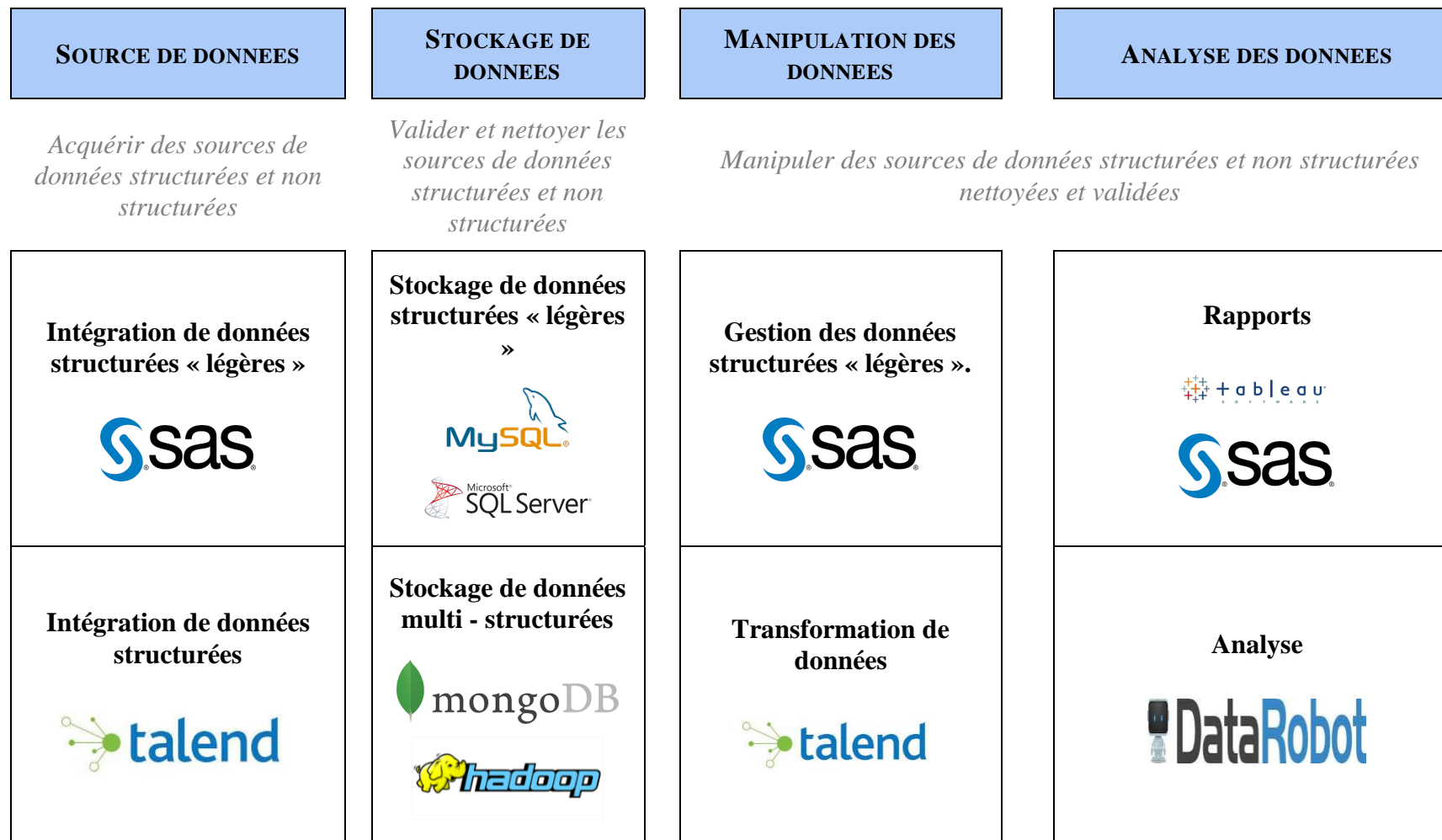
4.7 Les outils et solutions

4.7.1 Les outils

Les services Cloud ont fortement réduit les coûts de stockage car ils peuvent désormais stocker de grandes quantités de données qui ne sont pas forcément structurées dans des plates-formes telle que Hadoop : nous pouvons stocker de plus en plus de types de données, diversifier et élargir le périmètre d'analyse.

Les technologies :

- Technologie de source de données (base de données, application, Marché des données, ...) permet la collecte de données ;
- Technologie de traitement des données (enrichissement, traitement, devis, Sécurité ...) permet d'organiser et de gérer les données ;
- Les technologies d'application des données (insights, modèles, etc.) permettent d'analyser les données ;
- Les modèles analytiques permettent d'agencer et de mettre en œuvre le traitement.









<p>Intégration de données non - structurées</p> 	<p>Stockage de données structurées « lourdes »</p>  <p>Stockage de données non - structurées</p> 	<p>Gestion des données non - structurées.</p> 	<p>Exploitation / prédictive / statistique.</p>  <p>Gestion des données non structurées.</p> 
----------------------------------------------------------------------------------------------------------------------------------------------	-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	----------------------------------------------------------------------------------------------------------------------------------------------	-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Tableau 4 - Logiciels pour la gestion et l'analyse des données.

4.8 Conclusion

Si les données sont bien reconnues comme un actif stratégique clé par les entreprises, leur échéance en termes de données précieuses reste très inégale. Cela soulève des problèmes importants que ce rapport a essayé de mettre en évidence, partageant certaines méthodes et bonnes pratiques :

- L'auto-évaluation de la maturité des entreprises qui permet à chaque entreprise d'identifier leurs forces et leurs zones d'amélioration de la valeur des données ;
- Méthodes et exemples d'administration et d'organisation autour des données ;
- Un exemple de modèle général comportant les différentes étapes clés nécessaires à l'accès à la notation de service en fonction de la fourniture d'écosystèmes de données ;
- Les quelques cas d'utilisation qui ont été divisés et ceux présentés ici, selon le développement de modèles commerciaux, l'amélioration des connaissances clients et l'écosystème, la création et la valeur de partage autour des données.

Le thème de l'évaluation des données, leurs paris, mais aussi leur maturité toujours insuffisante illustre l'extension de la révolution compromise, à laquelle aucune industrie, aucune entreprise, ne sera en mesure d'échapper.

Dans de nombreuses grandes organisations, ces masses de données sont exploitées par des technologies de renseignement artificielles telle que la machine d'apprentissage afin d'innover en permanence, d'optimiser les usages et de les services liée.

Selon l'enquête de terrain, les entreprises exploitent partiellement le potentiel latent de leurs données. Pourtant, l'impact de la donnée dans le développement de l'activité est nombreux :

- Gouvernance globale ;
- Partenariats ;
- Gestion du cycle de vie de la donnée ;
- Transversalité de l'organisation ;
- Méthodologies et outillages adaptés ;
- Éthique et conformité ;
- Nouveaux services au client ;
- Monétisation et partage de la valeur avec les partenaires ;
- Etc.

Mais toutes les innovations basées sur l'exploitation des données ne seront bonnes que si elles sont bien gérées, de qualité, partagées et que la société crée une confiance, sur la base de règles explicites et limpides.

Sur ce dernier point, le sujet le plus urgent à traiter aujourd'hui, est certainement celui de la mise en œuvre du Règlement sur la protection des données (Selon le GDPR : réglementation de la réglementation de la protection des données générales)

Désormais, les grandes structures sont de concert avec le CNIL¹³ depuis plusieurs années afin d'établissent des politiques de protection des données. Tel que la protection des données à caractère personnel pour les entreprises, en particulier en ce qui concerne l'analyse des risques, la mise en œuvre de normes concrètes liées à la « confidentialité de données », à la portabilité des données, etc.

Le problème est de savoir comment interprété correctement ces textes, en tenant compte des appareils déjà existant. Cette complication doit être traité par des fournisseurs de service. Car en particulier, les services généraux (messages, gestion des ressources humaines, relation avec le client, ...) font l'objet d'une large gamme de logiciels dit SAAS¹⁴ ou cloud public gérés par les éditeurs.

¹³ La Commission nationale de l'informatique et des libertés de France est une autorité administrative indépendante française.

¹⁴ Le software as a service ou logiciel en tant que service, est un modèle d'exploitation commerciale des logiciels dans lequel ceux-ci sont installés sur des serveurs distants plutôt que sur la machine de l'utilisateur.

5. PARAGON ID : LA MATURATION DES DONNÉES

5.1 Paragon ID

5.1.1 Vision

A l'origine, la société se nommait Paragon Check Book Company fut fondée en 1886 à Londres par l'imprimeur, Samuel J. Moore. Celui-ci fut le premier à percevoir le potentiel dans l'impression de formulaires commerciaux.

Aujourd'hui, le groupe Paragon est sur le marché européen de l'imprimerie. Son rapprochement avec ASK, l'a spécialisé dans l'identification des produits (étiquettes, tags RFID), l'identification de personnes (tickets magnétiques & sans contact) et les solutions et services numériques (Personnalisation courriers & cartes, web-to-print).

Paragon Id souhaite tout mettre en œuvre pour que les tâches qu'elle accomplit puissent profiter pleinement et spontanément aux clients auxquels elle offre ses services.

De plus, la société souhaite prendre la place du leader du marché dans les produits d'identification dans des sociétés de services.

5.1.2 Missions

Paragon ID est une entreprise innovante, spécialisée dans l'identification de produits (étiquettes), l'identification de personnes (billetterie, badges NFC), de solutions et des services autour de la relation client (web to post). Paragon ID est l'une des rares entreprises françaises à éditer des billets magnétiques et sans doute, l'une des entreprises leaders sur le marché à proposer des billets sans contact.

L'usine située à Argent-sur-Sauldre (Cher) est le plus gros site de production d'étiquettes de traçabilité en France. Paragon ID a notamment le monopole des tickets de métro de Paris mais aussi de plusieurs métropoles dans le monde (Mexico, Le Caire, Los Angeles...) et gère des programmes de fidélité comme SNCF Grand voyageur ou le PMU.

L'étiquette est un produit technique nécessitant un savoir-faire spécifique dont la conception comprend plusieurs étapes :

- Définition des formes/formats
- Connaissance du support sur laquelle elle sera collée
- Usage ou traçabilité demandés, etc.

Ces exigences sont reprises dans un cahier des charges finalisé avec le client en fonction de ses besoins. Paragon fournit également des étiquettes avec code à barres, numérotation, puce RFid ..., et produit sur la technologie sans contact, des couvertures de passeport électronique, des cartes d'identités, des permis de conduire, des cartes de santé et d'autres documents sécurisés intelligents.

5.1.3 Stratégie

Dans le cadre de toutes ses missions, Paragon Id souhaite se forger une opinion publique, définir son point de vue et prendre des décisions en conséquence. L'objectif est de créer les conditions qui vont lui permettre de réaliser ces tâches :

- Promouvoir ses produits d'identification
- Mettre à disposition des outils permettant la production de solutions d'identification chez le client.
- Fournir plusieurs prestations de services (fournitures, service après-vente, entretien et maintenance...)

5.1.4 Valeurs clés

L'apport de Paragon Id doit être perçu comme très enrichissant parce qu'il engendre des effets positifs dans les processus d'identification de ses clients. La contribution fondamentale des produits doit aider à améliorer le processus de d'identification dans un marché inexistant.

5.1.5 Les référentiels

- Le transport
 - RATP, VINCI, UNDERGROUND, SNCF, ...
- L'industrie
 - REXEL, MICHELIN, FORD, VALEO, DELPHI, ...
- Le parking
 - MARIE DE PARIS, INDIGO, METRIC, ...
- Sports et loisir
 - ACCORD, UGC, PMU, FDJ, ...
- Logistique
 - La POSTE, TNT, ROYAL MAIL ...



5.1.6 Les chiffres clés

Paragon Id fait partie du Paragon Groupe, fournisseur de premier plan d'identification et de services de communication clients, dont le chiffre d'affaires s'élève à plus de 108M €, et qui compte près de 6.000 employés.

Paragon Id emploie plus de 600 personnes, avec des sites de production basés en Europe, aux États-Unis d'Amérique et en l'Asie.

5.2 Politique de gouvernance des données

5.2.1 La donnée au centre de la transformation digitale

Paragon Id place la valorisation des données au cœur de l'entreprise. Sa culture et son mode de fonctionnement sont orientés dans ce sens. Elle a mis en place un programme global de transformation digitale, avec un volet spécifique sur le « Big Data » comprenant :

- Une acculturation dans tous les services pour que l'approche s'enracine durablement, au travers des « Journées de sensibilisation » pour l'ensemble des cadres.
- Une incitation à l'exploration transverse des données afin d'identifier de nouveaux cas d'usage pertinents pour proposer des services incomparables pour les clients.
- L'accompagnement agile de tous les services qui le souhaitent, par une équipe transverse d'experts (notamment le service « Solution & Services ») afin de développer, en collaboration, des cas d'usage et procéder ainsi au transfert opérationnel de compétences.
- L'accès aisé à des moyens et plates-formes communalisées, pour que chaque service accède à moindre coût aux meilleurs outils et que la « technique » ne soit pas un obstacle. Pour exemple, nous avons déployé le service Office 365 par Microsoft à l'ensemble du personnel du site.

Un changement culturel très fort impulsé par la transformation digitale chez Paragon Id a été l'acceptation du « droit à l'échec » et s'autoriser à progresser pas à pas.

En résumé, ces mesures ont permis aux différents services de se forger une opinion par l'exemple, sans besoin préalable d'investissements majeurs, sur la base de projets qui ont donné d'excellents résultats. Elles ont également permis d'essayer et de mûrir progressivement les modes de travail agiles ainsi que la collaboration renforcée entre les équipes métiers et techniques, démontrant ainsi à la fois leur efficacité et leur compatibilité avec la culture d'entreprise.

Cette transformation s'est diffusée rapidement au sein du groupe et touche désormais tous les services. À ce titre, plusieurs projets opérationnels de ce type sont actuellement en cours de co-développement.

5.2.2 L'organisation et responsabilisation

La DSI de Paragon ID a beaucoup œuvré pour expliquer, convaincre, et accompagner les services, avec l'objectif de tirer de la valeur, et savoir ce qu'il faut mettre en place en termes de process pour garantir le niveau de qualité et d'intégrité de la donnée.

Les services « solutions & services »

La position du service dans l'entreprise se situe entre la production et les clients. Elle a pour missions la gestion et l'intégration des données clients en production, le développement d'une solution applicative répondant à des besoins clients, mais également à des problématiques en interne. Elle est divisée en trois pôles :

➤ **PÔLE SERVICES ET SOLUTIONS :**

Les experts en développement logiciel sont en charge des projets clients (en interne et externe à l'entreprise), de la gestion des traitements des données et de la mise en forme d'impression avec le logiciel métier Quadient. Ils expriment les besoins clients à la production et réalisent les solutions correspondantes.

➤ **PÔLE RFID :**

Les experts sont en charge de la personnalisation électrique des supports et des puces RFID. Ils travaillent conjointement avec le laboratoire de production RFID et réalisent des solutions en interne.

➤ **PÔLE INFRASTRUCTURE RÉSEAUX :**

L'expert s'occupe de la gestion complète du parc informatique du site et également de toute la partie administration du réseau.

Rôles

Une équipe d'experts est responsable du développement logiciel des projets clients (en interne et externe à l'entreprise), de la gestion des traitements des données et de la mise en forme d'impression avec le logiciel métier. Pour résumer, leurs objectifs en terme plus génériques sont :

- De co-développer avec les services, les premiers cas d'usage afin d'assurer la diffusion du savoir-faire, pour que ces métiers puissent ensuite le faire par eux-mêmes,
- D'apporter une aide méthodologique et algorithmique dans l'analyse des données
- De développer des outils automatisant les analyses, afin que les solutions soient plus autonomes (outils de Machine Learning, typiquement).

Par ailleurs, une organisation spécifique a été mise en place pour garantir la qualité des données, à la base pour répondre aux attentes des clients, mais de fait, à des fins opérationnelles :

- Au niveau du service, le responsable s'assure de la qualité des données. Il est le garant de la protection des données personnelles et s'appuie sur les pôles de service infrastructure, RFID et logiciel.
- Dans chaque pôle, un pilote mène, pour le service, la vérification de la pertinence des données tout au long de leur cycle de vie.
- Une organisation opérationnelle est en charge des contrôles effectifs : performance, test de non-régression et d'intégration.

5.2.3 Le cycle de vie des données pour l'innovation

Les aspects techniques (les plates-formes de traitements, des volumes de stockage, changement de format, ...) peuvent paraître complexes, au point qu'ils monopolisent parfois l'attention, au détriment des questions fondamentales de valorisation de données. Pour couper court à cette difficulté, l'approche Paragon ID consiste à régler de façon anticipée et décisive tous les aspects techniques du Big Data pour permettre ensuite aux métiers de se concentrer sur la valeur à tirer des données.

Pour cela, un modèle unique, partagé et « clé en main » d'architecture complète de Big Data a été mis au point, exclusivement sur la base d'outils Open Source, puis progressivement déployé sur les plates-formes du groupe.

Ce modèle enrobe l'ensemble des préoccupations techniques : architecture, moteurs d'analyse de données, outils, interfaces, solutions, évolution capacitaire, mais également exigences de sécurité, recommandations sur les modes de développement d'analyses et d'applications ou catalogue des compositions architecturales pour une bonne intégration dans le système d'information existant. Il est utilisé aussi bien pour des environnements d'exploration que de production.

Ce banc technologique a permis d'accélérer la diffusion des savoir-faire de valorisation de données, en assurant cinq aspects essentiels :

- L'absence de perte de temps et d'investissements en cherchant des solutions auprès des différents éditeurs ou en ingénierie interne.
- Le fait que la technique n'est plus un problème et qu'elle permette de se focaliser sur les besoins, les usages et les opportunités métiers.
- La réutilisation, facilitée par les services, des cas d'usage développés grâce à l'interopérabilité native des plates-formes.
- L'industrialisation claire et le changement de format des cas d'usage.
- Point important, il faut prévoir la montée en puissance des plates-formes ainsi que les ressources nécessaires : les premiers développements faits par chaque métier amènent très rapidement beaucoup de bénéfices en production.

5.2.4 L'éthique

Paragon ID a défini des règles pour formaliser sa posture d'entreprise de confiance, respectueuse de la vie privée de ses clients. Cette première approche très accessible était bien entendu destinée aux clients, mais aussi et surtout au groupe, afin que chaque collaborateur puisse facilement comprendre les enjeux et repérer la délimitation claire concernant la valorisation de données.

Cette étape a permis au groupe d'établir une stratégie des données, d'identifier les différents cas nécessitant un approfondissement de la posture, de mieux comprendre les difficultés engendrées par une valorisation dans son ensemble, tous métiers confondus, des données de l'entreprise et de confirmer la nécessité d'une gouvernance de cette stratégie au plus haut niveau faisant écho sur l'ensemble du groupe.

5.2.5 Les outils

Talend est un logiciel multifonction de traitement de données, qui associe un environnement complet capable d'intégrer le volume et la variété des données, à grande vitesse de traitement.

Il faut préciser que les technologies de ce type sont des actrices de notre écosystème. Elles nous accompagnent dans nos évolutions de service afin de mieux répondre aux besoins de nos clients. Les besoins client sont en constante évolution, si bien que Paragon Id impose à nos services la migration de nos données dans un système stockage structuré tel que Microsoft SQL Server. Cette migration est une preuve que nos services métiers dans ce domaine doivent effectuer l'état de l'art des technologies régulièrement.

Et Microsoft SQL Server semble être l'écosystème le plus dynamique pour nos besoins. Pour autant, il semble que le choix dominant aujourd'hui pour le Big Data soit Hadoop, même si beaucoup de clients sont encore en phase de pilotes sur Microsoft SQL Server. La maturité est assez avancée sur Microsoft SQL Server et l'enjeu réside dans l'expertise (il est difficile pour une entreprise de s'engager rapidement sur Hadoop du fait des expertises que ces technologies requièrent.).

5.3 La connaissance client

5.3.1 La donnée autour de l'expérience client

La vision de Paragon Id est d'être leader du marché dans les produits d'identification dans des sociétés de service. Pour cela, l'entreprise a considéré que sa différenciation pouvait se faire sur la personnalisation et l'interaction avec le client. Toutes les informations récoltées peuvent permettre à Paragon Id de proposer des services spécifiques au client pour qu'il personnalise leurs produits d'identification, sans que cela soit intrusif.

L'environnement de donnée d'un client est très vaste : données clients classiques (données personnelles, les données produites, ...) mais aussi des données externes, plus ou moins personnalisées, sur les médias sociaux, qui peuvent enrichir la connaissance cliente (donnée des réseaux sociaux, les données fournisseur...) des équipes Marketing notamment.

Toutes ces dimensions peuvent être traitées par des technologies Big Data pour avoir des résultats pertinents. Le travail de récupération des données étant facilité en amont, il est possible pour les équipes de se concentrer à la valorisation de ces données, sur les cas d'utilisation et sur la mise à disposition des données au service métier.

L'ambition est de mettre la confidentialité client au cœur de chacune des interactions avec les clients. Il faut pour cela savoir restituer ces données aux différents services "web to print" et canaux digitaux d'informations (emailing, médias sociaux, promotions, sites internet...).

5.3.2 Les objectifs de Paragon Id

- ◆ Développer une vision complète du client ;
- ◆ Connaître les données « produits » et « usages » pour construire des stratégies marketing
- ◆ Établir des projets commerciaux adaptés aux besoins clients ;
- ◆ Améliorer la qualité des données grâce à une bonne gestion des référentiels ;
- ◆ Croiser les données en temps réel ;
- ◆ Accéder aux données des partenaires : achat, commande... ;
- ◆ Garder un contact permanent avec le client ;
- ◆ Développer des applications pour remonter des informations en temps réel ;
- ◆ Adapter les applications traditionnelles au temps réel ;

5.4 La valeur ajoutée et le partage des données

5.4.1 Partager la valeur

Paragon Id tire beaucoup de valeur des données disponibles à l'extérieur de l'entreprise telles que les données des prestataires, celles des fournisseurs ou encore celles du marché de l'impression.

Pour l'entreprise, c'est une révolution industrielle qui dépasse la seule valorisation standard des données : l'exploitation des machines d'impression et du trafic de livraison produite, de la maintenance des infrastructures (équipements au sol et matériels roulants), ou encore de l'optimisation des consommations d'énergie. Mais elle partage très peu ses données disponibles.

En effet, dans sa politique de valorisation des données, la dimension de l'ouverture des données n'est pas couverte au sein de l'entreprise. Les données collectées qui ne sont pas exploitées peuvent être utilisées pour permettre la création de nouveaux services par des développeurs tiers. Un dispositif d'ouverture des données peut également favoriser l'interopérabilité du SI avec notre écosystème.

Par exemple, la SNCF met à disposition sur son site¹⁵ de nombreuses données sur des événements de sécurité, des données de réparations, sur des objets trouvés, le trafic en temps réel, des données LIDAR (Scan en 3D de l'environnement qu'il y a autour des rails SNCF ; ces dernières peuvent être intéressantes pour des fournisseurs.). En partageant certaines données, la SNCF contribue à faire vivre un écosystème.

5.4.2 Partage en interne

Pour le groupe, le partage de la valeur se fait en interne : elle se partage entre les services. L'exploitation des données issues des différents projets est un moyen efficace pour optimiser les processus et réduire les coûts, en suivant et contrôlant mieux les frais, les délais, les plannings, etc.

Il est possible que la création de la valeur se fasse avec une machine dite "e-learning". Mais dans notre cas, la corrélation de la valeur ne sera pas utile. Au-delà des systèmes, les partages de valeurs fonctionnelles sont clairement identifiés, comme la connaissance des applications métiers dans les services Paragon Id.

Par ailleurs, la culture de l'ouverture des données présente un avantage en matière de développement informatique, puisqu'elle fournit des codes accessibles et réutilisables par tous les membres du service.

¹⁵ <https://data.sncf.com/>

Sur ces points, la mise en place du service Microsoft SharePoint¹⁶ représente une opportunité pour poursuivre la transformation numérique, puisqu'elle repose sur la transmission dématérialisée des données issues des projets et sur la transmission dématérialisée de signalements d'événements aux partis prenants. C'est également par ce biais que les membres vont récupérer ou/et partager les informations nécessaires pour le management (planning, Procès verbale CSE, ...). D'où l'importance de vérifier régulièrement les mises à jour effectuées sur le réseau interne.

5.5 Conclusion

Dans l'ensemble, Paragon Id gagne de plus en plus de valeur dans son développement autour de ses données. Cette valorisation lui permet d'en faire de nombreux usages. Néanmoins, on peut constater que l'ouverture de ses données est un des axes d'amélioration que l'entreprise devrait suivre. Avec la stratégie d'entreprise, les moyens mis à disposition et les pratiques que nous utilisons. Il reste encore à identifier les acteurs qui seraient les mieux disposés à faire bénéficier de nos données aux structures externes.

Le partage de valeur élaboré en interne constitue le noyau important pour l'évaluation de l'intérêt, de la faisabilité et de pérennité d'un service basé sur les données à partager que nous continuons à développer. Ce cadre ouvre également la voie à de nouvelles méthodes d'évaluation de l'impact des initiatives d'ouverture des données.

¹⁶ SharePoint est un logiciel de gestion de contenu, moteurs de recherche, gestion électronique de documents, les forums et des statistiques décisionnelles

6. CONCLUSION GÉNÉRALE

En conclusion, à l'heure de la prolifération du Big Data, il faut prendre conscience qu'il existe encore des potentiels très importants dans l'optimisation des processus de collecte des données. L'apparition de nouveaux outils, permet de réaliser des gains de productivité très avantageux tout en valorisant les investissements déjà faits par les services métiers.

Par exemple, les appareils connectés aux réseaux internet ouvrent de nouveaux champs d'activité. Les données collectées par les capteurs permettront un meilleur développement produit et une meilleure réponse à la demande.

L'analyse des données permet également de réduire les coûts. En effet, de nombreuses entreprises ont du mal à suivre leurs factures. Provenant de nombreux vendeurs, très difficiles à décrypter et sont par conséquent très souvent payées sans vérification.

Pourtant, les progrès récents de l'intelligence artificielle et de l'apprentissage automatique, basés sur l'agrégation de toutes les données des entreprises et de l'expertise dans une base de connaissances, permettent :

- Dans le domaine des achats, d'automatiser les processus de suivi et de contrôle, la lecture et l'extraction des données à valeur ajoutée (chaque ligne de facture est analysée et contrôlée) au paiement ;
- Dans de nombreux autres domaines, l'automatisation de toutes les tâches administratives pour optimiser les dépenses.

Cette approche numérique moderne de l'utilisation et de la valorisation des données est à la base de la transformation numérique des entreprises. Elle s'inscrit également dans une démarche collaborative entre entreprises.

En effet, des données non stratégiques peuvent être partagées sur des fichiers stockés dans un environnement totalement sécurisé (entraide administrative, règles de gestion d'auto-apprentissage, partage des connaissances sur la problématique rencontrée, échange de bonnes pratiques, etc.).

La valorisation des données constitue à la fois un vrai levier d'amélioration et de performance pour les entreprises. Elles peuvent se recentrer sur des missions métiers valorisantes. Et l'ouverture des données d'entreprises gagne de plus en plus de valeur et permet de nombreux usages. La qualité des services et produits dépendra de sa gestion.

Cette thèse a pour principal but de faire une analyse approfondie sur le sujet afin de mieux le situer dans le contexte d'entreprise. L'étude m'a permis d'éclaircir le potentiel latent autour de la question à laquelle une recherche minutieuse m'a apporté des réponses.

7. Annexe

7.1 Questionnaire

Développement des entreprises en matière de valorisation des données.

Évaluez le niveau de développement en matière de valorisation des données.

* Obligatoire

La stratégie de l'entreprise

1) Si vous possédez une base de données (tableau Excel, SGBD, ...), dans quelle proportion est-elle déployée ?

La stratégie de l'entreprise

- ☐ Inexistant
- ☐ Novice
- ☐ Avancé
- ☐ Confirmé
- ☐ N/A (Non appliqué / Ne sait pas)

2) Exploitez-vous vos données pour proposer de nouveau service ?

La stratégie de l'entreprise

- ☐ Inexistant
- ☐ Novice
- ☐ Avancé
- ☐ Confirmé
- ☐ N/A (Non appliqué / Ne sait pas)

3) Un nouveau modèle de service innovant d'entreprise ont été mis en place grâce à l'usage et l'analyse de vos données d'entreprise ?

La stratégie de l'entreprise

- ☐ Inexistant
- ☐ Novice
- ☐ Avancé
- ☐ Confirmé
- ☐ N/A (Non appliqué / Ne sait pas)

- 4) Dans quelle proportion les données (marketing, techniques, financières, RH ou autres) sont une composante essentielle aux développements de vos produits et services ?
*

La stratégie de l'entreprise

- ☐ Inexistant
- ☐ Novice
- ☐ Avancé
- ☐ Confirmé
- ☐ N/A (Non appliqué / Ne sait pas)

- 5) Avez-vous mise en place un système de rétribution au client qui partage ses données (système fidélité client)?

La stratégie de l'entreprise

- ☐ Inexistant
- ☐ Novice
- ☐ Avancé
- ☐ Confirmé
- ☐ N/A (Non appliqué / Ne sait pas)

- 6) Avez-vous des accords de partenariats sur le partage de vos données existent (sous-traitants, co-traitants, fournisseurs...)?

La stratégie de l'entreprise

- ☐ Inexistant
- ☐ Novice
- ☐ Avancé
- ☐ Confirmé
- ☐ N/A (Non appliqué / Ne sait pas)

- 7) Dans quelle proportion exploitez-vous vos données pour éclairer les prises de décisions ?

La stratégie de l'entreprise

- ☐ Inexistant
- ☐ Novice
- ☐ Avancé
- ☐ Confirmé
- ☐ N/A (Non appliqué / Ne sait pas)

Les connaissances clients

8) Avez-vous mis en place des actions pour identifier les clients/partenaires ?

Les connaissances clients

- ☐ Inexistant
- ☐ Novice
- ☐ Avancé
- ☐ Confirmé
- ☐ N/A (Non appliqué / Ne sait pas)

9) Collectez-vous les données sur les attentes des clients afin à personnaliser les services et l'expérience client?

Les connaissances clients

- ☐ Inexistant
- ☐ Novice
- ☐ Avancé
- ☐ Confirmé
- ☐ N/A (Non appliqué / Ne sait pas)

10) Au-delà des données clients, les données produits ou service sont t'elles exploitées pour améliorer la connaissance des clients ?

Les connaissances clients

- ☐ Inexistant
- ☐ Novice
- ☐ Avancé
- ☐ Confirmé
- ☐ N/A (Non appliqué / Ne sait pas)

11) Mobilisez-vous les données pour prédire ou anticiper les comportements des clients ? *

Les connaissances clients

- ☐ Inexistant
- ☐ Novice
- ☐ Avancé
- ☐ Confirmé
- ☐ N/A (Non appliqué / Ne sait pas)

12) Adaptez-vous des stratégies marketing en fonction des connaissances des données clients et/ou produits ?

Les connaissances clients

- ☐ Inexistant
- ☐ Novice
- ☐ Avancé
- ☐ Confirmé
- ☐ N/A (Non appliqué / Ne sait pas)

13) Analysez-vous des données des concurrents et des nouveaux entrants disrupteurs ?

Les connaissances clients

- ☐ Inexistant
- ☐ Novice
- ☐ Avancé
- ☐ Confirmé
- ☐ N/A (Non appliqué / Ne sait pas)

14) Assurez-vous une veille active sur les réseaux sociaux en temps réel pour anticiper, prendre des décisions ?

Les connaissances clients

- ☐ Inexistant
- ☐ Novice
- ☐ Avancé
- ☐ Confirmé
- ☐ N/A (Non appliqué / Ne sait pas)

La transversalité

15) Partagez-vous efficacement les données de valeur en interne ?

La transversalité

- ☐ Inexistant
- ☐ Novice
- ☐ Avancé
- ☐ Confirmé
- ☐ N/A (Non appliqué / Ne sait pas)

16) Diffusez-vous la connaissance des outils d'analyse des données de façon à faciliter l'organisation de l'entreprise? La transversalité

- ☐ Inexistant
- ☐ Novice
- ☐ Avancé
- ☐ Confirmé
- ☐ N/A (Non appliqué / Ne sait pas)

17) Avez-vous mise en place un systemes de collecte des données issu de capteurs et objets connectés internes afin de les analysée ?

La transversalité

- ☐ Inexistant
- ☐ Novice
- ☐ Avancé
- ☐ Confirmé
- ☐ N/A (Non appliqué / Ne sait pas)

18) Mobilisez-vous des données pour offrir des combinaisons de produits/services les plus adaptés? *

La transversalité

- ☐ Inexistant
- ☐ Novice
- ☐ Avancé
- ☐ Confirmé
- ☐ N/A (Non appliqué / Ne sait pas)

19) Collectez-vous des données pour mesurer l'usage réel des produits et services?

La transversalité

- ☐ Inexistant
- ☐ Novice
- ☐ Avancé
- ☐ Confirmé
- ☐ N/A (Non appliqué / Ne sait pas)

20) Partagez-vous des données avec des concurrents/partenaires pour améliorer la productivité de l'activité ?

La transversalité

- ☐ Inexistant
- ☐ Novice
- ☐ Avancé
- ☐ Confirmé
- ☐ N/A (Non appliqué / Ne sait pas)

21) Analysez-vous les données issues du processus R/H pour définir la valeur capitale humain?

La transversalité

- ☐ Inexistant
- ☐ Novice
- ☐ Avancé
- ☐ Confirmé
- ☐ N/A (Non appliqué / Ne sait pas)

Partage des valeurs

22) Organiser vous le management des données reconnue comme un actif clé au sein de l'entreprise ?

Partage des valeurs

- ☐ Inexistant
- ☐ Novice
- ☐ Avancé
- ☐ Confirmé
- ☐ N/A (Non appliqué / Ne sait pas)

23) Existe-t-il une stratégie et des politiques de valorisation des données ?

Partage des valeurs

- ☐ Inexistant
- ☐ Novice
- ☐ Avancé
- ☐ Confirmé
- ☐ N/A (Non appliqué / Ne sait pas)

24) Y a t'il un suivi régulier sur la cohérence de la politique de gestion des données?

*

Partage des valeurs

- ☐ Inexistant
- ☐ Novice
- ☐ Avancé
- ☐ Confirmé
- ☐ N/A (Non appliqué / Ne sait pas)

25) Avez-vous clairement définie l'organisation, les rôles et les responsabilités en matière de gestion des données ?

Partage des valeurs

- ☐ Inexistant
- ☐ Novice
- ☐ Avancé
- ☐ Confirmé
- ☐ N/A (Non appliqué / Ne sait pas)

26) Avez-vous mis en place une démarche d'identification des sources de données non-structurée ?

Partage des valeurs

- Inexistant
- Novice
- Avancé
- Confirmé
- N/A (Non appliqué / Ne sait pas)

27. Une politique d'ouverture est-elle définie pour les données (interne et externe) ?

Partage des valeurs

- Inexistant
- Novice
- Avancé
- Confirmé
- N/A (Non appliqué / Ne sait pas)

27) Existence-t-il des méthodes de préparation des données pour créer des ensembles de données cohérents ?

Partage des valeurs

- Inexistant
- Novice
- Avancé
- Confirmé
- N/A (Non appliqué / Ne sait pas)

L'éthique et conformité

28) Avez-vous défini, partagé et appliqué des règles éthiques sur l'usage et l'exploitation des données ?

L'éthique et conformité

- ☐ Inexistant
- ☐ Novice
- ☐ Avancé
- ☐ Confirmé
- ☐ N/A (Non appliqué / Ne sait pas)

29) Avez-vous défini un système de gestion de crise en cas de fuites et/ou d'usages détournés de vos données ?

L'éthique et conformité

- ☐ Inexistant
- ☐ Novice
- ☐ Avancé
- ☐ Confirmé
- ☐ N/A (Non appliqué / Ne sait pas)

30) Avez-vous placé un processus d'anonymisation des données personnelles pour permettre leur exploitation ?

L'éthique et conformité

- ☐ Inexistant
- ☐ Novice
- ☐ Avancé
- ☐ Confirmé
- ☐ N/A (Non appliqué / Ne sait pas)

31) Tenez-vous compte des règles de protection des données personnelles ?

L'éthique et conformité

- ☐ Inexistant
- ☐ Novice
- ☐ Avancé
- ☐ Confirmé
- ☐ N/A (Non appliqué / Ne sait pas)

32) Avez-vous clairement identifiés et encadrés les droits de propriété intellectuelle ?

L'éthique et conformité

- ☐ Inexistant
- ☐ Novice
- ☐ Avancé
- ☐ Confirmé
- ☐ N/A (Non appliqué / Ne sait pas)

Outils et solutions

33) Avez-vous mis en place de système de partage des données (Des connecteurs, APIs, et autres web services, etc..)?

Outils et solutions

- ☐ Inexistant Novice
- ☐ Avancé
- ☐ Confirmé
- ☐ N/A (Non appliqué / Ne sait pas)

34) Des systèmes/outils sont mis en place pour capter les données en temps réel?

Outils et solutions

- ☐ Inexistant Novice
- ☐ Avancé
- ☐ Confirmé
- ☐ N/A (Non appliqué / Ne sait pas)

35) Est-ce que vos système mis en place tient en compte l'interopérabilité les données ?

Outils et solutions

- ☐ Inexistant
- ☐ Novice
- ☐ Avancé
- ☐ Confirmé
- ☐ N/A (Non appliqué / Ne sait pas)

36) Disposez-vous d'une architecture et/ou des applications qui permet la collecte, le traitement et stockage d'un masse important de données ? *

Outils et solutions

- ☐ Inexistant
- ☐ Novice
- ☐ Avancé
- ☐ Confirmé
- ☐ N/A (Non appliqué / Ne sait pas)

8. Bibliographie

Breezcar. (2017, avril 17). *Voiture autonome : Tesla va-t-il bouleverser le monde de l'assurance ?* Récupéré sur Breezcar.com: <https://www.breezcar.com/actualites/article/contrat-assurance-specifique-voiture-electrique-et-autonome-0417>

CNIL. (2018, Mai 23). *CNIL.fr*. Récupéré sur Le règlement général sur la protection des données: <https://www.cnil.fr/fr/reglement-europeen-protection-donnees>

CNIL. (2019, Juin 17). *La loi Informatique et Libertés*. Récupéré sur CNIL.fr: <https://www.cnil.fr/fr/la-loi-informatique-et-libertes>

CNIL. (s.d.). *Donnée personnelle*. Récupéré sur CNIL.fr: <https://www.cnil.fr/fr/definition/donnee-personnelle>

Cookiebot. (2020, février 24). *Les cookies sur Internet – L'essentiel*. Récupéré sur Cookiebot.com: <https://www.cookiebot.com/fr/cookies-internet/>

Devillard, A. (2018, Janvier 01). *Handicap International teste l'impression 3D de prothèses sur mesure*. Consulté le Février 2020, sur Science & Avenir: https://www.sciencesetavenir.fr/high-tech/handicap-international-teste-l-impression-3d-de-protheses-sur-mesure_119704

Fullrama. (2020). *Fullrama*. Consulté le Février 2020, sur TRAITEMENT & ANALYSE : <https://fullrama.net/data-processing/>

Lafaye, L. (2019, Aout 30). *Les échanges de données, clé de voûte de l'économie*. Récupéré sur Journal du Net: <https://www.journaldunet.com/solutions/reseau-social-d-entreprise/1443538-les-echanges-de-donnees-cle-de-voute-de-l-economie/>

Largerion, M. (2016, 09 12). *Amazon: histoire et performances webmarketing*. Consulté le Février 2020, sur MLConseil Marketing & Communication Digitale: <https://www.mauricelargerion.com/amazon-l-inventeur-du-web-marketing/>

LesDefinitions.fr. (2013, juillet 30). *LesDefinitions.fr*. Récupéré sur Définition de marchandise : <https://lesdefinitions.fr/marchandise>

Raymond, J.-L. (2020, janvier 10). *DESI 2019 : LES CHIFFRES CLÉS DU NUMÉRIQUE EN EUROPE ET EN FRANCE*. Récupéré sur Francenum.gouv.fr: <https://www.francenum.gouv.fr/comprendre-le-numerique/desi-2019-les-chiffres-cles-du-numerique-en-europe-et-en-france>

Schonberger, V. M. (2013). *Big Data - La revolution des données est n marche*. Consulté le Mars 2020, sur Google Book: <https://books.google.fr/books?hl=fr&lr=&id=RZzbAgAAQBAJ&oi=fnd&pg>

=PT6&dq=big+data+exposé&ots=m3XmTA1yVC&sig=DEZvrGEcPwLfAi
CoRw9HptQ-IBU#v=onepage&q=big%20data%20exposé&f=false

Sobel, N. P. (2010, Février). *Le concept de « marchandise fictive », pierre angulaire de l'institutionnalisme de Karl Polanyi ?* Consulté le Mars 2020, sur Cairn Info: <https://www.cairn.info/revue-de-philosophie-economique-2010-2-page-3.htm>

Vincent, Y. (2015, 06 26). *Les volumes de production des usines françaises ne comptent plus pour PSA*. Récupéré sur L'Usine Nouvelle: <https://www.usinenouvelle.com/article/les-volumes-de-production-des-usines-francaises-ne-comptent-plus-pour-psa.N338566>

Wikipedia. (2017, janvier 23). *Base de données relationnelle*. Récupéré sur [Wikipedia.org: https://fr.wikipedia.org/wiki/Base_de_données_relationnelle](https://fr.wikipedia.org/wiki/Base_de_données_relationnelle)

Wikipedia. (2020, janvier 17). *Talend*. Récupéré sur Wikipedia: <https://fr.wikipedia.org/wiki/Talend>

9. Travaux cités

Breezcar. (2017, avril 17). *Voiture autonome : Tesla va-t-il bouleverser le monde de l'assurance ?* Récupéré sur Breezcar.com: <https://www.breezcar.com/actualites/article/contrat-assurance-specifique-voiture-electrique-et-autonome-0417>

CNIL. (2018, Mai 23). *CNIL.fr*. Récupéré sur Le règlement général sur la protection des données: <https://www.cnil.fr/fr/reglement-europeen-protection-donnees>

CNIL. (2019, Juin 17). *La loi Informatique et Libertés*. Récupéré sur CNIL.fr: <https://www.cnil.fr/fr/la-loi-informatique-et-libertes>

CNIL. (s.d.). *Donnée personnelle*. Récupéré sur CNIL.fr: <https://www.cnil.fr/fr/definition/donnee-personnelle>

Cookiebot. (2020, février 24). *Les cookies sur Internet – L'essentiel*. Récupéré sur Cookiebot.com: <https://www.cookiebot.com/fr/cookies-internet/>

Devillard, A. (2018, Janvier 01). *Handicap International teste l'impression 3D de prothèses sur mesure*. Consulté le Fevrier 2020, sur Science & Avenir: https://www.sciencesetavenir.fr/high-tech/handicap-international-teste-l-impression-3d-de-protheses-sur-mesure_119704

Fullrama. (2020). *Fullrama*. Consulté le Fevrier 2020, sur TRAITEMENT & ANALYSE : <https://fullrama.net/data-processing/>

Lafaye, L. (2019, Aout 30). *Les échanges de données, clé de voûte de l'économie*. Récupéré sur Journal du Net: <https://www.journaldunet.com/solutions/reseau-social-d-entreprise/1443538-les-echanges-de-donnees-cle-de-voute-de-l-economie/>

Largerion, M. (2016, 09 12). *Amazon: histoire et performances webmarketing*. Consulté le Fevrier 2020, sur MLConseil Marketing & Communication Digitale: <https://www.mauricelargerion.com/amazon-l-inventeur-du-web-marketing/>

LesDefinitions.fr. (2013, juillet 30). *LesDefinitions.fr*. Récupéré sur Définition de marchandise : <https://lesdefinitions.fr/marchandise>

Raymond, J.-L. (2020, janvier 10). *DESI 2019 : LES CHIFFRES CLÉS DU NUMÉRIQUE EN EUROPE ET EN FRANCE*. Récupéré sur Francenum.gouv.fr: <https://www.francenum.gouv.fr/comprendre-le-numerique/desi-2019-les-chiffres-cles-du-numerique-en-europe-et-en-france>

Schonberger, V. M. (2013). *Big Data - La révolution des données est n marche*. Consulté le Mars 2020, sur Google Book: <https://books.google.fr/books?hl=fr&lr=&id=RZzbAgAAQBAJ&oi=fnd&pg=PT6&dq=big+data+exposé&ots=m3XmTA1yVC&sig=DEZvrGEcPwLfAiCoRw9HptQ-IBU#v=onepage&q=big%20data%20exposé&f=false>

Sobel, N. P. (2010, Février). *Le concept de « marchandise fictive », pierre angulaire de l'institutionnalisme de Karl Polanyi ?* Consulté le Mars 2020, sur Cairn Info: <https://www.cairn.info/revue-de-philosophie-economique-2010-2-page-3.htm>

Vincent, Y. (2015, 06 26). *Les volumes de production des usines françaises ne comptent plus pour PSA*. Récupéré sur L'Usine Nouvelle: <https://www.usinenouvelle.com/article/les-volumes-de-production-des-usines-francaises-ne-comptent-plus-pour-psa.N338566>

Wikipedia. (2017, janvier 23). *Base de données relationnelle*. Récupéré sur [Wikipedia.org: https://fr.wikipedia.org/wiki/Base_de_données_relationnelle](https://fr.wikipedia.org/wiki/Base_de_données_relationnelle)

Wikipedia. (2020, janvier 17). *Talend*. Récupéré sur Wikipedia: <https://fr.wikipedia.org/wiki/Talend>