

**Haute Ecole
Groupe ICHEC – ECAM – ISFSC**



Enseignement supérieur de type long de niveau universitaire

**Big data & Transformation digitale des PME/TPE :
mise en perspective, opportunités, enjeux
et obstacles**

Mémoire présenté par :
Fernanda CORTES PALACIOS

Pour l'obtention du diplôme de :
**Master en sciences
commerciales**

Année académique 2020 – 2021

Promoteur :
Emilio VILLANO

Je soussigné, CORTES PALACIOS, Fernanda, master en Sciences commerciales 2020-2021, déclare par la présente que le mémoire ci-joint est exempt de tout plagiat et respecte en tous points le règlement des études en matière d'emprunts, de citations et d'exploitation de sources diverses signé lors de mon inscription à l'ICHEC, ainsi que les instructions et consignes concernant le référencement dans le texte respectant la norme APA, la bibliographie respectant la norme APA, etc. mises à ma disposition sur Moodle.

Par ma signature, je certifie sur l'honneur avoir pris connaissance des documents précités et que le travail présenté est original et exempt de tout emprunt à un tiers non-cité correctement.

Date et Signature.



15/08/2021

**Haute Ecole
Groupe ICHEC – ECAM – ISFSC**



Enseignement supérieur de type long de niveau universitaire

**Big data & Transformation digitale des PME/TPE :
mise en perspective, opportunités, enjeux
et obstacles**

Mémoire présenté par :
Fernanda CORTES PALACIOS

Pour l'obtention du diplôme de :
**Master en sciences
commerciales**

Année académique 2020 – 2021

Promoteur :
Emilio VILLANO

Remerciements

Je souhaite adresser mes remerciements les plus sincères à toutes les personnes qui m'ont aidé au cours de la réalisation de ce mémoire.

Mes remerciements particuliers vont :

- A mon promoteur, le professeur Emilio Villano pour ses conseils, sa disposition et son soutien lors de l'élaboration de cette recherche.
- Aux sept experts qui ont accepté d'enrichir cette étude à travers le partage de leurs connaissances et de leur expérience.
- Aux entrepreneurs, au sens large, qui ont participé à l'enquête quantitative.
- A Claire Duponcheel pour ses encouragements, son enthousiasme et le temps précieux qu'elle a consacré à la lecture et à la révision de mon travail.

Enfin, je tiens à exprimer ma reconnaissance envers ma famille pour son soutien inconditionnel.

Table de matières

Introduction générale	1
1. Introduction	1
1.1. Motivations	1
2. Positionnement du problème	2
3. Objectif général.....	3
4. Fil conducteur	4
Partie 1.....	6
Approche théorique : revue de la littérature	6
♦ Perspective technique ♦	6
1. Introduction	6
2. Big data	6
2.1. Définition.....	6
2.2. Caractéristiques du big data.....	8
3. L'analyse des données (Data Analytics).....	10
3.1. Les types d'analyse de données	12
4. Big data et l'intelligence artificielle (IA).....	14
4.1. Machine Learning (ML) ou apprentissage automatique.....	15
4.2. Deep Learning (DL) ou apprentissage profond.....	16
5. Donner du sens aux données.....	17
5.1. Business Intelligence (BI) ou informatique décisionnelle	18
5.2. Data Mining (DM) ou forage/exploration/fouille de données	19
5.3. Data Visualisation	20
6. Data Pipeline.....	21
7. Ecosystème big data.....	22
♦ Perspective Managériale ♦	26
1. La révolution des données qui transforme l'entreprenariat et les entreprises.....	26
2. Les modèles d'affaires et de revenus de l'ère numérique	27
3. Opportunités, enjeux et obstacles dans le cadre de projets (big) data / IA	32
3.1. Opportunités : finalités d'usage des projets (big) data / IA	32
3.2. Enjeux liés aux projets (big) data / IA	37
4. Obstacles structurels face aux projets (big) data / IA	39
5. Eviter le big échec	40

6.	Mener à bien un projet (big) data / IA.....	41
7.	Vers la transformation digitale des entreprises.....	44
7.1.	Dématérialisation, digitalisation, innovation et transformation digitale	44
7.2.	Transformation digitale : technologie ou culture ?	45
7.3.	Leviers stratégiques de la transformation digitale	48
7.4.	Se faire aider et accompagner.....	51
7.5.	Facteurs d'échec et défis de la transformation digitale.....	53
7.6.	Une vision, une stratégie et beaucoup de volonté.....	54
7.7.	La transformation digitale en action.....	55
	Partie 2	57
	Approche Pratique : études empiriques	57
1.	Etude quantitative.....	57
1.1.	Méthodologie	57
1.2.	Objectifs	57
1.3.	Limites.....	57
1.4.	Questionnaire.....	58
1.5.	Caractéristiques et proportions de l'échantillon.....	58
1.6.	Analyses des résultats.....	59
2.	Etude qualitative	64
2.1.	Méthodologie	64
2.2.	Objectifs	64
2.3.	Limites.....	64
2.4.	Entretiens individuels semi-directifs.....	65
	Partie 3	71
	Approche Analytique	71
	Conclusion	76
	Bibliographie.....	77

Liste des figures

Figure 1 : (Big) data – Cycle itératif d'analyse	10
Figure 2 : Types d'analyse de données.....	12
Figure 3 : Les approches de l'intelligence artificielle	14
Figure 4 : Degré d'intelligence	17
Figure 5 : Architecture Business Intelligence	18
Figure 6 : Etapes du Data Mining.....	20
Figure 7 : Data Pipeline	21
Figure 8 : Big Data (Hadoop) ecosystem	23
Figure 9 : Avantages globaux du (Big) data	32
Figure 10 : Enjeux organisationnels, technologiques et économiques du big data pour les entreprises.....	37
Figure 11 : Projets (big) data – Approche par critères de succès	43
Figure 12 : Modèle conceptuel - lien entre digitalisation, Innovation et Transformation digitale	45
Figure 13 : Les PME, un net retard dans l'adoption des technologies numériques	47
Figure 14 : Les 7 étapes vers la transformation digitale	56
Figure 15 : Résultats de questionnaire : familiarisation technologique	59
Figure 16 : Résultats de questionnaire : perception sur l'utilité des technologies numériques .	60
Figure 17 : Résultats de questionnaire : rapport utilité/profil	60
Figure 18 : Résultats de questionnaire : usage.....	61
Figure 19 : Résultats de questionnaire : limites	61
Figure 20 : Résultats de questionnaire : collecte de données internes	62
Figure 21 : Résultats de questionnaire : collecte de données externes	62
Figure 22 : Résultats de questionnaire : aide à l'accompagnement.....	63

Liste des tableaux

Tableau 1: Modèles de revenus de l'ère numérique - Crédit et captation de valeur	29
Tableau 2 : Modèles d'usages et ses bénéfices selon Bénavent (2014)	33
Tableau 3 : Les facteurs de succès d'une transformation digitale réussie.....	55
Tableau 4 : Echantillon : type	58
Tableau 5 : Echantillon : âge.....	58
Tableau 6 : Echantillon : siège	58
Tableau 7 : Echantillon : secteur d'activité.....	59
Tableau 8 : Entretiens : profils des interlocuteurs.....	65

Introduction générale

1. Introduction

Un smartphone en poche, une smartwatch au poignet ou encore un Google Home qui nous raconte des mauvaises blagues, nous sommes devenus des individus hyper connectés et ultra impatients, nous voulons que les informations soient toujours à disposition et de plus en plus rapidement.

C'est cette connexion permanente des individus et des entreprises, notamment par notre passage volontaire et manuel sur internet mais aussi instantanément à travers des objets connectés et autres technologies numériques, qui a contribué à la croissance incessante des données. C'est le big data.

Dans l'actualité, les données sont considérées comme un actif stratégique, un vecteur de valeur pour les organisations, peu importe leur taille. A priori utiles mais pas toujours utilisées, il convient de comprendre les opportunités et les enjeux liées ainsi que les raisons qui conduisent à les sous-estimer.

En plus de rassembler les apports et les challenges du big data pour les petites et moyennes structures, mon intention est surtout d'offrir une synthèse et une vulgarisation des aspects clés étroitement liés aux données, tels que l'intelligence artificielle, les méthodes d'analyse et l'intelligence décisionnelle. Je souhaite aussi aborder la transformation digitale des entreprises, concept également très lié aux données mais davantage sous un angle managérial que technologique.

Ce mémoire établi donc un état des lieux du big data, dans le but de saisir son utilité dans la création, la délivrance et la captation de valeur ainsi que l'amélioration continue non linéaire de PME/TPE.

1.1. *Motivations*

Motivée initialement par une récente formation de Busines Analyst, j'ai souhaité combler certains écarts de compréhension par rapport aux attentes du métier et aux opportunités professionnelles autour des données massives. Il m'a donc semblé évident d'un point de vue technique d'assimiler ne fût-ce que la pointe de l'iceberg.

Dans un deuxième temps, en tant que future entrepreneuse, je me suis également intéressée au sens et aux apports du big data au niveau de la différentiation, l'innovation et la croissance, applicables non seulement à mon projet d'entreprenariat dans le secteur FoodTech, mais également à tous les domaines d'activité.

2. Positionnement du problème

Le big data, les technologies et méthodes associées, notamment l'intelligence décisionnelle, l'intelligence artificielle et l'analyse prédictive sont à l'heure actuelle considérées comme les principaux leviers de progrès pour les organisations, peu importe leur nature.

Ce phénomène a autant d'avantages que d'inconvénients, d'ailleurs les petites structures ont souvent été écartées à cause de limitations généralement techniques et budgétaires. On peut donc se demander si c'est encore cela le cas ?

Il faut garder à l'esprit que le big data a beaucoup de potentiel, nombreuses sont les études qui le confirment. Les médias le nomment l'or noir du 21^e siècle, car il représente la matière première et la source de croissance des sociétés technologiques les plus rentables, les GAFAM (Google, Amazon, Facebook, Appel et Microsoft).

Les GAFAM exploitent si brillamment le pouvoir des données, qu'aujourd'hui elles détiennent une position dominante à l'échelle mondiale. Cette situation cause un déséquilibre en matière de concurrence, car leur positionnement fait d'office barrière à l'entrée pour les nouveaux acteurs.

Chaque société peut tirer des conclusions statistiques descriptives à partir de ses propres données en vue d'évaluer ses performances sous forme de KPI (Key Performance Indicator). Ces entreprises peuvent aussi avoir accès à des services tiers gratuits ou payants d'analyse et/ou d'accès aux données. Cependant ce sont les organismes qui disposent de cet or noir, avec la technologie d'analyse, l'infrastructure et l'expertise nécessaires qui tirent le majeur profit.

Grâce aux médias qui nous parlent quotidiennement du sujet, il est plutôt clair que le big data est une mine d'or exploitée principalement par les géants du web. Mais le grand public s'interroge encore sur la raison pour laquelle cette ressource est si utile ainsi que de l'importance des moyens pour pouvoir l'exploiter.

Afin de soutenir ma recherche, une série des questionnements concernant les données massives et la transition digitale en ressortent, parmi certaines de mes interrogations :

Comment les mégadonnées sont-elles créées et exploitées ? Comment se structure la filière big data ? Le big data est-il à la portée des PME/TPE et des non-spécialistes ? Quelle est la véritable valeur du big data pour les entreprises ? Quel est le rôle du big data dans la stratégie d'entreprises ? Quid du positionnement des PME/TPE en Belgique sur ce marché ? Est-ce qu'une micro entreprise ou un indépendant pourrait aussi en profiter ? Quelle est l'utilité de l'analyse prédictive et de l'intelligence artificielle ? Quels sont les enjeux et les risques de l'analyse prédictive et de l'automatisation ? Qu'en est-il des avantages et des inconvénients ? Quels sont les champs d'application et les domaines d'activité concernés ? Quelles sont les perspectives d'évolution quant au traitement des données compte tenu de l'humain et de la législation belge, européenne et internationale ? Comment passer au big data ? Quel est le rôle du big data dans

la transformation digitale ? Quels sont les leviers, les facteurs à échec et à succès de l'innovation digitale ? Quid des critiques ?

Finalement, la question de recherche qu'inclut les questions ci-dessus est la suivante :

Quel est l'intérêt du big data pour les PME/TPE et comment favoriser l'appropriation du big data par ces acteurs ?

L'hypothèse résultant de ma revue de la littérature est la suivante :

Le big data et les technologies associées (les nouvelles capacités d'analyse prévisionnelle, IA, etc.) sont d'un intérêt considérable pour les entreprises en général, cela inclut notamment les PME/TPE et tout entrepreneur, puisqu'elles apportent un avantage compétitif non négligeable au niveau commercial, en plus d'être une source d'inspiration pour les produits et services du futur. La règle du pauvre qui devient de plus en plus pauvre et du riche qui devient de plus en plus riche, est désormais plus aussi délimitée, l'accès et la manipulation du big data se démocratise et permet aux petites structures de s'intégrer elles aussi dans l'ère du numérique.

3. Objectif général

L'objectif général de ce mémoire consiste à tenter de répondre à la question de recherche précédemment mentionnée (cf. supra) afin d'aboutir à une compréhension éclairée de la valeur ajouté des données.

En ce compris l'intérêt des pratiques d'analyse et de transformation des données pour la prise des décisions ponctuelles et stratégiques. Mais aussi, pour la prédiction de modèles et l'automatisation de processus. En plus de l'intégration des technologies et des outils numériques de façon durable au sein et à tous les niveaux de l'organisation.

Concrètement, le but de ce mémoire est de :

Repérer, synthétiser, vulgariser et partager les points clés concernant le volet technique mais aussi le volet business, tels que les opportunités, les enjeux, les freins et les leviers stratégiques du big data et de la transformation digitale. De plus, présenter les moyens d'appropriation les plus efficaces et adéquats pour les PME/TPE.

Ce rapport se veut un manuel abrégé, permettant à tout décideur en entreprise et à toute personne qui débute dans un métier analytique, de comprendre les éléments les plus pertinents autour de l'exploitation des données et d'amélioration continue des entreprises.

4. Fil conducteur

Sous la forme d'une recherche déductive, ce mémoire est constitué de trois parties : (1) Approche théorique ; (2) Approche pratique et (3) Approche analytique.

Partie 1. Approche théorique

Sur base d'informations préexistantes, cette revue de la littérature dépeint globalement le phénomène big data (caractéristiques, business model...), compte tenu des différentes technologies (Hadoop, Kafka, Spark, NoSQL...), des méthodes d'analyse et des modèles associées (intelligence artificielle, analyse prédictive...), ainsi que des outils et infrastructures d'aide à la décision (Business Intelligence).

En outre, cette partie couvre tant les opportunités que les enjeux et les obstacles ; défis qui en dépit des idées préconçues, n'épargnent pas les petites et moyennes structures.

Pour terminer, les étapes à suivre dans un projet big data, ainsi qu'une section dédiée à la transformation digitale sont abordées en fin de ce premier chapitre. Cette dernière inclut les leviers stratégiques, les aides d'accompagnement et de financement, et finalement, les facteurs d'échec et de succès.

Parmi les différentes méthodes de recueil documentaire et d'analyse du même, je distingue :

1. La recherche documentaire des sources secondaires : articles et revues scientifiques, ouvrages, rapports statistiques, enquêtes, conférences, etc.
2. La participation à des évènements organisés par des institutions liées aux Technologies de l'information et de la communication (ITC) tel que Numeria, abordant des thèmes articulés autour du big data, de l'intelligence artificielle et de la transformation digitale des entreprises.

Partie 2. Approche Pratique

En vue d'apporter une perspective pratique, deux enquêtes de terrain ont été réalisées : l'une quantitative de type questionnaire en ligne et l'autre qualitative sous forme d'entretiens téléphoniques/appel vidéo individuels.

L'enquête quantitative était destinée aux PME/TPE, aux indépendants et porteurs de projets. Elle cherchait à mesurer le niveau de familiarisation par rapport aux technologies numériques et le degré de digitalisation de ces structures.

Quant aux entretiens, les interlocuteurs étaient principalement des experts en manipulation des données et en transformation digitale des PME/TPE. Ils ont partagé leur perception concernant les difficultés et les blocages les plus courants au sein de ces structures, ainsi que quelques conseils aux lecteurs afin de mieux entreprendre et de réussir ce genre de projets.

Une description détaillée de la méthodologie pratique est disponible dans la deuxième partie de ce travail (cf. Méthodologie des études de terrain).

Partie 3. Approche Analytique

Cette dernière partie couvre les difficultés rencontrées au cours de ce travail, ainsi que la question de recherche sous un angle analytique. Elle met en relation les deux chapitres précédents pour faire ressortir les ressemblances, les contradictions et de nouvelles pistes d'exploration.

Partie 1

Approche théorique : revue de la littérature

♦ Perspective technique ♦

1. Introduction

Autant convoitées que controversées, les données sont devenues une ressource à part entière.

Bien que, nous entendons régulièrement parler du big data et d'intelligence artificielle, entre autres, que ce soit, par le biais des médias ou sur notre lieu de travail, cela ne sont que de grands mots pour la plupart d'entre nous.

Il est donc indispensable de dédier un peu de temps de lecture aux concepts techniques les plus percutants. Ceci permettra de faciliter la compréhension des lecteurs quant à la finalité de ce travail, c'est-à-dire déterminer la portée que les PME/TPE ont sur le big data, et l'intérêt que les données représentent pour ces acteurs.

Ce premier cadre théorique combine la ***perspective technique*** (définitions, caractéristiques, techniques, méthodes de modélisation et d'analyse) et la ***perspective managériale*** (opportunités, enjeux, obstacles et les leviers stratégiques liées au Big data et à la transformation digitale).

2. Big data

2.1. *Définition*

En fonction des auteurs et des chercheurs ainsi que de leur discipline, contexte ou finalité d'étude, la définition du big data diffère, de manière plus ou moins importante.

Selon Dumbill (2013), les big data ce sont des données tellement grandes, que les systèmes de base de données relationnelles classiques (SGBDR) ne peuvent pas les traiter.

Plus tard, NIST (2014), ajoutera la nécessité d'une architecture puissante et évolutive afin de répondre aux besoins spécifiques de stockage, de manipulation et d'analyse.

D'après les notes de lecture de la revue I2D (2015), Davenport déclare que le big data est une « expression fourre-tout [qui] désigne des données dont la taille ne permet pas de les caser dans les conteneurs habituels » (Davenport, 2014, p.1).

Selon les mêmes notes de lecture, Cointot et Eychenne, définissent le big data comme un concept plus large recouvrant « l'ensemble des technologies, des métiers, des approches conceptuelles permettant d'exploiter l'ensemble des données générées par les hommes de façon consciente ou non et par tous les objets connectés ou non » (Eychenne et Cointot, 2014, p.3).

En réalité, le terme big data existait bien avant ces auteurs, il daterait même de 1997, selon l'Association for Computing Machinery. Mais ce n'est qu'en 2001 que le terme a pris force et s'est universalisé grâce aux travaux et à la granularité de la définition faite par Doug Laney, qui était alors analyste chez Gartner, cabinet américain de conseil et recherche.

Bien qu'il n'existe pas de définition officielle, l'interprétation de Doug Laney figure comme la définition la plus utilisée à l'heure actuelle pour se référer aux mégadonnées. Cette interprétation consiste à résumer le big data en trois piliers qui le caractérisent, appelés aussi les « 3V » : Volume, Variété et Vitesse (cité par Diebold, 2012).

Plus récemment, d'autres auteurs ont ajouté d'autres pilier ou des « V » à la définition de Laney, notamment Dijcks (2012), a rajouté une quatrième V de Valeur, Schroeck et al. (2012) ont rajouté une cinquième V, la Véracité (cité par De Mauro, Greco et Grimaldi, 2016). D'autres travaux plus récents intègrent la variabilité, la viscosité, la volatilité, la viabilité et la visualisation créant ainsi les 10 V (Khan, Alsaqer, Shah, Badsha, Abbasi et Salehian, 2018). Mais l'obsession pour le big data ne s'arrête pas là, Khan, Naim, Hussain, Naveed, Ahmad, Qamar (2019), dans leur étude ont recensé les caractéristiques et les challenges du big data en 51 « V ».

Enfin, d'autres aspects ont également été intégrés au concept, tels que la complexité et la non structure (Intel, 2012) ou encore, l'exhaustivité et le caractère relationnel des données (Kitchin, 2014).

« Certains considèrent finalement que ces distinctives relèvent plus des effets des données massives que de leur nature » (Bastin et Tubaro, 2018, p.9).

En quelques mots, les mégadonnées sont des données multi sources, en très grande quantité (en Pétaoctet équivalent à 1.000.000 de Go et plus), très variées (données structurées, semi-structurées et non-structurées) et qui nécessitent un traitement avancé ainsi qu'une structure de stockage conséquente et des outils adaptés aux types de données à analyser.

Ceci explique la faible performance des systèmes de gestion de bases de données classiques (SGBD), mises à l'épreuve du volume, de la variété de formats et de la vitesse de calcul ; en particulier si les analyses souhaitées concernent des flux de données en temps réel venant de l'IoT (Internet Of Things) comme les capteurs, les senseurs et les compteurs.

2.2. Caractéristiques du big data

Comme cité antérieurement, les caractérisés les plus remarquables du big data basées sur les 3V de Laney (2012), le travail de Khan et al. (2018) et sur d'autres sources complémentaires, comme Agoria (2021) sont :

Volume & Variété : il n'est pas difficile de comprendre que le volume de données accroît de manière exponentielle. Cependant, il est bien plus difficile d'imager l'ampleur que représente cette prolifération des données.

Les causes principales sont la facilité d'accès à internet, et les larges possibilités de stockage à moindre coût (Cloud Storage & Cloud Computing). Un des exemples les plus parlants est la messagerie instantanée qui permet de créer du contenu non-structuré (les messages) et de les stocker le plus souvent gratuitement et sans limite d'espace.

Une autre cause, qui d'ailleurs est devenue la plus significative ces dernières années, c'est l'IoT comme les capteurs, les senseurs et les compteurs, présents dans la domotique, dans les smartphones et dans les wearables tels que les fitness trackers. L'IoT se retrouve aussi dans certaines technologies comme le RFID (Radio-Frequency Identification), fréquemment utilisée pour l'étiquetage des produits dans le retail.

Les QR code, les logs (journal de documentation automatiquement générée et horodatée), les ELD (Electronic Logging Device) souvent utilisés pour les camions routiers, etc., sont aussi des technologies issues de l'IoT.

Les principaux acteurs dans la création des données sont les sociétés, les institutions publiques et les utilisateurs comme vous et moi. En effet, il faut reconnaître qu'aujourd'hui quoi que l'on fasse, nous laissons des traces numériques sans même nous en rendre compte. Par exemple : lorsque nous effectuons de simples requêtes sur un navigateur produisant ainsi un historique de navigation, des cookies ou que nous postons des photos sur les réseaux sociaux.

Mais aussi lorsque nous utilisons des cartes de fidélité des supermarchés. Ces cartes connectées, vont capter des informations clés et générer des données sur des données (métadonnées) comme ce sont les identifiants, la taille d'un fichier, le nombre de caractères, la géolocalisation, les heures d'utilisation, etc.

Par ailleurs, le big data est caractérisé de non-structuré car, contrairement aux informations structurées c'est-à-dire, les données facilement repérables par des requêtes simples, dû au fait qu'elles sont organisées dans des tableaux Excel ou dans des tables relationnelles, les mégadonnées ne sont pas organisées. Par exemple : les messages, les images, les coordonnées GPS, etc.

Ces données issues de sources hétérogènes aboutissent le plus souvent dans des bases de données conçues à cet effet.

Selon un article publié par Statista dans sa revue Digital Economy Compass (2019), le volume des données mondiales est passé de 2 zettaoctets ($Zo = 10^{21}$) en 2010 à près de 50 zettaoctets en 2020. Cette étude prévoit 2142 zettaoctets pour 2035.

Mais, à quoi équivaut un zettaoctet ? A mille milliards de gigaoctets ($Go = 10^9$) ou à un milliard de téraoctet ($To = 10^{12}$). Afin de mieux visualiser, il est bon de savoir qu'un seul terra disponible dans un disque dur communément vendu en magasin, peut contenir jusqu'à 250 000 photos de 12MP, 250 films en HD ou 500 heures vidéo ou encore 6,5 millions de pages A4 numérisées (Dropbox, 2020).

Valeur : c'est la clé du big data, car si elles ne sont pas exploitables, ces données ne valent rien. Au contraire, elles peuvent coûter très cher, à cause des frais de stockage, mais aussi des frais de maintenance et d'infrastructure.

La vision et les objectifs de chaque organisation guideront l'utilisation des données pour la compréhension d'une situation et/ou pour la prise de décisions ; le plus souvent d'ordre commercial, marketing, politique ou scientifique.

En résumé, la valeur est un attribut majeur, car il impacte directement le profit ; qu'elle soit créée par la propre entreprise ou captée d'une source externe. D'une manière primitive, elle se traduit par la monétisation des données qui croisées entre elles révèlent des informations exploitables en interne et/ou commercialisables en externe, ainsi que la mesure du retour sur l'investissement de la mise en œuvre du big data.

Vélocité : les exemples les plus parlants sont Google et les réseaux sociaux. Ces applications sont capables de traiter un flux hallucinant de chargements et téléchargements de données en temps réel. De la même manière, les voitures modernes qui disposent des dizaines de capteurs capables de mesurer à l'instant le niveau d'essence, si une porte est restée ouverte ou encore activer le freinage automatique, etc.

Véracité : plus qu'une caractéristique, c'est un réel challenge. L'intérêt des organisations est de réaliser des analyses ayant un impact durable et positif en fonction de leurs objectifs, mais cela n'est possible que si les résultats d'analyses sont fondés sur des données fiables et de qualité.

Or, l'internet regorge des fausses informations, certaines illégales, voire même dangereuses pour la sécurité des citoyens et des systèmes informatiques privés.

Les données des réseaux sociaux sont notamment regardées d'un œil critique. Les profils fictifs y sont nombreux. La grande difficulté pour les data scientist est donc de détecter le faux du vrai et de certifier les données légitimes. « C'est en croisant les compétences de cybersécurité et de sciences des données que l'on fera émerger des expertises de qualiticien de la donnée. Le qualiticien de la donnée s'appuiera sur des architectures algorithmiques capables d'évaluer en temps réel la légitimité d'une donnée et d'alerter lorsque le faux numérique aura été détecté. » (Teboul et Berthier, 2015).

Viscosité : centrée sur la complexité et la résistance de manipulation des données de masse.

Variabilité : fait référence aux résultats de calcul incohérents. « ...différents types de données similaires (voire identiques) mais dont les résultats peuvent varier d'une expérience à l'autre ou d'une plateforme à l'autre. » « ...les expériences nécessitent d'être reproductibles pour être considérer comme « valides ». (...) Il s'agit donc dans ce genre de cas d'analyser statistiquement la variance. » (Amir-Aslani et Bhajun, 2016).

Statistiquement, une variabilité faible indique que les résultats tendent à se reproduire contrairement à une variabilité forte. Dans ce dernier cas, les résultats ne peuvent pas être intégrés directement, ce qui n'empêche pas pour autant d'être utilisables d'une autre manière.

Volatilité : est fortement liée à la notion du temps, cet attribut se centre sur « ...la durée pendant laquelle les données sont valides et la durée pendant laquelle elles doivent être stockées » (cité par Khan et al., 2018).

Viabilité : c'est la sélection du data conformément aux attributs spécifiques de chaque projet. Les données choisies seront celles considérées comme les plus viables pour prédire et aboutir sur des résultats à valeur ajoutée.

Validité : c'est la compréhension et la vérification des hypothèses et des corrélations entre les éléments d'un ensemble de données.

Visualisation : il s'agit de rendre l'information la plus visuelle possible pour faciliter la compréhension des résultats et la prise de décisions grâce à l'identification des éléments à améliorer, à remplacer ou à implémenter. Cela inclut, notamment le web design, les dashboard, les cartes, les graphiques, etc.

En résumé, la véracité, le niveau de stockage et la vitesse des calculs, sont les enjeux capitaux directement liés au big data. De plus, l'objectif ultime est celui de donner la bonne information au bon moment et via le bon canal.

3. L'analyse des données (Data Analytics)

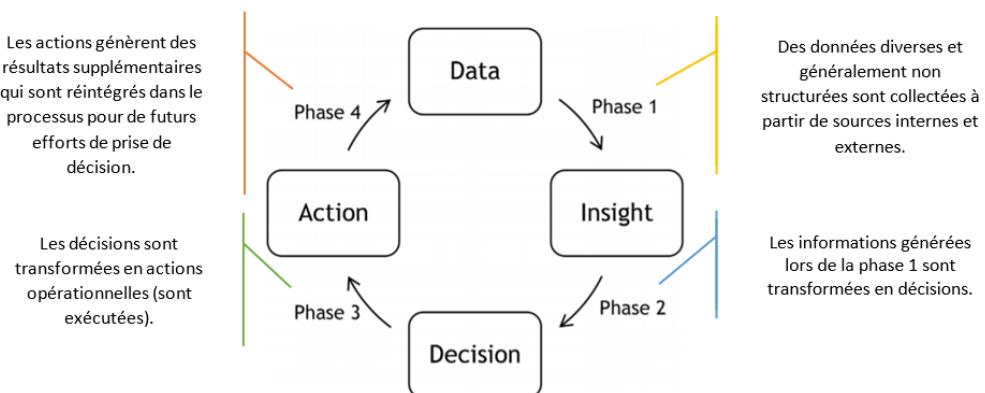


Figure 1 : (Big) data – Cycle itératif d'analyse

Source : Tabesh, P., Mousavardin, E., & Hasani, S. (2019). Implementing big data strategies : A managerial perspective. Business Horizons, 62(3), 347-358.

La notion de valeur, qui figure parmi les caractéristiques du big data, constitue l'objectif principal d'un projet d'analyse des données. Cette valeur est souvent interprétée par le gain ou l'épargne économique. Néanmoins, pour certains auteurs, chef d'entreprise et managers cette notion de valeur peut aller plus loin. Elle peut, entre autres, se traduire par de la satisfaction client, un gain de compétitivité, l'amélioration des partenariats et l'anticipation de comportements.

Plusieurs auteurs, notamment Bénavent (2014) s'accordent sur le fait que les données n'ont de valeur que si elles ont un sens correspondant aux objectifs des projets de l'entreprise.

Cependant, ce sont les objectifs projets et stratégiques de l'entreprise qui donneront un sens aux data et qui permettront de les exploiter efficacement en vue d'obtenir le retour escompté.

Une exploitation efficace des données s'aligne donc avec les finalités d'utilisation de celles-ci et aide à définir les moyens par lesquels les efforts vont être récompensés.

Selon Chen, Roger, Chiang et Storey (2012), l'analyse des données est un procédé où les données extraites à partir de diverses sources sont traitées dans le but de trouver des liens représentatifs entre elles et de les mettre à disposition des parties prenantes (cité par Ponsard, Touzani et Majchrowski, 2018).

3.1. Les types d'analyse de données

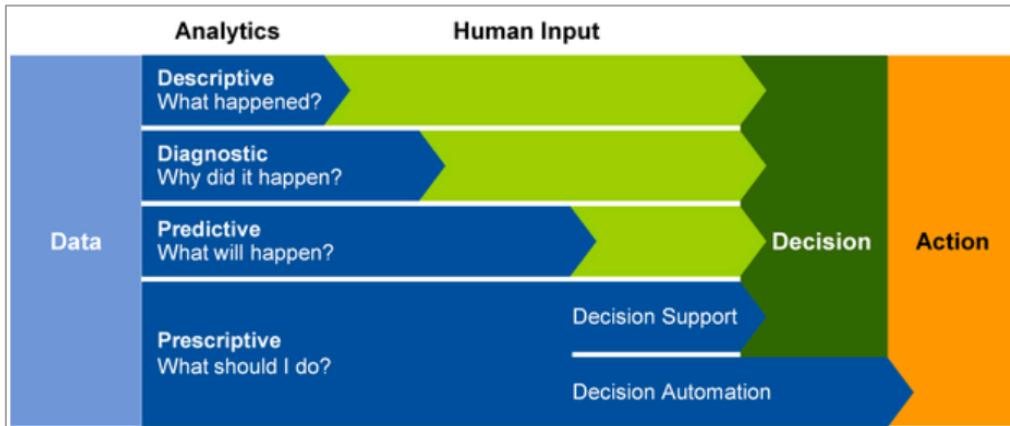


Figure 2 : Types d'analyse de données

Source : Gartner Says Advanced Analytics Is a Top Business Priority. (n.d.). Gartner. Récupéré le 13 avril 2021 de <https://www.gartner.com/en/newsroom/press-releases/2014-10-21-gartner-says-advanced-analytics-is-a-top-business-priority>

Les analyses des données reposent sur une série de méthodes, c'est-à-dire des modèles mathématiques comme les statistiques et des algorithmes qui ne sont rien d'autre que des programmations ou ensemble d'instructions composées des conditions et des ordres précis sur des données.

Quelles que soient la ou les méthodes et les types d'analyses combinées, celles-ci sont appliquées par les organisations en vue de faire un inventaire de leurs activités, de leurs projets et problématiques. Ceci leur permet de mieux connaître leurs performances, leurs faiblesses, mais aussi de saisir des opportunités, faire face aux imprévus et soutenir la prise des décisions (Ponsard et al., 2018).

3.1.1. L'analyse descriptive

Selon Ponsard, Touzani et Majchrowski (2018), l'analyse descriptive s'interroge sur le passé, elle répond à la question : Que s'est-il passé ? mais aussi s'interroge sur les causes. C'est ici qu'entre en jeu l'analyse de diagnostic : Pourquoi est-ce arrivé ?

Les techniques ou méthodes communément utilisées sont : l'analyse statistique, la classification et la catégorisation des données. Ces procédés aboutissent à des visualisations et des rapports historiques.

3.1.2. L'analyse prédictive

L'analyse prédictive s'intéresse à l'avenir. Les questions auxquelles elle répond sont les suivantes : Que va-t-il se passer ? Pourquoi cela risque-t-il de se produire ? Bien qu'elle cherche à prédire le futur, cette analyse est aussi fondée sur l'historique de données. Les approches ou méthodes utilisées sont le forage ou exploration de données (Data Mining) et l'apprentissage automatique (Machine Learning - ML), (Ponsard *et al.*, 2018), concepts que seront repris un peu plus tard dans ce document.

L'analyse prédictive modélise les réponses des situations anticipées, l'idée étant de favoriser ou éviter ces situations. Pour ce faire, des données de qualité et vérales sont nécessaires. En général, les entreprises sont attirées par l'identification des tendances et l'anticipation des problèmes qui leur permet d'augmenter leurs marges ou diminuer les coûts.

Pour Junghans (2015), une prédiction efficace résulte de la combinaison de logiciels réactifs capables de traiter en temps réel des grands volumes de données tant statiques que dynamiques (streaming data) et de trouver des corrélations pertinentes entre les informations. Cette phase analytique est la plus ardue.

Une autre singularité de l'analyse prédictive appliquée au ML est que contrairement à l'analyse clasico-descriptive, ce n'est plus l'humain qui impose sa propre interprétation ; guidé par son intuition, par ses propres compétences, son expérience, son entourage, ou encore la situation au temps présent pouvant introduire des facteurs perturbateurs tels que les sentiments, les émotions et les intérêts personnels.

Ainsi, l'analyse par la prédiction pourrait apporter des solutions plus neutres, fondées sur des données passées et leurs effets corrélatés. Une neutralité tout de même critiquée, notamment dans les domaines éthiquement sensibles comme la justice, la santé, l'emploi, entre autres.

Dans le domaine du Machine Learning, imaginez seulement laisser une machine décider d'engager ou non un candidat sur base des données historiques ou encore la laisser décider du sort d'un accusé. Est-ce que les données ayant alimenté la machine ont été vérales ? Est-ce vraiment juste ?

3.1.3. L'analyse prescriptive

« Elle fournit aux organisations des séquences d'actions opérationnelles adaptatives, automatisées et fiables dans le temps. » (Deshpande, Sharma et Peddoju, 2019, p.75).

Cette technique s'intéresse aussi à l'avenir, à différence qu'elle se focalise sur les recommandations les plus viables pour répondre aux questionnements suivants : Que devrait-il être fait ? Pourquoi cela devrait-il être fait ? Les conseils peuvent porter par exemple sur l'optimisation du business, la simulation des comportements futurs, entre autres. (Ponsard *et al.*, 2018).

En bref, cette analyse cherche à déterminer la meilleure solution ou le meilleur choix, parmi diverses options, compte tenu des paramètres connus.

4. Big data et l'intelligence artificielle (IA)

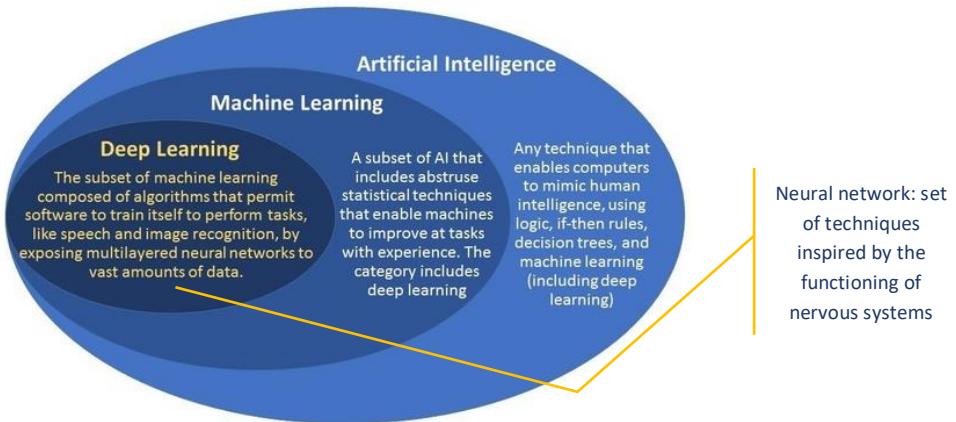


Figure 3 : Les approches de l'intelligence artificielle

Source : Quelles sont les différences entre l'Intelligence Artificielle, le Machine Learning et le Deep Learning ? –

Quora. (n.d.). Récupéré le 16 août 2021 de <https://fr.quora.com/Quelles-sont-les-diff%C3%AArences-entre-Intelligence-Artificielle-le-Machine-Learning-et-le-Deep-Learning#:~:text=Le%20Machine%20Learning%20est%20un,de%20neurones%20dit%20%C2%ABprofonds%C2%BB>.

L'IA est une discipline qui a démarré en 1956 suite à la conférence scientifique de Dartmouth organisée par Marvin Minsky et John McCarthy, considérés depuis comme les pionniers de l'IA.

La conférence mettait en avant certains travaux scientifiques, notamment les publications d'Alan Turing en 1950, exprimant l'idée qu'il est possible de doter les ordinateurs d'intelligence ; le Test de Turing (confronter verbalement l'humain et la machine) et les programmes de jeux aux échecs (homme vs machine) sont ses travaux d'IA les plus notables (Buchanan, 2005).

L'intelligence artificielle ou plutôt les intelligences artificielles au pluriel, puisqu'il en existe plusieurs. En effet, chaque système capable de simuler sous certaines conditions l'esprit humain se distingue par sa capacité d'autonomie à traiter des tâches hyper spécifiques, grâce à des méthodes et ordres algorithmiques. La machine n'est donc pas si intelligente que ça, car elle reproduit certes d'une manière très efficace des actions précises qui dépassent souvent les actions de l'homme mais c'est bien ce dernier qui établit les règles.

Par conséquent, la croyance que l'intelligence artificielle remplace l'humain est inexacte, du moins jusqu'à présent l'IA remplace des tâches répétitives et chronophages qui font partie d'un monde qui évolue et se transforme. L'homme n'étant pas un spécialiste par nature, contrairement aux machines, il s'adapte et des nouveaux métiers émergent (Julia, 2019).

Voici un aperçu de deux sous-disciplines de l'IA très utilisées grâce aux big data : le Machine Learning et le Deep Learning.

4.1. Machine Learning (ML) ou apprentissage automatique

« ...un jeune enfant n'a pas besoin de connaître les règles : il apprend seul, à partir de ses expériences. » (Georges 2019, p.14).

Le ML est « le domaine d'étude qui donne aux ordinateurs la capacité d'apprendre sans être explicitement programmés » (Arthur Samuel, 1959).

« Machine Learning est l'application de logiciels et de matériel informatique pour reconnaître des modèles et développer des actions pour répondre à ces modèles. » (Pries et Dunningan, 2015).

D'après Pries et Dunningan (2015), c'est cette capacité à reconnaître de modèles qui donne toute la richesse à cette technologie. Ainsi, des patrons d'intérêt vont pouvoir être sélectionnés et programmés de manière à fonctionner en un temps record et en toute autonomie.

D'une manière simple, au lieu de donner des instructions explicites, parfois presque infinies, pour résoudre un problème, indépendamment de son degré de complexité, des algorithmes vont permettre au programme de base à apprendre à partir des exemples (jeux de données ou datasets), et ainsi accomplir une tâche déterminée ou résoudre un problème (Georges, 2019).

Pour l'humain reconnaître un visage est aisément naturel. Mais pour un ordinateur, un problème même élémentaire, peut être très difficile. En effet écrire un algorithme capable de lui faire reconnaître des formes complexes se révèle extrêmement compliqué. Dans ce contexte, le ML devient un atout majeur et la raison pour laquelle cette technique est omniprésente dans le monde moderne.

Concrètement, comment un ordinateur peut apprendre à conduire une voiture, à identifier des maladies ou à nous donner des prédictions quant aux nombres de ventes espérées ?

D'après l'article écrit par Bheemaiah, Esposito et Tse (2021) et la plateforme DataScientest (2021), les ordinateurs apprennent à apprendre grâce aux algorithmes et aux modèles suivants :

Apprentissage supervisé : Arbre de décision, Régression linéaire, Support Vector Machine...

L'ordinateur reçoit une question, une réponse attendue et des exemples (data) déjà catégorisées sur lesquels l'ordinateur va s'entraîner, pour ensuite faire des prédictions.

Exemple : sur un programme statistique open source comme RStudio ou un logiciel intuitif comme Anatella, l'on peut analyser le taux d'attrition (churn ou proportion de clients/abonnées perdus) et grâce au modèle d'arbre de décisions prédire les clients qui en fonction des variables choisies risqueraient de partir. Le but est d'identifier proactivement ces clients pour tenter de les convaincre de rester.

Apprentissage non supervisé : Clustering...

L'ordinateur ne reçoit pas de réponse espérée, il va résoudre seul le problème en identifiant des similitudes entre les données pour ensuite classer et regrouper celles qui partagent des caractéristiques communes. Fréquemment utilisée en marketing pour la segmentation client et le profilage, aussi pour la détection d'anomalies comme les fraudes, les maladies, etc.

Dans tous les cas des variables d'entrée « x » sont nécessaires, par exemple pour l'analyse client, les variables peuvent être l'âge, le sexe et le revenu, dans l'apprentissage supervisé la variable « y » est connue et les variables « x » sont déjà catégorisées. On connaît donc dès le départ l'information que l'on veut obtenir. Par contre pour l'apprentissage non supervisé, la variable « y » est inconnue et l'ordinateur se chargera de classifier les variables « x ».

Apprentissage par renforcement : Réseau de neurones...

Sur un principe de récompense, l'ordinateur apprend de ses propres erreurs par rapport à un environnement donné et les évitera à l'avenir. Par exemple, dans le cas de la voiture connectée, une des actions possibles est de tourner à gauche, mais si la voiture tourne alors qu'il y a un mur celle-ci s'écrase. La voiture apprend donc qu'il ne faut pas tourner lorsqu'il y a un obstacle.

4.2. Deep Learning (DL) ou apprentissage profond

Dans son article, Georges (2014) souligne qu'en plus du big data grandissant et les méthodes de calcul améliorées, tant en termes de vitesse que de coût, les modèles statistiques et mathématiques du Machine Learning se modifient et s'améliorent jusqu'au moment où le système aboutit au résultat optimal.

Ceci se compare avec les interconnections neuronales du cerveau et c'est d'ailleurs la source d'inspiration des chercheurs Yann LeCun, Yoshua Bengio et Geoffrey Hinton qui ont travaillé sur la technique de réseaux de neurones (neural network) pour développer la leur, nommée Deep Learning.

Cette branche du ML s'impose et est utilisée pour la reconnaissance de sons, d'images, de l'écriture manuscrite, mais aussi pour la détection d'objets dans une image, pour la traduction, etc. (Georges, 2014).

La différence entre le modèle réseaux de neurones du DL et les modèles présents dans le ML, est que le ML travaille avec des données structurées et semi-structurées (Excel, CVS, XML, JSON...) contrairement au DL. De plus les réseaux de neurones du DL sont composés de plusieurs fonctions qui communiquent entre elles pour aboutir à un résultat, alors que les méthodes ML appliquent une seule fonction par résultat. Ces fonctions en analogie avec le cerveau viendraient être des signaux électriques déclenchés par nos sens et nos connaissances au sein du système nerveux (NeurIPS, 2007 ; Zhang, 2019).

5. Donner du sens aux données

Trois métiers principalement sont chargés de faire parler les données, chacun nécessitant des compétences et méthodes particulières.

Le graphique ci-dessous montre le degré ascendant de complexité dans l'analyse données.

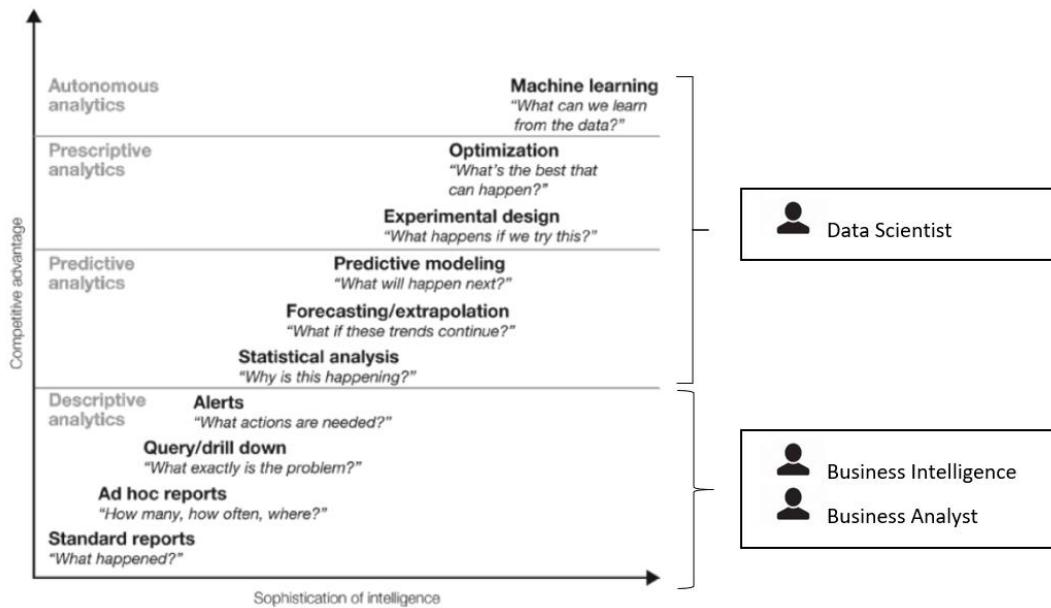


Figure 4 : Degré d'intelligence

Source : Davenport, T., & Harris, J. (2017). *Competing on Analytics : Updated, with a New Introduction: The New Science of Winning*. Harvard Business Press. Récupéré de [https://books.google.be/books?hl=fr&lr=&id=LW9GDqAAQBAJ&oi=fnd&pg=PT20&dq=intelligence+degre+Davenport,+T.,+%26+Harris,+J.+\(2017\).&ots=d0j3e0M_rS&sig=D4RuhvzEUPLVA5A_qSq64TsR0M8&redir_esc=y#v=onepage&q=&f=false](https://books.google.be/books?hl=fr&lr=&id=LW9GDqAAQBAJ&oi=fnd&pg=PT20&dq=intelligence+degre+Davenport,+T.,+%26+Harris,+J.+(2017).&ots=d0j3e0M_rS&sig=D4RuhvzEUPLVA5A_qSq64TsR0M8&redir_esc=y#v=onepage&q=&f=false)

5.1. Business Intelligence (BI) ou informatique décisionnelle

La BI est l'ensemble de moyens et des méthodes par lesquels les données sont collectées depuis des sources hétérogènes, et ensuite stockées et analysées. Une fois les données transformées en informations, ces dernières seront transmises en vue d'aider à la prise des décisions (Negash et Grey 2008 ; Chen et al. 2012).

Selon les mêmes auteurs, la gestion de données et des bases de données constituent les fondations de la BI Analytique. Applicable et pertinente pour l'ensemble de l'entreprise, peu importe sa taille, elle apporte une visualisation éclairée des indicateurs de performance (KPI).

Elle permet donc de mieux appréhender les différentes problématiques au sein des l'entreprise et de fixer des objectifs et des stratégies pour les corriger.

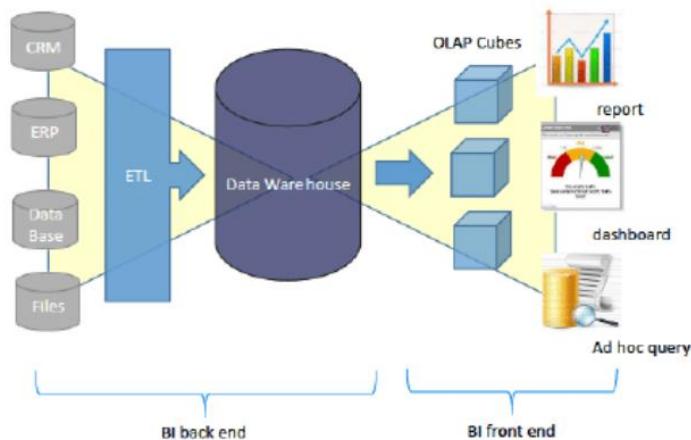


Figure 5 : Architecture Business Intelligence

Source : Habibu, T. (2013). Parallel Data Analytics for Business Intelligence Real-Time Online Analytical Processing (OLAP) For Multi-Core and Cloud Architectures. Récupéré de https://www.researchgate.net/publication/319458909_Parallel_Data_Analytics_for_Business_Intelligence_Real-Time_Online_Analytical_Processing_O LAP_For_Multi-Core_and_Cloud_Architectures

En effet les analystes BI, en fonction des exigences métier, vont extraire les données de plusieurs data base (internes et externes), pour ensuite les centraliser et les regrouper de manière logique et classifiée (modélisation) au sein d'une autre data base (data warehouse). Ce procédé est nommé ETL ou Extract-Transform-Load.

Une fois les données nettoyées, des requêtes SQL propres à des SGBD relationnelles vont être lancées, les résultats se présentent ensuite sous forme de tableaux de bord (dashboards) et de rapports. Ce processus permet aux entreprises de mieux comprendre leurs processus métier, les points sur lesquels ils sont performants et les points qu'il faudrait améliorer.

Aujourd’hui nous en sommes à la troisième génération BI. Elle s’applique aussi aux données massives et à une grande variété de silos de contenu créés par les différents métiers ou départements au sein d’une même entreprise. Par conséquent, les données à gérer peuvent provenir des sources et silos différents (réseaux sociaux, Google Analytics, plusieurs solutions SaaS, ERP, RH, marketing, production, logistique, etc.) et sous formats distincts (photo, vidéo, JSON, CSV, XML, etc.). Cette nouvelle génération de systèmes BI cherche à être plus intuitive pour faciliter la montée en compétences autour du traitement, l’analyse et la compréhension de données, c’est-à-dire la Data Literacy (Qlik, 2017/2018).

5.2. Data Mining (DM) ou forage/exploration/fouille de données

Le Data Mining ainsi que les différents types d’analyses évoqués plus tôt, est une discipline qui s’intègre à la séquence d’étapes de la BI, pour l’aider à transmettre des informations pertinentes.

Dans ce processus c’est le Data Scientist qui prend le relais. Sa fonction consiste à explorer les données en profondeur. Pour cela les étapes de nettoyage et de préparation de données sont essentielles, sauf lorsqu’il s’agit d’appliquer des méthodes propres au big data avec des données non structurées comme c’est le cas du Deep Learning.

La DM à différence de la BI, travaille avec des données volumineuses, applique des algorithmes, c’est-à-dire que s’appuie sur les techniques d’IA et vise à détecter des corrélations et modèles (patterns) pour établir des prédictions. Elle travaille donc avec une complexité d’analyse supérieure et des outils plus performants (Burquier, 2008 ; Fernandez, 2013).

Il est inévitable de remarquer des similitudes avec l’analyse prédictive. En effet, cette dernière est une forme de DM.

Certains des algorithmes utilisés sont : association, analyse de séquence, classification, clustering, prédition (Fernandez, 2013).

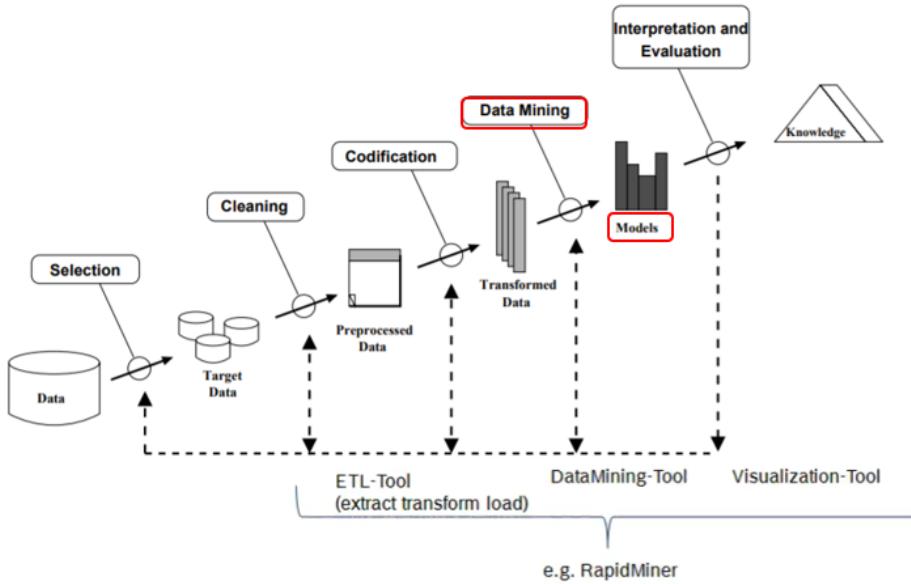


Figure 6 : Etapes du Data Mining

Source : Marbán, O., de Amescua, A., Cuadrado, J., & García, L. (2002, janvier 1). Cost Drivers of a Parametric Cost Estimation Model for Data Mining Projects (DMCOMO). Récupéré de https://www.researchgate.net/publication/220958274_Cost_Drivers_of_a_Parametric_Cost_Estimation_Model_for_Data_Mining_Projects_DMCOMO

5.3. Data Visualisation

« La représentation et la présentation des données pour en faciliter la compréhension. » (Kirk, 2016).

Cette discipline est un sujet brulant, l'augmentation dramatique des données exige une représentation utile des informations, un histogramme illisible avec des milliers des barres n'aurait aucun sens. Des outils très performants permettent aujourd'hui de visualiser des informations de manière propre, dynamique et même en temps réel. A titre d'exemple, Google Studio est un outil BI freemium assez complet.

6. Data Pipeline

Data Pipeline Patterns

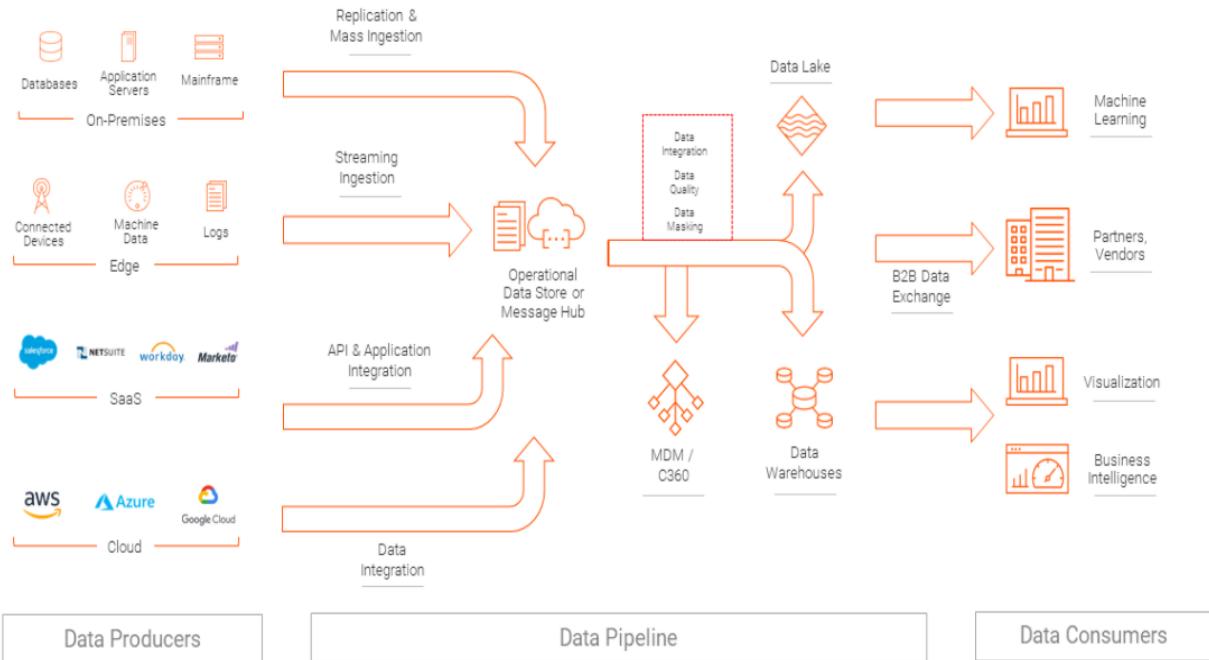


Figure 7 : Data Pipeline

Source : Darshan, J. (2019, August 20). Data Processing Pipelines: An Explainer with Examples. The Informatica Blog - Perspectives for the Data Ready Enterprise. Récupéré le 16 août 2021 de <https://blogs.informatica.com/2019/08/20/data-processing-pipeline-patterns/>

Le pipeline de données correspond à la séquence de processus à laquelle les données sont soumises. Divisée globalement en trois dimensions : Data Producers, Data Pipeline et Data Consumers.

Data Producers (collecte de données) : les données peuvent être variées, provenir de différentes sources et peuvent être issues de sources internes ou externes. Cette production est constituée de deux types de données : premièrement, les fichiers structurels classiques comme Excel, CSV, EDI (Electronic Data Interchange), SQL, ERP ou tout autre fichier en provenance d'une quelconque base de données. Ce type d'intrant est appelé Batch Data. La seconde catégorie regroupe les données en temps réel fournies par des senseurs et des objets connectés sous forme des messages. Ce sont les Streaming Data.

Data Pipeline (flux de données) : Processing, Data Cleaning et Data Gouvernance. Le choix du Data Pipeline dépendra des besoins et des objectifs organisationnels. Traditionnellement, les

données sont stockées en local (ordinateur privé), sur un serveur interne ou encore placées dans une centrale (Data Center), option destinée à une époque qu'aux plus grandes sociétés.

Avec la croissance des données, les serveurs étaient vite limités en stockage ralentissant le traitement des informations. Suite à ces contraintes les Data Wharehouse et les ETL (Extract-Transform-Load) ont émergés, et font maintenant partie du Pipeline.

En effet, en fonction du type de données à traiter, une architecture spécifique sera nécessaire. Pour les Batch Data, un outil ETL suffit par contre, pour les Streaming Data, il faudra utiliser un outil de traitement de messages. Dans le cas où il est nécessaire de traiter les deux types de données en même temps, une architecture Lambda est la solution préférée.

En résumé, le Pipeline montre le flux des manipulations des données, le but consiste à respecter le principe de véracité, c'est-à-dire obtenir des données de qualité qui seront ensuite utilisées pour les analyses, pour l'intelligence artificielle, pour le Machine Learning et bien sûr pour le reporting.

Pour obtenir des données de qualité, il faut donc rassembler, explorer, nettoyer, standardiser et intégrer les données. Ensuite, elles doivent être centralisées dans un ODS (base de données opérationnelle) s'il s'agit d'un système ETL, dans un Message Hub pour le système de Streaming, ou aller plus loin dans l'amélioration de ces données et les stocker dans un Data Warehouse.

Dans le cas d'une exploitation ultérieure, une solution pourrait être de les stocker dans un Data Lake, mais qui fonctionne à l'inverse, suivant un modèle ELT (Extract-Load-Transform).

Data Consumers (analyse de données) : c'est la phase où les données sont consommées, c'est-à-dire modélisées, analysées et interprétées. A ce stade, des modèles statistiques, d'apprentissage automatique, de prédiction, de reporting et de visualisation sont réglés, testés et déployés pour être exécutés en temps réel ou par lots via des automatismes programmés.

Il est évident qu'un pipeline d'informations n'est pas nécessaire pour chaque entreprise, le choix d'en instaurer un est défini par le nombre et la variété de données à traiter, plus ils sont importants et hétérogènes, plus il est utile d'avoir une architecture nette, afin d'éviter les silos d'informations et au contraire les classer pour mieux les exploiter.

7. Ecosystème big data

De toute évidence, notre manière de créer et consommer de l'information a véritablement changé grâce à la technologie, autour de laquelle un écosystème s'est développé. Celui-ci est composé d'une série d'acteurs qui interagissent entre eux pour nous délivrer des supports numériques et des infrastructures capables de faire face aux challenges du big data, particulièrement en termes de volume, variété et vitesse (Oliveira et Loscio, 2018).

Hadoop est la principale plateforme big data open source. Crée initialement pour accélérer les résultats de moteurs de recherche, elle est par la suite venue au secours de l'immensité de

données et du défaut de puissance de calcul de serveurs classiques. Ces derniers devaient en permanence être augmentés en capacité physique (RAM, espace de stockage, processeurs, etc.). Hadoop répond donc aux besoins de volume et de vélocité.

Il existe une autre solution, plutôt sur le plan du traitement des données non structurées et en temps réel (variété), ce sont les bases de données non relationnelles NoSQL.

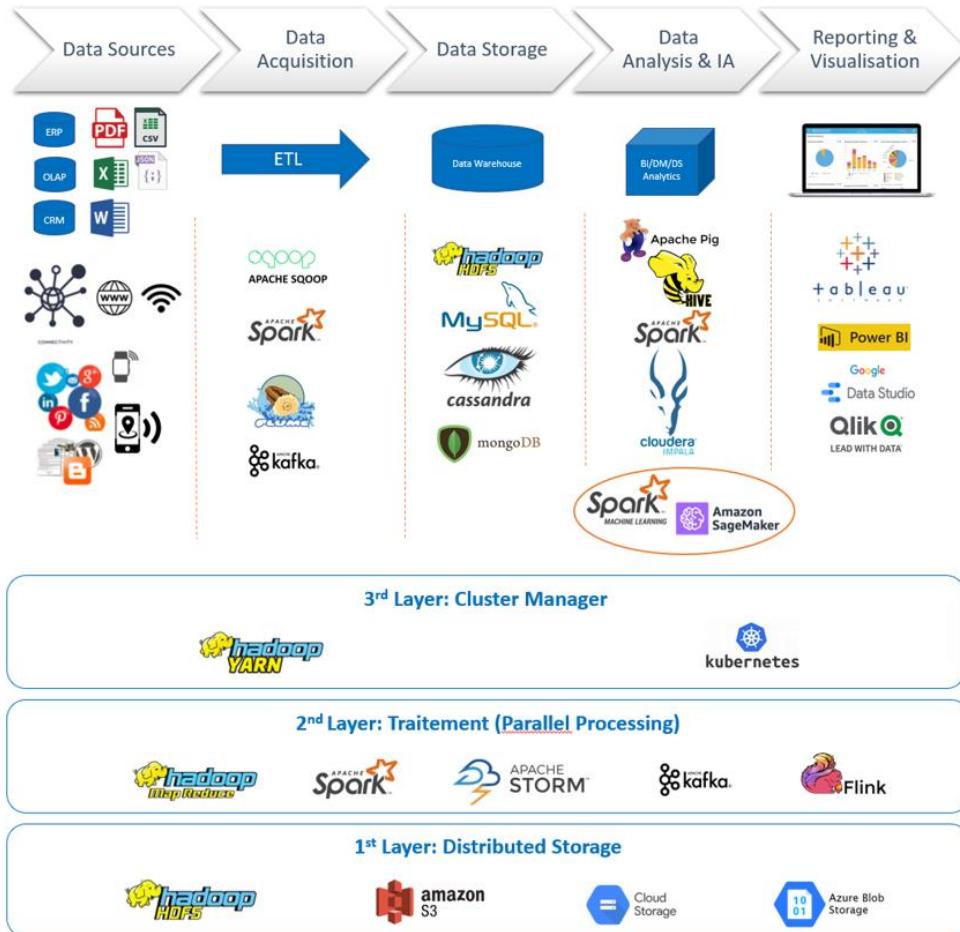


Figure 8 : Big Data (Hadoop) ecosystem

Source : Fernanda Cortés. Illustration inspirée des travaux de Mazumdar et Dhar (2015) : Hadoop as Big Data Operating System--The Emerging Approach for Managing Challenges of Enterprise Big Data Platform.

Cette figure illustre dans les grandes lignes l'écosystème big data ou Hadoop. Bien entendu, c'est un exemple assez réducteur de l'ensemble des acteurs mais elle reprend les plus connus.

Pour comprendre comment l'écosystème fonctionne, il faut passer en revue les trois modules (couches) principaux le composant :

HDFS (Hadoop Distributed File System) : il s'agit de la première étape, une fois que données ont intégré l'écosystème grâce à des connecteurs, les données sont divisées et stockées en petits blocs. Chaque bloc est distribué dans un serveur différent (nœud ou DataNode du cluster). Les blocs de données sont protégés contre les défaillances hardware, si l'un tombe en panne, le bloc sera automatiquement dupliqué et redistribué vers un autre serveur (DataNode), cette fonctionnalité fait d'Hadoop un système « fault-tolerant » (Mazumdar et Dhar, 2015 ; DataScientest, 2021).

Hadoop MapReduce (Parallel Processing) : après le stockage, place au traitement ultra rapide grâce au traitement d'ensembles de données en parallèle. Cette méthode permet de partager les tâches de calcul entre plusieurs ordinateurs puisqu'un seul ne serait pas capable de gérer le volume de données. Comme les données ont été divisées, chaque DataNode va s'occuper d'un bloc, puis les résultats individuels vont être assemblés pour obtenir un résultat final (Mazumdar et al., 2015 ; DataScientest, 2021).

Hadoop YARN (Yet Another Resource Negotiator) : permet de gérer les ressources des clusters. Les ressources étant : processeurs CPU, mémoire vive RAM, processeur graphique GPU, disques durs, bande passante réseau, etc. Ceci permet de réduire les coûts puisqu'ils sont partagés et d'assurer la continuité du système (Mazumdar et al., 2015 ; DataScientest, 2021).

Au-delà de ces trois modules d'autres se sont ajoutés, certains créés aussi par la communauté Apache.

Services d'ingestion ou Connecteurs ETL :

- **Sqoop** : pour les données chargées en lots ou en paquets entre bases de données relationnelles et Hadoop.
 - **Chargement des données par lots (batch)** : mode traditionnel de l'approche ETL. Il s'agit de charger un grand volume de données en une fois. Le traitement et l'analyse de données ne sera possible qu'une fois chargées. Il faut donc tenir compte du délai entre le chargement et la disponibilité des données. Les lots se chargent automatiquement les uns à la suite des autres, selon les paramètres choisis, exemple : une fois par semaine, le lundi à 22 heures.
Souvent utilisé pour des fichiers plats, de texte, etc., comme les factures, les commandes, mettre à jour la base de données, etc. (Hub-Charleroi, 2021).
- **Flume et Kafka** : permettent d'intégrer des données non structurées en temps réel (fichier log ou streaming data).
 - **Chargement des données en temps réel (streaming data)** : ce mode traite les données en continu. Les systèmes ne doivent pas conserver toutes les données. Utile pour les données venant des senseurs et d'algorithmes permettant détecter les fraudes, les attaques numériques, le niveau du stock, les infos des réseaux sociaux, etc.

- **Flink et Apache Storm** : aussi pour des fichiers log.
- **Spark** : gère les deux types, batch et streaming data. La solution la plus rapide à l'heure actuelle (Hub-Charleroi, 2021).
- En effet, Spark est une plateforme composée d'un écosystème de composants Spark. Similaire à Hadoop DDFS, ce module de calcul distribué travaille en mémoire (in-memory) c'est-à-dire sur la RAM., ce qui la rend dix fois plus rapide. Spark est aussi fault-tolerant et nécessite moins de puissance de calcul (donc moins d'infrastructure). Spark solutionne l'incompatibilité de MapReduce avec les calculs itératifs et interactifs, communément utilisés par les algorithmes d'apprentissage statistique (Chokogoue, 2018, p.58).
Elle est composée de différents modules : Spark SQL pour les requêtes, Spark Streaming connecteur des fichiers log, SparkMLlib pour le Machine Learning, GraphX pour analyser et exécuter des requêtes sur un graphique. Les langages de programmation supportés sont Scala, Java, Python, R et SQL (Chokogoue, 2018, p.58).

Stockage de données :

- **Hadoop HDFS, MySQL en SQL** : ODS (Operational Data Store) pour les données en batch.
- **Cassandra, MongoDB et HBase en NoSQL** : Message Hub pour les données en temps réel (streaming data).

Traitement et analyse via requêtes SQL :

- **Pig, Hive**
- **Spark SQL et Cloudera** : de faible latence, capables de traiter de centaines de requêtes en millisecondes.

Services pour l'IA et les prévisions (ML et DL)

- **SparkML, AmazonSageMaker...**

Logiciels de reporting et visualisation

- **Tableaux, Power BI, Google Studio, Qlik...**

En résumé, Hadoop est un excellent environnement de stockage à moindre coût et rapide qui permet de tirer profit de ses données.

♦ Perspective Managériale ♦

1. La révolution des données qui transforme l'entreprenariat et les entreprises

Les PME (micro, petites et moyennes entreprises de 1 à 249 personnes) sont au cœur de notre économie. Selon un rapport publié par Agoria Wallonie (2019), 70% de l'emploi dans le secteur privé est généré par les PME, lesquelles représentent 99% des entreprises en Belgique.

De plus, depuis la dernière réforme du Code des sociétés et associations entrée en vigueur le 1^{er} mai 2019, le nombre de nouvelles entreprises créées dans l'année qui suit la réforme a augmenté de 31,4%. Les nouvelles entreprises ont été principalement créées sous le statut SRL, lequel a été assoupli (notaire.be, 2020).

La même tendance a été enregistrée entre le 1^{er} mai 2020 et le 30 avril 2021 à hauteur de +15,1% (notaire.be, 2020).

En outre, le paysage techno-numérique se développe fortement, avec l'e-mail dans les années 80, l'émergence des progiciels de gestion (CRM, ERP, des outils BI d'analyse...) dans la décennie suivante et plus récemment l'IoT, l'IA, etc. Il semble évident que cette tendance digitoo-analytique ne risque pas de s'arrêter. Elle aura donc inévitablement un impact sur la société et les organisations indépendamment de leur taille et de leur industrie (Paschek, Luminosu et Draghici, 2017).

De ce fait, l'intégration des systèmes d'information (big data, intelligence artificielle, cloud, réseaux sociaux, cybersécurité, business intelligence, etc.) sont utilisés dans le but de gagner en compétitivité, développer une offre plus adaptée, accroître et sécuriser son business, tout en réduisant les limites temporelles et spatiales.

L'Indice d'intensité digitale (IID) qui mesure le taux de pénétration et de vitesse d'adoption de ces technologies est net : plus les investissements dans ce domaine sont importants, plus les entreprises sont considérées innovantes et résilientes (European Commission, 2020).

D'un autre côté, cet engouement pour les données et les technologies s'explique en partie, comme déjà évoqué, par l'utilisation effrénée de l'internet et des objets connectés. En 2019, 85% des citoyens européens utilisent l'internet pour toute une série d'activités et de transactions y compris commerciales. L'utilisation des services internet et des services en ligne augmentent eux aussi d'année en année (European Commission, 2020).

Selon l'étude « Digital is a mindset » réalisée par Agoria (2019), la Belgique a fait des progrès intéressants et se positionne dans la moyenne européenne. Notre pays reste tout de même, à la traîne par rapport à d'autres états notamment les pays nordiques et les Pays-Bas. De plus, cette étude a révélé que 45% des PME investissent dans le digital pour être plus efficaces, 30% pour répondre aux exigences des clients et 25% pour être plus compétitives.

De même, l'intérêt pour l'innovation digitale des 400 interviewés est important, soit huit PME sur dix. Mais étonnement, pour plus de la moitié d'entre elles, l'exploitation des données au profit de l'innovation et de l'optimisation des processus ne figure pas parmi leurs priorités actuelles (Agoria, 2019).

Le big data et l'IA sont les principaux protagonistes de cette révolution 4.0. Environ 20% des entreprises belges se servent de l'analyse des données massives comme levier à la prise de décisions à court et moyen terme. Dans l'UE, 33% des grandes entreprises utilisent le big data contre 12% pour les PME (EPRI, 2019 ; European Commission, 2020).

Parmi les PME industrielles belges, seulement 12% des entreprises ont des business modèles propres de l'ère digitale (abonnement et market place), 13% font usage ou sont en préparation d'utiliser l'IA, 14% ont des projets IA en cours, mais ça veut dire que 73% des entreprises industrielles belges n'en profitent pas encore des opportunités et des capacités d'innovation de l'IA. Globalement, 80% des industriels belges n'envisagent pas s'appliquer dans l'IA à court terme (Agoria, 2019).

Le Baromètre de l'Agence du numérique (AdN-Digital Wallonia), se base sur quatre axes d'analyse chaque année : infrastructure numérique ; organisation et collaboration ; digitalisation des processus et stratégie numérique de l'entreprise. Corrélé avec les études menés par Agoria, tous deux confirment que les entreprises belges progressent.

Le Baromètre a constaté que les firmes flamandes et wallonnes sont quasi équivalentes au niveau de la maturité numérique, bien que la capitale les devance, principalement dans les secteurs de services (Agoria, 2019 ; AdN 2020).

Finalement, la principale faiblesse pointée du doigt par ces organismes interprofessionnels et par les sociétés professionnelles spécialistes du numérique est la culture d'entreprise, encore plutôt précaire au sein des PME (Agoria, 2019).

2. Les modèles d'affaires et de revenus de l'ère numérique

L'internet, a permis aux entreprises de s'adresser à une plus large audience et de ne pas se contenter de sa zone de chalandise habituelle. La technologie associée à l'internet et aux données des utilisateurs/clients a changé le monde des affaires.

La monétisation des données a fait émerger des business modèles novateurs tout en forçant les anciens à s'adapter au paradigme digital, mêlant numérique et tradition.

Les caractéristiques du big data nous rappellent souvent que la valeur fait figure d'objectif ultime poussant les entreprises à entreprendre des projets de digitalisation, d'analyse des données et de robotisation.

Il est dès lors intéressant de connaître les modèles d'affaires et de revenus liés au phénomène big data.

Le modèle de revenus à ne pas confondre avec le modèle d'affaires, également appelé business model (BM) ou modèle économique, ce dernier est le terme le plus couramment utilisé pour exprimer la ou les sources de revenus d'une entreprise ou de ses projets. Il est néanmoins important de nuancer, car en réalité le BM peut combiner plusieurs modèles de revenus et se développe en tenant en compte d'autres facteurs tels que l'offre proposée, les moyens de promotion et de distribution ainsi que les différents coûts liés à l'exploitation de la ou des activités.

D'ailleurs, Lecocq, Demil et Warnier (2006), ont déterminé les éléments qui selon eux constituent le business model d'une entreprise et de ses projets, ils distinguent : (1) les ressources et compétences déployées ; (2) le produit/service en question ; (3) l'organisation nécessaire et le choix des activités à mettre en place ; et (4) la marge résultante des revenus et les dépenses liées à l'activité ou projet, autrement dit le modèle de revenus.

Les sociétés les plus rentables ont créé des modèles d'affaires disruptifs. Elles sont en permanence à l'affût de nouvelles sources de croissance et de gain de compétitivité.

Pour comprendre comment certaines sociétés ont réussi à tirer parti des données qu'elles génèrent mais aussi que d'autres engendrent, il faut revenir à la base, et définir la notion de business model.

Goethals (2011) a cité dans l'un de ses articles la définition de business model comme « *the firm's logic for creating and commercializing value* » (cité par Osterwalder, Pigneur et Tucci, 2005). Ce qui en français se traduit par la logique d'une entreprise à créer et à commercialiser de la valeur.

Le big data naît justement de sa puissante création de profit, telle que mentionné précédemment. La valeur se concrétise par la monétisation des données en interne et/ou suite à leur vente ainsi que par le retour sur l'investissement de la mise en œuvre des stratégies big data.

Historiquement, la création de valeur vient du fait de répondre à un besoin, au bon moment et par le bon canal, afin de créer de l'engagement par le consommateur/client ou de générer un retour positif en vue d'une fidélité pour le service/produit offert.

Paul Timmers (1988), suite à un travail de recherche au sein de la Commission Européenne, introduisit un cadre regroupant onze modèles d'affaires associés au boom de l'internet : *E-shop* ; *E-enchères* ; *Marketplace* ; *Communauté virtuelle* ; *Fournisseur de services de chaîne de valeurs* (par exemple Fedex et UPS qui offrent des services de logistique, de paiements, etc.) ; *Plateforme de collaboration* (Teams par Microsoft...) ; *Courtage d'informations* (Google, Firefox, Wikipédia...) ; entre autres.

Au début du 21^e siècle, Rappa (2001) établissait une classification affinée des nouveaux business model de l'ère digitale. Ces nouveaux modèles, tout comme pour les modèles décrits par Timmers, se distinguent des classiques par leur orientation web. Selon Zott, Amit et Massa

(2011), cela inclue les commerces dont l'activité se déroule principalement par voie électronique (cité par Soumaya et Jaussaud, 2018).

Les neuf catégories de business model identifiés par Rappa (2001), par lesquels les données, les produits et les services en ligne sont monétisés, sont les suivants :

Tableau 1: Modèles de revenus de l'ère numérique - Crédit et captation de valeur

1. Modèle de courtage	Les courtiers sont des intermédiaires qui moyennant une commission ou un forfait pré-déterminé, mettent en relation l'offre et la demande et facilitent les échanges entre vendeurs et acheteurs. Types : <ul style="list-style-type: none"> - Courtier de ventes aux enchères : eBay... - Courtier des transactions financières : PayPal... - Agent de recherche : Skyscanner, Kayak... - Marketplace virtuel : Amazon...
2. Modèle publicitaire	Pour atteindre une rentabilité suffisante uniquement avec la publicité en ligne, il faut une audience et un volume de trafic importants et très ciblés. Ce modèle profite de l'espace de pages web pour y placer des bannières publicitaires, pop-up, liens d'enregistrements, etc. Types : <ul style="list-style-type: none"> - Portail ou moteur de recherche : Yahoo, Google... - Classifieds : Craigslist... (sites dédiés aux petites annonces) - User registration : Researchgate... (l'utilisateur doit s'inscrire gratuitement pour accéder à l'information) - Query-based Paid Placement : Google Adwords (positionnement privilégié suite à une requête) - Publicité contextuelle : AliExpress, Facebook... (annonces ciblées en fonction du comportement de navigation de l'utilisateur) - Publicité liée au contenu de page web : un blog de beauté affichera des publicités de produits de beauté
3. Modèle infomédiaire	Ce sont des intermédiaires qui utilisent l'information (base de données) comme moyen de relation entre vendeurs et acheteurs potentiels ou comme source d'analyse d'un marché donné. Types : <ul style="list-style-type: none"> - Réseaux publicitaires : des sites membres qui enregistrent des données sur l'utilisateur - Mesure d'audience : Nielsen, Bisnode... (agences d'études) - Marketing incitatif : MyShopi... (axé sur la fidélisation client via des coupons, des points échangeables...) - Métamédiaire : 2ememain... (plateformes d'échange entre utilisateurs sans implication réelle de la plateforme/app)
4. Modèle marchand	Grossistes et détaillants de biens et services. Types : <ul style="list-style-type: none"> - Vente en ligne (e-commerce) : Yoursuper.com, Zalando... (opèrent uniquement en ligne)

	<ul style="list-style-type: none"> - Vente sur catalogue : 3 Suisses... - Vente au détail : Decathlon... (combine la vente en magasin et par internet) - Vente de produits numériques : iTunes Music Store, Fnac...
5. Modèle du fabriquant direct	<p>L'acheteur peut directement se fournir auprès du fabriquant. Le numérique met en avant le passage de la propriété à l'usage (Delorme et Djellalil, 2015).</p> <p>Types :</p> <ul style="list-style-type: none"> - Vente directe : The Ordinary, Dell... (combine la vente en ligne, avec les magasins propres et/ou distributeurs tiers) - Location : Keystone (économie circulaire de location de tapis) - Licence : Adobe (licences de programmes tels que Photoshop) - Contenu intégré à la marque : publicités liées au contenu créées par le même producteur <p>Deux autres modèles plus récents qui n'ont pas été définis par Rappa mais qui en raison de leur nature, peuvent s'inscrire dans cette catégorie sont :</p> <ul style="list-style-type: none"> - App payante (au téléchargement) : Procreate... - Achats intégrés (In-App) : TikTok... (microtransactions pour l'achat des fonctionnalités, de contenu numérique et des biens virtuels ou pour désactiver la publicité)
6. Modèle d'affiliation	<p>Deux agents, les affiliateurs (entreprises) et les affiliés (intermédiaires : autres entreprises ou des particuliers).</p> <p>L'affilié propose dans son espace web (page, réseaux sociaux...) le(s) produit(s)/service(s) de l'affiliateur en contrepartie d'une rémunération.</p> <p>Types :</p> <ul style="list-style-type: none"> - Rémunération au clic : ScarletClicks, Gab... - Partage de revenus (si la vente a lieu) : Amazon
7. Modèle communautaire	<p>Le facteur d'ancrage de ce modèle repose sur la fidélité des utilisateurs, ce qui transforme ce modèle dans une cible marketing extrêmement bien définie. La création de valeur peut provenir de la vente de produits/services, de la publicité contextuelle, des donations volontaires et de la contribution bénévole.</p> <p>Types :</p> <ul style="list-style-type: none"> - Open Source (Logiciels à usage libre) : Gimp, VLC Media Player, Firefox... - Open Content : Wikipédia, ... - Public Broadcasting : All shows, PBS Arts... - Social Networking services (services de réseautage) : Flickr, LinkedIn, groupes Facebook... (les utilisateurs vont connecter entre eux par le partage d'un intérêt commun)
8. Modèle par abonnement	<p>Un O</p> <ul style="list-style-type: none"> - Fournisseurs de services internet : Proximus, Orange (offrent de la connectivité à internet et d'autres services associés)
9. Modèle utilitaire	<p>Comme pour les services traditionnels d'eau et d'électricité, ces services sont proposés à la demande et facturés à l'utilisation effective de l'utilisateur.</p> <ul style="list-style-type: none"> - Utilisation mesurée : la facturation hors forfait en téléphonie, c'est à la minute ou au mégabits consommés.

	- Abonnements avec compteur : Uber, Cairn.info, sites de presse... (achat d'un chapitre au lieu d'un ouvrage entier)
--	---

Il est important de souligner qu'une même entreprise peut opérer sous la combinaison de différents modèles de revenus. Mais pour ces entreprises, il est crucial de bien établir son cœur métier, donc son business model. Par exemple, Proximus combine la vente directe de services téléphoniques (métier historique) et de services internet mais aussi de services de contenu (TV à la demande), en point de vente physique et en ligne, par abonnement au forfait et à l'utilisation mesurée.

Contrairement à d'autres auteurs qui développent une approche holistique, Michael Rappa prend une approche très économique de la création de valeur, centrée sur la mission commerciale.

Selon Timmers (1988) pour qu'une entreprise concrétise comment elle va produire du profit, le business model à lui seul ne suffit pas, il doit s'accompagner d'objectifs et d'une stratégie marketing autour du produit/service.

De leur côté, Schmid et Scaringella (2001), dans leur travail sur les modèles d'affaires disruptifs, distinguent les six éléments les différentiant : (1) la mission ; (2) la structure ; (3) les processus ; (4) les revenus et (5-6) les dimensions juridique et technologique (cité par Voelpel, Leibold et Tekie, 2004).

On constate donc que la définition de business model varie elle aussi relativement selon les auteurs.

Quoi qu'il en soit, un élément très pertinent évoqué par Hamel (1988) ainsi que par Senge et Carstedt (2001), est l'innovation comme source de valeur (cité par Voelpel, Leibold et Tekie, 2004). Cet aspect est clé dans le contexte changeant et perturbateur qui caractérise cette ère digitale, où l'accès à l'information a été simplifié et grandement démocratisé ; ouvrant aux entreprises la possibilité de développer leur portée à l'internationale, mais aussi incrémentant la concurrence et donc le besoin de se différencier pour se démarquer.

Les investissements R&D (recherche et développement), la compréhension client, l'identification de tendances et la prise de décisions stratégiques visant l'amélioration continue sont donc indispensables pour se différencier et renforcer sa compétitivité.

Finalement d'autres approches plus pragmatiques existent pour positionner son entreprise, car avant de produire de revenus, il est important de connaître son environnement et ses capacités, en vue de créer de la valeur tout en apportant un retour. Lecocq, Demil et Warnier (2004) notamment ont listé six questions déterminant un modèle d'affaires centré sur les revenus :

1. Quelles sont les ressources et les compétences susceptibles de générer des revenus ? (Matériel, personnel, savoir-faire...)

2. Quels sont les acteurs qui exploitent ces ressources et compétences ? (Définir les intervenants et les relations entre eux)
3. Qui paie pour acquérir et utiliser de ces ressources ? (Quelle est la cible client/consommateur ?)
4. Quelle est la structure (répartition) des revenus ? (Structurer les clients, les produits et services du plus rentable au moins rentable)
5. Quelles sont les modalités de rémunération de la vente ? (Prix, délais de paiement, moyens de paiement...)
6. Quels sont les coûts et quelle est la structure organisationnelle du business model en question ? (Calcul des coûts fixes et variables, pour s'assurer que l'activité es profitable).

Enfin, un entrepreneur à succès dans le domaine de la réservation en ligne parle même d'une équation à respecter : « Monétisation = big data + objectifs + technologie + personnalisation » (Lundø Nielsen, 2016).

3. Opportunités, enjeux et obstacles dans le cadre de projets (big) data / IA

3.1. Opportunités : finalités d'usage des projets (big) data / IA

<ul style="list-style-type: none"> Smarter Decisions Greater Knowledge Better Products Optimal Solutions Deeper Insights 	<ul style="list-style-type: none"> Customer Centric Products Increased Customer Loyalty Data-to-Discovery, Data-to-Decisions Data-to-Dollars
---	---

Figure 9 : Avantages globaux du (Big) data

Source : Khan, N., Naim, A., Hussain, M. R., Naveed, Q. N., Ahmad, N., & Qamar, S. (2019). *The 51 V's Of Big Data: Survey, Technologies, Characteristics, Opportunities, Issues and Challenges*. Proceedings of the International Conference on Omni-Layer Intelligent Systems, 19–24. Récupéré de <https://dl.acm.org/doi/10.1145/3312614.3312623>

Les données forment un afflux d'opportunités pour les entreprises, qu'elles soient grandes, petites, privées ou publiques. En réalité, les usages et les bénéfices du (big) data et de l'IA sont vastes, ils touchent tous les secteurs (recherche scientifique, santé, éducation, sécurité...) et l'ensemble des unités d'activité (marketing, RH, R&D...).

L'étude réalisée par Soumaya et Jaussaud (2018), met en exergue les modèles d'usage du big data (Bénavent, 2014), dont l'application concerne essentiellement des projets autour d'une activité ou stratégie particulière.

Tableau 2 : Modèles d'usages et ses bénéfices selon Bénavent (2014)

Finalités d'usage	Bénéfices d'usage
Améliorer la connaissance et la relation client (Knowledge Model)	<ul style="list-style-type: none"> - Obtenir une meilleure segmentation et un meilleur ciblage ; - Proposer une réponse supérieure à la demande ; - Augmenter ses ventes ; - Réduire ses coûts ; - Développer des nouveaux produits/services.
Augmenter les gains de productivité grâce aux algorithmes d'Intelligence artificielle (Automation Model)	<ul style="list-style-type: none"> - Réduire les coûts ; - Augmenter sa performance ; - Améliorer sa réactivité ; - Améliorer des processus
Améliorer la prise des décisions (Empowerment Model)	<ul style="list-style-type: none"> - Des tableaux de bord ; - Des informations personnalisées ; - La rétrospection ; - L'amélioration des relations avec ses partenaires/fournisseurs.
Services par les données (Platform Model) ⇒ <i>Sites collaboratifs</i> (mobilisation des actifs sous-utilisés des tiers via des plateformes de mise en relation de l'offre et la demande.), <i>marketplaces</i> , <i>réseaux sociaux</i> , <i>e-commerce</i> ...	<ul style="list-style-type: none"> - Répondre à une grande hétérogénéité des services ; - Se passer d'investissements pour l'acquisition d'actifs ; - Augmenter la relation, la confiance et le sens du partage entre utilisateurs.

La suite fournit les détails concernant les finalités d'usage du (big) data / IA et certains des bénéfices.

Knowledge Model – Modèle de la connaissance

1. ***L'anticipation de la demande, l'optimisation des ventes et du stock*** : des solutions existent que ce soit en open source avec du code que l'on développe soi-même ou des produits clés en main, pour prédire le nombre de ventes, appliquer des stratégies cross selling et up selling. Il existe également des solutions pour une gestion de stocks plus performante ainsi qu'une diminution du gaspillage d'espace, d'énergie, de matières premières, etc.
2. ***L'amélioration de prédictions*** : l'analyse des données pour autant qu'elle soit bien faite peut aider à atténuer les incertitudes en connaissant en temps réel les effets des stratégies commerciales et publicitaires. Ce qui permet donc de les améliorer constamment (Mejia, 2015).

Par exemple, les réseaux sociaux et les assistants virtuels ou Chatbots, entre autres, permettent des interactions constantes avec les clients, tout en recueillant des informations utiles pour l'analyse de leurs préférences et besoins.

3. **L'amélioration de stratégies segmentaires** : connaitre la segmentation du marché est un atout mais la micro segmentation est encore meilleure. S'adresser à la bonne audience facilite la compréhension des besoins individuels, l'identification des nouveaux segments et le profilage sur base des réactions comportementales. Cela permet aux entreprises de proposer une offre adaptée et concurrentielle ainsi que d'optimiser les échanges et les canaux de communication afin d'avoir un impact majeur sur l'expérience client, la fidélisation et l'engagement (Mejia, 2015).

Mejia (2015) cite l'exemple de l'utilisation du big data par China Merchants Bank (CMB) pour identifier les profils les plus susceptibles de quitter la banque et éviter leur départ en leur proposant des conditions et des offres ciblées.

Empowerment Model – Modèle décisionnel

1. **Le soutien à la prise de décisions** : aux niveaux opérationnel, commercial ou administratif, toute la chaîne de valeurs est concernée. L'analyse et la visualisation des données supporte les avantages futurs des projets, tout en diminuant le risque d'incertitude.

Des études montrent que les entreprises mettant en œuvre l'analyse des données sont plus performantes. Dans ce cas particulier, la difficulté se situe au niveau de la qualité des données, après leur nettoyage, préparation et regroupement.

Par exemple, suite à l'achat d'un article sur un site d'e-commerce, le client pourrait se voir proposer d'autres articles similaires ou complémentaires. Le choix d'articles qui seront mis en avant va se baser sur ses précédents comportements d'achat, dans le cas contraire un système de dépréciation d'intérêt pourrait se mettre en place. (Mayer-Schönberger et Cukier, 2013, pp. 16, 42, 110, 145).

Les algorithmes de la plateforme Mozzeno constituent un autre exemple. Cette start-up belge âgée de 3 ans, a créé un système de Crowdlending pour particuliers qui filtre les profils des demandeurs et leur donne une réponse immédiate. Les pessimistes pensaient que la plateforme n'attirerait que les mauvais payeurs. Leur prédiction est erronée car les algorithmes refusent 85% des demandes grâce à une analyse de quelques secondes.

Un dernier exemple : le « Matching » ou la sélection avancée des candidats à travers LinkedIn et d'autres outils sur lesquels nombreux managers de ressources humaines s'appuient.

Automation Model – Modèle de l'automatisation

- 1. L'amélioration des performances, l'augmentation de la productivité et le suivi des processus :** dès la prise de commandes jusqu'aux machines industrielles, l'analyse des données permet d'identifier les facteurs entravants ou ralentissant le flux du travail.

Cette analyse permet de les éliminer et d'automatiser ce qui peut l'être afin de gagner du temps, de diminuer les coûts et le taux d'erreurs mais aussi de valoriser le capital humain pour les activités essentielles propres à chaque business.

Il existe une multitude de tâches simples mais chronophages qui peuvent tirer profit des projets de digitalisation, en voici une liste non exhaustive : le déploiement automatique des campagnes publicitaires ; le traitement des documents comptables tels que les factures ; les rappels de paiement ou encore dans la logistique, l'établissement automatisé des bons de livraison. (SPF économie, 2021).

- 2. L'amélioration de la rentabilité des produits/services et des politiques de flux tendu (zéro-stock) :** via une meilleure gestion des ressources, des budgets et des déchets (compétences, matières premières, énergie, temps...). Outre la raison économique, ceci est aussi bon pour l'environnement et contribue donc à la responsabilité sociétale des entreprises (SPF économie, 2021).
- 3. Prévention et sécurité :** un autre usage non explicitement couvert par Bénavent, mais sur lequel les gouvernements se focalisent fortement est celui de la prévention des phénomènes nuisibles (épidémies, criminalité, etc.) par la gestion des risques ou Risk Management (Marguerit et Hamel, 2013). Ce principe s'applique également aux entreprises, puisque les cyberattaques, les fraudes et les captations des données par des logiciels de rançonnage sont de plus en plus fréquentes (Ghernaouti, 2016).

Platform Model – Modèle plateforme collaborative

- 1. L'innovation :** l'identification des tendances permet d'aboutir à des innovations avec ou sans intention initiale.

Les plateformes licornes tels qu'Uber, AirBnB et autres, sont-elles mêmes des innovations avec un modèle d'affaires inversé. Bien entendu, l'innovation ne concerne pas que les plateformes, mais toute nouvelle idée ou méthodologie cherchant à créer des océans bleus, c'est-à-dire des marchés tout à fait nouveaux où la compétition n'existe pas encore. Cette stratégie permet de dégager de nouvelles sources de profit à travers des innovations entre autres dues aux données captées et transformées en solutions.

- 2. Les Open Data :** ce sont des plateformes qui favorisent la transparence, l'amélioration et la création de nouveaux produits/services. Leur but le partage du savoir. La connaissance comme bien commun, pour Idriss Aberkane se multiplie, se partage sans la perdre.

Ces bases de données ouvertes à tous sont nombreuses : ODWB (Open Data Wallonie Bruxelles), ares-digitalwallonia, databank.worldbank, data-publica, Open Data France, data.world, datasf.org, data.europa.eu, data.gov, eosc-portal (European Open Science Cloud)...

Encore sous exploités, elles sont pourtant la boîte à idées des plusieurs entreprises/projets. Les Open Data incluent entre autres des données de géographiques, environnementales, climatiques, financières, statistiques, scientifiques, culturelles.

Des nombreux cas d'usage appliquant des modèles web tels que le freemium, modèle de réseaux et collaboratif, se basent sur ces données de libre usage. L'application Yuka en est un exemple, elle se nourrit des données mises à disposition par la Open Food Facts. Depuis cette source elle capte des informations concernant les allergènes, les ingrédients et la composition nutritionnelle des aliments (EGE, 2020). En effet, cette app permet de scanner les produits alimentaires et cosmétiques pour connaître leur composition.

En résumé les données sont omniprésentes et peuvent être utilisées dans toutes les industries, à partir du moment qu'une entreprise, même de très petite taille, dispose d'un capital de données.

Grâce à l'implémentation des projets de Data Analytics et d'IA, tout en mettant l'accent sur la sécurité des données et leur conformité, les sociétés peuvent utiliser leurs données pour une meilleure compréhension du marché et optimiser leurs processus. Ce qui leur permet d'améliorer leur fonctionnellement interne de manière à assurer leur pérennité.

3.2. Enjeux liés aux projets (big data / IA)

Les enjeux sont organisationnels, technologiques et économiques.

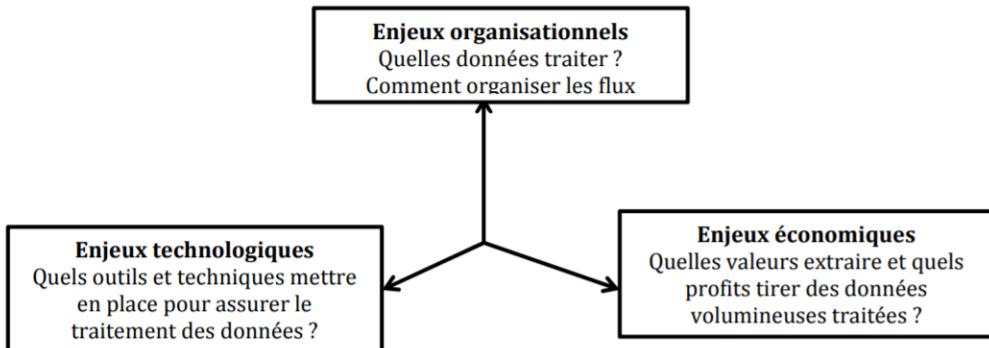


Figure 10 : Enjeux organisationnels, technologiques et économiques du big data pour les entreprises

Source : Karoui M., Devauchelle G., Dudezert A. (2014), « Big Data : Mise en perspective et enjeux pour les entreprises », N° Spécial ‘Big Data’, Revue Ingénierie des Systèmes d’Information, 19 (3), 73-92 Hermès. Ingénierie Des Systèmes d’Information, 19, 73–92. DOI : 10.3166/isi.19.3.73-92

1. **Quantité vs qualité :** le terme big data fait principalement allusion au volume, néanmoins la littérature met surtout en avant la qualité « véracité » des données. La densité fait partie des enjeux technologiques alors que la qualité fait partie des enjeux organisationnels.

Certes, plus les données sont nombreuses, plus elles vont nourrir les analyses croisées, mais c'est leur qualité qui déterminera la pertinence des résultats. Il est donc impératif disposer de données propres et de les traiter avec des solutions suffisamment performantes en vitesse d'analyse et en capacité de stockage (Karoui, Davauchelle et Dudezert, 2014).

2. **La protection de la vie privée (RGPD et ePrivacy) :** les propriétaires de données doivent respecter le Règlement général sur la protection de données (RGPD), lequel encadre le traitement des données à caractère personnel, sous peine de lourdes amendes allant jusqu'à 4% du chiffre d'affaires global.

Les sociétés qui conservent et utilisent des données sensibles doivent d'une part prouver qu'elles répondent aux principes d'une collecte précise, explicite et légitime et d'une autre part, qu'elles disposent des processus et des outils qui assurent la sécurité et la privacité.

Par exemple, pour les données partagées avec des tiers, comme les fournisseurs Cloud, les entreprises doivent s'assurer que ces tiers sont aussi en conformité. Enfin, le GDPR

inclus aussi les règles sur le consentement, la disponibilité ou la durée de conservation et la récupération de données par les individus. (Neamah, 2016).

L'anonymisation des données, leur cryptage, le contrôle et la surveillance d'accès figurent parmi les solutions possibles présentées par Lafuente (2015), (cité par Neamah, 2016).

Par exemple, en ce qui concerne la durée de conservation des données, leur anonymisation peut faire en sorte que les informations ne soient plus attachées à une personne en particulier, ce qui assure la conservation de la donnée pour une durée indéterminée.

Néanmoins, ce cryptage est aussi critiqué puisque certaines études révèlent aussi par les médias nous parlent du fait qu'à partir d'un jeu de données très important même anonymisé il serait possible d'identifier les personnes en question.

En outre, l'ePrivacy (ERP – Règlement sur la confidentialité électronique) qui est lié au RGPD, définit les directives à respecter quant à l'utilisation des cookies. Bien que les cookies soient des petits outils très intéressants pour les entreprises afin d'avoir un aperçu de l'activité des internautes sur leur applications et pages web, l'information qu'ils peuvent contenir, et qui est enregistrée par les navigateurs (Google, Safari...) peut potentiellement identifier les utilisateurs sans leur consentement (GDPR, 2021 ; CNIL, 2021).

Il existe différents types de cookies : Des cookies temporaires (effacement après fermeture des pages) et persistantes (restent jusqu'un an ou plus dans les appareils) mais aussi de cookies nécessaires, de préférence, pour les analyses statistiques et marketing. Dans tous les cas, cette réglementation vise à simplifier les refus sauf pour ceux nécessaires au fonctionnement des navigateurs et à l'amélioration de l'ergonomie des sites. Elle a également pour ambition d'informer explicitement le consommateur sur la nature et l'objectif des cookies utilisés (GDPR, 2021 ; CNIL, 2021).

3. ***La sécurité des données et des réseaux*** : Elle permet d'éviter la perte, le vol ou la destruction des données principalement des plus sensibles telles que les données à caractère privé des utilisateurs/clients ou encore les secrets industriels (propriété intellectuelle), les informations commerciales, etc.

Les PME sont particulièrement exposées, car elles ne disposent pas du même type d'environnement et d'expertise dont bénéficient les grandes entreprises. Les bases de données obsolètes figurent parmi les failles les plus courantes (Iqbal, Kazmi, Manzoor, Soomrani, 2018).

4. ***La disponibilité des données et la dépendance aux services cloud*** : de plus en plus de PME/TPE utilisent les services de Cloud Computing (IaaS – Infrastructure as a Service, PaaS – Platform as a Service, SaaS – Software as a service), grâce aux coûts très compétitifs et car ils demandent des connaissances techniques moindres. Cependant

cela implique une certaine dépendance envers ces fournisseurs au niveau du rendement, de la disponibilité et de la sécurité.

La crise sanitaire a accéléré cette tendance imposant des changements dans les flux de travail. Cependant, de telles transformations méritent d'établir des stratégies afin de choisir les bons systèmes d'information (SI), d'assurer la transition et le retour sur l'investissement. En outre, cette urgence de survie amplifie les risques, d'autant plus pour les petites et moyennes structures. En effet en cas d'erreur, les conséquences financières pour ces dernières peuvent être importantes, parfois même jusqu'à devoir cesser leurs activités. (OECD, 2021).

5. ***L'émergence des nouvelles technologies et méthodes d'application*** : une veille technologique et les conseils des organismes spécialisés s'avèrent fondamentaux lorsqu'il s'agit d'adapter un système d'information impliquant des changements décisifs pour maximiser leurs bénéfices (Digital Wallonia, 2020).

4. Obstacles structurels face aux projets (big) data / IA

Comme expliqué antérieurement, les PME/TPE sont à la traîne quant à la transformation digitale. L'adoption des projets big data et d'IA reste un challenge conséquent pour ces organisations, bien que ces technologies aient démontré un potentiel important en termes de création d'emploi et de richesse.

Mais alors, pourquoi l'adoption du big data et de l'IA n'est pas encore une réalité pour beaucoup d'entreprises ?

Schwertner (2017), mentionne l'humain comme principale barrière à l'adoption de moyens d'amélioration continue. Elle se focalise surtout sur le corps organisationnel, mettant en avant la culture, les traditions, la résistance au changement, le manque de sensibilisation mais aussi de connaissances, de ressources, de motivation et le risque.

Les freins structurels et managériaux persistent, c'est ce que les études de l'OECD (2021) et les informations publiées par SPF Finances (2021) à propos du Flash Eurobaromètre 486 (2020).

En voici les points principaux :

1. ***Carence de compétences analytiques, techniques et managériales*** : Dans les grands groupes et les start-ups disruptives, des métiers hyperspecialisés émergent (Data Officer, Data Scientist, Data Ingineer, etc.) et les offres de travail « globale » sont de plus en plus courantes. Le travail à distance permettrait donc aux entreprises de choisir les meilleurs talents, alors que les PME/TPE se voient restreintes par le nombre réduit de leurs effectifs, attachés à une zone géographique déterminée, ce qui implique un portefeuille de compétences amoindri. Par conséquent, la définition des besoins et la résolution des problèmes au sein des PME/TPE se voient défavorisées.

2. **Déficit de marges budgétaires et de financement**: logiciels de traitement, équipements, formation, recrutement, etc. Tous ces éléments pèsent lourd sur le budget des sociétés. L'accès au financement des investissements immatériels par des organismes bancaires s'avèrent difficiles puisqu'ils ne constituent pas une garantie.
3. **Manque d'infrastructures**: une connexion internet à haut débit est une nécessité fondamentale pour le big data. D'après l'OCD, les taux de pénétration sont en augmentation, mais les inégalités entre les grandes et les petites-moyennes entreprises sont considérables, les mettant encore une fois en position de faiblesse.
4. **Problèmes de sécurité et de confidentialité des données**: décider de tout centraliser en interne ou choisir un fournisseur Cloud (SaaS - Software As A Service), (IaaS - Infrastructure As A Service) ou (PaaS - Platform As A Service). La diversité des canaux de collecte et d'échange de données ainsi que la protection des informations est un des freins les plus marqués.
5. **Leadership et manque de culture digitale**: la vision et l'acceptance d'une organisation digitalisée divergent entre dirigeants et collaborateurs. Ces derniers souvent peu encadrés peuvent éprouver une résistance due au manque d'accompagnement méthodologique et à la peur de perdre leur emploi. Les dirigeants, parfois trop passifs et prudents, tardent à faire de la digitalisation une priorité. Les investissements risquent de ne pas fructifier si la gestion du changement est négligée (Capgemini, 2021).
6. **Limites réglementaires**: les contraintes croissantes liées à la manipulation et le stockage des données rendent peu aisée l'identification des normes auxquelles les PME/TPE sont soumises, sans oublier, les actions à mettre en place par la suite pour les respecter.

5. Eviter le big échec

« ...jusqu'en 2022, 85 % des projets d'IA donneront des résultats erronés en raison de biais dans les données, les algorithmes ou les équipes chargées de les gérer. » (Gartner, 2018)

Une étude mené en Belgique publiée 2018 par Ponsard et al., indiquait que la moitié des projets IT tombaient à l'eau. Gartner (2016) constate que 85% des projets big data restaient au stage pilote. Les raisons étaient technologiques mais surtout méthodiques.

Nous l'avons vu, le big data est doté des grandes qualités mais les challenges quant à la stratégie, au pilotage et au déploiement sont encore fortement présents, en particulier au sein des plus petites structures.

Sans aucun doute le big data et l'intelligence artificielle sont d'une façon ou d'une autre au service de toutes les entreprises. Par exemple, leurs fournisseurs peuvent disposer de systèmes informatiques orientés vers ces technologies.

Toutefois, l'idée est d'habiliter les PME/TPE pour qu'elles puissent aussi bénéficier directement.

Les petites structures ont une singularité en leur faveur par rapport aux grands groupes. Bien que quelque peu contradictoire, le fait de posséder plutôt des « Small Data » leur permet de plus facilement les traiter afin d'en extraire des « Smart Data », c'est-à-dire des informations suffisamment intéressantes pour établir des conclusions et prendre des décisions. Le taux de conversion d'un site internet est par exemple un bon indicateur. De plus, leur proximité avec leurs clients s'avère un atout en termes de confiance.

Il est connu que les entreprises dans une optique de continuité doivent être au courant des tendances pour s'adapter, en particulier aux générations Y (Millennials) et Z. Ces groupes composent le nouveau pouvoir d'achat, plus conscients de ce qu'ils consomment et à qui ils achètent, plus informés, hyperconnectés et plus difficiles à fidéliser. Dans ce contexte, les données marketing peuvent déjà faire toute la différence.

Pour rappel, les usages les plus courants aujourd'hui du (big) data et de l'IA sont principalement au niveau marketing, administratif et opérationnel avec les indicateurs d'évaluation technologique à travers les IoT (RFID, senseurs...).

Les PME/TPE peuvent donc d'une part analyser leurs données *a posteriori* comme c'est traditionnellement le cas, mais aussi en temps réel.

Au niveau marketing, pour lequel l'utilisation du big data est le plus répandu, les données servent principalement à comprendre le marché dans lequel l'entreprise évolue. Ces informations combinées avec des données internes vont pouvoir apporter des « insights » c'est-à-dire des connaissances à la base inconnues pour les enrichir davantage avec le temps et les nouvelles données. Ce qui permettra progressivement prendre des décisions et ensuite entreprendre des actions. Les entreprises évoluent ainsi d'un modèle « Data-Driven » à celui de « Insight-Driven ».

6. Mener à bien un projet (big) data / IA

Eychenne et Cointot (2014) mettent en évidence deux approches pour implémenter le big data.

Premièrement, il existe le modèle disruptif sur lequel se basent toutes les start-ups pour qui les mégadonnées font partie par défaut de leur modèle économique.

La seconde approche, évolutive et mieux adaptée aux PME/TPE, vise à utiliser le big data pour enrichir certaines activités précises. Encore une fois au niveau marketing, les réseaux sociaux s'avèrent très utiles pour connaître la perception client d'une marque, déterminer ce qui les intéressent, quels produits/services ils achètent, trouver les influenceurs idéaux pour promotionner un produit, etc.

Les étapes décrites par Eychenne et Cointot (2014) ainsi que celles recommandés par nombreux cabinets spécialistes du numérique et articles indépendants sont :

- 1) **Définir les besoins** : (Quels problèmes résoudre ? Quels sont les attentes ? Où l'on veut ajouter de la valeur ? ...)

Il est risqué de choisir directement un outil informatique sans connaître l'utilité que celui-ci aura pour l'entreprise.

- 2) **Identifier les raisons** : (Pourquoi il faut résoudre tel ou tel problème ? Pourquoi l'on veut ajouter de la valeur à tel ou tel niveau ? ...)

Ces interrogations permettent de découvrir le déclencheur d'un projet big data et donc, définir plus clairement ses besoins et objectifs.

- 3) **Etablir des objectifs mesurables** : la combinaison du quoi et du pourquoi génère des objectifs, lesquels doivent être explicites et progressifs (Cetic, 2017).

- 4) **Identifier les données nécessaires** : (internes, externes, directes et indirectes) (Quelles données ait-je besoin ? Quel volume de données faut-il calculer ? L'on veut obtenir des analyses à priori ou posteriori ? ...)

Si les données sont disponibles dans plusieurs logiciels et dispositifs, sous différents formats, il faut alors commencer par une phase de préparation, c'est-à-dire :

- a. **Les trier** : séparer les données pertinentes pour le projet en question ;
- b. **Les nettoyer** : se débarrasser des doublons, des données inutiles comme les fausses adresses email ;
- c. **Les segmenter** : catégoriser les données pour faciliter leur traitement.

Cette étape vitale est pourtant la plus complexe et longue. Les analystes et data scientist y consacrent près de 80% de leur temps contre seulement 20% pour l'analyse (Forbes, 2021). Plus récemment les études d'Anaconda (2021) et Matillon (2020) rapportent que cette phase reste encore très chronophage à l'ordre de 45%.

Un jeu de données devient exploitable une fois cette phase terminée. Il peut alors contenir divers types de données : financières, commerciales, marketing, production, approvisionnements, etc. Le business analyst ou le data scientist détermine les données qui vont influencer les analyses et les prédictions. Exemple : pour les prédictions d'achat, l'âge, le sexe et le revenu peuvent être de variables correctes.

Il s'avère difficile d'identifier les données à utiliser sans disposer d'une combinaison d'expertise métier et des compétences analytiques.

- 5) **Identifier les ressources nécessaires** : (Quels métiers impliquer ? Quels sont les compétences requises ? Quel est le budget ? Quels outils utiliser ? ...)

Une fois le projet bien défini, il peut s'avérer qu'un outil de base tel qu'un simple fichier Excel avec macros répond aux besoins. Excel n'a toujours pas été détrôné, au contraire tous les outils informatiques de gestion permettent de communiquer avec Microsoft Office.

- 6) Identifier les risques :** (Est-ce que les règles RGPD sont bien respectées ? Quid de la sécurité des données ? ...)

Négliger les risques techniques et réglementaires intrinsèques à la manipulation de données peut avoir des lourdes conséquences affectant la confiance des consommateurs, des partenaires et des investisseurs.

- 7) Vérifier les résultats :** (ils répondent bien aux questions de départ ? l'interprétation est fiable ? Les résultats sont-ils opérationnels ? ...)

- 8) Evaluer le retour sur investissement :** comme tout projet, il est indispensable d'établir le ratio entre son coût et ce qu'il rapporte à court et moyen terme. Les bénéfices vont varier en fonction des objectifs. Ils peuvent se traduire par une économie sur les frais de stockage, une diminution de la consommation d'énergie, l'amélioration du taux de conversion ou la diminution du taux d'abandon panier d'un e-shop.

ROI projet big data = Bénéfices – Ressources déployées

Dimension	Critères
Données	qualité, sécurité, niveau de structuration dans les données
Gouvernance	support à la gestion organisation bien définie culture axée sur les données
Objectifs	valeur de l'entreprise identifiée (KPI) Rentabilité taille réaliste du projet
Processus	agilité, conduite de changement, maturité, volumétrie des données
Equipe	compétences en ingénierie des données Multidisciplinarité
Outils	infrastructures IT, stockage capacité de visualisation des données contrôle des performances

Figure 11 : Projets (big) data – Approche par critères de succès

Source : Ponsard, C., Touzani, M., & Majchrowski, A. (2018). Synthèse des méthodes de conduite de projets Big Data et des retours collectés lors de pilotes industriels. *Ingénierie des systèmes d'information*, 23(1), 9-33. Récupéré le 10 Juin 2021 de <https://www.cetic.be/Comment-deployer-avec-succes-un-projet-Big-Data>

Enfin, Ponsard et al. (2018), rappellent comme pour tous les projets IT, les projets big data / IA doivent être souples. Les méthodes **agiles** sont particulières bien adaptées. Ce qui peut constituer un obstacle pour les PME/TPE qui n'ont pas cette culture de projet à leur fonctionnement.

7. Vers la transformation digitale des entreprises

7.1. Dématérialisation, digitalisation, innovation et transformation digitale

Certains emploient digitalisation pour se référer à la numérisation et d'autres utilisent interchangeablement transformation digitale et innovation digitale. Ce sont pourtant des concepts bien distincts, mais des amalgames sont régulièrement commis.

Pour ne plus faire des erreurs en voici les définitions :

- **La numérisation (dématérialisation)** est la transformation des objets (image, son, texte...) en informations numériques. (L'Académie française, 2021). Le remplacement des supports papier pour des supports digitaux facilite la collecte, et la structuration et le control des données, à travers notamment des formulaires, très pratiques pour l'entreprise et pour les utilisateurs (Denoix, 2018).
- **La digitalisation** comprend les technologies digitales capables de rassembler des ressources numériques hétérogènes et d'estomper les frontières (Yoo, 2010 ; Osmundsen, Iden, Bygstad, 2018). En effet, la digitalisation s'opère sur les outils devenant les moyens par lesquels les données sont véhiculées, créés, manipulées et stockées. Exemple : passer au stockage cloud et la signature électronique.
- **L'innovation digitale (ID)** a comme objectif de créer et de mettre à disposition les produits et services informatiques de demain, et ainsi trouver des nouvelles sources de croissance (Hinings, Gegenhuber et Greenwood, 2018). Pour que cela soit possible, il est nécessaire d'avoir des équipes avec des compétences pointues d'ingénierie et d'analyse. Dans ce contexte, l'IoT et l'IA (ML et DL) sont particulièrement utilisés. Exemple : AmazonGo (magasins sans caisse).
- **La transformation digitale (TD)** a pour objectif de moderniser le quotidien des organisations. Si l'on fait une analogie, ce concept pourrait être considéré comme un nouveau style de vie, où la digitalisation et le capital humain sont au cœur de l'organisation. Pour cela, il est fondamental de revisiter l'ensemble des savoir-faire et des savoir-être de façon continue. En bref, la TD correspond plutôt à un concept combinant la technologie avec des éléments socio-culturels, organisationnels et de gouvernance des données (Hinings et al., 2018 ; Osmundsen, Iden, Bygstad, 2018).

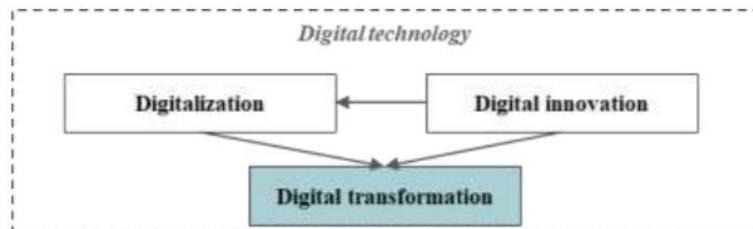


Figure 12 : Modèle conceptuel - lien entre digitalisation, Innovation et Transformation digitale

Source : Osmundsen, K., Iden, J., & Bygstad, B. (2018). *Digital Transformation: Drivers, Success Factors, and Implications*. *Digital Transformation*, 16. Récupéré de <https://aisel.aisnet.org/cgi/viewcontent.cgi?article=1004&context=mcis2018>

L'adoption des technologies numériques est une réalité omniprésente, suite aux retombées sur la croissance et la performance des entreprises qui les mettent en pratique. D'ailleurs, nombreuses publications scientifiques cataloguent les outils d'aide à la décision et à la planification (Mejia, 2015).

L'automatisation industrielle et de processus, le paperless, l'administration et la gestion numérique de flux de données, des matières et des produits ainsi que la désintermédiation, sont les principaux mouvements permettant des gains de temps, d'argent et de productivité considérables engendrés par la digitalisation (Fayon et Tartar, 2019).

Dès lors, une entreprise digitalisée ne se limite pas à une simple présence sur internet à travers une page Facebook ou via l'échange d'e-mails. Il s'agit bien d'avoir une forte présence sur internet mais aussi de disposer d'outils de gestion, de communication et d'analyse permettant d'alléger les tâches répétitives, de diminuer les erreurs et d'historiser les informations utiles. Le fait d'agiliser et harmoniser les flux de travail de manière à éviter les silos d'informations et enfin exploiter les données dont elle dispose, font partie d'une maturité digitale.

Par exemple, la digitalisation de processus métier d'une micro entreprise traditionnelle pourrait inclure l'intégration d'outils simples voire gratuits mais limités, tels qu'un logiciel d'e-mailing pour automatiser les newsletters. D'autres applications sont possibles telles que l'utilisation des moteurs sémantiques comme Google Trends pour analyser le marché ou encore l'introduction la facturation électronique. Cette dernière est d'ailleurs obligatoire depuis 2020 pour les fournisseurs des pouvoirs publics à Bruxelles et deviendra la norme pour tous les entreprises B2B en France à partir de 2023 (CEF Digital, 2021).

Dans le cas d'une organisation plus grande, cela pourrait aller jusqu'à la création de tout un écosystème d'outils transversaux, facilitant les flux et la synergie entre départements internes mais aussi avec les partenaires externes comme les fournisseurs. Un tel niveau de digitalisation caractérise la maturité digitale, qui au final regroupe des dimensions d'ordre socio-culturelles, organisationnelles et technologiques.

Il est certain que plus les flux des données sont importants, plus les projets d'analyse avancée sur des data internes et surtout des big data prennent du sens.

7.2. Transformation digitale : technologie ou culture ?

Des entreprises plus intelligentes et performantes, qui comprennent les subtilités de la demande et qui sont capables d'agir dans l'urgence, c'est ça le principe de la transformation digitale.

La transformation digitale est un mouvement guidé par les données, les modèles, les algorithmes et les technologies numériques. Les données sont quant à elles des facilitateurs à la croissance et à l'innovation.

Au-delà de la digitalisation, il s'agit de métamorphoser l'offre même et la façon dont celle-ci est produite, promotionnée, mise à disposition, évaluée, contrôlée, etc. Ceci implique forcément un changement aussi au niveau des mentalités pour adopter cette métamorphose. (Schwertner, 2017).

La transformation digitale s'accompagne d'une diminution des coûts et une plus grande accessibilité aux technologies numériques, que ce soit par la démocratisation du Cloud Computing, par l'augmentation des puissances CPU, par la visualisation dynamique et le stockage NoSQL, entre autres.

La servicisation numérique permet de transformer son offre. Elle peut être complémentaire au modèle d'affaires existant ou bien complètement le remplacer. En effet, la servicisation consiste à passer d'un modèle de vente de bien vers la vente d'offres intégrées produit-service (Peillon, 2016).

Concrètement, cela pourrait correspondre un changement de la vente de moquettes à l'offre d'un service de mise à disposition, d'installation et de maintenance de ces moquettes où le consommateur fera toute la démarche du choix et des dimensions à travers une application dédiée. Un autre exemple plus parlant et connu de tous est Netflix.

Les grandes entreprises et les Start-ups « natives numériques » sont les plus concernées par la transformation digitale, néanmoins les digital nomades le sont aussi, ce n'est donc pas d'une question de taille, ni d'argent mais de culture digitale.

Nonobstant, plusieurs études confirment que les plus petites structures sont les plus vulnérables, en retard sauf sur le plan marketing et administratif où l'écart est moins perceptible. Elles restent toutefois fragiles car la transformation digitale contrairement à la digitalisation implique un processus profond qui touche l'ensemble du business model (produits/services, collaborateurs, partenaires, coûts, revenus...) mais aussi la culture d'entreprise, les habitudes organisationnelles, les pratiques managériales, etc., et ce dans un souci d'étendre le spectre de possibilités (ESPR, 2021 ; OECD, 2018 ; OECD, 2021 ; Agoria, 2019...).

En Belgique, malgré un intérêt soutenu pour les PME belges, seulement 22% des entreprises de minimum 10 employés analysent en interne des données provenant de tout type de dispositifs (smartphone, site web...) et sources (base de données, open data, réseaux sociaux...) (Eurostat, 2021).

En Europe et plus particulièrement en Belgique, 7% des sociétés de 10 employés et plus emploient de l'intelligence artificielle et 2% utilisent le machine learning pour analyser le big data (Eurostat, 2021).

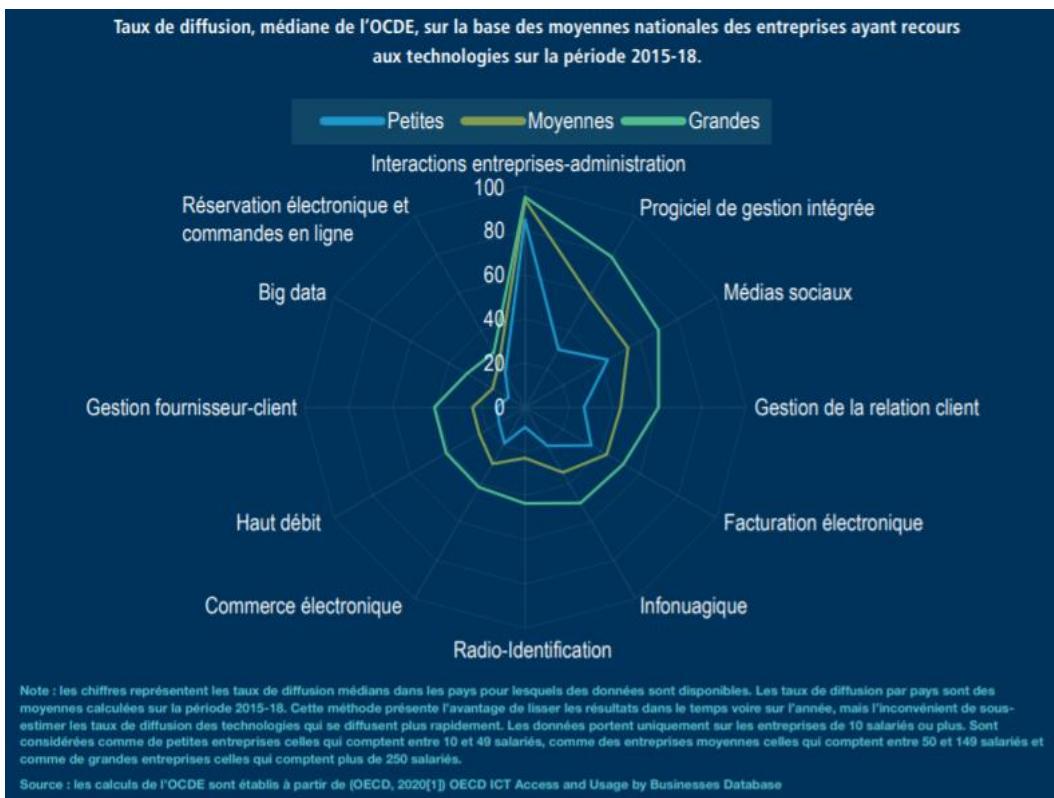


Figure 13 : Les PME, un net retard dans l'adoption des technologies numériques

Source : OECD. (2021). *The Digital Transformation of SMEs*. OECD. Récupéré le 23 juin 2021 de https://www.oecd-ilibrary.org/industry-and-services/the-digital-transformation-of-smes_bdb9256a-en

En conclusion, les PME/TPE selon leur secteur d'activité peuvent parfaitement se digitaliser davantage voire totalement. C'est avant tout une question de volonté.

7.3. Leviers stratégiques de la transformation digitale

Tel qu'évoqué précédemment, la transformation digitale est un véritable défi qui s'accompagne de grandes promesses de développement. Mais pour cela les entreprises doivent se préparer.

D'après Delorme et al. (2015), cette révolution implique :

- De nouveaux business model axés sur le digital ;
- Une nouvelle gouvernance axée sur les données ;
- Des entreprises ouvertes et collaboratives au sein d'un écosystème ;
- Moins de hiérarchie, mais plus d'organisation ;
- Un management plus démocratique favorisant l'innovation ;
- Des leaders plutôt que des dirigeants ;
- Des changements constants ;
- Une meilleure gestion de ressources ;
- La fluidité de l'information ;
- L'engagement des collaborateurs ;
- Etc.

Les 6 leviers stratégiques de Fayon et Tartar (2019) à tenir en compte avant de se lancer dans la transformation digitale sont :

1. La stratégie

- a. Le plan stratégique**
- b. La veille économique**
- c. L'excellence économique**

Pour devenir un acteur numérique, il est nécessaire de repenser la stratégie et de l'adapter afin de la rendre plus souple pour absorber les chocs c'est-à-dire les changements constants dans le monde digital. Pour cela, il faut prendre du recul sur le business model existant et identifier les atouts de l'entreprise qui lui permettrait d'engendrer plus de valeur. Il convient également de faire une veille économique et technologique, de penser à la manière de combler l'écart car les clients, eux sont quotidiennement dans le digital.

Les auteurs recommandent d'avoir une présence omnicanale, de renforcer l'exploitation des technologies numériques et d'envisager des collaborations avec des entreprises innovantes.

2. L'organisation

- a. La gouvernance**

- b. Le management*
- c. Les structures numériques*

Les entreprises doivent se tourner vers l'agilité en intégrant des méthodes de travail telles que le Lean Management, méthodes Agile et DevOps tout en implémentant des TIC adéquates à la hauteur des ambitions, des motivations et des capacités de l'entreprise. Un management ajusté aux changement économiques actuels pourra faire augmenter la valeur Time to Market, et livrer des plus rapidement et mieux.

3. Le personnel

- a. La formation*
- b. L'intelligence collective*
- c. L'expertise*

L'humain est plus que jamais un élément fondamental. Les dirigeants et les managers ont la responsabilité de transmettre aux collaborateurs une vision commune et de cultiver un climat de confiance. Il est donc essentiel de les encourager la formation et la maîtrise des outils collaboratifs pour ainsi favoriser l'innovation, l'engagement des salariés, la fluidité de la communication et l'élimination de silos.

De plus, la carence des profils analytiques mène les entreprises à tout faire pour attirer et développer des talents, mais surtout à valoriser les individus pour leur donner envie de s'engager.

4. L'offre

- a. L'innovation et la conception*
- b. Le commercial*
- c. Le marketing*
- d. L'expérience client*

L'engagement d'une entreprise qui se veut digitale passe avant tout par son offre, celle-ci doit être facilement et rapidement accessible (time to market). Elle doit aussi être personnalisable voire même sur mesure. Il est vivement conseillé aux entreprises d'être présentes sur les réseaux sociaux car ce sont des sources riches en indicateurs. Bien que cela parait évident, la soigner sa e-réputation est également un élément clé.

En outre, la publicité n'est plus la seule arme pour vendre, d'autant plus qu'il existe des solutions informatiques pour les bloquer. Elle est donc devenue plutôt un argument d'attraction et de curiosité qui sera complété par le partage de contenu de qualité. Les blogs sont un bon moyen pour communiquer permettant aux clients d'enrichir leurs connaissances sur certains sujets tout en améliorant l'expérience client.

5. La technologie et l'innovation

- a. L'architecture*
- b. L'orientation client*
- c. Les standards*
- d. La productivité*

L'entreprise doit identifier les technologies et les outils qui sont les mieux adaptés pour elle. Comme que mentionné plus haut, ce choix se fait en fonction des motivations, des ambitions et du budget. Il est important pour les PME de se digitaliser de plus en plus, à leur rythme mais en ne perdant pas de vue cet objectif. A titre d'exemple, une PME pourrait commencer par instaurer un CRM cloud, puis optimiser le référencement de son site web, et enfin opter pour des solutions plus techniques comme IA et l'analyse prédictive.

6. L'environnement

- a. L'environnement économique**
- b. L'environnement réglementaire**
- c. Le lobbying**

Les entreprises doivent constamment s'informer pour décrypter les opportunités et menaces. Il est essentiel pour elles de s'adapter et donc adopter une culture évolutive.

Au niveau réglementaire, nous avons vu précédemment le RGPD (cf. 3.2 Les enjeux) ainsi que d'autres enjeux juridiques, fiscaux et éthiques. Les auteurs évoquent aussi l'utilisation des API (interface de programmation d'applications) permettant la communication entre différentes applications même celles des partenaires, toujours dans l'optique de faciliter les échanges et d'augmenter l'intégration.

7.4. Se faire aider et accompagner

« Seul on va plus vite, ensemble on va plus loin » (auteur inconnu).

Que ce soit pour définir la vision et des besoins de l'entreprise, choisir ou mettre en place des nouvelles technologies et des solutions numériques, plusieurs alternatives sont à disposition des PME/TPE, des indépendants et des entrepreneurs. Elles permettant aux différentes structures de se faire aider dans leur volonté et leur démarche de digitalisation, de projets (big) data/IA, d'innovation et de transformation digitale.

Consciente du retard numérique de ces entités, la Commission coordonne toute une série d'initiatives européennes, régionales et nationales portant sur la numérisation du secteur public, la connexion à haut débit, l'internet ouvert, la cybersécurité, la protection des données à caractère personnel, la libre circulation des données non personnelles, entre autres (EPoS, 2019).

La Belgique tente d'y remédier via les Régions qui proposent :

- Séances de sensibilisation et d'information :
 - o Outils numériques ;
 - o Nouvelles technologie (Big data, IA...)
 - o RGPD
- Conseil et accompagnement pour concrétiser un projet Big Data et IA, ainsi que développer les compétences des collaborateurs ou les réorienter :
 - o Wallonie-Bruxelles : Digital Wallonia, Cetic...
 - o Flandre : Digitale Toekomst

- Bruxelles : Hub.brussels
- Formations gratuites ou à faible coût (en ligne ou non) :
 - Wallonie : Numéria
 - Flandre : IA Academy
 - Bruxelles : BE Central, Digitalcity.brussels, MolenGeek
- Subsides et chèques formation/consultance
 - Bruxelles : Prime web et appels à projet
 - Wallonie : Chèques-Entreprises
 - Intégration des outils numériques et des nouvelles technologies
 - Audit de la maturité numérique
 - Formations
 - Consultance
- Espaces de co-travail et de co-création
- Un large réseau des partenaires

En plus des nombreuses formations, plus d'une trentaine de Digital Innovation Hubs (DIH) partout en Belgique, soutiennent les petites et moyennes structures ainsi que les entrepreneurs en leur offrant des environnements de co-working et co-création, des opportunités de réseautage, des infrastructures pour la mise au point et le prototypage des projets, des conseils sur les technologies et les outils les plus adaptés en fonction de leur secteur et capacités.

Les différents services ci-dessus permettent entre autres de faciliter l'accès aux ressources stratégiques comme des alternatives de financement. Ils permettent également l'accès à des laboratoires d'expérimentation et l'utilisation des outils de prototypage comme les imprimantes 3D, et favorisent les partenariats et les liaisons (OECD, 2020).

Les DIH les plus connus sont Agoria, SIRRIS, Digital Wallonia, entre autres. La liste complète se trouve sur le site : www.intelligenceartificiellepourpme.be

En outre, de manière plus ciblée, en région Wallonne, le programme Digital Wallonia met à disposition depuis 2017 les chèques-Entreprises, regroupant en une seule plateforme plusieurs aides autour du numérique. Celles-ci cherchent notamment à améliorer sa sécurité informatique, favoriser et accompagner les projets des performances des processus opérationnels, faire auditer sa maturité numérique.

Aujourd'hui, il existe une foule d'outils gratuits en ligne pour mesurer sa maturité digitale. DigiScan proposé par Agoria, DIMM.UP outil développé par Michaël Tartar et intelligenceartificiellepourpme.be en sont quelques exemples.

De la même manière, l'administration dématérialisée est promue par les gouvernements. Cela inclut la facture électronique, la signature électronique et les déclarations en ligne (OECD, 2020). De nombreux projets de sensibilisation sont organisés depuis 2017.

Les pouvoirs publics cherchent également à créer et renforcer une culture des données. Des avocats experts par l'initiative « Académie digitale des avocats » transmettent leurs savoirs quant à l'utilisation et la gestion de données ainsi que le respect des règles RGPD.

Digital Commerce est un autre projet qui vise la digitalisation des commerces, spécialement par l'utilisation des moyens de paiement efficaces.

Enfin, la plateforme Open Data Digital Wallonia et une multitude d'autres (cf. 3.1. Les opportunités : finalités d'usage dans le cadre des projets big data – *Platform Model – L'innovation*) encouragent l'innovation, grâce à la mise à disposition des données en libre accès et libres d'utilisation.

7.5. Facteurs d'échec et défis de la transformation digitale

Davenport et Westerman (2018) ont analysé les profils des plusieurs entreprises ambitieuses qui ont échoué dans leur tentative de transformation digitale. Malgré des investissements colossaux, pressées de tout digitaliser très rapidement, elles se sont cognées au mur des imprévus, des départs anticipés des cadres, l'abandon des projets, la diminution des performances, etc.

Davenport et Westerman (2018) ont remarqué que les facteurs suivants avaient une influence :

1. ***La conjoncture et les persona*** : les auteurs citent l'économie et la désirabilité des produits comme possibles facteurs d'échec, si ceux-ci sont mal anticipés.
2. ***Les outils digitaux ne sont pas des baguettes magiques*** : considérer la transformation digitale comme un concept impliquant seulement la technologie et qui résout tous les problèmes ; les travaux de Fayon et Tartar ainsi que ceux de Delorme et Djellalil le confirment. Ils soulignent également l'impact de la complexité dans la gestion et le traitement des données, qui impliquent des processus de préparation des données avant de les utiliser.
3. ***Le manque d'introspection*** : du management par rapport à la maturité numérique de ces entreprises et de l'environnement au moment de leur mésaventure. Des investissements moins nombreux mais plus ciblés auraient pu les aider.
4. ***Perte d'objectivité et de rationalité des dirigeants*** : dernier facteur concernant également les dirigeants, est de partir à la dérive avec le risque de couler, dû à une sur-focalisation sur la technologie. Les PDG de ces grands groupes ont oublié les autres vecteurs de profit, tels que les clients et les collaborateurs et sont tombés dans le piège d'investir sans mesure pour obtenir des rendements ultra rapides.

En résumé, le digital n'est pas une solution magique. Il faut bien considérer les risques et surtout il faut bien travailler sa stratégie en amont. Pour les PME/TPE cela veut dire qu'il vaut mieux commencer par de petites étapes en testant différentes solutions. D'ailleurs, nombreuses sont les start-ups qui mettent à disposition de PME/TPE des modules de test pour se familiariser avec leurs outils, des solutions sans investissements (frais fixes) permettant d'évoluer progressivement.

7.6. Une vision, une stratégie et beaucoup de volonté

Les projets de digitalisation et de transformation digitale, big data et d'Intelligence artificielle inclus, sont des décisions stratégiques dans les mains de la direction et liées à sa volonté de faire évoluer les choses.

Justement cette volonté fait souvent défaut. Plusieurs éléments en témoignent (Fayon et Tartar, 2019 ; Bellalij, 2019 ; Delorme et Djellalil, 2015, Schwerther, 2017) :

- Des PME encore trop traditionnelles et lentes craignant la vague digitale
- Une culture d'entreprise et des habitudes difficiles à faire évoluer sans un réel engagement « top-down »
- Une culture projet peu développée
- La crainte d'une perte de pouvoir, dû au fait que la TD prône une communication ouverte et étendue à tous les niveaux de l'organisation, des collaborateurs conscients de leur rôle et des enjeux de l'entreprise, autonomes et proactifs.
- Une faible valorisation de l'innovation de :
 - o Outils, produits, services, marchés ;
 - o Procédés internes : nouveaux flux de travail, nouveaux métiers, télétravail... ;
 - o Modèle de revenus : des nouvelles manières de monétiser son offre ;
 - o Sources d'approvisionnement : louer du matériel informatique en lieu d'investir ou passer aux énergies vertes... ;
 - o Nouvelles activités à part entière : la création d'une start-up pour la gestion des déchets de l'entreprise...

Subir ou agir ? Tel est le dilemme qu'il faut mieux éviter. La crise Covid l'a bien démontrée, et d'ailleurs les études économiques financées par les pouvoirs publics et par les sociétés de consultance ainsi que les médias nous le confirment, les PME belges comptent investir davantage dans le digital (Agoria, 2019). Il est désormais clair, que la culture digitale est une nécessité qui ne tient pas compte des frontières. Toutes les entreprises sont concernées afin d'assurer leur croissance et leur rentabilité à long terme.

Cependant, investissement n'est pas synonyme de réussite. Il n'existe aucune garantie et c'est pour cela que la transformation digitale nécessite d'adopter un état d'esprit ouvert à l'évolution constante.

7.7. La transformation digitale en action

7.7.1. Facteurs de succès

Tableau 3 : Les facteurs de succès d'une transformation digitale réussie

Résultats – Etude PME Osmundsen, Iden et Bygstad (2018)	Résultats – Etude PME Industrielles Belges Agoria (2019)
Développer une stratégie d'entreprise numérique	Intégration de la TD dans la stratégie de l'entreprise
Une culture organisationnelle favorable	Approche Top-down et Bottom-up simultanée (lorsque le niveau de maturité est plus élevé les projets commencent aussi par initiative de ceux qui sont sur le terrain)
Des activités de transformation bien gérées	
Engager les managers et les employés	Engagement et accompagnement des collaborateurs (ne pas imposer mais travailler ensemble, l'avis des collaborateurs est précieux car ce sont eux qui vont appliquer les outils et méthodes au quotidien)
Tirer parti des connaissances externes et internes	Collaboration avec des partenaires digitaux externes : Consultants experts, Digital Hubs, Agoria... (cf. Se faire aider et accompagner)
Développer les capacités des SI	Désignation d'un responsable de la TD
Développer des capacités dynamiques	
Aligner les projets et les SI	
Implications	
Réforme de l'organisation des SI	
Nouveaux modèles d'affaires	
Effets sur les résultats et la performance	

7.7.2. Conseils à suivre

Afin de parvenir à créer et maintenir une culture digitale au sein des PME/TPE, Agoria suite à son étude sur les PME en Belgique recommande :



Figure 14 : Les 7 étapes vers la transformation digitale

Source : Agoria (2019). *Etude Digital is a mindset*. Bruxelles : Agoria Récupéré le 27 juin 2021 de <https://www.agoria.be/fr/Livre-blanc-Le-digital-un-etat-d-esprit>

L'étude *Strategy, not Technology, Drives Digital Transformation* été mené par Kane, Palmer, Phillips, Kiron, Buckley, réalisée en 2015 auprès de 4.800 dirigeants d'entreprise, gestionnaires et analystes d'organisations du monde entier. Elle converge avec plusieurs études à grande et plus petite échelle qui au cours des dernières années ont démontré l'importance de la transformation digitale et les grands avantages du big data et de l'intelligence artificielle.

Cette étude soulève un point particulièrement intéressant : l'innovation. Selon les résultats, les bonnes idées n'éclosent pas par magie un matin au réveil. La plupart des idées intéressantes et qui aboutiront à des résultats tangibles naissent de la collaboration inter-fonctionnelle entre personnes d'horizons différents.

En effet, la façon dont les organisations à un stade de maturité numérique élevé développent une offre, des processus et des business modèles innovants, est en grande partie due à un environnement de travail collaboratif. Cet exemple à suivre est donc tout à fait applicable aux PME/TPE.

Partie 2

Approche Pratique : études empiriques

1. Etude quantitative

Pour comprendre les opportunités et les moyens par lesquels les PME/TPE peuvent tirer profit de leur données, une enquête par questionnaire et des entretiens ont été menés.

La mise en œuvre de l'échantillon probabiliste nécessiterait des moyens et des ressources, financières, humaines et temporelles qui dépassent mes capacités. Dès lors, c'est la méthode empirique (non-probabiliste) qui a été retenue, plus précisément la méthode de volontaires aussi appelé de convenance.

1.1. Méthodologie

Déployé en ligne, le questionnaire était destiné aux chefs d'entreprise des PME/TPE (0 à 250 employées) ainsi qu'aux indépendant et porteurs de projet.

Afin d'augmenter le taux de réponse, l'enquête a été publiée plusieurs fois dans des groupes Facebook et LinkedIn rassemblant des communautés d'indépendants, des entrepreneurs et de chefs d'entreprise.

La durée totale de l'enquête a été de quatre semaines, du 12 juillet au 9 août.

1.2. Objectifs

Dans un souci descriptif et évaluatif, cette enquête apporte une idée du niveau de familiarisation des répondants par rapport aux technologies et opportunités associées au big data.

De même, cette enquête tente d'évaluer niveau de maturité numérique des répondants, c'est-à-dire le degré d'application et d'implication par rapport aux technologies numériques.

1.3. Limites

Puisque la méthode de volontaires n'a pas besoin d'une base de sondage, elle s'adapte particulièrement bien aux enquêtes en ligne, néanmoins cette enquête est non représentative et ne permet pas l'extrapolation de résultats. Par conséquent, un échantillon plus grand ne serait pas forcément plus fiable.

De plus, le risque de biais est très important puisque les répondants sont à priori ceux qui s'intéressent au sujet. Les individus visés n'ont pas tous la même chance d'y participer.

1.4. Questionnaire

L'enquête a permis de recueillir le ressentit de 33 individus afin d'évaluer leur maturité numérique.

Des dirigeants d'entreprise, des indépendants et des porteurs de projet figurent parmi les répondants. Les différents éléments permettant les caractériser sont le secteur d'activité, l'âge de l'entreprise/activité et la région

Place à la description des phénomènes :

1.5. Caractéristiques et proportions de l'échantillon

Tableau 4 : Echantillon : type

Type	Echantillon	Répartition %
En création	11	33%
Indépendant	8	24%
Micro	11	33%
Petite	2	6%
Moyenne	0	0%
Grande	1	3%
Total	33	100%

33 personnes ont répondu au questionnaire, dont la plupart étaient des porteurs de projet et de micro entreprises (0-10 employés). Ces deux catégories représentent chacune 33% de l'échantillon. Suivis des indépendants qui représentent 24%.

Tableau 5 : Echantillon : âge

Age	Echantillon	Répartition %
0	8	24%
< 1	3	9%
1-3	3	9%
3-5	7	21%
> 5	12	36%
Total	33	100%

La plupart des répondants sont des entreprises jeunes de moins de 5 ans ou en phase de création. Seulement 36% sont des entreprises de plus de 5 ans d'activité.

Tableau 6 : Echantillon : siège

Siège	Echantillon	Répartition %
Bruxelles	6	18%
Flandre	1	3%
Wallonie	23	70%
France	3	9%
Total	33	100%

La région de Bruxelles-Capitale et la région Wallonne représentent ensemble 88% de l'échantillon. Il est nécessaire de souligner que le sondage a été déployé en français. La Flandre sous-représenté ne figurera pas dans les analyses graphiques.

Tableau 7 : Echantillon : secteur d'activité

Secteur	Echantillon	Répartition %
Administration et soutien	1	3%
Agriculture / Sylviculture / pêche	1	3%
Banque / Assurance	1	3%
Bâtiment / Construction	3	9%
Commerce de gros (B2B)	1	3%
Culture / Evénementiel	1	3%
Édition / Communication / Marketing	1	3%
Horeca	1	3%
Informatique / télécoms	3	9%
Professions libérales	1	3%
Professions libérales	1	3%
Retail (B2C)	6	18%
Santé	1	3%
Santé	3	9%
Sport / Loisirs	1	3%
Autres	7	21%
Total	33	100%

Le commerce de détail est le secteur le plus représenté avec 18%.

1.6. Analyses des résultats

Etes-vous familier avec les concepts suivants ?

Les micros entreprises et les porteurs de projets figurent parmi les plus familiarisés avec les technologies et méthodes permettant la valorisation des données.

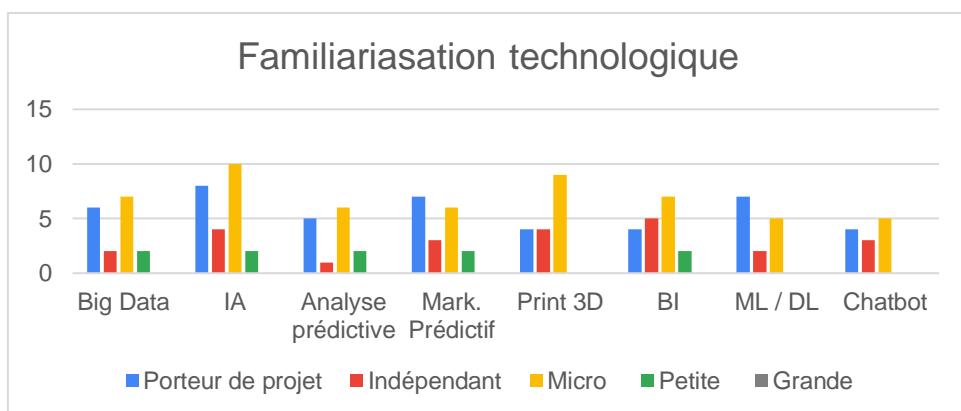


Figure 15 : Résultats de questionnaire : familiarisation technologique

Que représentent les technologies citées précédemment pour votre projet d'entrepreneuriat / entreprise ?

Pour 33% des répondants ces technologies et pratiques représentent une opportunité, et une nécessité pour la même proportion.

Le graphique ci-dessous montre aussi que 21% des répondants ne savent pas ce que l'exploitation de données à travers ces pratiques pourrait leur apporter.

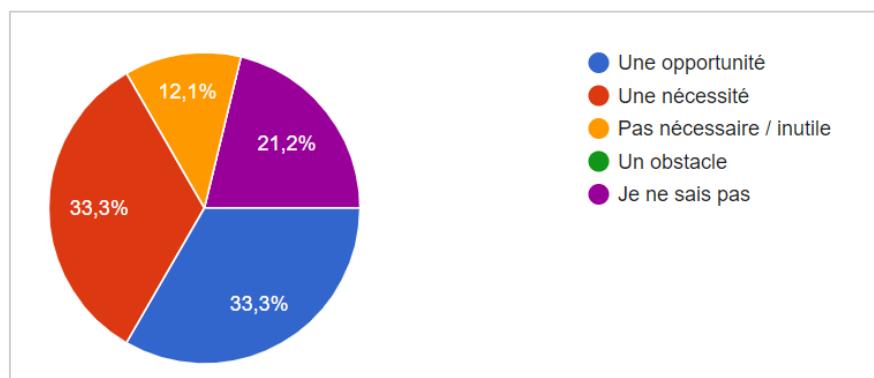


Figure 16 : Résultats de questionnaire : perception sur l'utilité des technologies numériques

Le graphique ci-après montre que les plus intéressés sont les micros entreprises et les porteurs de projet.

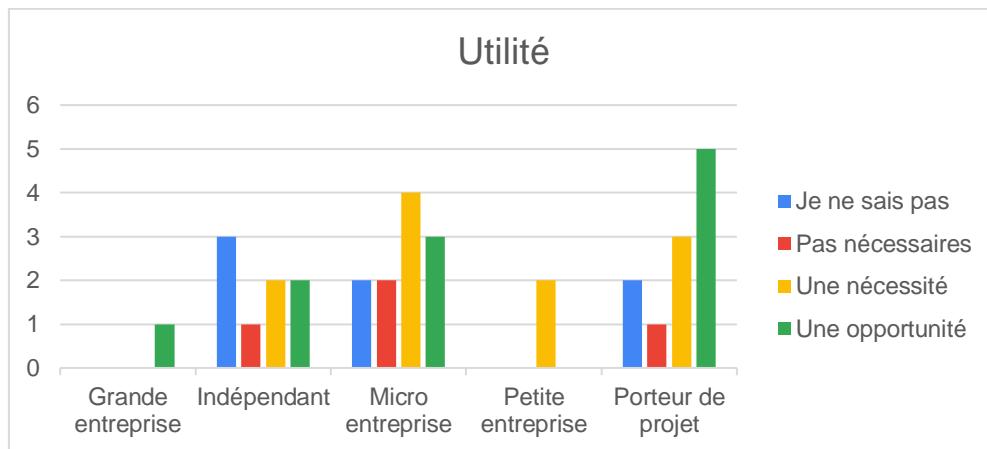


Figure 17 : Résultats de questionnaire : rapport utilité/profil

Le graphique ci-dessous montre que les répondants cherchent à optimiser principalement des fonctions support et marketing.

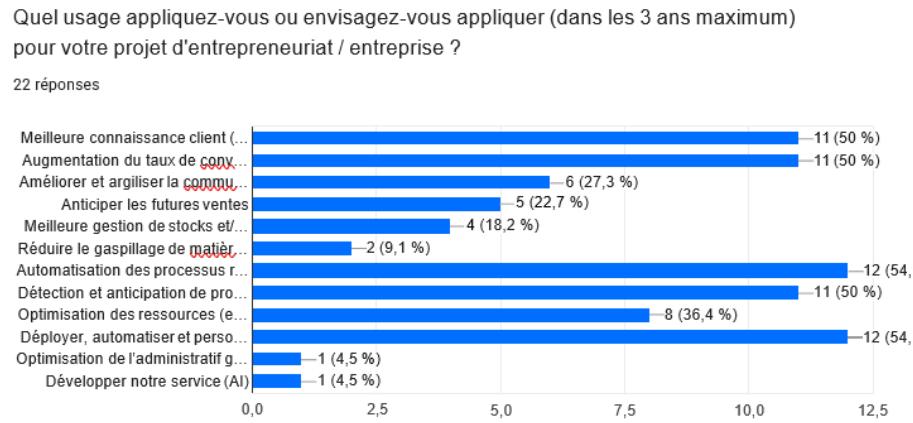


Figure 18 : Résultats de questionnaire : usage

Le graphique suivant montre que les limites se trouvent au niveau du budget, des connaissances technico/analytiques et la difficulté à identifier ses besoins.

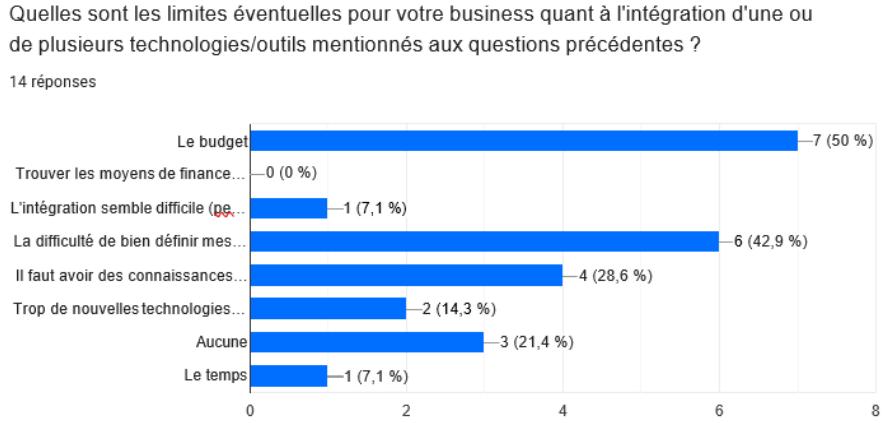


Figure 19 : Résultats de questionnaire : limites

Le graphique ci-après indique que près de 40% des personnes conservent leurs données, cependant ils n'ont pas un format standardisé. Pour 30% des répondants, les outils d'aide à la gestion conforment leur endroit principal de stockage.

Rassembliez-vous des données internes (production, achats, clients, prospects, ventes, etc.) dans des fichiers informatisés ?

33 réponses

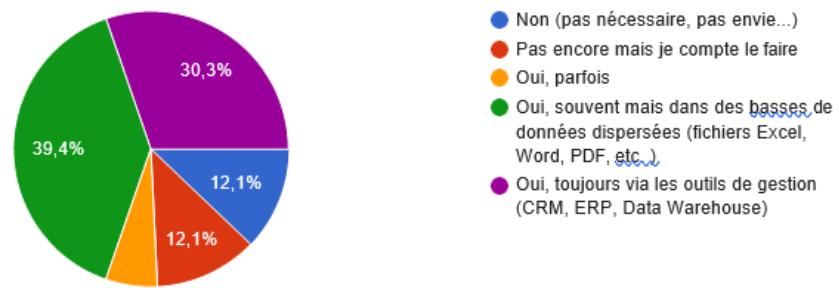


Figure 20 : Résultats de questionnaire : collecte de données internes

Le graphique suivant, montre que la majorité des répondants 51% ne font pas usages des données externes.

Collectez-vous des données externes gratuites ou payantes pour l'analyse et la prise de décisions (statistiques, prix, marché, concurrents, innovations, etc.)

33 réponses

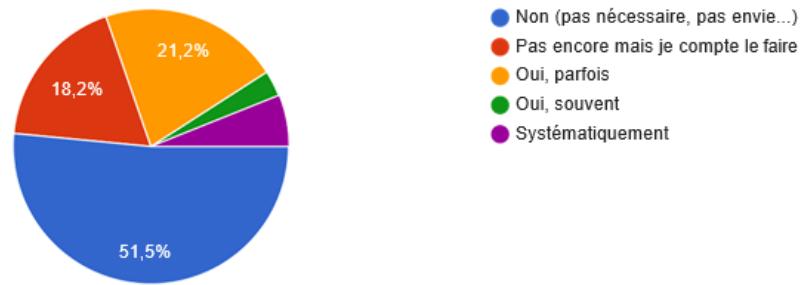


Figure 21 : Résultats de questionnaire : collecte de données externes

Pour finir, la plupart de répondants n'étaient pas au courant du soutien des régions, mais se montrent intéressés.

Saviez-vous qu'il existe des organisations d'accompagnement, de formations mais aussi des aides financières et des subsides pour les PME, TPE et Start-up aux niveaux fédéral, régional et européen ? (Ex : Digital Innovation Hubs)

33 réponses

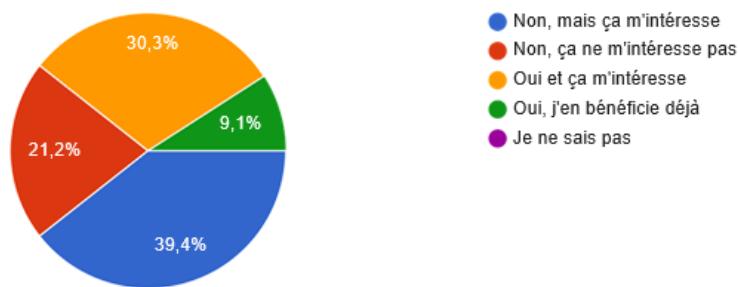


Figure 22 : Résultats de questionnaire : aide à l'accompagnement

2. Etude qualitative

2.1. Méthodologie

Sur base d'un guide d'entretien, des discussions avec six professionnels et experts ont eu lieu via Teams ou par téléphone.

Le guide d'entretien repose sur les apprentissages issus d'une revue de littérature faite en amont.

La prise de contact s'est effectuée via LinkedIn en fonction de leur profil, tenant compte de leurs métiers et centres d'intérêts.

Des questions ouvertes allant du général au spécifique ont été posées, cependant la priorité était de favoriser la discussion et de relancer l'interlocuteur au besoin. De ce fait, chaque entretien a sa propre personnalité et le guide n'a pas été toujours respecté.

2.2. Objectifs

Les entretiens étaient centrés sur les bénéfices, les difficultés et des exemples pratiques liés aux projets (big) data et de transformation digitale des PME/TPE.

Le but était de récolter des informations nouvelles via le développement du point de vue des interviewés et de faire surgir de nouvelles pistes de travail.

Les entretiens ne visent en aucun cas à généraliser les informations collectées. Ils servent plutôt de moyen pour approfondir et enrichir l'étude, tout en analysant les avis et les attitudes des interlocuteurs.

2.3. Limites

La principale limite se situe au niveau du nombre réduit d'interviewés et de leur temps d'engagement, en général entre 25 et 45 minutes par entretien.

Les éléments recueillis dépendent des connaissances, de l'expérience et du vécu des personnes interrogées ainsi que de leur volonté à répondre.

Enfin, la subjectivité de l'interprétation est également une limite bien que j'ai tenté de la réduire au maximum par la reformulation et la récapitulation des déclarations des participants, tout en suivant les règles d'analyse des études quantitatives et en évitant toute généralisation absolue.

2.4. Entretiens individuels semi-directifs

2.4.1. Profil des interlocuteurs

Tableau 8 : Entretiens : profils des interlocuteurs

Nom	Prénom	Fonction	Entreprise	Expertise
Minard	Benjamin	Digital Expert	Agoria	Transformation digitale, management, Design Thinking et marketing
Maréchal	Jean-François	Digital Advisor	GreenWin	Transformation digitale, consulting et formation
VandeVyver	Hubert	Digital Advisor	W2.0	Transformation digitale, consulting, formation et réseaux sociaux.
Stromboni	Carole	CEO – Innovation Advisor	Lab2034	Transformation et innovation digitale, Design Thinking, IT Project Management et expérience utilisateur.
Sterken	Vincent	IT Project Manager	Delaware	Digital Workplace Expert, BI, Information Management et analyse fonctionnelle, méthodes agiles.
Couvreur	Aurélie	CEO - Digital Transformation Expert	Strategy Wizard	Stratégie, innovation digitale, communication, marketing, Lean Management

2.4.2. Discussions

Le présent chapitre aborde, une par une, les thématiques des enquêtes. Ceci permet d'obtenir de plus amples informations au sujet des problématiques pratiques auxquelles les PME/TPE sont confrontées.

Les arguments convergents et divergents ainsi que les sujets évoqués spontanément par les interlocuteurs sont repris dans cette synthèse.

Favoriser l'appropriation technologique

Mots-clés : accompagnement, maturité, fracture digitale.

L'ensemble des interrogés considère que le fait d'être bien entouré c'est un atout majeur, peu importe la taille de l'organisation.

Faire appel à des experts est la manière la plus efficiente et moins risquée pour réussir un projet digital et d'exploitation des données. D'autant plus, si le projet concerne des technologies peu

récentes mais sous-exploitées car elles sont peu connues, surtout dans l'environnement des PME/TPE.

Il est particulièrement souhaitable pour les petites structures comptant avec peu ou pas de ressources techniques, d'oser demander de l'aide.

Les projets big data et IA dans le tissu économique du pays restent encore peu nombreux. Seulement 8% des données sont exploités dans l'industrie manufacturière et 73% des industriels belges n'utilisent pas l'IA. De plus, le taux d'échec pour les entreprises les plus audacieuses est encore très élevé, ce qui décourage les équipes et les dirigeants à de défaut d'une rentabilité palpable.

Selon Benjamin Minard, malgré les idées préconçues, la majorité des PME industrielles ne sont pas encore à l'aise avec ces technologies, encore moins avec la culture de projet orienté vers les utilisateurs finaux. Paradoxalement, 80% des PME voient le digital comme une opportunité pour optimiser leur processus mais elles ne consacrent pas assez leur temps et de ressources au sujet.

Une minorité travaille avec le Machine Learning car pour cela il faut disposer d'un énorme jeu de données. Par contre l'intelligence artificielle est un peu plus utilisée, néanmoins les entreprises restent méfiantes. Minard a partagé le cas de la société KSP qui aurait des difficultés à convaincre certains de ses prospects quant aux avantages de leurs produits et services, lesquels sont dotés d'intelligence artificielle. Les industriels sont retissant non par peur de l'espionnage industriel mais à cause de la perte de pouvoir de négociation puisque tout est contrôlé.

Agoria l'a compris grâce à son étude « Le digital : un état d'esprit » réalisée auprès de 400 PME industrielles belges. Aussi surprenant que cela puisse paraître, l'organisation sectorielle n'est pas parvenue à trouver de corrélations hormis la dimension culturelle.

Par conséquent, l'étude met l'accent sur l'importance de l'attitude et de la prédisposition à entamer et poursuivre des changements importants, parfois radicaux au sein des organisations.

Sans le bon état d'esprit, ni la taille ni les ressources suffisent pour bâtir d'elles-mêmes la maturité d'une entreprise.

Pour Benjamin Minard, avant d'envisager tout changement digital, le dialogue est la clé. Il commence par l'identification des problématiques business internes et externes, afin de déterminer la maturité des entreprises pour ensuite leur proposer les technologies et les outils les plus appropriés.

D'après Aurélie Couvreur, l'appropriation des technologies et des projets autour des données est une question de collaboration, pour améliorer le partage et la qualité des données ainsi que la lutte contre l'héritage des silos. Elle n'oublie pas de rajouter la formation.

Selon Jean-François Maréchal, l'appropriation des technologies et des outils démarre par une maîtrise du flux des activités et des informations qu'elles génèrent. D'après son expérience, dans l'accompagnement des PME/TPE, rares sont celles qui démontrent cette maturité.

De plus, Maréchal affirme qu'il est courant de se retrouver avec des entreprises se targuant d'être digitalisées. En effet, elles utilisent toute une série d'outils informatiques mais ces organisations n'échangent pas entre elles. Cette situation renvoie à la critique souvent évoquée par les interrogés sur les silos contre-productifs.

Afin de solutionner ce genre d'écueil, il précise qu'il est important d'accorder du temps à la communication. Dans son cas particulier, il tente d'apporter des conseils au-delà de l'accompagnement ponctuel. En effet, il cherche à couvrir l'ensemble des hypothèses pour construire des plans d'actions, et ainsi sensibiliser les entreprises à l'anticipation des problèmes.

« Il ne faut pas non plus numériser à tout prix » défend Maréchal. Les services de Cloud Computing par exemple, répondent à un ensemble de besoins très pertinents, mais ils sont devenus le talon d'Achille des entreprises sans plan B. Idem, pour les réseaux sociaux, il pense spécialement aux micros entreprises et affirme qu'il n'est pas judicieux de dépendre complètement des externes, qui du jour au lendemain pourraient disparaître ou changer leurs conditions d'utilisation.

Les points de vue de Hubert VandeVyver, de Carole Stromboni et de Vincent Sterken, confirment que l'introspection est l'étape sine qua non par laquelle il faut passer en premier, indifféremment de la nature du projet.

Favoriser la vision des entreprises

Mots-clés : vision, stratégie, dirigeant.

Benjamin Minard déclare que le retard des PME est d'environ 3 ans par rapport aux grandes entreprises.

D'après les enquêtés, la vision de l'entreprise concerne principalement les dirigeants et leur volonté à vouloir changer les choses, ce qui après une analyse approfondie de leur situation peut devenir une véritable stratégie.

Les avis convergent également à ce sujet. Selon les interlocuteurs, il n'y a aucun intérêt à proposer des plans d'action standards ni d'instaurer des technologies ou des méthodes par plaisir. A contrario, une définition précise des besoins, une feuille de route et l'engagement des dirigeants et des collaborateurs figurent parmi les facteurs de succès.

Différents programmes tentent d'aider les PME/TPE, les Indépendants et les entrepreneurs au sens large à démarrer leurs projets digitaux dans les meilleures conditions : DigiCoach, DigiSprint, DigiAcceleration et Data Coach proposés par Agoria mais aussi différents programmes organisés par des centres d'innovation et d'entreprenariat dans toute la Belgique. La définition des besoins, l'expérience client, l'innovation, la réinvention de l'offre et la captation

de valeur grâce aux données figurent parmi les principales thématiques couvertes par ces programmes.

« La vision c'est un thème sur lequel il faut souvent intervenir », commente VandeVyver. A ses yeux, les arguments rationnels coûts/bénéfices ne fonctionne pas toujours au sein des PME/TPE.

D'après Aurélie Couvreur, la favorisation de la vision commence par la saisie de leur utilité, l'un de ses rôles principaux c'est l'aide à la prise de conscience souvent à travers d'exemples et des cas à succès qui matérialisent les bénéfices potentiels. Pour ce qui est de l'exploitation des données, elle commente qu'il faut ajouter de l'intelligence aux données et c'est exactement ça dont les dirigeants et cadres des PME/TPE ont du mal à appliquer. Selon elle, très peu de gens ont les compétences pour exploiter les données pour les transformer en valeur ajoutée, et pire encore peu en savent qu'elles données faut-il saisir.

Aurélie Couvreur a aussi partagé deux exemples : le premier concerne Tartes de Françoise, qui pendant le confinement ont été forcés de fermer pendant plusieurs semaines leurs ateliers, ils se sont tournés vers les commandes en ligne grâce à l'habilitation du service sur leur plateforme web, rien que ça leur a permis de s'en sortir. Aujourd'hui ils font des prédictions pour mieux estimer la demande et diminuer le gaspillage.

Le deuxième exemple qu'elle a partagé c'est celui d'AXA Assurances, qui utilise les données pour agiliser le processus et la communication avec ses partenaires et clients. Avant la paperasse et les délais étaient très lourds, aujourd'hui il est très simple de déclarer un sinistre complètement en ligne, l'assuré peut directement voir sur leur application une estimation des frais accordés, ceci est possible grâce à l'exploitation des données historiques. Le montant est ensuite automatiquement versé sur le compte client dans les 48 heures.

Tout comme Sterken, ils différencient les petites et les grandes structures. Tous deux mettent en avant les habitus. Les grandes sociétés ont plus souvent l'habitude de faire appel à des externes pour se centrer sur leur métier. Tandis que les plus petites structures, même si elles veulent se digitaliser davantage, font moins confiance et tendent à vouloir tout faire par elles-mêmes.

Depuis le confinement, où la fracture digitale a mis en évidence l'écart de maturité des entreprises, VandeVyver essaie d'encourager davantage une stratégie flexible et réfléchie sur les éventualités à risque. Ceci en vue d'éviter toute paralysie opérationnelle.

Les difficultés et les blocages

Mots-clés : dirigeant, budget.

Une fois de plus, la plupart des interlocuteurs convergent. Ils soulèvent la responsabilité du dirigeant, sa méconnaissance et le budget comme les principaux freins à la mise en marche de projets numériques et de transformation digitale.

Vincent Sterken qui travaille avec des entreprises de taille moyenne mais surtout avec des grands groupes, pense que le facteur financier est une entrave majeure.

Jean François Maréchal et Hubert VandeVyver pensent aussi que le budget est un frein. Ils vont même plus loin : le budget est souvent sous-estimé contrairement à l'ampleur des projets qui est surestimée.

Tous trois affirment que le problème ne vient pas directement du coût lié aux solutions numériques, dans le cas où il s'agirait des licences, mais il se situe dans le paramétrage de celles-ci et la formation à leur utilisation. C'est à ce niveau que le chef d'entreprise est confronté au risque mais aussi au bénéfice d'investir dans ses collaborateurs.

D'un autre côté, l'implication et la volonté du chef d'entreprise sont tout aussi importantes. Traditionnellement dans les PME/TPE, c'est le dirigeant qui va initier les projets impliquant des changements conséquents. C'est sur lui que repose non seulement la responsabilité d'avancer ou de stagner, mais aussi de montrer l'exemple, en adoptant lui-même les instruments et les méthodes de manière adéquate.

Carole Stromboni a dit : « ...il faut balayer les escaliers par le haut ». Elle reste convaincue que le manager peut-être le meilleur initiateur de changement et d'innovation ou le plus grand blocage.

Ils ont tous revenus sur le fait qu'il faut prendre le temps de bien analyser ses besoins, définir sa vision, l'accompagner d'une stratégie solide et ne pas hésiter à se faire accompagner. La sélection des moyens et des méthodes sera plus objective et propice à leur développement.

L'humain au centre

Mots-clés : collaborateurs, utilisateurs, clients, collaboration, formation, culture.

Tout au long des discussions, des allusions concernant l'humain et sa subjectivité à s'adapter, à s'impliquer et à adopter de nouveaux comportements ont été abordées. A nouveau, la culture d'entreprise et les habitudes sont mises en avant. En effet, Benjamin Minard souligne que les projets impliquant des changements dans les comportements sont des projets multidisciplinaires où le collectif prime sur l'individuel.

Que ce soit une solution technologie ou une nouvelle approche, il est crucial d'inclure dès que possible les utilisateurs finaux à défaut de risquer son investissement mais aussi pour améliorer la qualité de la solution et surtout pour assurer son adoption.

Un autre élément évoqué par Minard c'est la stratégie « Top down - Bottom up », fortement recommandé, pour soutenir la co-création et la cohésion, mais cela requiert effectivement un changement de mentalité. Cette approche consiste à disposer d'un cadre et véhiculer un message clair à la fois sur la vision et sur les défis à moyen et long terme, tout en valorisant les retours des collaborateurs et leur prise d'initiative.

Les avis des répondants concernant la formation et le suivi sont unanimes. Ils considèrent que c'est l'une des meilleures manières pour combler l'écart et lutter contre la fracture numérique.

En résumé, quelques conseils

- Se rendre compte que le digital ce n'est pas un « nice to have » mais une urgence.
- Investir dans les personnes qui ont les connaissances pour accélérer la progression des projets et diminuer le risque d'échec.
- Instaurer une culture du changement transversale
- Promouvoir l'intelligence collective, indispensable pour développer les aptitudes nécessaires au changement et pour cocréer plus et mieux.
- Définir une gouvernance autour des données sans tarder pour les valoriser.
- Il est inutile de brûler les étapes, la base demeure dans la gestion et l'accessibilité des données avant tout data insight.
- Mener une gouvernance hybride « Top down – Bottom up » pour enrichir les échanges.
- Adopter l'agilité à travers une culture de projet.
- Eliminer les silos et favoriser la mobilité des données à travers toute l'organisation.
- Respecter les données sensibles des utilisateurs.

Partie 3

Approche Analytique

Avant de procéder à la confrontation de la théorie et de l'empirie, il convient de rappeler la question de recherche ainsi que les hypothèses qui ont guidé cette étude.

Question de recherche :

Quel est l'intérêt du big data pour les PME/TPE et comment favoriser l'appropriation du big data par ces acteurs ?

Hypothèses :

Le big data et les technologies associées (les nouvelles capacités d'analyse prévisionnelle, IA, etc.) sont d'un intérêt considérable pour les entreprises en général, cela inclut notamment les PME/TPE et tout entrepreneur, puisqu'elles apportent un avantage compétitif non négligeable à niveau commercial, en plus d'être une source d'inspiration pour les produits et services du futur. La règle du pauvre qui devient de plus en plus pauvre et du riche qui devient de plus en plus riche, est désormais plus aussi délimitée, l'accès et la manipulation du big data se démocratise et permet aux petites structures de s'intégrer elles aussi dans l'ère du numérique.

Les données et la course technologique dans laquelle nous évoluons ont été constamment placées au centre de cette recherche.

La revue de la littérature a repris les éléments techniques les plus percutants, une théorie indispensable pour la bonne compréhension du monde numérique qui nous enveloppe. De la même manière, les éléments managériaux directement affectés tant positivement que négativement ont été mis en lumière.

Aujourd'hui, les avancées technologiques estompent de plus en plus la limite entre le physique et le virtuel. Tant des changements intéressants depuis quelques décennies à peine, font que l'on s'interroge sur leurs effets présents et futurs.

Dans le cadre de ce travail, l'intention s'est orientée vers la situation organisationnelle et numérique auxquelles font face les PME/TPE.

L'approche théorique a révélé que malgré une remarquable démocratisation des technologies et d'outils digitaux (logiciels SaaS) de support et de reporting, les PME/TPE se retrouvent dans une condition défavorable par rapport aux grands groupes, en particulier les GAFAM qui indiscutablement sont les grands dominants.

Sans surprise, la source de valeur des GAFAM, ce sont les données dont ils exploitent avec brillance. Nos interactions sociales et professionnels sont en quelque sorte manœuvrées par ces

quelques acteurs. Certains sont contre d'autres pour, j'éviterai rentrer dans ce débat de dimension plutôt éthico-social. Mais d'un point de vue économique et managérial, il y en a de quoi s'inspirer.

A l'heure actuelle, ce n'est pas l'accès aux outils d'aide à la gestion et à la décision qui pose problème, puisqu'ils sont devenus accessibles. Des alternatives aux développement à partir des rien et aux licences, comme le sont les formules par abonnement, permettent de ne pas avoir à réaliser des investissements considérables et de garder la liberté s'arrêter à tout moment.

En revanche, la littérature tout comme les avis des experts interviewés, affirment que la problématique s'attache à deux facteurs principaux : les compétences et la culture d'entreprise.

En effet, les faibles compétences analytiques figurent comme un obstacle, car elles déterminent quelles données extraire et comment les analyser. De même, la communication peu aisée entre les équipes métier et IT représentent un frein au bon flux d'informations.

De plus, les décideurs sont souvent signalés soit comme facilitateurs soit comme un frein eux aussi, dû à une méconnaissance des opportunités et des moyens pour profiter de leur capital de données. Selon l'expérience des interviewés, peut d'entreprises exploitent leur données internes avec intelligence, et profitent encore moins des données externes provenant des portails Open Data ou des réseaux sociaux.

Leur volonté des décideurs est aussi critiquée, car dans les PME/TPE traditionnelles ce sont eux qui prennent les décisions stratégiques et sur qui repose la responsabilité de vouloir changer les comportements, cela inclus la culture d'entreprise, la vision, la stratégie et la prise de risque.

En effet, les données sont au cœur de la transformation digitale, car elles jouent un rôle important au niveau de la stratégie d'entreprise. Les objectifs déterminent leur usage, ça peut aller de la simple analyse des historiques, jusqu'à la codification des algorithmes prédictifs puissants.

Une entreprise digitalisée n'est pas forcément numériquement mature, tout dépendra de l'usage effectif et de la synergie des ressources humaines et numériques internes et externes à l'entreprise.

Un signe de maturité digital que l'on retrouve dans la littérature mais aussi à travers du partage des experts, c'est le travail collaboratif, c'est-à-dire un milieu hautement intégré, ne fonctionnant plus en silos. De cette manière, les données de l'ensemble des départements peuvent être agrégées, à fin d'alimenter d'autres cas d'usage.

En vue de favoriser l'appropriation des projets digitaux, mais surtout des projets à haute valeur comme le sont l'exploitation des données et l'intelligence artificielle, la montée en compétences des collaborateurs est fondamentale, ainsi que les efforts pour la part de la direction pour faciliter la formation, encourager une organisation aplatie qui soutient les échanges entre les différentes équipes. Synergies étant confirmées comme source d'innovation et d'efficacité.

Les connaissances « digito-analytiques » et techniques sont devenues des réels atouts. Pour que la passerelle vers une culture de projets soit la plus effective possible, les équipes métier et tech doivent être en mesure de parler un langage commun.

Les PME/TPE comptant avec peu ou pas de ressources techniques, la meilleure manière de réussir et accélérer leur projets digito-analytiques et de se tourner vers des organisations subventionnées et des experts qui peuvent les épauler à différents niveaux, notamment pour la définition de leur vision, de leurs objectifs, dans le choix des outils et des technologies les plus adaptés à leurs ambitions et capacités.

Pour cela, toute une série d'organisations avec le soutien des pouvoirs publics, sont à disposition pour aider même les plus petites structures (cf. 7.4. Se faire aider et accompagner).

Un autre aspect qui est ressortit dans la littérature et dans les entretiens, toujours sur la dimension humaine, c'est l'intelligence collective, laquelle est profondément basée sur la communication, la proactivité et la confiance. En effet, il s'agit de créer une ambiance de travail ouverte et agréable, facilitant encore une fois les flux d'informations et l'innovation puisque la prise des initiatives et le partage de nouvelles idées sont aussi encouragées. L'engagement, le bien-être et la valorisation des individus se voient améliorées, ce qui permet entre autres de garder son équipe motivée et d'attirer des nouveaux talents.

Revisiter son offre, améliorer les services rendus, connaître davantage ses clients, capitaliser ses données, rendre les processus plus efficaces, anticiper les problèmes et identifier des nouvelles opportunités business, ces quelques exemples sont souvent le résultat d'un changement de perspective grâce à des techniques des design thinking, et bien entendu à de l'intelligence appliquée aux données. Bref, il suffit sortir de sa zone de confort et entamer une démarche proactive.

En outre, la crise du Covid a bien mis en lumière la fracture numérique. Les entreprises les plus en adéquation avec le paradigme digital ont subi moins les effets néfastes du confinement, d'autres ont même augmenté leurs marges, contrairement à celles qui ont été obligés d'agir dans l'urgence, ce qui certes a accéléré leur digitalisation mais sous des conditions oppressives et non froidement réfléchies.

Selon les informations obtenues grâce à la recherche documentaires et aux entretiens individuels, il est clair que la plupart des PME sont encore loin de la transformation digitale. De plus les projets (big) data et IA sont assez précaires. Une autre cause, la préparation des données en amont. Une phase obligatoire avant d'exploiter quelque donnée, mais jusqu'à présent fortement handicapé par la culture en silos.

Malgré cette pénible étape, les PME/TPE ont un avantage, elles possèdent des small data, ce qui facilite leur traitement pour finalement obtenir des smart data, c'est-à-dire des informations stratégiquement et fonctionnelles prêtes à l'emploi.

Par ailleurs, les résultats du sondage en ligne, montrent que pour la plupart des répondants les données et les technologies associées ainsi que les outils de support, sont une opportunité voire même une nécessité. Nonobstant, les supports de stockage restent très dispersés, de même que le format des données. Pour 51% des répondants l'analyse de données externe n'a aucune utilité. Intéressant, après avoir plongé dans une documentation riche sur l'exploitation des Open Data et des données web/sociaux.

Bien que l'enquête ne puisse pas être extrapolée à l'ensemble de la population visée, puisque l'échantillon était non représentatif. Il est toutefois, intéressant de remarquer que l'intérêt pour les nouvelles technologies associées au big data s'est fait ressentir principalement auprès de micro-entreprises et des porteurs de projets.

L'enquête quantitative figure de base exploratoire très réduite. Cependant, elle permet de s'interroger sur les moyens d'attendre les populations jusqu'ici les moins impliqués, notamment les indépendants. Ainsi que les moyens en général pour répandre d'avantage le message : le digital est une urgence et les données une source de croissance non négligeable.

Derniers conseils pour faciliter l'appropriation du (big) data par les PME/TPE

Enfin, pour revenir à la question de recherche, l'appropriation du (big) data, commence par la prise de conscience de sa valeur ajoutée et par un changement de culture, qui tendra à promouvoir la collaboration transversale.

Ensuite, se fixer des objectifs réalistes (Quelles données exploiter ? Pourquoi faire ?).

Bien s'entourer et encourager la formation.

Pour finir, garder à l'esprit que les contraintes techniques et réglementaires ne sont pas insurmontables.

Nouvelles pistes d'exploration

Nous avons vu que le big data et l'exploitation des données tout court, est effectivement à la portée des PME/TPE. Cependant la culture, les habitudes, les dirigeants et les compétences limitent ces structures de profiter efficacement des nouvelles technologies.

Bien que, les logiciels (SaaS) permettent déjà grandement simplifier et améliorer le quotidien des entreprises, le travail en silos persiste encore, ce qui empêche l'optimisation des activités, l'exploitation des données ainsi que l'innovation.

De plus, en fonction de sa taille, de son activité et de ses besoins, l'entreprise pourrait en accumuler une multitude d'outils, qui ne communiquent pas entre eux, et pour solutionner le problème elle aurait sûrement besoin d'engager des consultants ou des développeurs, pour développer des applications spécifiques, ce qui pourrait très rapidement faire exploser les coûts.

Alors, les outils No-code sont-ils la meilleure alternative ? Voilà un sujet extrêmement intéressant, qui mérite une lui seul une étude approfondie. Des sociétés encore jeunes affirment

que leurs solutions No-code révolutionnera le mode de fonctionnement PME/TPE. Rapide, pas cher, intuitif et sur mesure. Tant de promesses dont j'ai hâte de découvrir.

Conclusion

De toute évidence, les données sont une source à haut potentiel pour les entreprises. Les petites et moyennes structures qui n'ont pas encore décidé à s'aventurer sur ce terrain, sont malheureusement en train de laisser passer des belles opportunités.

Parmi l'infinité de bénéfices : amélioration de la connaissance et le l'expérience client, anticiper la demande et les risques, gérer les stocks, prendre des meilleures décisions marketing et budgétaires, automatisation des tâches chronophages et récurrentes, etc.

Les données constituent la base de l'économie numérique et très probablement le phénomène ne fera que s'accentuer. Des nouvelles technologies et usages surgiront, et chaque entreprise devrait être en mesure de les utiliser intelligemment, autrement les données resteront inexploitées et tous les bénéfices potentiels seront ignorés.

Au fil du temps, seules les structures ayant fait preuve d'adaptation au paradigme numérique et qui auront mis en place des stratégies visant valoriser leurs données, seront plus enclins à assurer leur pérennité.

Pour les PME/TPE encore sceptiques, il faut se dire que les données sont là, les outils aussi ainsi les aides à l'accompagnement pour les aider à surmonter les contraintes techniques et réglementaires.

Si aujourd'hui la digitalisation et le travail collaboratif est une urgence, ce sont les projets les plus audacieux qui représentent un réel impact économique et temporaire.

Des cas d'usages probants à l'échelle des petites et moyennes structures nous montrent que valoriser ses données c'est possible. D'une part, grâce aux systèmes cloud, les logiciels en SaaS, les API et à la capacité des processeurs. En effet, tout l'écosystème informatique rend le traitement des données possible. D'un autre côté, la montée exponentielle des données le permet aussi.

Les PME/TPE qui n'exploitent pas leurs données sont tout simplement en situation de déséquilibre. Par, conséquent elles sont plus vulnérables aux variations micro et macroéconomiques.

Nonobstant, la capacité d'une entreprise à utiliser les data comme un actif stratégique est étroitement lié à sa maturité en termes de culture numérique et de gouvernance des données.

La manière la plus simple, pour que les PME/TPE puissent commencer à tirer profit de leur capital de données, c'est commencer par une stratégie de digitalisation où les data et la communication intégrée seront au cœur.

Ensuite entamer une stratégie d'analyse, privilégiant les données en interne de qualité (small data). Enfin progresser sur la route de la transformation digitale continue.

Bibliographie

- AdN - *Baromètre Entreprises*. (s. d.). Consulté 16 août 2021, à l'adresse <https://www.odwb.be/explore/dataset/entreprises-adn/information/>
- Amir-Aslani, A., & Bhajun, R. (2015). *Les 7 «V» Piliers du Big Data*.
- Anaconda / State of Data Science 2021. (2021). Anaconda. <https://www.anaconda.com/state-of-data-science-2021>
- Baromètre des entreprises : +15,1% de créations de sociétés en un an malgré le coronavirus*. (s. d.). Consulté 16 août 2021, à l'adresse <https://www.notaire.be/nouveautes/detail/barometre-des-entreprises-15-1-de-creations-de-societes-en-un-an-malgre-le-coronavirus>
- Bastin, G., & Tubaro, P. (2018). Le moment big data des sciences sociales. *Revue française de sociologie*, Vol. 59(3), 375-394.
- Benavent, C. (2014). Big data : No best way. *Libellio d'Aegis*, 10, 5-14.
- Bheemaiah, K., Esposito, M., & Tse, T. (s. d.). *What is machine learning?* The Conversation. Consulté 27 juin 2021, à l'adresse <http://theconversation.com/what-is-machine-learning-76759>
- Buchanan, B. G. (2005). A (very) brief history of artificial intelligence. *Ai Magazine*, 26(4), 53-53.
- Burquier, B. (2009). *Business Intelligence avec SQL Server 2008 : Mise en oeuvre d'un projet décisionnel*. Dunod.
- <https://books.google.be/books?hl=fr&lr=&id=m71bVkzbF70C&oi=fnd&pg=PA1&dq=burquier+business+intelligence&ots=6JBbrfdMRx&sig=vHCsTX4L0E8HRvicFJh7aMBvGDc#v=onepage&q=burquier%20business%20intelligence&f=false>
- Chen, H., Chiang, R. H. L., & Storey, V. C. (2012). Business Intelligence and Analytics : From Big Data to Big Impact. *MIS Quarterly*, 36(4), 1165-1188. <https://doi.org/10.2307/41703503>
- Chokogoue, J. (2018). *Maîtrisez l'utilisation des technologies Hadoop : Initiation à l'écosystème Hadoop*. Editions Eyrolles.
- https://books.google.be/books?id=JMFeDwAAQBAJ&printsec=frontcover&dq=Chokogoue&hl=fr&sa=X&redir_esc=y#v=onepage&q=Chokogoue&f=false
- Comment déployer avec succès un projet Big Data ? - CETIC - Your connection to ICT research.* (s. d.). Consulté 16 août 2021, à l'adresse <https://www.cetic.be/Comment-deployer-avec-succes-un-projet-Big-Data>
- Commité économique et social européen. (2017). *Etude sur les mégadonnées (big data)*. <https://www.eesc.europa.eu/sites/default/files/resources/docs/qe-04-17-306-fr-n.pdf>
- Data Literacy Program. (s. d.). Qlik. Consulté 16 août 2021, à l'adresse <https://www.qlik.com/fr-fr/services/data-literacy-program>
- Database—Eurostat. (s. d.). Consulté 16 août 2021, à l'adresse <https://ec.europa.eu/eurostat/web/main/data/database>
- Davenport, T. (2014). *Stratégie Big Data*. Pearson.
- Davenport, T. H., & Westerman, G. (2018a). *Why So Many High-Profile Digital Transformations Fail*. 5.

- Davenport, T. H., & Westerman, G. (2018b, mars 9). Why So Many High-Profile Digital Transformations Fail. *Harvard Business Review*. <https://hbr.org/2018/03/why-so-many-high-profile-digital-transformations-fail>
- Davenport, T., & Harris, J. (2017). *Competing on Analytics : Updated, with a New Introduction: The New Science of Winning*. Harvard Business Press. [https://books.google.be/books?hl=fr&lr=&id=LW9GDgAAQBAJ&oi=fnd&pg=PT20&dq=intelligence+degre+Davenport,+T.,+%26+Harris,+J.+\(2017\).&ots=d0j3e0MrS&sig=D4RuhvzEUPLVA5AqSg64TsR0M8&redir_esc=y#v=onepage&q&f=false](https://books.google.be/books?hl=fr&lr=&id=LW9GDgAAQBAJ&oi=fnd&pg=PT20&dq=intelligence+degre+Davenport,+T.,+%26+Harris,+J.+(2017).&ots=d0j3e0MrS&sig=D4RuhvzEUPLVA5AqSg64TsR0M8&redir_esc=y#v=onepage&q&f=false)
- De Mauro, A., Greco, M., & Grimaldi, M. (2016). A formal definition of Big Data based on its essential features. *Library Review*, 65(3), 122-135. <https://doi.org/10.1108/LR-06-2015-0061>
- Delorme, P., & Djellalil, J. (2015). *La transformation digitale : Saisir les opportunités du numérique pour l'entreprise*. Dunod. https://books.google.be/books?hl=fr&lr=&id=jknpCQAAQBAJ&oi=fnd&pg=PP3&dq=delorme+et+djellalil+la+transformation+digitale&ots=3RZXfOXUcg&sig=ugrDToPvScgH2Bg1ExIJVMmRHQ&redir_esc=y#v=onepage&q=delorme%20et%20djellalil%20la%20transformation%20digitale&f=false
- Denoix, A. (2018). *Big Data, Smart Data, Stupid Data... : Comment (vraiment) valoriser vos données*. Dunod.
- Deshpande, P. S., Sharma, S. C., & Peddoju, S. K. (2019). Predictive and Prescriptive Analytics in Big-data Era. In P. S. Deshpande, S. C. Sharma, & S. K. Peddoju, *Security and Data Storage Aspect in Cloud Computing* (Vol. 52, p. 71-81). Springer Singapore. https://doi.org/10.1007/978-981-13-6089-3_5
- Diebold, F. X. (2012). *On the Origin(s) and Development of the Term « Big Data »* (SSRN Scholarly Paper ID 2152421). Social Science Research Network. <https://doi.org/10.2139/ssrn.2152421>
- Digital Economy Compass* 2019. (2019). Statista. <https://www.statista.com/study/52194/digital-economy-compass/>
- Digital Wallonia. (s. d.). *Economie numérique*. Consulté 30 juillet 2021, à l'adresse <https://www.digitalwallonia.be/fr/strategie-numerique/economie-numerique>
- Dropbox. (s. d.). À quoi correspond 1 To de stockage ? Consulté 16 août 2021, à l'adresse <https://www.dropbox.com/fr/features/cloud-storage/how-much-is-1tb>
- Dulal, K., & Binkele, C. (2021). *Disruptive Business Model Innovation : Incumbent's perspectives and approaches to competitive advantage generation*. <https://lup.lub.lu.se/student-papers/search/publication/9058204>
- Dumbill, E. (2013). Making Sense of Big Data. *Big Data*, 1(1), 1-2. <https://doi.org/10.1089/big.2012.1503>
- EGE Open Data. (2020). Ecole de Guerre Economique. https://www.ege.fr/home_eInvoicing_in_Belgium—CEF_Digital. (s. d.). CEF Digital. Consulté 27 juillet 2021, à l'adresse <https://ec.europa.eu/cefdigital/wiki/cefdigital/wiki/display/CEFDIGITAL/eInvoicing+in+Belgium>
- Étude : « Le digital: un état d'esprit ». (s. d.). Agoria. Consulté 16 août 2021, à l'adresse <https://www.agoria.be/fr/Livre-blanc-Le-digital-un-etat-d-esprit>
- Eychenne, Y., & Cointot, J.-C. (2014). *La Révolution Big data : Les données au cœur de la transformation de l'entreprise*. Dunod.

- Fernandez, A. (2013). *Les nouveaux tableaux de bord des managers : Le projet Business Intelligence clés en main.*. Editions Eyrolles.
<https://books.google.be/books?hl=fr&lr=&id=Nbs5CM15JgYC&oi=fnd&pg=PR5&dq=fernandez+les+nouveaux+tableaux+de+bord+des+managers+le+projet&ots=mXYZDY0hpq&sig=Bz-AIWcjqxgAHVuAh0GvXevUfMc#v=onepage&q=fernandez%20les%20nouveaux%20tableaux%20de%20bord%20des%20managers%20le%20projet&f=false>
- Française, A. (2021). *Dictionnaire de l'Académie française*. <https://www.dictionnaire-academie.fr/article/A9N0819>
- Full Text.* (s. d.). Consulté 16 août 2021, à l'adresse https://www.cubility.com.au/wp-content/uploads/2018/11/dup_strategy-not-technology-drives-digital-transformation.pdf
- Gartner Says Advanced Analytics Is a Top Business Priority.* (s. d.). Gartner. Consulté 16 août 2021, à l'adresse <https://www.gartner.com/en/newsroom/press-releases/2014-10-21-gartner-says-advanced-analytics-is-a-top-business-priority>
- Georges, B. (2019). Intelligence artificielle : De quoi parle-t-on? *Constructif*, 3, 5-10.
- Ghernaouti, S. (2016). *Cybersécurité - 5e éd. : Sécurité informatique et réseaux*. Dunod.
https://books.google.be/books?hl=fr&lr=&id=kYdIDQAAQBAJ&oi=fnd&pg=PR3&dq=Ghernaouti_+2016&ots=JnX_i4XJAH&sig=-4iI2q_zjX3B0Pi4yx5Sv_88XXM&redir_esc=y#v=onepage&q=Ghernaouti%2C%202016&f=false
- Habibu, T. (2013). *Parallel Data Analytics for Business Intelligence Real-Time Online Analytical Processing (OLAP) For Multi-Core and Cloud Architectures*.
https://www.researchgate.net/publication/319458909_Parallel_Data_Analytics_for_Business_Intelligence_Real-Time_Online_Analytical_Processing_OLAP_For_Multi-Core_and_Cloud_Architectures
- Hadoop : Qu'est-ce que c'est et comment apprendre à l'utiliser ?* (2021).
<https://datascientest.com/hadoop>
- Hamel, M.-P., & Marguerit, D. (2013). Quelles possibilités offertes par l'analyse des big data pour améliorer les téléservices publics ? *Revue française d'administration publique*, N° 146(2), 437-448.
- Hinings, B., Gegenhuber, T., & Greenwood, R. (2018). Digital innovation and transformation : An institutional perspective. *Information and Organization*, 28(1), 52-61.
<https://doi.org/10.1016/j.infoandorg.2018.02.004>
- Home | Hub Créatif Charleroi Métropole.* (s. d.). Consulté 16 août 2021, à l'adresse <https://www.hub-charleroi.be/>
- Indice relatif à l'économie et à la société numériques (DESI) 2020 : Questions et.* (2020). [Text]. European Commission - European Commission.
https://ec.europa.eu/commission/presscorner/detail/fr/QANDA_20_1022
- Intel IT Center. (2012). *Big Data 101 : Unstructured Data Analytics*. 4.
- Iqbal, M., Kazmi, S. H. A., Manzoor, A., Soomrani, A. R., Butt, S. H., & Shaikh, K. A. (2018). A study of big data for business growth in SMEs : Opportunities amp; challenges. *2018 International Conference on Computing, Mathematics and Engineering Technologies (iCoMET)*, 1-7.
<https://doi.org/10.1109/ICOMET.2018.8346368>
- Julia, L. (2019). *L'intelligence artificielle n'existe pas.* edi8.

- Junghans, P. (2015). Le Big data pour construire une information d'anticipation. *I2D - Information, donnees documents, Volume 52(4)*, 12-14.
- Kane, G. C., Palmer, D., Phillips, A. N., Kiron, D., & Buckley, N. (2015). Strategy, not technology, drives digital transformation. *MIT Sloan Management Review and Deloitte University Press*, 14(1-25). https://www.cubility.com.au/wp-content/uploads/2018/11/dup_strategy-not-technology-drives-digital-transformation.pdf
- Karoui, M., Davauchelle, G., & Dudezert, A. (2014). Karoui M., Devauchelle G., Dudezert A. (2014), « Big Data : Mise en perspective et enjeux pour les entreprises », N° Spécial « Big Data », Revue Ingénierie des Systèmes d'Information, 19 (3), 73-92 Hermès. *Ingénierie des systèmes d'information*, 19, 73-92. <https://doi.org/10.3166/isi.19.3.73-92>
- Khan, N., Alsaqer, M., Shah, H., Badsha, G., Abbasi, A. A., & Salehian, S. (2018). The 10 Vs, Issues and Challenges of Big Data. *Proceedings of the 2018 International Conference on Big Data and Education*, 52-56. <https://doi.org/10.1145/3206157.3206166>
- Khan, N., Naim, A., Hussain, M. R., Naveed, Q. N., Ahmad, N., & Qamar, S. (2019). The 51 V's Of Big Data : Survey, Technologies, Characteristics, Opportunities, Issues and Challenges. *Proceedings of the International Conference on Omni-Layer Intelligent Systems*, 19-24. <https://doi.org/10.1145/3312614.3312623>
- Kirk, A. (2016). *Data visualisation : A handbook for data driven design*. Sage. https://books.google.be/books?hl=fr&lr=&id=wNpsDAAAQBAJ&oi=fnd&pg=PP1&dq=kirk+data+visualisation+a+handbook+for+data+&ots=AGraYsKch7&sig=AdP-xsFyyQtnyHUGLDNYhsvn9dw&redir_esc=y#v=onepage&q=kirk%20data%20visualisation%20a%20handbook%20for%20data&f=false
- Le règlement général sur la protection des données—RGPD / CNIL*. (s. d.). Consulté 16 août 2021, à l'adresse <https://www.cnil.fr/fr/reglement-europeen-protection-donnees>
- Lecocq, X., Demil, B., & Warnier, V. (2006). Le business model, un outil d'analyse stratégique. *L'Expansion Management Review*, N° 123(4), 96-109.
- L'économie collaborative en Belgique et en Europe / SPF Economie*. (s. d.). Consulté 16 août 2021, à l'adresse <https://economie.fgov.be/fr/themes/entreprises/economie-durable/economie-collaborative/leconomie-collaborative-en-0>
- L'intelligence artificielle (IA) pour entreprises et PME en Belgique (SPF economie)*. (2021). <https://www.intelligenceartificiellepourpme.be/>
- Matillion. (2020). *Gaining Time, Savings, and Insights via Cloud-Powered Data Transformation*. <https://www.matillion.com/resources/ebooks/gaining-time-savings-and-insights-via-cloud-powered-data-transformation>
- Mayer-Schönberger, V., & Cukier, K. (2013). *Big Data : A Revolution that Will Transform how We Live, Work, and Think*. Houghton Mifflin Harcourt. https://books.google.be/books?hl=fr&lr=&id=uy4lh-WEhhIC&oi=fnd&pg=PP1&dq=Big+Data:+A+Revolution+that+Will+Transform+how+We+Live,+Work,+and+Think&ots=JuebboCXEM&sig=7j8fQXW4cncwO21aiXCDFsPVmd0&redir_esc=y#v=onepage&q=Big%20Data%3A%20A%20Revolution%20that%20Will%20Transform%20how%20We%20Live%2C%20Work%2C%20and%20Think&f=false

- Mazumdar, S., & Dhar, S. (2015). Hadoop as Big Data Operating System—The Emerging Approach for Managing Challenges of Enterprise Big Data Platform. *2015 IEEE First International Conference on Big Data Computing Service and Applications*, 499-505. <https://ieeexplore.ieee.org/abstract/document/7184922/>
- Neamah, M. (2016, mars 4). *Big Data Security*. https://www.researchgate.net/publication/296702778_Big_Data_Security
- Negash, M. (2017, août 18). *Wayback Machine*. <https://web.archive.org/web/20170818083206/http://ageconsearch.umn.edu/bitstream/126793/2/Negash.pdf>
- Negash, S., & Gray, P. (2008). Business Intelligence. In F. Burstein & C. W. Holsapple (Éds.), *Handbook on Decision Support Systems 2 : Variations* (p. 175-193). Springer. https://doi.org/10.1007/978-3-540-48716-6_9
- NIPS. (2007). *NeurIPS 2007 Overview*. <https://nips.cc/Conferences/2007>
- NIST Big Data Public Working Group Definitions and Taxonomies Subgroup. (2015). *NIST Big Data Interoperability Framework : Volume 1, Definitions* (NIST SP 1500-1; p. NIST SP 1500-1). National Institute of Standards and Technology. <https://doi.org/10.6028/NIST.SP.1500-1>
- Notes de lecture. (2015). *I2D - Information, donnees documents, Volume 52(2)*, 80-87. <https://doi.org/10.3917/i2d.152.0080>
- OECD. (2021). *The Digital Transformation of SMEs*. OECD. <https://doi.org/10.1787/bdb9256a-en>
- Oliveira, M. I. S., & Lóscio, B. F. (2018). What is a data ecosystem? *Proceedings of the 19th Annual International Conference on Digital Government Research: Governance in the Data Age*, 1-9. <https://dl.acm.org/doi/abs/10.1145/3209281.3209335>
- Osmundsen et al. - 2018—*Digital Transformation Drivers, Success Factors, .pdf*. (s. d.). Consulté 1 août 2021, à l'adresse <https://aisel.aisnet.org/cgi/viewcontent.cgi?article=1004&context=mcis2018>
- Osmundsen, K., Iden, J., & Bygstad, B. (2018a). Digital Transformation : Drivers, Success Factors, and Implications. *Digital Transformation*, 16.
- Osmundsen, K., Iden, J., & Bygstad, B. (2018b). Digital Transformation : Drivers, Success Factors, and Implications. *MCIS*, 37. https://books.google.be/books?hl=fr&lr=&id=fS2eDwAAQBAJ&oi=fnd&pg=PT15&dq=fayon+tar+2019&ots=7O9tPcrTKk&sig=nHtYlrtKsjiBwtdsQSEZRhUs&redir_esc=y#v=onepage&q=fayon%20tar+2019&f=false
- Osterwalder, A., Pigneur, Y., & Tucci, C. L. (2005). Clarifying Business Models : Origins, Present, and Future of the Concept. *Communications of the Association for Information Systems*, 16(1). <https://doi.org/10.17705/1CAIS.01601>
- Paschek, D., Luminosu, C. T., & Draghici, A. (2017). Automated business process management—in times of digital transformation using machine learning or artificial intelligence. *MATEC Web of Conferences*, 121, 04007. https://www.matec-conferences.org/articles/matecconf/abs/2017/35/matecconf_mse2017_04007/matecconf_ms_e2017_04007.html
- Peillon, S. (2016). La servicisation des entreprises industrielles. *La Revue des Sciences de Gestion*, N° 278-279(2), 131-140.
- Polo Mejía, O. (2015). *BIG DATA : Une révision*. <https://doi.org/10.13140/RG.2.1.3752.7760>

- Ponsard, C., Touzani, M., & Majchrowski, A. (2018). Synthèse des méthodes de conduite de projets Big Data et des retours collectés lors de pilotes industriels. *Ingenierie des Systèmes d'Information*, 23(1), 9.
- Press, G. (s. d.). *A Very Short History Of Big Data*. Forbes. Consulté 11 avril 2021, à l'adresse <https://www.forbes.com/sites/gilpress/2013/05/09/a-very-short-history-of-big-data/>
- Press, G. (2016). *Cleaning Big Data : Most Time-Consuming, Least Enjoyable Data Science Task, Survey Says*. Forbes. <https://www.forbes.com/sites/gilpress/2016/03/23/data-preparation-most-time-consuming-least-enjoyable-data-science-task-survey-says/>
- Pries, K. H., & Dunnigan, R. (2015). *Big Data Analytics : A practical guide for managers*. CRC Press.
- Quelles sont les différences entre l'Intelligence Artificielle, le Machine Learning et le Deep Learning ? - Quora*. (s. d.). Consulté 16 août 2021, à l'adresse <https://fr.quora.com/Quelles-sont-les-differences-entre-lIntelligence-Artificielle-le-Machine-Learning-et-le-Deep-Learning>
- Rappa, M. A. (2004). The utility business model and the future of computing services. *IBM Systems Journal*, 43(1), 32-42. <https://doi.org/10.1147/sj.431.0032>
- Samuel, A. L. (1959). Some studies in machine learning using the game of checkers. *IBM Journal of research and development*, 3(3), 210-229.
- Schiuma, G., & Carlucci, D. (2018). *Big Data in the Arts and Humanities : Theory and Practice*. CRC Press.
- Schmidt, A. L., & Scaringella, L. (2020). Uncovering disruptors' business model innovation activities : Evidencing the relationships between dynamic capabilities and value proposition innovation. *Journal of Engineering and Technology Management*, 57, 101589.
- Schwertner, K. (2017). Digital transformation of business. *Trakia Journal of Sciences*, 15(1), 388-393.
- Tabesh, P., Mousavidin, E., & Hasani, S. (2019). Implementing big data strategies : A managerial perspective. *Business Horizons*, 62(3), 347-358. <https://doi.org/10.1016/j.bushor.2019.02.001>
- Teboul, B., & Berthier, T. (2015). Valeur et Véracité de la donnée : Enjeux pour l'entreprise et défis pour le Data Scientist. In *Actes du colloque " La donnée n'est pas donnée " École Militaire – 23 mars 2015*. <https://hal.archives-ouvertes.fr/hal-01152219>
- Timmers, P. (1998). Business Models for Electronic Markets. *Electronic Markets*, 8(2), 3-8. <https://doi.org/10.1080/10196789800000016>
- Transformation digitale : Les entreprises peinent à avancer malgré leurs investissements. (2018, juillet 3). *Capgemini France*. <https://www.capgemini.com/fr-fr/news/transformation-digitale-les-entreprises-peinent-a-avancer-malgre-leurs-investissements/>
- Warnier, V., Lecocq, X., & Demil, B. (2004). Le business model : L'oublié de la stratégie? *Actes de la 13ième conférence AIMS, Le Havre*. http://www.businessmodelcommunity.com/fs/Root/8u9mp-Warnier_Lecocq_Demil.pdf
- Yoo, Y. (2010). *Digitalization and Innovation*. https://www.researchgate.net/publication/48926382_Digitalization_and_Innovation
- Zhang, S. (2019). *Deep Learning for Unstructured Data by Leveraging Domain Knowledge—ProQuest*. <https://www.proquest.com/openview/d35b7e5973d58cab8473a6b3c28b92bd/1?pq-origsite=gscholar&cbl=18750&diss=y>
- Zott, C., Amit, R., & Massa, L. (2011). *The Business Model : Recent Developments and Future Research*. <https://journals.sagepub.com/doi/abs/10.1177/0149206311406265>