



Trendletter Big Data

Inhoud

1.	Big What?	2
1.1	Wat is 'Big Data'?	2
1.2	Hoe werkt Big Data?	3
1.3	Hoe nieuw is Big Data?	3
1.4	Open Data versus Big Data	4
1.5	Privacy	4
2.	Big Data in Antwerpen	4
2.1	Binnen de stad	4
a.	Energie	4
b.	Luchtkwaliteit	5
c.	Mobiliteit	5
2.2	Binnen de Antwerpse sectoren	6
a.	Chemie	6
b.	E-health	6
c.	Big Data en creativiteit	7
3.	De toekomst van Big Data	8
3.1.	Overstromende meren van data	8
3.2.	Nieuwe jobprofielen	8
4.	Kansen voor de stad	10
5.	Volgende trendletter: City of Things	12





1. Big What?

1.1 Wat is 'Big Data'?

Big Data: een grote hoeveelheid informatie, altijd digitaal. Video, audio, tekst en cijfers. Wanneer noem je informatie of data 'big'? De definitie is vaag. De hoeveelheid moet gewoon 'groter zijn dan gangbaar'.

De ware betekenis van Big Data zit hem in de verwerking van al die info. De hoeveelheid is zo groot dat je ze niet zomaar verwerkt. Het vraagt nieuwe, aangepaste analysesystemen om ze te doorploeteren en er zinvolle verbanden uit te filteren.

Ook de snelheid waarmee de data binnenkomen of opgevraagd kunnen worden, speelt een belangrijke rol. Soms gaat het zo snel dat men spreekt over real-time data. Hoe sneller er nieuwe data beschikbaar zijn, hoe meer rekenkracht je nodig hebt om ze onmiddellijk mee op te nemen.

Ander belangrijk punt: de diversiteit van de informatie. Hoe meer verschillende soorten data, hoe ingewikkelder de analyse ervan.

De grote uitdaging van Big Data? Heldere conclusies filteren uit een onoverzichtelijk kluwen informatie.^{1 2}

Supermarktketen Target: pionier in klantenbinding dankzij Big Data

Retailers, en vooral grote supermarktketens, houden massa's gegevens bij over hun klanten. Om dan gerichte reclameacties te voeren. De Amerikaanse supermarktketen Target is daarin een voorloper.

Target verkoopt ongeveer alles, van kleding over voeding tot schoonmaakproducten en multimedia. Het probleem: winkelgedrag is vooral routine. Je winkelt met één doel voor ogen. Klanten komen naar Target om – bijvoorbeeld – schoonmaakproducten te kopen. Het is moeilijk om hen aan te zetten ook andere aankopen te doen.

Maar ... Er zijn ook momenten waarop je geneigd bent om je winkelgedrag aan te passen. De komst van een kind is zo'n moment. Bestaande gewoontes veranderen plots, je hebt nieuwe producten nodig: luiers, fopspenen, speelgoed, ... Je hebt minder tijd om te winkelen. En daarom koop je voortaan verschillende producten in *dezelfde* winkel.

Een zwangerschap is een buitenkans om klanten te binden. Daarvan zijn winkelketens zich bewust. Bovendien weten ze die nieuwe winkelroutine snel te verankeren in de gewoontes van klanten. Daarom ging Target op zoek naar patronen in databestanden die wijzen op zwangere vrouwen. Die bereikten ze met specifieke promo-acties, nog voor de geboorte van hun kind. En dat dankzij Big Data.

Het verhaal van de tienerdochter

Een winkelverantwoordelijke van Target kreeg een boze vader over de vloer. Zijn tienerdochter had een reclamefolder ontvangen voor babyspullen. Persoonlijk aan haar gericht. De winkelverantwoordelijke stond met de mond vol tanden, want hij was niet betrokken bij de marketingstrategie. Hij bood zijn excuses aan. Net als de winkelmanager een paar dagen later, via de telefoon. De vader was ondertussen óók van slag: zijn dochter had net opgebiecht dat ze zwanger was.

¹ <https://decorrespondent.nl/296/Wat-is-Big-Data-/14414312-a609db7d>

² https://nl.wikipedia.org/wiki/Big_data



1.2 Hoe werkt Big Data?

De grote uitdaging bij Big Data is de hoeveelheid informatie verwerken. Big Data gebruikt een groot aantal computers voor de opslag en de verwerking van data, de zogenaamde *distributed storage* en *distributed processing*. Het platform Hadoop koppelt bijvoorbeeld duizenden computers aan elkaar en genereert zo een immense rekenkracht.

Maar Big Data is meer dan rekenkracht. Het betekent ook geavanceerde analysetechnieken, en data effectiever en efficiënter laten verwerken. Google is daarvan een goed voorbeeld. De zoekmachine bundelt en categoriseert een enorme hoeveelheid info. Het maakt daarbij gebruik van een leger aan servers. Het grote verschil met oudere zoekmachines zit hem in de algoritmes die Google loslaat op die info. Naast de totale computercapaciteit is de aansturing ervan dus minstens even belangrijk.^{3 4 5 6}

Snel, sneller, FLITS

Hoe lang duurt het voor een gebruiker een resultaat krijgt? Dat hangt af van de reactiesnelheid van een computersysteem. Die bepaalt in grote mate de gebruiksvriendelijkheid ervan. Maar in bepaalde sectoren draait snelheid om meer dan gebruiksgemak. Typisch voorbeeld: de financiële sector, en dan in het bijzonder de beurshandel.

‘Buy low, sell high’, klinkt het beursdevies. Koop aandelen laag, en verkoop ze wanneer de koers gestegen is. Strijk dan de winst op. Aandelenkoersen zijn jammer genoeg grillig en fluctueren de hele dag door. Een speciale tak van de financiële sector legde zich toe op die kleine fluctuaties: de flitshandel of *high-frequency trading*.

Bij flitshandel koopt en verkoopt de computer automatisch aandelen. Hij beslist op basis van algoritmes. Op een fractie van een seconde spelen flitshandelaars in op prijsfluctuaties. Om orders supersnel te verwerken, investeren flitshandelaars in de meest geavanceerde computers en dataverbindingen. Om die computers aan te sturen en maximaal te laten presteren, gebruiken ze inzichten en algoritmes uit Big Data. Snelheid is van het grootste belang: een tijdswinst van 1 milliseconde - $1/1000^{\text{ste}}$ van een seconde - in de verwerking van de aan- of verkooporders ... levert een flitshandelaar ongeveer 100 miljoen dollar op per jaar.⁷

1.3 Hoe nieuw is Big Data?

Big Data is niets nieuws: computers gebruiken we al zo'n 75 jaar om grote hoeveelheden informatie te verwerken. De Britten gebruikten tijdens de Tweede Wereldoorlog Colossus om de Duitse Enigma-code te kraken. Colossus, de eerste elektronische computer uit 1943, onderzocht de Duitse codes op terugkerende patronen.

De oorzaak van de huidige hype rond Big Data zit hem in de explosieve groei van de hoeveelheid data. Die neemt zo snel toe dat ook de verwerkingsystemen moeten mee-evolueren. De toename in internetverkeer is maar één element in die groeiende berg data. Toch biedt volgende tabel een goede illustratie van de groeisnelheid:

Groei van het internetverkeer (Cisco, 2016)

³ http://www.sas.com/en_us/insights/analytics/big-data-analytics.html

⁴ <http://bigdata.teradata.com/US/Big-Data-Quick-Start/How-Big-Data-Works/>

⁵ <http://www.theatlantic.com/sponsored/delta-view-from-the-top-2014/how-big-data-works/115/>

⁶ <http://computerworld.nl/big-data/75185-wat-is-hadoop>

⁷ <http://www.sciencedirect.com/science/article/pii/S2405918815000045>



Jaar	Wereldwijd internetverkeer
1992	100 GB per dag
1997	100 GB per uur
2002	100 GB per seconde
2007	2.000 GB per seconde
2015	20.235 GB per seconde
2020	61.386 GB per seconde*

* schatting

1.4 Open Data versus Big Data

Open Data en Big Data zijn twee verschillende concepten. In de praktijk komen ze in de stad samen. De stad stelt haar data ter beschikking van het brede publiek. Dat zijn dan Open Data. De hoeveelheid of aard van de data doet er niet toe. Hoewel: de stad streeft er wel naar om die data bruikbaar aan te bieden.

Omdat de stad wettelijk verplicht is alle data ter beschikking te stellen, valt daar per definitie ook de Big Data onder. Belangrijk aandachtspunt: de stad mag alleen anonieme data vrijgeven. Ze moet altijd de privacy van burgers garanderen.

1.5 Privacy

Door de complexiteit van Big Data is het soms moeilijk om privacy te garanderen. Wie verschillende databronnen met elkaar verbindt, kan mogelijks toch nog persoonlijke gegevens filteren uit geanonimiseerde data. Dat moet uiteraard koste wat het kost uitgesloten worden. Niet alleen om juridische redenen, maar ook om het vertrouwen te behouden. Want als mensen vermoeden dat hun privacy onvoldoende beschermd is, stellen ze hun gegevens niet langer ter beschikking. En dat ondergraaft betekenisvolle analyses.

2. Big Data in Antwerpen

2.1 Binnen de stad

a. Energie

Door gebouwen uit te rusten met sensoren verzamel je een schat aan waardevolle informatie. Met realltime data stel je verwarmings- en verluchtingssystemen af op het aantal aanwezige mensen en de weersomstandigheden. Met een doorgedreven analyse van de data schat je de noden in, op verschillende momenten van de dag. En dat altijd maar accurater. Koppel je – bijvoorbeeld – het verwarmingssysteem aan het boekingsysteem van vergaderzalen? Dan verwarm je die zalen volgens het geplande gebruik.

Bij een samenwerking tussen de Amerikaanse stad Seattle, Microsoft en dienstenbureau Accenture werden 90 gebouwen uitgerust met sensoren. De samenwerking verminderde het energieverbruik met 10% tot 25% per gebouw.⁸

⁸ <https://gcn.com/articles/2016/07/01/gis-mosquito-zika.aspx><https://gcn.com/Articles/2013/08/02/big-data-cities-seattle.aspx?Page=2>



Een andere mogelijkheid zijn de slimme meters. Zij registreren het energieverbruik én stellen de info tegelijk ter beschikking van de gebruiker. Zo stuur je je energieverbruik bij. Als je die individuele gegevens verzamelt in een overkoepelend platform, komt er een schat aan informatie vrij. Daarmee voer je dan een gericht energiebeleid.

Slimme meters zijn een geschenk uit de hemel voor netwerkbeheerders. Energie uit duurzame bronnen, zoals wind- en zonne-energie, is onvoorspelbaar. Vraag en aanbod zijn moeilijk af te stemmen. Als de vraag naar energie in real time beschikbaar is via zo'n platform, vereenvoudigt dat de taak van netbeheerders. Dat komt het hele proces van duurzame energie ten goede.

Om mensen te overtuigen om mee te stappen in zo'n systeem, is het cruciaal om de privacy van de individuele energieverbruiker te garanderen, onder meer door geaggregeerde data te gebruiken.

b. Luchtkwaliteit

Luchtkwaliteit meten we in Vlaanderen met vaste meetstations van de Vlaamse Milieumaatschappij (VMM). Dat levert gedetailleerde en wetenschappelijk bruikbare metingen op. Maar die vaste meetpunten zijn duur.

In Antwerpen meten we ook mobiel, onder meer in samenwerking met B-post. Sensoren op hun voertuigen maken voortdurend metingen door heel de stad. Het probleem? Die metingen zijn niet kwalitatief genoeg om wetenschappelijk te zijn.

Meer betaalbare sensoren zouden meer betrouwbare meetresultaten opleveren, voor een lagere prijs. Zo wordt het mogelijk om beter in te spelen op de actuele luchtkwaliteit. En met die gegevens neem je verkeersmaatregelen die mobiliteit én luchtkwaliteit verbeteren.

Een andere mogelijkheid, die al in Elsene is gelanceerd: een mobiele app voor voetgangers en fietsers. De app biedt hen de snelste en gezondste route aan, op basis van realtime-informatie.⁹

c. Mobiliteit

Veel steden doen nu al een beroep op parkeergeleidingssystemen. Die verwerken gegevens van parkeergarages in real time. Dynamische verkeersborden leiden bestuurders naar vrije parkeerplaatsen. Bijkomende sensoren zijn niet nodig: de systemen meten alleen het in- en uitgaande verkeer van de garages.

Het vraagt wel bijkomende investeringen om elke aparte parkeerplaats in de stad op te nemen in zo'n parkeergeleidingssysteem. Want elke parkeerplaats krijgt een eigen sensor.

Een andere, meer gerichte toepassing van parkeersensoren zie je bij laad- en loszones. Zo loopt er in de Van Schoonhovestraat en de Van Wesenbkestraat een proefproject om foutparkeerders te detecteren. Wanneer de metingen wijzen op een mogelijke foutparkeerder, krijgen parkeewachters in de buurt een verwittiging. Zij gaan dan ter plaatse kijken.

Een andere toepassing binnen Antwerpen: de metingen van fietsers. Zowel via vaste als mobiele telpunten. Het grote nadeel van mobiele telpunten? Ze geven een momentopname en zijn weersafhankelijk. Bij regen nemen er minder mensen de fiets. Meer vaste meetpunten op strategische plaatsen bieden in de nabije toekomst meer gegevens. Die zorgen ervoor dat we het beleid beter afstemmen op de noden van de Antwerpenaar.

⁹ <https://www.youtube.com/watch?v=61B5t212bz8&index=24&list=PLTdre3HM8tOrA9y5qmqIc56dxwmie8EFI>



2.2 Binnen de Antwerpse sectoren

a. Chemie

Antwerpen huisvest de grootste chemische cluster van Europa. De chemische installaties in de haven zijn enorm. Ze tellen tot 500.000 sensoren om temperatuur, druk, debiet en andere waarden te meten. Die gegevens gaan naar een centrale databank. Medewerkers bekijken in de controlekamers schermen vol tijdseries. Zij gaan op zoek naar onregelmatigheden en gevaren.

De software van Trendminer, een Big Data-bedrijf uit Hasselt, gaat aan de slag met de data. Het indexeert, labelt, analyseert en monitort. Het voorspelt incidenten aan de hand van algoritmes.

Trendminer heeft al een mooie portefeuille aan (petro)chemische mastodonten: Total, Solvay, Evonik Borealis, ... Maar ook farmabedrijven (Janssen Pharma), grondstoffenverwerkers (Umicore) en industriële spelers (ThyssenKrupp, Stora Enso) zijn er klant.¹⁰

De software automatiseert een deel van de taken. Gaan er dan jobs verloren? De operatoren behouden een cruciale rol bij een incident. En er is sowieso een tekort aan technisch geschoold personeel. De gedeeltelijke automatisatie ondersteunt de verdere groei van de chemische sector in Vlaanderen.

b. E-health

E-health, dat is het gebruik van informatie- en communicatietechnologie in de gezondheids- en welzijnssector.¹¹ Een typevoorbeeld? De elektronische uitwisseling van gezondheidsgegevens. Ziekenhuisartsen, huisartsen en specialisten delen patiëntgegevens met elkaar. In Antwerpen trekken startende e-healthbedrijven naar BlueHealth Antwerp. Deze incubator¹² coacht en ondersteunt startups, zowel financieel als inhoudelijk. Zij verwerken uitgebreide, complexe datasets uit de gezondheidssector. Waarom? Om bijvoorbeeld efficiënter te werken, of de evolutie van een patiënt nauwer op te volgen.

Het Antwerpse Appi@Home helpt dokter en patiënt

Appi@Home staat voor 'Antwerp Personalized Pain Initiative at Home'. Het platform bundelt patiëntendata, en zet ze om in een vlot leesbaar en bruikbaar rapport. Waar haalt het die data? Uit pijndagboeken, slaapmonitors, zuurstofsaturatiemeters, bloeddrukmeters, en uit de fysieke activiteiten van de patiënt via een 'activity tracker'. Appi@Home geeft zowel patiënt als zorgverstreker een duidelijk inzicht in de impact van de behandeling op het klachtenpatroon.

Het pilootproject geeft puike resultaten. Patiënten raken meer betrokken en geresponsabiliseerd. Met een kortere behandelingsperiode en een snellere socio-economische re-integratie als gevolg.

Appi@Home won de Agoria e-Health Awards 2016. Het mag zich het beste e-Health project van 2016 noemen.¹³

¹⁰

http://homes.esat.kuleuven.be/~bdmdotbe/bdm2013/documents/Chemiebonzen%20scharen%20zich%20achter%20Vlaamse%20'Google%20van%20de%20industrie'%20%20De%20Tijd_pdf.pdf

¹¹ <https://ec.europa.eu/digital-single-market/en/eu-policy-ehealth>

¹² <http://www.bluehealthantwerp.be/>

¹³ <http://www.112antwerpen.be/index.php?mact=News,cntnt01,detail,0&cntnt01articleid=11099&cntnt01returnid=16>



c. Big Data en creativiteit

Big Data draait om patronen vinden in een grote hoeveelheid info. Dat maakt het heel geschikt voor 'droge' analyses. Voor creatieve toepassingen heeft het weinig meerwaarde. Want een computer werkt wiskundig, en is niet creatief. Of toch?

Laurence Van Elegem, marketing- en communicatiemanager van @nexxworks, breekt in een interview in Trends een lans voor de creatieve mogelijkheden van computers¹⁴. Vandaag zijn er al voorbeelden. Denk aan Chef Watson (IBM), die nieuwe recepten voorstelt door wiskunde los te laten op een gigantische berg data. Automatische vertalingen zijn steeds accurater. De creatie van nieuwe, literaire teksten? Zelfs dat lijkt te lukken. Een computersysteem schreef een kortverhaal ... en haalde de tweede ronde in een Japanse literaire wedstrijd. Toegegeven, dat is nog ver van de eerste plaats. Maar het is alweer een nieuwe stap in de richting van creatieve, intelligente computers.

Creativiteit koppelt bestaande elementen. Laat dat nu net de sterkte van Big Data zijn. De mens evalueert nu nog zelf het resultaat ... Maar misschien lukt dat de computer morgen ook? Hoe meer feedback een computer registreert, hoe kwalitatiever de voorstellen.¹⁵

Antwerpen telt verschillende bedrijven die zijn gespecialiseerd in het ontwerp van altijd maar intelligentere software. Die levert op een creatieve manier inzichten in complexe zaken, zoals mobiliteit en gezondheidstoepassingen.

Sentiance personaliseert marketing met Big Data

Het bedrijf Sentiance is specialist in contextuele marketing. Die gaat nog een stap verder dan gepersonaliseerde marketing. Op basis van realltime data onderzoekt contextuele marketing de context van een persoon. En dan biedt het specifieke reclame aan. De software schat bijvoorbeeld in of iemand op zoek is naar een restaurant. Heeft hij net gegeten? Misschien zoekt hij dan een cinemavoorstelling, of de diensttabel van De Lijn ...

¹⁴ http://trends.knack.be/economie/e-business/computers-woorden-creatiever-en-daar-zijn-we-niet-klaar-voor/article-opinion-725265.html?utm_source=Newsletter-06/07/2016&utm_medium=Email&utm_campaign=Newsletter-RNBTRKZ&M_BT=11330606509953

¹⁵ <http://www.digitaltrends.com/cool-tech/japanese-ai-writes-novel-passes-first-round-national-literary-prize/>



3. De toekomst van Big Data

3.1. Overstromende meren van data

Big Data ging vroeger vaak over vaste datasets. Vandaag is er dan weer een exponentiële toename van *streaming data*: een continue stroom aan informatie, afkomstig van – bijvoorbeeld – verkeerscamera's, weerstations of financiële bewegingen op de beurs. De voortdurende groei van databronnen vergroot tegelijk de hoeveelheid ruwe data. Soms spreken we zelfs niet meer over een dataset of databank, maar over een *data lake* – zeg maar een datameer.

'Data lake' betekent ook data anders structureren ... Maar de techniciteit daarvan zou ons te ver leiden. Belangrijker is dat de analogie mooi weergeeft waar de uitdaging ligt voor Big Data. De groeiende stroom aan nieuwe data neemt bovendien exponentieel toe. Het datameer overstroomt. Neem dat gerust letterlijk: het is onmogelijk om alle geproduceerde data bij te houden. Er gaan voortdurend data verloren. De uitdaging voor Big Data, of beter gezegd voor de analisten die met Big Data werken? Op tijd zinvolle informatie filteren uit het datameer.

3.2. Nieuwe jobprofielen

Belgische werkgevers gaven hun top tien van meest gevraagde vaardigheden. Technische skills die te maken hebben met data staan met stip bovenaan. En helemaal bovenaan staat statistische analyse en data mining.¹⁶ In Vlaanderen behoren die ICT-profielen al jaren tot de lijst van knelpuntberoepen. Bedrijfsanalist ICT, databankbeheerder, ... De *boom* in Big Data-toepassingen doet dat tekort nog toenemen, samen met alle andere trends richting digitalisering.

“In sommige bedrijven ligt het loon van de datawetenschapper drie keer hoger dan dat van de CEO.”

Citaat van Jacques Bughin (McKinsey) in Trends

De VS verwachten een tekort van 140.000 tot 190.000 analisten met *deep analytical skills* tegen 2018. Samen met een tekort van 1,5 miljoen managers en analisten, die de knowhow hebben om de geanalyseerde big data om te zetten in effectieve beslissingen. Voor Vlaanderen bestaat er geen schatting.^{17 18}

Gespecialiseerde dienstverleners in Big Data zijn een groeiende markt. Gezien de specifieke expertise besteden meer en meer bedrijven hun data-analyse uit. Tegen 2017 zou meer dan 30% van de bedrijven die werken met big data, een beroep doen op een gespecialiseerde tussenpartij.

3.3. Nieuwe business modellen

Meer en meer sensoren in producten leveren informatie. De consument die online bestelt krijgt juistere info over de leveringstermijn via *track & trace*-toepassingen.

Maar ook na levering bezorgt de groeiende hoeveelheid sensoren nuttige info: waar staat mijn wagen? Moet de olie vervangen worden? Is de bandendruk in orde? Als die informatie rechtstreeks naar de garage gaat, ontstaan misschien nieuwe diensten. Zoals een totaal onderhoudscontract, waarbij de eigenaar van de wagen niets meer zelf moet controleren.

Op consumentenniveau gaat het over een relatief beperkte hoeveelheid data. Maar voor een volledig klantenbestand geeft dat een schat aan informatie. De analyse daarvan helpt een bedrijf om betere

¹⁶ <http://www.bloovi.be/nieuws/detail/dit-zijn-de-digitale-skills-die-werkgevers-in-belgie-zoeken>

¹⁷ https://www.vdab.be/sites/web/files/doc/trends/knelpuntberoepen_2016.pdf

¹⁸ <https://www.vdab.be/trendsdoc/vacatureanalyse/knelpuntberoepenlijst2016.pdf>



inzichten te krijgen. Hoe gebruiken verschillende consumenten het product? Hoe stemmen ze hun aanbod daarop af? Big Data biedt een voedingsbodemp voor nieuwe businessmodellen. Die sluiten dan weer nauwer aan bij de nood van de consument.

Cumul.io visualiseert data

Ook in Antwerpen zijn er Big Data-analisten. Cumul.io startte in 2015 in Leuven. In 2016 opende het al een kantoor in Antwerpen.

Cumul.io specialiseert zich in de visualisatie van grote hoeveelheden data. Ze bieden een analysesysteem waarbij de klant zélf met de data aan de slag kan gaan. Van grafieken tot de koppeling aan geografische locatiedata. Cumul.io maakt daarbij ook gebruik van open data, waarbij bedrijfseigen data onmiddellijk gekoppeld worden aan bijvoorbeeld bevolkingscijfers of verkeerscondities.



4. Kansen voor de stad

Big Data op zich is niet meer dan een *asset*: een hulpmiddel of een grondstof. Verwaarloost een stad haar data? Dan kan ze die even goed niét bijhouden. De meerwaarde zit hem in de verwerking ervan. De exponentieel toenemende hoeveelheid en variëteit aan data vraagt een geëvolueerde data-analyse. We onderscheiden vier types¹⁹ data-analyse, in toenemende orde van complexiteit:

- Descriptive analysis: *wat* gebeurde er?
- Diagnostic analysis: *waarom* gebeurde het?
- Predictive analysis: wat zal er (waarschijnlijk) gebeuren?
- Prescriptive analysis: wat is nu de beste actie?

De typologie schuift op van pure constatactie naar de aanlevering van een te volgen koers. Die bepaalt in de toekomst de meest optimale resultaten. Duidelijk een moeilijke oefening. Onderzoek naar predictieve en prescriptieve analysemethoden behoort tot de voorhoede van de wetenschap.

Antwerpen zit daarin in de kop van het peloton: de stad en Digipolis onderzoeken – samen met onder meer imec – doorgedreven analysemogelijkheden. Het doel? Altijd betere analysetools ter beschikking stellen van de stad. Zulke hoogstaande data-analytics ogen soms bedrieglijk eenvoudig. Want de kunst bestaat er net in om complexe materie bevattelijk voor te stellen.

Ken je de app Antwerp-Taks? Die toont wanneer je taxi aankomt, en wanneer je op je bestemming bent. Twee tijdstippen weergeven: simpel. Maar denk even terug aan ‘Slim naar Antwerpen’. Ook daar was een eenvoudige uitkomst de vrucht van een enorme analyse.

¹⁹ <http://www.gartner.com/it-glossary/predictive-analytics/>



Peter Crombecq, directeur Digipolis:

“Big Data scheidt fantastische mogelijkheden”



Wat betekent Big Data voor stad Antwerpen?

“Een organisatie als stad Antwerpen móet met Big Data aan de slag gaan. Maar Big Data op zich betekent niets: een stad die over massa’s data beschikt is nog geen *smart city*. Een stad wordt pas slim als ze iets intelligents doet met data. Alles staat of valt met de analytics die men loslaat op de data.”

Geef eens een voorbeeld van zo’n nieuwe analytics.

“De meest gebruikte vorm van analytics is de beschrijvende vorm, de ‘Descriptive Analytics’: de temperatuur is nu 15°, het is nu druk op de weg, er zijn nu plaatsen vrij op de parking. De volgende stap in analytics is de voorspellende vorm, de ‘Predictive Analytics’: straks wordt het 20°, binnen een uurtje is het veel drukker op de weg, de parking zal volzet zijn. De ultieme vorm is de aanbevelende vorm, de ‘Prescriptive Analytics’: als u straks naar de stad rijdt, dan staan er files op de Frankrijklei, en bovendien zijn alle parkings volzet. U parkeert beter aan het justitiepaleis. Daar neemt u tram 4, dan bent u sneller op uw bestemming.”

“Daarnaast willen we met ²⁰‘Slim naar Antwerpen’ evolueren. Overigens is ‘Slim naar Antwerpen’ nu al heel slim. Want het houdt rekening met verwachte files en beveelt aan of je beter met de fiets, de auto of het openbaar vervoer gaat.”

Voor welke uitdagingen staat de stad inzake Big Data?

“Specifiek rond Big Data zijn er twee belangrijke evoluties. De eerste is het *data lake*. Waar data vroeger in gestructureerde databanken werd bewaard, is er vandaag meer en meer sprake van *data lakes*. Een meer van ongestructureerde en gestructureerde data, waarvoor aangepaste analytics nodig zijn.”

“Het *data lake* is het gevolg van de tweede evolutie: de toename van zogenaamde *streaming data*. Dat slaat op de toevloed aan realltime data, bijvoorbeeld uit nummerplaattherkenningscamera’s. Het is een grote uitdaging om die stroom van data nuttig aan te wenden.”

“Daarnaast is het een blijvende uitdaging om de nodige technische expertise in huis te halen. Een goede data-analyse bestaat erin om de essentie van een complexe materie op een simpele, bevattelijke manier weer te geven. Een datawetenschapper, de data scientist, moet daarom over een waaier aan competenties beschikken: van doorgedreven technische knowhow tot een grote dosis creativiteit. Die profielen zijn voorlopig dun gezaaid.”

Wat brengt de toekomst voor Big Data?

“Het is belangrijk om voor ogen te houden dat Big Data een middel is, maar geen doel op zich. Wat Big Data betekent, hangt dus af van waarvoor we het gebruiken. Daarom kan je Big Data niet los zien van de andere technologische ontwikkelingen die meer en meer realiteit worden, zoals het *Internet of Things*, *deep learning*, *pattern recognition* en artificiële intelligentie.”

²⁰ <https://www.slimnaarantwerpen.be/nl>



“Waar dat allemaal samenkomt, krijg je fantastische mogelijkheden. Een duidelijk voorbeeld is mobiliteit. Meer en meer databronnen zullen beschikbaar worden op een geaggregeerd niveau. De toepassingen zijn eindeloos.”

“Daarbij moeten we wel blijven waken over de mogelijke impact op de privacy. Individuele gegevens mag je niet uit de data kunnen filteren. Anders heb je een vertrouwensprobleem. Mensen zullen hun data niet meer delen. Daarom is privacy een apart aandachtspunt binnen de nieuwe tools die we ontwikkelen. Daarvoor werken we specifiek samen met de privacy-experts van de KU Leuven.”

5. Volgende trendletter: City of Things

Big Data is maar het topje van de ijsberg, een stap op weg naar de stad van de toekomst. We raakten het in deze trendletter al aan: een waaier aan voorwerpen geraakt altijd meer verbonden met het internet. De evolutie naar een Internet of Things is een uitdaging, ook voor de stad. Wat kan het Internet of Things betekenen voor Antwerpen? Welke stappen worden er al gezet? Dat lees je in de volgende trendletter.

