



Stichting NIOC en de NIOC kennisbank

Stichting NIOC (www.nioc.nl) stelt zich conform zijn statuten tot doel: het realiseren van congressen over informatica onderwijs en voorts al hetgeen met een en ander rechtstreeks of zijdelings verband houdt of daartoe bevorderlijk kan zijn, alles in de ruimste zin des woords.

De stichting NIOC neemt de archivering van de resultaten van de congressen voor zijn rekening. De website www.nioc.nl ontsluit onder "Eerdere congressen" de gearchiveerde websites van eerdere congressen. De vele afzonderlijke congresbijdragen zijn opgenomen in een kennisbank die via dezelfde website onder "NIOC kennisbank" ontsloten wordt.

Op dit moment bevat de NIOC kennisbank alle bijdragen, incl. die van het laatste congres (NIOC2023, gehouden op donderdag 30 maart 2023 jl. en georganiseerd door NHL Stenden Hogeschool). Bij elkaar bijna 1500 bijdragen!

We roepen je op, na het lezen van het document dat door jou is gedownload, de auteur(s) feedback te geven. Dit kan door je te registreren als gebruiker van de NIOC kennisbank. Na registratie krijg je bericht hoe in te loggen op de NIOC kennisbank.

Het eerstvolgende NIOC vindt plaats op donderdag 27 maart 2025 in Zwolle en wordt dan georganiseerd door Hogeschool Windesheim. Kijk op www.nioc2025.nl voor meer informatie.

Wil je op de hoogte blijven van de ontwikkeling rond Stichting NIOC en de NIOC kennisbank, schrijf je dan in op de nieuwsbrief via

www.nioc.nl/nioc-kennisbank/aanmelden_nieuwsbrief

Reacties over de NIOC kennisbank en de inhoud daarvan kun je richten aan de beheerder:

R. Smedinga kennisbank@nioc.nl.

Vermeld bij reacties jouw naam en telefoonnummer voor nader contact.



Minor Data Science 2015-2016

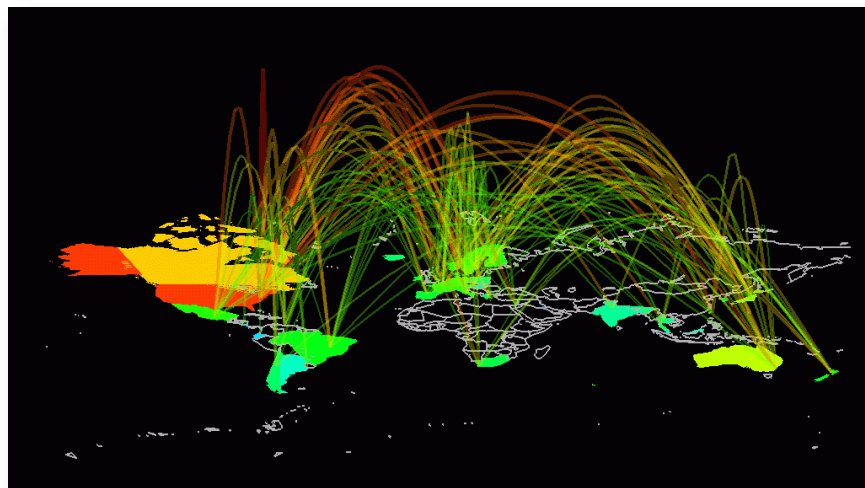
(NIOC 2015 - 23 april)

Tony Busker, Jan Kroon en Sunil Choenni (Instituut voor Communicatie,
Media- en Informatietechnologie, Hoogeschool Rotterdam)
Sunil Choenni en Mortaza S. Bargh (WODC, min V&J)

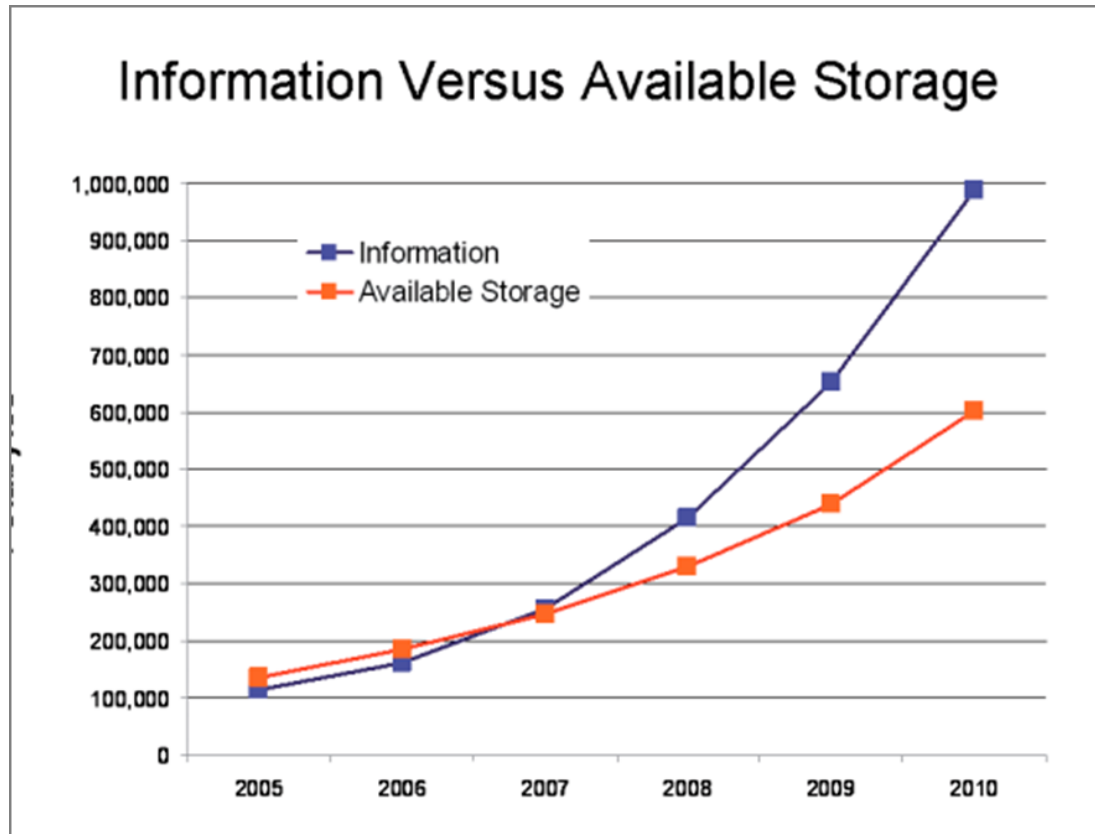
overtref jezelf



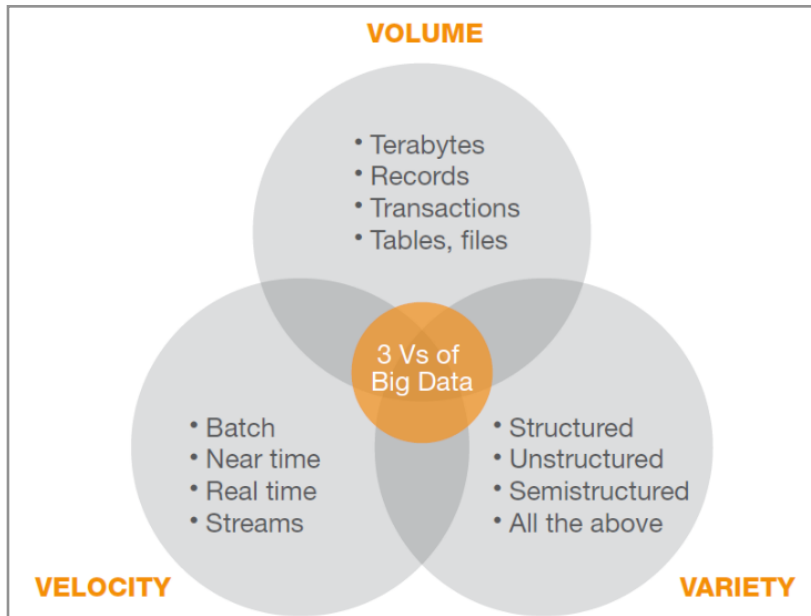
Volume



Big Challenges: Volume



Big data: What is it?



&



The (r)evolution of data

New techniques

BIG OPPERTUNITIES (ten dienste van de mens)

- Automatisch signaleren wat er in je koelkast moet
- Simuleren alsof je thuis bent terwijl je op vakantie bent
- Auto's die met elkaar communiceren
- Betere en actueler voorspellingen (griep epidemieën)
- Ouderen helpen navigeren
- Veiligheid
- Serious Games, benadert de werkelijkheid

Big Challenges

- Filesystemen (Google)
- Gegevens beheersystemen (BigTable)
- Geavanceerde redeneertechnieken, redeneren met onzekerheid, incompleetheid en onwetendheid
- Programmeerparadigma's (Functional MapReduce)
- Nieuwe programmeertalen (Sawzall?)

New techniques

Spark



Not Only SQL

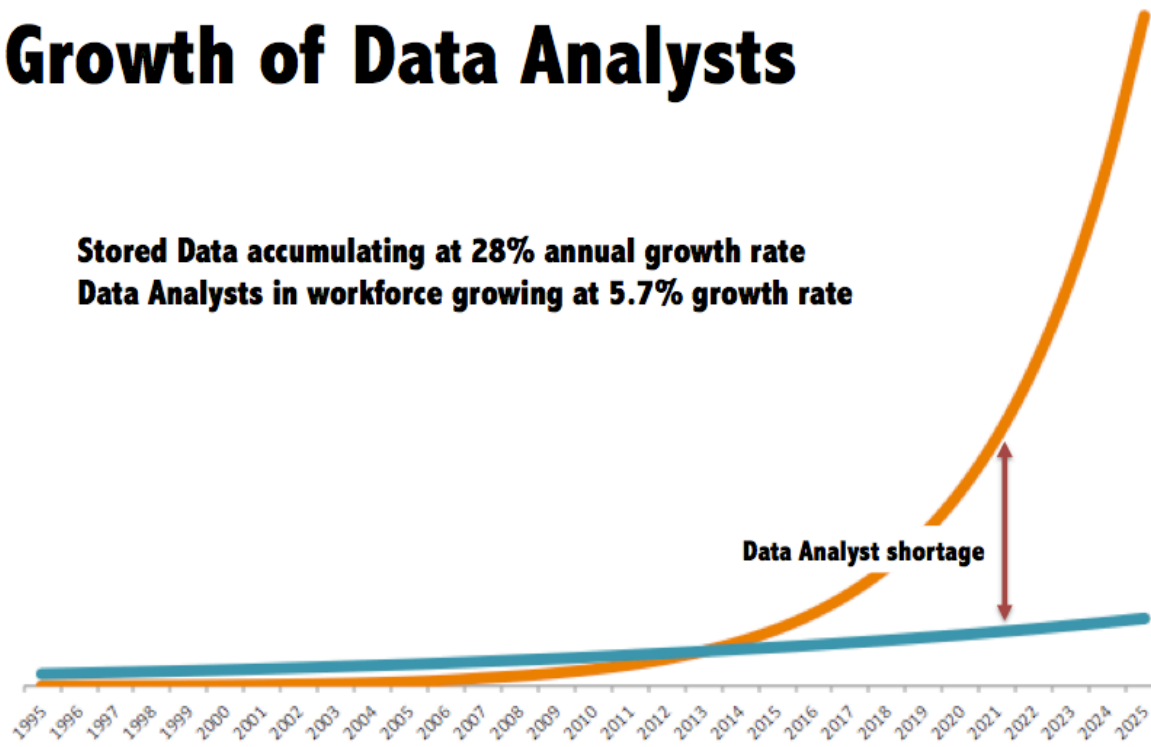


BIG DATA

- Combinatie (real ...) time, verschillende types en veel data
 - sensor en streaming data
 - audio, video, natuurlijke taal
 - terra- of pettabytes
- Integratie van (real ...) time, multi media and very large database technologie

Growth of Data vs. Growth of Data Analysts

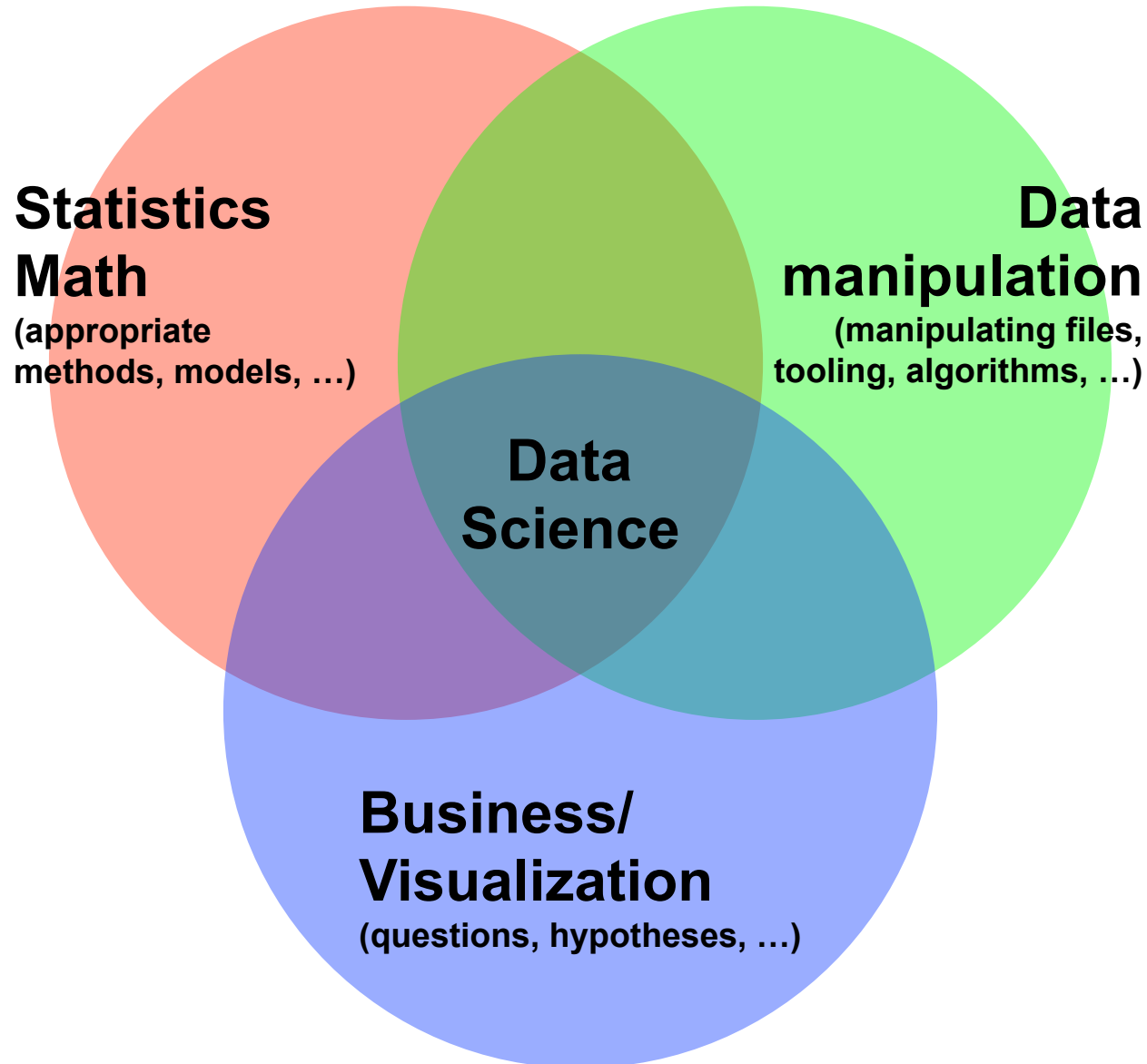
Stored Data accumulating at 28% annual growth rate
Data Analysts in workforce growing at 5.7% growth rate



Data Science



The Data Scientist



Minor Big Data

- Informatie als vierde productiefactor
- Nieuwe toepassingsmogelijkheden met besluitvorming/bewustwording
- Belangrijke actuele vraagstuk in de minor: Hoe ga je om met (grote) complexe bestanden afkomstig van verschillende organisaties?

Big Challenges: softside

- Welke data mag je met elkaar combineren?
- Hoe ga je verkregen combinaties duiden
- Hoe behoud je de relatie met de 'echte' werkelijkheid
- Privacy wordt een issue:
prijisdifferentiatie (AH, Booking.com, Bol)

Het doel is om de student een verzameling van methoden, technieken en tools aan te bieden zodat hij/zij is in staat om een toegevoegde waarde te creëren door het gebruik van grote hoeveelheden (open) data.



Aandachtsgebieden (1)

- met de opdrachtgever een goede vraagstelling kunnen formuleren
- inzicht hebben in de mogelijkheden om het (pre)processen van data te paralleliseren (map/reduce)
- voldoende kennis hebben van statistiek, data mining en machine learning incl. de beschikbare tools (R, RHadoop, machine learning libraries, Spark)

Aandachtsgebieden (2)

- Verzamelen van data met eigentijdse (smart) devices
- Kernconcepten m.b.t. verzamelen van data (privacy, transparantie, accountability, identiteit en identificatie, keuzevrijheid en efficiëntie en effectiviteit)
- Opslaan van data in data spaces, Big Data infrastructuren zoals Hadoop (incl. complexiteit en benodigde activiteiten zoals opschonen en uniformeren van data)

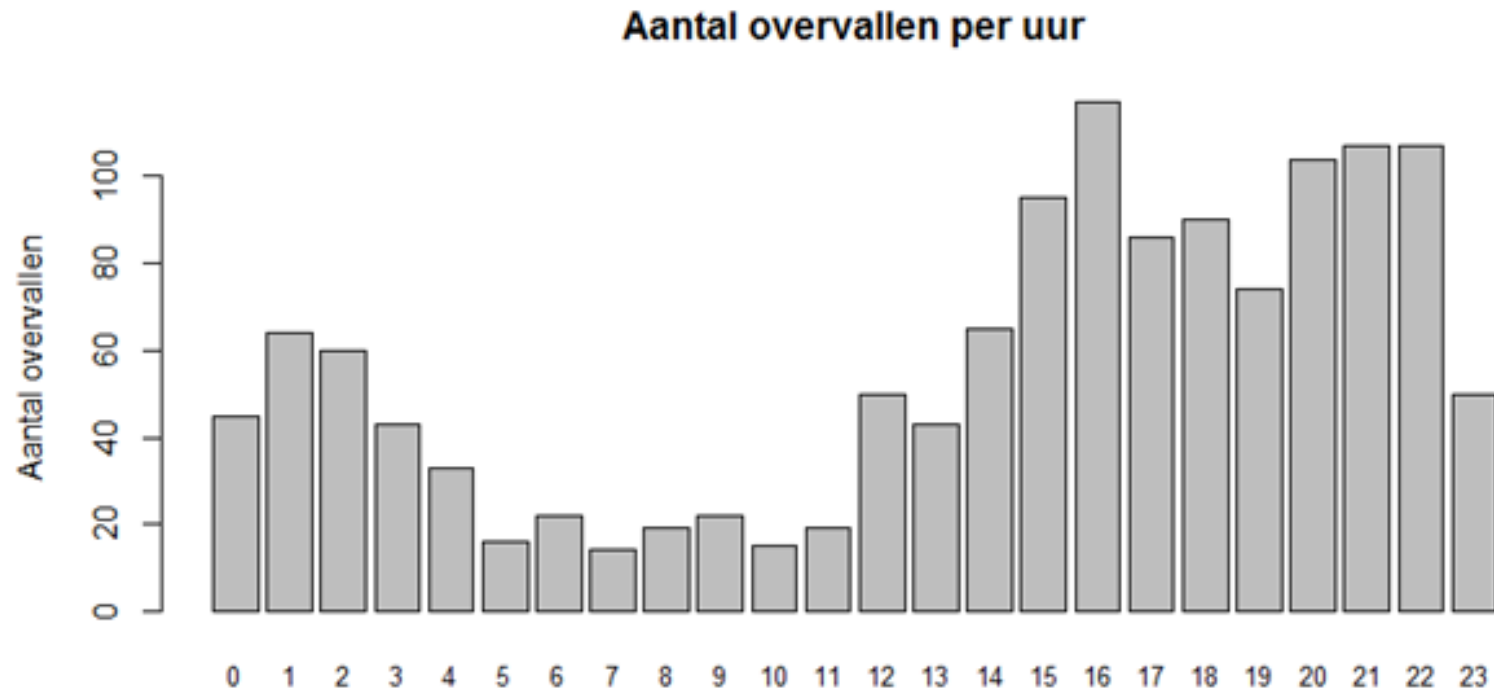
Aandachtsgebieden (3)

- Bevragen en analyseren van data (m.b.v. technieken zoals data mining, machine learning, ...)
- Aanbieden platforms aan gebruiker om hun eigen applicaties te genereren, zoals **mashups** (verzameling van data, analyse- en visualisatietechnieken)
- Gebruik van methoden en/of technieken voor visualisatie van informatie wat voor de doelgroep nieuwe inzichten oplevert bij complexe vraagstukken

Minor BD - programma (30 cp)

- Onderwijsperiode 1
 - Statistiek (met R) (4 cp)
 - Big Data, privacy, security analytics (3 cp)
 - Project (8 cp)
- Onderwijsperiode 2
 - Data Mining, Machine learning (4 cp)
 - Visualisatie/Business case/Reporting (4cp)
 - Project (7 cp)
- Toetsing (tentamens, opdrachten, assessments)

Berovingen op straat in Rotterdam



Voorbeelden Visualisatie

[TweetPing](#)

[Tweets na GermanWings tragedie](#) *

[200 years 200 countries](#) (video) *

[Twitter Sentiment analyse](#) (video)

[Browser Market Share](#) *

[wolfram data-science-platform/](#)

[TimeMap OV](#) *

overtref jezelf

