

Transformation Digitale : les nouvelles compétences à l'épreuve des Big data

Digital Transformation: new skills to test the big data

Ghanouane Karim¹, Benkaraache Taoufik²

*1*Doctorant, laboratoire de recherche en intelligence Stratégique, Université Hassan 2 Faculté des sciences juridiques, économiques et sociales Mohammedia – Maroc, ghanouanekarim@gmail.com.

*2*Professeur encadrant, Directeur du laboratoire de recherche en intelligence Stratégique, Université Hassan 2 Faculté des sciences juridiques, économiques et sociales Mohammedia – Maroc, benkaraache@gmail.com.

RESUME

Le passage de l'ère industrielle à l'ère des produits intelligents et connectés a permis aux décideurs de disposer de volumes importants de données. Les Big data font référence à un groupement de données non seulement volumineux, mais également riche en variété et en vitesse, ce qui les rends difficiles à manipuler à l'aide des systèmes de gestion des bases de données traditionnels. Bien que de nombreuses études sur les techniques d'analyse des Big data existent, l'étude des compétences requises pour gérer ces données volumineuses demeure à un état embryonnaire.

Par la combinaison des techniques d'extraction des données « Web Scraping » et la méthodologie de l'analyse de contenu ; sur des moteurs de recherche d'offres d'emploi au Maroc ; nous développons une taxonomie des compétences pour le Big data. Ce papier a pour finalité de guider les praticiens et la sphère académique au Maroc dans l'évaluation et la progression des compétences en Big Data.

Mots clés :

Science des données, Transformation Digitale, Compétences, Analyse de contenu, Web Scraping.

ABSTRACT

The transition from the industrial era to the era of intelligent and connected products has provided decision-makers with large volumes of data. Big data refers to a grouping of data that is not only voluminous but also rich in variety and speed, making it difficult to manipulate using traditional database management systems. Although many studies on big data analysis techniques exist, the study of the skills required to manage this large data set is still in its infancy.

By combining "Web Scraping" data extraction techniques and content analysis methodology; on job search engines in Morocco; we are developing taxonomy of skills for big data. This paper aims to guide practitioners and the academic sphere in Morocco in the evaluation and progression of Big Data skills.

Key words:

Data Science, Digital Transformation, Skills, Content Analysis, Web Scraping.

INTRODUCTION

La collecte, le traitement, l'analyse et le stockage de l'information est un processus qui ne se limite pas à définir des idées mais surtout à les concrétiser avec pour finalité, une meilleure production de connaissances qui engendrent des innovations. Ce processus permet à l'entreprise d'optimiser son offre vis-à-vis du marché, quantitativement et qualitativement, ainsi qu'optimiser ses procédés de production (Monino, 2015).

L'introduction et la vulgarisation de nouvelles technologies digitales, telles que les réseaux sociaux, les technologies mobiles, les Big Data, l'internet des objets et le Cloud Computing, ont poussé les entreprises de presque tous les secteurs à mener de nombreuses initiatives pour explorer et exploiter les avantages de telles technologies (Fitzgerald et al., 2013 ; Ross et al., 2016). Cela passe souvent par des transformations externes centrées client et des transformations internes centrées sur les produits et les processus, voir même les modèles d'affaires (Matt et al., 2015). Ainsi, l'entreprise dans son ensemble fait face à un changement rapide et radical dû à la maturation des technologies digitales et à leur pénétration omniprésente sur tous les marchés (Ebert et al., 2016). Les big data font référence à un groupement de données non seulement volumineux, mais également riche en variété et en vitesse, ce qui les rends difficiles à manipuler à l'aide des systèmes de gestion des bases de données traditionnels. Cette nouvelle ère est en train de bouleverser les pratiques classiques du management, l'entreprise moderne est contrainte d'adopter de nouvelles visions afin de disposer de « information stratégique » (Martinet. B et Marti. Y.M, 2001).

Bien que de nombreuses études sur les techniques d'analyse des big data existent, l'étude des compétences requises pour gérer ces données volumineuses demeure à un état embryonnaire.

Dans ce contexte, notre problématique sera comme suit :

Quelles compétences les entreprises ont besoin à l'épreuve des Big data ?

Par la combinaison des techniques d'extraction des données « Web Scraping » et la méthodologie de l'analyse de contenu ; sur des moteurs de recherche d'offres d'emploi au Maroc ; nous développons une taxonomie des compétences pour les big data. Ce papier a pour finalité de guider les praticiens et la sphère académique au Maroc dans l'évaluation et la progression des compétences en Big Data.

La suite du papier se déroule comme suit. La deuxième section présente une revue de la littérature des concepts de transformation digitale, des Big data et des nouvelles compétences qu'exige ces données volumineuses. Dans la section suivante, nous présenterons notre démarche méthodologique. Ensuite, une analyse des résultats des données empiriques compte tenu de la littérature. Les conclusions de la revue de la littérature et données empiriques sont tirées dans la dernière section du papier.

I- BIG DATA ET COMPETENCES HUMAINES : OUVRIR LA VOIE VERS LA TRANSFORMATION DIGITALE

1. Conception de la transformation digitale

Reis et al (2018) identifient les aspects de la transformation digitale en trois classes distinctes :

Un aspect technologique où la transformation digitale repose sur l'utilisation de nouvelles technologies digitales telles que les réseaux sociaux, la technologie mobile, les outils analytiques ou intégrés (Fitzgerald et al., 2013).

Un aspect organisationnel où la transformation digitale nécessite un changement des processus opérationnels ou la création de nouveaux modèles d'affaire (Ross et al., 2016).

Un dernier aspect social où la transformation digitale est un phénomène qui influence tous les aspects de la vie humaine (Matt et al., 2015).

Ces différents aspects sont repris par quasiment tous les chercheurs [voir tableau 1] dans la conception de la transformation digitale. Reis et al (2018) synthétisent tous ces aspects et proposent que la transformation est l'utilisation de nouvelles technologies digitales qui permettent des améliorations majeures de l'activité de l'entreprise et influence tous les aspects de la vie des clients.

Tableau 1 : Evolution du concept de la transformation digitale¹

Auteur (s)	Aspect (s)
Collin et al. (2015); Gimpel et Röglinger (2015); Kane et al. (2015)	Alors que la digitisation décrit généralement la simple conversion d'informations analogiques en informations digitales, les termes Transformation digitale et digitalisation sont utilisés de manière interchangeable et font référence à un concept large affectant les questions politiques, commerciales et sociales.
Solis et al. (2014)	La réorientation de ; ou nouveaux investissements dans ; la technologie et des modèles d'affaire pour engager plus efficacement les clients numériques à

¹ Source: Reis, João & Amorim, Marlene & Melao, Nuno & Matos, Patrícia. (2018). *"Digital Transformation: A Literature Review and Guidelines for Future Research"*. Springer International Publishing AG 2018.

	chaque point de contact du cycle de vie de l'expérience client.
Fitzgerald et al. (2013); McDonald et Rowsell-Jones (2012)	Utilisation de nouvelles technologies digitales, telles que les réseaux sociaux, les technologies mobiles, outils analytiques ou intégrés, afin de permettre des améliorations majeurs des activités de l'entreprise telles que l'amélioration de l'expérience client, la rationalisation des opérations ou la création de nouveaux modèles d'affaire (2013). En tant que telle, la transformation digitale va au-delà de la simple numérisation des ressources et génère de la valeur et des revenus à partir d'actifs digitales (2013).
Westerman et al. (2011)	La transformation digitale est définie comme l'utilisation de la technologie pour améliorer radicalement les performances ou la portée des entreprises.
Martin (2008)	La transformation digitale est maintenant communément interprétée comme un tel usage des technologies de l'information et de la communication, lorsque l'on n'effectue pas une automatisation triviale, mais que de nouvelles capacités sont créées dans les entreprises, les administrations publiques, ainsi que dans la vie des personnes et de la société.
Stolterman et Fors (2004)	La transformation digitale correspond aux changements que la technologie digitale provoque ou influence dans tous les aspects de la vie humaine.

2. Les leviers de la transformation digitale

Les technologies facilitatrices de la transformation digitale comprennent la technologie mobile, les réseaux sociaux, le Cloud computing, les Big Data et l'internet des objets. Celles-ci sont explorées ci-dessous.

2.1. La Technologie mobile

Dans leur article, Reddi et Zhu (2017) avancent qu'en 2015, il existe plus de 3 milliards d'appareils mobiles connectés au Web. D'ici 2020, il y aura 50 milliards d'appareils de ce

type. Cette croissance ne concerne pas seulement le nombre d'appareils mais également le volume de trafic internet mobile, qui dépasse souvent le trafic traditionnel.

La technologie mobile (Bucki, 2016) se définit comme étant «un terme générique utilisé pour désigner une variété de dispositifs permettant aux utilisateurs d'accéder à des données et à des informations d'où qu'ils se trouvent. Parfois appelée "interaction homme-machine", la technologie mobile transporte des données, de l'audio et de la vidéo sur un réseau via un appareil mobile.»

La croissance du marché de l'informatique mobile permet l'accès à un large éventail de technologies et l'intégration de divers environnements, tels que le Cloud computing, les réseaux sociaux, les Big Data et l'analytique (Ferreira et al., 2017).

2.2. Les réseaux sociaux

Les réseaux sociaux englobent tous les médias, technologies et applications digitales qui permettent aux utilisateurs d'Internet d'échanger des opinions, des expériences, des informations et des connaissances (Schlesinger, 2014). A la différence des réseaux traditionnels, Les réseaux sociaux englobent tous les médias, technologies et applications digitales qui permettent aux utilisateurs d'Internet d'échanger des opinions, des expériences, des informations et des connaissances (Schlesinger, 2014).. Selon Statista en 2017 ; en termes d'usage de services offerts ; Facebook est la plus grande plate-forme de réseaux sociaux avec plus de 2 milliards d'utilisateurs, suivie de YouTube avec 1,5 milliard d'utilisateurs et de WhatsApp et Facebook Messenger avec 1,3 milliard d'utilisateurs.

Pour Kaschny et Nolden (2018), un nombre important et croissant d'utilisateurs s'organisent au sein de réseaux sociaux et échangent des informations sur des produits, des services et des entreprises. L'époque où les réseaux sociaux avaient un caractère purement privé est révolue. De plus en plus d'entreprises adhèrent aux réseaux et les utilisent généralement pour la publicité et le recrutement. Pour ces entreprises, les réseaux sociaux permettent de se connecter quotidiennement avec les clients actuels et potentiels. Ces réseaux jouent un rôle clé dans l'identification des sujets en cours de discussion et des opinions qui circulent. Il y a aussi la possibilité de participer activement aux discussions et de se positionner. En conséquence, l'entreprise est en contact direct avec le groupe cible, ce qui renforce la perception du client et peut avoir des effets positifs sur le développement.

2.3. Le Cloud Computing

Pour Armbrust et al (2010) « le Cloud Computing désigne à la fois les applications livrées sous forme de services sur Internet ainsi que le matériel et les logiciels système au sein des centres de données qui fournissent ces services. Les services eux-mêmes ont longtemps été appelés logiciel en tant que service (SaaS). Certains fournisseurs utilisent des termes tels que l'IaaS (infrastructure en tant que service) et PaaS (plate-forme en tant que service) pour décrire leurs produits. »

Dans son article, Schwertner (2018) définit le Cloud Computing comme étant un modèle permettant l'accès à la demande à un pool partagé de ressources informatiques configurables (réseaux, serveurs, stockage, applications et services) pouvant être rapidement mis en service et libéré avec un minimum d'effort de la part du fournisseur de services. Ce modèle se compose de cinq caractéristiques essentielles (libre-service à la demande, accès réseau étendu, mise en commun des ressources, élasticité rapide, service mesuré); trois modèles de services (logiciel en tant que service (SaaS), plate-forme en tant que service (PaaS), infrastructure en tant que service (IaaS)); et quatre modèles de déploiement (Cloud privé, Cloud communautaire, Cloud public, Cloud hybride).

Les avantages du Cloud Computing sont généralement la réduction des dépenses, du personnel technique et des efforts. Autrement, ces avantages peuvent être scindés comme suit (Schwertner, 2018) :

Coûts prévisibles : Les services du Cloud sont généralement payés sur une base mensuelle ou sur la base d'une utilisation avec des coûts initiaux minimes ou nuls

Coût total réduit : Réaliser des économies d'échelle suite aux services qu'un fournisseur peut réaliser.

L'accès à la meilleure technologie : Les services du Cloud Computing permettent aux entreprises de bénéficier des meilleures technologies.

Facturation selon l'usage : Dans le modèle des services du Cloud Computing, les entreprises paient pour une utilisation réelle plutôt que maximale.

2.4. Internet des objets

Miorandi et al. (2012) définissent l'Internet des objets en étant : « terme habituellement utilisé pour désigner la connexion d'objets physiques usuels équipés de capteurs et capables de mesurer diverses caractéristiques de leur environnement et de partager ces mesures via un réseau de capteurs sans fil. Concrètement, de tels objets, également appelés «objets intelligents» ou «agents IoT», doivent être identifiables de manière unique, doivent pouvoir effectuer des calculs simples à complexes, pouvoir communiquer et requièrent, bien entendu, des fonctions capables de détecter et de mesurer les paramètres de leur environnement »

L'Internet des objets représente une évolution du concept Machine-to-Machine (M2M) grâce à la coordination des machines, appareils et applications de plusieurs fournisseurs connectés à Internet via plusieurs réseaux. Ces dispositifs intelligents combinés avec d'autres technologies telles que l'intelligence artificielle, l'apprentissage automatique, le Cloud computing modifient radicalement le monde en développant de nouveaux services intelligents (Miorandi et al., 2012).

Si Bucherer et Uckelmann (2011) affirment que l'Internet des objets fournit aux entreprises un moyen de transformer progressivement leurs opérations commerciales. Roussos et

Kostakos (2009) quand à eux limitent l'impact de l'Internet des objets dans un premier lieu aux processus opérationnels des entreprises.

2.4. Les Big Data

Les Big Data représentent des groupements de données dont la taille dépasse la capacité; à capturer, stocker, gérer et analyser; des systèmes de gestion des bases de données classiques. Ce phénomène des Big Data varie selon les secteurs (Benkarrache et Ghanouane, 2020 ; Manyika et al., 2011). Pour laney (2001) faire face aux Big Data revient à assimiler ses trois dimensions (3V) à savoir : le volume, la vélocité et la variété.

Le volume concerne l'énorme quantité de données générées. La vélocité désigne la rapidité nécessaire pour traiter ces données en temps voulu. Enfin, la variété peut être comprise comme une mesure de l'hétérogénéité des données (Benkarrache et Ghanouane, 2020 ; Gandomi et Haider, 2015). D'autres V peuvent s'ajouter aux précédentes dimensions comme le soulève Mc Nulty (2014) à savoir :

La variabilité c'est-à-dire des données dont la signification est en constante évolution ;

La véracité indique le résultat de l'usage des données, où les données sont pratiquement sans valeur si elles ne sont pas exactes ;

Et enfin, la visualisation qui est l'un des défis de « Big data », se trouve dans la façon de présenter les résultats de traitement des données (Information) qui rend les conclusions plus claires.

Le concept de transformation digitale est la réingénierie de tous les processus de production et de vente, afin qu'ils deviennent plus efficaces, bien que cela n'entraîne pas nécessairement la modification des services et produits existants, ni la création automatique de nouveaux. La transformation digitale couvre bien plus que la simple mise en œuvre des nouvelles technologies (cloud computing, IoT, Big data, technologies mobiles) (Petkovics, 2018). **Si les collaborateurs ne soutiennent pas les approches suggérées, ces technologies digitales recommandées pour atteindre l'objectif de la digitalisation pourraient même être contre-productives** (Petkovics, 2018).

3. Ressources humaines pour les métiers du Big Data : Une classification systématique des compétences requises

Pour Davenport (2013) et Granville (2014), donner du sens aux données et leur transformation en connaissances ont longtemps été sujet de débat entre scientifiques, statisticiens, informaticiens et autres disciplines. Avec l'avènement des Big Data, le titre de scientifique de données « Data scientists » a été utilisé pour qualifier les professionnels qui produisent des connaissances en analysant les données.

Au début, le titre de scientifique de données « Data scientists » était courant dans des structures telles que Google, eBay, LinkedIn et Facebook, dans lesquelles les big data analytiques étaient le seul centre d'analyse. (Davenport 2013). Cependant dans de nombreuses autres entreprises, les Big Data doivent s'intégrer à des sources de données traditionnelles en tenant compte de ces dernières dans le processus d'analyse. Ces entreprises ont recours à d'autres compétences en analyse de données, tels que les statisticiens, les ingénieurs logiciels, les développeurs, les analystes marchés, les architectes des Big data, les ingénieurs de données, etc. (Hammerbacher 2009 ; Davenport et Dyche 2013 ; Granville 2014 ; Miller 2014).

Les scientifiques de données sont généralement formés en informatique, en mathématiques ou en économie, mais ils peuvent émerger de tout domaine qui est fortement axé sur les données et l'informatique (Patil 2011 ; Davenport et Patil 2012). Harris, Murphy et Vaisman (2013) ajoutent à la liste des domaines plus spécifiques tels que les statistiques, l'apprentissage automatique, les bases de données, la recherche opérationnelle, la veille stratégique et les sciences sociales et physiques.

La littérature nous permet de distinguer trois catégories de compétences liées aux Big Data :

3.1. Les compétences techniques

Au niveau des compétences techniques la compétence la plus fondamentale et la plus universelle du scientifique des données est le codage afin de pouvoir développer des prototypes (Patil, 2011). Au même registre le besoin de plus en plus important des entreprises ; à mettre en place des systèmes de décision automatisés et prédictifs pour comprendre les Big Data ; rend l'apprentissage automatique « Machine learning » une compétence de base. Dhar (2013) liste les compétences fondamentales en plus de l'apprentissage automatique dont un scientifique des données a besoin pour bien mener ses missions outre :

- Les statistiques bayésiennes, qui nécessitent une connaissance pratique de la probabilité, des distributions, des tests d'hypothèses et de l'analyse multi-variée ;
- La connaissance des structures de données, des algorithmes et des systèmes, y compris l'informatique distribuée, les bases de données, l'informatique parallèle et l'informatique tolérante aux pannes ;
- Les langages de script (par exemple, Python et Perl) ;
- La connaissance de la corrélation et de la causalité ;
- Les techniques et Les outils issus de l'informatique, de la linguistique, de l'économétrie, de la sociologie et d'autres disciplines.

Miller (2014) souligne que les scientifiques de données doivent maîtriser les mathématiques et les statistiques, l'apprentissage automatique, l'analyse prédictive, l'informatique et la programmation, ainsi que des compétences de base en systèmes d'information, bases de données, entreposage de données « Data Warehouse » et fouille de données « Data Mining » .

Au sein d'un marché de plus en plus concurrentiel, les entreprises s'arrachent les profils spécialisés dans l'apprentissage automatique capable d'inventer des algorithmes innovants (Miller, 2014).

L'étude menée par Pinola (2015) sur la base des offres d'emploi a révélé que les compétences les plus fréquemment mentionnées sont les statistiques, l'apprentissage automatique, la programmation, l'exploration de données, les algorithmes et l'analyse prédictive.

3.2. Les compétences professionnelles

Un scientifique de données doit disposer de la capacité de formuler des problèmes de manière à aboutir à des solutions efficaces (Dhar, 2013). Wing (2006) et Dhar (2013) utilisent le terme de La pensée computationnelle pour décrire les compétences de base dont les scientifiques de données auront besoin pour faire face à ce tsunami informationnel. Ces compétences ont fait l'objet d'étude de la part de Barr et Stephenson (2011). Ils ont recensé les capacités suivantes : collecte de données, analyse de données, visualisation et analyse de données, abstraction, analyse et validation de modèles, automatisation, tests et vérifications, algorithmes et procédures, décomposition de problèmes, structures de contrôle, parallélisation et simulation.

Au-delà des compétences analytiques, les scientifiques des données doivent avoir une bonne compréhension du monde des affaires et de la gestion, une expertise en matière d'algorithmes d'exploration de données et être en mesure de cerner les problèmes de l'entreprise d'un point de vue données (Dhar 2013 ; Waller et Fawcett, 2013 ; Provost et Fawcett, 2013). Pour Davenport et Patil(2012), les scientifiques de données doivent amener les décideurs à un dialogue orienté données, ils doivent de même communiquer de manière compréhensible pour tous les acteurs, verbalement et visuellement. Les scientifiques de données doivent être en mesure de rendre l'analyse possible en associant des données à d'autres éléments et à des sources de données incomplètes. Ils nettoient également le jeu de données résultant. (Patil 2011; Davenport et Patil 2012.)

Toutefois, les compétences spécifiques aux Big Data doivent être combinées aux compétences traditionnelles en matière d'administration des bases données à savoir : les architectures de données, les métadonnées, la qualité des données et les processus de correction, les tableaux de bord d'administration et de gestion des données, master data management, des algorithmes de correspondance et de nombreux autres sujets spécifiques aux données (Davenport et Dyché, 2013).

Du fait que les modèles analytiques sont souvent intégrés aux processus opérationnels, Davenport et Dyché (2013) évoque la nécessité de maîtriser la gestion du changement afin d'apporter les adaptations nécessaires aux niveaux des rôles, la conception des processus et les compétences.

3.3. Les compétences personnelles

Les compétences personnelles qui caractérisent le scientifique de données dans tous les domaines sont la curiosité et la pensée associative. Ces caractéristiques sont nécessaires pour déceler les questions et les hypothèses que les scientifiques de données tentent de résoudre. Ces derniers doivent faire preuve de créativité pour pouvoir aborder les problèmes de manière innovante. (Patil 2011 ; Davenport et Patil 2012.)

Dans leur article, Davenport et Patil (2012) se focalisent sur la formation et la curiosité comme facteur clé pour tout scientifique de données afin de faire des découvertes dans le monde des Big Data. De même, les analystes de données ont également besoin de compétences analytiques, d'autonomie et doivent également être sur la même longueur d'onde avec leurs collègues gestionnaires en temps réel pour répondre aux problèmes de gestion. Pour faire face aux limites techniques, les scientifiques de données doivent être créatifs, déterminés et tolérants à la pression lorsqu'on attend beaucoup d'eux. (Davenport et Patil 2012.)

Toutefois, il faut signaler que la littérature fournit plusieurs descriptions concordantes des caractéristiques [voir tableau 2]² que les scientifiques de données devraient afficher.

Tableau 2 : liste non exhaustive des nombreux «traits» qu'un scientifique des données

Traits du scientifique de données	Source
Expert en outils Big Data	(Miller, 2014 ; Song et Zhu, 2015 ; Provost et Fawcett, 2013)
Codeur	(Davenport et Patil, 2012)
Statisticien / Analyste quantitatif	(Davenport, 2014, Provost et Fawcett, 2013)
Chercheur	(Davenport, 2014)
Data Hacker	(Davenport et Patil, 2012 ; Conway, 2010)
Auditeur	(Mayer-Schönberger et Cukier, 2013)
Responsable éthique des données	(Miller, 2014)
Gestionnaire de données et stratège	(Miller, 2014, Song et Zhu, 2015, Wixom et al., 2014)

² Source : Andrea De Mauro, Marco Greco, Michele Grimaldi, Paavo Ritala (2018). " *Ressources humaines pour les professions Big Data: une classification systématique des rôles et des compétences requises*", Journal Information Processing and Management.

Expert en visualisation	(Provost et Fawcett, 2013, Davenport et Patil, 2012)
Communicateur	(Wixom et al., 2014, Chen et al., 2012, Song et Zhu, 2015, Davenport et Patil, 2012)
Gestionnaire de projet	(Song et Zhu, 2015)
Expert / conseiller en affaires	(Chen et al., 2012, Davenport, 2014, McAfee et Brynjolfsson, 2012)

II- L'ENTREPRISE MAROCAINE ET LES COMPETENCES EN BIG DATA : UNE ETUDE BASEE SUR LES ANNONCES D'EMPLOI

1. Méthodologie : une double approche Web Scraping - Analyse du contenu

Cette étude a été conçue selon une méthodologie en deux étapes :

Premièrement, nous avons téléchargé un nombre considérable d'offres d'emploi comme données d'étude. Ces offres correspondent aux attentes des entreprises par rapport aux profils qu'elles recherchent. Cet article porte sur les compétences en matière des Big data, le mot clé « big data » a été utilisé pour identifier les offres objet de notre étude. Cette première étape a nécessité le recours à la technique d'extraction du contenu de sites Web « Web scraping » via des programmes gratuits (octoparse, parsehub, web scraper). La technique d'extraction du contenu consiste à rechercher des éléments de données spécifiques présentant un intérêt dans une série de pages Web semi-structurées, à les extraire à l'aide de robots d'exploration et à les stocker dans des ensembles de données plus structurés (Vargiu et Urru, 2012).

La collecte des données a porté sur les offres postées sur des sites de recrutement marocains durant 2019. Après avoir supprimé toutes les données en double et incomplètes, il nous restait un ensemble de données de 80 postes, que nous avons utilisé comme données pour notre analyse.

La deuxième étape a nécessité le recours à la méthodologie de l'analyse de contenu. Pour Berelson (1952) : « L'analyse de contenu est une technique de recherche pour la description objective, systématique, et quantitative du contenu manifeste de la communication »

L'objet central de toute analyse de contenu est le codage et la catégorisation. La méthodologie consiste à classer en catégories, les différentes positions ou attitudes que reflètent les réponses, pour permettre une présentation quantifiée des résultats (Mucchielli, 1984).

2. Résultats : un examen des compétences et des familles d'emplois dans le domaine des données

Les résultats ci-dessous, représentent une description de chaque catégorie tenant compte de la littérature.

La majorité des offres d'emploi requièrent une formation de base orientée ingénieur ou master. Les formations doctorales restent non significatives dans les données empiriques. Sur la base de ces résultats, nous pouvons conclure que plus la formation académique est élevée plus les chances d'insertion professionnelle sont élevées. En outre, les entreprises marocaines sont exigeantes en termes d'expérience. En général, les durées varient entre 2 ans et 5 ans.

Les résultats montrent que 77 % des annonces sont orientées vers les activités de conseil en systèmes d'information, le secteur financier occupe une deuxième position avec 16 %.

2.1. Compétences techniques

Un totale de 180 occurrences a été identifié dans la catégorie des compétences techniques dont la majorité se réfère au processus Big Data allant de collecte des données jusqu'à la visualisation, relevant ainsi une large similitude entre la littérature et les offres d'emploi au niveau marocain. De même, dans leur quête des scientifiques des données, les entreprises marocaines recherchent des profils combinant les nouvelles compétences de l'univers des Big data mais aussi une maîtrise des systèmes d'information décisionnels traditionnels [voir tableau 3].

Une comparaison entre la littérature et les offres d'emploi des entreprises marocaines pour certaines compétences techniques montre qu'en pratique la recherche des entreprises est plus détaillée comme par exemple les compétences en termes des statistiques, maths et algorithmes [voir tableau 3].

Tableau 3 : Compétences techniques

Compétences techniques	Sources	Occurrences
Compétences de programmation	(Patil 2011; Davenport et Patil 2012; Davenport et Dyché 2013; Miller 2014)	14
Apprentissage automatique	(Conway 2010; Dhar 2013; Harris, Murphy et Vaisman 2013; Miller 2014)	16
Statistiques	(Conway 2010, Dhar 2013; Miller 2014)	11
Maths, algorithmes	(Dhar 2013; Miller 2014)	8
Intégration des données traditionnelles et Big data, Analyse en temps réel et visualisation	(Davenport et Dyché 2013; Davenport 2013)	79
Analyses prédictives	(Dhar 2013; Miller 2014)	3

Analyse de réseau	(Dhar 2013; Davenport 2013)	3
Cloud computing	(Pinola, 2015)	16
Système d'exploitation	(Pinola, 2015)	13
Construire des systèmes de recommandation	(Pinola, 2015)	1
Big data-systèmes	(Pinola, 2015)	16

Sources : auteurs

2.2. Outils et technologies

La majorité des offres listées requière des compétences en langages de programmation (Python, R, Java, Spark, C++, NoSQL). De même, La plateforme ; de traitement des Big Data ; Hadoop apparait plus que les systèmes de gestion de base des données traditionnels au niveau des annonces d'emploi [voir tableau 4].

Tableau 4 : Outils et technologies

Outils et technologies	Sources	Occurrences
Hadoop	(Conwey 2011; McAfee et al. 2012; Davenport 2013; Davenport et Dyche 2013)	31
R	(Davenport et Dyche 2013)	16
Python	(Davenport et Dyche 2013)	25
SQL et NoSQL	(Davenport et Dyche 2013)	23
Hive	(Davenport et Dyche 2013)	9
Excel	(Kiron et al. 2012)	1
SAS	(Davenport et Dyche 2013)	4
IBM DB2	(Davenport et Dyche 2013)	1
Java	(Pinola, 2015)	18
C++	(Pinola, 2015)	3
Spark	(Pinola, 2015)	26

Sources : auteurs

2.3. Compétences fonctionnelles

Les offres d'emploi listées au niveau des compétences fonctionnelles incluent les savoirs faire techniques. Les compétences analytiques et pensée computationnelles, conformément à la littérature, constituent une exigence pour les entreprises marocaines. Autres compétences, les

scientifiques de données aussi bien au niveau de la littérature que les profils demandés doivent disposer de compétences en communication, en gestion ; en travail en équipe et surtout en Reporting. Cependant, contrairement aux auteurs scientifiques, la majorité des entreprises marocaines ne sont pas exigeantes en termes de connaissance du monde des affaires ou des secteurs d'activité, durant notre traitement des données, nous n'avons remarqué aucune exigence des compétences en conduite de changement ou de gestion stratégique des données [voir tableau 5].

Enfin, l'étude des offres d'emploi des entreprises marocaines a permis de déceler une compétence non mentionnée par les auteurs, celle de la veille technologique [voir tableau 5].

Tableau 5 : Compétences fonctionnelles

Compétences fonctionnelles	Sources	Occurrences
Connaissance du sujet / compréhension de l'entreprise	(Davenport et Patil 2012; Chen ym.2012; Conway 2011; Dhar 2013; van den Driest 2016; Provost et Fawcett 2013; Waller et Fawcett 2013)	22
Compétences analytiques, Pensée computationnelle	(Wing 2006; Barr et Stephenson 2011; Davenport et Patil 2012; Davenport et Dyche 2013; Dhar 2013; Provost et Fawcett 2013; Waller et Fawcett 2013)	67
Compréhension de la gestion, gestion de la décision	(Waller et Fawcett 2013; Provost et Fawcett 2013; Miller 2014)	43
Habiletés de rapport, compétences de documentation	(Davenport et Patil 2012; Davenport et Dyche 2013; Driest et al 2016)	33
Compétences en communication	(Davenport et Patil 2012; Chen ym.2012)	32
Gestion stratégique des données	(Miller 2014)	0
Méthodes agiles	(Patil 2011)	10
La gestion du changement	(Davenport et Dyche 2013)	0
Planification stratégique	(Pinola, 2015)	11
Compétences de travail en équipe	(Pinola, 2015)	26
Veille technologique		13

Sources : auteurs

2.4. Compétences personnelles

Pour Davenport et Patil (2012), le scientifique des données est une personne créative, proactive, innovante, passionnée par son métier, souple dans sa prise de contact. De même, ces profils doivent être rigoureux et déterminés à résoudre les problèmes. En outre, l'habilité d'être autonome est importante (Pinola, 2015). Toutes ces compétences recensées dans la littérature ont été observées au niveau des offres d'emploi des entreprises marocaines [voir tableau 6].

Tableau 6 : Compétences personnelles

compétences personnelles	Sources	Occurrences
Passionné de résolution de problèmes, Compétences en résolution de problèmes	(Davenport et Patil 2012; Dhar 2013; Patil 2011)	5
Créatif, innovant	(Davenport et Patil 2012; Van den Driest 2016; Patil 2011)	19
Bon esprit d'équipe, fortes compétences sociales	(Davenport et Patil 2012; Davenport 2012)	15
Tolérance à la pression	(Davenport et Patil 2012)	15
Prise de conscience en temps réel	(Davenport et Patil 2012)	3
Curieux	(Davenport et Patil 2012)	5
Proactif	(Davenport et Patil 2012)	17
Productif	(Davenport et Patil 2012)	25
Capacité d'apprendre	(Davenport et Patil 2012)	6
Déterminé	(Davenport et Patil 2012)	
Capable de travailler de façon autonome	(Pinola, 2015)	21

Sources : auteurs

CONCLUSION

Pour la communauté scientifique, L'analyse des compétences en Big Data a été approchée à travers deux différentes perspectives, une revue de la littérature et une étude empirique par le biais des annonces d'emploi des entreprises marocaines. Nous avons eu recours à deux méthodologies de recherche, une technique d'extraction du contenu de sites Web et une méthodologie de l'analyse de contenu. Les compétences collectées à travers la revue de la littérature et les données empiriques ont été comparées sur la base d'un codage et d'une catégorisation des auteurs scientifiques. Les résultats de la recherche ont permis de démontrer

que les exigences des entreprises marocaines vont dans le même sens que la littérature à la différence près concernant la connaissance du monde des affaires ou des processus métiers qui peut s'expliquer du fait que la majorité des annonces émanent des entreprises spécialisées dans le conseil en système d'information. Par ailleurs, cette spécialisation pourrait être un argument pour expliquer l'exigence des entreprises pour la compétence veille technologique.

Selon les résultats, les compétences les plus importantes d'un scientifique des données sont les compétences analytiques, les compétences de programmation et de communication, l'apprentissage automatique et l'exploration de données. En outre, la passion et la capacité de résoudre des problèmes d'entreprise en recourant aux Big Data.

Pour les praticiens, cet article pourrait constituer une référence afin de s'aligner avec les normes internationales relevées au niveau de la revue de littérature et les orientations du marché marocain.

REFERENCES

1. Armbrust, M., Fox, A., Griffith, R., et al. (2010). *A View of Cloud Computing*. Communications of the ACM, vol. 53, pp. 50- 58.
2. Barr, V., & Stephenson, C. (2011). *Bringing computational thinking to K-12: what is involved and what is the role of the computer science education community?*. ACM Inroads, 2(1), 48-54."
3. Albig, W. (1952). BERELSON, BERNARD. *Content Analysis in Communication Research*. Pp. 220. Glencoe, Ill.: The Free Press.
4. BENKARAACHE, T. and GHANOUANE, K. 2020. Modèle théorique d'évaluation de l'apport de la transformation digitale à la chaîne de valeur des entreprises. *International Journal of Management Sciences*. 3, 2
5. Boyd, D., & Crawford, K. (2012). *Critical questions for big data: Provocations for a cultural, technological, and scholarly phenomenon*. *Information, communication & society*, 15(5), 662-679.
6. Bucki, J., (2016). *Definition and Explanation of Mobile Computing*. Repéré à <https://www.thebalance.com/definitionof-mobile-computing-2533640>
7. Chen, H., Chiang, R. H., & Storey, V. C. (2012). *Business Intelligence and Analytics: From Big Data to Big Impact*. *MIS quarterly*, 36(4), 1165-1188."
8. Chen, H., Zhang, J., Ma, S., Yang, S., Li, L., Huang, X., & Xiao, J. (2015). *Confined flares in solar active region 12192 from 2014 October 18 to 29*. *The Astrophysical Journal Letters*, 808(1), L24.
9. Conway, D. (2010). *The data science Venn diagram*. Repéré à <http://drewconway.com/the-lab/>"
10. Davenport, T. H., & Dyché, J. (2013). *Big data in big companies*. International Institute for Analytics.
11. Davenport, T.H. and Patil, D.J. (2012), *Data Scientist: The Sexiest Job of the 21st Century*, Harvard Business Review, Harvard Business School Publication Corp., Vol. 90 No. 10, pp. 70–76.
12. De Mauro, A., Greco, M., Grimaldi, M., & Ritala, P. (2018). *Human resources for Big Data professions: A systematic classification of job roles and required skill sets*. *Information. Processing & Management*, 54(5), 807-817.
13. Dhar, V. (2013). *Data science and prediction*. Communications of the ACM, 56(12), 64-73.
14. Ebert, C., & Duarte, C. H. C. (2016, September). *Requirements engineering for the digital transformation: Industry panel*. In 2016 IEEE 24th International Requirements Engineering Conference (RE) (pp. 4-5). IEEE.
15. Ferreira et Niels, van Berkel, Denzil et Kostakos, Vassilis. (2017). *The Experience Sampling Method on Mobile Devices*. *ACM Computing Surveys*, vol. 50, pp. 1-40.

16. Fitzgerald, M., Kruschwitz, N., Bonnet, D., and Welch, M. (2013). *Embracing Digital Technology*, MIT Sloan Management Review, pp. 1-12.
17. Gandomi, A. and Haider, M. (2014), *Beyond the hype: Big data concepts, methods, and analytics*, International Journal of Information Management, Elsevier Ltd, Vol. 35 No. 2, pp. 137–144.
18. Granville, V. (2014). *Developing analytic talent: Becoming a data scientist*. John Wiley & Sons.
19. El Haddadi, A., El Haddadi, A., Dousset, B., & Fennan, A. (2015). *Le système d'intelligence économique-XEW*. Revue Electronique en Technologies de l'Information, 1(8), 20-35.
20. Hammerbacher, J. (2009). *Information platforms and the rise of the data scientist*. Beautiful Data, 73-84.
21. Harold Wilensky, dans *la définition de l'Organizational Intelligence dans l'International Encyclopedia of Social Sciences* (David L. Sills), editor, New York: The Macmillan Company & The Free Press, 1967.
22. Harris, H., Murphy, S., & Vaisman, M. (2013). *Analyzing the Analyzers: An Introspective Survey of Data Scientists and Their Work*. O'Reilly Media.
23. Kaschny, Martin, Nolden, Matthias (2018). *Innovation and Transformation Basics, Implementation and Optimization*. Springer
24. Kiron, D., Shockley, R., Kruschwitz, N., Finch, G., & Haydock, M. (2012). *Analytics: The widening divide*. MIT Sloan Management Review, 53(2), 1.
25. Kwon, O., & Sim, J. M. (2013). *Effects of dataset features on the performances of classification algorithms*. Expert Systems with Applications, 40(5), 1847-1857.
26. Laney D (2001) *3D Data Management: Controlling Data Volume, Velocity, and Variety*. Repéré à <http://blogs.gartner.com/doug-laney/files/2012/01/ad949-3D-Data-Management-Controlling-Data>.
27. Manyika, J., Chui, M., Brown, B., Bughin, J. and James Manyika, Michael Chui, Brad Brown, Jacques Bughin, Richard Dobbs, Charles Roxburgh, A.H.B. (2011), *Big data: The next frontier for innovation, competition, and productivity*, McKinsey Global Institute.
28. Martinet, B., & Marti, Y. M. (2001). *L'intelligence économique : comment donner de la valeur concurrentielle à l'information*. Editions d'organisation.
29. Matt, Christian; Hess, Thomas; and Benlian, Alexander (2015). *Digital Transformation Strategies*, Business & Information Systems Engineering: Vol. 57: Iss. 5, pp. 339-343.
30. Mayer-Schönberger, V. and Cukier, K. (2013), *Big Data: A Revolution that Will Transform how We Live, Work, and Think*.
31. McAfee, A., Brynjolfsson, E., Davenport, T. H., Patil, D. J., & Barton, D. (2012). *Big data. The management revolution*. Harvard Bus Rev, 90(10), 61-67.
32. McNulty. E, (2014), *Big Data: The Seven V's*, Repéré À [Http://Dataconomy.Com/Seven-Vs-Big-Data/](http://Dataconomy.Com/Seven-Vs-Big-Data/)
33. Miller, S. (2014). *Collaborative Approaches Needed to Close the Big Data Skills Gap*. Journal of Organization Design, 3(1), 26-30.
34. Monino. L.L, Sedkaoui.S Et Matouk. J, (2014). *Big Data, Ethique Des Donnees, Et Entreprises*, Les Cahiers Du Cedimes, Dossier: « Economie Et Gouvernance », Vol. 8, No2 Repéré A [Http://Cedimes.Com/Images/Documents/Cahiers_Cedimes/Cahier_Cedimes_Vol_8_N2_20_4.Pdf](http://Cedimes.Com/Images/Documents/Cahiers_Cedimes/Cahier_Cedimes_Vol_8_N2_20_4.Pdf)
35. Moody, D. L., & Walsh, P. (1999, June). *Measuring the Value of Information-An Asset Valuation Approach*. In ECIS (pp. 496-512).
36. Miorandi, Daniele & Sicari, Sabrina & De Pellegrini, Francesco & Chlamtac, Imrich. (2012). *Internet of Things: Vision, Applications and Research Challenges*. Ad Hoc Networks.
37. Patil, D. J. (2011). *Building data science teams*. O'Reilly Media, Inc.
38. Petkovics, I. (2018). Digital transformation in higher education. Journal of Applied Technical and Educational Sciences, 8(4), 77-89.
39. Pinola, T. (2015). *Datatieteilijän kompetenssien määrittelyminen*. "Provost, F. & Fawcett, T (2013A). *Data Science and its Relationship to Big Data and Data-driven Decision Making*. Big Data Journal, Vol. 1, No. 1, March 2013

40. Reddi, V., Zhu, Y. (2017). *The Red Future of Mobile Web Computing*. Communications of the ACM, vol. 60(1), pp. 52- 53.
41. Reis, J., Amorim, M., Melão, N., & Matos, P. (2018, March). *Digital transformation: a literature review and guidelines for future research*. In World conference on information systems and technologies (pp. 411-421). Springer, Cham.
42. Ross, J.W., Sebastian, I.M., Beath, C., Mockler, M., Moloney, K., & Fonstad, N. (2016). *Designing and Executing Digital Strategies*. ICIS.
43. Rouse, M. (2013, October 1). *What is advanced analytics? – Definition from WhatIs.com*. Repéré à <http://searchbusinessanalytics.techtarget.com/definition/advanced-analytics>
44. Roussos, George & Kostakos, Vassilis. (2009). *RFID in Pervasive Computing: State-of-the-art and Outlook*. Pervasive and Mobile Computing. pp. 110-131.
45. Schlesinger, Philip, and Doyle, Gillian (2014). *from organizational crisis to multi-platform salvation? Creative destruction and the recomposition of news media*. Journalism: Theory, Practice and Criticism.
46. Schwertner Charão A., Marin F., D. Lima J., Cortez da Rocha C. and Steffenel L. (2018). *On the Performance of Cloud-based Spreadsheets as a Backend for View-only Web Applications*. In Proceedings of the 20th International Conference on Enterprise Information Systems, Vol 1, pp 642-647.
47. Song, I.-Y. and Zhu, Y. (2015). *Big data and data science: what should we teach?*. Expert Systems.
48. Van den Driest, F., Sthanunathan, S., & Weed, K. (2016). *Building an insights engine*. Harvard business review, 94(9), 15.
49. Vargiu, E. and Urru, M. (2012). *Exploiting web scraping in a collaborative filtering- based approach to web advertising*. Artificial Intelligence Research, Vol. 2 No. 1, pp. 44–54.
50. Waller, M. A., & Fawcett, S. E. (2013). *Data science, predictive analytics, and big data: a revolution that will transform supply chain design and management*. Journal of Business Logistics, 34(2), 77-84.
51. Wamba, S. F., Akter, S., Edwards, A., Chopin, G., & Gnanzou, D. (2015). *How ‘big data’ can make big impact: Findings from a systematic review and a longitudinal case study*. International Journal of Production Economics, 165, 234-246.
52. Wing, J. M. (2006). *Computational thinking*. Communications. ACM 49, 33–35.
53. Wixom, B., Ariyachandra, T., Douglas, D., Goul, M., Gupta, B., Iyer, L., Kulkarni, U., et al. (2014). *The current state of business intelligence in academia: The arrival of big data*. Communications of the Association for Information Systems, Vol. 34 No. 1, pp. 1–13.