Livre **Blanc**

IDENTIFIONS

ENSEMBLE LES IMPACTS SUR

LE SI DANS 3 ANS ET DES SOLUTIONS ENVISAGEABLES !

30 Novembre 2018

Auteurs : Nicolas Carpentier, Nicolas Huet, Nicolas Olive, Arielle Petit, Matthieu Poletti, Emmanuel Real, Guillaume Real, Nicolas Roman.

cesi ÉCOLE D'INGÉNIEURS

sopra steria

Préambule

Etudiants en dernière année en école d'ingénieur en informatique, nous avons réalisé cette veille technologique pour présenter nos recherches lors d'une conférence visant un public de professionnels et d'étudiants dans le domaine de l'informatique.

Dans notre école, le CESI, cet évènement est une tradition et a plusieurs buts.

Le premier est de promouvoir notre école auprès des entreprises et des étudiants afin de gagner en notoriété.

Le second est de nous entraîner à effectuer de la veille technologique pour être informés des dernières nouveautés.

Cette année le sujet de la conférence porte sur le SI et son évolution d'ici 3 ans avec des problématiques qui sont très variées que ce soit en fonction de l'environnement ou de la stratégie d'entreprise. Nous avons décidé de nous concentrer sur les problèmes que les SIs seront amenés à surmonter et les évolutions majeures qu'ils vont rencontrer.

Droit de propriété intellectuelle

Table des Matières

Table des figures

1 Les technologies qui transforment le SI

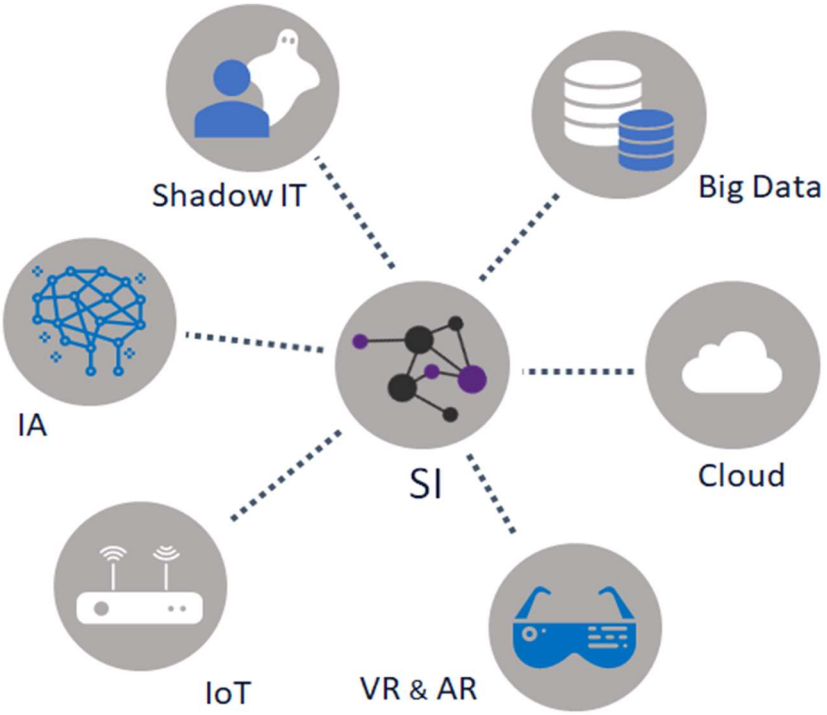

Figure 1 - Les technologies qui influenceront les SIs

1.1 Le cloud

1.1.1 Bref historique

Le terme “Cloud” ou “Nuage”, apparu dans les années 90, désigne toutes les technologies permettant de dématérialiser l’informatique et mettre à disposition des “ressources” et des “services” via les réseaux de communication à la demande. Les ressources informatiques sont ainsi dématérialisées, mutualisées et en libre accès.

Grâce au Cloud il est devenu possible de virtualiser des composants matériels et les allouer à loisir, l’informatique devient ainsi adaptable et évolutive au gré des besoins de l’entreprise.

Faisons le point des diverses possibilités utilisables par une entreprise.

Le Cloud Computing, tel que défini par le NIST (National Institute of Standards and Technology), est constitué de différentes composantes possiblement complémentaires :

SaaS (Software as a Service)

Mise à disposition d'applications d'entreprise : ERP, CRM, progiciels spécialisés, outils collaboratifs, messagerie, Business Intelligence, etc. Le fournisseur offre une fonction opérationnelle et gère de façon transparente pour l'utilisateur l'ensemble des aspects techniques requérant des compétences informatiques. *Le client a la possibilité d'effectuer des paramétrages de l'application.*

PaaS (Platform as a Service)

Mise à disposition de plates-formes de middleware, de développement, de test, d'exécution d'applications… Le fournisseur gère et contrôle l'infrastructure technique (réseau, serveurs, OS, stockage…). *Le client garde la main sur le déploiement des applications, sur leur paramétrage.*

IaaS (Infrastructure as a Service)

Mise à disposition de ressources informatiques (puissance CPU, mémoire, stockage…). Le client dispose de ressources virtualisées et déportées. *Celui-ci garde le contrôle sur le système d'exploitation (OS), le stockage, les applications déployées ainsi que sur certains composants réseau (pare-feu, par exemple).*

1.1.2 Les diverses possibilités de disposer d'un Cloud pour une entreprise

Fruit de nos éléments de veille, imaginons que votre entreprise envisage de se mettre au Cloud, voici quelques critères ou questionnements qui pourraient s'avérer pertinents.

Point d'attention	Cloud privé internalisé	Cloud privé externalisé	Cloud Public
Responsabilité de gestion du Cloud	L'entreprise elle même *Attention ampleur prévisible des enjeux stockage et sécurité*	Un fournisseur spécialisé *Enjeux grandissant autour de la mesure de la qualité du service (SLA)*	Un fournisseur si possible très pérenne *Enjeux grandissant autour de la mesure de la qualité du service (SLA)*
Localisation de la plateforme Cloud	En entreprise	En entreprise ou chez le fournisseur	Serveurs et stockage sont détenus chez le tiers et mutualisés
Gestion de la scalabilité de l'architecture	Fait par l'entreprise	Sous responsabilité du fournisseur *(Prudence contractuelle à avoir sur les critères de scalabilité et les couts associés !)*	Sous responsabilité du fournisseur Facile *car les moyens sont la plupart du temps déjà disponibles et extensibles, bien penser à faire déposer les tarifs d'extension)*
Importance de l'investissement	Conséquent	Très dépendant des services associés et de leur valeur ajoutée	Attractif, "location" de plateforme Coûts de mise en place faible puisque la principale forme de facturation se fait pour "louer" la plateforme
Mode de connexion	Via moyens sécurisés comme l'utilisation de Réseaux Privés Virtuels (ou Virtual Private Network) permettant un échange protégé entre l'utilisateur et la ressource hébergée	Via Internet ou un réseau privé	
Pourcentage d'adhésion à ce genre de solution Cloud au monde. Évolution planifiée à 3 ans (sur un parc Cloud très élargi en périmètre !)	10% environ 5% réservée à des situations d'entreprise particulières	40% environ 15% environ	60% environ 80% environ

Notez qu'il est possible en situation transitoire de faire du Cloud Hybride : assemblage de plusieurs Clouds (public et privé) amenés à coopérer, partager applications et données. Ce genre de Cloud est mis en place lorsqu'une entreprise cherche à migrer d' un Cloud privé vers un Cloud public de manière progressive. Des variantes intéressantes pourraient naître pour faciliter des entreprises en partenariat ou en approche de fusion de fonctionner ensemble...

1.1.3 Qu'est ce qui change dans le Cloud ?

1.1.3.1 Docker et le Cloud :

Dès 2013, Docker révolutionne le monde de la virtualisation en proposant une nouvelle façon de virtualiser son environnement de travail : des « conteneurs » (briques logicielles) contenant applications et librairies.

Chaque conteneur est indépendant des autres et peut fonctionner en parallèle sur le même système. Chaque conteneur est représenté dans la mémoire par un processus propre, garantissant le cloisonnement de l'environnement.

Prêtez attention au fait que contrairement à une Machine Virtuelle classique, Docker démarre en quelques secondes et prend très peu d'espace mémoire.

La piste : Avec l'arrivée de solution de clouding comme Kubernetes, il est possible de mettre en place une architecture complexe de conteneurs, dispatchés sur des serveurs (appelés des nodes) chacun contenant une instance de Docker ainsi que les différentes applications souhaitées. L'ensemble forme un "Cluster" géré par une instance de Kubernetes. Kubernetes gère l'instanciation et la suppression de conteneurs (appelés des 'Pods') garantissant ainsi le load balancing et l'optimisation de l'utilisation des ressources.

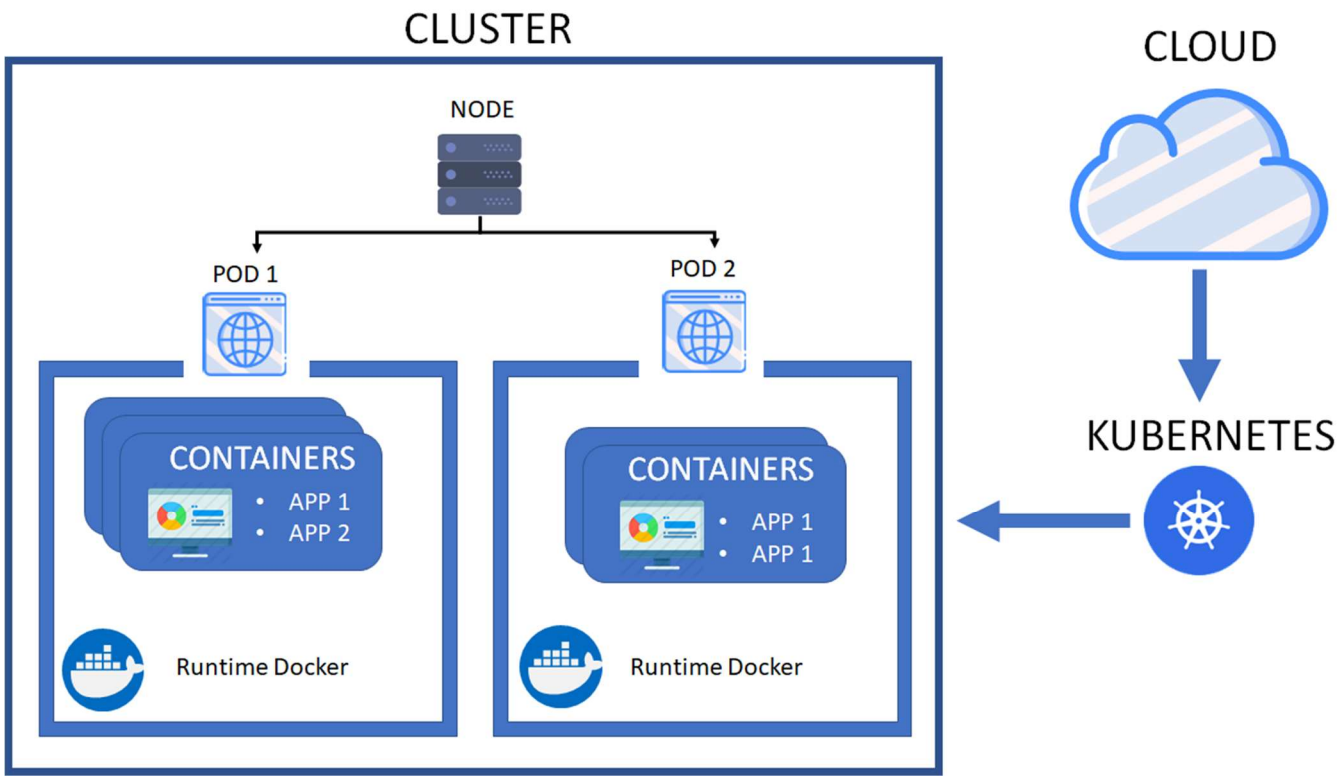

Figure 2 - Kubernetes

Ce que l'on retrouve dans une architecture Kubernetes :

- **Le Cluster** : Il englobe la totalité d'une architecture Kubernetes. Un Cluster est notamment composé d'une ou plusieurs Nodes.
- **La Node** : Elle représente un serveur physique ou une machine virtuelle dans lequel est exécuté Docker. La Node contient toutes les bibliothèques et fichiers de configuration pour le bon fonctionnement de Kubernetes.
- **Le Pod** : Il contient un ou plusieurs conteneurs dans lesquels peuvent fonctionner une ou plusieurs instances d'une application. Les Pods peuvent être créés et supprimés à loisir, offrant ainsi une flexibilité accrue. Par exemple, lors d'un pic de charge d'une application, Kubernetes peut automatiquement créer une nouvelle POD dans lequel une instance supplémentaire de l'application 1 sera créée. Une fois le pic de charge passé, Kubernetes supprimera l'instance de l'application et libérera les ressources associées.
- **Le Cloud** : L'architecture Kubernetes permet d'exposer via des Services, chacune des applications virtualisées dans leur conteneurs. Les applications peuvent même communiquer entre elles via les services.

1.1.3.2 Edge computing

L'Edge computing est une mouvance, un phénomène ponctuel, allant à l'encontre des tendances actuelles.

Il s'agit de traiter directement traiter la donnée depuis le périphérique l'ayant recueillie et non pas récupérer la donnée d'un périphérique pour ensuite la centraliser vers un serveur pour enfin la traiter.

Cette architecture a pour vocation d'alléger la charge des serveurs en précalculant les données qui leur seront envoyés. De plus ce système permet également de réduire la nécessité des IoT d'être connectés en permanence au réseau puisque les données n'ont plus besoin d'être envoyées en flux tendu.

Par exemple, une application sur un smartphone évaluant la luminosité, effectue directement le traitement sur cette donnée avant de l'envoyer dans un Datacenter. Ce phénomène a pour but de réduire le nombre de transfert de données lors du processus de centralisation et de traitement.

Solution intéressante mais à réserver dans un premier temps à des cas adaptés car il reste 2 inconvénients :

- La donnée peut être détournée après avoir été traitée car les périphériques en bout de chaîne sont par définition beaucoup plus vulnérables qu'un serveur distant protégé derrière un pare-feu.
- L'appareil effectuant le traitement peut et sera moins puissant qu'un serveur dédié à des calculs exigeants.

1.1.3.3 Serverless Cloud

Le "Cloud Serverless" ou le "nuage sans serveur" en Français, repose sur l'idée d'une architecture ne reposant plus sur des serveurs à proprement parler, en effet les serveurs existent toujours mais l'organisme mettant en place la solution serveur ne facture plus pour l'architecture mais uniquement pour le service, la fonctionnalité, la ressource proposée sur la plateforme. En d'autres termes, le modèle consiste à facturer la ligne de code exécutée lors de l'appel à la plateforme Serverless.

Ce mode de fonctionnement entraîne une réduction des coûts, une simplification de la facturation et une hausse de la flexibilité du SI.

1.2 La réalité augmentée

1.2.1 Introduction

Selon Ronald T Azuma, chercheur à l'Université de Caroline du Nord et auteur d'une des premières études sur la réalité augmentée intitulée *"A survey of Augmented reality"*, la réalité augmentée est un ensemble de technologies combinant le monde réel et des ressources numériques à la réalité, offrant à l'utilisateur des possibilités d'interaction en temps réel et reposant généralement sur un environnement 3D.

L'une des premières applications est l'invention de la "NaviCam", premier système embarqué capable de lire des marqueurs présents dans l'environnement de l'utilisateur. Le système permettait d'afficher des informations textuelles sur un HUD porté par l'utilisateur en fonction des marqueurs détectés.

D'autres Framework de réalité augmentée existent comme la librairie Vuforia couplée au moteur "Unity3D" permettant d'implémenter en 30 minutes un premier test de réalité augmentée. Pour les amateurs de JavaScript la librairie "AR.js" permet de directement disposer d'une solution de réalité augmentée sur tout appareil capable d'interpréter le javascript

1.2.2 Quel devenir ?

Dans un premier temps, la suite de l'aventure **ARKit** de Microsoft et **ARCore** de Google pour la réalité augmentée sur mobile semble être la solution à court terme la plus prometteuse. Apple est également extrêmement impliqué dans la réalité augmentée et améliore continuellement son outil. La dernière grosse mise à jour datant de septembre 2018, disponible sur iOS 12, apporte d'ailleurs plusieurs fonctionnalités comme le multijoueur, ainsi qu'une nouvelle application intégrée au nouvel OS.

La réalité augmentée progresse à grande vitesse dans tous les secteurs, que ce soit dans l'architecture, l'immobilier, la télévision et l'industrie. A titre d'exemples :

- Elle va être de plus en plus utilisée dans les usines 4.0 afin de faciliter le travail de salariés en production tout à côté des "Cobots".
- Plusieurs industriels de l'automobile et des fabricants de simulateurs s'intègrent pour augmenter les possibilités pédagogiques et la précision d'usage.

1.3 Le Shadow IT

1.3.1 Introduction

Le Shadow IT est un phénomène faisant référence à l'utilisation de logiciels ou de matériel non validés par la Direction Informatique d'une entreprise. Le simple fait d'utiliser un éditeur de texte non validé en amont peut s'avérer être du Shadow IT. Le Shadow IT se traduit comme étant un grand facteur de problème de sécurité dans les entreprises actuelles, selon une étude menée par Gartner.

Le premier Shadow IT était dû à un stagiaire ou un employé impatient qui voulait utiliser une ressource ou installer une application sans attendre le feu vert de sa direction Informatique. Avec l'arrivée du Cloud Computing et la globalisation de l'Informatique, le terme englobe désormais l'utilisation de ressources web et autres manières d'amener ou de retirer de l'information depuis et vers l'extérieur. Par exemple, un employé hébergeant des informations relatives à son travail sur un Dropbox personnel peut être considéré comme du Shadow IT et donc une faille de sécurité pouvant mener à une fuite d'information.

En 2018, 38% des cyber-attaque à l'encontre d'une entreprise ont eu lieu via des failles de sécurité liées au Shadow IT, cette statistique dépassera les 50% d'ici 2020.

1.3.2 Comment endiguer le Shadow IT à très court terme ?

1.3.2.1 Détecter le Shadow IT

Il appartient à la DSI de faire l'inventaire de qui utilise quoi dans l'entreprise. Cet inventaire va permettre d'identifier les risques et de proposer des solutions plus adaptées. Pour y parvenir, il est possible de monitorer le réseau à l'aide de « sniffers » et autres outils de scan de sécurité. Même si ces scans ne résolvent pas les failles de sécurité, ils font apparaître les flux des applications inconnues ou nouvelles, lesquels constitueront une base de connaissances pour que la DSI puisse déterminer les meilleures alternatives possibles.

1.3.2.2 Besoins du terrain et propositions de solutions maîtrisées

Être ouvert, conscient des besoins du terrain et proposer des solutions maîtrisées par la DSI pour prévenir l'usage ou remplacer l'utilisation du Shadow IT

La plupart des shadow IT révèlent un besoin utilisateur mal pourvu ou non pourvu par la DSI envers les utilisateurs métier. Il s'agit souvent là du point de départ du shadow IT : face à la pression à l'efficacité, une attente non traitée, il y a recherche d'une solution pour arriver à générer de la valeur en entreprise ou éviter d'en fuiter. L'utilisateur est donc indirectement poussé au Shadow IT pour s'acquitter de ses tâches.

A court terme prévoir pour éviter de revoir va consister à projeter le besoin des utilisateurs et apporter le nécessaire au bon moment.

Prenons par exemple les télétravailleurs. Si l'entreprise ne leur propose pas spontanément les moyens pour collaborer à distance, ils en trouveront eux-mêmes, sans même se demander s'il était pertinent d'en parler à la DSI. Et c'est à partir de ce moment-là que les complications commencent. À ce titre, il est important de créer le dialogue en écoutant leurs retours d'expérience et les problèmes qu'ils essaient de résoudre.

2 Big Data

2.1 Introduction

Depuis plusieurs années le Big Data fait parler de lui. Considéré comme une mode pour certains, il est vu comme le nouvel or noir pour d'autre. En effet il représente une mine d'informations inépuisable qui va permettre à certaines entreprises de prospérer, grâce à l'utilisation de ces milliards de Gigaoctet de données. Mais toutes ces données n'ont pas une très grande valeur dans leur état brut, il faut donc mettre en place des projets coûteux et parfois longs pour pouvoir valoriser ces données et peut-être en extraire la pépite qui vous rendra riche.

Malheureusement d'après une étude de Matt Asay, effectuée en 2017, 85 % des projets Big Data sont un échec, et 12% des entreprises qui utilisent le Big Data en tirent bénéfice. Les raisons de ces échecs sont rarement en lien avec la technologie. Nous allons donc découvrir ensemble comment maximiser les chances de réussite d'un projet Big Data.

2.2 Définition

Commençons par répondre à une question qui semble plutôt simple Qu'est-ce que le Big Data ?

Le Big Data fait référence à l'explosion du volume de données informatiques qui sont collectées dans le monde.

"Des ensembles de données devenus si volumineux qu'ils dépassent l'intuition et les capacités humaines d'analyse et même celles des outils informatiques classiques de gestion de base de données ou de l'information."

Citation par Stéphan Clémençon, professeur et titulaire de la Chaire Machine-Learning for Big Data à Télécom ParisTech

En effet 90% des informations numériques que nous avons engrangées depuis le commencement de l'humanité a été produite ces 2 dernières années.

2.3 Les raisons d'être du Big Data.

Quand une entreprise décide de s'attaquer à la question du Big Data, c'est généralement pour l'une des raisons suivantes :

- Elle souhaite améliorer ses analyses, mais nécessite une quantité de données plus importante.
- Elle se rend compte qu'elle peut réaliser des analyses en temps réel dans le but de fournir un service d'aide à ses clients pour une utilisation plus efficace.
- Elle veut pouvoir réagir plus vite et de façons efficaces à moindre coût.
- Elle peut vouloir augmenter l'apport de sa solution BI en y intégrant des données provenant de l'extérieur.

2.4 Quels types de données peut-on y trouver ?

Une des raisons de l'essor de la Big Data vient de l'utilisation des objets connectés qui ne fait que croître ces dernières années, comme vous pouvez le constater dans l'histogramme ci-dessous, il y a environ 23 Milliards d'objets connectés présent dans le monde et ce chiffre pourrait tripler d'ici 2025.

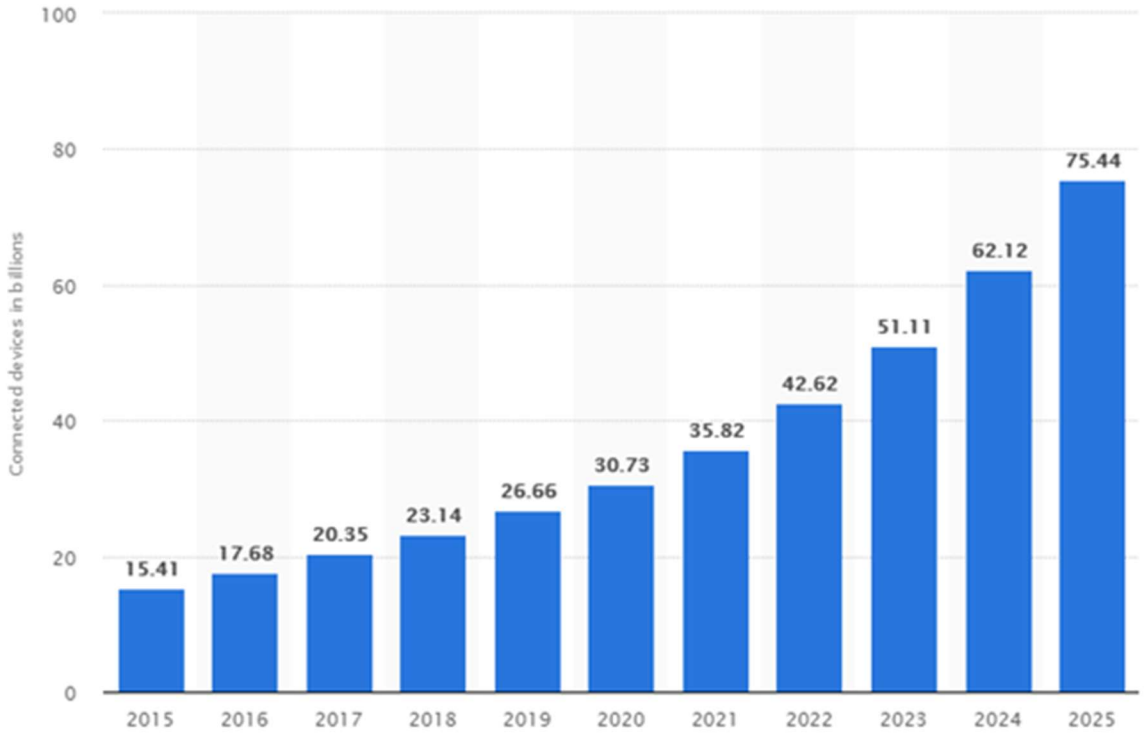

Figure 3 - Nombre d'objets connectés dans le monde

Une autre raison de l'essor du Big Data est l'augmentation des capacités de stockage pour des volumes physiques qui restent constants, voire qui diminuent. Les précédents facteurs, couplés avec l'augmentation du trafic sur internet, entraînent une augmentation logique de création et stockage de la donnée. Pour justifier nos propos, sachez que nous créons et sauvegardons chaque jour plus de 2 500 000 000 de gigaoctet de données.

Figure 4 - Création de grand volume de données

Ces données proviennent en grande partie des utilisateurs de produits numériques, comme le montre l'infographie précédente. Les données que les utilisateurs génèrent grâce aux réseaux sociaux, avec par exemple 293000 statuts publiés sur Facebook, aux plateformes de blogs et aux forums, ainsi que les données relatives à l'utilisation du smartphone, de la tablette ou de l'ordinateur. Une grande partie des données provient aussi des achats effectués, sur internet ou en magasin. Enfin, une dernière partie des données est produite par les objets connectés.

2.5 Secteurs qui utilisent le Big Data

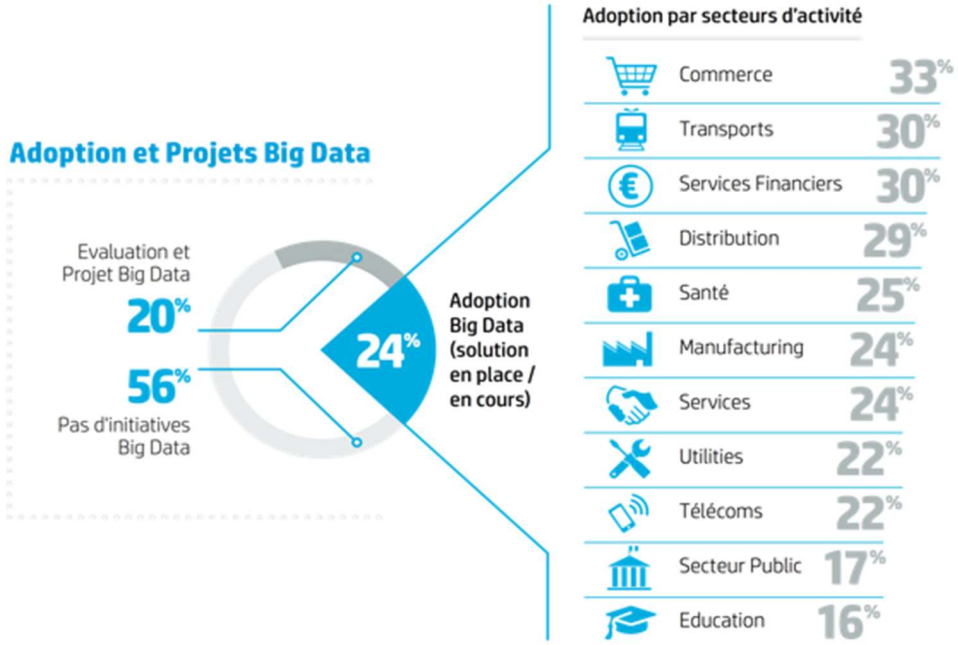

Figure 5 - Les initiatives Big Data

Beaucoup d'entreprises prennent le virage du Big Data. Elles l'utilisent pour améliorer leur productivité ou aider leurs clients. Comme le montre le schéma précédent, les secteurs qui utilisent le plus le Big Data sont le commerce - pour suggérer des produits par exemple -, les transports - pour prédire le trafic ou optimiser le trajet -, et les services financiers - pour détecter la fraude.

Bien des domaines gagneraient à utiliser le Big Data dans leurs activités, comme par exemple le secteur de la santé, afin de mettre en corrélation les causes et les symptômes d'une pathologie, ou encore la distribution pour les services basés sur la localisation, ou encore l'industrie avec la maintenance prédictive.

2.6 Les Projets Big Data

Comme dit précédemment de nombreuses entreprises n'arrivent pas à mener à bien leur projet Big Data. Les raisons peuvent être diverses mais comme pour tous les projets, la partie la plus cruciale reste sa définition. Pour bien définir un projet il faut commencer par définir des critères qui vont définir votre projet Big Data.

2.6.1 Les critères d'un projet Big Data : les 5V

Volume

Le Volume est la caractéristique la plus importante d'un projet Big Data. C'est ce Volume qui détermine la quantité de données que votre projet va pouvoir recevoir, il va donc définir votre architecture. En effet si la quantité de donnée est largement supérieur aux prévisions, vous risquez d'avoir des problèmes avec votre architecture, mais inversement si la quantité n'est pas suffisante vous prenez le risque d'avoir des résultats moins précis.

Vélocité

La Vélocité peut être simplement définie comme la vitesse d'actualisation des données. En d'autres termes la Vélocité définit à quelle vitesse les nouvelles données doivent être accessibles. Dans certain cas, les données doivent être mise à jour en temps réel pour un suivi direct, mais elles peuvent aussi être mise à jour en bloc tous les jours. Si ce critère est mal défini il peut mettre en péril votre projet car il est possible que les données montrées deviennent moins pertinentes, ou bien encore demander des ressources supplémentaires dont vous auriez pu vous passer.

Variété

La variété peut être simplement définie comme le fait d'avoir différentes sources de données. Mais ces différentes sources de données peuvent prendre plusieurs formes. Par exemple des fichiers vidéo, des messages textuels, des tableurs ou bien encore des informations provenant d'autres bases de données. Toutes ces données peuvent être potentiellement intéressantes mais ne seront pas forcements structurés. Elles devront donc subir un traitement en amont permettant d'extraire les données pertinentes.

Véracité

La véracité fait référence à la fiabilité des données utilisées. Votre projet Big Data va utiliser de très nombreuses données provenant de sources diverses, il faut donc être certain que les données qui vont être exploitées sont exactes et ne vont pas compromettre la fiabilité des analyses futur. Par exemple les données météorologiques doivent provenir d'un laboratoire et non d'une source non vérifiée telle qu'un amateur qui recueillerait ces données par passion avec des moyens limités.

Valeur

La Valeur d'un projet big data représente sa raison d'être, quels problèmes ce projet va permettre de résoudre, ce que cela peut apporter comme bonnes idées, facilités, idées de services, éléments de benchmarking....

Effectivement, il faut spécifier les caractéristiques des problèmes auxquels doit répondre votre projet Big Data et tout faire pour orienter les données vers la résolution de ce problème. En revanche, il ne faut pas oublier que votre projet n'a pas pour but de valider une idée préconçue ou une théorie que vous auriez dès le début de votre projet, il faut rester objectif dans l'utilisation des données pour découvrir de nouvelles possibilités, qui évolueront sûrement au cours du temps.

Le choix de l'architecture d'un projet Big Data est une étape importante. On pense à trois solutions principales en matière de big data : Base de Données Relationnelle Classique (MySQL, SQL Server) versus NoSQL (HBase, MongoDB, Cassandra) versus Cloud (public, hybride, privée).

Les bases de données NoSQL conviennent davantage aux grands ensembles de données fréquemment exposés à de nouvelles informations, au sein desquels les enregistrements ont des structures variables.

Les bases de données, présentes sur le Cloud, ont pour avantage d'éliminer les infrastructures physiques (location), de réduire les coûts, d'avoir une scalabilité instantanée et grâce aux fournisseurs, de profiter des dernières technologies en date.

2.6.2 Méthodes de traitements de données

Une autre erreur souvent faite par les entreprises, est de considérer les projets Big Data comme un projet lié à une base de données classique et de ne pas faire appel à du personnel qualifié. De plus, ce personnel est souvent coûteux mais il est capable de manipuler les données pour mettre en évidence les informations que recherche l'entreprise, pour cela il va utiliser plusieurs méthodes.

2.6.2.1 Data Cleansing

Le Data Cleansing consiste à supprimer de la base de données toutes les données potentiellement incorrectes, incomplètes, mal formatées ou dupliquées.

Les erreurs dans les données coûtent aux entreprises l'équivalent de 10 à 20 % de leur budget d'implémentation. De plus, on estime que 40 à 50 % du budget temps d'un projet est dépensé dans la correction d'erreurs dans les données.

Les différents types d'erreurs sont :

- Erreurs de syntaxe (lexicales, formatage, irrégularité).
- Erreurs sémantiques (Violation des contraintes d'intégrité, contradiction, duplication, donnée invalide).
- Erreurs de couverture (Valeur manquante, Donnée manquante).

On distingue deux types d'approches :

- Celle des entreprises, qui utilisent généralement des nettoyeurs de type ETL. Ce type d'approche a été le premier à avoir vu le jour.
- Celle des scientifiques, qui explorent de nouvelles approches qui se basent sur les contraintes d'intégrité, les statistiques, l'apprentissage automatique ou encore le crowdsourcing.

Le nettoyage de données se décompose en 3 phases :

1. Analyser les données afin de détecter les potentiels problèmes.
2. Choisir le type de transformations à effectuer.
3. Appliquer ces informations aux données.

Méthodes existantes :

- Parsing (détection d'erreurs de syntaxe)
- Transformation de donnée (regroupe plusieurs champs en un seul)
- Renforcement des contraintes d'intégrité (trouver et modifier les contraintes d'intégrité douteuses afin qu'elles s'accordent mieux avec les données)
- Méthode statistique (pour l'analyse des données et/ou leur correction)
- Crowdsourcing (multiplier les sources)

2.6.2.2 Data Analytics

L'analyse des données est le processus d'examen des ensembles de données afin de tirer des conclusions sur l'information qu'ils contiennent, de plus en plus à l'aide de systèmes et de logiciels spécialisés.

- L'Exploratory Data Analysis (analyse de donnée exploratoire), abrégée par EDA, permet de découvrir de nouveaux éléments dans les données.
- La Confirmatory Data Analysis (analyse de données confirmatoire), abrégée par CDA, permet de prouver si des hypothèses existantes sont vraies ou fausses.
- La Qualitative Data Analysis (analyse de données qualitative), abrégée par QDA, est utilisée dans les sciences sociales pour tirer des conclusions de données non numériques telles que les mots, les photographies ou la vidéo.
- Real-Time Analytics est le traitement de petits ensembles de données qui nécessitent d'être traités et analysés en temps réel dès qu'ils sont générés.

Une autre étape importante d'un projet Big Data, consiste à définir de quelle manière l'entreprise va utiliser cette mine d'informations.

L'analyse descriptive

Elle permet d'examiner les données et l'information pour définir l'état actuel d'une entreprise, de telle sorte que les développements, les tendances et les exceptions deviennent évidents, sous la forme de rapports standard et d'alerte. On peut classer les individus dans des catégories, trouver les individus les plus proches ou les plus éloignés entre eux ; mais aussi trouver les exceptions ou les cas atypiques. On peut également voir si des variables sont proches, expliquer une variable en fonction des autres ou encore repérer les variables les plus influentes.

L'analyse investigatrice

Elle consiste à sonder les données pour comprendre pourquoi un événement s'est produit dans une entreprise, grâce à des analyses statistiques des données et des analyses factorielles.

Analyse prédictive, qui prédit les tendances et les modèles de comportement

Elle consiste à analyser les données actuelles afin de faire des hypothèses sur des comportements futurs. On se sert des données que l'on possède déjà pour extrapoler et anticiper le comportement de nouveaux individus mais également l'évolution des individus déjà présents.

Analyse prescriptive, qui se concentre sur la meilleure action pour une situation donnée

Elle permet de croiser les informations et l'action. Il ne s'agit plus de livrer uniquement des indicateurs mais de se focaliser sur une proposition de décision. L'analyse prescriptive n'est cependant pas infaillible. Elle est sujette aux mêmes distorsions que les analyses descriptives et prédictives, notamment les limites des données et les forces impondérables extérieures. Elle doit sa viabilité aux avancées en matière de vitesse de traitement et au développement d'algorithmes mathématiques complexes appliqués aux ensembles de données.

Analyse normative :

Elle vise l'optimisation et les tests randomisés pour évaluer comment les entreprises améliorent leurs niveaux de service tout en réduisant leurs dépenses.

Analyse préventive :

Elle consiste à avoir la capacité de prendre des mesures de précaution à l'égard d'événements susceptibles d'influer de façon indésirable sur le rendement organisationnel, par exemple en déterminant les dangers possibles et en recommandant des stratégies d'atténuation à long terme.

2.6.3 Cycle de vie des données

La gestion du cycle de vie des données traite la question du stockage des données en fonction de leur caractère critique. L'automatisation est importante à cause du gros volume de données.

La première étape est d'identifier - quel type, quelle date, quel émetteur - et caractériser les données - et éventuellement mettre un tag sur ces données. La seconde étape est de créer des règles destinées à préciser l'évolution de leur valeur. Si la criticité est importante, les données pourront être sur une base donnée haute performance, et si la criticité est moindre, les données pourront être migrées sur une base de données plus lente.

2.6.4 DO & DON'T

DO	DON'T
• Contrôlez les changements (matériels, architecture, jeux de données, algorithmes, et leurs impacts positifs/négatifs). • Préférez la méthode Agile et ses itérations modulables. • Faites attention à la qualité du code (première cause de mauvaise performance des algorithmes). • Pensez à une stratégie de gestion du cycle de vie des données.	• Ne misez pas tout sur un aspect (seulement les serveurs, seulement la base de données, seulement les algorithmes, seulement votre/vos data scientists, …). • N'omettez pas la sécurité de vos données. • Ne négligez pas la composition de votre équipe. • Ne considérez pas le Big Data comme un projet basique de base de données.

2.7 Le Big Data est-il objectif ?

Le Big Data s'appuie sur des données passées, et la prévision et prédiction ne peuvent pas innover en extrapolant sur ces données. Toute nouveauté est par conséquent imprévisible et elle ne pourrait se déduire que par le déclin de l'existant, si cela se produit. Le big data ne peut pas prédire le hasard.

Les données sont partielles :

Seules sont enregistrées les données à portée du capteur. Si une personne n'est pas sur les réseaux sociaux, ou utilise peu internet, ou fait ses achats au commerçant en liquide, c'est toute une partie de ses actions qui échappent à la collecte.

Les données ne sont pas éthiques ou morales :

Dans le cas où il y aurait une quelconque forme de partialité ou discrimination des informations, le tri ou le résultat de l'algorithme sera en conséquence biaisé. Les algorithmes ne représentent pas la vérité, mais le point de vue de celui ou celle qui l'écrit.

2.8 Conclusion

Que le Big Data soit une mode ou une pratique qui s'installe, il est important de se positionner. La décision ou non de se lancer dans un projet doit être mûrement réfléchie, pour ne pas se lancer dans un projet coûteux qui a beaucoup de chance d'échouer. Chaque étape doit être pensée, rien ne doit être laissé à plus tard ou au hasard.

Le Big Data peut aussi servir à alimenter la Business Intelligence de l'entreprise, dans la mesure où il peut aider à comparer ses résultats avec ceux des concurrents. Il peut également aider à découvrir des problèmes et aider à optimiser des processus existants.

3 IOT

3.1 Introduction

Depuis plusieurs années, on assiste à une explosion du nombre d'objets connectés dans la vie de tous les jours. Ce changement touche également les entreprises et les industries (Industrial Internet of Things) ce qui entraîne une évolution du mode de gestion de ces dernières. Selon *Jupiner Research*, le nombre d'objets connectés devrait tripler d'ici 2020, jusqu'à atteindre 38 milliards d'objets. De quoi donner le vertige...

Comme tout changement, cette évolution s'accompagne de nombreux défis. Afin d'intégrer sereinement des objets connectés dans un SI d'entreprise, il convient de s'intéresser en premier lieu aux risques sécuritaires. De ces risques découlent des conséquences pouvant être dévastatrices si non anticipées. C'est pourquoi nous vous présenterons diverses solutions et bonnes pratiques afin de répondre à ces nouvelles problématiques.

3.2 Risques et conséquences

Aujourd'hui, les constructeurs d'IoT fabriquent à grande échelle une multitude d'objets différents qui peuvent s'implémenter dans des SI de toutes catégories. Le principal risque concerne les politiques de sécurité car comment établir des standards et des règles pour protéger les entreprises si l'hétérogénéité des objets connectés n'est pas régulée ? En effet, si les constructeurs ne respectent pas des normes techniques et technologiques similaires, il est difficile voire impossible de se protéger correctement avec un ensemble d'objets connecté qui disposent d'un ensemble hardware et software hétéroclite.

Un autre risque soulevé par cette diversité concerne les systèmes d'exploitation. En effet, plus il y a d'OS différents, plus il y a de failles de sécurité potentielles. C'est pourquoi, les entreprises peuvent être confrontées à des enjeux de sécurité importants pouvant mettre en péril leur activité. Une fois un système intégré au sein de l'entreprise, son objectif premier est d'être fonctionnel et d'assurer un certain nombre de fonctionnalités. Cependant, il est nécessaire que les objets

connectés disposent d'un protocole de mises à jour sans quoi il pourrait perdre en efficacité et devenir une menace pour le SI et l'activité de l'entreprise.

Dans ce monde ultra-connecté, la sensibilisation et la formation des utilisateurs d'objets connectés est primordiale. Le risque de phishing ou de mauvaise manipulation peut endommager sérieusement la structure informatique d'une entreprise ou être un frein à son activité en nuisant notamment à son image ou en permettant la fuite d'informations.

Les conséquences liées aux risques de l'IoT peuvent être sévères pour une entreprise et ses partenaires. Les plus importantes sont le vol de données qui peut concerner à la fois les particuliers ou les professionnelles, qui est apparenté à de l'espionnage industriel. D'autre part, les cas les plus extrêmes peuvent entraîner tout simplement l'arrêt de l'activité de l'entreprise voire son dépôt de bilan.

Un autre problème sont les botnet qui peuvent être utilisés pour les attaques de déni de service. Il est important de comprendre que de tels réseaux peuvent être composés de n'importe quel objet connecté disposant d'une adresse IP, tel qu'un réfrigérateur ou une machine à café.

Les problèmes IoT peuvent également impliquer les entreprises sur le plan juridique. En effet, si une entreprise possède des objets connectés, elle en est responsable et doit rendre des comptes à la Justice au cas où des actions malveillantes ont été commises via les objets.

3.3 Sécuriser les équipements

3.3.1 Gestion de projet

La gestion de projet a une importance cruciale dans la sécurité de l'IoT. Imaginons une entreprise qui travaille sur un projet de réfrigérateur connecté, elle favorisera certainement l'ergonomie et l'expérience utilisateur à la sécurité. Son objectif est de vendre son produit est d'attirer le consommateur en se distinguant notamment sur des aspects visibles et pratiques. En effet, l'aspect sécurité IT pour un réfrigérateur semble anodin. Pourtant, c'est un objet connecté comme un autre et il peut donc servir à des pirates informatiques lors d'attaques massive de DDOS par exemple. Or la conception et l'intégration des points de sécurité de

base dans un projet d'objets connectés nécessite d'être abordées au début de la conception du produit sans quoi l'entreprise perdra du temps et de l'argent à essayer d'ajouter au mieux des éléments essentiels qu'il aurait pu être ajouté à moindre coût.

La sécurité du système doit être intégrée dès la phase de définition des besoins. Sans quoi, l'entreprise prend le risque de devoir appliquer des correctifs plus tard.

Il ne faut pas considérer que la sécurité n'est qu'un coût sans intérêt pour le produit. Si une faille est découverte et exploitée, l'entreprise devra peut-être payer des dédommagements, remboursements, ou bien rappeler des produits. De plus, l'image et la réputation de l'entreprise va être dégradée et les clients n'auront plus confiance et à terme l'entreprise perdra encore plus d'argent que si la sécurité avait été intégrée dès le début de la conception.

3.3.2 Commodités de sécurité basique

Certains dispositifs d'IoT peuvent fonctionner en continu sans surveillance et ne nécessite pas d'être surveillés fréquemment par des humains. Il faut garder ces objets relativement isolés de façon que seules quelques personnes autorisées aient un accès physique. Cette mesure de sécurité aide à empêcher les intrus potentiels d'accéder aux données du device.

La sécurité physique des terminaux peut inclure, par exemple, des ports sécurisés physiquement et des couvercles qui couvrent les webcams. Les verrouillages de ports aident à empêcher l'ajout de logiciels malveillants direct depuis des clés USB. Certaines approches vont jusqu'à désactiver l'appareil lorsqu'il est manipulé.

Imaginons un détecteur de fumée disposant d'un port opérationnel. Un pirate pourrait accéder au logiciel embarqué et l'extraire. Ainsi il pourrait soit modifier le logiciel ou en injecter un autre ou bien étudier le logiciel récupéré et s'épargner des phases de R&D onéreuses et récupérer des informations fonctionnelles à moindre coûts.

C'est pourquoi, les objets connectés doivent embarquer des systèmes de protection qui empêchent la récupération des logiciels une fois l'objet sur le marché.

Des procédures de démarrage telles que des mots de passe forts ou le fait d'exiger que le périphérique démarre uniquement à partir de son stockage local peuvent être des approches judicieuses. Il y a énormément de moyen d'attaquer des IoT, c'est pourquoi les vulnérabilités de bases doivent être traitées (les ports TCP/UDP ouverts, les ports physiques, les mots de passe par défaut, les injections de code, les communications non chiffrées, les connexions radio).

Un dernier point important se concentre lors de la phase d'expédition. En effet, l'emballage doit être inviolable afin de permettre au client de savoir si l'appareil a été ouvert avant son arrivée.

3.3.3 Vulnérabilités, mises à jour et obsolescence

Inévitablement, les vulnérabilités seront découvertes après le déploiement des dispositifs. Les dispositifs doivent pouvoir être patchés ou mis à niveau. Naturellement, le micrologiciel de l'appareil ne devrait être modifiable qu'avec une signature numérique appropriée. Dans l'état actuel des choses, les vendeurs et les fabricants d'appareils n'ont guère d'incitation financière à assurer la mise à niveau continue des correctifs IoT, puisque les recettes proviennent de la vente de l'appareil et non de la maintenance. L'entretien des appareils IoT peut nuire aux recettes.

Imaginons des ampoules connectées qui présenteraient une faille pouvant empêcher l'utilisateur de les allumer. Une fois la faille connue, les futures ampoules pourront être corrigées, cependant si les ampoules déjà vendues ne disposent pas d'un système de mises à jour alors elles devront continuer à fonctionner avec une vulnérabilité rendue publique.

Ceci souligne qu'une procédure permettant la mise à jour sécurisée de l'objet est obligatoire sans quoi le consommateur n'en achètera plus ou l'objet connecté pourrait nuire à la sécurité d'individu. La procédure doit pouvoir établir une communication authentifiée et chiffrée, afin que l'objet connecté obtienne des mises à jour officielles.

De plus, les fournisseurs ne sont pas tenus légalement responsables de l'entretien continu des dispositifs au-delà des ventes initiales et la concurrence pousse les fournisseurs à faire des économies, ce qui nuit à la qualité pour l'efficacité et la rapidité du lancement sur le marché. Bien que ces facteurs n'aient peut-être pas été critiques avant l'IoT, la nature interconnectée de l'interdépendance de l'IoT des dispositifs d'IoT élève la barre à un nouveau niveau en termes de fonctionnalité et de responsabilité.

La tendance des fournisseurs à l'obsolescence planifiée des appareils est également préjudiciable afin de maximiser les profits par la poursuite des ventes plutôt que par l'entretien des appareils existants. En outre, les dispositifs d'IoT ne sont pas efficacement conçus ou configurés pour répondre aux mises à jour, ce qui entraîne, au mieux, des procédures coûteuses et, au pire, ingérables. En l'état actuel des choses, de nombreux dispositifs d'IoT sont impraticables et, en tant que tels, ne peuvent être sécurisés.

Outre l'obsolescence planifiée, de nombreux dispositifs d'IoT ont simplement des cycles de vie limités. Les entreprises doivent être tenues légalement responsables de la surveillance et de l'entretien des dispositifs au cours des cycles de vie prescrits et convenus. Pour ce faire, des normes doivent être établies et une législation doit être mise en place. En outre, les fournisseurs doivent rester transparents et ouverts sur le cycle de vie des appareils, en particulier en termes de politiques de service et d'entretien, y compris la durée pendant laquelle ils prévoient de prendre en charge leurs appareils. Ils doivent jouer un rôle actif en fournissant des détails sur les correctifs et les mises à niveau ainsi que sur les risques de sécurité et les préoccupations en matière de protection de la vie privée, en s'assurant que le consommateur est informé des changements de politique, de fonctionnalité et de sécurité.

Un dernier problème apparaît lorsque le fournisseur d'origine n'existe plus, il devient compliqué voire impossible de corriger des vulnérabilités et des failles de sécurité car les équipes pour développer et déployer les mises à jour n’existent plus.

3.3.4 Fabricants : essais dynamiques

Il est essentiel que les dispositifs IoT soient soumis à des tests pour détecter des failles de sécurité. Les tests statiques ne sont pas conçus pour détecter les vulnérabilités qui existent dans les composants standards tels que les processeurs et la mémoire. En revanche, les tests dynamiques sont capables d'exposer à la fois les faiblesses du code et les vulnérabilités du matériel. Les tests dynamiques peuvent découvrir des vulnérabilités qui sont créées lorsque du nouveau code est utilisé sur d'anciens processeurs. Les fabricants qui achètent du matériel et des logiciels d'autres fabricants doivent effectuer des essais dynamiques pour s'assurer que les articles sont sécurisés.

3.3.5 Fin de vie des objets connectés

Les appareils finissent par devenir obsolètes et les utilisateurs peuvent décider de les jeter. Les dispositifs devraient être jetés sans exposer les données privées. Il s'agit d'un problème de sécurité, car des dispositifs jetés de manière incorrecte peuvent être utilisés à des fins malveillantes.

Il s'agit d'une question de respect de la vie privée car le matériel peut contenir des informations personnelles sur l'utilisateur ou d'autres parties prenantes.

C’est pourquoi, il faut se débarrasser de ses produits directement auprès du fabricant. Lors de l'achat d'un produit IoT usagé, les anciennes informations d'identification personnelle telles que le nom d'utilisateur et le mot de passe peuvent être stockées sur l'appareil et sont donc accessibles.

Aujourd’hui, les utilisateurs sont mal accompagnés et équipés pour comprendre le système de stockage des mots de passe dans les objets connectés. Par exemple, certains photocopieurs possèdent des disques durs qui conservent les copies des documents ce qui représente une faille de sécurité conséquente pour l’entreprise.

3.4 Sécurisation des réseaux

3.4.1 Utiliser l'authentification forte

Les périphériques IoT ne doivent pas utiliser de nom d'utilisateur/mot de passe faciles à deviner, tels que admin/admin. Les périphériques ne doivent pas utiliser les informations d'identification par défaut et ne doivent pas inclure de back-door ainsi que de paramètres de débogage (identifiant secret du constructeur) car, une fois connues, elles peuvent être utilisées pour pirater plusieurs périphériques.

Chaque appareil doit avoir un nom d'utilisateur/mot de passe par défaut qui doit être réinitialisable par l'utilisateur et suffisamment sophistiqués pour résister aux méthodes de brute-force. De plus, lors de la première connexion, il faut imposer à l'utilisateur de changer les identifiants de connexion immédiatement. Pour cette réinitialisation, il est préférable d'utiliser un minimum de huit caractères de longueur avec des lettres minuscules, lettres majuscules, chiffres et caractères spéciaux.

Il faut privilégier l'authentification à 2 niveaux, c'est-à-dire qui exige qu'un utilisateur utilise à la fois un mot de passe et un autre formulaire d'authentification, comme un code aléatoire généré par SMS. Quant aux applications IoT, l'utilisation de l'authentification contextuelle est à privilégier. C'est une authentification adaptative, des algorithmes de machine learning évalue en permanence le risque sécuritaire. Si le risque est élevé, l'utilisateur se verrait demander un token pour continuer à utiliser l'application.

3.4.2 Chiffrement, encodage et protocoles sécurisés

Même si les mots de passe des périphériques sont sécurisés, les communications entre les périphériques peuvent être piratées. Dans l'IoT il y a beaucoup de protocoles. Selon le protocole et les ressources informatiques disponibles, tous les objets connectés ne sont pas capables d'utiliser un chiffrement fort. L'objectif est d'utiliser le chiffrement le plus fort possible, tel que TLS/SSL.

Selon le contexte le chiffrement n'est pas toujours souhaitable. En effet, les voitures autonomes peuvent utiliser des messages de sécurité pour éviter les collisions. Dans ce cas, les messages peuvent être envoyés en clair et vérifiés à l'aide de signatures numériques. Toutefois, il faut tenir compte des conséquences

de l'omission du chiffrement. Il n'y a pas de solution générique pour faire face à des menaces et des vulnérabilités identifiées. Si les données sont transmises non chiffrées et non signées, les précautions à prendre doivent être appliquées pour s'assurer que les fausses données ont peu ou pas de chance de causer du tort.

L'encodage n'apporte aucune sécurité. Ces algorithmes ne servent qu'à optimiser ou adapter un flux de données à un support de transport.

De plus, il faut être vigilant quant à la fausse impression de sécurité apportée par l'utilisation d'un algorithme faible tel que DES, SHA1, car ils ne sont plus fiables.

Un autre risque est l'utilisation d'un algorithme conçu en interne. Les standards de chiffrement ont été mathématiquement éprouvés. Cela est très rarement le cas pour les algorithmes « maison » qui peuvent embarquer des faiblesses ou des vulnérabilités. La création d'algorithmes de chiffrement est une activité scientifique spécialisée.

Différents objets connectés ne doivent pas partager une même clé de chiffrement. Une règle très importante dans le milieu des systèmes IoT est qu'un produit compromis ne doit pas permettre de les compromettre tous.

Une clé de chiffrement ne doit pas servir à sécuriser les communications avec un objet pendant toute sa durée de vie. Les clés doivent être changées régulièrement.

3.4.3 Minimiser la bande passante de l'appareil

Récemment, des attaques DDoS ont été menées dans une large mesure par des armées de dispositifs d'IoT mal protégés qui sont devenus des systèmes zombies (botnet) lors de campagnes mondiales massives. La plupart des dispositifs d'IoT sont constitués de composants de base qui ont des capacités de réseau surpuissantes pour la fonction qu'ils sont censés exécuter, ce qui provoque une congestion sur les réseaux domestiques et peut contribuer à des coûts énormes pour les cibles subissant des attaques DDoS transmises par IoT.

Il est fortement recommandé de réduire le trafic malveillant produit par ces systèmes en limitant le trafic réseau que les appareils d'IoT peuvent générer à des niveaux raisonnables par rapport à leurs fonctions. Une utilisation restreinte de bande passante au niveau du matériel permet de limiter les débits de

transmission du réseau à des niveaux raisonnables. Ces limitations rendent beaucoup plus difficile une attaque DDoS.

Une autre contre-mesure aux attaques DDoS concerne la surveillance des objets connecté entre eux. En effet, ils doivent être programmés pour s'auto-surveiller et si des comportements inhabituels sont détectés, soit le périphérique se réinitialise aux réglages d'usine, soit il doit redémarrer pour effacer le code malveillant.

3.4.4 Diviser les réseaux en segments

Une autre bonne pratique est de séparer le réseau en réseaux locaux plus petits en utilisant des VLAN, des plages d'adresses IP ou une combinaison des deux. Les segmentations de réseau sont utilisées dans les politiques de sécurité de pare-feu de nouvelle génération pour identifier clairement une ou plusieurs interfaces source et destination sur la plate-forme. Chaque interface du pare-feu doit être affectée à une zone de sécurité avant de pouvoir traiter le trafic. Cela permet aux organisations de créer des zones de sécurité pour représenter les différents segments connectés au pare-feu et contrôlés par celui-ci et permet d'élaborer des politiques de sécurité qui contrôlent l'accès à des applications en fonction des groupes utilisateurs ce qui empêche ainsi tout accès interne ou externe non autorisé aux données enregistrées dans ce segment.

Ce type de solution est plus courant dans les applications industrielles, mais peut être utilisé pour un réseau domotique.

3.4.5 Authentification des serveurs

Un système IoT peut utiliser le protocole TLS pour sécuriser les serveurs et le micrologiciel. Cependant lors de l'initiation des communications, le certificat envoyé par le serveur n'est pas toujours authentifié c'est pourquoi si l'objet connecté et un pirate sont connectés sur le même réseau le pirate pourrait lire et modifier toutes les informations échangées. Il est important de s'assurer que le dispositif IoT utilise des certificats officiels tel que le x.509 afin d'éviter ce genre de faille de sécurité. De même qu'il faut proscrire l'utilisation de certificats auto-signés. L'idéal est de favoriser l'usage de certificats clients en remplacement d'une authentification par simple mot de passe.

3.5 Sécuriser l'ensemble du système IoT

3.5.1 Stockage local sécurisé

Un point essentiel est de qualifier la sensibilité des informations et d'y adapter un niveau de protection correspondant. Toutes les données pouvant aider à compromettre le système et les données personnelles doivent impérativement être au plus haut niveau de sensibilité.

Un moyen efficace de protéger des données est de les chiffrer via l'utilisation d'un algorithme de chiffrement fort (par exemple AES-256).

Pour des raisons évidentes, les clés de chiffrement utilisées pour le stockage ne peuvent être sauvegardées sur le même terminal. Il est recommandé de ne sauvegarder que l'empreinte numérique (hash) d'une clé de chiffrement ou d'un mot de passe.

Dans l'idéal, il est recommandé de définir un mot de passe sur chaque clé de chiffrement. Ce mot de passe doit être déterminer par l'utilisateur afin de s'assurer qu'il soit le seul à pouvoir accéder aux données.

3.5.2 Encourager le piratage éthique

Une façon de faire la différence entre la recherche et le piratage informatique contraire à l'éthique est d'exiger la divulgation responsable des vulnérabilités découvertes. La divulgation responsable exige que la personne avise d'abord le fabricant ou les autorités gouvernementales et accorde un délai raisonnable pour que la vulnérabilité soit vérifiée et corrigée de façon indépendante avant de rendre publique l'information. Une autre approche, moins souhaitable, pourrait consister à exiger des chercheurs qu'ils s'enregistrent d'abord auprès d'un bureau gouvernemental ou du fabricant.

La législation devrait éviter de criminaliser les activités qui contribuent à promouvoir la sûreté et la sécurité. Les fabricants ne tirent aucun avantage financier de l'exposition des défauts de leurs produits, mais ces défauts doivent être identifiés pour améliorer la fonctionnalité et la sécurité. Les systèmes de primes d'erreurs payés par le fabricant peuvent permettre aux fabricants

d'atténuer la mauvaise presse tout en améliorant la qualité du produit à un coût inférieur à celui de l'embauche de testeurs d'intrusion payés. Les législateurs devraient prendre des précautions pour empêcher les poursuites en représailles contre les pirates informatiques éthiques. Des dispositions explicites devraient être inscrites dans la législation pour permettre la recherche et le piratage éthique. Les clauses libératoires devraient être évitées pour les fabricants qui pourraient mettre en œuvre des produits nocifs et peu sûrs à des fins lucratives. Les fabricants, les utilisateurs/citoyens et surtout les ingénieurs/chercheurs devraient rester au courant de la législation en cours et faire connaître leur position aux législateurs.

L'utilisation d'un objet connecté ou de son application peut mener à la découverte de failles passées inaperçues durant les phases de test. Ces découvertes peuvent être réalisées par des utilisateurs lambda, il est nécessaire d'être à l'écoute et de prévoir une procédure claire, simple à utiliser et anonyme de remontée des failles de sécurité.

Une fois la faille découverte, un correctif doit être rapidement déployé sur tous les exemplaires de ce produit afin d'en éviter l'exploitation par un pirate.

En mettant en place un programme de « bug bounty », une entreprise incite les utilisateurs à rechercher les failles et à les rapporter en échange d'une récompense, ce qui permet de les corriger avant que le grand public n'en soit informé.

Les hackers éthiques ont pour philosophie de rendre plus sûrs les systèmes qu'ils analysent. Ils contactent systématiquement les éditeurs et les industriels. C'est pourquoi il est important de prévoir une plateforme de mise en relation entre les acteurs. Enfin, les personnes remontant les failles peuvent vouloir protéger leur anonymat, il ne faut donc pas prévoir une procédure de remontée dont les éléments envoyés seraient rendus publics.

3.5.3 Instituer un conseil de certification de la sécurité et de la protection de la vie privée de l'IoT

En raison des problèmes de sécurité et de respect de la vie privée déjà causés par les appareils IoT, les ingénieurs doivent accepter la responsabilité de leurs créations. L'IEEE ou une organisation internationale devrait fournir un programme de certification professionnelle pour les concepteurs, les constructeurs et les fournisseurs de nouvelles technologies IoT qui s'engagent à respecter les meilleures pratiques établies pour la création des nouveaux dispositifs décrits dans ce document et d'autres sources dans leur création des produits IoT.

Le comité du programme devrait être habilité à vérifier si le fournisseur respecte les pratiques d'ingénierie responsables (en particulier les pratiques qui permettent la sécurité et la confidentialité de l'IoT), et à donner son aval aux fournisseurs qui sont liés par elles. Les mesures négatives constitueraient une autre dimension de ce programme de certification et devraient se limiter à la perte du statut de certification et à la possibilité de faire un rapport à la FTC ou à un autre organisme gouvernemental pour que d'autres mesures soient prises.

L'organisme de certification devrait vérifier au moins les éléments suivants des produits, protocoles et documents d'un fournisseur :

1. Les données sont traitées, utilisées, protégées et partagées de manière responsable.
2. Les protocoles utilisés ou recommandés ne divulguent pas d'informations sur les utilisateurs au-delà de l'intention explicite de ces utilisateurs.
3. Lorsque des problèmes de protection de la vie privée surviennent, le fournisseur certifié répond rapidement aux préoccupations.
4. L'authentification est suffisamment forte et suit des protocoles éprouvés.
5. Les appareils ne sont pas suramplifiés ou sous-protégés.
6. Les dispositifs doivent porter une étiquette d'identification difficilement falsifiable qui contient un lien Web où les clients peuvent trouver l'état de certification du dispositif ainsi qu'une description du dispositif (modèle et numéro de série, etc.). Cela peut se faire en coopération avec la FTC ou d'autres organismes nationaux.

De tels programmes de certification réduisent l'incertitude et fournissent aux fabricants de dispositifs, aux ingénieurs et aux auteurs les meilleures pratiques à suivre. Les tribunaux peuvent considérer la certification comme une preuve que

les pratiques acceptables qui sont généralement suivies. En cas de litige, un fournisseur peut indiquer la certification et dire qu'il a suivi de bonnes pratiques d'ingénierie.

3.6 La place juridique de l'IoT

3.6.1 Protection des données

Aujourd'hui, la protection de la vie privée est d'une importance primordiale pour les utilisateurs, c'est pourquoi la protection des données doit être intégrée dès la phase de conception pour tout objet connecté. De plus, le Règlement général sur la protection des données est entré en vigueur le 25 mai 2018.

Il est important que chaque entreprise désigne un responsable et mette en place des processus techniques et organisationnels tout au long des projets. Il faut pouvoir garantir que les données récoltées serviront uniquement les finalités de départ du produit et uniquement celles-là. Dernier point, en cas d'incident, il est important que des systèmes de secours permettent de garantir la sécurité des données tout au long du projet quoi qu'il arrive.

3.6.2 Informations, droits utilisateur et finalités des collectes de données

Les informations récupérées par les objets connectées peuvent être très variées avec un degré de confidentialité avancé, c'est pourquoi l'utilisateur doit être accompagné et informé sur ses données personnelles qui pourraient être collectées. Il est important que l'information transmise soit accessible, claire, précise et compréhensible dès le début de l'utilisation de l'appareil.

Il est important que l'utilisateur garde la maîtrise de ses données et puisse changer d'avis. Cela implique également un système qui garantisse à l'utilisateur que ses données puissent lui être accessibles, modifiables et supprimables. De plus, ce droit doit pouvoir être exercé directement et rapidement sans passer par des tiers.

L'utilisateur doit également pouvoir contacter un responsable de la protection des données s'il souhaite plus d'informations quant à la gestion et l'utilisation de ses données.

Les finalités de l'utilisation des données doivent être déterminées, compréhensibles et immuables. Il est obligatoire que ces finalités soient établies avant la collecte des données. Les données ne doivent pas être utilisées a posteriori pour d'autres finalités.

3.7 Conclusion

La forte croissance des objets connectés s'accompagne d'opportunités attrayantes pour les entreprises. Cependant, afin d'en profiter pleinement et sereinement, il convient d'évaluer les risques associés, qu'ils soient matériels, logiciels ou humains. Une méconnaissance de ces derniers entraînerait de graves conséquences pour le SI et donc l'entreprise.

Ainsi, la mise en place et l'application de mesures sécuritaires rigoureuses et adaptées se veut obligatoire afin de tirer parti au mieux de ces nouvelles technologies.

4 L'IA un allié de choix

4.1 Introduction

Aujourd'hui la complexité des raisonnements ne cesse d'augmenter notamment avec la quantité de données générée par la Big Data et les objets connectés. Avec l'augmentation des capacités de nos machines, l'IA progresse de façon fulgurante et pourrait bientôt s'imposer comme la solution évolutive à la plupart des problèmes de notre société moderne.

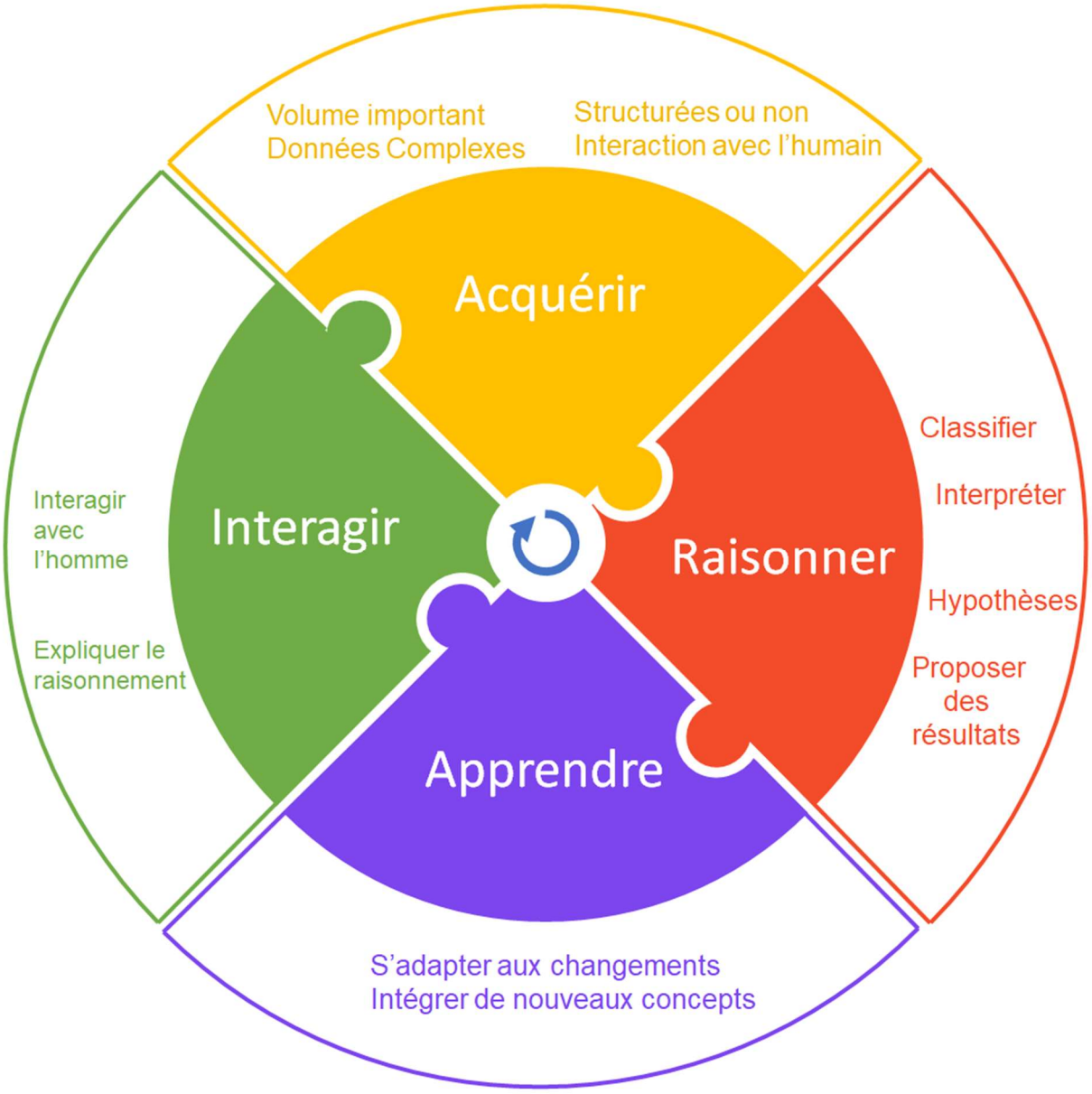

Figure 6 - Caractéristiques de l'IA

L'IA est aujourd'hui au cœur de nombreuses discussions et est très prisée par les entreprises pour gagner en performance et ainsi rester compétitives vis à vis de nos principaux concurrents qui sont les Etats-Unis et la Chine.

Nous pouvons catégoriser un programme comme étant une IA si elle a les 4 principales capacités suivantes :

- Elle doit pouvoir accéder à de grands volumes de données de types différents (structurées ou non).
- Elle doit avoir la capacité de raisonner, c'est à dire pouvoir travailler sur les données récoltées (Classifier, interpréter, obtenir des résultats, …)
- Elle doit apprendre grâce aux données qui lui sont fournies, c'est cette partie qui déterminera si l'IA est capable de s'adapter aux changements.
- Pour finir elle doit pouvoir interagir avec l'homme afin de nous donner toutes informations susceptibles de nous intéresser.

Chacune de ces parties est vitale pour réaliser une bonne IA et être serein de l'utilité de celle-ci dans le temps.

Avec ce modèle en tête nous pouvons réfléchir à comment l'IA peut nous aider en l'intégrant dans notre SI.

4.2 La place de l'IA dans le quotidien de l'entreprise

"L'intelligence artificielle sera l'avantage commercial du futur."

C'est le constat du rapport de PWC « A Revolutionary Partnership: How AI Is Pushing Man And Machine Closer ». En effet, le marché est aujourd'hui en grande conquête de l'IA. Deep learning (apprentissage profond), machine learning (apprentissage automatique), réseaux de neurones, entreprises et processus apprenants, la vague de l'IA déferle sur bon nombre d'industries et de secteurs d'activité, signant l'avènement d'une 4e Révolution industrielle.

Le marché de l'IA est florissant et comprend aujourd'hui une variété de technologies et d'outils qui facilitent la décision humaine : agent virtuel, reconnaissance faciale, reconnaissance d'images etc… ce qui révolutionne l'expérience et les services clients. L'impact à terme : une véritable transformation

des interactions clients/entreprises, des gains de productivité et une expérience cliente « augmentée », personnalisée et simplifiée.

Le HUB Institute - Digital Think-Tank, en partenariat avec IBM, vient de publier (en juin 2018) son HUBREPORT Future of Artificial Intelligence. Ce rapport fournit un état des lieux de l'IA et se penche sur cinq secteurs clés où l'IA révolutionne l'expérience client, tout en portant une réflexion sur le potentiel vertueux de la technologie au-delà de la sphère strictement marchande.

Cela met en avant une vision de l'IA où celle-ci ne vient pas remplacer l'humain mais améliorer grandement ses capacités et sa productivité. Le directeur de ce HUB institute expose par ailleurs clairement l'objectif de l'IA selon lui : "l'Augmentation de l'Intelligence Humaine", dont l'IA est un "moyen".

Si l'IA se développe actuellement de manière exponentielle, c'est qu'elle bénéficie d'une accélération intensifiée, qui se traduit par des levées de fonds et des financements sans précédents : les montants levés par des startups IA seront multipliés par 4 entre 2010 et 2020. De plus, l'IA promet une rentabilité accrue des entreprises, avec une prévision globale de +38 % en 2035 et un impact sur le profit des entreprises de tous les secteurs.

Un regain de financement sans précédent : Entre 1993 et 2017, 1980 levées de fonds ont été réalisées sur les cinq continents pour accompagner des projets intégrant l'intelligence artificielle. Entre 2010 et 2016, les montants levés par les jeunes pousses du secteur ont été multipliés par trois, passant de 0,6 à 1,8 milliard de dollars.

4.2.1.1 L'IA pour révolutionner l'expérience client

Selon l'étude réalisée par PwC, l'IA sera l'avantage commercial du futur. En effet, elle augmenterait significativement l'engagement client et aiderait à sa prise de décision.

D'après les 500 décideurs interviewés au sein de l'étude, voici dans quel ordre l'IA est susceptible d'impacter leur quotidien :

1. Assistants personnels virtuels (31%)
2. Analyses de données automatisées (29%)
3. Communications automatisées telles que les emails et les agents conversationnels. (28%)

4. Rapports de recherche automatisés et regroupement des informations (26%)
5. Analyses de l'efficacité opérationnelle automatisée (26%)
6. Analyses prédictives (26%)
7. Systèmes utilisés lors de prise de décision (21%)
8. Robotique (19%)
9. Analyses de vente automatisées (18%)
10. Apprentissage automatique (16%)

4.2.1.2 Quelques exemples d'applications dans différents secteurs d'activités :

La banque et l'assurance

- Le Crédit Mutuel utilise l'IA d'IBM, Watson, pour déterminer le niveau d'urgence des emails et répondre à certaines questions posées en langage naturel à partir d'une base d'information. Grâce à cela, les conseillers ont pu augmenter leur vitesse de réponse de 60%.
- Orange Bank et la Royal Bank of Scotland utilisent quant à elles Watson pour répondre directement aux questions des clients afin d'assurer un premier service client disponible 24h/24, 7j/7.

La distribution

Prédire la demande, automatiser des tâches et délivrer une expérience client personnalisée.

- La société Bazarchic utilise ainsi l'IA pour personnaliser son parcours client et cibler sa base de données lors de l'envoi de sa newsletter.
- Deliveroo et 1-800-Flowers.com s'appuient également sur de l'IA pour améliorer leurs services, que ce soit pour optimiser la livraison ou pour conseiller les clients.

Les télécoms

Baisser les coûts de gestion d'un client, faire progresser la qualité du service et créer des nouveaux services à forte valeur ajoutée.

- Bouygues Télécom utilise l'IA pour répondre aux questions les plus fréquentes et ainsi optimiser le temps de travail de ses employés.

L'industrie 4.0

- L'IA permet à la SNCF d'améliorer sa maintenance et de réduire ses coûts opérationnels tout en optimisant son service client.
- Schneider Electric utilise l'IA pour gérer à distance une ferme solaire au Nigéria.
- Les constructeurs automobiles investissent massivement sur l'IA pour leurs usines 4.0.

L'hôtellerie

- L'IA permettra d'automatiser certaines tâches et de repenser les process pour fluidifier l'expérience client et personnaliser le service.
- AccorHotels tire bénéfice de l'IA pour déterminer les meilleurs tarifs à pratiquer en fonction de la période de l'année.

4.3 L'intervention de l'IA pour nous aider dans nos tâches

Il y a une certaine tendance à voir l'IA comme une avancée dangereuse, susceptible de nous causer plus de tort que de bien tel que nous "voler" nos emplois.

Et si, au contraire, l'IA était le meilleur allié des salariés de votre entreprise ?

Peut-elle soulager les collaborateurs de leurs tâches les plus contraignantes, voire parfois inutiles, en les faisant à notre place ?

Peut-elle jouer un rôle d'assistant au travail qui nous faciliterait certaines actions ?

Est-ce que certaines IAs pourrait être construites spécialement dans le but d'améliorer le quotidien des employés ?

4.3.1 L'IA pour améliorer votre bonheur au travail

La jeune startup Bleexo a en effet décidé de s'attaquer à cette dernière question.

Au moyen d'une application, les collaborateurs se voient proposer à intervalle régulier, quatre ou cinq affirmations, telles que "Je pense que la stratégie de mon entreprise est pertinente", auxquelles ils doivent indiquer leur degré d'adhésion en 45 secondes. L'IA adapte les questions posées à l'expérience propre de chacun. Les réponses sont ensuite anonymisées et agrégées par Bleexo, qui joue un rôle de tiers de confiance, puis analysées et transmises aux managers. Ces derniers peuvent comprendre à tout instant où le bât blesse et quelles sont les priorités à traiter. Surtout, ils vont se voir proposer des actions correctives via des modules de micro Learning. La plateforme les accompagne ensuite dans leur mise en œuvre.

4.3.2 L'IA au service de tous les métiers

L'objectif est de libérer le plus possible les salariés des tâches répétitives et automatisables. Cela permettra de les rendre plus productifs et créatifs.

L'IA au sein des entreprises devrait doubler la croissance économique de la France d'ici à 2035, assure même Laurent Stefani, directeur exécutif pour l'intelligence artificielle au sein d'Accenture Technology en France. Voici quelques exemples :

Les ressources humaines

Les recruteurs passent près de 60% de leur temps à lire des CV. Pourquoi une personne devrait-elle lire 300 curriculums vitae si une machine peut en quelques secondes lui suggérer les 10 meilleurs ?

Comme Amazon, de nombreuses entreprises commencent à utiliser ce genre d'outils dopés à l'intelligence artificielle pour aider les ressources humaines à trouver les meilleurs talents.

Droits et comptabilité

Il existe déjà des logiciels qui peuvent "relire" des contrats de dizaines de pages en quelques secondes et les comparer avec une base de données de milliers

d'autres pour détecter des anomalies. L'IA permet de les perfectionner et de nouveaux systèmes vont maintenant permettre de les rédiger entièrement en fonction des objectifs de l'entreprise et de différents critères prédéfinis.

Lutter contre le surplus de formalisme en entreprise

Trois professionnels sur dix estiment que leurs réunions sont inutiles. Plus des deux tiers sont régulièrement agacés par les problèmes technologiques qu'ils rencontrent au cours de leurs réunions. Grâce à l'Intelligence Artificielle, ces frustrations pourraient bientôt être évitées. Par exemple, lors de sa Build Conference de 2018, Microsoft a dévoilé un appareil d'IA qui utilise la technologie de reconnaissance faciale pour scanner la salle de réunion et identifier les participants.

L'IA pourra ensuite changer la configuration des PC et des écrans en fonction des personnes qui les utilisent. Une innovation encore plus impressionnante serait **la mise en place automatique d'une salle de réunion** avec la présentation, les logos et la charte graphique **propre à chaque** client qui entrerait dans la salle.

De plus, Microsoft a également présenté un assistant virtuel pour accompagner des réunions. Cet assistant est capable de **retranscrire en temps réel toutes les conversations**. Cela sera particulièrement utile et permettra un important gain de temps. Elle pourra **assigner automatiquement des tâches** et ainsi envoyer des rappels de suivi et ajouter des éléments au calendrier des personnes concernées.

L'utilisation de la reconnaissance vocale pour fournir des sous-titres dans les réunions améliorera les collaborations entre bureaux et rendra les réunions plus efficaces. Les réunions internationales ne seront plus dépendantes de tiers traducteurs et les employés pourront interagir en temps réel avec leurs collègues ou clients étrangers.

En bref, l'intelligence artificielle est en train de prouver sa légitimité à remplacer l'homme pour les tâches rébarbatives sans pour autant complètement se substituer à lui. L'homme se différencie toujours de la machine grâce à ses émotions, essentielles au bon fonctionnement des relations au sein d'une entreprise.

4.4 L'IA dans nos infrastructures informatiques

4.4.1 Gestion de nos ressources informatiques

L'IA n'est pas seulement utile pour nous aider dans nos tâches mais est particulièrement habile pour optimiser l'utilisation de ressources qui sont la consommation énergétique et la disponibilité de nos équipements.

Les GAFAs ont bien compris cet aspect et l'ont d'ores-et-déjà implémenté dans leurs infrastructures.

Dès 2014, Google a utilisé une IA avec un réseau neuronal pour optimiser l'efficacité énergétique (PUE) de son installation IT et a obtenu des résultats plus que prometteur. Le groupe a réussi à optimiser son PUE de 15% et a également diminué l'énergie utilisée par leurs systèmes de refroidissement de 40%. Ce résultat a été possible grâce à la collecte de millions de données techniques issues de capteurs répartis sur l'ensemble de leurs data centers. Avec cette masse de données l'IA a pu proposer des points d'amélioration qui ont été appliqués par la suite.

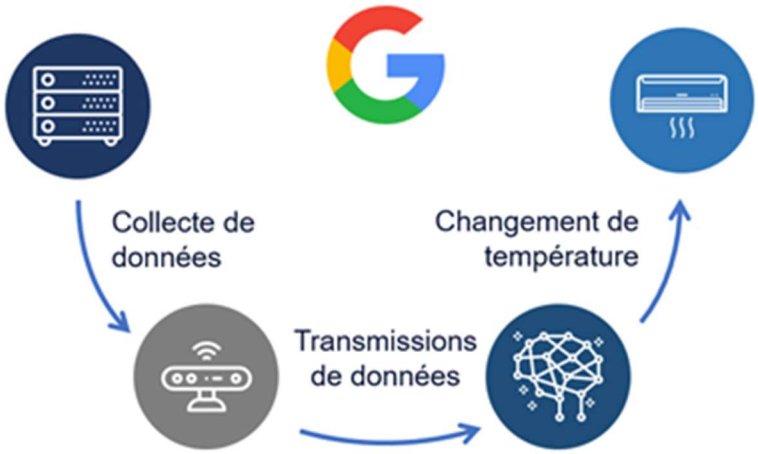

Figure 7 - Optimisation énergétique d'un Datacenter

La diminution des coûts n'est pas le seul avantage offert par l'IA. Amazon a choisi d'utiliser une IA pour anticiper les pics d'utilisation de ses ressources IT. L'IA utilisée prend en compte les historiques commerciaux du groupe pour savoir à quel moment les clients seront susceptibles d'augmenter l'utilisation des services d'Amazon, ce qui peut engendrer des pics de demande conséquents. Avec ces prévisions, Amazon peut créer de nouvelles ressources à allouer avant que les pics d'utilisation ne surviennent. La société a rendu ses services beaucoup plus disponibles et peut faire face à une forte demande plus facilement.

Figure 8 - Création de ressources pour anticiper les pics d'utilisations

4.4.2 Une maintenance corrective et préventive

L'IA a fait ses preuves auprès des cloud publics mais elle pourrait très bien rendre service à nos cloud privés pour optimiser la consommation d'énergie, tout comme Google, et nous aider à maintenir notre infrastructure.

Actuellement plus de 80% des pannes d'équipements sont prédictibles grâce à l'intervention de l'IA sur des sites physiques. Nous pourrions avoir un résultat similaire pour les cloud privés dans quelques années.

VMware a compris cet enjeu et améliore actuellement l'un de ces produits qui est utilisé pour le monitoring des infrastructures de Clouds privés. Dans cette nouvelle version du produit, une IA y sera incorporée afin de détecter les anomalies pour ensuite les corriger. Cette IA sera également capable de prédire certaines pannes futures et d'effectuer une maintenance préventive permettant de réduire au maximum l'indisponibilité de l'équipement.

Ce produit permettra donc d'avoir une infrastructure logicielle intelligente qui évoluera et se soignera sans l'intervention de l'humain.

Pour la partie physique, l'IA sera capable de détecter l'obsolescence des composants ou leurs dysfonctionnements afin de notifier un technicien pour qu'il change les pièces nécessaires.

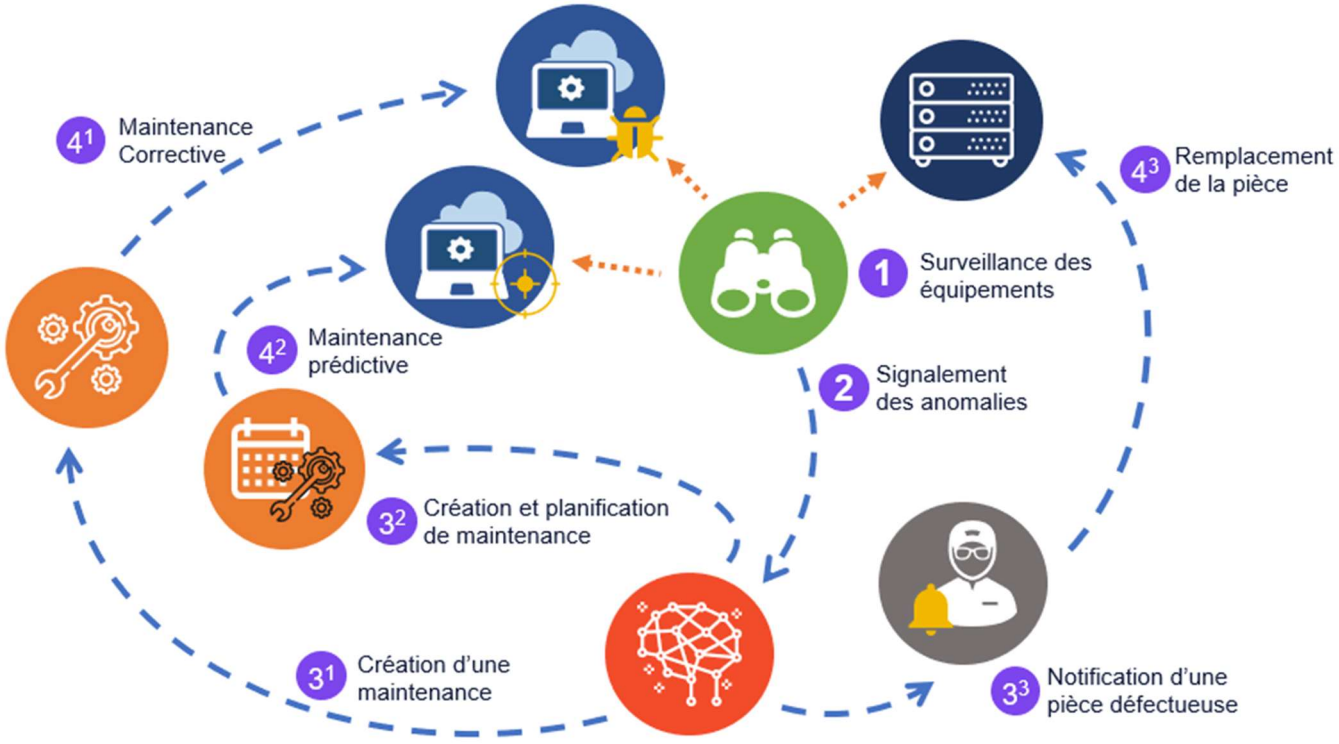

Figure 9 - Architecture intelligente

4.5 L'IA accessible par tous

4.5.1 L'IA disponible depuis le cloud

Aujourd'hui, les personnes capables de créer des intelligences artificielles sont peu nombreuses, qui sont généralement employées par de grandes entreprises qui possèdent le matériel nécessaire et qui leurs offrent des salaires élevés. Par conséquent, de plus en plus d'entreprises se renouvellent en devenant des sous-traitants de l'IA via le cloud.

Nous avons aujourd'hui plusieurs géants du web qui ont incorporé des services d'IA dans leurs offres cloud. Par exemple, Google a ajouté un service de Machine Learning avec des librairies open source et des outils pour mettre en place un réseau neuronal. Ces services permettent aux entreprises plus modestes d'avoir accès à l'IA avec pour atout majeur la qualité de l'IA. Cette qualité est due à l'expertise et la quantité de données détenues par ces grands groupes.

Les outils offerts par ces fournisseurs d'IA sont construits pour être simple d'utilisation afin que les ingénieurs n'ayant pas les compétences requises pour concevoir une IA puissent tout de même en créer et l'utiliser ce qui ajoute de la

valeur à leurs différents projets. Pour accompagner ses clients sur comment utiliser ces outils, Google et Amazon ont ouvert un département consacré au conseil sur l'IA.

4.5.2 Un système multi-agents intelligents

L'IA devient de plus en plus accessible grâce aux solutions cloud mais les besoins auxquels elle doit répondre ne cesse de se complexifier. Pour pallier à cela, une nouvelle approche permet de créer une Intelligence Artificielle Distribuée (IAD) que l'on nomme Système Multi-Agents (SMA).

Cette approche se focalise sur la concurrence et la distribution des expertises multiples. Cela permet de découper les problèmes que l'IA doit résoudre en plusieurs morceaux qui seront traités par les agents avec les capacités nécessaires. Chaque agent a une tâche qui lui est attribuée et le travail de chacun de ces agents contribue à un but commun.

Avec cette architecture d'IA, chaque IA a son rôle à jouer et peut être utilisée par d'autres IAs pour effectuer le même type de tâche. L'IA devient plus simple grâce à la séparation des expertises requises mais un nouveau problème se pose donc : Comment faire communiquer les différents agents (IA) et sécuriser ces interactions ?

Si nos agents sont utilisés uniquement à l'intérieur de l'entreprise, la solution est simple, il faut restreindre l'accès au réseau interne et définir une convention de langage entre les IAs. Cependant cette solution délaisserait une possible communication externe avec des clients ou des fournisseurs.

Si nos IAs sont amenées à communiquer avec d'autre IAs externes, plusieurs enjeux apparaissent en plus de la communication.

Rappelons comment nos IAs doivent communiquer. Il faut que chaque puisse communiquer avec n'importe quelle autre IA, il est donc nécessaire de définir des règles conversationnelles ainsi que des modalités de partage d'information. A l'avenir, il faudra convenir d'une norme ou d'un standard afin de faciliter les communications. A terme, une IA pourra découvrir de nouvelles IAs avec lesquelles elle n'avait jamais été en contact au préalable.

Ensuite, il faut définir avec quelles IAs nous partageons des informations confidentielles et à qui nous demandons des services. Il faut donc établir une restriction d'utilisation. Pour ce faire, il faut créer un réseau de confiance. Ce réseau peut être interne ou externe et doit accélérer les processus entre les acteurs majeurs liés à notre entreprise. Cependant, des informations dites publiques doivent toujours être accessibles pour des IAs externe à notre réseau de confiance. Le comportement de l'IA variera en fonction de son interlocuteur et des informations manipulées.

Une fois cela mis en place, les interactions entre IAs seront possibles et seront incroyablement nombreuses. Avec ce système, les échanges se feront très rapidement et si une erreur survient il faudra comprendre à quel moment elle est apparue et la rectifier.

Afin de garder une trace de toutes ces communications, la technologie de blockchain semble être une des solutions la plus adaptée. La blockchain permettrait de garder une trace de toutes les interactions grâce à son aspect base de données décentralisée et « l'impossibilité » de l'altérer.

Pour utiliser cette technologie, nous utiliserons le même réseau de confiance où chaque nœud du réseau aurait une copie de la chaîne permettant de détecter rapidement les anomalies apparues sur l'un des nœuds et ainsi les corriger.

Le fait que chaque nœud est un réplica de la chaîne permet de multiplier les validations. Cela permet également que chaque acteur puisse vérifier les données précédentes et de proposer une action corrective si nécessaire. Grâce à ces correctifs, les IAs pourrait également apprendre de leurs erreurs et ainsi ne plus les reproduire.

4.5.3 Une IA avec de l'empathie

L'IA est de plus en plus accessible auprès des professionnels grâce au cloud. Elle aide l'humain dans ses tâches et le remplace pour des tâches à faible valeur ajoutée.

Aujourd'hui une nouvelle génération d'IA a permis la création de chat bot pour communiquer avec le client. Malgré les progrès dans la compréhension et le langage, les clients privilégient le contact humain à l'interaction avec la machine.

Pourtant au sein des services clients, l'IA est utilisée comme assistant en communiquant directement avec l'utilisateur sans que ce dernier ne se rende compte. Cela a été possible grâce un traitement de donnée si efficace que l'IA a pu comprendre les questions qui lui étaient posées mais également les émotions que présentait l'utilisateur.

Certes l'IA occupe de plus en plus de place dans la communication avec les clients mais reste utilisée seulement pour des missions relativement simples. Pour des missions complexes, il est préférable de les confier à des humains qui pourront faire un accompagnement personnalisé.

Bien que l'IA comprenne des concepts complexes comme les émotions, cette dernière sera généralement utilisée en tant que support. Dans le cas du chat bot, l'IA a pour objectif de comprendre les intentions du client et de le rediriger vers une personne pouvant le conseiller davantage.

Ce concept d'IA empathique pourrait également se retrouver du côté client. Il est très probable que dans un futur proche, chaque individu est un double virtuel représenté par son IA personnelle. Cette dernière se chargerait de faire les premiers contacts commerciaux en négociant avec d'autres IAs d'entreprises. À ce stade, nous aurions 2 IAs devant protéger les intérêts de son propriétaire sans toutefois oublier l'aspect commercial.

L'IA personnelle doit comprendre le client et ses problèmes directes et indirectes. Pour comprendre ce point, prenons l'exemple d'un client qui veut voyager. L'IA va analyser les problèmes "directes" tels que la volonté de voyager, les préférences, la capacité financière, etc... puis l'IA devra comprendre quels sont les problèmes indirects liés à l'objectif principal, ici la volonté de voyager :

- L'envie de se retrouver seul ou au contraire de voir des personnes différentes
- Vouloir voir des horizons divers
- Se ressourcer
- etc...

Les problèmes indirects sont très complexes à cerner même pour l'humain durant la communication puisqu'ils ne sont généralement pas abordés et doivent être devinés par le vendeur.

La première IA personnelle permettra d'exprimer l'ensemble des problèmes et la seconde lui répondra en adéquation en lui proposant une offre "idéale".

4.6 Conclusion

Aujourd'hui l'IA est un élément omniprésent dans nos vies personnelles et professionnelles. La tendance va s'accentuer dans les prochaines décennies, c'est pourquoi il est primordial de connaître les différents domaines dans lesquels l'IA peut intervenir et imaginer de nouveaux domaines d'applications. Les entreprises, qu'elles soient de petite ou de grande taille, ont tout intérêt à intégrer ses technologies à leur SI en implémentant des solutions maisons ou en utilisant des services clouds.

Aux vues des avantages offerts par l'IA et l'accès à des infrastructures dédiées facilité, il est évident que la place de l'IA dans la Société va exploser et il est crucial que les ressources matérielles, logicielles et humaines des entreprises y soient préparées.

Bibliographies :

Ahmed, M. I. (2015). *Internet of Things (IOT) Standards, Protocols and Security Issues*. Récupéré sur https://ijarcce.com/wp-content/uploads/2015/12/IJARCCE-109.pdf

Asay, M. (2017). *85% of big data projects fail, but your developers can help yours succeed*. Récupéré sur https://www.techrepublic.com/article/85-of-big-data-projects-fail-but-your-developers-can-help-yours-succeed/

Aubay, R. (2016). *L'analyse de données*. Récupéré sur https://www.aubay.com/wp-content/uploads/2016/07/Analyse-de-donnes-VF.pdf

Bendor-Samuel, P. (2017). *How to eliminate enterprise shadow IT*. Récupéré sur https://www.cio.com/article/3188726/it-industry/how-to-eliminate-enterprise-shadow-it.html

Bertier, L. (2018). *Demain, l'intelligence artificielle mettra fin à vos frustrations en réunion !* Récupéré sur https://startco.lesechos.fr/posts/demain-lintelligence-artificielle-mettra-fin-a-vos-frustrations-en-reunion/

Boccara, G. (2017). *L'intelligence artificielle : quel impact sur le monde du travail ?* Récupéré sur https://www.welcometothejungle.co/articles/intelligence-artificielle-quel-impact-sur-le-monde-du-travail

CIGREF. (s.d.). *Mise en œuvre opérationnelle de l'intelligence artificielle dans les grandes entreprise*. Récupéré sur https://www.cigref.fr/wp/wp-content/uploads/2017/10/CIGREF-Cercle-IA-2017-Mise-en-oeuvre-operationnelle-IA-en-Entreprises.pdf

Claveria, K. (s.d.). *13 stunning stats on the Internet of Things*. Récupéré sur 2017: https://www.visioncritical.com/internet-of-things-stats/

Corser, G. (2017). *Internet of things (iot) security best practices*. Récupéré sur https://internetinitiative.ieee.org/images/files/resources/white_papers/internet_of_things_feb2017.pdf

Crochet-Damais, A. (2017). *Avec l'IA, l'automatisation des clouds s'accélère*. Récupéré sur https://www.journaldunet.com/solutions/cloud-computing/1194830-avec-l-ia-l-automatisation-des-clouds-s-accelere/

DJEDOUR, R. (2017). *Big data : du prédictif au prescriptif, une nouvelle ère pour les entreprises*. Récupéré sur https://www.lesechos.fr/08/06/2017/lesechos.fr/030373431440_big-data---du-predictif-au-prescriptif--une-nouvelle-ere-pour-les-entreprises.htm

Grandmontagne, Y. (2018). *Les 5 tendances qui façonneront le cloud pour 2020*. Récupéré sur https://itsocial.fr/enjeux/cloud-computing/cloud-public-prive-hybride/5-tendances-faconneront-cloud-2020/

Grosdidier, A. (2017). *Des algorithmes peuvent-ils être éthiques ?* Récupéré sur https://usbeketrica.com/article/des-algorithmes-peuvent-ils-etre-ethiques

Grosjeanne, O. (2018). *2018 marquera l'avènement de l'IA, de l'informatique hybride, distribuée et automatisée*. Récupéré sur https://www.journaldunet.com/solutions/expert/68253/2018-marquera-l-avenement-de-l-ia--de-l-informatique-hybride--distribuee-et-automatisee.shtml

HUB Institute - Digital Think-Tank. (2018). *L'IA, au service de l'entreprise augmentée*. Récupéré sur http://www.up-magazine.info/index.php/intelligence-artificielle/intelligence-artificielle/7864-l-ia-au-service-de-l-entreprise-augmentee

IBM. (2017). *10 Key Marketing Trends for 2017*. Récupéré sur https://public.dhe.ibm.com/common/ssi/ecm/wr/en/wrl12345usen/watson-customer-engagement-watson-marketing-wr-other-papers-and-reports-wrl12345usen-20170719.pdf

inprincipio. (2017). *Système multi-agents et IA distribuée*. Récupéré sur https://www.inprincipio.xyz/ia-distribuee/

ioTRUST. (2018). *Guide de bonnes pratiques - Développement de système IoT*. Récupéré sur https://www.fondation-maif.fr/up/pj/20180612_BBU_Guide-bonnes-pratiques-V7_WEB.pdf

L, B. (2018). *Intelligence Artificielle et Big Data : une convergence révolutionnaire*. Récupéré sur https://www.lebigdata.fr/intelligence-artificielle-et-big-data

MARTIN, P. (2018). *Le Plan Copenhague : Comment nous avons migré la plateforme 6play vers Le Cloud*. Récupéré sur https://leanpub.com/6cloud

Michael Page. (s.d.). *L'intelligence artificielle : jusqu'à l'empathie automatisée ?* Récupéré sur https://www.michaelpage.fr/actualit%C3%A9s/%C3%A9tudes-barom%C3%A8tres/fw-travail-de-demain/intelligence-artificielle-empathie

Nick, I. (2017). *Why do big data projects fail?* Récupéré sur https://www.information-age.com/big-data-projects-fail-123468000/

NOMIOS. (2018). *APPLICATION DE LA BLOCKCHAIN À LA SÉCURITÉ INFORMATIQUE*. Récupéré sur https://www.nomios.fr/application-blockchain-securite/

PEYRE, M. (2018). *IA : Taxonomie et cas d'usages*. Récupéré sur http://www.timspirit.com/ia-intelligence-artificielle/

Pierre, F. (2017). *De l'application du Machine Learning à la virtualisation d'applications*. Récupéré sur https://www.systancia.com/machine-learning-virtualisation/

Renouard, G. (2018). *L'INTELLIGENCE ARTIFICIELLE, UN SERVICE COMME LES AUTRES ?* Récupéré sur https://atelier.bnpparibas/prospective/article/l-intelligence-artificielle-service-autres

SINGH, P. (2017). *10 Reasons Why Big Data And Analytics Projects Fail*. Récupéré sur https://www.analyticsindiamag.com/10-reasons-big-data-analytics-projects-fail/

Uthayasankar, S., Muhammad , M., Zahir , I., & Vishanth , W. (2017). *Critical analysis of Big Data challenges and analytical methods*. Récupéré sur https://ac.els-cdn.com/S014829631630488X/1-s2.0-S014829631630488X-main.pdf?_tid=5feabd26-fdc7-4b6d-8b1d-f486e0fc7598&acdnat=1543414975_00619fde9a96387cc09bc5a7c13753ab

Wastie, S. (2018). *Is The Promise of Big Data Hampered by Skills Shortages and Poor Performance?* Récupéré sur https://unraveldata.com/uk-research-big-data/