



Le **Big Data**

à Bruxelles aujourd'hui.
Et demain ?





"Big data is like teenage sex: Everyone talks about it, nobody really knows how to do it, everyone thinks everyone else is doing it, so everyone claims they are doing it..."¹

Professor Dan Ariely (Duke University)

1. Source: <https://whatsthebigdata.com/2013/06/03/big-data-quotes/>



Table des matières

Préface	7
1. Introduction et méthodologie	9
1.1 Méthodologie du rapport	10
2. Le Big Data : délimitation du sujet et origine	10
2.1 Qu'est-ce que le Big Data ?	10
2.2 Les atouts du Big Data	14
2.3 Les limites du Big Data	15
2.4 L'aspect juridique du Big Data	15
3. Le Big Data en Europe	18
4. Le Big Data à Bruxelles	19
4.1 L'aspect politico-économique	19
4.2 Les entreprises et le Big Data	19
5. Les métiers du Big Data	21
5.1 Le Data Architect	22
5.2 Le Data Scientist	22
5.3 Le Data Analyst	23
5.4 Des Disciplines transversales	24
5.5 Les compétences clés	24
6. Des formations Big Data	25
6.1 Les formations scolaires	25
6.2 Les formations universitaires et Hautes Ecoles	26
6.3 Les formations qualifiantes	27
6.4 Autres initiatives	28
6.5 Enseigner le Big Data à Bruxelles	28
7. Conclusions et recommandations	30
7.1 Recommandations	32
7.1.1 La sensibilisation des entreprises et du grand public	32
7.1.2 Former plus de spécialistes des données en Belgique	32
7.1.3 Sensibiliser les entreprises sur l'aspect juridique	33
8. Remerciements	34
9. Glossaire	38
10. Evoliris	40



Préface

Le monde de la technologie évolue à grande vitesse. Evoliris en tant que Centre de Référence Professionnelle ICT bruxellois a pour mission de suivre les tendances technologiques, de les analyser et se positionner en tant que prescripteur dans le secteur de l'IT bruxellois. Cette mission s'adresse tant aux acteurs du monde IT qu'aux demandeurs d'emploi.

Dans le contexte d'une société numérique de plus en plus informatisée, Evoliris se doit d'analyser l'actualité au niveau national mais aussi et surtout au niveau bruxellois afin de déterminer les tendances dans les entreprises IT bruxelloises et de proposer ou suggérer aux acteurs de l'emploi et de la formation des pistes de travail et des recommandations.

Cette veille sectorielle s'étoffe par la conception de rapport tel que celui-ci. L'objectif de ce document consiste à présenter la situation à Bruxelles d'une tendance technologique qui fait beaucoup parler d'elle dans l'actualité IT : le Big Data.

Sciences des données, données massives, le Big-Data est de plus en plus considéré comme un or noir car il permet aux entreprises de développer leur potentiel économique. Dans cette réflexion, le débat politique et public a toute son importance. En effet, le Big Data trouve également sa place dans les administrations publiques et permet d'améliorer le quotidien des citoyens.

Le présent document se veut également vivant et au plus proche des réalités du terrain. C'est pourquoi, nous avons sollicité l'avis de plusieurs experts sur leur connaissance pratique et quotidienne du Big Data. Ainsi, dès le début de ce projet nous avons souhaité sortir du cadre théorique et présenter une analyse réaliste de la situation à Bruxelles. Dans ce sens, nous avons rédigé un dossier qui présente et analyse les points de vue de spécialistes sur le plan politico-économique et légal pour en présenter les perspectives dans le cadre des problématiques de l'emploi et de la formation à Bruxelles. Le point fort de cette démarche est de consolider le lien entre les différentes parties et positionner Evoliris en tant que médiateur et facilitateur.

Nous tenons à remercier tous les contributeurs à ce rapport car ce travail n'aurait pas eu la qualité qu'il a aujourd'hui sans leur expertise.

Vous tenez donc entre vos mains le premier pilier d'une série de rapports qui se succéderont à l'avenir et que nous appellerons « Les Cahiers d'Evoliris ». Celui-ci concerne le Big Data, les suivants analyseront d'autres thématiques IT tendances sur le marché belge.



Jean-Pierre Rucci
Directeur d'Evoliris



Floriane de Kerchove
Présidente d'Evoliris & directrice d'Agoria



Dans le but d'améliorer sa réponse face au flux de touristes, la ville de Bruxelles utilise le Big Data pour mieux connaître ses visiteurs. Chaque année, la ville de Bruxelles accueille 5,1 millions de touristes, et 1,5 millions de vacanciers². Afin de mesurer le flux de visiteurs, Bruxelles a décidé d'utiliser les services de géolocalisation. Avec l'appui de Proximus, Visit.brussel a analysé les données téléphoniques de 10 lieux touristiques bruxellois. Ces données dépersonnalisées ont permis de comprendre la logique du touriste et notamment de calculer le temps de visite d'un lieu, le temps de trajet et le lieu où les visiteurs dorment de préférence. D'un point de vue plus général, ces données ont permis à la ville de Bruxelles de mieux gérer ce flux et d'améliorer les services des organisateurs d'événements³.

Cet exemple est une belle application de l'utilisation des données Big Data dans le but d'améliorer le quotidien d'un groupe d'individus. Il existe de nombreux autres cas dans le public et le privé qui démontrent que le Big Data peut être une source d'amélioration importante qui contraste avec la vision négative du Big Data.

Le Big Data ou *la science des données* est un sujet complexe et passionnant. Son application est multiple et permet aux entreprises et institutions d'élaborer une stratégie efficace afin d'améliorer leur rentabilité et leurs actions. Le terme « Big Data » est très tendance depuis plusieurs années. En effet, son application est de plus en plus utile. Néanmoins, la communauté du Big Data est divisée : certains en font l'apologie, d'autres estiment que ce n'est rien d'autre qu'une évolution logique de l'analyse des données. Il ne nous appartient pas dans ce rapport de trancher sur la question.

Ce rapport a pour objectif de mettre en perspective cette technologie dans le contexte bruxellois. Nous y interprétons ce concept dans le contexte socioéconomique de Bruxelles, ainsi qu'en termes d'emploi et de formation. Nous analysons également son impact sur les entreprises bruxelloises, au sein de la Commission Européenne, ainsi que sur le plan juridique.

Le Big Data attire depuis longtemps l'attention des grosses sociétés qui doivent brasser de grandes quantités de données. On peut d'ailleurs traduire littéralement « Big Data » par « données massives ». Toutefois, Le Big Data ne peut pas être réduit aux données massives. Nous verrons plus tard qu'il n'est pas nécessaire d'avoir de grandes quantités de données à analyser pour faire du Big Data. Dans cette perspective, Il est essentiel de se poser la question de l'intérêt du Big Data et de son avenir à Bruxelles.

2. http://www.proximus.be/fr/id_b_cl_tourism/grandes-entreprises-et-secteur-public/decouvrir/blog/one/customer-stories/le-tourisme.html#

3. http://www.proximus.be/fr/id_b_cl_tourism/grandes-entreprises-et-secteur-public/decouvrir/blog/one/customer-stories/le-tourisme.html#

1.1 Méthodologie du rapport

La partie préparatoire de ce rapport a été divisée en deux étapes. Une première étape consistait à faire une enquête sur le sujet par des recherches. Cette étape avait pour but de délimiter le cadre du rapport : comprendre et définir le concept, savoir qui contacter pour la phase d'entretien, etc. Après avoir délimité le sujet, nous avons commencé la phase d'entretien. En effet, pour avoir un aperçu au plus proche de la réalité du terrain, il nous a paru essentiel de compléter nos recherches par des interviews avec des spécialistes du Big Data. Ces spécialistes sont issus du monde de l'entreprise, mais sont également des acteurs de l'emploi et de la formation, etc. (Pour plus d'informations sur ces personnes, rendez-vous sur la page consacrée aux «Remerciements »).

Ainsi, le rapport ne représente qu'un aperçu de la situation. Nous avons cherché à présenter une analyse au plus proche de la réalité en interrogeant des acteurs du secteur ayant une expérience probante dans le Big Data et la science des données.

Au fil de nos enquêtes, nous nous sommes aperçus d'une réalité propre à Bruxelles : le Big Data est séduisant mais peu d'entreprises osent s'y lancer. **Il y a un vrai problème de sensibilisation.** Beaucoup de sociétés pensent à tort que le Big Data n'est pas intéressant pour leur stratégie marketing. On constate également un retard important en termes de formation malgré des initiatives de plus en plus nombreuses et une pénurie de spécialistes belges du Big Data.

Le Big Data : délimitation du sujet et origine

2

2.1 Qu'est-ce que le Big Data ?

On traduit généralement cette technologie par *l'ensemble des données recueillies, nettoyées, traitées, analysées dans le but de fournir une analyse et des recommandations en termes de stratégie.*

L'objectif de ce rapport n'est pas de définir en profondeur le concept du Big Data. Il est néanmoins important de commencer ce document par une courte définition avant d'évoquer les problématiques qui y sont liées.

Le Big Data est souvent mentionné dans le contexte de la lutte contre le terrorisme. Les technologies Big Data peuvent, en effet, être utilisées pour améliorer la sécurité de la collectivité. Dans ce sens, Microsoft a développé des outils (Microsoft Azure Data Lake⁴, Microsoft Cortana Analytics⁵) permettant d'appliquer une analyse prédictive sur la criminalité. Ces outils très poussés permettent d'acquérir des informations détaillées afin de dégager un schéma criminel. On appelle cette technologie : l'analyse prédictive⁶. Le terrorisme est également un sujet très sensible en Belgique. Le Big Data permet aux enquêteurs d'analyser et trier des informations potentiellement importantes en étudiant des conversations téléphoniques, des données issues de Smartphones, en analysant les déplacements de

4. Microsoft Azure Data Lake : voir glossaire

5. Microsoft Cortana Analytics : voir glossaire

6. <https://enterprise.microsoft.com/en-us/industries/government/fighting-crime-with-big-data-analytics/>

possibles terroristes à l'aide d'images de caméras publiques, etc. D'ailleurs, le CIRB⁷ a mis en place une « *plateforme mutualisée d'échange d'images de vidéoprotection. Ces images sont issues des 6 zones de polices bruxelloises, de la Société des transports intercommunaux de Bruxelles (STIB), du Service d'incendie et d'Aide Médicale urgente (SIAMU), de Bruxelles-Mobilité et du Port de Bruxelles* »⁸.

À première vue, on a tendance à associer Big Data au **marketing**. C'est, en effet, la première application à laquelle on pense. Le Big Data est largement utilisé dans ce contexte et a permis à de grandes sociétés d'analyser les besoins de leurs clients et d'adopter une stratégie de communication ciblée. En recueillant toute une série d'informations sur les clients à travers des sources différentes (réseaux sociaux, cartes de fidélité, navigation sur le site, ..), ces entreprises peuvent adapter leur stratégie au cas par cas. Facebook, par exemple, utilise les technologies du Big Data pour déployer de manière plus ciblée ses publicités sponsorisées⁹.

Gardons en tête que le Big Data n'a pas que des applications dans le marketing. Le Big Data est une véritable source de progrès dans d'autres secteurs comme par exemple **les statistiques sportives**. L'article sur l'EURO 2016 « *l'Allemagne joue au Big Data sportif pour l'emporter (encore)* » dans l'Usine Digitale¹⁰, évoque un fait qui n'a rien d'anecdotique. L'équipe d'Allemagne a déclaré avoir engagé des experts en Big Data sportif. Les résultats issus de ces analyses des jeux dans les matchs permettent à l'Allemagne d'adapter sa stratégie. Ils affirment qu'un match doit se préparer aussi avec la science des données. On ne sait pas si le recours à cette technologie a influencé leur progression mais on a retrouvé l'équipe allemande en demi-finale face à la France. Finalement, on sait tous que ça ne leur a pas porté chance, mais qui sait ?! Peut-être pour un prochain match. Comme nous le voyons avec cet exemple, même les équipes de football peuvent profiter des technologies Big Data pour améliorer leur stratégie de jeu et bien mieux connaître leurs adversaires.

Si on entend beaucoup parler du Big Data dans l'actualité, ce n'est néanmoins pas une technologie récente. Déjà, dans les années 1970, l'idée du Big Data est mise sur la table. C'est en 1997 que le terme « Big Data » est utilisé pour la première fois dans un article – « *Application-controlled demand paging for out-of-core visualization* » - par Michael Cox et David Ellsworth¹¹. À l'époque, c'était encore une théorie car la technologie ne permettait pas de mettre en place des outils. Il a fallu attendre les années 2010 pour envisager concrètement l'application du Big Data.

En 2001, la société Gartner¹² définissait le Big Data avec trois caractéristiques :

1. **Le volume** : l'une des caractéristiques les plus connues du Big Data est que l'algorithme traite une quantité volumineuse de données. La quantité massive des données est une problématique d'actualité : le monde a produit ces dernières années plus d'informations que toutes celles produites depuis l'invention de l'écriture.
2. **La variété** : la variété distingue le Big Data de l'analyse de données dite « classique ». En effet, le Big Data analyse des jeux de données en provenance de sources différentes.

7. CIRB : Voir glossaire

8. Pour plus d'information sur ce projet et d'autres : consulter le cahier du CIRB 35 : <http://cirb.brussels/fr/quoi-de-neuf/publications/cahiers/nr-35-4-projets-cles-de-smartcity-brussels>

9. Pour en savoir plus sur l'utilisation du Big Data par Facebook : <http://www.simplilearn.com/how-facebook-is-using-big-data-article>

10. Nathalie Steiwer, EURO 2016 : l'Allemagne joue au Big Data sportif pour l'emporter (encore) Dans L'usine Digitale [en ligne] <http://www.usine-digitale.fr/article/euro-2016-l-allemande-joue-au-big-data-sportif-pour-l-emporter-encore.N396112> (Consulté le 16 juin 2016).

11. Gil Press, A very short History of Big Data Dans Forbes [en ligne] <http://www.forbes.com/sites/gilpress/2013/05/09/a-very-short-history-of-big-data/#6d76ff6155da> (consulté le 17 juin 2016).

12. Gartner : Voir glossaire.

On traite donc des données structurées (tableaux croisés dynamiques du type Excel, bases de données traditionnelles, etc.) et des données non-structurées (textes libres, informations issues des réseaux sociaux, vidéos, etc.). A l'aide des outils de base de données classiques, il est donc difficile de rassembler des données non-structurées. Le Big Data, lui, a les capacités de recueillir ces données très disparates afin de les analyser.

3. **La vélocité** : Le troisième critère peut être défini par la rapidité avec laquelle les données sont et doivent être collectées pour être classées et analysées. Dans le Big Data, les données sont souvent très nombreuses : il faut donc rassembler ces données rapidement pour pouvoir toutes les traiter.

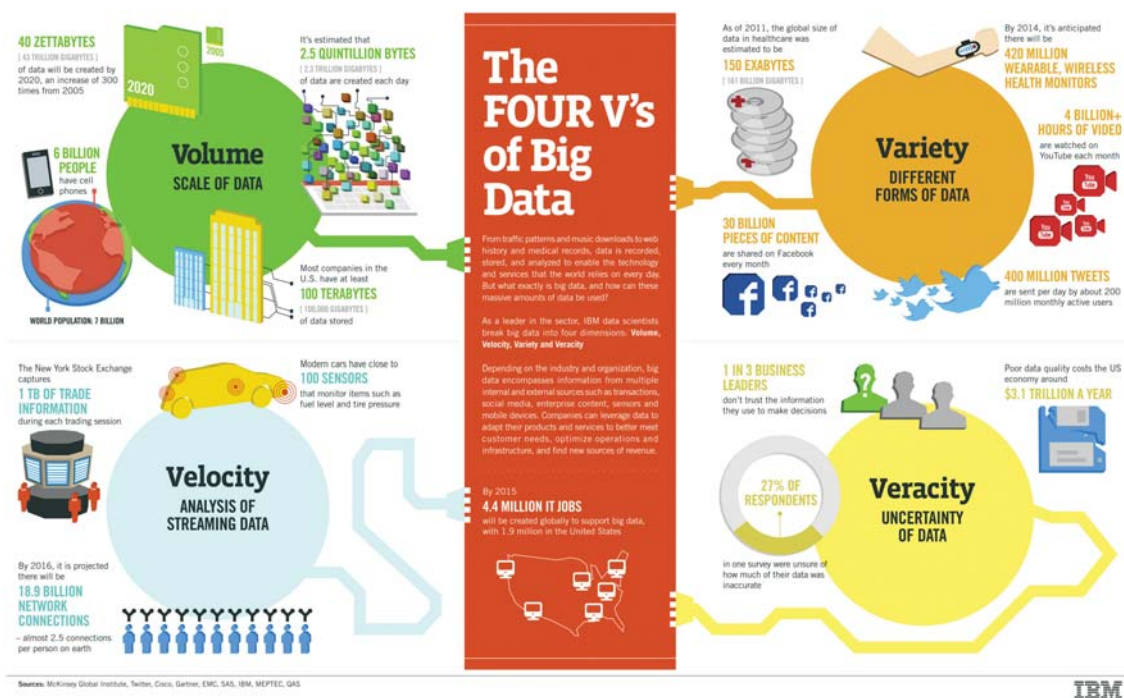


Figure 1 The four V Of Big Data¹³

Par après, la communauté scientifique ajouta à ces trois caractéristiques, d'autres critères comme :

4. **la véracité** : « La notion de véracité met en avant la dimension qualitative des données nécessaire au fonctionnement des outils de Big Data¹⁴ ».
5. **la visibilité** : « la visibilité proposée par les plates-formes de Big Data efficaces permettent aux utilisateurs de bien utiliser les données et informations publiées en effectuant des recherches pertinentes¹⁵ ».
6. **la valeur** : « pour une entreprise, elle se mesure à l'avantage compétitif qu'elle en aura dégagé. Cela dépend notamment de la qualité des analytics et de la compétence des « Data Scientists » chargés de leur donner du sens¹⁶. »
7. **la marge d'erreur** (penalty ranking qu'on appelle aussi « snow »)¹⁷.

13. <http://www.ibmbigdatahub.com/infographic/four-vs-big-data>

14. <http://www.zdnet.fr/actualites/big-data-3-4-ou-5-v-39832210.htm>

15. <http://www.zdnet.fr/actualites/big-data-3-4-ou-5-v-39832210.htm>

16. <https://www.digitalwallonia.be/big-data-la-revolution-des-donnees/>

17. Pour aller plus loin : <https://campus-bruxelles.com/2014/02/12/big-data-formation-informatique-bruxelles/>

Dans la compréhension générale, on a tendance à associer Big Data et technologies avancées. Ce n'est pas faux mais très réducteur, car le Big Data ce n'est pas qu'une question de technologies. Depuis quelques années, on parle des technologies suivantes :

- Hadoop : « *framework Open Source conçu pour réaliser des traitements sur des volumes de données massifs, de l'ordre de plusieurs petaoctets (soit plusieurs milliers de To)*¹⁸ »
- Le logiciel R : analyse et traite les données à l'aide du langage de programmation S¹⁹.
- Map Reduce : « *permet de manipuler de grandes quantités de données en les distribuant dans un cluster de machines pour être traitées* »²⁰.
- Et d'autres encore, ...

Ces **technologies** évoluent très vite et sont rapidement remplacées par d'autres plus évoluées, plus pertinentes ; Hugues Bersini en tant que responsable de l'IRIDIA²¹ à l'ULB insiste sur ce fait lors de notre conversation. Ainsi, il est évident que le Big Data ne doit pas être réduit à des technologies. Ce sont certes des technologies de pointe mais encore instables en raison de leur évolution rapide. Le Big Data, c'est avant tout un **ALGORITHME**. L'algorithme constitue le cœur de l'enjeu. Il recueille les data, les traite, etc. La formule statistique est souvent conçue en fonction de la demande du client, de la question qu'il veut poser à l'algorithme. Un algorithme mal conçu est une mauvaise application de Big Data. C'est pourquoi les sociétés accordent beaucoup d'importance au profil métier qui s'occupe de la conception de l'algorithme : **le Data Scientist**. Malheureusement, nous verrons dans le chapitre consacré aux profils métier du Big Data (Les métiers du Big Data) que ce type de professionnel est encore trop rare en Belgique.

Une autre problématique liée au Big Data est celle de l'association exclusive du Big Data à la notion de **données massives**. Peut-être le terme est-il mal choisi mais cela donne l'impression que l'enjeu de cette technologie se résume à la quantité volumineuse de données brassées par l'algorithme. Or, le volume n'est pas une fin en soi. Le Big Data ce n'est pas qu'une question de volume. Il faut penser le Big Data en se détachant du « Big », car il ne concerne pas que les grosses sociétés qui peuvent se permettre de brasser une quantité « hallucinante » de données. Lors des divers entretiens que nous avons réalisés, il apparaît que le Big Data c'est aussi et surtout une question de variété. Il est d'ailleurs de moins en moins fréquent d'avoir recours au terme « Big Data ». Mr. Philippe Van Impe, fondateur de l'European Data Innovation Hub et de Brussels Data Science Community²², explique que le terme Big Data est dépassé. La communauté scientifique préfère désormais parler de « **Science des données** ».

Le Big Data, enfin, est lié au **stockage des données**. En effet, la quantité de données engendrée de nos jours requiert un lieu de stockage performant et adapté. C'est depuis l'apparition du Cloud Computing et sa maturation que le Big Data a réellement pu se développer²³.

18. <http://www.journaldunet.com/solutions/systemes-reseaux/definition-d-hadoop.shtml>

19. [https://fr.wikipedia.org/wiki/R_\(langage_de_programmation_et_environnement_statistique\)](https://fr.wikipedia.org/wiki/R_(langage_de_programmation_et_environnement_statistique))

20. <https://fr.wikipedia.org/wiki/MapReduce>

21. IRIDIA : voir glossaire.

22. Site web : <https://datasciencebe.com/>

23. <http://www.zdnet.fr/actualites/comment-le-big-data-impacte-le-cloud-computing-39829534.htm>

2.2 Les atouts du Big Data

On entend énormément parler du Big Data dans la presse, chez les spécialistes des données, les entreprises ; tout le monde s'y met. Or, nous l'avons déjà signalé ; ce n'est pas un concept nouveau. Pourtant, ce qui fait le succès du Big Data actuellement, est sa maturation dans le monde. Dans son Hype Cycle de 2014, Gartner classe le Big Data au début de la courbe du « *trough of Disillusionment* » juste après le « *Peak of Inflated Expectations* ».

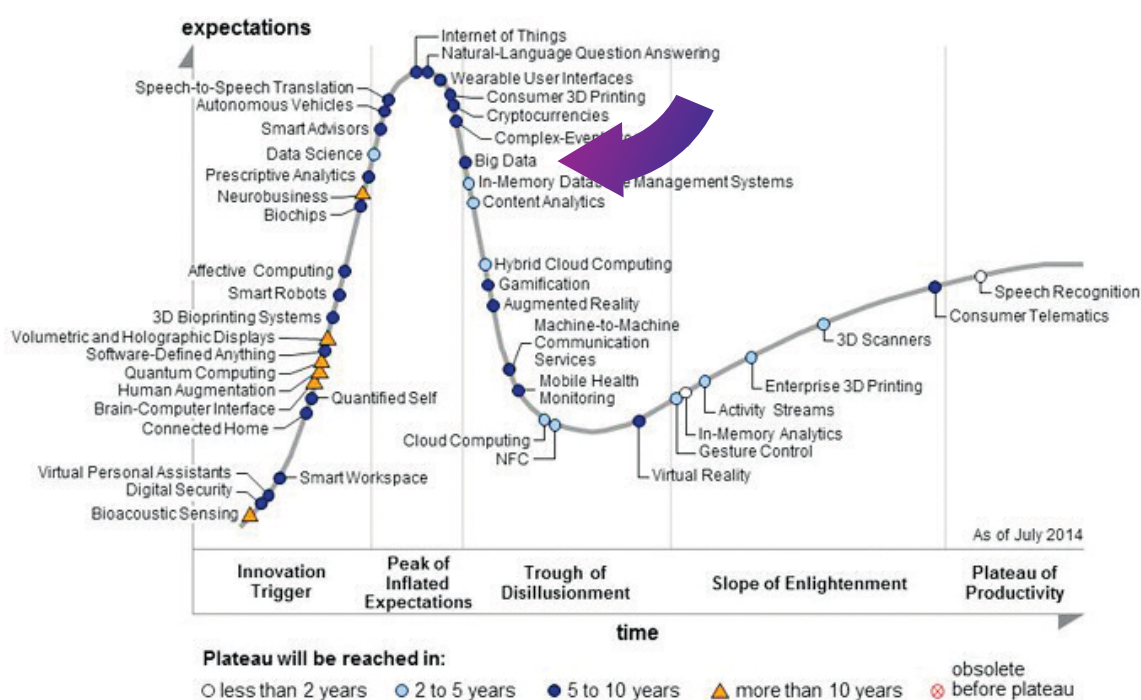


Figure 2: le 'Hype Cycle' 2014 de Gartner

C'est bien entendu en lien avec sa maturation. Les entreprises doivent passer à un niveau supérieur en termes d'analyse de données et le Big Data exprime bien cette évolution. L'un de ses atouts est donc sa consolidation sur le marché international et son évolution très rapide. On peut interpréter sa position en 2014 par le fait que les technologies associées au Big Data évoluent et se remplacent très vite. Bruxelles est à la traîne mais il ne serait pas étonnant que le Big Data finisse par concrètement s'y implanter de manière durable.

En raison de cette maturité et de son application de plus en plus facilitée, on peut également considérer que le Big Data doit et sera un critère de mise à l'emploi. La donnée est devenue l'or noir que les entreprises veulent exploiter à tout prix. Afin d'exploiter les données, il faut engager. Nous verrons plus loin qu'il y a un besoin urgent de former des spécialistes au Big Data. Ces jobs sont particulièrement demandés par les recruteurs belges.

On ne le répètera jamais assez ; les données ont un réel impact dans toute une série de domaines d'application : criminalité, agriculture, gestion des stocks d'une industrie, secteur de la santé, etc.

2.3 Les limites du Big Data

La principale contrainte du Big Data est juridique (voir point 2.4). Les notions associées aux problématiques du Big Data sont la protection des données à caractère privé. La science des données demande de traiter parfois des données personnelles. Par exemple : un opérateur de téléphone qui voudrait analyser les habitudes de communication d'un groupe cible, a besoin des données de ses clients. Or, cela pourrait s'apparenter au genre tant apprécié par les films futuristes : le concept de « Big Brother ». C'est un sujet très sensible dans l'actualité (Affaire Snowden, etc.). Il faut néanmoins se rendre compte que la plupart du temps cette technologie ne va pas surveiller l'individu et sa vie personnelle. L'outil va juste analyser le comportement dans un échantillon de personnes afin de dégager une estimation des attitudes sans rentrer dans le détail, dans la donnée personnelle. C'est une chose qui peut également avoir son utilité publique, par exemple dans une société de transport en commun²³. Si celle-ci veut analyser la pertinence et la fréquence des trajets de ses usagers dans ses véhicules, elle peut le faire en étudiant les données transmises par la carte d'abonnement des usagers. Cette technique n'est pas seulement un moyen de surveillance mais un outil pour améliorer les services proposés aux utilisateurs.

Un autre frein à l'ancrage du Big Data à Bruxelles est un manque de connaissance. Les dirigeants et les chefs d'entreprises ne connaissent pas le Big Data. Ils n'osent pas l'utiliser pour leur société ou ne pensent pas en avoir besoin. Nous aborderons ce sujet plus en détail dans la partie consacrée aux entreprises (Le Big Data à Bruxelles).

2.4 L'aspect juridique du Big Data

Comme pour toute évolution technologique, il ne faut pas oublier de s'intéresser aux lois qui régulent son application.

Maitre Philippe Laurent, avocat au barreau de Bruxelles, spécialiste en droit de la propriété intellectuelle et en droit des nouvelles technologies de l'information et de la communication, explique au cours de notre entretien que « *certaines entreprises s'intéressent mal ou trop tard à l'aspect juridique. C'est une nécessité d'y penser le plus tôt possible car la viabilité d'un projet peut parfois reposer sur une pratique qui, après analyse, peut s'avérer illégale* ».

D'un point de vue juridique, le Big Data est un concept assez récent mais très vaste. C'est pourquoi dans la législation il n'en existe pas encore de définition concrète. Il est donc très facile de rendre floue la frontière entre la pratique éthique et la fraude (intentionnelle ou non). Comme le précise M. Laurent, les entreprises qui touchent au Big Data doivent se poser une série de questions avant de commencer le projet : Quelles données sont traitées ? Quelle est leur origine ? Pour quelle finalité et comment sont-elles traitées ? Ces données sont-elles « à caractère personnel » ? Toutes ces questions permettent de délimiter un cadre au sein duquel elles pourront jouer.

Il y a des limitations dans l'utilisation des données. On évoque souvent un réel problème de protection des données privées et de la propriété intellectuelle.

24. Exemple concret : " un réseau d'autobus redessiné grâce au téléphone mobile » In La recherche – décembre 2013 N° 482 [en ligne] http://www.ffue.org/wp-content/uploads/2015/09/FU-US_150927_Blondel_LaRecherche_BigData.pdf [consulté le 11 juillet 2016].

Selon la Commission Vie Privée, les **données à caractère personnel** sont « *des données concernant une personne directement ou indirectement identifiée ou identifiable*²⁵ ». Il existe donc des lois qui régissent l'utilisation des données privées. Une entreprise ne peut pas se permettre d'utiliser comme il le souhaite les données de ses clients potentiels ou existants.

« *La **propriété intellectuelle** désigne l'ensemble des droits exclusifs accordés sur les créations intellectuelles*²⁶ » explique le Parlement Européen. De ce côté également, une législation est en vigueur. Le terme est entré dans le traité sur le Fonctionnement de l'Union Européenne (traité FUE) en 2009²⁷.

En Europe, la protection des données est un véritable enjeu. C'est une problématique qui est prise très au sérieux. D'ailleurs, la législation est assez stricte sur l'utilisation des données privées à des fins commerciales. Au contraire, les Etats-Unis sont moins exigeants quant à l'aspect juridique.

Depuis 2015 environ, nous avons beaucoup entendu parler du **Safe Harbor** et de **Privacy Shield** dans les médias. Il s'agit de lois légiférant sur le transfert de données entre l'Europe et les Etats-Unis. En effet, les deux continents n'ont pas les mêmes priorités, il est donc important, à l'heure de la mondialisation, de trouver une législation commune qui puisse convenir tant à l'Europe qu'aux Etats-Unis.

Le **Safe Harbor**, créé en 2015, consiste en l'adoption d'une série de lois visant à régir l'utilisation des données des citoyens européens par les entreprises américaines ayant des succursales en Europe. Un cadre a été imposé par l'Union Européenne afin d'éviter les dérives de l'utilisation des données.

Il faut savoir qu'une directive « vie privée » existe seulement depuis 1998. Il s'agit de la directive 95/46/CE²⁸. Néanmoins, elle ne légifère pas sur le transfert de données entre l'Union Européenne et les Etats-Unis. C'est pourquoi le Safe Harbor a été mis en place. Très vite, cette législation a été contestée et les discussions ont été remises sur le tapis.

De ces discussions est né l'**EU/US Privacy Shield** qui vise à encadrer le transfert de données entre l'UE et les Etats-Unis. C'est une refonte du Safe Harbor avec une législation plus stricte.

« *Les mesures prises sont les suivantes :*

- *Les entreprises seront soumises à des obligations fermes, assorties d'une mise à exécution rigoureuse,*
- *Un accès par les autorités américaines étroitement encadré et transparent,*
- *Une protection effective des droits des citoyens de l'Union et plusieurs possibilités de recours,*
- *Mécanisme de réexamen annuel conjoint.*²⁹»

Dès la publication de cette nouvelle directive, le « groupe 29³⁰ » qui travaille sur le respect de la protection des données à caractère privé, a contesté le Privacy Shield. Il note qu'il y a une avancée mais que trop de zones d'ombres persistent. En effet, selon lui, il manque de « *garanties plus strictes concernant l'indépendance et le pouvoir du médiateur américain* »³¹.

25. <https://www.privacycommission.be/fr/que-sont-les-donnees-a-caractere-personnel>

26. http://www.europarl.europa.eu/aboutparliament/fr/displayFtu.html?ftulId=FTU_3.2.4.html

27. http://www.europarl.europa.eu/aboutparliament/fr/displayFtu.html?ftulId=FTU_3.2.4.html

28. <https://www.cnil.fr/fr/communiqu%C3%A9-29-publication-de-lavis-du-g29-sur-laccord-privacy-shield>

29. http://europa.eu/rapid/press-release_IP-16-433_fr.htm

30. Pour en apprendre plus sur le Groupe 29 : http://ec.europa.eu/justice/data-protection/article-29/index_fr.htm

31. <http://www.europaforum.public.lu/fr/actualites/2016/07/article29-privacy-shield/index.html>

La sécurité est également un point essentiel dans ce contexte. Pouvoir assurer au client que ses données personnelles seront protégées est essentiel. Ne pas pouvoir s'engager dans ce sens peut atteindre l'e-réputation (la réputation de l'entreprise sur internet, son image) de l'entreprise.

Au final, la transparence dans l'utilisation des données Big Data est fondamentale. Il ne peut que se développer dans le contexte d'une Europe qui souhaite protéger le transfert des données privées. Penser à la sécurisation des données et être au plus proche des lois européennes le plus tôt possible participent à la réalisation d'un projet d'analyse des données à grande échelle.

Il faut sensibiliser les entreprises à la législation qui encadre l'utilisation des données. Sensibiliser aussi les clients et citoyens sur l'impact de la publication de leurs données personnelles. Cette problématique est au cœur des discussions européennes. Dans le cadre du projet européen « Horizon 2020³² », une telle réflexion sur le Big Data nécessite des projets de recherches. Nous aborderons ce sujet dans la partie « le Big Data en Europe » (voir Le Big Data en Europe.)

Le CRIDS – UNAMUR

Le centre de Recherches en Information, Droit et Société se consacre à l'étude de questions liées à la société de l'information et ses problématiques sociétales, juridiques, éthiques et philosophiques. Le centre s'intéresse à des sujets tels que la vie privée, la propriété intellectuelle, l'e-gouvernement, les données volumineuses, la cyber-sécurité, la culture numérique, la gouvernance algorithmique, d'autres.

Créé en 2010 au sein de l'université de Namur (UNAMUR), le CRIDS est le résultat de la fusion entre le centre de recherche en droit informatique (le CRID), le centre d'évaluation des technologies interdisciplinaire (le CITA) et le « *Groupe de Recherche Interdisciplinaire en Communication et Internet* » (le GRICI).

Le CRIDS est composé d'une cinquantaine de chercheurs ayant pour la plupart une expérience professionnelle en rapport avec le monde la société de l'information.

Le CRIDS est divisé en sept unités de recherches :

- L'unité « libertés et société de l'information »
- L'unité « propriété intellectuelle »
- L'unité « E-commerce »
- L'unité « communication électronique »
- L'unité « communication et internet »
- L'unité « technologies et société »
- L'unité « e-gouvernement ».

Le CRIDS propose également des programmes de cours notamment sous la forme de master ou de certificat :

- Une spécialisation en Droit des Technologies de l'Information et de la Communication
- Un master de Business Analysis et Gouvernance IT (présenté dans ce rapport dans le chapitre formation)
- Un certificat en management de la sécurité des systèmes d'information
- Un certificat Interuniversitaire en Protection des données à caractère personnel (CRIDS ET ICHEC).

Pour plus d'informations : <http://www.crids.eu>

32. Pour plus d'information sur les présentations du plan Big Data à la Commission Européenne : <https://ec.europa.eu/digital-single-market/en/news/big-data-information-and-networking-days-horizon-2020-topics>

Le Big Data en Europe

3

Le Big Data a également attiré l'attention de la Commission Européenne en termes de recherche et de développement. Dans le contexte d'« Horizon 2020 » et plus précisément dans la partie spécifique au Big Data, la Commission a lancé un appel à projet sur la valorisation du Big Data. Elle investit 2,5 milliards d'euros dans le cadre d'un protocole d'accord avec PPP (partenariat public-privé). La Commission a découvert le potentiel d'expansion de l'économie des données. Cette économie ne cesse de se développer et les exemples d'utilisation du Big Data se multiplient. L'Europe a besoin de prendre le train en marche pour mettre en place des projets favorables à la communauté européenne. Le marché a en effet de beaux jours devant lui.

La Commission Européenne a fait le constat que le Big Data a un impact favorable sur l'économie d'un pays³³ :

- Augmentation de la productivité de 5-7 %,
- 100 000 nouveaux emplois potentiellement créés d'ici à 2020,
- Diminution de la consommation d'énergie de 10%,
- Amélioration des soins de santé,...

La Commission souhaite que les projets Big Data issus de « Horizon 2020 » se concentrent sur les besoins réels et la création d'un « écosystème de la donnée efficace³⁴ ». L'Europe doit réfléchir à une politique de régulation de la donnée, c'est-à-dire, créer un cadre d'utilisation sans aller dans l'utilisation de données privées à des fins commerciales. Ce dernier point est l'un des plus sensibles car il faut trouver l'équilibre entre le progrès de la « data économie » et le respect de l'éthique avec une législation sur l'utilisation des données à caractère privé.

Les sujets d'études cités comme pistes de réflexion sont notamment l'Open Data³⁵, les outils Big Data, l'infrastructure, le cadre légal.

Les domaines d'applications peuvent être la domotique, l'énergie, la mobilité, la télécom, la santé, et d'autres encore.

Le site *datalandscape*³⁶ présente les différents programmes initiés dans tous les pays d'Europe pour « Horizon 2020 ». On remarque dans la liste quelques projets belges.

33. http://europa.eu/rapid/press-release_IP-14-1129_fr.htm

34. http://europa.eu/rapid/press-release_IP-14-1129_fr.htm

35. OPEN DATA : voir glossaire

36. <http://www.datalandscape.eu/>

4 Le Big Data à Bruxelles

4.1 L'aspect politico-économique

Pour la région de Bruxelles Capitale, le sujet du Big Data et, plus particulièrement des Open Data, a son importance. La Région souhaite mettre en place une stratégie numérique pour faire de Bruxelles une Smart City³⁷. Elle a décidé de placer certaines technologies en avant comme les Données, la réalité virtuelle et l'Internet Of Things. L'objectif de la Région est de mettre à disposition un maximum de données publiques, via un portail donnant accès aux données de la ville de Bruxelles³⁸. Un autre objectif de Bruxelles est de mettre à disposition des start-ups et autres jeunes entreprises un incubateur public. Cet incubateur permettrait d'utiliser des technologies et un accompagnement d'experts.

Le vrai défi de Bruxelles est législatif et économique. La société doit valoriser les données et avoir un rôle de sensibilisation au travers de diverses actions. Ces actions peuvent se concrétiser par la création de clusters ou incubateurs et un message fort au niveau politique. On est encore très loin de l'utilisation par le grand public des données.

Mr Nicolas Harmel, attaché au Cabinet du Ministre Didier Gosuin, Ministre de l'Economie et de l'Emploi à Bruxelles déclare : Le secteur public doit montrer que la donnée peut servir. Il faut montrer l'impact positif du Big Data. D'autant plus qu'on utilise tous les jours le Big Data sans s'en rendre compte à travers nos smartphones, PC portable et autres.

Aujourd'hui, il y a un frein à la libération des données empêchant tout un chacun de les utiliser à 100%. Il faut que les administrations jouent le jeu et montrent au grand public le pouvoir positif du Big Data. Dans une ville comme Bruxelles, il y a de véritables enjeux dans l'utilisation de la donnée et de l'Open Data au travers de la mobilité, de la sécurité, etc. Tout cela peut être développé et des entreprises, grandes ou petites, peuvent y répondre.

4.2 Les entreprises et le Big Data

Dans le contexte du Big Data et son implantation en Belgique, on peut distinguer deux cas de figure.

- Les entreprises qui ont intégré ou qui pourraient intégrer dans leur département marketing ou communication une section dévolue au Big Data.
- Les sociétés spécialisées dans l'analyse des données qui proposent leurs services et une stratégie d'analyse Big Data aux autres entreprises. Ces dernières sont souvent des start-ups majoritairement basées à Bruxelles.

Dans ce rapport, nous avons déjà évoqué une attitude réticente au Big Data de la part des entreprises bruxelloises. Or, nous constatons que les sociétés de Data Analyse sont florissantes et ont beaucoup de clients. D'où tirent-elles leur part de marché ? Face à ce problème de sensibilisation des entreprises belges, les sociétés spécialisées en analyse des données se

37. SMART CITY : voir glossaire

38. Le portail : <http://opendata.bruxelles.be/page/home/>

tournent volontiers vers des clients étrangers. Bruxelles a l'avantage d'être la capitale de l'Europe, déclare Claude Delcour, administrateur et Chef de la direction chez Vadis.

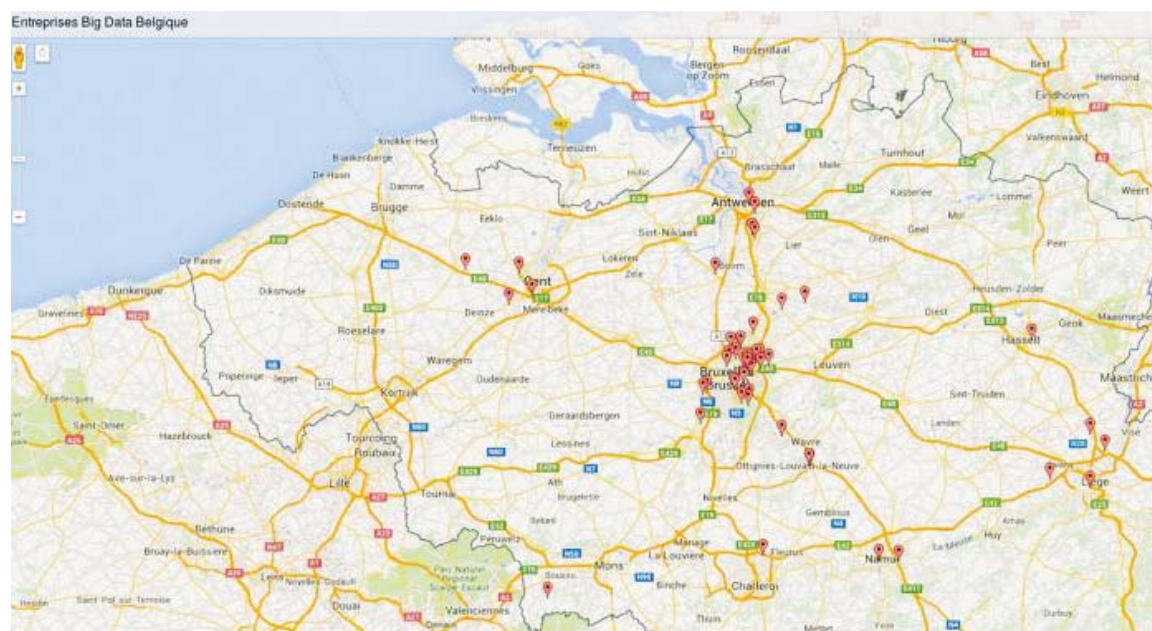


Figure 3 - Robert VISEUR; L'écosystème régional du Big Data. Retours d'expérience sur les technologies Big Data. Big Data Week 2015: Bridging data events all over the world – CETIC Gosselies, 24/11/2015.

« Des dizaines d'entreprises bruxelloises sont assises sur des données qu'elles ne valorisent pas », affirme Juan Bossicard de chez Software.brussels. C'est une préoccupation pour les sociétés d'analyse des datas. Pierre-Nicolas Schwab, responsable de la mise en place de la stratégie Big Data à la RTBF estime que l'une des raisons de la frilosité des entreprises belges au Big Data est « la méconnaissance du top management pour cette matière ».

« C'est un constat récurrent. Il faut communiquer ! Il faut démarcher les entreprises pour qu'elles aient conscience du potentiel des datas et des analyses qu'elles pourraient en tirer » commente Laurent Fayet, Managing Partner chez LBFayet Consulting³⁹.

Cette dualité entre ces start-up spécialisées dans l'analyse des données Big Data florissantes et ces entreprises dans les autres secteurs (industrie, médical, etc.) réticentes à se lancer dans le Big Data caractérise donc Bruxelles.

39. LBFayet Consulting : Voir glossaire

5

Les métiers du Big Data



L'économie de la data est en constante évolution. Le marché international ne cesse d'évoluer et de prendre de l'importance. Ainsi, on constate l'apparition de nouveaux profils d'emploi. Il est intéressant de s'attarder sur les profils métiers qui ont évolué et qui sont apparus grâce au secteur de la data.

Avant d'analyser les différents profils liés du Big Data, il est important de mettre le doigt sur une spécificité liée aux dénominations. Sur la toile, on voit beaucoup de profils différents apparaître : Data Scientist, Data analyste, Data manager, Chief Data Architect, Chief Technology Officer et d'autres encore. Difficile de se retrouver dans tous ces termes. D'autant plus que ces profils ont tous plus ou moins les mêmes compétences. Dans l'intérêt d'une analyse raisonnée, nous avons pris la décision de n'aborder dans ce rapport que les quelques noms les plus courants - c'est-à-dire : Data Architect, Data Scientist et Data Analyst. Bien entendu, il existe d'autres profils plus orientés Business ou technologies, etc., mais nous allons voir que ces trois profils résument bien les compétences générales demandées pour travailler dans le domaine des données.

Alors ! Est-ce que les métiers du Big Data sont sources d'emploi et seront sources d'emploi dans le futur à Bruxelles ? Cette question est finalement l'une des plus importantes de ce dossier. Par cette problématique, on s'interroge sur la pérennité du Big Data en termes de sources de revenus et d'innovation pour la Capitale. Au final, la grande question est : **Existe-t-il une demande des entreprises pour utiliser la science des données ?**

Oui et sans doute de plus en plus. Les entreprises ne peuvent pas passer à côté de cet outil qui entre désormais dans les pratiques courantes. Néanmoins en Belgique, l'impact du Big Data est encore peu retentissant. *A contrario*, on peut découvrir de plus en plus d'offres d'emploi en lien avec la data. L'analyse des données est donc une compétence de plus en plus recherchée par les entreprises et le sera plus encore dans les années à venir.

Au niveau international, les métiers de la data et plus précisément du Big Data ont le vent en poupe. En 2014, LinkedIn, cette grande base de données des professions a listé les métiers les plus tendances du moment. Il n'est pas surprenant d'y retrouver deux profils métiers liés au Big Data : le Data Scientist et le Big Data Architect⁴⁰. Néanmoins, dans la réalité belge, il faut nuancer. Bien que ces profils soient très demandés, il ne faut pas croire que le Big Data suscitera des milliers de recrutement. Pourquoi ?

- Tout d'abord, ce sont des profils très spécifiques qui demandent une expertise très poussée dans le monde des statistiques, des mathématiques, des technologies et du business. Pour prétendre à un poste de Data Scientist, un diplôme universitaire au minimum est nécessaire avec parfois une expérience professionnelle dans le monde de l'analyse des données.
- Ensuite, pour recruter des spécialistes belges de la data, il faudrait une prise de conscience de la part des institutions académiques et autres du potentiel du Big Data. Nous l'aborderons ci-dessous dans le chapitre consacré à la formation. Les programmes consacrés aux formations dans le domaine du Big Data existent mais sont très récents ou en projet et sont peu nombreux.

40. <https://business.linkedin.com/talent-solutions/blog/2014/01/top-10-job-titles-that-didnt-exist-5-years-ago-infographic>

- Dans ce contexte de sensibilisation, les entreprises doivent être informées sur la problématique du Big Data. Nous avons déjà discuté de ce sujet dans la partie « le Big Data à Bruxelles ».

On peut donc dire sans se tromper que le domaine de la science des données va générer de l'emploi car ses perspectives à Bruxelles sont très bonnes et en voie de maturation. D'autant plus que la Région Bruxelles-Capitale, dans sa visée Smart City, souhaite promouvoir une utilisation publique et positive des données. Les technologies seront présentes, il faut pouvoir engager des ressources humaines capables de les utiliser. Si on suit cette logique, on peut estimer qu'il y aura plus d'embauche.



5.1 Le Data Architect

Le travail du Data Architect consiste à organiser les données, les trier et les assembler. Il définit des standards permettant de faciliter la récolte et le tri⁴². C'est un profil plus orienté informatique que celui du Data Scientist. Il a donc besoin d'avoir de solides connaissances en base de données et des outils technologiques utilisés pour le Big Data (Hadoop, langage R, ..).

Lors de notre discussion, Laurent Fayet, Managing Partner chez LBFayet Consulting, divise ce profil métier en deux catégories de compétences :

1. « Un **profil orienté technologies pures**. C'est un ingénieur en infrastructure qui réfléchit aux technologies qui devront être utilisées dans la stratégie de récolte des données,
2. Un **profil orienté application**. Il s'occupe de la structure qui se charge de lier les données entre elles pour dégager un schéma. »

En termes de formation : le Data Architect doit avoir de solides connaissances en programmation et technologies des données. Les mathématiques et les statistiques sont nécessaires même si ce ne sont pas des compétences principales dans son profil métier. Il doit s'y connaître en data visualisation et en machine learning. Certaines soft skills sont nécessaires : curiosité par rapport aux données, passion pour le domaine d'application, communication, etc.

5.2 Le Data Scientist

Le Data Scientist est celui qui conçoit l'algorithme qui va traiter les données. C'est lui qui réalise des prédictions par rapport aux résultats qu'il va recevoir (analyse prédictive). On peut dire que c'est le profil métier le plus souvent évoqué quand on parle de Big Data. C'est un peu le pilier central du processus. Il est très recherché par les entreprises de data analyse.

41. "I'm a little surprised, with such extensive experience in predictive analysis, you should've known we wouldn't hire you."; "Je suis un peu étonnée qu'avec autant d'expérience en analyse prédictive, vous n'ayez pas deviné que nous n'allions pas vous engager. »

42. Fabrice Coudray, Focus sur un job en or : le Data Architect, Dans JDN [en ligne] <http://www.journaldunet.com/solutions/expert/59345/focus-sur-un-job-en-or---le-la-data-architect.shtml> (consulté le 21 juin 2016).

MODERN DATA SCIENTIST

Data Scientist, the sexiest job of the 21st century, requires a mixture of multidisciplinary skills ranging from an intersection of mathematics, statistics, computer science, communication and business. Finding a data scientist is hard. Finding people who understand who a data scientist is, is equally hard. So here is a little cheat sheet on who the modern data scientist really is.

MATH & STATISTICS

- ☆ Machine learning
- ☆ Statistical modeling
- ☆ Experiment design
- ☆ Bayesian inference
- ☆ Supervised learning: decision trees, random forests, logistic regression
- ☆ Unsupervised learning: clustering, dimensionality reduction
- ☆ Optimization: gradient descent and variants

PROGRAMMING & DATABASE

- ☆ Computer science fundamentals
- ☆ Scripting language e.g. Python
- ☆ Statistical computing packages, e.g., R
- ☆ Databases: SQL and NoSQL
- ☆ Relational algebra
- ☆ Parallel databases and parallel query processing
- ☆ MapReduce concepts
- ☆ Hadoop and Hive/Pig
- ☆ Custom reducers
- ☆ Experience with xaaS like AWS

DOMAIN KNOWLEDGE & SOFT SKILLS

- ☆ Passionate about the business
- ☆ Curious about data
- ☆ Influence without authority
- ☆ Hacker mindset
- ☆ Problem solver
- ☆ Strategic, proactive, creative, innovative and collaborative

COMMUNICATION & VISUALIZATION

- ☆ Able to engage with senior management
- ☆ Story telling skills
- ☆ Translate data-driven insights into decisions and actions
- ☆ Visual art design
- ☆ R packages like ggplot or lattice
- ☆ Knowledge of any of visualization tools e.g. Flare, D3.js, Tableau



MarketingDistillery.com is a group of practitioners in the area of e-commerce marketing. Our fields of expertise include: marketing strategy and optimization; customer tracking and on-site analytics; predictive analytics and econometrics; data warehousing and big data systems; marketing channel insights in Paid Search, SEO, Social, CRM and brand.

Marketing
DISTILLERY
(c) Krzysztof Zawadzki

David Majorel dans l'article « Salaire: Experts cybersécurité et data scientists sont les rois du pétrole » résume bien le profil :

"Le Data Scientist est le poste le plus demandé, car il est rarissime. On est sur des profils qui sont hyper pointus techniquement, sur les maths et les statistiques. Ce sont des profils rares car il n'y a pas vraiment de formation, sauf des ingénieurs spécialisés ou qui ont fait des écoles de stats. Ils doivent connaître beaucoup d'outils, mais aussi avoir une vision, être créatifs et tenter des choses⁴³."

Le Data Scientist est issu d'une formation en mathématiques et statistiques mais a aussi des compétences pointues en informatique (notamment pour concevoir les algorithmes). Il est souvent amené à faire des probabilités.

On a tendance à considérer le Data Scientist comme une évolution du Data Analyst et du Business Analyst⁴⁴. Il se distingue des profils métiers classiques de la data en analysant des données massives en provenance de sources diversifiées (donc Big Data).

5.3 Le Data Analyst

Souvent confondu avec le Data Scientist, il faut le considérer comme son pendant. Ils travaillent généralement en collaboration. Le Data Analyst extrait des schémas, des modèles dans les données récoltées, triées par le Data Architect et passées dans l'algorithme du Data Scientist. C'est souvent le rôle du Data Analyst de synthétiser les résultats, de communiquer avec la direction et de fournir des recommandations.

Le Data Analyst est issu d'une filière en statistiques : il maîtrise les statistiques, les maths, l'informatique comme les profils métiers précédemment cités. Il a également étudié la Business Intelligence (ou BI) et le marketing. C'est le spécialiste qui rend des comptes, il doit donc avoir des connaissances en communication marketing ; il doit savoir extraire une stratégie pour l'entreprise afin d'engendrer des bénéfices.

43. Julien Bergounhoux, Salaire : Experts cybersécurité et data scientists sont les rois du pétrole Dans l'Usine Digitale [en ligne] <http://www.usine-digitale.fr/editorial/salaire-experts-cybersecurite-et-data-scientists-sont-les-rois-du-petrole.N383849>

44. <http://www.kdnuggets.com/2016/01/modern-data-science-evolution-bi.html>

5.4 Des Disciplines transversales

Il est très réducteur de simplifier les métiers liés au Big Data aux trois profils métiers développés précédemment. Néanmoins, ce sont les plus récurrents dans les débats sur les métiers de l'analyse des données. On peut citer d'autres termes : Project manager, Chief Data Officer (le responsable des données⁴⁵), Chief Technology Officer (il s'occupe des questions techniques et scientifiques⁴⁶), etc.

5.5 Les compétences clés

Dans les profils développés, vous avez sans doute remarqué la récurrence de certaines compétences. Ces compétences sont essentielles pour travailler dans l'univers de la donnée.

- Les statistiques et les mathématiques
- L'informatique et la maîtrise de certaines technologies spécifiques (logiciel R, Map Reduce, Hadoop, etc),
- La data visualisation,
- Le marketing et la communication,
- Connaissance du domaine d'application : cette dernière compétence est moins évoquée précédemment mais il va de soi qu'un spécialiste de la donnée, qu'il soit Data Scientist, Data Analyste ou autre, doit être conscient des problématiques du secteur de l'entreprise pour concevoir une stratégie pertinente.

On notera enfin que le niveau de compétences est très élevé, universitaire voir plus avec une expérience dans le domaine du traitement des données. Il n'est pas impossible qu'un niveau bachelier puisse accéder à un poste dans le Big Data, peut-être data Architect, mais avant cela il faut que les mentalités changent et que le secteur de l'enseignement, et des Hautes Ecoles en particulier, prenne conscience du potentiel du Big Data. De plus, en complémentarité à ces profils élevés, on peut estimer qu'avec l'utilisation plus accrue des technologies du Big Data, au niveau public et privé, les administrations devront engager des profils de techniciens. Ces techniciens doivent être capables d'assurer l'installation et la maintenance des capteurs, capteurs et autres outils servant à recueillir les données. Ces profils spécifiques sont liés à l'IT et nécessitent parfois des compétences particulières. Néanmoins, dans la société en pleine évolution qui caractérise la Belgique, il est intéressant de se poser la question de l'intégration de cours d'informatique et de sensibilisation à l'informatique dans les cursus scolaires.

45. Définition complète : https://fr.wikipedia.org/wiki/Chief_Data_Officer

46. Définition complète : https://fr.wikipedia.org/wiki/Directeur_de_la_technologie

6

Des formations Big Data



Le constat est criant et unanime : en Belgique, on manque de profils métiers liés au Big Data! Pourquoi ? A ce jour, les formations sont peu nombreuses voire rares.

Hugues Bersini, responsable de l'IRIDIA, professeur de l'ULB et membre de l'Académie Royale des Sciences de Bruxelles résume bien le paysage académique Big Data en Belgique « *Il y a une demande importante de formation dans l'enseignement universitaire, Haute Ecole mais aussi des organismes comme EVOLIRIS. Le marché n'est pas encore tout à fait mature mais nous sommes en transition. On manifeste un besoin d'approcher le Big Data dans toutes les disciplines.* »

« *En terme de formations, il faut investir dans le marché de la data* », déclare Stéphane Faulkner, professeur à l'université de Namur et professionnel dans l'IT. On constate, ainsi, une lente prise de conscience des institutions académiques sur les enjeux du Big Data. De nouvelles formations sont en train d'éclore dans diverses institutions, d'autres sont en projet. Nous constatons également que certaines disciplines ont fait le choix d'inclure un ou plusieurs cours liés au Big Data dans le cursus de formation. Parmi ces initiatives, on retrouve des formations universitaires, Hautes Ecoles et des formations qualifiantes.

Dans ce chapitre du rapport, nous souhaitons présenter quelques exemples de formations proposées au sein de l'enseignement officiel ou non. Signalons cependant que cette liste est sans doute lacunaire et ne représente qu'un échantillon de ce que se retrouve sur le marché de la formation en Big Data.

6.1 Les formations scolaires



Nous allons vers une société de plus en plus informatisée et les profils liés à l'IT sont de plus en plus demandés. En dehors du débat sur la création de cours lié au Big Data au niveau supérieur, il est intéressant de débattre sur l'importance d'intégrer des cours d'informatique ou de sensibilisation à l'information en secondaire. À ce jour, il existe peu de cours d'informatique pour les jeunes dans l'enseignement officiel. Etant donné que les jeunes sont nés et ont grandi baignés dans la culture de la technologie, il semble important de les préparer à un avenir professionnel où la technologie aura une part importante. Cette sensibilisation peut se manifester de diverses manières : l'intégration de cours d'informatique qui vont au-delà de la sensibilisation aux données privées, et l'utilisation d'outils administratifs tels Windows Office. Une campagne de sensibilisation vers des initiatives privées et extra-scolaires tel que Coderdojo⁴⁷, les coding sessions du MIC⁴⁸, etc. et surtout la mise en place de formations destinées aux enseignants qui à ce jour sont intéressés de sensibiliser les étudiants à l'informatique mais n'ont pas les outils pédagogiques pour cela.

47. <http://www.coderdojobru.com/#!/home-fr/cv7i>

48. <http://www.mic-brussels.be/en/for-schools/coding-sessions-fr/>

Mr Tanguy De Lestré, Conseiller en Informatique et Transition numérique au Cabinet du Secrétaire d'Etat à la Région de Bruxelles-Capitale, explique lors de notre entretien: « De nos jours (et cela sera de plus en plus important au fil des années), il est important que tout le monde puisse avoir une base pour coder car l'économie mondiale et belge est de plus en plus virtuelle. »

6.2 Les formations universitaires et Hautes Ecoles

Lors de notre entretien avec M. Bersini, nous avons appris qu'un **master spécialisé en Big Data** était en préparation à l'ULB : ce master comprendra un tronc consacré aux mathématiques et aux statistiques, un autre consacré à l'informatique et aux technologies Big data, et des cours spécifiques en rapport avec des domaines d'application divers (marketing, économie, etc.). Ce master est toujours en préparation, l'ULB espère pouvoir l'inclure dans le programme de la rentrée 2016.

En Flandre, à l'université de Gand (UGENT), il existe un **master en «science de l'analyse Marketing»**. Ce programme est divisé en trois parties : le marketing et les recherches marketing, l'informatique, (notamment les bases de données) et enfin les statistiques et le Data Mining⁴⁹. C'est un master d'une année qui permet de découvrir le potentiel du Big Data dans le marketing⁵⁰.

Toujours en Flandre, la Haute Ecole « Thomas More » propose un **bachelier en « applied computer science»⁵¹** qui comprend des compétences en lien avec le Big Data : la structuration des données, le design et le management de banques de données, l'analyse orienté-objet et la communication de données. Lors de la première année de cours, les professeurs enseignent les bases de la connaissance et c'est à partir de la deuxième année du bachelier que les étudiants étudient la gestion des datas.

À l'UNamur, en Wallonie, il existe un master d'un an accessible au détenteur d'un autre master : « **Business Analyste et Gouvernance IT**⁵² ». Celui-ci propose quatre unités : Organisation et Innovation, Business Analysis (gestion de projet, ingénierie des exigences,), Gouvernance IT (BI, Droit des TIC, ...) et le mémoire. Ce master semble orienté business, marketing avec une orientation analyse de données. L'étudiant acquiert des compétences en analyse stratégique et en organisation.

À la faculté des sciences de l'UCL (Université Catholique de Louvain), le LSBA⁵³ et la plateforme SMCS⁵⁴ ont lancé un **certificat en « Junior Data Analyst »** depuis octobre 2016. Ce certificat est proposé aux demandeurs d'emploi et « jeunes diplômés » au niveau master ou aux personnes travaillant dans une entreprise et souhaitant acquérir de nouvelles compétences technologiques. Cette formation est donnée 2 à 3 jours par semaine d'octobre à décembre 2016. Elle est divisée en plusieurs modules :

- Analyse exploratoire et introduction aux enquêtes,
- Gestion de données comme partie intégrante du processus du traitement statistique,
- Rappels et renforcement en probabilité et statistique,
- Programmation statistiques,
- Analyse exploratoire multivariée,
- Modélisation,
- Ateliers.

49. Le data mining ou exploration de données : voir glossaire.

50. <http://www.mma.ugent.be/>

51. <http://www.thomasmore.be/welcome/bachelor-applied-computer-science>

52. <http://msbagi.unamur.be/presentation-du-ms-bagi/>

53. L'Ecole de statistique, biostatistiques et sciences actuariales

54. Plateforme technologiques de Support en Méthodologie et Calcul Statistiques

On remarque également cette prise de conscience dans l'inclusion de problématiques Big Data dans certains cursus :

- Un cours de « visualisation des données » dans le master Sciences et Technologies de l'Information et de la Communication à l'ULB,
- Un cours de « Business Intelligence and Data science » dans le master en ingénieur de gestion à Solvay, ULB⁵⁵,
- Etc, ...

6.3 Les formations qualifiantes

Dans le monde de la formation qualifiante, le Big Data fait également une percée. On relève notamment deux formations chez Technofutur à Charleroi.

Une formation qualifiante de 50 jours avec 6 semaines de stage en entreprise pour Data Analyst. Cette formation a été donnée pour la première fois en février 2016. Le public visé est un public de chercheurs d'emploi ayant un esprit logique et de fortes connaissances en math et en statistiques.

Le cours a été divisé en plusieurs modules :

- Une module de base : introduction aux datas, Excel, ...
- Un module de manipulation des données : SQL, ...
- Un module de datawarehousing⁵⁶ : BI⁵⁷, ...
- Un module d'analyse métier : UML⁵⁸, BPMN⁵⁹, ...
- Un module de reporting : visualisation des données,...
- Un module d'analyse des données : Data Mining, ...
- Et un laboratoire datawarehousing, BI, Data Mining

Comme indiqué sur le site web du centre de formation, le futur Data Analyst devrait à la fin de sa formation être capable « d'élaborer des rapports de visualisation » ... « *afin de mettre en évidence les résultats de ses analyses*⁶⁰ ». Il aura les capacités de comprendre les processus de l'entreprise.

Toujours chez Technofutur se déroule depuis le 13 juin 2016 une formation de 58 jours en **Big Data Technical Consultant**⁶¹. Cette formation est destinée aux chercheurs d'emploi ayant au moins un bachelier en informatique ou une expérience professionnelle de 3 ans dans le domaine de l'informatique de gestion. Le programme se divise en plusieurs modules :

- Un module d'introduction : Data Mining, cloud, gestion des données, stockage, traitement, etc.,
- Un module de modélisation, manipulation et stockage des données,
- Un module de BI,
- Un module de Data Mining,
- Un laboratoire final testant la capacité à trouver une solution Big Data,
- Un module de recherche d'emploi.

55. http://banssbfir.ulb.ac.be/PROD_frFR/bzscsrse.p_disp_course_detail?cat_term_in=201617&subj_code_in=GEST&crse_num_in=5123&PPAGE=ESC_PROG-CAT_AREREQ&PPROGCODE=MA-INGE&PAREA=M-INGEF&PARETERM=201516&PTERM=201617

56. Le datawarehouse : voir glossaire

57. Business Intelligence

58. UML : voir glossaire

59. BPMN : voir glossaire

60. <http://www.technofutur.be/Formation/Data-Analyst?ft=5&a=1&ihh=1&cp=0>

61. <http://www.technofutur.be/Formation/Big-Data-Technical-Consultant?ft=5&a=1&ihh=1&cp=1>

Le Data Innovation Hub : Data Innovation academy



Le Data Innovation Hub (DIH) de Bruxelles est un espace networking Big Data dirigé par Philippe Van Impe. Le DIH souhaite promouvoir le Big Data à Bruxelles en présentant le sujet et ses meilleurs pratiques. Il se veut le centre des différents acteurs de la donnée : entreprises, start-ups, politiques, institutions académiques, etc. et est également l'animateur d'une communauté; la *Data Science Community*, qui organise régulièrement des réunions entre membres pour des partages d'idées.

Le DIH organise des conférences, propose ses services aux entreprises, et met un espace de networking à disposition des start-ups. Il organise également le Data Innovation Summit tous les ans en invitant des acteurs dynamiques du secteur.

De plus, le Data Innovation Hub est également le fondateur de la di-Academy. Celle-ci propose une série de formations sur le Big Data et ses technologies accessibles aux employés et chercheurs d'emploi. Ces formations sont certifiantes. De même, des boot camps et hackathons sont organisés avec diverses sociétés partenaires.

Pour plus d'informations : <https://datasciencebe.com/> ou <http://di-academy.com/>

6.4 Autres initiatives

Sur internet, on trouve également toute une série de formations et cours non qualifiants. Ces cours et tutoriels sont parfois très poussés et permettent aux individus d'appréhender par eux-mêmes les caractéristiques techniques du Big Data. On peut citer la « **Big Data university**⁶² » qui propose des cours orientés exclusivement Big Data (les fondamentaux, Initiation à Hadoop, Map Reduce, ...).

Il existe également un **MOOC**⁶³ **d'initiation au Big Data** lancé en juin 2016 par UTT de Troyes en France⁶⁴.

Et il en existe encore d'autres sur la toile.

Enfin, certaines institutions (comme le DIH) organisent des Hackathon ou Boot camps afin de permettre aux candidats de se former de manière intensive et pratique à des outils technologiques de la science des données.

6.5 Enseigner le Big Data à Bruxelles

Toutes ces initiatives permettent de nous rendre compte que la prise de conscience du potentiel emploi du Big Data est bien réelle. On sent une volonté affirmée de produire des Data Spécialistes capables de travailler dans un environnement Big Data. Ces formations, qu'elles soient qualifiantes ou non, cherchent à introduire sur le marché de l'emploi belge des spécialistes maîtrisant les statistiques, les technologies d'analyse de la data, les bases de données, la programmation, la BI et le Data Mining.

62. <https://bigdatauniversity.com/>

63. MOOC : voir glossaire

64. UTT : <http://www.utt.fr/fr/toute-l-actualite/ouverture-mooc.html?platform=hootsuite>

Les avis sont partagés sur la mise en œuvre de ces cours. Que choisir entre la création d'un cursus « Big Data » ou l'intégration du Big Data dans un cursus en médecine par exemple ?

Selon Max De Wilde et Raphael Hubain, post-doctorant et doctorant à l'ULB dans la filière Sciences et technologies de l'Information et de la Communication : « *Il est moins intéressant de créer un master en data science que d'intégrer cette discipline dans le cadre d'un cursus comme par exemple en médecine, économie ou autre* ».

Le Big Data aurait donc plus d'impact pour certains si le cours est intégré à un domaine d'application en particulier : la médecine, l'économie, l'ingénierie, etc. C'est également un constat qui a été fait chez Evoliris lors de la table ronde sur le développement organisée en 2015 par Evoliris avec le concours d'experts en la matière : Ivan Jureta (chercheur au FNRS), Stéphane Faulkner⁶⁵, Bruno Schröder⁶⁶, et Nicolas Bauwens (conseiller pédagogique et formateur au pôle management et multimédia TIC à Bruxelles Formation).

Finalement, à l'instar de son utilisation en entreprise, le Big Data doit également avoir sa place dans le cursus scolaire. La Belgique et Bruxelles ne sont pas encore matures en matière d'enseignement Big Data mais il y a une volonté de faire avancer les choses. Comme nous l'avons vu, des initiatives très intéressantes sont développées un peu partout et à tous niveaux de compétences.

65. Voir partie remerciement

66. Voir partie remerciement

Conclusions et recommandations

7

Est-ce que le Big Data est une technologie tendance aujourd'hui ? Oui mais il évolue.

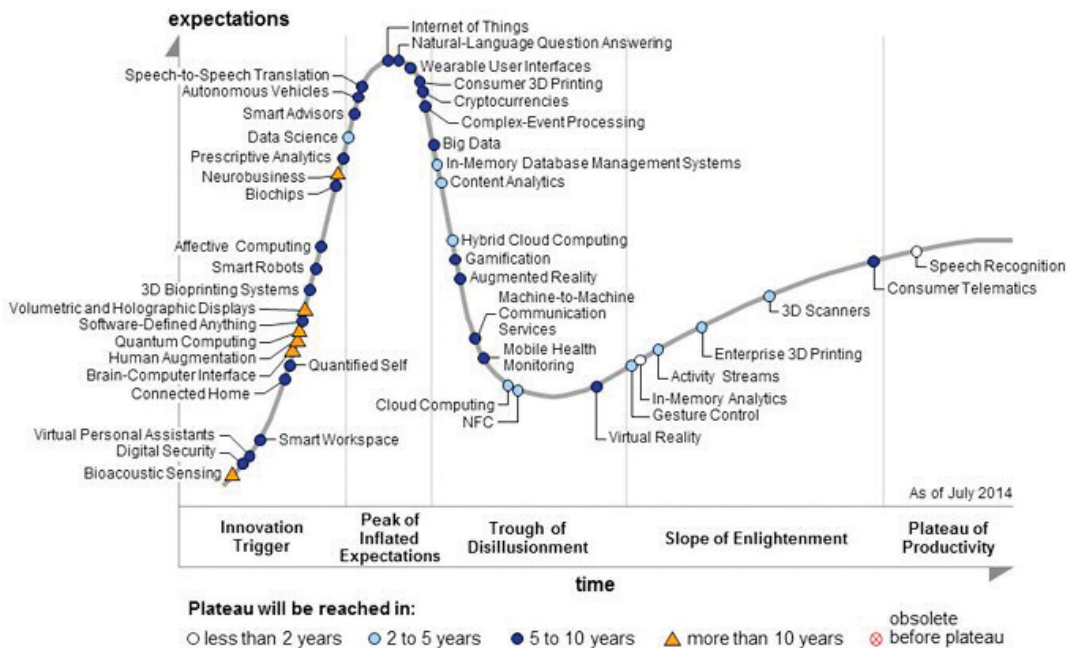


Figure 4: le 'Hype Cycle' 2014 de Gartner

Dans son Hype Cycle de 2014, Gartner positionne le Big Data dans la catégorie « *trough of Disillusionment* », proche de la catégorie « *peak of inflated Expectations* ». C'est donc un concept en pleine maturation. La société américaine estime que le « plateau de productivité » sera atteint dans 5 à 10 ans. Le Big Data a donc de beaux jours devant lui.

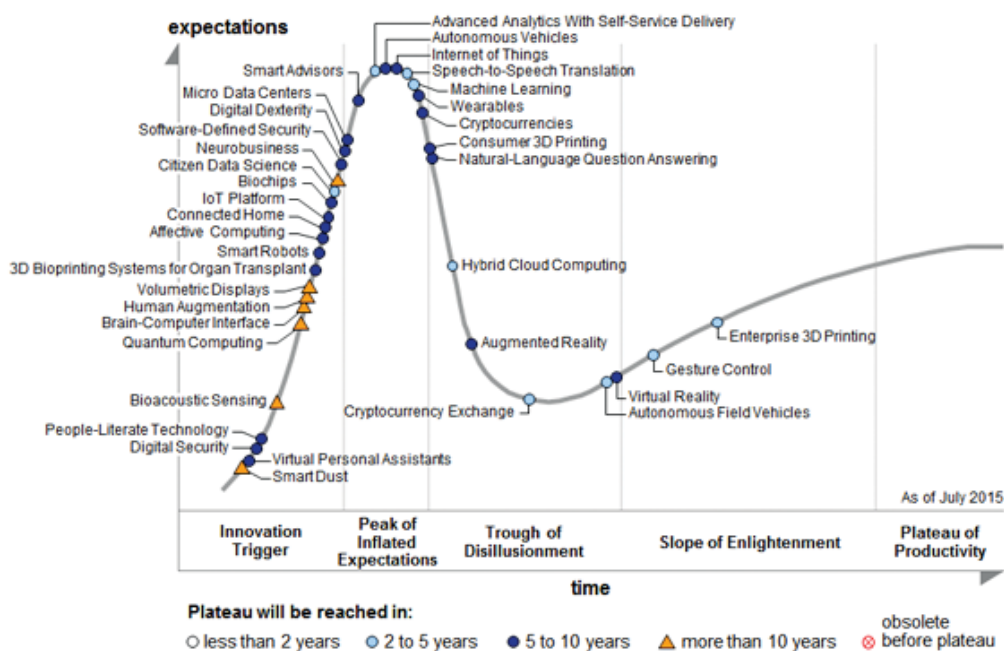


Figure 5: le 'Hype Cycle' 2015 de Gartner

Si le Big Data fait bien partie du Hype Cycle en 2014, force est de constater qu'en 2015 il brille par son absence. Pourquoi ? Betsy Burton⁶⁷, responsable de la réalisation du Hype Cycle chez Gartner, nous apprend que le Big Data n'a plus de raison de figurer dans ce schéma. Il est devenu une technologie à part entière qui figure dans d'autres cycles (plus uniquement dans celui qui représente les technologies tendances). Le Big Data n'est plus, en 2015, une technologie TENDANCE puisqu'il est désormais une technologie en pleine maturité au niveau international qui s'intègre parfaitement dans la société et dans tous les secteurs économiques d'un pays⁶⁸.

La maturité de cette technologie apporte une sécurité en termes d'investissements. Il y a moins de risques de s'y lancer. C'est désormais un outil très utile pour faire de Bruxelles une Smart City, une ville d'innovation. Les différents projets du CIRB sont des exemples de la volonté d'aller vers une Bruxelles intelligente⁶⁹.

Quelle est la situation à Bruxelles et en Belgique ?

Une ville comme Bruxelles est productrice de beaucoup de données. Lors de l'écriture de ce rapport, nous avons noté qu'un fil rouge, un point commun, relie tous les thèmes (entreprise, emploi, enseignement, juridique). Ce fil rouge se caractérise de manière assez cohérente dans la réalité du terrain par un manque de prise de conscience de la part des acteurs et utilisateurs de la data Analytique. Le marché du Big Data n'est pas encore tout à fait mature en Belgique et particulièrement à Bruxelles. On n'a cessé tout au long de ce rapport de mettre l'accent sur cette problématique. Pourquoi ? Parce qu'il y a un manque de sensibilisation sur les aspects positifs, et surtout sur ce qu'est véritablement le Big Data. Les entreprises bruxelloises n'ont pas encore saisi les opportunités qu'elles peuvent faire en l'utilisant. Beaucoup estiment à tort que cette technologie n'est qu'une affaire de grosses, très grosses sociétés travaillant principalement dans le marketing. Or, nous avons présenté une série d'exemples concrets d'utilisation dans des domaines d'activité n'ayant aucun lien avec le commerce. De plus, la capitale, dans sa volonté d'aller vers une Smart Bruxelles, souhaite sensibiliser sur le positif de l'utilisation des données au niveau public et privé. Avec du recul, on peut noter une incohérence. D'une part, le Big Data se développe bien à l'international, les entreprises de data analytiques belges et bruxelloises fleurissent et ont beaucoup de clients car elles travaillent avec des clients étrangers. D'autre part, les entreprises bruxelloises sont frileuses avec l'analyse des données à grande échelle. C'est une réalité préoccupante mais symptomatique de la situation ; les entreprises belges spécialisées dans l'analyse de données n'ont pas ou peu de clients belges. Pourtant, les métiers en lien avec le Big Data sont en plein boom. Les profils Data Analyst, Data Scientist s'arrachent à la pelle ; ils valent de l'or !

Il apparaît de plus en plus urgent que la tendance s'inverse et que les entreprises comprennent le bénéfice d'engager des spécialistes de la donnée. Malheureusement, on retient ce frein en Belgique : les spécialistes belges sont rares. Pourquoi ? Il existe à ce jour très peu de cursus intégrant les matières Big Data à Bruxelles et en Belgique. Les programmes intègrent trop peu de matières liées à l'analyse des données. Cependant, on constate une lente prise de conscience.

Finalement, tout est lié : enseignement, emploi, entreprises. Cette maturation doit venir de tous les fronts et de tous les acteurs du secteur pour avancer vers une technologie porteuse d'innovation et d'emploi. On peut espérer que d'ici 4-5 ans, avec la finalisation du projet

67. VIDEO : explication de Betsy Burton sur le cycle de Hype en 2015 : <http://www.gartner.com/technology/research/hype-cycles> .

68. Alex Woodie, Why Gartner Dropped Big Data Off the Hype Curve Dans Datanami [en ligne] <http://www.datanami.com/2015/08/26/why-gartner-dropped-big-data-off-the-hype-curve/> (consulté le 17 juin 2016).

69. Pour plus d'information sur ce projet et d'autres : consulter le cahier du CIRB 35 : <http://cirb.brussels/fr/quoi-de-neuf/publications/cahiers/nr-35-4-projets-cles-de-smartcity-brussels>

européen « Horizon 2020 », des formations sérieuses en Big Data approvisionneront le monde professionnel de l'emploi avec des travailleurs compétents Big Data.

En conclusion, on peut être sûr que le Big Data a un bel avenir à Bruxelles. C'est un outil très efficace qui a déjà fait ses preuves partout dans le monde. Cette technologie d'avenir va se développer et s'adapter aux contraintes du temps.

7.1 Recommandations

Nous pouvons distinguer trois axes de recommandations.

7.1.1 La sensibilisation des entreprises et du grand public

Le Big Data est plus qu'un outil destiné aux plus grandes entreprises telles que Facebook et Google. Les entreprises bruxelloises, grosses ou petites, peuvent en profiter également. C'est pourquoi la science des données s'adapte aux caractéristiques de chaque entreprise. Il faut arrêter d'avoir peur du Big Data et dépasser le phénomène de mode qui lui colle à la peau. Se renseigner sur des cas concrets belges et sur la réalité de terrain est important et permet de prendre conscience de l'intérêt.

« L'analytics c'est l'augmentation des revenus et la diminution des coûts. Ce sont des données auxquelles les entreprises ne pensent pas et dont elles font une mauvaise utilisation : le Big Data permet de regarder différemment », Laurent Fayet, pour LBFayet Consulting⁷⁰ lors de notre entretien.

La sensibilisation passe également par une discussion avec les entreprises. L'approche consiste à expliquer le Big Data et à donner des cas concrets et simples d'application. Il faut sortir des débats complexes et exposer simplement et efficacement la réalité du sujet. Les entreprises veulent savoir deux choses :

- À quoi va me servir le Big Data dans mon entreprise qui fonctionne déjà très bien sans ?
- Est-ce que cela va me demander beaucoup d'investissement pour le mettre en place ?

Il faut répondre à ces questions de la manière la plus simple et adaptée aux besoins. Il est donc nécessaire de travailler au cas par cas. Le Big Data c'est aussi ça : exprimer ses besoins individuels et y répondre par un algorithme et une infrastructure qui apportent des réponses et des solutions stratégiques.

Il faut également sensibiliser le grand public sur l'aspect positif de la science des données, lui montrer des exemples de cas concrets ayant permis d'améliorer un aspect de leur vie grâce au Big Data. Une communication doit être mise en place pour montrer que le grand public est déjà touché par le Big Data dans son quotidien. Cette communication doit être axée sur une explication concrète des points positifs et négatifs de cette technologie.

7.1.2 Former plus de spécialistes des données en Belgique

Il faut initier une série d'initiatives pour « apprendre » le Big Data. Créer un cursus Big Data ou intégrer cette matière dans un cursus. Le Big Data doit être une matière qu'il ne faut pas négliger. On demande des Data Scientist et les entreprises vont chercher leurs spécialistes à l'étranger. Claude Delcour, fondateur de l'entreprise Vadis⁷¹, insiste sur le fait qu'il y a actuellement **en Belgique un problème important pour trouver des ressources humaines pour travailler dans le domaine des données**. Il semble également plus qu'important de

70. <http://www.lbfc.eu/>

71. <http://www.vadis.com/>

mettre en valeur les cursus déjà présents sur le marché de la formation. Présenter les avantages en termes d'emploi de ces cours et communiquer à propos de ces formations. Néanmoins, le Big Data ne doit pas seulement figurer dans les cursus d'informaticiens, ingénieurs, ou mathématiciens. En effet, il faut qu'il s'intègre dans d'autres disciplines comme l'économie, la médecine, la gestion culturelle, etc. Cela va faciliter l'intégration et la prise de conscience de son potentiel.

En sensibilisant aux problématiques du Big Data, il est possible de former des niveaux bacheliers à travailler dans le futur sur cette thématique. Actuellement, nous l'avons vu, seuls des niveaux très qualifiés sont demandés.

7.1.3 Sensibiliser les entreprises sur l'aspect juridique

Finalement, il faut faire attention à l'aspect juridique et bien se renseigner dans quoi on s'engage en utilisant des données. La transparence est fondamentale et permet d'être efficace, fiable et intègre. Il faut également sensibiliser sur la notion de données personnelles. Le grand public a tendance à associer le Big Data à cette idée de grandes compagnies utilisant les données personnelles des utilisateurs pour faire du profit. À travers des campagnes de communication, les administrations doivent expliquer que les données recueillies sont dépersonnalisées et peuvent parfois aider à résoudre un problème d'utilité public comme au niveau de la mobilité ou de la sécurité.

8

Ferdinand Cassier

Jeroen Franssen

Juan Bossicard

Max de Wilde

Raphael Hubain

 34

Céline Vanderborght

CIRB - <http://cirb.brussels/>

Smart City manager au CIRB

Pour en savoir plus :

https://www.linkedin.com/in/c%C3%A9line-vanderborght95a06254?authType=NAME_SEARCH&authToken=gchN&locale=en_US&trk=tyah&trkInfo=clickedVertical%3Amynetwork%2CclickedEntityId%3A191232214%2CauthType%3ANAME_SEARCH%2Cid%3A1-1-1%2CtarId%3A1468504449829%2Ctas%3Ac%C3%A9line%20vander

Laurent Fayet

LBFayet Consulting - <http://www.lbfc.eu/>

"Managing Partner" chez LBFayet Consulting

Pour en savoir plus :

https://www.linkedin.com/in/laurentfayet?authType=NAME_SEARCH&authToken=GQkg&locale=en_US&trk=tyah&trkInfo=clickedVertical%3Amynetwork%2CentityType%3AentityHistoryName%2CclickedEntityId%3Amynetwork_7127546%2Cid%3A

Claude Delcour

VADIS - www.vadis.com

Administrateur et Chef de la Direction chez Vadis

Maitre Philippe Laurent

Cabinet MVVP - <http://www.mvvp.be/>

LegalICT - <http://www.legalict.be>

Avocat spécialisé en droit des technologies de l'information et de la propriété intellectuelle – IP/IT & Commercial Lawyer chez MVVP – formateur chez Legalict.be

Pour en savoir plus :

https://www.linkedin.com/in/phlaurent?authType=NAME_SEARCH&authToken=bZ4r&locale=en_US&trk=tyah&trkInfo=clickedVertical%3Amynetwork%2CclickedEntityId%3A11788190%2CauthType%3ANAME_SEARCH%2Cid%3A1-3-3%2CtarId%3A1468505304990%2Ctas%3Aphili

Maitre Jean-Marc Van Gyseghem

Cabinet Paul Muylaert Lan Officer – CRIDS

<http://www.paulmuylaert.be/home.php>

<http://www.crids.eu/>

Avocat spécialiste des technologies de l'information et de la communication. Chercheur au CRIDS dans l'unité de recherche « Liberties in the information society ».

Pour en savoir plus : <https://directory.unamur.be/staff/jmvangys>

<http://www.rawlingsgiles.be/team/jean-marc-van-gyseghem/>

Hugues Bersini

ULB – Académie royale de Science de Bruxelles – IRIDIA (laboratoire de recherche sur l'intelligence artificielle) - <http://iridia.ulb.ac.be/bersini/>

Responsable de l'IRIDIA, professeur de l'ULB et membre de l'académie royale de science de Bruxelles

Pour en savoir plus : https://fr.wikipedia.org/wiki/Hugues_Bersini

Cécile Jabaudon

MIC - <http://www.mic-brussels.be/en/home/>

Operational Director au MIC

Pour en savoir plus : https://www.linkedin.com/in/c%C3%A9cile-jabaudon-915a914a?authType=NAME_SEARCH&authToken=P6Hd&locale=en_US&trk=tyah&trkInfo=clickedVertical%3Amynetwork%2CclickedEntityId%3A175865057%2CauthType%3A%2CNAME_SEARCH%2Cid%3A1-1-1%2CtarId%3A1468506565728%2Ctas%3Ac%C3%A9cile%20jabaudon

Stephane Faulkner

Technofutur, Cognitic, professeur à UNamur
<http://www.technofutur.be/> - <http://cognitic.be/>

Professor in Technologies and Information Systems Engineering

Pierre-Nicolas Schwab

RTBf

Fondateur IntotheMinds, Big Data / CRM Manager à la RTBf

Pour en savoir plus : https://www.linkedin.com/in/pnschwab?authType=NAME_SEARCH&authToken=qBp4&locale=en_US&trk=tyah&trkInfo=clickedVertical%3Amynetwork%2CclickedEntityId%3A835527%2CauthType%3A%2CNAME_SEARCH%2Cid%3A1-1-1%2CtarId%3A1468507366634%2Ctas%3APierre-Nicolas%20Schwab

Bruno Schröder

Microsoft, MIC, - <http://www.mic-brussels.be/en/home/>

Microsoft Technology Officer Belgium and Luxembourg, Directeur technologique du MIC Belgium ; administrateur des Innovation Centers

Pour en savoir plus : https://www.linkedin.com/in/bruno-schroder-bb57481?authType=NAME_SEARCH&authToken=PS49&locale=en_US&trk=tyah&trkInfo=clickedVertical%3Amynetwork%2CclickedEntityId%3A4826297%2CauthType%3A%2CNAME_SEARCH%2Cid%3A1-1-1%2CtarId%3A1468507330896%2Ctas%3ABruno%20sch

Géraud Guilloud

NCP-Wallonie - <http://www.ncpwallonie.be/fr/>

European Project Developer
Horizon 2020 ICT & Security NCP

Pour en savoir plus : <http://www.ncpwallonie.be/fr/equipe>

Philippe Van Impe

DIH
<https://datasciencebe.com/>
<http://di-academy.com/>

Fondateur du DIH et de la data science community à Bruxelles

Pour en savoir plus : <https://datasciencebe.com/>

Tanguy De Lestré

Cabinet de la Secrétaire d'Etat à la Région Bruxelles-Capitale Bianca Debaets

<http://be.brussels/a-propos-de-la-region/le-gouvernement-regional/bianca-debaets>

Conseiller en informatique et transition numérique

Pour en savoir plus : <https://www.linkedin.com/in/delestre/fr>

Nicolas Harmel

Cabinet du Ministre Didier Gosuin, Ministre de L'Economie, l'Emploi de Bruxelles

<http://be.brussels/a-propos-de-la-region/le-gouvernement-regional/didier-gosuin>

Attaché au Cabinet du Ministre Didier Gosuin

Pour en savoir plus : <http://be.brussels/a-propos-de-la-region/le-gouvernement-regional/didier-gosuin>

BPMN : Le Business Process Model and Notation est un modèle de processus métier et une notation graphique standardisée pour modéliser le savoir-faire d'une organisation à travers l'approche processus. https://fr.wikipedia.org/wiki/Business_Process_Model_and_Notation

CIRB : Centre d'Informatique pour la Région Bruxelles-Capitale.

Le site web : <http://cirb.brussels/>

DATA MINING : a pour objet l'extraction d'un savoir ou d'une connaissance à partir de grandes quantités de données, par des méthodes automatiques ou semi-automatiques. https://fr.wikipedia.org/wiki/Exploration_de_donn%C3%A9es

DATAWAREHOUSE : qu'on peut traduire par "entrepôt de données" désigne une base de données utilisée pour collecter, ordonner et stocker des informations provenant de bases de données opérationnelles et fournir ainsi un socle à l'aide à la décision en entreprise. https://fr.wikipedia.org/wiki/Entrep%C3%B4t_de_donn%C3%A9es

Gartner : Gartner est une société qui a su s'imposer en mentor fournissant des conseils sur des technologies avancées. Gartner est particulièrement connue pour son « Hype Cycle » (analyse des tendances technologiques par année sous la forme d'une courbe de vie) et son « Magic Quadrant » (analyse des atouts des meilleures sociétés travaillant dans le domaine d'une certaine technologie de pointe en comparant l'exhaustivité de la vision et la capacité d'exécution des technologies). Le site web : <http://www.gartner.com/technology/home.jsp>

IOT : contraction de "Internet of Things" ou "Internet des objets" représente l'extension d'Internet à des choses et à des lieux du monde physique. L'internet des objets connectés représente les échanges d'informations et de données provenant de dispositifs présents dans le monde réel vers le réseau Internet. Ex : frigo intelligent, montres connectées,... https://fr.wikipedia.org/wiki/Internet_des_objets

IRIDIA : Institut de Recherches Interdisciplinaires et de Développement en Intelligence Artificielle à l'Université Libre de Bruxelles. Le site web : <http://iridia.ulb.ac.be/bersini/>

LBFayet Consulting : services de conseils en définition et implémentation de stratégies Big Data et d'activités de placement de personnel spécialisé en Big Data.

Le site web : <http://www.lbfc.eu/>

Microsoft Azure Data Lake : outil permettant de stocker tout type de données aux formats différents dans un référentiel unique.

Microsoft Cortana Analytics : analyse les informations à l'aide d'un algorithme intelligent issu du Machine Learning

MOOC: *massive open online course*. Le MOOC est un cours ouvert et massif en ligne que l'on peut associer à une formation ouverte à distance.

OPEN DATA : ou *donnée ouverte*, est une donnée numérique d'origine publique ou privée. Elle peut être notamment produite par une collectivité, un service public (éventuellement délégué) ou une entreprise. Elle est diffusée de manière structurée selon une méthode et

une licence ouverte garantissant son libre accès et sa réutilisation par tous, sans restriction technique, juridique ou financière. http://fr.wikipedia.org/wiki/Open_data

Smart City (ville intelligente) : désigne une ville utilisant les technologies de l'information et de la communication (TIC) pour « améliorer » la qualité des services urbains ou encore réduire ses coûts. https://fr.wikipedia.org/wiki/Ville_intelligente

UML : Le langage de modélisation unifié, de l'anglais Unified Modeling Language (UML), est un langage de modélisation graphique à base de pictogrammes conçu pour fournir une méthode normalisée pour visualiser la conception d'un système. Il est couramment utilisé en développement logiciel et en conception orientée objet.
[https://fr.wikipedia.org/wiki/UML_\(informatique\)](https://fr.wikipedia.org/wiki/UML_(informatique))

Rédacteur

Christina Galouzis

Post-éditeur

Luc Huygh

Éditeur responsable

Jean-Pierre Rucci

Conception graphique

www.sergeantpaper.be

Impression

DADDY KATE

Albert Van Cotthemstraat 54

1600 Sint-Pieters-Leeuw

www.daddykate.be

Contact

Evoliris Asbl

Rue de la Borne, 14 A

1080 Bruxelles

02/475 20 00

info@evoliris.be

www.evoliris.be



- Introduction et méthodologie
- Le Big Data : délimitation du sujet et origine
- Le Big Data en Europe
- Le Big Data à Bruxelles
- Les métiers du Big Data
- Des formations Big Data
- Conclusions et recommandations