Le rôle de la gouvernance des données numériques dans le renforcement des avantages compétitifs des entreprises: une mise en perspective des facteurs d'impact

Mehdi El Yahyaoui
Docteur es-sciences économiques
Faculté des sciences juridiques, économiques et sociales – Souissi
Université Mohamed V, Rabat, Maroc

Abstract

The development of new information and communication technologies (NICTs) over the last few decades is unprecedented in several respects. This transformation has emerged from digitisation. The convergence of ICTs, the proliferation of digital technologies, Internet of Things devices and digital platforms, social networking applications, geolocation techniques, etc., have set in motion a process of capturing, extracting, collecting, exploiting and analysing mass data that is unprecedented in human history. The aim of this article is to focus on the new value chain models associated with Big Data. These new business models are designed and developed from and around data. They are indeed Big Data value chains. This article also looks at the phenomenon of the simultaneous automation of tasks, which follows the adoption of data as the main source of value, and leads to the obsolescence of traditional value chain models, i.e. those in force before the advent of digital technology.

Key words : Digital data; Big Data; Value chain; Digital technologies.

Résumé

Le développement des nouvelles technologies de l'information et de la communication (NTIC), au cours des dernières décennies, est inédit par plusieurs de ses aspects. Cette mutation a émergé de la numérisation. La convergence des TIC, la démultiplication des technologies numériques, des dispositifs de l'Internet des Objets et des plateformes numériques, des applications de réseaux sociaux, des techniques de géolocalisation…etc., ont enclenché un processus de captation, d'extraction, de collecte, d'exploitation et d'analyse des données de masse sans précédent dans l'histoire humaine. Le présent article vise à mettre l'accent sur les nouveaux modèles de chaîne de valeur relatifs au Big Data. Ces nouveaux modèles d'affaires sont conçus et développés à partir et autour des données. Ce sont bel et bien les chaînes de valeur du Big Data. Il est de surcroît question dans cet article, de décrypter le phénomène d'automatisation simultanée des tâches, qui fait suite à l'adoption des données comme principale source de valeur, et entraîne l'obsolescence des modèles de chaîne de valeur traditionnels, c'est-à-dire ceux en vigueur avant l'avènement du numérique.

Mots clés : Données numériques ; Données de masse ; Chaîne de valeur ; Technologies numériques.

1. Introduction

Le développement des nouvelles technologies de l'information et de la communication (NTIC), au cours des dernières décennies, est inédit par plusieurs de ses aspects. Cette mutation a émergé de la numérisation. La convergence des TIC, la démultiplication des technologies numériques, des dispositifs de l'Internet des Objets et des plateformes numériques, des applications de réseaux sociaux, des techniques de géolocalisation…etc., ont enclenché un processus de captation, d'extraction, de collecte, d'exploitation et d'analyse des données de masse sans précédent dans l'histoire humaine.

Ce ne sont plus les procédés «statiques», longtemps en vigueur en matière de réseaux et d'applications, qui déterminent de nos jours, les choix stratégiques des entreprises. Le Big Data, l'informatique en nuage, les réseaux virtuels, l'intelligence artificielle et les algorithmes de prédiction et de recommandation, les ont largement supplantés dans la production de la valeur, tout autant que dans les stratégies de compétitivité, de performance et de veille concurrentielle.

Les configurations de marchés qui allaient en découler, en portent foncièrement la marque et les caractéristiques.

Ainsi, le stockage centralisé des données massives, provenant de sources différentes multiples, à la faveur de l'informatique en nuage (Cloud Computing), tend par exemple à écarter, voire à supplanter les anciens systèmes d'information internes ne pouvant s'adapter aux solutions informatiques distribuées, évolutives, flexibles et fiables.

En effet, ces données, qu'elles soient ou non structurées, et stockées de façon distribuée, tout autant que les sources variées dont elles proviennent, interpellent le mode de leur représentation. Il ne s'agit pas uniquement de les collecter et stocker, encore faudrait-il les analyser et les interpréter pour en extraire du «sens».

La chaîne de valeur des données dont il est assez souvent question en milieu professionnel, pourrait être considérée ainsi comme l'élément clé pour la conception d'un modèle de gestion efficace des processus de création de la valeur au sein des organisations.

En effet, le phénomène d'automatisation simultanée des tâches (la numérisation des processus de bout en bout, diraient d'aucuns), faisant suite à l'adoption des données comme principale source de valeur, a entraîné l'obsolescence des modèles de chaîne de valeur traditionnels, c'est-à-dire ceux en vigueur avant l'avènement du numérique.

Les nouveaux modèles de chaîne de valeur en oeuvre depuis le début du siècle courant, sont conçus et développés à partir et autour des données. Ce sont bel et bien les chaînes de valeur du Big Data.

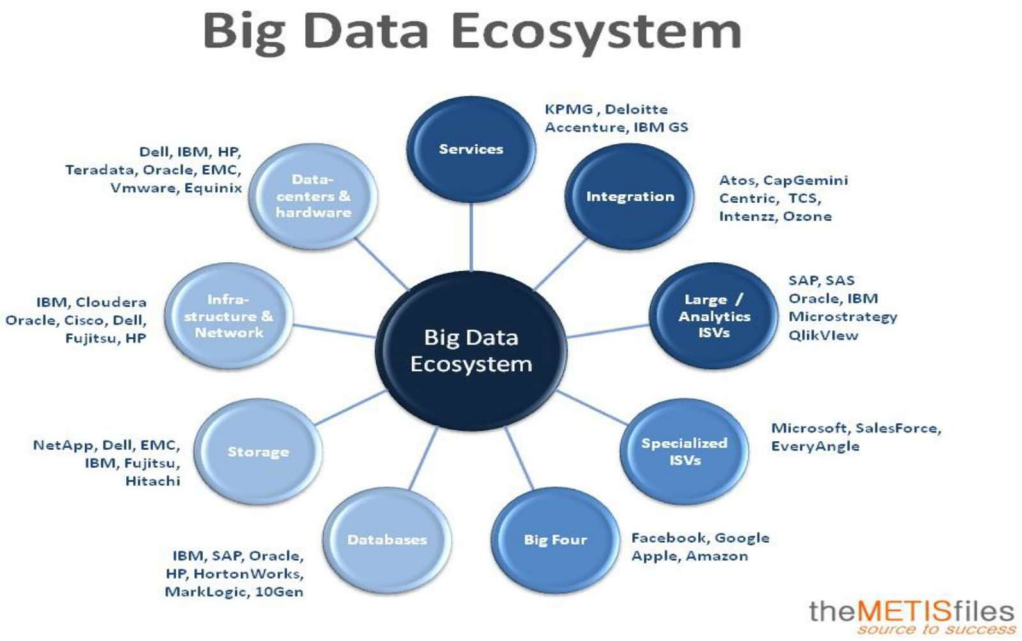
Elles le sont ainsi, parce qu'elles portent sur les caractéristiques du Big Data dont plus particulièrement le volume, la variété et la vélocité. Elles le sont davantage, car elles s'appuient sur les flux de données et traces de données pour extraire des informations portant un sens.

C'est que les chaînes de valeur dans ce cas, représentent les différentes étapes, ordonnées et cohérentes, d'utilisation des dispositifs et des outils d'analyse des «données massives» (Big Data Analytics) pour passer de la capture, la génération et le traitement des données à la «production du savoir et de la connaissance», autrement à la production du sens. La monétisation dont on parle à la fin du cycle, n'est en dernière instance, que la composante visible de l'écosystème d'ensemble.

2. Ecosystème des données massives

Quand il avait à parler des données dans le contexte d'une économie de plus en plus numérisée, Douglas Laney ne fait en réalité que confirmer notre assertion de départ : les trois dimensions qui permettent de caractériser à la fois les défis et les opportunités relatives aux quantités de plus en plus énormes des «données de masse», sont incontestablement le volume, la variété et la vitesse, autrement les fameux «3V» (Laney, 2001).

Figure n° 1 : Ecosystème des données de masse



Source : www.themetisfiles.com

Le «Big Data» est ainsi, cet ensemble d'informations dont le volume est phénoménal, extrêmement varié et provenant de sources multiples. Les caractéristiques qu'il porte, tout autant que le mode de structuration dont il devrait faire l'objet, interpellent le rapport coût/efficacité s'agissant du traitement et de l'exploitation des données dont il est l'émanation. L'opération de prise de décision en est largement conditionnée, non seulement pour ce qui est de l'organisation interne de l'entreprise, mais aussi pour ce qui est de ses grands choix stratégiques.

Or, une telle perception est en partie mise en doute, car la définition même du Big Data est évolutive, dynamique et subjective par certains de ses aspects (Manyika 2011).

En effet, les «données massives» pour des firmes comme Google ou Amazon par exemple, ne sont pas de même nature, encore moins de même portée que celles des compagnies d'assurance ou des opérateurs de télécommunications de petite et de moyenne taille. Ces entreprises ne rivalisent ni par la taille, ni par l'innovation, ni par la place sur le marché, encore moins par le pouvoir de disruption qu'elles pourraient enclencher dans le temps et dans l'espace.

Les définitions généralement adoptées ne visent par conséquent, pas à remettre en question le coeur de la problématique posée, car toujours centrée sur l'impact des technologies d'information et de communication, mais l'ont juste perceptiblement déplacé pour faire place à ces ensembles de données dont les quantités et les tailles sont telles qu'elles dépassent les seuils de «l'entendement» et deviennent disproportionnées par rapport aux dispositifs et outils logiciels de base de données «typiques».

C'est que la taille du Big Data dépasse généralement la capacité de ces outils à la saisir, la traiter et l'analyser. Elle la dépasse aussi quand il s'agit de gérer ces données, les traiter et les analyser pour en dégager des informations fiables et en extraire de la valeur. L'extraction du sens à laquelle on faisait allusion plus haut, se ramène en fait à cette valeur dont il n'est pas toujours aisé d'en dégager la teneur, notamment à partir de données de masse appelées incessamment à être structurées.

C'est la raison pour laquelle, des «V» additionnels, plus synthétiques et plus explicatifs, devraient être ajoutés, pensaient d'aucuns, à ceux jusqu'ici en vigueur, et ce pour rendre compte justement de cette capacité disproportionnée à prendre en charge le cycle de vie des données de masse.

3. De l'idée d'écosystème des données massives

Le terme d'«écosystème» est emprunté au domaine scientifique, pour désigner ce réseau complexe de systèmes interconnectés et interactifs.

Il exprime la réalité selon laquelle si les systèmes traditionnels tendent à être rapidement supplantés par de nouveaux et puissants systèmes d'information (telle l'informatique en nuage par exemple), les gros ensembles de données, comme les fichiers journaux, les traces de «sentiments» ou les flux de clic sur les

réseaux sociaux, ne pourraient être stockés, encore moins protégés dans de simples serveurs centraux ou dans un quelconque endroit fixe du Cloud.

Autrement dit, le traitement et l'exploitation des données massives requièrent des outils de plus en plus avancés, qui soient à même de traiter et de stocker en temps réel, les milliards d'octets de données, avec simultanément ou en parallèle, des milliers voire même des millions de transactions s'effectuant par seconde.

Ce traitement, tout autant que les possibilités d'exploitation des données, font partie intégrante de cet écosystème, aux dimensions composites et variées, mais fédérées par la logique de cohérence qu'impose la configuration systémique : sources de données, gestion de données, stockage de données et intégration de ces dernières dans la rétroaction d'ensemble.

C'est au niveau de la phase du traitement des flux que s'opère réellement le cycle d'analyse et d'extraction de la valeur, du sens et de ce que certains appellent l'«intelligence économique».

3.1. Des sources et de la gestion des données massives

Dans les écosystèmes d'information modernes, les sources de données se rapportent à des couches de sources multiples et diverses, mais appartenant pour la quasi-majorité d'entre elles, à des entités publiques ou privées, ou toutes les deux à la fois.

Les données publiques sont celles qui émanent généralement des organisations et des entreprises publiques. Elles proviennent de leurs systèmes internes traditionnels, ou de systèmes leur appartenant fondés sur le nuage (le Cloud). Elles pourraient être externes aussi et provenir de leurs partenaires ou de divers tiers. Leurs architectures de stockage sont assez souvent larges et flexibles pour pouvoir accueillir différents types de données.

Or, quelle qu'en soit l'origine, la capture de ces ensembles de données, de plus en plus hétérogènes et aux sources de plus en plus variées, est sinon le coeur de l'écosystème des données, du moins son pilier central.

Dans le même temps, le développement et la mise en place des applications informatiques à base de données ouvertes (Open Data), requérant des formats qui soient lisibles par la machine, ont précipité l'avènement des langages informatiques XML et JSON, devenus depuis quelques années seulement, les formats les plus adéquats à la fois pour l'Internet et pour les applications

mobiles. Leur souplesse, tout autant que leur ouverture, leur a permis d'intégrer facilement les technologies des navigateurs d'Internet et des serveurs.

Mais la phase d'acquisition et de capture des données ne pourrait générer de la valeur pour l'écosystème, qu'associée à la phase d'interconnexion à établir et/ou à maintenir entre les diverses sources de données. Cette phase est d'autant plus complexe qu'elle met en relation des données générées ou acquises sous une forme non structurée. L'essor, ces dernières années, des technologies sémantiques et des principes qui leur sont associées, traduit en fait, une telle complexité (Berners-Lee 2006, Janev et Vranes, 2011).

C'est la raison pour laquelle, l'unification de la représentation des informations fiables tirées des données et de leur traitement, appelle une normalisation des schémas, leur hiérarchisation et l'adaptation graduelle de leur langage.

Une telle évolution est dictée non seulement par le besoin d'interconnexion qu'exige la nature intrinsèque de l'écosystème des données, mais aussi parce que l'on assiste sur le réseau Internet notamment, à des centaines de milliers de dépôts de données, utilisés par les entreprises et les organisations dans l'objectif d'améliorer leurs produits et processus (Berbers et Zwaenepoel, 2019).

Par ailleurs, les origines des données publiques, qu'elles soient des données statistiques, des conversations, des images, des vidéos, des podcasts ou des audio, sont chargées d'informations fiables et crédibles, offertes dans l'immédiat et en temps réel. Elles permettent aux entreprises de puiser d'une mine contenant de multiples interactions effectuées sur les plateformes numériques, les réseaux sociaux et les applications de discussion instantanée et à distance. Leur analyse est fort conséquente pour l'écosystème dans son ensemble (Marr, 2018).

Il faudrait préciser à ce niveau, que les outils et dispositifs de l'Internet des Objets sont porteurs de réelles opportunités, mais sont aussi inducteurs de risques et de défis, plus particulièrement pour les industries ou secteurs d'activité dont les modèles d'affaires sont déficients ou peu enclins au changement.

Les modèles déficients dont il s'agit sont ceux qui n'ont pas réussi à s'approprier la logique de gestion des données, et à fortiori celle des données de masse : leur sélection, leur collecte, leur combinaison, leur stockage, tout autant que l'analyse qui devrait en découler. Ils ne sont par conséquent, pas inscrits

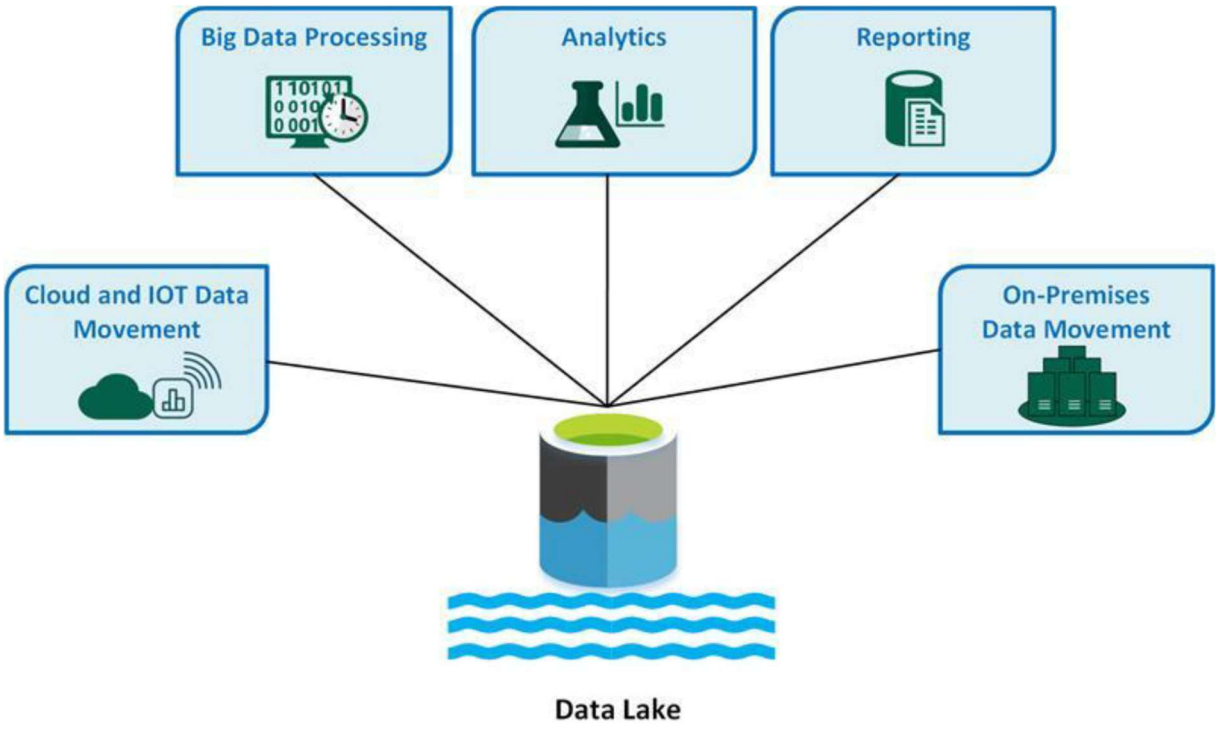
dans une perspective de Big Data, pour mettre en place les dispositifs et structures de traitement et d'analyse des données.

Les défis qu'ils affrontent tiennent en d'autres termes, à leur incapacité à ériger un écosystème de données, avec des canaux de traitement et d'analyse des données de bout en bout. Une telle orientation est d'autant plus vertueuse qu'elle prend de plus en plus appui sur les solutions permises par l'informatique en nuage et les lacs de données.

L'informatique en nuage, expression fort médiatisée ces dernières années, est ce «paradigme» qui porte sur le partage des données et les calculs sur un réseau de noeuds évolutif, comprenant des centres de données, des sites web, voire des ordinateurs personnels (Assuncaoa, 2015).

Les lacs de données sont, quant à eux, cette nouvelle architecture de stockage des données à l'état brut, sans grande considération de leurs sources, leurs structures ou leurs tailles.

Figure n° 2 : Les lacs de données



Source : www.attassian.com

Force est de relever ici que les approches dites d'entreposage (Data Warehousing), fondées sur le traitement préalable des données pour des buts spécifiques (avec un référentiel de données structurées), sont considérées

comme étant obsolètes, ou à tout le moins dépassées. Elles le sont encore davantage qu'elles posent des problèmes se rapportant à l'intégration des données ou à l'ajout de sources nouvelles de données. La question de la qualité et de la fiabilité des données est fort interpellée à ce niveau.

Or, les quantités énormes de données que traduit l'expression données massives, sont-elles aussi entachées de biais de qualité et de fiabilité, auquel cas il faudrait les traiter aussi, avant d'en faire usage.

En effet, les interférences éventuelles ou les dysfonctionnements imprévus des capteurs par exemple, pourraient donner lieu à des données compromises à l'origine, marquées d'incertitudes ou faisant l'objet de « bruits intrinsèques » inattendus. Ce qui induirait inévitablement des données contradictoires les unes par rapport aux autres, ou engendrer des données biaisées et peu opératoires. Dans ce cas, l'on sera astreint à recourir et à appliquer des méthodes systématiques de contrôle et d'évaluation de la qualité de l'ensemble des données de l'écosystème (Zaveri et al., 2016).

Ces méthodes sont par ailleurs en vigueur, épaulées en cela par les développements réalisés dans le domaine de la représentation sémantique des sources et des processus de données, l'objectif étant d'assurer une interaction et une interopérabilité entre les processus et les systèmes interconnectés. Il en va ainsi de l'analyse qui s'ensuivrait pour la génération de la valeur dans l'écosystème.

3.2. De l'analyse des données massives

L'analyse des données est une dimension centrale de l'écosystème des données massives. Elle comprend les technologies, les systèmes et les process qui s'appuient sur les techniques d'exploration des données et leur analyse statistique (Chen et al., 2012). Le choix des modèles de traitement, tout autant que celui des solutions analytiques, dépend largement des enjeux économiques et commerciaux à relever, mais aussi de la nature du domaine ciblé (Janev et Vranes, 2005).

Le domaine en question pourrait être le commerce électronique, l'intelligence et la veille économiques, l'administration en ligne, les soins relatifs à la santé, comme il pourrait concerner l'efficience énergétique, la gestion des urgences,

de la production ou de la sécurité (Sun et al. 2014, Janev et al. 2018, Berbakov et al. 2018, Mijovic et al., 2019).

Il faudrait préciser que les solutions de problèmes fondées sur l'analyse des données massives reposent et dans une large mesure, sur les services d'analyse mobiles, textuels, de réseau ou de web.

La mobilisation de telle ou telle solution d'analyse de données dépend naturellement, des types de problèmes à résoudre et de défis à relever. Ceux-ci pourraient être de nature financière par exemple, comme l'évaluation des risques dans les établissements bancaires et financiers. Comme ils pourraient être de nature industrielle, en permettant par exemple la maintenance prédictive, la cognition, la détection des anomalies au sein des usines de production, ou les réponses automatiques dans les salles de contrôle … etc.

Dans ces cas, les techniques relevant de l'apprentissage automatique sont déployées et mises en oeuvre : arbres de décision, régression, machines à vecteurs de support, règles d'association…etc. (Upadhyaya, 2013).

Toute la philosophie dans l'analyse des données massives, est de choisir les meilleures méthodes et techniques, pour non seulement réduire les incertitudes qui pourraient entacher l'extraction des données, mais aussi pour faire de cette dernière un outil de résolution des problèmes et de perfectionnement du processus de prise de décision. La sémantique est ici d'un apport appréciable.

En effet, la variété des sources des données massives appelle la nécessité de concevoir et de mettre en oeuvre l'architecture qui assurerait leur traitement en temps réel. Leur volume, tout autant que la vélocité qui les caractérise, incitent par ailleurs, les entreprises à renforcer davantage leur capacité d'«interception», mais aussi de stockage.

Car, si les données de masse émanent assez souvent, d'une multitude de sources, variées et hétérogènes (capteurs, appareils et dispositifs intelligents, réseaux sociaux…etc.,), elles le sont généralement dans un état brut, semi-structuré, ou systématiquement non structuré. Le niveau de structuration réduit d'une partie d'entre elles, ne signifie nullement qu'elles sont dépourvues de valeur ou d'intérêt pour l'entreprise. Elles font toutes l'objet d'un stockage, mais ne pourraient rendre compte de leur «richesse» que mises à l'épreuve du traitement.

Les «offres» pertinentes de stockage sont généralement connues. Elles concernent notamment les solutions de stockage distribué et de fusion des données, mais concernent surtout l'informatique en nuage évoquée plus haut (Liu et al., 2015).

Ces méthodes relèvent du souci, trop souvent émis par les entreprises, pour réduire la complexité inhérente aux données et rétrécir leurs échelles pour les rendre aptes à exploitation, à traitement et à analyse.

Il s'agit, autrement dit, des méthodes liées à la réduction dimensionnelle, à l'échantillonnage et au codage. Simultanément, ces méthodes permettent d'opérer un prétraitement des données massives, en l'occurrence l'extraction approfondie des données, leur nettoyage, leur intégration ainsi que leur transformation (Wang, 2017).

Quant à la variété ou l'hétérogénéité des données massives, il faudrait rappeler à nouveau, qu'elles tiennent généralement à quatre dimensions principales : dimensions liées à la variété des structures des données, à la variété des médias, aux variétés sémantiques et aux variations liées à la disponibilité :

°- La variété des structures concerne à ce propos, la représentation des données, car les formats des images provenant des satellites par exemple, diffèrent de ceux représentés par des tweets ou des contenus circulant sur le web. La différence des formats est proportionnelle à la différence des capacités de stockage.

°- La variété des médias. C'est la variété qui tient essentiellement aux supports via lesquels les données massives allaient être véhiculées, transmises et livrées. Ainsi, les audio des discours et leur transcription par exemple, représentent la même information, mais empruntant chacun son propre support ou son propre média.

°- Les variétés sémantiques comprennent les significations des unités destinées à mesurer et à décrire les données nécessaires à l'exploitation, à l'analyse et à l'interprétation des données capturées.

°- Les variations de disponibilité signifient que l'accès aux données est soit permanent et durable (cas des caméras de surveillance de la circulation par exemple) soit intermittent, comme le cas des satellites survolant un espace géographique quelconque.

3.3. Interopérabilité et types d'analyse des données massives

L'interopérabilité des données (l'on parlera plus tard de la mise en données) est l'un des grands défis de l'écosystème des données massives.

Lorsqu'on prend le cas de l'Europe par exemple, l'on ne manquera pas de relever la portée dudit défi. Dans leur vision de l'UE 2030, les pays de l'Union européenne prévoient la mise en place d'un marché unique à la faveur de la normalisation d'un ensemble de lois relatives aux données, devant faire l'objet d'application dans tous les Etats membres de l'Union. La philosophie était de créer «un espace européen unique des données» (European Commission, 2020).

La finalité de ladite vision était en effet, d'ériger un véritable marché unifié des données, qui soit caractérisé par l'accès rapide et sans contraintes, des entreprises européennes à un volume massif de données à caractère industriel. Le programme de normalisation des technologies de l'information et de la communication adopté, soutenait à l'évidence cette vision commune (European Commission, 2022).

La vision ainsi tracée concerne les cinq grands domaines en rapport avec les données : le Big Data, l'Internet des Objets, l'informatique en nuage, la 5G et la cybersécurité.

Or, pour pouvoir garantir une interopérabilité à grande échelle des données (ou systèmes de données), leur intégration progressive et leur interactivité en termes d'échange, et pour encourager l'innovation à base de données massives, la normalisation s'avérait incontournable. Elle s'est faite en plusieurs étapes et selon différents processus et niveaux d'action : schémas relatifs aux métadonnées, formats relatifs à la représentation des données, conditions de délivrance des licences relatives aux données ouvertes…etc.

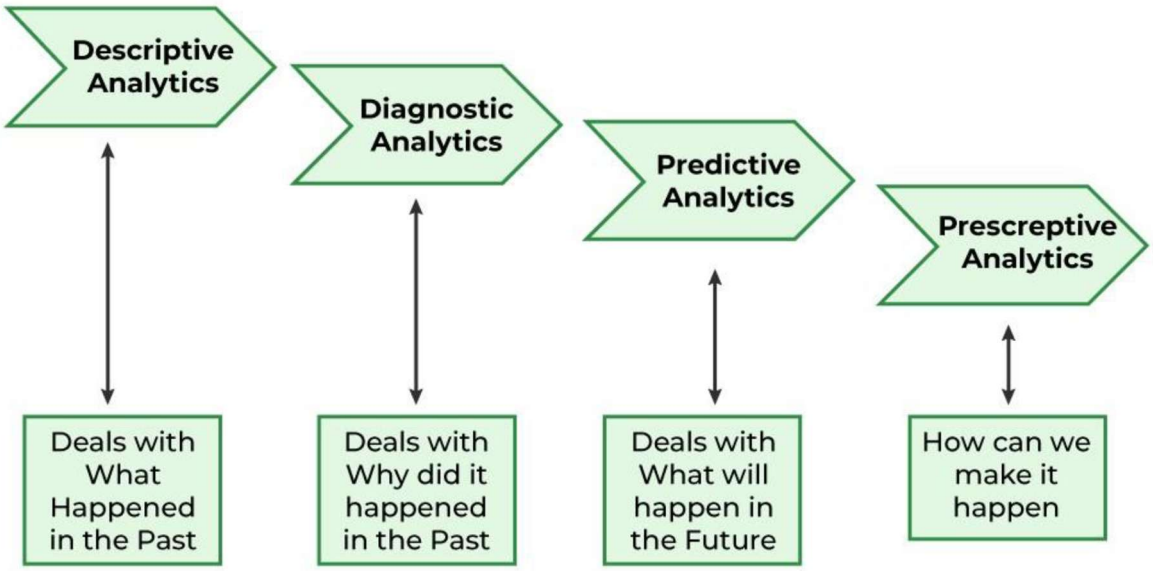
Tous les types de données ont été concernés (données structurées comme non-structurées), et ce au-delà de leurs sources ou de leur nature: données géo-spatiales, données statistiques, données météorologiques, données de recherche… etc.

La normalisation, pour l'expérience de l'Union européenne, est centrale. Car, elle a non seulement permis l'interopérabilité de systèmes de données «fermés et se suffisant à eux-mêmes», mais a permis aussi l'émergence d'un «espace de

données» unique et cohérent, nécessaire pour la compétitivité des entreprises et des économies européennes.

D'un autre côté, et s'agissant des techniques d'analyse des données massives, il faudrait préciser qu'elles sont fort tributaires des types d'analyse recherchés : descriptive, « diagnostique », prédictive, prescriptive ou cognitive (Cf. figure ci-après).

Figure n° 3 : Processus d'analyse des données de masse



Source : www.geeksforgeeks.org

+ L'analyse descriptive se focalise en économie des données, sur l'analyse des données historiques. Son objectif consiste à concevoir des modèles rétrospectifs de nature à mettre en perspective les tendances ou scénarios de la question posée.

Si les méthodologies relatives aux statistiques descriptives ont largement fait l'objet de travaux académiques ou de rapports des organisations internationales spécialisées, il n'en est pas toujours de même pour les méthodes d'analyse portant sur les données massives ou l'informatique en nuage (Agarwal, 2015).

+ L'analyse diagnostique, complémentaire de l'analyse descriptive par certains de ses aspects, vise à identifier les causes réelles ayant présidé à l'émergence d'un problème de quelque nature que ce soit (Pun cochar et Skach, 2018). Elles sont

orientées par l'unique et grande question : «Pour quelles raisons cela s'est-il produit ?».

+ Les méthodes fondées sur l'analyse prédictive procèdent par contre, par une démarche éminemment prospective. Elles se basent sur l'application des techniques de prévision et de modélisation mathématique et statistique, pour établir des scénarios d'évolution ou des perspectives d'avenir.

C'est par ailleurs le type d'analyses qui s'appuient sur les modèles d'apprentissage automatique en vigueur, selon qu'ils soient supervisés, semi-supervisés ou non supervisés.

+ Les méthodes d'analyse dite prescriptive permettent, quant à elles, de répondre à la question principale suivante : «Que devrait-on faire ?» (Lepenioti et al., 2020). Cette approche d'analyse est fondée dans sa démarche, sur des outils logiciels adoptant l'intelligence artificielle, les algorithmes d'optimisation et les systèmes experts.

Une telle démarche est pertinente, parce qu'elle prend en compte le contexte tout autant que les contraintes. Elle favorise par ailleurs, les prises de décision automatisées, le plus souvent optimales et temporelles.

+ L'analyse cognitive est, à la différence des autres méthodes, un concept nouveau, car récemment introduit dans le sillage de ce qui est appelé l'informatique cognitive (Cognitive Computing). Elle se base sur plusieurs facteurs dits de «motivation» : évolution des technologies de l'information, dispositifs et outils intelligents, capteurs et potentialités de l'informatique en nuage.

L'objectif principal de cette analyse est la conception et le développement de services fondés sur les technologies de l'intelligence artificielle, qui soient capables d'interagir avec les êtres humains, de décrypter le sens des différents contextes et de tirer des conclusions en analysant les soi-disant «antécédents» des utilisateurs (Gudivada et al., 2016, Gupta et al., 2018).

Toutes ces méthodes et approches d'analyse des données de masse renvoient directement ou indirectement, à des modèles d'affaires qui en portent la marque, et dont il faudrait dès présent, en définir les contours, la nature et les caractéristiques.

4. Modèles d'affaires des données massives : les modèles de l'innovation

Les organisations (entreprises en premier) sont le plus souvent objet de contraintes et de pressions de toute nature. Elles le sont encore davantage de nos jours, non seulement en raison des avantages compétitifs qui risqueraient de s'éroder rapidement, mais aussi en raison d'un contexte complexe, incertain et en mutation régulière.

De telles conditions astreignent à l'évidence, ces organisations à concevoir (ou à adopter) des solutions ou à définir des processus qui soient à même de relever les défis posés. Les solutions envisagées, tout autant que les processus adoptés, portent assez souvent la marque d'un mot clef : l'innovation.

Ce dont il est de plus en plus question autour des «nouveaux modèles économiques fondés sur l'innovation», s'inscrit dans la droite ligne de cette perspective. Les mutations de l'environnement externe, tout autant que l'émergence des nouvelles attentes des clients (Jensen et Sund, 2017), en sont le facteur accélérateur.

Les nouveaux modèles économiques ne sont pas toujours le prolongement des anciens modèles, encore moins leur reconduction élargie. Ce sont bien au contraire, ceux qui cherchent en permanence à expérimenter les méthodes nouvelles d'organisation et à concevoir les nouveaux services, générateurs de valeur et susceptibles de leur permettre maintien et/ou renforcement de leur position compétitive et concurrentielle sur le marché.

L'innovation paraît à ce niveau, comme l'une des grandes opportunités (des grandes portes d'entrée, diraient d'aucuns) pour faire émerger de nouveaux modèles d'affaires plus fédérateurs et plus intégrateurs de l'écosystème des données (Fjeldstad et Snow, 2018).

Les solutions d'innovation auxquelles on fait allusion, peuvent être introduites au niveau des produits (innovation de produits), comme au niveau des processus (innovation de procédés) et des systèmes d'organisation (innovation de process).

Les systèmes de liens ou de connexions, de même que les relations de causalité entre les différentes composantes du modèle d'affaires ainsi défini, devraient être suffisamment souples et flexibles, pour permettre d'apporter les modifications

nécessaires à la mise en place de l'innovation, tout autant que celle du modèle d'innovation s'y rapportant.

Un modèle d'affaires fondé sur l'innovation n'est pas toujours synonyme d'une quelconque mutation en profondeur, sauf s'il est porteur d'une innovation (dite de rupture) pouvant l'enclencher. Il n'est pas toujours non plus, synonyme d'un quelconque bouleversement de marché, sauf s'il s'agit là encore d'une innovation de produits ou de services foncièrement disruptive.

Les solutions dans les modèles d'affaires des organisations, aussi nouvelles et originales que celles-là soient-elles, n'aboutissent que rarement et par épisodes, à mutation, encore moins à bouleversement de marché. Car, dans un contexte incertain, les décisions des acteurs économiques en faveur de l'adoption de telle ou telle innovation, sont le plus souvent dictées par le souci d'augmenter sensiblement la valeur générée par le modèle d'innovation précédent, ou à défaut la maintenir à son niveau antérieur.

Dans l'un des cas, l'innovation est un simple facteur de stabilisation, dans l'autre, elle est une «barricade de défense» provisoire. Ce que ne traduisent que partiellement le rôle et la portée des modèles d'affaires fondés sur l'innovation.

C'est la raison pour laquelle, certains pensent que les modèles d'affaires n'ont pris réellement forme et consistance qu'à la faveur de l'émergence et du développement de l'Internet, de la croissance «nouvelle» enclenchée par les «marchés émergents», et de l'essor des industries et entreprises fort dépendantes des technologies post-industrielles (Sorrentino et Smarra, 2015). Le modèle d'affaires qui en découle, porte nécessairement la marque combinée de ces divers facteurs.

D'autres estiment que ledit modèle (tout autant que la structure le caractérisant) suppose l'association de trois facteurs d'ordre contingent: l'interaction avec les consommateurs, l'acquisition des ressources et la veille informationnelle. Le développement des connaissances, tout autant que la «robustesse» des plateformes relatives aux systèmes et technologies de l'information, en sont l'ossature de base (Venkatraman et Henderson, 1998).

L'on identifierait par ailleurs, les modèles d'affaires à l'ensemble de ses principales composantes, à savoir ses ressources stratégiques, ses réseaux de valeur, ses relations avec les clients et ses stratégies clés (Hamel, 2000).

L'on l'inscrit aussi dans une perspective plus large, pour dire simplement que les modèles d'affaires permettent de comprendre et de mettre l'accent sur le fonctionnement des organisations et des entreprises (Magretta, 2002).

A partir de là, l'on identifie les modèles d'affaires efficients comme étant ceux qui répondent de manière efficiente aux questions essentielles suivantes : qui sont les clients ? Comment accéder et investir dans un secteur d'activité ? Quels sont les principes et logiques économiques à même d'apporter de la valeur à des coûts appropriés ? (Amit et Zott, 2001).

Les modèles d'affaires comprennent de ce point de vue, les logiques permettant de déterminer les sources de création de la valeur et les possibilités offertes pour les maintenir (Schafer et al., 2005). Mais ils obéissent, d'un autre côté, à une logique de nature heuristique, celle qui tend à combiner les potentiels des technologies et les résultats obtenus en termes de valeur économique (Chesbrough et Rosenbloom, 2002).

Par ailleurs, ces «business models» renseignent sur les techniques que les organisations mettent en oeuvre pour créer, capturer et fournir de la «valeur» (Osterwalder et Pigneur, 2010). La valeur dont il s'agit ne profite pas uniquement aux entreprises l'ayant générées et rendues disponible, mais aussi à ceux qui s'en trouvent destinataires, c'est à dire les clients et les autres entreprises.

Les modèles en question décrivent ainsi les logiques, les facteurs et les éléments à l'origine de la création de la valeur et sa «médiatisation » par le marché. Ils décrivent, parallèlement à cela, les modes de répartition et de redistribution des revenus, tout autant que les coûts supportés par les entreprises y ayant présidé (Teece, 2010).

La valeur autour de laquelle s'articulent ces modèles contribue par conséquent, à une amélioration des relations entre les entreprises et leurs clients, mais aussi au renforcement de la compétitivité et de la position concurrentielle de ces dernières.

Force est de noter à ce propos, qu'il va de soi que les entreprises optent pour des modèles aux solutions de plus en plus innovantes. Les capacités des entreprises à générer de la valeur, dans un contexte mouvant et instable, en dépendent grandement.

Une telle perspective est d'autant plus prometteuse qu'elle prend appui sur une économie fondée sur les données de masse, dont l'extraction de la valeur conditionne et la chaîne de cette dernière et le modèle d'innovation lui-même.

Car, l'innovation dans les modèles d'affaires, est significative d'un changement considérable dans les logiques opérationnelles et stratégiques des entreprises (Teece, 2010), les solutions innovantes introduites dans lesdits modèles étant considérées comme des changements dans la sphère de création et de fourniture de la valeur.

Ce dont il s'agit concrètement, c'est cette tendance des modèles d'affaires à concevoir des mécanismes qui visent à créer le «besoin d'acquisition» des biens et services chez les utilisateurs, qui est en quelque sorte une acquisition de la valeur ou d'une partie de celle-ci (Baden-Fuller et Morgan 2010, Teece, 2010).

Cette innovation orientée/valeur fait de ces modèles d'affaires un puissant levier sinon de compétitivité, du moins d'adaptation raisonnable aux environnements de la concurrence et du marché (Johnson et al. 2008).

Rares, de nos jours, sont les acteurs économiques (dans les industries de réseaux notamment), qui ne conçoivent ou déclinent leur modèles d'affaires, dans le sens de la création et de la captation de la valeur à partir des processus d'innovation qu'ils initient (Casadesus-Masanell et Ricart, 2010).

Il convient de préciser ici, encore une fois, que l'innovation dans les modèles d'affaires pourrait concerner les entreprises en position de force sur le marché, comme elle pourrait être l'apanage des petites structures, quasi-absentes des marchés. L'exemple des start-ups dans le domaine des logiciels et des applications informatiques en est la grande démonstration.

Or, «l'activisme» des start-up sur certaines niches de marché n'est pas toujours déstabilisateur des grandes entreprises. Ces dernières, contrairement aux premières, ont les capacités suffisantes pour soit concevoir de nouveaux modèles, soit améliorer les anciens, soit combiner les deux, du moins jusqu'à ce qu'une solution innovante radicale ne vienne «chasser» celui en situation de déclin continu.

Tout le défi, pour les grandes structures notamment, consiste à opérer les choix nécessaires, à définir les horizons de temps et à paramétrer les coûts qui en résultent : quand, comment et à quel prix concevoir et développer un nouveau modèle d'affaires ? Pour combien de temps et à quel prix, se limiter à apporter les modifications et les améliorations à caractère incrémental des produits ? Dans

ce dernier cas, pour quel horizon de temps se limiter à de telles «retouches» ? (Friis-HolmEgfjord et Sund, 2020).

De tels questionnements n'excluent nullement ce que l'on appelle les «barrières cognitives», celles qui pourraient entraîner la pénétration imprévue de solutions inefficientes dans les activités et services en question. Une telle éventualité est d'autant plus sérieuse que les orientations d'innovation propres aux modèles d'affaires sont en grande partie, conditionnées par la perception des facteurs externes pouvant impacter les décisions des entreprises.

C'est la raison pour laquelle l'échec des solutions (ou activités) d'innovation, pourrait provenir d'une divergence de perception, voire d'interprétation des exigences qu'imposent les changements dans les modèles d'affaires.

Mais le changement pourrait être associé aux risques d'éventuelles inadéquations entre les modèles d'affaires en vigueur ou en gestation et les exigences de marché, dont sans nul doute les conditions de l'environnement.

De telles inadéquations pourraient être par ailleurs, à l'origine d'une éventuelle perte d'un ou de plusieurs segments de marché. Il est par conséquent nécessaire de déterminer quand et comment introduire et mettre en oeuvre lesdits changements, notamment quand l'environnement est incertain, instable et peu porteur. Car ce dernier n'est pas souvent un facteur d'impulsion pour l'innovation dans les modèles d'affaires (McDonald et al. 2019).

5. Conclusion

L'économie numérique est une économie des données par excellence. Elle est fondée sur la production, la génération, l'extraction, la collecte, l'exploitation, le traitement et l'analyse des données. Elle en fait un outil clef des processus de décision des organisations et de prédiction des tendances. Les données dont il s'agit, ce ne sont pas uniquement les données contenues dans les sources traditionnelles, publiques comme privées, encore moins celles contenues dans les archives.

Il s'agit de ces énormes quantités de données (dites Big Data), générées par les technologies nouvelles dont l'Internet des Objets, via ces centaines de millions de capteurs, les traces que laissent les utilisateurs des réseaux sociaux et des plateformes numériques, que les algorithmes de diverses puissances collectent et/ou traitent et que les technologies d'intelligence artificielle «profilent» pour en inférer des recommandations, des informations et du «sens».

Les modèles d'affaires fondés sur les données, tout autant que la grille et chaîne de valeur qui en découlent, portent profondément la marque et la caractéristique principales de cette économie toujours en gestation.

Ce sont des modèles d'affaires innovants et disruptifs. Ils caractérisent les modes d'organisation, tout autant que les nouvelles configurations des structures des différents marchés.

Ils sont conçus pour permettre aux entreprises de s'inscrire dans la droite ligne de la transformation numérique, horizon objectif central des évolutions en cours. De la seule observation de la nature et des caractéristiques desdits modèles dans une entreprise donnée, l'on pourrait aisément présumer de son positionnement dans l'échelle des processus de transformation numérique en vigueur.

Les plateformes sont les entreprises de l'économie numérique par excellence. Elles sont articulées autour de modèles d'affaires dont la chaîne de valeur est centrée sur les données. Elles en font un instrument de décision et de prédiction, mais aussi une marchandise qu'elles placent auprès des annonceurs. Les données pour elles, sont par conséquent, un «input» tout autant qu'un «output» Elles sont simultanément le produit et le facteur de production.

Références

-° Agarwal, B.L. (2015). «Basic Statistics», 6th edition. New Age International, New Delhi.

-° Amit, R., Zott, C. (2001). «Value creation in e-business». Strategic Management Journal, 22, pp. 493–520.

-° Assun¸c˜aoa, M.D. (2015). «Big data computing and clouds: trends and future directions». J. Parallel Distrib. Comput. 79–80, 3–15.

-° Baden-Fuller, C., Morgan, M.S. (2010). «Business models as models». Long Range Planning, 43 (2–3), pp. 156–171.

-° Berbakov, L., Tomasevic, N., Batic, M. (2018). «Architecture and implementation of IoT system for energy efficient living». In: 26th Telecommunications Forum, TELFOR 2018, Belgrade, Serbia.

-° Berbers, Y., Zwaenepoel, W. (eds.). (2019). «Google Dataset Search: Building a Search Engine for Datasets in an Open Web Ecosystem». ACM.

-° Berners-Lee, T. (2006). «Linked data». https://www.w3.org/DesignIssues/ LinkedData.html

-° Casadesus-Masanell, R., Ricart, J.E. (2010). «From strategy to business models and onto tactics». Long Range Planning, 43 (2–3), pp. 195–215.

-° Chen, H., Chiang, R.H., Storey, V.C. (2012). «Business intelligence and analytics: From big data to big impact». MIS Quarterly, 36(4).

-° Chesbrough, H., Rosenbloom, R.S. (2002). «The role of the business model in capturing value from innovation: Evidence from Xerox Corporation‘s technology spin-of companies». Industrial and Corporate Change, 11(3), pp. 529–555.

-° European Commission (2022), «Rolling Plan for ICT standardisation : Rolling Plan 2020», Brussels, February.

-° Fjeldstad, Ø., Snow, C.C. (2018). «Business models and organization design». Long Range Planning, 51(1), pp. 32–39.

-° Friis-Holm Egford, K., Sund, K.J. (2020). «Do you see what I see? How differing perceptions of the environment can hinder radical business model innovation». Technological Forecasting & Social Change, 150(119787), pp. 1–10.

-° Gudivada, V., Irfan, M., Fathi, E., Rao, D.L. (2016). «Cognitive analytics: going beyond big data analytics and machine learning», Chap. 5. Elsevier.

-° Gupta, S., Arpan Kumar Kar, A., Baabdullah, A., Al-Khowaiter, W. (2018). «Big data with cognitive computing: a review for the future». Int. J. Inf. Manage. 42, 78–89.

-° Hamel, G. (2000). «Leading the revolution». Boston: Harvard Business School Press.

-° Janev, V., Vrane˘s, S. (2005). «The role of knowledge management solutions in entreprise business processes». Journal of Universal Computer Science. 11(4), 526–546.

-° Janev, V., Vrane˘s, S. (2011). «Applicability assessment of semantic web technologies». Information Processing and Management. 47, 507–517.

-° Janev, V., Mijovic, V., Vranes, S. (2018). «Using the linked data approach in Europeane-government systems: example from Serbia». International Journal of Semantic Web Information System. 14(2), 27–46.

-° Jensen, H., Sund, K.J. (2017). «The journey of business model innovation in media agencies: Towards a three-stage process model». Journal of Media Business Studies, 14(4), pp. 282–298.

-° Johnson, M.W., Christensen, C.M., Kagermann, H. (2008). «Reinventing your business model». Harvard Business Review, 87 (December 12), pp. 51–59.

-° Laney, D. (2001). «3D data management: controlling data volume, velocity, and variety. Application Delivery Strategies». Meta Group.

-° Lepenioti, K., Bousdekis, A., Apostolou, D., Mentzas, G. (2020). «Prescriptive analytics: literature review and research challenges». International Journal of Information Management. 50, 57–70.

-° Liu C-H, Chen C-L. (2015). «A review of data monetization: strategic use of Big Data». In: The fifteenth international conference on electronic business (ICEB 2015). p. 7.

-° Magretta, J. (2002). «Why business models matter?» Harvard Business Review, 80(5), pp. 86–92.

-° Manyika, J. (2011). «Big data: the next frontier for innovation, competition, and productivity». The McKinsey Global Institute.

-° Marr, B. (2018). «Big data and AI: 30 amazing (and free) public data sources for 2018». https://www.forbes.com/sites/bernardmarr/2018/02/26/big-dataand-ai-30-amazing-and-free-public-data-sources-for-2018

-° McDonald, R.E., Masselli, J.J., Chanda, B. (2019). «Nonproft business model innovation as a response to existential environmental threats: Performing arts in the United States». Journal of Business Research. Doi.org/10.1016/j. jbusres.2019.12.022.

-° Mijovic, V., Toma˘sevi´c, N., Janev, V., Stanojevi´c, M., Vrane˘s, S. (2019). «Emergency management in critical infrastructures: a complex-event-processing paradigm». Journal of Systems Science and Systems Engineering. 28, 37–62 (2019). https://doi.org/10.1007/s11518-018-5393-5

-° Osterwalder, A., Pigneur, Y. (2010). «Business model canvas». Self-Published. Last. Parker, G.G., Van Alstyne, M.W., Choudary, S.P. (2017). Plataforma: A revolução da estratégia. São Paulo: Casa Educação.

-° Puncochar, I., ˘Skach, J. (2018). «A survey of active fault diagnosis methods». IFACPapersOnLine 51(24), 1091–1098.

-° Sorrentino, M., Smarra, M. (2015). «The term —Business Model‖ in financial reporting: Does it need a proper definition?» Open Journal of Accounting, 4, pp. 11–22.

-° Sun, C., Gao, R., Xi, H. (2014). «Big data based retail recommender system of non e-commerce». In: Fifth International Conference on Computing, Communications and Networking Technologies (ICCCNT), pp. 1–7.

-° Teece, D.J. (2010). «Business models, business strategy and innovation ». Long Range Planning, 43, pp. 172–194.

-° Upadhyaya, S.R. (2013). «Parallel approaches to machine learning - a comprehensive survey». Journal of Parallel and Distributer Computing. 73, 284–292.

-° Venkatraman, N., Henderson, J.C. (1998). «Real strategies for virtual organizing». Sloan Management Review, 40(1), pp. 33–48.

-° Wang, L. (2017). «Heterogeneous data and big data analytics». Automatic Control and Computer Sciences. 3(1), 8–15.

-° Zaveri, A., Rula, A., Maurino, R., Pietrobon, R., Lehmann, J., Auer, S. (2016). «Quality assessment for linked data: a survey». Semantic Web- Interoperability, Usability, Applicability. 7(1), 63–93.

Webographie

www.themetisfiles.com

www.attassian.com

www.geeksforgeeks.org