

EVOLIRIS ZOOMT IN OP ...



Big Data



in Brussel vandaag
En morgen ?





"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. Bron: <https://whatsthebigdata.com/2013/06/03/big-data-quotes/>



Inhoudsopgave

Voorwoord	7
1. Inleiding en methodologie	9
1.1 Methodologie van het rapport	10
2. Big Data: afbakening van het onderwerp en herkomst	10
2.1 Wat verstaan we onder Big Data?	10
2.2 Troeven van Big Data	14
2.3 De beperkingen van Big Data	15
2.4 Het juridische aspect van Big Data	15
3. Big Data in Europa	18
4. Big Data in Brussel	19
4.1 Het politiek-economisch aspect	19
4.2 De bedrijven en Big Data	19
5. Big Data-beroepen	21
5.1 De Data Architect	22
5.2 De Data Scientist	23
5.3 De Data Analyst	24
5.4 Transversale disciplines	24
5.5 De voornaamste competenties	24
6. Big Data-opleidingen	25
6.1 Schoolopleidingen	25
6.2 Opleidingen aan universiteiten en hogescholen	26
6.3 Kwalificerende opleidingen	27
6.4 Andere initiatieven	28
6.5 Big Data onderwijzen in Brussel	28
7. Besluiten en aanbevelingen	30
7.1 Aanbevelingen	32
7.1.1 Sensibiliseren van ondernemingen en van het grote publiek	32
7.1.2 Meer gegevensspecialisten opleiden in België	33
7.1.3 Bedrijven bewust maken van het juridisch aspect	33
8. Dankwoord	34
9. Glossarium	38
10. Evoliris	40

Voorwoord

De wereld van de technologie evolueert razendsnel. Als Brussels ICT beroepenreferentiecentrum bestaat de opdracht van Evoliris erin om de technologische tendensen te volgen, ze te analyseren en zich te positioneren in de Brusselse IT-sector. Deze opdracht richt zich zowel tot de actoren van de IT-wereld, als tot de werkzoekenden.

Binnen de context van een alsmaar meer geïnformatiseerde maatschappij moet Evoliris de actualiteit analyseren op nationaal niveau, maar ook en vooral op Brussels niveau, om zodoende de trends te bepalen in de Brusselse IT-bedrijven, en om aan de job- en opleidingsactoren jobpistes en aanbevelingen voor te stellen of te suggereren.

Er wordt gekeken naar wat er binnen de sector gebeurt en dit brengen we onder in het rapport dat u nu voor ogen hebt. Het doel van dit document bestaat erin een technologische tendens te beschrijven waarover binnen de Brussels IT-wereld momenteel veel wordt gesproken: de Big Data.

Data sciences, massale gegevens, Big Data; ze worden steeds vaker beschouwd als het zwarte goud voor ondernemingen die hun economisch potentieel willen ontwikkelen. Een politiek en publiek debat is in dit kader echt op zijn plaats. Big Data is trouwens ook aanwezig bij de overheidsinstanties en biedt de mogelijkheid om het dagelijks leven van de burgers te verbeteren.

Met onderhavig document willen we tonen wat er leeft en willen we zo dicht mogelijk bij de realiteit op het terrein staan. Daarom hebben we de mening gevraagd van verschillende deskundigen over hun praktische en dagelijkse kennis van Big Data. Vanaf het begin van dit project wilden we dit uit het theoretische kader halen, en een realistische analyse tonen van de situatie in Brussel. In die zin hebben wij een dossier samengesteld dat de standpunten van specialisten weergeeft op politiek-economisch en wettelijk vlak, om zo de vooruitzichten te tonen in het kader van de tewerkstellings- en opleidingsproblematiek in Brussel. De sterke troef van deze aanpak bestaat erin om de band tussen de verschillende partijen te consolideren en Evoliris te positioneren als bemiddelaar en stimulerende partij.

Wij houden eraan iedereen te bedanken die heeft meegewerkt aan dit rapport, want zonder hun expertise zou het nooit het huidige niveau gehaald hebben.

U hebt dus het eerste nummer in handen van een reeks rapporten die in de toekomst zullen volgen en die wij 'Evoliris zoomt in op...' noemen. Deze editie gaat over Big Data, de volgende zullen andere IT-thema's en trends van de Belgische markt analyseren.



Jean-Pierre Rucci
Directeur van Evoliris



Floriane de Kerchove
Voorzitter van Evoliris & directeur van Agoria

1

Inleiding en methodologie

Om beter te kunnen inspelen op de toeristenstroom gebruikt de stad Brussel Big Data om zijn bezoekers beter te leren kennen. Elk jaar onthaalt de stad Brussel 5,1 miljoen toeristen,



en 1,5 miljoen vakantiegangers². Om de stroom bezoekers te meten heeft Brussel beslist om geolokalisatie-diensten te gebruiken. Met de steun van Proximus heeft Visit.brussel de telefoongegevens geanalyseerd van 10 Brusselse toeristische trekpleisters. Deze gedepersonaliseerde gegevens hebben de mogelijkheid geboden om de logica van de toerist te begrijpen, en om in het bijzonder de bezoektijd van een plaats te berekenen en de plaats waar de bezoekers bij voorkeur overnachten. Vanuit een meer algemeen standpunt, hebben deze gegevens de stad Brussel de mogelijkheid gegeven om deze stroom beter te beheren, en om de diensten te verbeteren van de evenementenorganisatoren³.

Dit voorbeeld is een mooie toepassing van het gebruik van Big Data-gegevens om zo het dagelijks leven van een groep mensen te verbeteren. Er bestaan tal van andere gevallen, zowel publiek als privé, die aantonen dat Big Data een belangrijke aanzet kunnen vormen om zaken te verbeteren, wat contrasteert met de negatieve visie die men heeft omtrent Big Data.

Big Data of *de wetenschap van de gegevens* is een ingewikkeld en boeiend onderwerp. Er zijn meerdere toepassingen mogelijk en big data bieden bedrijven en instellingen de mogelijkheid om een efficiënte strategie uit te werken om hun rendabiliteit en hun acties te verbeteren. De term 'Big Data' is al verschillende jaren heel populair. De toepassing ervan wordt dus ook alsmat nuttiger. Maar de Big Data-gemeenschap is niettemin verdeeld: er zijn mensen die big data verheerlijken en anderen zijn van mening dat dit niets meer is dan een logische evolutie van de gegevensanalyse. Het komt ons in dit rapport niet toe om in deze een beslissing te nemen.

Dit rapport is bedoeld om deze technologie te kaderen in de Brusselse context. We zullen dit concept interpreteren in de socio-economische context van Brussel, maar ook inzake tewerkstelling en opleiding. Wij zullen eveneens de impact analyseren op de Brusselse bedrijven, binnen de Europese Commissie, evenals op juridisch vlak.

Big Data trekt al lang de aandacht van grote bedrijven die grote hoeveelheden gegevens moeten combineren. Letterlijk vertaald, betekent 'Big Data' ook 'massagegegevens'. Maar Big Data beperken zich niet tot massagegegevens. Wij zullen later zien dat het niet nodig is om over grote hoeveelheden data te beschikken om van Big Data te spreken. In dit kader is het essentieel om zich de vraag te stellen wat het belang is van Big Data, en welke toekomst er voor Big Data is weggelegd in Brussel.

2. http://www.proximus.be/nl/id_b_cl_tourism/grote-bedrijven-en-overheden/discover/blog/one-magazine/klantenverhalen/het-toerisme-.html
3. http://www.proximus.be/nl/id_b_cl_tourism/grote-bedrijven-en-overheden/discover/blog/one-magazine/klantenverhalen/het-toerisme-.html

1.1 Methodologie van het rapport

De voorbereiding van dit rapport werd opgedeeld in twee fasen.

Een eerste fase bestond erin om een *peiling* te doen over het onderwerp via onderzoek. De bedoeling van deze fase bestond erin om het kader van het rapport af te bakenen: het concept begrijpen en definiëren, weten wie men moet contacteren voor de gespreksfase, enz.

Nadat we het onderzoek hebben afgebakend, zijn we gestart met de *gespreksfase*. Om een zo goed mogelijk overzicht te hebben van de realiteit op het terrein, leek het ons essentieel om ons onderzoek aan te vullen met interviews met Big Data-specialisten. Deze specialisten komen uit de bedrijfswereld, maar het zijn ook tewerkstellings- en opleidingsactoren, enz. Voor meer informatie over deze mensen kunt u terecht op de pagina met het dankwoord.

Maar deze aanpak levert niet meer dan een glimp van de eigenlijke situatie. Wij hebben geprobeerd om te komen tot een analyse die zo dicht mogelijk aanleunt bij de realiteit door vragen te stellen aan actoren uit de sector die voldoende ervaring hebben inzake Big Data en de gegevenswetenschap.

Tijdens ons onderzoek hebben wij iets vastgesteld dat eigen is aan Brussel: Big Data is heel aantrekkelijk maar weinig bedrijven durven zich te lanceren. **Er is een echt sensibiliseringsprobleem.** Veel bedrijven denken ten onrechte dat Big Data niet interessant zijn voor hun marketingstrategie. We stellen ook een aanzienlijke vertraging vast inzake opleidingen, ondanks alsmatig talrijker initiatieven en een tekort aan Belgische Big Data-specialisten.

Big Data : afbakening van het onderwerp en herkomst

2

2.1 Wat verstaan we onder Big Data?

Meestal vertalen wij deze technologie met *het geheel van enerzijds de gegevens die werden ingezameld, gesorteerd, verwerkt en geanalyseerd teneinde te komen tot een analyse, en anderzijds aanbevelingen inzake strategie.*

De bedoeling van dit rapport is niet een uitgebreide definitie te geven van het concept Big Data. Het is wel belangrijk om van meet af aan van wal te steken met een korte definitie om de problematiek aan te halen die ermee verband houdt.

Big Data worden vaak vermeld in de context van de strijd tegen het terrorisme. Big Data-technologieën kunnen immers gebruikt worden om de veiligheid van de gemeenschap te verbeteren. In die zin heeft Microsoft tools ontwikkeld (Microsoft Azure

Data Lake⁴, Microsoft Cortana Analytics⁵) die de mogelijkheid bieden om een voorspellende analyse toe te passen inzake criminaliteit. Deze hoogtechnologische tools bieden de mogelijkheid om gedetailleerde informatie in te zamelen om een crimineel profiel tot stand te brengen. We noemen deze technologie: de voorspellende analyse⁶. Terrorisme is ook een onderwerp dat heel gevoelig ligt in België. Big Data bieden de onderzoekers de mogelijkheid om potentieel belangrijke informatie te analyseren en te sorteren door telefoongesprekken te onderzoeken, gegevens afkomstig van Smartphones, door verplaatsingen te analyseren van mogelijke terroristen met behulp van openbare camerabeelden, enz. Bovendien heeft het CIBG⁷ een *'platform in het leven geroepen voor het uitwisselen van videobeschermingsbeelden. Deze beelden zijn afkomstig van 6 Brusselse politiezones, van de Maatschappij voor het Intercommunaal Vervoer te Brussel (MIVB), de Dienst voor Brandbestrijding en Dringende Medische Hulp (DBDMH), Brussel-Mobiliteit en de Haven van Brussel*⁸.

Op het eerste gezicht heeft men de neiging om Big Data te associëren met **marketing**. Het is inderdaad de eerste toepassing waaraan men denkt. Big Data wordt vaak gebruikt binnen deze context en heeft grote bedrijven de mogelijkheid geboden om de noden van hun klanten te analyseren en om een doelgerichte communicatiestrategie aan te nemen. Door een hele reeks informatie in te zamelen over klanten via verschillende bronnen (sociale netwerken, klantenkaarten, surfen op een website, ...), kunnen deze bedrijven hun strategie geval per geval aanpassen. Facebook gebruikt bijvoorbeeld Big Data-technologieën om op een meer gerichte manier de gesponsorde reclameboodschappen⁹ te ontwikkelen.

Laten we voor ogen houden dat Big Data niet alleen in de marketing kan worden toegepast. Big Data zijn een echte bron van vooruitgang in andere sectoren zoals bijvoorbeeld **de sportstatistiek**. Het artikel over EURO 2016 *'l'Allemagne joue au Big Data sportif pour l'exporter (encore)'* in *'Usine Digitale'*¹⁰, haalt een feit aan dat niet bepaald een historische kanttekening is. Het Duitse team heeft verklaard deskundigen te hebben aangeworven inzake Big Data op het gebied van sport. De resultaten afkomstig van deze analyses van de gespeelde matches bieden Duitsland de mogelijkheid om haar strategie aan te passen. Zij bevestigen dat een match ook moet worden voorbereid met de wetenschap van de gegevens. We weten niet of het feit dat men een beroep heeft gedaan op deze technologie een invloed heeft gehad op hun vooruitgang, maar het is wel een feit dat het Duitse team in de halve finale stond tegenover Frankrijk. Uiteindelijk weten we allemaal dat dit hen geen geluk heeft gebracht, maar wie weet? Misschien voor een volgende match. Zoals we dit zien met dit voorbeeld kunnen zelfs voetbalteams genieten van Big Data-technologieën om hun spelstrategie te verbeteren en hun tegenstanders beter te leren kennen.

We mogen dan momenteel wel horen praten over Big Data, de technologie is niet recent. Reeds in de jaren 1970 werd het idee van Big Data op tafel gelegd. Het is in 1997 dat de term *'Big Data'* voor het eerst wordt gebruikt in een artikel – *'Application-controlled demand paging for out-of-core visualization'* - door Michael Cox en David Ellsworth¹¹. In die tijd was het nog een theorie, want de technologie bood niet de mogelijkheid om tools aan te bieden. Het was wachten tot de jaren 2010 vooraleer er concreet kon worden gedacht aan een Big Data-toepassing.

4. Microsoft Azure Data Lake: zie glossarium

5. Microsoft Cortana Analytics: zie glossarium

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

7. CIBG: zie glossarium

8. Voor meer informatie over dit project en over andere projecten: raadpleeg het dossier van het CIBG 35: http://cibg.brussels/nl/nieuws_publicaties/publicaties/katernen/nr-35-vier-centrale-projecten-van-smartcity-brussels

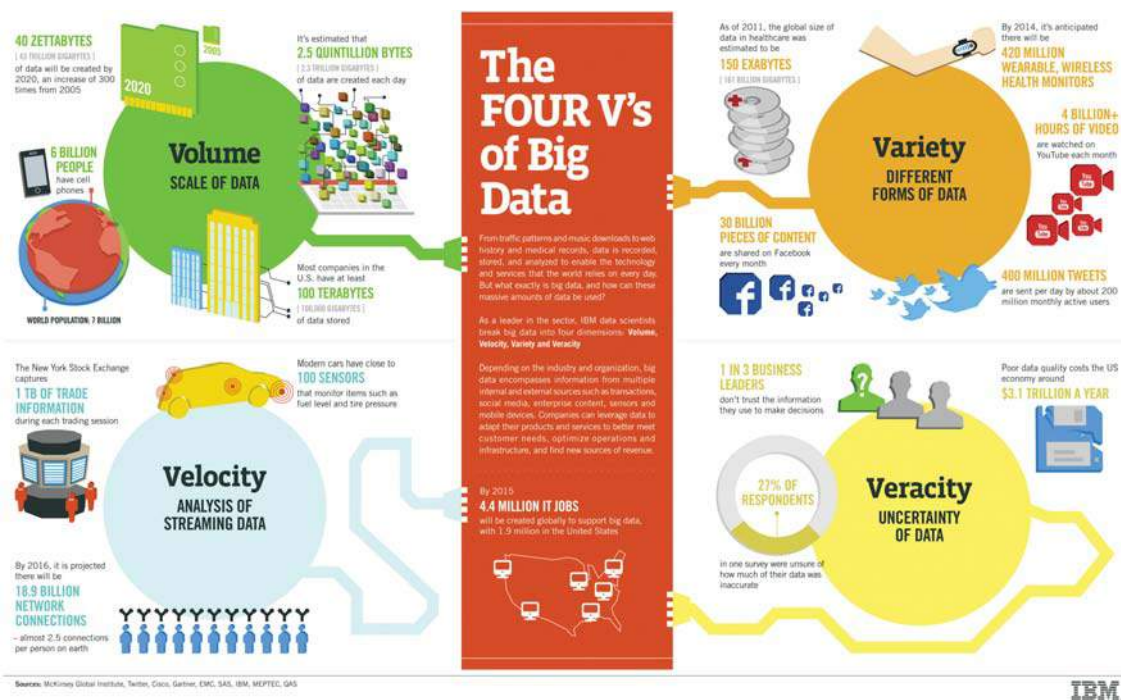
9. Om meer te weten over het gebruik van Big Data door Facebook: <http://www.simplilearn.com/how-facebook-is-using-big-data-article>

10. Nathalie Steiwer, EURO 2016: l'Allemagne joue au Big Data sportif pour l'exporter (encore) In L'usine Digitale [online] <http://www.usine-digitale.fr/article/euro-2016-l-allemaigne-joue-au-big-data-sportif-pour-l-exporter-encore.N396112> (Geraadpleegd op 16 juni 2016) (in het Frans).

11. Gil Press, A very short History of Big Data in Forbes [online] <http://www.forbes.com/sites/gilpress/2013/05/09/a-very-short-history-of-big-data/#6d76ff6155da> (geraadpleegd op 17 juni 2016).

In 2001 definieerde het bedrijf Gartner¹² Big Data aan de hand van drie kenmerken:

1. **Het volume** : één van de meest gekende kenmerken van Big Data is dat het algoritme een gigantische hoeveelheid gegevens verwerkt. De gigantische hoeveelheid gegevens is een actueel probleem: de wereld heeft de voorbije jaren meer informatie geproduceerd dan alle informatie samen sinds het schrift werd uitgevonden.
2. **De verscheidenheid**: de verscheidenheid onderscheidt de Big Data van de zogenaamde 'klassieke' analyse. Big Data analyseren immers gegevens afkomstig van verschillende bronnen. Men verwerkt dus gestructureerde gegevens (dynamische tabellen met dubbele ingangen van het type Excel, traditionele gegevensbanken, enz.) en niet gestructureerde gegevens (vrije teksten, informatie afkomstig van sociale netwerken, video's, enz.). Met behulp van klassieke gegevensbanktools is het dus moeilijk om niet gestructureerde gegevens in te zamelen. Big Data hebben de capaciteit om deze heel ongelijke gegevens in te zamelen teneinde ze te analyseren.
3. **De snelheid**: het derde criterium kan worden bepaald door de snelheid waarmee de gegevens worden en moeten worden ingezameld om geklasseerd en geanalyseerd te worden. Bij Big Data zijn de gegevens vaak heel talrijk: men moet dus al deze gegevens snel kunnen verzamelen om ze allemaal te kunnen verwerken.



Illustratie 1 De vier V's van Big Data¹³

Nadien zal de wetenschappelijke gemeenschap aan deze drie kenmerken nog andere criteria toevoegen zoals:

4. **de betrouwbaarheid**: « de notie van betrouwbaarheid benadrukt de kwalitatieve dimensie van de gegevens die nodig zijn voor de werking van Big Data¹⁴tools'».
5. **de zichtbaarheid** : « de zichtbaarheid die wordt geboden door de platforms van bruikbare Big Data biedt de gebruikers de mogelijkheid om de gegevens en de gepubliceerde informatie goed te gebruiken door gerichte zoekopdrachten¹⁵».

12. Gartner: Zie glossarium.

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> (in het Frans)

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

6. **de waarde:** « voor een bedrijf wordt dit gemeten aan de hand van het competitief voordeel dat eruit wordt gehaald. Dit hangt vooral af van de kwaliteit van de analytics en van de competentie van de 'Data Scientists' die hieraan zin geven¹⁶. »

7. **de foutenmarge** (penalty ranking qu'on appelle aussi « snow »)¹⁷.

Over het algemeen heeft men de neiging om Big Data te associëren met geavanceerde technologieën. Het is niet verkeerd maar heel beperkend, want Big Data is geen kwestie van technologieën. Sinds enkele jaren praat men over de volgende technologieën:

- Hadoop : « framework Open Source ontworpen om bewerkingen te realiseren op grote volumes gegevens, in de grootorde van verscheidene petabytes (m.a.w. ettelijke duizenden Terabytes¹⁸)»
- R-software: analyseert en verwerkt de gegevens met behulp van de programmeertaal S¹⁹.
- Map Reduce : « biedt de mogelijkheid om grote hoeveelheden gegevens te bewerken door ze te verdelen in een machinecluster om verwerkt te worden²⁰».
- En nog veel meer, ...

Deze **technologieën** evolueren heel snel en worden snel vervangen door andere meer geavanceerde en meer relevante technologieën; Hugues Bersini wijst tijdens ons gesprek op dit feit als verantwoordelijke van de IRIDIA²¹ aan de ULB. Het spreekt zo voor zich dat Big Data niet mogen beperkt worden tot technologieën. Het zijn wel degelijk geavanceerde technologieën, maar ze zijn nog onstabiel, omdat ze zo snel evolueren. Big Data, dat is in de eerste plaats een **ALGORITME**. Het algoritme vormt de kern van het geheel. Het algoritme zamelt gegevens in, verwerkt ze, enz. De statistische formule is vaak ontworpen in functie van de vraag van de klant, van de vraag die hij wil stellen aan het algoritme. Een slecht ontworpen algoritme is een slechte Big Data-applicatie. Dat is de reden waarom de bedrijven veel belang hechten aan het beroepsprofiel dat het ontwerp van het algoritme behartigt: **de Data Scientist**. Jammer genoeg zullen we in het hoofdstuk over de beroepsprofielen van Big Data zien (Big Data-beroepen) dat dit type professional nog veel te weinig te vinden is in België.

Een andere problematiek inzake Big Data is dat Big Data exclusief geassocieerd worden met **massale gegevens**. Misschien is de term slecht gekozen, maar het geeft de indruk dat deze technologie draait rond een gigantische hoeveelheid gegevens die wordt verwerkt door een algoritme. Het volume is echter geen doel op zich. Big Data zijn meer dan volume alleen. Big Data dient men te abstraheren van het 'Big'-concept, want het gaat niet enkel om grote bedrijven die het zich kunnen permitteren om een 'hallucinante' hoeveelheid gegevens te verwerken. Tijdens verschillende gesprekken die wij hadden, blijkt dat Big Data ook en vooral een kwestie van variëteit is. De term 'Big Data' wordt trouwens steeds minder gebruikt. Philippe Van Impe, stichter van de European Data Innovation Hub en van de Brussels Data Science Community²², legt uit dat de term Big Data voorbijgestreefd is. De wetenschappelijke gemeenschap praat voortaan liever over '**Gegevenswetenschap**'.

Big Data gaat ten slotte over **gegevensopslag**. De hoeveelheid gegevens die tegenwoordig wordt gegenereerd, vereist immers een performante en aangepaste opslagplaats. Het is sinds het verschijnen van Cloud Computing en zijn rijping dat Big Data zich werkelijk kon ontwikkelen²³.

16. <https://www.digitalwallonia.be/big-data-la-revolution-des-donnees/> (in het Frans)

17. Voor meer info: <https://campus-bruxelles.com/2014/02/12/big-data-formation-informatique-bruxelles/> (in het Frans)

18. <http://hadoop.apache.org/>

19. [https://nl.wikipedia.org/wiki/R_\(softwarepakket\)](https://nl.wikipedia.org/wiki/R_(softwarepakket))

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

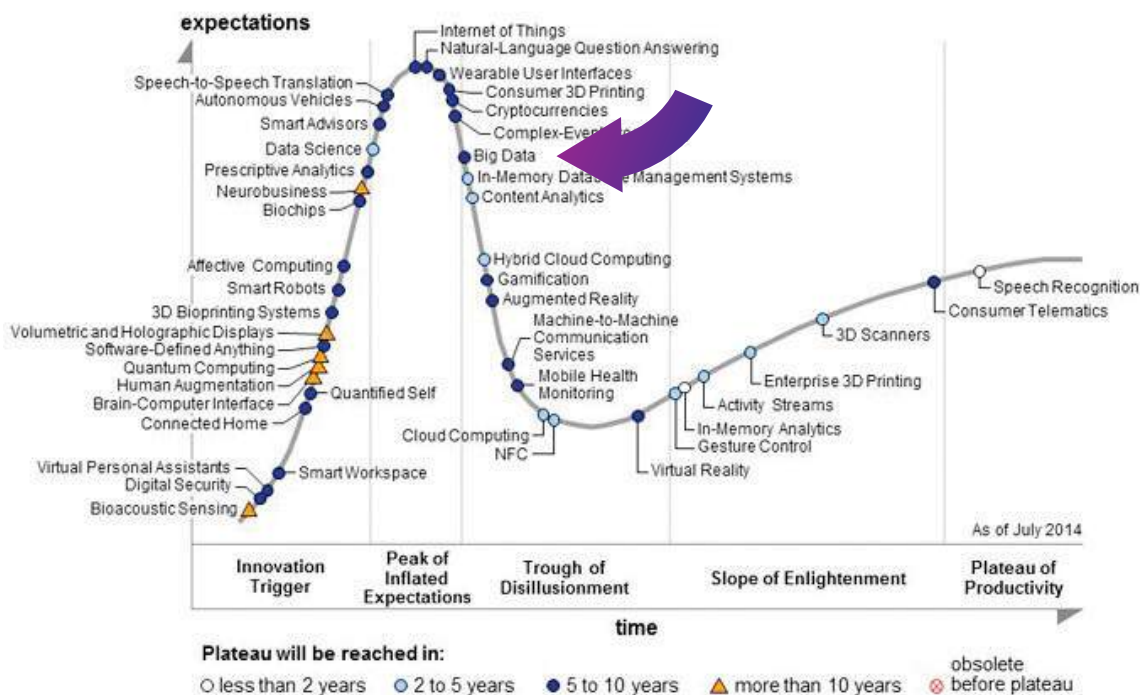
21. IRIDIA: zie glossarium.

22. Website: <https://datasciencebe.com/>

23. <http://www.technologischeontwikkeling.nl/ict/de-goede-match-van-cloud-computing-en-big-data>

2.2 Troeven van Big Data

Er wordt enorm veel gesproken over Big Data in de pers, bij gegevensspecialisten, bedrijven; iedereen doet zijn best. Het is inderdaad geen nieuw concept. Wat momenteel het succes uitmaakt van Big Data is de mate waarin Big Data wereldwijd geëvolueerd is. In zijn Hype Cycle van 2014, klasseert Gartner de Big Data aan het begin van de curve van de « *trough of Disillusionment* » net na de « *Peak of Inflated Expectations* ».



Illustratie 2: de 'Hype Cycle' 2014 van Gartner

Dit houdt uiteraard verband met het rijpingsproces (maturity process). Bedrijven moeten een hoger niveau qua gegevensanalyse bereiken en Big Data toont deze evolutie heel goed aan. Een van de troeven is dus de consolidering op de internationale markt en de heel snelle evolutie. Men kan de positie van Big Data in 2014 interpreteren door het feit dat de technologieën die ermee worden geassocieerd evolueren en heel snel worden vervangen. Brussel blijft achter maar het zou niet verwonderlijk zijn indien Big Data er uiteindelijk op duurzame wijze wel terug te vinden zullen zijn.

Omdat Big Data al een heel eind gevorderd is, steeds makkelijker toegepast wordt, kan en moet deze een criterium zijn voor tewerkstelling. Gegevens zijn het zwarte goud geworden die bedrijven kost wat kost willen benutten. En om de gegevens te kunnen gebruiken, moet men gaan aanwerven. We zullen verder zien dat het dringend nodig is om Big Data-specialisten op te leiden. Deze jobs worden enorm veel gevraagd bij de Belgische rekruteerders.

We kunnen het nooit genoeg herhalen; gegevens hebben een reële impact binnen een hele reeks toepassingsdomeinen: criminaliteit, landbouw, stockbeheer, gezondheidssector, enz.

2.3 De beperkingen van Big Data

De grootste uitdaging inzake Big Data is juridisch (zie punt 2.4). De Big Data problematiek houdt voornamelijk verband met de bescherming van privé-gegevens van private aard. De gegevenswetenschap vraagt soms om persoonlijke gegevens te verwerken. Zo heeft een telefoonoperator de gegevens van zijn klanten nodig, wil hij de communicatiegewoonten van zijn doelgroep analyseren. En dat doet natuurlijk veel denken aan het genre dat zoveel bijval kent in futuristische films: het 'Big Brother'-concept. Het is momenteel een heel gevoelig onderwerp (het Snowden schandaal, enz.). Men mag toch niet uit het oog verliezen dat die technologie meestal niet het individu of zijn persoonlijk leven zal controleren; de tool zal enkel het gedrag analyseren bij een aantal personen om zich zo een idee te vormen van de gewoontes zonder in detail te treden, zonder verder te gaan met die persoonlijke gegevens. Het is iets wat ook publiek van nut kan zijn, bijvoorbeeld bij een openbare vervoersmaatschappij²⁴. Indien zo een maatschappij de relevantie en de frequentie van haar gebruikers in haar voertuigen wil analyseren, kan ze dat doen door de gegevens te bestuderen afkomstig van de abonnementskaart van de gebruikers. Deze techniek is niet alleen een controlemiddel maar ook een tool om de diensten die worden aangeboden aan de gebruikers te verbeteren.

Een andere rem voor de verankering van Big Data in Brussel is een gebrek aan kennis. Bedrijfsleiders weten niet wat Big Data zijn. Zij durven die niet te gebruiken voor hun bedrijf of denken dat ze die niet nodig hebben. Wij zullen daar meer in detail op ingaan in het deel over de bedrijven (Big Data in Brussel).

2.4 Het juridische aspect van Big Data

Voor elke technologische evolutie geldt dat de wetten die de toepassing regelen in acht genomen moeten worden. Laten we hier even bij stilstaan.

Meester Philippe Laurent, advocaat aan de balie van Brussel, expert intellectueel eigendomsrecht en recht inzakenieuwe informatie- en communicatietechnologieën, legt tijdens ons gesprek uit dat *'bepaalde bedrijven niet voldoende of te laat belangstelling tonen voor het juridische aspect. Men moet daar zo vroeg mogelijk aan denken want de leefbaarheid van een project kan soms berusten op beslissingen die na analyse onwettig blijken te zijn'*.

Vanuit een juridisch standpunt is Big Data een vrij recent maar zeer ruim concept. Daarom bestaat er in de wetgeving nog geen concrete definitie van. Het is dus heel gemakkelijk om de grens tussen ethisch handelen en fraude (bewust of niet) heel vaag te houden. Zoals dhr. Laurent preciseert, moeten de bedrijven die werken met Big Data zich een aantal vragen stellen alvorens het project aan te vatten: welke gegevens worden verwerkt? Van waar komen ze? Waarom en hoe worden zij verwerkt? Zijn deze gegevens 'persoonlijk'? Al deze vragen maken het mogelijk een kader af te bakenen waarbinnen zij zullen kunnen spelen.

Er zijn beperkingen inzake het gebruik van de gegevens; de bescherming van de prive-sfeer en van de intellectuele eigendom.

24. Concreet voorbeeld: 'un réseau d'autobus redessiné grâce au téléphone mobile' in La recherche – december 2013 Nr. 482 [online] http://www.ffue.org/wp-content/uploads/2015/09/FU-US_150927_Blondel_LaRecherche_BigData.pdf [geraadpleegd op 11 juli 2016].

Volgens de Privacycommissie zijn **gegevens van persoonlijke aard** 'gegevens betreffende een persoon waarmee hij rechtstreeks of onrechtstreeks kan worden geïdentificeerd of identificeerbaar is'²⁵. Er bestaan dus wetten die het gebruik van de private regels regelen. Een bedrijf kan zich niet permitteren om de gegevens van zijn potentiële of bestaande klanten te gebruiken zoals het wenst.

'Met **intellectuele eigendom** bedoelen we het geheel van rechten die exclusief worden toegekend aan intellectuele creaties'²⁶ legt het Europees Parlement uit. Een specifieke wetgeving werd hiervoor ook uitgevaardigd. De term duikt nu ook op in het Verdrag betreffende de werking van de Europese Unie in 2009²⁷.

In Europa is het beschermen van gegevens een echte uitdaging. Het is een problematiek die heel ernstig wordt genomen. Bovendien is de wetgeving inzake het gebruik van private gegevens voor commerciële doeleinden heel strikt. De Verenigde Staten zijn veel minder veeleisend op dat gebied.

We horen al een tijdje spreken over **Safe Harbor** en **Privacy Shield** in de media; een wetgeving die de overdracht van gegevens tussen Europa en de Verenigde Staten tracht te regelen. De twee continenten hebben immers niet dezelfde prioriteiten. In tijden van mondialisering is het belangrijk om een gezamenlijke wetgeving te vinden die zowel geschikt is in Europa als in de Verenigde Staten.

Safe Harbor werd opgericht in 2015; het regulariseert het gebruik van gegevens van Europese burgers door Amerikaanse bedrijven met filialen in Europa. Er werd een kader opgelegd door de Europese Unie om afwijkingen inzake dit gegevensgebruik te vermijden.

Een specifieke wetgeving inzake 'privacy' bestaat slechts sinds 1998; het gaat hier om richtlijn 95/46/CE²⁸. Deze zegt echter niets over de overdracht van gegevens tussen de Europese Unie en de Verenigde Staten. Daarom werd Safe Harbor in het leven geroepen. Heel snel werd deze wetgeving betwist en de discussie kwam weer op tafel.

Uit deze discussies ontstond het **EU/VS Privacy Shield** dat een kader wil bieden voor de overdracht van gegevens tussen de EU en de Verenigde Staten. Het is een herziening van Safe Harbor met een striktere wetgeving.

« De volgende maatregelen worden genomen :

- De bedrijven moeten zich schikken naar vaste verplichtingen, met een navenante, strikte uitvoeringswetgeving,
- Toegang door de Amerikaanse autoriteiten die sterk omkaderd is en transparant,
- Een effectieve bescherming van de rechten van de burgers van de Unie en verschillende mogelijkheden om in beroep te gaan,
- Gezamenlijk mechanisme voor een jaarlijkse herziening.²⁹»

Vanaf de publicatie van deze nieuwe richtlijn heeft de 'groep 29'³⁰ die werkt aan het naleven van de bescherming van de gegevens van private aard, het Privacy Shield betwist. Er is een zekere vooruitgang maar er zijn nog teveel duistere zones. Volgens hem ontbreken er 'striktere garanties inzake de onafhankelijkheid en de macht van de Amerikaanse bemiddelaar'³¹.

25. <https://www.privacycommission.be/nl/wat-zijn-persoonsgegevens>

26. http://www.europarl.europa.eu/atyourservice/nl/displayFtu.html?ftuld=FTU_3.2.4.html

27. http://www.europarl.europa.eu/atyourservice/nl/displayFtu.html?ftuld=FTU_3.2.4.html

28. <http://eur-lex.europa.eu/legal-content/NL/TXT/?uri=CELEX:32016D1250>

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

30. Voor meer info over Groep 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>

Veiligheid is ook in deze context een essentieel punt. Aan de klant kunnen garanderen dat zijn persoonlijke gegevens beschermd zijn, is essentieel. Zich niet kunnen engageren in die zin kan de e-reputatie van het bedrijf beschadigen (de reputatie van het bedrijf op internet, het imago).

Ten slotte is de transparantie inzake het gebruik van Big Data-gegevens van fundamenteel belang. Alleen een zich ontwikkelend Europa dat de overdracht van private gegevens wenst te beschermen, kan hiervan een realiteit maken. Denken aan de beveiliging van de gegevens en zo snel mogelijk de Europese wetten naleven, helpt bij het realiseren van een project voor het analyseren van gegevens op grote schaal.

De bedrijven moeten gesensibiliseerd worden voor de wetgeving die het gegevensgebruik regelt. Ook klanten en burgers moeten worden gesensibiliseerd inzake de impact van de publicatie op hun persoonlijke gegevens. Deze problematiek staat centraal bij Europese besprekingen. Het Europees project '*Horizon 2020*³²' opent bij een dergelijke reflectie over Big Data de deur naar onderzoeksprojecten mogelijk. Wij zullen dit onderwerp behandelen in het deel 'Big Data in Europa' (zie Big Data in Europa).

Het CRIDS – UNAMUR

Het onderzoekscentrum inzake Informatie, Recht en Maatschappij wijdt zich aan thema's die verband houden met de informatiemaatschappij en meer bepaald haar maatschappelijke, juridische, ethische en filosofische aspecten. Het centrum interesseert zich voor onderwerpen zoals de persoonlijke levenssfeer, intellectuele eigendom, e-government, massale gegevens, de cyberveiligheid, de digitale cultuur, algoritmisch government, en andere.

Het centrum werd in 2010 opgericht binnen de universiteit van Namen (UNAMUR), en is het resultaat van de fusie tussen het onderzoekscentrum naar informaticarecht (het CRID), het interdisciplinair evaluatiecentrum voor technologieën (CITA) en de '*Interdisciplinaire onderzoeksgroep inzake Communicatie en Internet*' (het GRICI).

Het CRIDS bestaat uit een vijftigtal vorsers die meestal professionele ervaring hebben met de informatiemaatschappij.

Het CRIDS is opgesplitst in zeven onderzoeksunits:

- De unit 'vrijheden en informatiemaatschappij',
- De unit 'intellectuele eigendom',
- De unit 'E-commerce',
- De unit 'elektronische communicatie',
- De unit 'communicatie en internet',
- De unit 'technologieën en maatschappij',
- De unit 'e-overheid'.

Het CRIDS biedt eveneens lessenreeksen aan in de vorm van een master of certificaat:

- Een specialisatie Recht inzake de Informatie- en Communicatietechnologieën,
- Een master Business Analysis en IT Governance (voorgesteld in dit rapport in het hoofdstuk opleiding),
- Een certificaat inzake het veiligheidsbeheer van de informatiesystemen,
- Een interuniversitair certificaat inzake de bescherming van de gegevens van persoonlijke aard (CRIDS en ICHEC).

Voor meer informatie : <http://www.crids.eu>

32. Voor meer informatie over de voorstellingen van het Big Data-plan in de Europese Commissie: <https://ec.europa.eu/digital-single-market/en/news/big-data-information-and-networking-days-horizon-2020-topics>

Big Data in Europa

3

Big Data trok ook de aandacht van de Europese Commissie inzake onderzoek en ontwikkeling. In de context van 'Horizon 2020' en meer precies in het specifieke deel over Big Data, lanceerde de Commissie een projectoproep inzake de valorisatie van Big Data. Zij investeert 2,5 miljard euro in het kader van een overeenkomstprotocol met een PPP (publieke-private partner). De Commissie heeft het uitbreidingspotentieel ontdekt van de gegevenseconomie. Deze economie blijft zich maar ontwikkelen en er zijn alsmar meer voorbeelden van het gebruik van Big Data. Europa mag de trein niet missen om projecten uit te werken die gunstig zijn voor de Europese Unie. Er wachten mooie dagen op deze markt.

De Europese Commissie heeft vastgesteld dat Big Data een gunstige impact hebben op de economie van een land³³ :

- *Verhoging van de productiviteit met 5-7 %,*
- *100.000 potentiële nieuwe jobs tegen 2020,*
- *Vermindering van het energieverbruik met 10 %,*
- *Verbetering van de gezondheidszorgen, ...*

De Commissie wil dat Big Data-projecten afkomstig van 'Horizon 2020' zich richten op de reële noden en het creëren van een 'ecosysteem van het efficiënte gegeven'³⁴. Europa moet nadenken over een reguleringsbeleid van het gegeven, dit wil zeggen dat Europa een gebruikskader zal moeten creëren zonder dat private gegevens worden gebruikt voor commerciële doeleinden. Dat laatste punt is één van de meest gevoelige, want je moet een evenwicht vinden tussen de vooruitgang van de 'data-economie' en ethisch handelen, overeenkomstig een wetgeving op het gebruik van gegevens van private aard.

De onderzoeksonderwerpen die worden aangehaald als reflectiepistes zijn vooral de Open Data³⁵, de Big Data-tools, de infrastructuur, en het wettelijke kader.

De toepassingsdomeinen kunnen domotica zijn, energie, mobiliteit, telecommunicatie, gezondheid en nog andere.

De site *datalandscape*³⁶ vermeldt verschillende programma's die worden gelanceerd in alle Europese landen voor 'Horizon 2020'. De lijst telt ook enkele Belgische projecten.

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

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

35. OPEN DATA: zie glossarium

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

4

Big Data in Brussel

4.1 Het politiek-economisch aspect

Voor het Brussels Hoofdstedelijk Gewest is het onderwerp Big Data en meer specifiek Open Data belangrijk. Het Gewest wil een digitale strategie uitwerken om van Brussel een Smart City³⁷ te maken. Het heeft beslist om bepaalde technologieën naar voren te schuiven, zoals de gegevenstechnologie, de virtuele realiteit en het Internet Of Things. De bedoeling van het Gewest bestaat erin om zoveel mogelijk publieke gegevens ter beschikking te stellen, via een portaal dat toegang biedt tot de gegevens van de stad Brussel³⁸. Een andere bedoeling van Brussel bestaat erin om voor de startups en andere jonge bedrijven een publiek onderzoekslabo ter beschikking te stellen. Dit onderzoekslabo zou de mogelijkheid bieden om technologieën te gebruiken en om begeleid te worden door deskundigen.

De echte uitdaging voor Brussel is van wetgevende en economische aard. De maatschappij moet de gegevens valoriseren en een sensibiliseringsrol spelen via verschillende acties. Deze acties kunnen verwezenlijkt worden door het creëren van clusters of incubatoren en een sterke boodschap op politiek vlak. We zijn nog ver verwijderd van een gegevensgebruik door het grote publiek.

Dhr. Nicolas Harmel, attaché bij het kabinet van minister Didier Gosuin, Minister van Economie en Tewerkstelling in Brussel, verklaart: de openbare sector moet aantonen dat het gegeven zijn nut heeft. De positieve impact van Big Data moet worden aangetoond. Des te meer daar men elke dag Big Data gebruikt zonder zich daarvan bewust te zijn, via zijn smartphones, laptop, enz.

Tegenwoordig wordt het vrijgeven van gegevens geremd, waardoor een hinderpaal ontstaat voor diegenen die ze voor 100 % wil gebruiken. De overheden moeten het spel meespelen en zij moeten de positieve impact tonen van Big Data. In een stad als Brussel zijn er echte uitdagingen inzake het gebruik van gegevens en Open Data zoals mobiliteit, veiligheid, enz. Dit alles kan worden ontwikkeld en zowel grote als kleine bedrijven kunnen hierop inspelen.

4.2 De bedrijven en Big Data

Binnen de context van de Big Data en het gebruik ervan in België kan men twee typevoorbeelden onderscheiden.

- Bedrijven die een Big Data-afdeling hebben geïntegreerd of dit zouden kunnen integreren in hun marketing- of communicatieafdeling.,
- Bedrijven gespecialiseerd in het analyseren van gegevens die hun diensten aanbieden en een Big Data-analysestrategie voor bedrijven. Die laatsten zijn vaak startups die hoofdzakelijk gevestigd zijn in Brussel.

In dit rapport hadden we het al over een terughoudende houding bij de Brusselse bedrijven. Wij stellen echter wel vast dat de Data Analyse-bedrijven bloeiend zijn en veel klanten hebben. Waar halen zij hun marktaandeel?

37. SMART CITY: zie glossarium

38. Portaal: <http://opendata.brussel.be/page/home/>

Gezien het sensibiliseringsprobleem bij de Belgische bedrijven, richten bedrijven die gespecialiseerd zijn in gegevensanalyse zich heel graag tot hun buitenlandse klanten. Brussel heeft het voordeel dat het de hoofdstad is van Europa, verklaart Claude Delcour, beheerder en directiehoofd bij Vadis.



Illustratie 3 - Robert VISEUR; Het Big Data regionaal ecosysteem. Ervaringen inzake Big Data-technologieën. Big Data Week 2015: Bridging data events all over the world – CETIC Gosselies, 24/11/2015.

'Tientallen Brusselse bedrijven beschikken over gegevens die ze niet gebruiken', stelt Juan Bossicard bij Software.brussels. Dat is een verontrustend gegeven voor de bedrijven die gegevens analyseren. Pierre-Nicolas Schwab, die verantwoordelijk is voor de Big Data-strategie bij de RTBF is van mening dat een van de redenen voor het lauwe onthaal bij de Belgische bedrijven van Big Data berust op het feit 'dat het topmanagement die materie niet wil kennen'.

'Het is een terugkerende vaststelling. We moeten communiceren! We moeten dit bij de bedrijven aan de man brengen opdat zij zich bewust zouden worden van het potentieel van de data en van de analyses die ze eruit kunnen halen' zegt Laurent Fayet, Managing Partner bij LBFayet Consulting³⁹.

39. LBFayet Consulting Zie glossarium

5

Big Data-beroepen



De data-economie evolueert constant. De internationale markt houdt niet op met evolueren en aan belang te winnen. Zo stellen we vast dat er nieuwe jobprofielen opduiken. Het is interessant om stil te staan bij de beroepsprofielen die geëvolueerd zijn en die opdoken dankzij de datasector.

Vooraleer de verschillende profielen te analyseren die verband houden met de Big Data is het belangrijk te wijzen op een specifiek gegeven dat te maken heeft met terminologie. Op het wereldwijde web zien we heel wat verschillende profielen opduiken: Data Scientist, Data Analyst, Data manager, Chief Data Architect, Chief Technology Officer en tal van andere. Moeilijk om daar nog wijs uit te geraken. Des te meer omdat al deze profielen allemaal min of meer dezelfde competenties dekken. Met het oog op een beredeneerde analyse hebben wij beslist om in dit rapport enkel de meest courante benamingen te gebruiken – namelijk: Data Architect, Data Scientist en Data Analyst. Er bestaan uiteraard andere profielen die eerder business- of technologie-gericht zijn maar wij bekijken alvast of deze drie profielen goed de algemene competenties samenvatten die het meest worden gevraagd.

Wel dan; zijn de Big Data-beroepen vandaag een bron van tewerkstelling en zullen zij in de toekomst werkgelegenheid creëren in Brussel? Die vraag is uiteindelijk één van de meest belangrijke in dit dossier. Omwille van deze problematiek stellen wij ons vragen over de duurzaamheid van Big Data als bron van inkomsten en innovatie voor de hoofdstad. En dan rest er nog dé grote vraag: **zijn bedrijven vragende partij om de gegevenswetenschap te gebruiken?**

Ja, en ongetwijfeld alsmaar meer. Bedrijven kunnen niet meer naast deze tool kijken, die in de praktijk alsmaar meer wordt gebruikt. In België is de impact van Big Data echter weinig opzienbarend. *A contrario*, we zien alsmaar meer jobaanbiedingen opduiken die verband houden met data. Gegevensanalyse is een competentie die alsmaar meer wordt gevraagd door de bedrijven en dit zal in de komende jaren nog toenemen.

Internationaal hebben de databeroepen en meer precies de Big Data-beroepen de wind in de zeilen. In 2014 stelde LinkedIn, die grote gegevensbank van de beroepswereld, een lijst op met de meest trendy beroepen van het moment. Het is niet verrassend dat daar twee beroepsprofielen bijzitten die verband houden met Big Data: Data Scientist en Big Data Architect⁴⁰. In België moet men dit weliswaar nuanceren. Niettegenstaande deze profielen heel erg in trek zijn, mag men hier niet onmiddellijk uit afleiden dat Big Data duizenden aanwervingen zullen opleveren. Waarom?

- Eerst en vooral gaat het over heel specifieke profielen die een grondige expertise vergen op het vlak van statistiek, wiskunde, technologieën en business. Om een Data Scientist-job in de wacht te slepen hebt u minstens een universitair diploma nodig en soms vraagt men ook beroepservaring op het vlak van gegevensanalyse,
- Om Belgische dataspecialisten te rekruteren, zouden academische instellingen en andere potentiële Big Data-actoren zich bewust moeten zijn van waar Big Data voor staan. Wij zullen dit hieronder behandelen in het hoofdstuk over opleidingen. De programma's gewijd aan opleidingen op het vlak van Big Data bestaan, maar zijn heel recent of ze zijn in voorbereiding en niet zo talrijk,

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

- Binnen deze sensibiliseringscontext moeten de bedrijven geïnformeerd worden over de Big Data-problematiek. Wij hebben reeds over dit onderwerp gepraat in het deel 'Big Data in Brussel'.

Men kan dus met zekerheid stellen dat het domein van de wetenschap van gegevens werk zal genereren want de vooruitzichten in Brussel zijn heel goed en aan het rijpen. Des te meer daar het Brussels Hoofdstedelijk Gewest dat een Smart City wil zijn, een openbaar en positief gebruik wil stimuleren van de gegevens. De technologieën zullen er komen, men moet mensen kunnen inzetten die in staat zijn ze te gebruiken. Volgens deze logica komt er tewerkstelling.



41

5.1 De Data Architect

Het werk van de Data Architect bestaat erin om gegevens te organiseren, ze te sorteren en ze samen te brengen. Hij definieert standaarden die de mogelijkheid bieden om het verzamelen en het sorteren te vergemakkelijken⁴². Het is een profiel dat meer informatica-gericht is dan dat van de Data Scientist. Hij moet dus een degelijke kennis hebben van gegevens en technologische tools die worden gebruikt voor de Big Data (Hadoop, langage de R-taal, ...).

Tijdens ons gesprek deelt Laurent Fayet, Managing Partner bij LBFayet Consulting dit beroepsprofiel op in twee competentie categorieën:

1. "Een **profiel gericht op zuivere technologieën**. Dat is een ingenieur infrastructuur die nadenkt over de technologieën die zullen moeten gebruikt worden in de strategie voor het verzamelen van gegevens,
2. Een **toepassingsgericht profiel**. Dat staat in voor de structuur die de gegevens onderling verbindt om een schema uit te werken."

Inzake opleidingen: de Data Architect moet een degelijke kennis hebben van programmeren en gegevenstechnologieën. Wiskunde en statistiek zijn nodig zelfs al zijn het geen kerncompetenties. Hij moet kennis hebben van data visualisation en machine learning. Er zijn bepaalde soft skills nodig: men moet nieuwsgierig zijn naar gegevens, men moet een passie hebben voor het toepassingsveld, communicatie, enz.

41. "I'm a little surprised, with such extensive experience in predictive analysis, you should've known we wouldn't hire you."; 'Ik ben een beetje verbaasd. Met zoveel ervaring en voorspellende analyse had u toch kunnen raden dat wij u niet zouden aanwerven.'

42. Fabrice Coudray, Focus op een gouden job: de Data Architect, in Journal Du Net [online] <http://www.journaldunet.com/solutions/expert/59345/focus-sur-un-job-en-or---le-la-data-architect.shtml> (geraadpleegd op 21 juni 2016).

5.2 De Data Scientist

De Data Scientist is diegene die het algoritme ontwerpt dat de gegevens zal behandelen. Hij doet de voorspellingen in functie van de resultaten die hij zal ontvangen (predictive analysis). We zouden kunnen zeggen dat dit het beroepsprofiel is dat het vaakst wordt aangehaald als het over Big Data gaat. Hij is zo een beetje de draaischijf van het proces. Deze pijler is zeer gegeerd door data-analyse-bedrijven.

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 withaaS 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
© Krzysztof Zawadzki

David Majorel in het artikel 'Salaire: Experts cybersécurité et data scientists sont les rois du pétrole' vat het profiel goed samen:

'De Data Scientist is de job die het vaakst wordt gevraagd, omdat hij heel zeldzaam is. Het gaat hier om profielen die technisch heel geavanceerd zijn inzake wiskunde en statistiek. Het zijn zeldzame profielen want er is niet echt een opleiding, behalve van gespecialiseerde ingenieurs of ingenieurs die een opleiding volgden bij statistiekscholen. Zij moeten heel wat tools kennen, maar ook een visie hebben, creatief zijn en zaken uitproberen⁴³.'

De Data Scientist genoot een opleiding wiskunde en statistiek maar heeft ook goede competenties inzake informatica (vooral om algoritmes te ontwerpen). Hij moet vaak probabiliteiten uitwerken.

De Data Scientist kan een beetje beschouwd worden als de next step, na de Data Analyst en de Business Analyst⁴⁴. Hij onderscheidt zich van de klassieke databeroepsprofielen omdat hij analyses uitvoert op basis van massale gegevens afkomstig van uiteenlopende bronnen (Big Data dus).

43. Julien Bergounhoux, Salaire: Experts cybersécurité et data scientists sont les rois du pétrole in l'Usine Digitale [online] (artikel in het Frans) <http://www.usine-digitale.fr/editorial/salaire-experts-cybersecurite-et-data-scientists-sont-les-rois-du-petrole.N383849> (artikel in het Frans)

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

5.3 De Data Analyst

De Data Analyst wordt vaak verward met de Data Scientist; men moet hem beschouwen als zijn tegenhanger. Zij werken meestal samen. De Data Analyst maakt overzichten en modellen op basis van de ingezamelde gegevens die worden gesorteerd door de Data Architect en die in het algoritme van de Data Scientist worden gegoten. Het is vaak de taak van de Data Analyst om de resultaten samen te vatten, om te communiceren met de directie en om aanbevelingen te formuleren.

De Data Analyst heeft een achtergrond in statistieken: hij kent statistieken, wiskunde en informatica net als de eerder vermelde beroepsprofielen. Hij heeft ook Business Intelligence (of BI) en marketing gestudeerd. Hij is de specialist die verslag uitbrengt, hij moet dus kennis hebben van marketingcommunicatie; hij moet voor het bedrijf een strategie kunnen afleiden om winst te genereren.

5.4 Transversale disciplines

Het is heel beperkend om de Big Data-beroepen terug te brengen tot de drie beroepsprofielen die eerder werden beschreven. Niettemin zijn het diegene die het vaakst terugkeren in de debatten over de beroepen van de gegevensanalyse. Men kan andere termen vermelden: Project manager, Chief Data Officer (de "gegevensverantwoordelijke"), Chief Technology Officer (staat in voor de technische en wetenschappelijke zaken⁴⁵), enz.

5.5 De voornaamste competenties

Bij de hierboven beschreven profielen hebt u ongetwijfeld de recurrentie opgemerkt van bepaalde competenties. Deze zijn essentieel om te werken met gegevens.

- Statistiek en wiskunde,
- Informatica en het beheersen van bepaalde specifieke technologieën (software R, Map Reduce, Hadoop, enz.),
- Datavisualisatie,
- Marketing en communicatie,
- Kennis van het toepassingsdomein: die laatste competentie werd eerder minder aangehaald maar het spreekt voor zich dat een gegevensspecialist, of dat nu een Data Scientist is, een Data Analyst of een andere specialist, zich bewust moet zijn van de problematiek van de bedrijfssector om een relevante strategie uit te werken.

Ten slotte; het competentieniveau ligt heel hoog, zodanig zelfs dat men een universitair diploma moet hebben en zelfs meer, en ervaring moet hebben met het verwerken van gegevens. Het is niet onmogelijk dat een bachelor een Big Data-job kan hebben, misschien als Data Architect, maar eerst en vooral moet de mentaliteit veranderen en de onderwijssector en de hogescholen in het bijzonder moeten zich bewust zijn van het potentieel van Big Data. Naast die hoogopgeleide profielen mag men ervan uitgaan dat met het toegenomen gebruik van de Big Data-technologieën, zowel publiek als privé, de overheden technische profielen zullen moeten aanwerven. Deze technici moeten in staat zijn om de installatie en het onderhoud te garanderen van de receptoren, censoren en andere tools die dienen om gegevens in te zamelen. Deze specifieke profielen houden verband met IT en vereisen vaak specifieke competenties. In een maatschappij die volop in evolutie is, wat kenmerkend is voor België, is het interessant om zich de vraag te stellen of het niet nuttig zou zijn om lessen informatica en sensibilisering voor informatica op te nemen in de schoolcursussen.

45. Volledige definitie: https://nl.wikipedia.org/wiki/Technisch_directeur

6

Big Data-opleidingen



Iedereen is het er roerend over eens: in België hebben we een tekort aan beroepsprofielen die verband houden met Big Data! Waarom?

Hugues Bersini, verantwoordelijke van IRIDIA, professor aan de ULB en lid van de Koninklijke Academie voor Wetenschappen in Brussel vat het academisch Big Data-landschap in België als volgt samen *'Er is een grote vraag naar opleidingen in het universitair onderwijs, in hogescholen maar ook bij instellingen als EVOLIRIS.*

De markt is nog niet helemaal rijp; we zitten in een overgangsfase. We hebben er nood aan om Big Data aan te wenden in alle disciplines.'

«Qua opleidingen moet men investeren in de datamarkt», verklaart Stephane Faulkner, professor aan de universiteit van Namen en IT-professional. We stellen ook een trage bewustwording vast van de academische instellingen inzake de Big Data-uitdagingen. Men is nieuwe opleidingen aan het opstarten in verschillende instellingen, andere is men nog aan het ontwikkelen. Wij stellen ook vast dat bepaalde disciplines ervoor hebben gekozen om een of meerdere lessen te integreren met betrekking tot Big Data in de opleidingscursus. Bij die initiatieven vinden we ook universitaire opleidingen, hogescholen en kwalificerende opleidingen.

In dit hoofdstuk van het rapport willen wij enkele voorbeelden van opleidingen voorstellen die al dan niet binnen het officieel onderwijs werden aangeboden. Wij weten echter dat deze lijst niet volledig is en niet representatief is voor wat er op de markt te vinden is aan Big Data-opleidingen.

6.1 Schoolopleidingen



Wij evolueren naar een alsmat meer digitale maatschappij en de profielen die verband houden met de IT zijn alsmat meer gegeerd. Buiten het debat over het creëren van lespakketten over Big Data in het hoger onderwijs, is het ook interessant om te praten over het belang van het integreren van lessen informatica of van het sensibiliseren voor informatica in het lager onderwijs. Op dit moment bestaan er weinig lessen informatica voor jongeren in het officieel onderwijs. Aangezien jongeren geboren en opgegroeid zijn in een technocultuur, lijkt het belangrijk

om hen voor te bereiden op een professionele toekomst waarbij technologie een belangrijk aspect zal vormen. Die sensibilisering kan zich uiten op verschillende manieren: het integreren van informaticalessen die verder gaan dan het sensibiliseren voor privégegevens, en het gebruik van administratieve tools zoals Windows Office. Een sensibiliseringscampagne voor private en buitenschoolse initiatieven zoals Coderdojo⁴⁶, de coding sessies van MIC⁴⁷, enz. en vooral de invoering van opleidingen die bestemd zijn voor lesgevers die op dit moment geïnteresseerd zijn in het sensibiliseren van studenten voor informatica, maar die daar de pedagogische tools niet voor hebben.

46. <http://www.coderdojobru.com/home-nl>

47. <http://www.mic-brussels.be/>



Dhr. Tanguy De Lestré, Adviseur Informatica en Digitale Overgang bij het Kabinet van Staatssecretaris bij het Brussels Hoofdstedelijk Gewest, mevr. Bianca Debaets, geeft uitleg tijdens ons gesprek: 'Momenteel (en dat zal jaar na jaar alsmat belangrijker worden) is het belangrijk dat iedereen een basis kan hebben om te coderen want de internationale en Belgische economie wordt alsmat virtueler.»

6.2 Opleidingen aan universiteiten en hogescholen

Tijdens ons gesprek met dhr. Bersini hebben wij vernomen dat er een **Big Data Master** wordt voorbereid aan de ULB: die master zal een basis zijn met wiskunde en statistiek, en een ander deel zal worden gewijd aan informatica en Big data-technologieën, en specifieke lessen die verband houden met uiteenlopende toepassingsdomeinen (marketing, economie, enz.). Deze master wordt nog altijd voorbereid, de ULB hoopt de opleiding te kunnen integreren in het programma van het schooljaar 2016.

In Vlaanderen bestaat er aan de universiteit van Gent (UGENT), een master in '**wetenschap van de Marketinganalyse**'. Dit programma wordt onderverdeeld in drie delen: marketing en marketingresearch, informatica, (met name databanken) en ten slotte statistiek en Data Mining⁴⁸. Het is een master van één jaar die de mogelijkheid biedt om kennis te maken met het potentieel van de Big Data inzake marketing⁴⁹.

In Vlaanderen biedt de Thomas More Hogeschool eveneens een **bachelor aan inzake 'applied computer science'**⁵⁰ die Big Data-gerelateerde competenties omvat: het structureren van gegevens, het ontwerp en het beheer van bankgegevens, de objectgerichte analyse en het communiceren van gegevens. Tijdens het eerste lesjaar onderwijzen de lesgevers de basiskennis en vanaf het tweede jaar van de bachelor leren studenten hoe ze gegevens kunnen beheren.

In Namen, in Wallonië bestaat er een master van één jaar die toegankelijk is voor iemand die een andere master heeft: '**Business Analyst en IT-Governance It**⁵¹'. Deze stelt vier modules voor: organisatie en Innovatie, Business Analysis (projectbeheer, requirements engineering, ...), IT-governance (BI, Recht inzake ICT, ...) en het eindwerk. Deze master lijkt business- en marketinggericht met specifieke aandacht voor gegevensanalyse. De student verwerft competenties inzake strategische analyse en organisatie.

Aan de faculteit voor wetenschappen van de UCL (Université Catholique de Louvain), organiseren LSBA⁵² en het platform SMCS⁵³ een **certificaat 'Junior Data Analyst'** sinds oktober 2016. Dit certificaat wordt aangeboden aan werkzoekenden en 'jonge gediplomeerden' op masterniveau of voor mensen die werken in een bedrijf en die nieuwe technologische competenties wensen te verwerven. Die opleiding wordt 2 tot 3 dagen per week gegeven van oktober tot december 2016. Deze opleiding wordt opgesplitst in verschillende modules:

- Verkennende analyse en inleiding tot het onderzoek,
- Gegevensbeheer als integraal onderdeel van het statistisch verwerkingsproces,
- Herhalingen en uitbreiding van kansrekenen en statistiek,
- Statistiekprogrammering,
- Multigevarieerde verkennende analyse,
- Modelvorming,
- Workshops.

48. Data mining of verkennen van gegevens: zie glossarium.

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

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

51. <http://msbagi.unamur.be/presentation-du-ms-bagi/> (link in het Frans)

52. L'École de statistique, biostatistiques et sciences actuarialles

53. Plateforme technologique de Support en Méthodologie et Calcul Statistiques

We merken eveneens deze bewustwording door het feit dat Big Data-problemen worden geïntegreerd in bepaalde cursussen:

- Een cursus 'gegevensvisualisatie' in de master Wetenschappen en Technologieën van de Informatie en de Communicatie aan de ULB,
- Een cursus 'Business Intelligence and Data science' bij de master engineering management aan de Solvay, ULB⁵⁴,
- Enz, ...

6.3 Kwalificerende opleidingen

Op het vlak van kwalificerende opleidingen is Big Data ook aan een doorbraak toe. Er zijn met name twee opleidingen bij Technofutur in Charleroi.

Een kwalificerende opleiding van 50 dagen met 6 weken stage in een bedrijf voor **Data Analyst**. Deze opleiding werd voor het eerst gegeven in februari 2016. Het beoogde publiek bestaat uit werkzoekenden met een logische geest en een grondige kennis van wiskunde en statistiek. De cursus werd opgesplitst in verschillende modules:

- Een basismodule: inleiding tot de data, Excel, ...
- Een module om gegevens te verwerken: SQL, ...
- Een module datawarehousing⁵⁵ : BI⁵⁶, ...
- Een module beroepsanalyse : UML⁵⁷, BPMN⁵⁸, ...
- Een module reporting: visualisering van de gegevens,...
- Een module gegevensanalyse : Data Mining, ...
- En een laboratorium datawarehousing, BI, Data Mining

Zoals vermeld op de website van het opleidingscentrum, zou de toekomstige Data Analyst op het einde van zijn opleiding moeten in staat zijn om '*visualiseringsrapporten uit te werken*' ... '*teneinde de resultaten van de analyses te tonen*⁵⁹'. Hij zal de capaciteiten hebben om het ondernemingsproces te begrijpen.

Bij Technofutur gaat sinds 13 juni 2016 een opleiding van 58 dagen door als **Big Data Technical Consultant**⁶⁰. Deze opleiding is bestemd voor werkzoekenden die minstens in het bezit zijn van een bachelor informatica of die 3 jaar beroepservaring hebben op het vlak van computergestuurd gegevensbeheer. Het programma kan worden opgesplitst in verschillende modules:

- Een inleidingsmodule: Data Mining, cloud, gegevensbeheer, opslag, verwerking, enz.,
- Een module modellering, bewerken en opslaan,
- Een module BI,
- Een module Data Mining,
- Een slotsessie die de capaciteit test om een Big Data-oplossing te vinden,
- Een module om een job te zoeken.

54. http://banssbf.ulb.ac.be/PROD_frFR/bzscrse.p_disp_course_detail?cat_term_in=201617&subj_code_in=GEST&crse_num_in=S430&PPAGE=ESC_PROG-CAT_AREREQ&PPROGCODE=MA-INGE&PAREA=M-INGEF&PARETERM=201516&PTERM=201617 (link in het Frans)

55. Datawarehouse: zie glossarium

56. Business Intelligence

57. UML: zie glossarium

58. BPMN: zie glossarium

59. Zie ook het programma van de Data Academy die onlangs werd opgericht bij Technofutur: <http://www.technofutur.be/Nosservices/P%C3%B4lesRessourcesDiffusion/DataAcademy.aspx> (link in het Frans)

60. <http://www.technofutur.be/Formation/Big-Data-Technical-Consultant?fth=5&a=1&ihh=1&cp=1> (link in het Frans)



Data Innovation Hub: Data Innovation academy



De Data Innovation Hub (DIH) van Brussel is een Big Data networkingruimte geleid door Philippe Van Impe. De DIH wil Big Data stimuleren in Brussel door het onderwerp en de beste praktijkvoorbeelden voor te stellen. De DIH wil het centrum zijn van verschillende gegevensactoren: ondernemingen, start-ups, beleid, academische instellingen, enz. en is eveneens animator van een gemeenschap; de Data Science Community, die regelmatig vergaderingen organiseert tussen leden om ideeën te delen.

De Hub organiseert conferenties, stelt haar diensten voor aan bedrijven, en stelt een networkingruimte ter beschikking van de startups. De networkingruimte organiseert eveneens elk jaar de Data Innovation Summit door dynamische actoren uit de sector uit te nodigen.

Bovendien is de organisatie de stichter van de di-Academy. Deze biedt een reeks opleidingen aan over Big Data en technologieën die toegankelijk zijn voor werknemers en werkzoekenden. Na deze opleidingen krijgt men een bewijs dat men de opleiding volgde. Er worden ook bootcamps en hackathons georganiseerd met verschillende partnerbedrijven.

Voor meer informatie: <https://datasciencebe.com/> of <http://di-academy.com/>

6.4 Andere initiatieven

Op internet vindt men eveneens een hele reeks opleidingen en niet kwalificerende lessenreeksen. Deze zijn vaak heel uitgebreid en bieden de mogelijkheid om zich de Big Data-lespakketten eigen te maken. De '**Big Data university**⁶¹' biedt cursussen aan die exclusief gericht zijn op Big Data (de beginselen, Inleiding tot Hadoop, Map Reduce, ...).

Er bestaat eveneens een **MOOC**⁶² **ter inleiding tot de Big Data** die werd gelanceerd in juni 2016 door UTT van Troyes in Frankrijk⁶³.

En er bestaan nog andere op het wereldwijde web.

Bepaalde instellingen (zoals de DIH) organiseren Hackathons of Boot camps om de kandidaten de mogelijkheid te bieden zich intensief en praktisch op te leiden voor technologische tools voor gegevensbeheer.

6.5 Big Data onderwijzen in Brussel

Al deze initiatieven laten ons toe om vast te stellen dat de bewustwording van het jobpotentieel van Big Data heel reëel is. We voelen dat de wil bestaat om Data Specialisten op te leiden die in staat zijn om te werken in een Big Data-omgeving. Deze opleidingen die al dan niet kwalificerend zijn, willen op de Belgische arbeidsmarkt specialisten introduceren die op de hoogte zijn van statistiek, data-analysetechnologieën, gegevensbanken, programmeren, BI en Data Mining.

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

62. MOOC: zie glossarium

63. UTT: <http://www.utt.fr/fr/toute-l-actualite/ouverture-mooc.html?platform=hootsuite> (link in het Frans)

De meningen zijn verdeeld over de werking van de lessen. Kiest men best voor een 'Big Data-cursus' of integreert men die Big Data niet beter in een cursus geneeskunde bijvoorbeeld?

Volgens Max De Wilde en Raphael Hubain, post-doc en doctorandus aan de ULB bij het departement Wetenschappen en Informatie- en Communicatietechnologieën *'is het minder interessant om een master data science te creëren dan deze discipline te integreren in het kader van een cursus zoals bijvoorbeeld geneeskunde, economie of een andere cursus'*.

Big Data zou dus voor sommigen meer impact hebben indien de cursus wordt geïntegreerd in een specifiek toepassingsveld: geneeskunde, economie, engineering, enz. Het is ook een vaststelling die werd gedaan bij Evoliris tijdens het rondetafelgesprek over de ontwikkeling georganiseerd in 2015 door Evoliris met een aantal deskundigen inzake deze materie: Ivan Jureta (onderzoeker bij het FNRS), Stéphane Faulkner⁶⁴, Bruno Schröder⁶⁵, en Nicolas Bauwens (pedagogisch adviseur en opleider bij het departement management en multimedia ICT bij Bruxelles Formation).

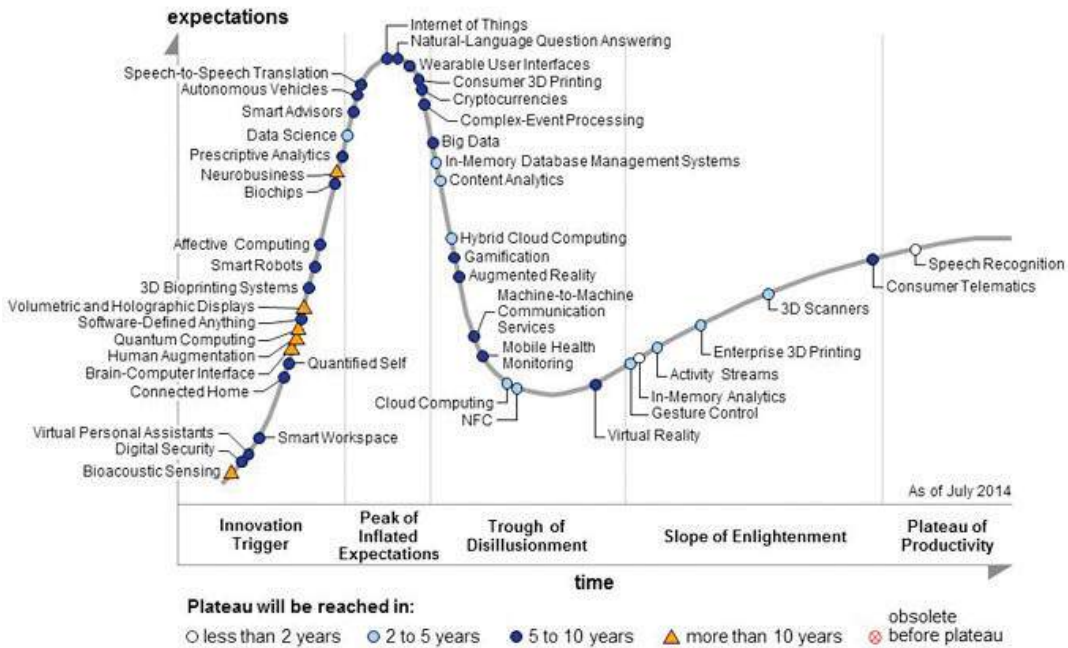
Naar het voorbeeld van het gebruik binnen het bedrijf moet Big Data ook een plaats krijgen in de schoolcursus. België en Brussel zijn nog niet rijp inzake het onderwijzen van Big Data, maar de wil bestaat om de zaken te doen vooruitgaan. Zoals wij hebben gezien worden er zowat overal heel interessante initiatieven ontwikkeld en op alle competentieniveaus.

64. Zie dankwoord/
65. Zie dankwoord

Besluiten en aanbevelingen

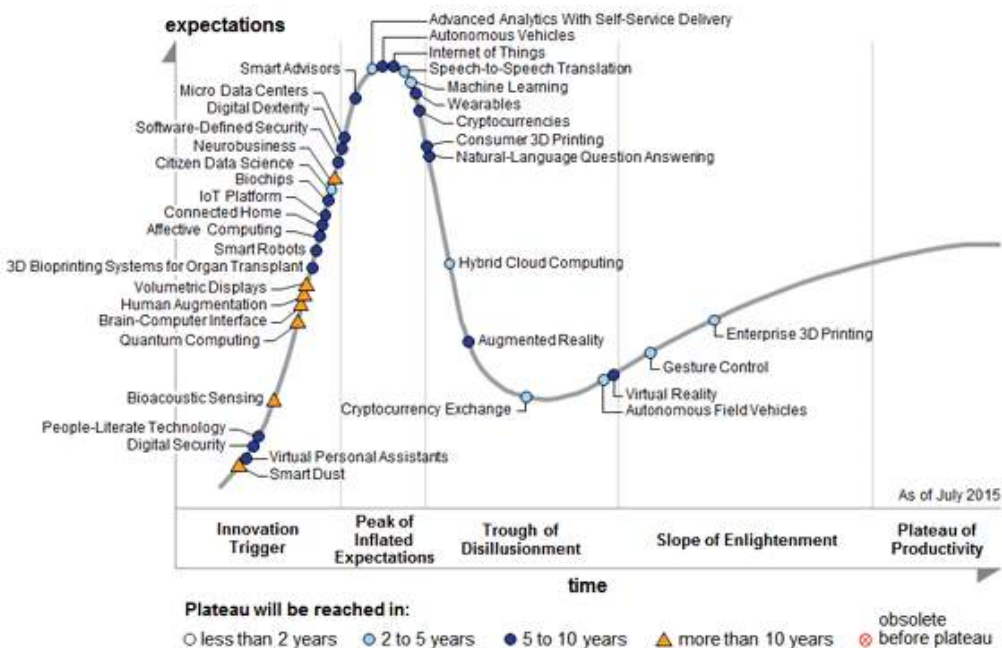


Is Big Data vandaag de dag een trendy technologie? Ja, maar die evolueert.



Afbeelding 4: de 'Hype Cycle' 2014 van Gartner

In de Hype Cycle (zie ook de nota onderaan pagina 12 en pagina 7) van 2014, positioneert Gartner de Big Data in de categorie 'trough of Disillusionment', vlakbij de categorie 'peak of inflated Expectations'. Het is dus een concept dat volop aan het rijpen is. De Amerikaanse maatschappij is van mening dat de 'productiviteitsgrens' binnen 5 tot 10 jaar zal bereikt worden. Er wachten Big Data dus mooie dagen.



Afbeelding 5: de 'Hype Cycle' 2015 van Gartner

Big Data mogen dan wel deel uitmaken van Hype Cycle in 2014, toch schitteren ze in 2015 door hun afwezigheid. Waarom? Betsy Burton⁶⁶ die verantwoordelijk is voor de realisatie van de Hype Cycle bij Gartner leert ons dat Big Data geen reden meer hebben om voor te komen in dit schema. Het is een volwaardige technologie geworden die voorkomt in andere cycli (niet alleen meer in de trendy technologieën). Big Data is in 2015 geen TRENDY technologie meer, aangezien het voortaan een technologie is die volop aan het rijpen is op internationaal niveau en die perfect kan worden geïntegreerd in alle economische sectoren van een land⁶⁷.

De maturiteit van deze technologie biedt veiligheid inzake investeringen. Er zijn minder risico's om van start te gaan. Het is inmiddels een heel nuttige tool om van Brussel een Smart City te maken, een innovatieve stad. De verschillende projecten van het CIBG getuigen van de wil om te evolueren naar een intelligent Brussel⁶⁸.

Hoe zit het in Brussel en in België?

Een stad als Brussel levert heel wat gegevens op. Tijdens het schrijven van dit rapport hebben wij vastgesteld dat er een rode draad is die alle thema's samenbrengt; bedrijf, tewerkstelling, onderwijs, juridische aspecten. Deze rode draad is op het terrein vrij coherent aanwezig en houdt verband met een ontbrekend bewustzijn bij actoren en gebruikers van analytische data. De markt van de Big Data is nog niet helemaal rijp in België en in het bijzonder in Brussel. Tijdens de opmaak van dit rapport hebben we telkens geprobeerd om het accent te leggen op deze problematiek. Waarom? Omdat er te weinig wordt gesensibiliseerd over de positieve aspecten, en vooral over wat Big Data werkelijk zijn. De Brusselse bedrijven hebben nog niet de opportuniteiten benut die ze zouden kunnen benutten door er gebruik van te maken. Velen zijn ten onrechte van mening dat deze technologie slechts een zaak is van grote bedrijven, van heelgrote bedrijven die hoofdzakelijk werkzaam zijn in de marketingsector. Wel, wij hebben een reeks concrete voorbeelden voorgesteld binnen het activiteitendomein die niet commercieel zijn. De hoofdstad die wil evolueren in de richting van een Smart Brussel wenst trouwens te sensibiliseren voor het positieve van het gebruik van gegevens, zowel voor openbare doeleinden als privé. Op de keper beschouwd, stellen we een zekere incoherentie vast. Enerzijds ontwikkelen Big Data zich goed internationaal, de Belgische en Brusselse data-analysebedrijven floreren en hebben veel bijval omdat ze werken met buitenlandse klanten. Anderzijds staan Brusselse bedrijven huiverig tegenover het analyseren van gegevens op grote schaal. Het is een verontrustende maar kenmerkende vaststelling; de Belgische bedrijven die gespecialiseerd zijn in het analyseren van gegevens hebben geen of weinig Belgische klanten. De Big Data-beroepen scheren nochtans hoge toppen. De profielen Data Analyst en Data Scientist worden weggeplukt; ze zijn goud waard!

Het is hoog tijd dat het tij keert en dat bedrijven het nut inzien van het aanwerven van gegevensspecialisten. Jammer genoeg is er een rem in België: Belgische specialisten zijn zeldzaam. Waarom? Er bestaan tegenwoordig in Brussel en in België heel weinig cursussen waar de Big Data-materie is opgenomen. De programma's bieden te weinig materie die verband houdt met gegevensanalyse. We stellen niettemin een trage bewustwording vast.

Ten slotte is alles verbonden met elkaar: onderwijs, tewerkstelling, bedrijven. Dat rijpingsproces moet op alle fronten gebeuren en komen van alle actoren uit de sector om te evolueren naar een technologie die innovatie en tewerkstelling genereert.

66. VIDEO: uitleg van Betsy Burton over de cyclus van Hype in 2015: <http://www.gartner.com/technology/research/hype-cycles>.

67. Alex Woodie, Why Gartner Dropped Big Data Off the Hype Curve Dans Datanami [online] <http://www.datanami.com/2015/08/26/why-gartner-dropped-big-data-off-the-hype-curve/> (geraadpleegd op 17 juni 2016).

68. Voor meer informatie over dit project en over andere projecten: raadpleeg katern 35 van het CIBG hier: http://cibg.brussels/nl/nieuws_publicaties/publicaties/katernen/nr-35-vier-centrale-projecten-van-smartcity-brussels

We mogen verwachten dat er binnen 4-5 jaar, met de voltooiing van het Europees project 'Horizon 2020' ernstige opleidingen inzake Big Data zullen worden gegeven zodat er op de professionele arbeidsmarkt competente Big Data-werknemers zullen te vinden zijn.

We kunnen besluiten door te zeggen dat Big Data een mooie toekomst tegemoet gaat in Brussel. Het is een hele efficiënte tool die al overal ter wereld zijn sporen heeft verdiend. Deze toekomsttechnologie zal zich ontwikkelen en zal zich aanpassen aan de vereisten van de tijd.

7.1 Aanbevelingen

Er zijn drie aanbevelingspunten.

7.1.1 Sensibiliseren van ondernemingen en van het grote publiek

Big Data is meer dan een tool bestemd voor grote bedrijven zoals Facebook en Google. Ook de grote of kleine Brusselse bedrijven kunnen daarvan genieten. Dat is de reden waarom de gegevenswetenschap zich aanpast aan de kenmerken van elk bedrijf. Men moet stoppen met bang te zijn voor die Big Data en we moeten het modiefenomeen dat eraan vasthangt, overschrijden. Zich informeren over concrete Belgische gevallen en over de realiteit op het terrein is belangrijk; het biedt de mogelijkheid zich bewust te worden van het belang.

'Analytics impliceert een stijging van de inkomsten en een daling van de kosten. Het zijn gegevens waaraan bedrijven niet denken en die ze slecht gebruiken: Big Data biedt de mogelijkheid om anders te kijken', aldus Laurent Fayet, voor LBFayet Consulting⁶⁹ tijdens ons gesprek.

Sensibiliseren gebeurt door in dialoog te treden met bedrijven. De aanpak bestaat erin om uit te leggen waaruit Big Data bestaat en om te praten over concrete gevallen en eenvoudige toepassingen. Big Data is een gegeven dat moet weggehaald worden van die complexe debatten en er moet eenvoudig en efficiënt verteld worden over de realiteit van het onderwerp. De bedrijven willen twee zaken weten:

- *Waarom dienen Big Data in mijn bedrijf, dat vandaag heel goed werkt zonder?*
- *Zal dit veel investeringen van mij vergen?*

Er moet op deze vragen worden geantwoord op de meest eenvoudige manier, en aangepast aan de noden. Het is dus nodig om geval per geval te werken. Big Data dat is ook: zijn individuele noden te kennen geven en erop antwoorden met een algoritme en een infrastructuur die antwoorden en strategische oplossingen biedt.

We moeten ook het grote publiek sensibiliseren voor het positieve aspect van de gegevenswetenschap, voorbeelden tonen van concrete gevallen die de mogelijkheid hebben geboden om een aspect van hun leven te verbeteren dankzij de Big Data. Er moet worden gecommuniceerd om aan te tonen dat het grote publiek al dagelijks geconfronteerd wordt met Big Data. Deze communicatie moet gericht zijn op een concrete uitleg van positieve en negatieve punten van deze technologie.

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

7.1.2 Meer gegevensspecialisten opleiden in België

Er moet een reeks van initiatieven gelanceerd worden om meer te weten te komen over Big Data. Een cursus Big Data creëren of deze materie integreren in een cursus. Big Data moet iets zijn dat men niet mag negeren. Er is vraag naar Data Scientists en bedrijven zoeken hun specialisten in het buitenland. Claude Delcour, stichter van het bedrijf Vadis⁷⁰, benadrukt dat het momenteel **in België heel problematisch is om mensen te vinden die aan de slag kunnen in de gegevenssector**. Het lijkt dan ook meer dan aangewezen om de cursussen die al aanwezig zijn op de opleidingsmarkt onder de aandacht te brengen. De voordelen van deze cursussen aantonen qua werkgelegenheid en communiceren over deze opleidingen. Big Data mag echter niet alleen maar voorkomen in cursussen van informatici, ingenieurs, of wiskundigen. Big Data moet ook worden opgenomen in andere disciplines zoals economie, geneeskunde, cultureel beheer, enz. Dat zal de integratie en de bewustwording van het potentieel vergemakkelijken.

Door de mensen te sensibiliseren voor de Big Data-problemen is het mogelijk om in de toekomst bachelors in die richting af te leveren. Zoals we zagen worden momenteel heel gekwalificeerde niveaus gevraagd.

7.1.3 Bedrijven bewust maken van het juridisch aspect

Ten slotte moet men letten op het juridisch aspect en moet men zich goed informeren waartoe men zich verbindt als men gegevens gebruikt. De transparantie is fundamenteel en biedt de mogelijkheid om efficiënt, betrouwbaar en integer te zijn. Er moet ook worden gesensibiliseerd over de notie van persoonlijke gegevens. Het grote publiek heeft de neiging om Big Data te associëren met grote bedrijven die persoonlijke gegevens van gebruikers gebruiken om winst te maken. Via communicatiecampagnes moeten de overheden uitleggen dat de ingezamelde gegevens gedepersonaliseerd worden en vaak kunnen helpen om een probleem op te lossen van openbaar nut, zoals inzake mobiliteit of veiligheid.

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

Dankwoord

8

Voor het opmaken van dit rapport deden wij een beroep op deskundigen uit het domein die hun kennis en knowhow wilden delen inzake Big Data. Wij danken hen voor de tijd en de aandacht die zij voor ons vrijmaakten.

Ferdinand Cassier

AGORIA - <http://agoria.be/>

Business Group Leader bij Agoria, Belgische Federatie van de Technologische Industrie.

Voor meer info :

<https://www.linkedin.com/in/ferdinandcassier/en>

<https://twitter.com/ferdicassier>

Jeroen Franssen

AGORIA - <http://agoria.be/>

Deskundige: Talent & Labour Market op het vlak van ICT - AGORIA. Hij wordt gevraagd voor zijn kennis van ICT in academische middens

Voor meer info :

<http://agoria.be/www.wsc/rep/prg/AplContent?SessionLID=3&vUserID=999999&ENewsID=105062>

Juan Bossicard

Software.Brussels - <http://www.softwareinbrussels.be>

Manager ICT & Creative Industries van de startup-cluster Impulse.Brussels

Voor meer info :

https://www.linkedin.com/in/juan-bossicard5435b72?authType=NAME_SEARCH&authToken=9HaS&locale=en_US&trk=tyah&trkInfo=clickedVertical%3Amynetwork%2CentityType%3AentityHistoryName%2CclickedEntityId%3Amynetwork_7457091%2Cidx%3A2

Max de Wilde

ULB [Departement Informatie- en Communicatiewetenschappen, Afdeling Informatie- en Communicatiewetenschappen en -technologieën] – Europese Commissie
<http://mastic.ulb.ac.be/>

Post-Doctoraat aan de ULB. Thesis: 'From information extraction to knowledge discovery: Semantic enrichment of Multilingual content with linked Open Data'. IT-taalkundige bij de Europese Commissie.

Raphael Hubain

ULB [Departement Informatie- en Communicatiewetenschappen, Afdeling Informatie- en Faculteit en -technologieën] – Biotech Quality Group
<http://mastic.ulb.ac.be/>
<http://www.thebiotechqualitygroup.com/en/>

Doctoreert aan de ULB. Data Science and Information Governance Specialist at Biotech Quality Group.

Voor meer info :

https://www.linkedin.com/in/raphaelhubain?authType=NAME_SEARCH&authToken=Oode&locale=en_US&trk=tyah&trkInfo=clickedVertical%3Amynetwork%2CclickedEntityId%3A249377895%2CauthType%3ANAME_SEARCH%2Cidx%3A1-1-1%2CtarId%3A1468505062258%2Ctas%3Ar

Céline Vanderborght

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

Smart City manager

Voor meer info :

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%2Cidx%3A1-11%2CtarId%3A1468504449829%2Ctas%3Ac%C3%A9line%20vander

Laurent Fayet

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

'Managing Partner' bij LBFayet Consulting

Voor meer info :

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

Claude Delcour

VADIS - www.vadis.com

Bestuurder en Hoofd van de Directie bij Vadis

Meester Philippe Laurent

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

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

Advocaat gespecialiseerd in recht inzake informatietechnologieën en intellectuele eigendom– IP/IT & Commercial Lawyer bij MVVP – opleider bij Legalict.be

Voor meer info :

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

Meester Jean-Marc Van Gyseghem

Kabinet Paul Muylaert Lan Officer – CRIDS

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

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

Advocaat gespecialiseerd in de informatie- en communicatietechnologieën. Onderzoeker bij CRIDS, onderzoekseenheid 'Liberties in the information society'.

Voor meer info :

<https://directory.unamur.be/staff/jmvangys>

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

Hugues Bersini

ULB – Koninklijke academie voor wetenschappen van Brussel – IRIDIA (onderzoekslaboratorium over artificiële intelligentie) - <http://iridia.ulb.ac.be/bersini>

Verantwoordelijke van IRIDIA, professor aan de ULB en lid van de koninklijke academie voor wetenschappen van Brussel.

Voor meer info : https://fr.wikipedia.org/wiki/Hugues_Bersini



Cécile Jabaudon

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

Operational Director

Voor meer info :

https://www.linkedin.com/in/c%C3%A9cile-jabaudon915a914a?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%3A%20j%20jabaudon

Stephane Faulkner

Technofutur, Cognitic, professor aan UNamur

<http://www.technofutur.be/>

<http://cognitic.be/>

Professor Technologies and Information Systems Engineering

Pierre-Nicolas Schwab

RTBf

Stichter IntotheMinds, Big Data / CRM Manager bij de RTBf

Voor meer info :

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%3A%20Pierre-Nicolas%20Schwab

Bruno Schröder

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

Microsoft Technology Officer Belgium and Luxembourg, Technologisch Directeur van MIC Belgium; beheerder van de Innovation Centers

Voor meer info :

https://www.linkedin.com/in/bruno-schroderbb57481?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%3A%20Bruno%20sch

Géraud Guilloud

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

European Project Developer

Horizon 2020 ICT & Security NCP

Voor meer info : <http://www.ncpwallonie.be/fr/equipe>

Philippe Van Impe

DIH

<https://datasciencebe.com/>

<http://di-academy.com/>

Stichter van de DIH en van de data science community in Brussel

Voor meer info : <https://datasciencebe.com/>

Tanguy De Lestré

Kabinet van de Staatssecretaris van het Brussels Hoofdstedelijk Gewest Bianca Debaets
<http://be.brussels/a-propos-de-la-region/le-gouvernement-regionallbianca-debaets>

Adviseur informatica en digitale overgang

Voor meer info : <https://www.linkedin.com/in/delestre/fr>

Nicolas Harmel

Kabinet van minister Didier Gosuin, minister van economie en tewerkstelling van Brussel
<http://be.brussels/a-propos-de-la-region/le-gouvernement-regionalldidier-gosuïn>

Attaché bij het Kabinet van minister Didier Gosuin

Voor meer info : <http://be.brussels/a-propos-de-la-region/le-gouvernement-regionalldidier-gosuïn>

BPMN : Business Proces Model and Notation (BPMN) is een standaard voor het modelleren van bedrijfsprocessen dat een grafische weergave geeft om bedrijfsprocessen weer te kunnen geven in een Business Proces Diagram (BPD). https://nl.wikipedia.org/wiki/Business_Process_Model_and_Notation

CIB : Centrum voor Informatica voor het Brussels Hoofdstedelijk Gewest.
Website: http://cibg.brussels/nl?set_language=nl

DATA MINING : de bedoeling bestaat erin om knowhow of kennis te extraheren uit grote hoeveelheden gegevens, op automatische of semi-automatische wijze.
<https://nl.wikipedia.org/wiki/Datamining>

DATAWAREHOUSE : wat zoveel betekent als 'opslag van gegevens' wijst op een gegevensbank die wordt gebruikt om informatie in te zamelen, te ordenen en op te slaan afkomstig van operationele databanken en om zodoende een basis aan te bieden om beslissingen te helpen nemen in een bedrijf. <https://nl.wikipedia.org/wiki/Datawarehouse>

Gartner : Gartner is een bedrijf dat erin slaagde om zich op te werpen als mentor die advies geeft over geavanceerde technologieën. Gartner is bijzonder gekend voor zijn 'Hype Cycle' (analyse van de technologische trends per jaar in de vorm van een levenscurve) en de 'Magic Quadrant' (analyse van de troeven van de beste bedrijven die werken binnen het domein van een bepaalde geavanceerde technologie door de volledigheid van de visie en de uitvoeringscapaciteit van de technologieën te vergelijken).
De website: <http://www.gartner.com/technology/home.jsp>

IOT : de samentrekking van 'Internet of Things' of 'Internet van objecten' staat voor de uitbreiding van Internet met dingen en plaatsen uit de fysieke wereld. Het internet van objecten staat voor de uitwisseling van informatie en gegevens afkomstig van middelen aanwezig in de real world, naar het internet. Bijv.: intelligente koelkast, interactieve uurwerken,... https://nl.wikipedia.org/wiki/Internet_der_dingen

IRIDIA : Instituut voor Interdisciplinair Onderzoek en Ontwikkelingen inzake Artificiële Intelligentie aan de Université Libre de Bruxelles.
Website: <http://iridia.ulb.ac.be/bersini/>

LBFayet Consulting : advies inzake het definiëren en implementeren van Big Data-strategieën en activiteiten om personeel te plaatsen inzake Big Data. De website: <http://www.lbfc.eu/>

Microsoft Azure Data Lake : tool die de mogelijkheid biedt om elk type gegevens op te slaan in de verschillende formaten in een uniek referentiesysteem.

Microsoft Cortana Analytics : analyseert de informatie met behulp van een intelligent algoritme afkomstig van Machine Learning

MOOC: *massive open online course*. De MOOC is een online open cursus op grote schaal die men kan associëren met een open afstandsonderwijs.

OPEN DATA : is een digitaal gegeven van publieke of private aard. Het gegeven kan worden geproduceerd door een gemeenschap, een openbare dienst (eventueel afgevaardigde) of een bedrijf. Dit gegeven wordt verspreid op gestructureerde wijze volgens een open methode en licentie die de vrije toegang garandeert en het hergebruik door iedereen, zonder technische, juridische of financiële beperkingen.

https://nl.wikipedia.org/wiki/Open_data

Smart City : is een stad die gebruik maakt van informatie- en -communicatietechnologieën (ICT) om de kwaliteit te 'verbeteren' van de stedelijke diensten of om de kosten te beperken.

https://nl.wikipedia.org/wiki/Slimme_stad

UML : de eengemaakte modelleringstaal staat voor Unified Modeling Language (UML). Een grafische modelleringstaal... op basis van pictogrammen ontworpen om een genormaliseerde methode te bieden voor het visualiseren van het ontwerp van een systeem. Het systeem wordt courant gebruikt in de software-ontwikkeling en voor het gericht objectontwerp. https://nl.wikipedia.org/wiki/Unified_Modeling_Language

EVOLIRIS vzw is het Brusselse Beroepenreferentiecentrum binnen het domein van de Informatie- en Communicatietechnologieën (ICT). Vanaf de oprichting in 2006 bestond de opdracht van het centrum erin om de ICT-beroepen te sensibiliseren, te informeren en op te leiden. EVOLIRIS is hoofdzakelijk werkzaam bij drie doelgroepen: **leerlingen en lesgevers**, voornamelijk in het kader van de sensibiliseringscampagne voor beroepen, **de werkzoekenden**, dankzij ons opleidingsaanbod en onze hulp bij begeleiding tijdens de zoektocht naar een job. En ten slotte **alle werknemers**, in het bijzonder de werknemers uit de sector die dankzij onze opleidingen kunnen evolueren binnen hun beroepsdomein.

Het Team

Jean-Pierre RUCCI

Directeur

Jp.Rucci@evoliris.be | 02 475 20 02

Latifa BENYAHIA

Directieassistente

Latifa.Benyahia@evoliris.be | 02 475 20 03

Myriam POELMAN

Projectleider

Myriam.Poelman@evoliris.be | 02 475 20 05

Luc HUYGH

ICT-coördinator. Projectleider

Luc.Huygh@evoliris.be | 02 475 20 06

Naïma ZEGOUD

Projectassistente (Office manager)

Naima.Zegoud@evoliris.be | 02 475 20 04

Florence TUTS

Secretariaat en onthaal

Florence.Tuts@evoliris.be | 02 475 00

Christina GALOUZIS

Projectleider

Christina.Galouzis@evoliris.be | 02 475 20 08

Kelly VANPUTTE

Referente-adviseur ACTIRIS

Kelly.Vanputte@evoliris.be | 02 475 20 01

Alexandre HEMERIJCK

Referent-adviseur ACTIRIS

Alexandre.Hemerijck@evoliris.be | 02 475 20 07

Voor meer inlichtingen : www.evoliris.be

Contact :

Paalstraat, 14 - Gebouw A | 1080 Brussel | 02 475 20 00 | info@evoliris.be

Redacteur

Christina Galouzis

Post-editor

Luc Huygh

Verantwoordelijke uitgever

Jean-Pierre Rucci

Grafisch ontwerp

www.sergeantpaper.be

Druk

DADDY KATE

Albert Van Cotthemstraat 54

1600 Sint-Pieters-Leeuw

www.daddykate.be

Contact

Evoliris vzw

Paalstraat, 14 A

1080 Brussel

02/475 20 00

info@evoliris.be

www.evoliris.be



- Inleiding en methodologie
- Big Data: afbakening van het onderwerp en herkomst
- Big Data in Europa
- Big Data in Brussel
- Big Data-beroepen
- Big Data-opleidingen
- Besluiten en aanbevelingen