JADS

Jheronimus
Academy
of Data Science

Creating value with data

Overcoming the Hype

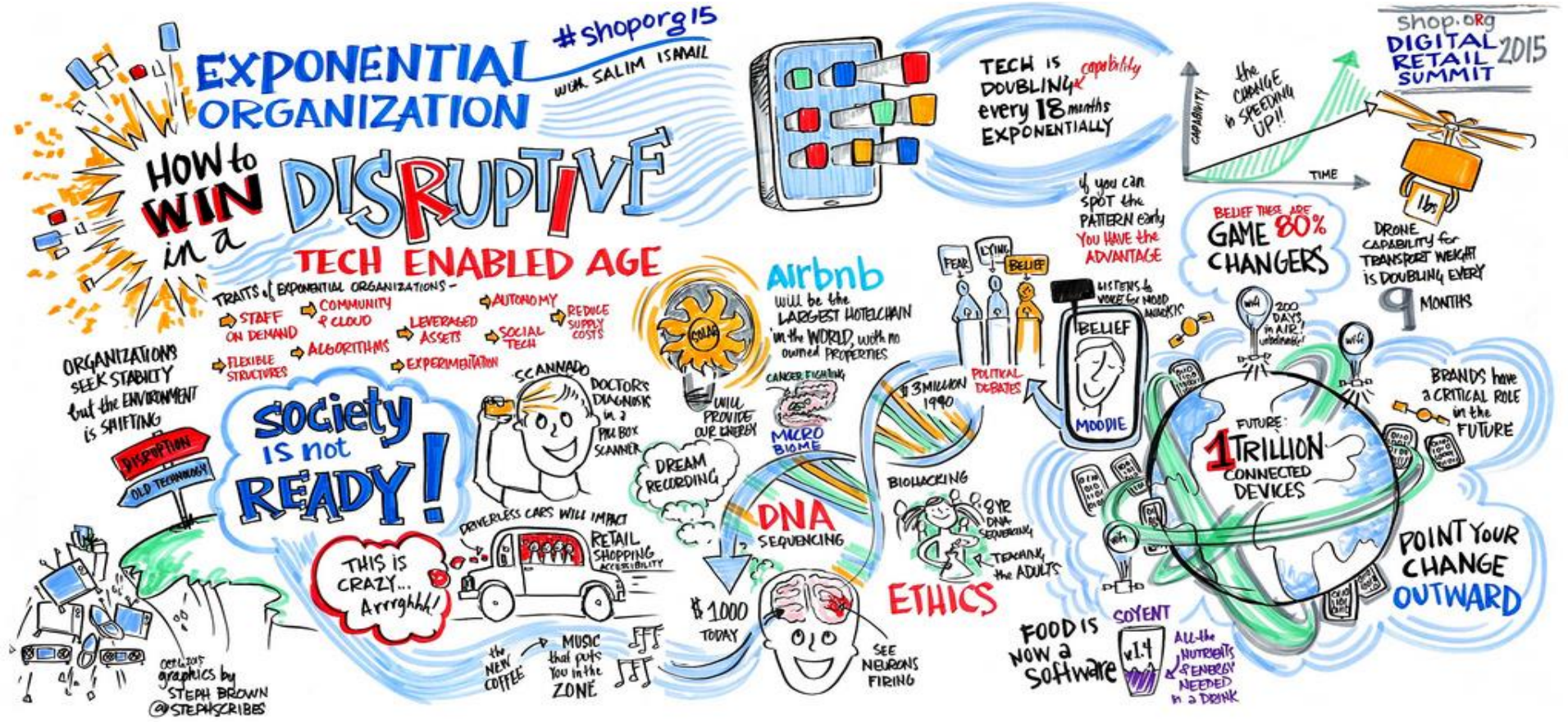
September 15th 2016

Arjan van den Born

CONTENT

- The disruptive force called data science
- Cool examples
- JADS – Bringing business and IT together
- The Graduate School Mariënburg – Data, Entrepreneurship & Innovation
- Unlocking the Value of Data Science – The Data Driven Organization

The Usual Story

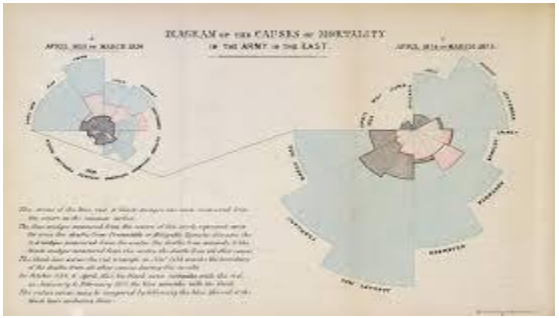


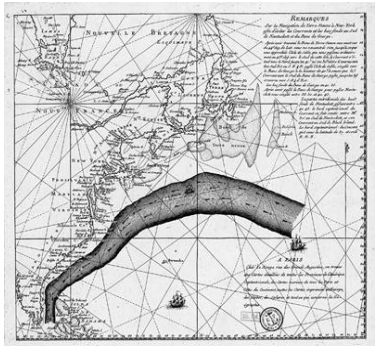
Life changes…so what

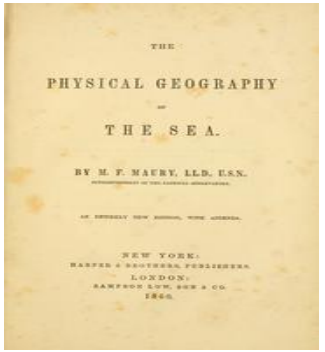
We trained hard, but it seemed that every time we were beginning to form up into teams, we would be reorganized. I was to learn later in life that we tend to meet any new situation by reorganizing; and a wonderful method it can be for creating the illusion of progress while producing confusion, inefficiency, and demoralization.

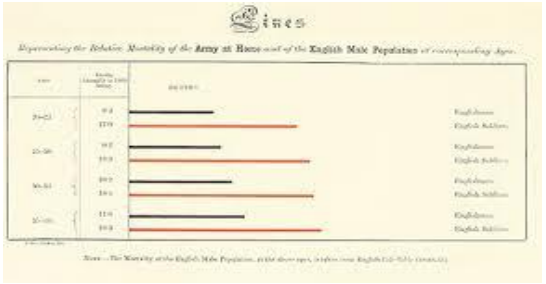
This quote is attributed to a disgruntled soldier in Ancient Rome, one Petronius Arbiter, in 210 BC.

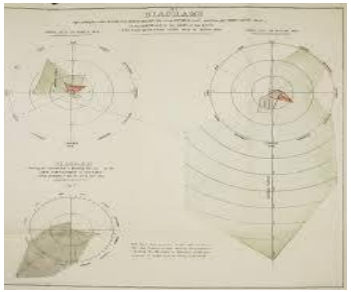
The Datafication era started in the early 1800's

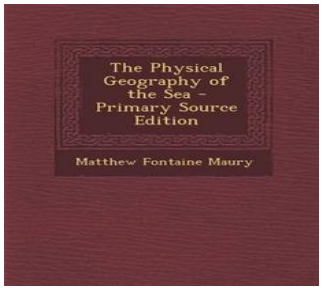












Florence Nightingale , London 1820-1910

Mathew Fontaine Maury, Virginia 1806-1873

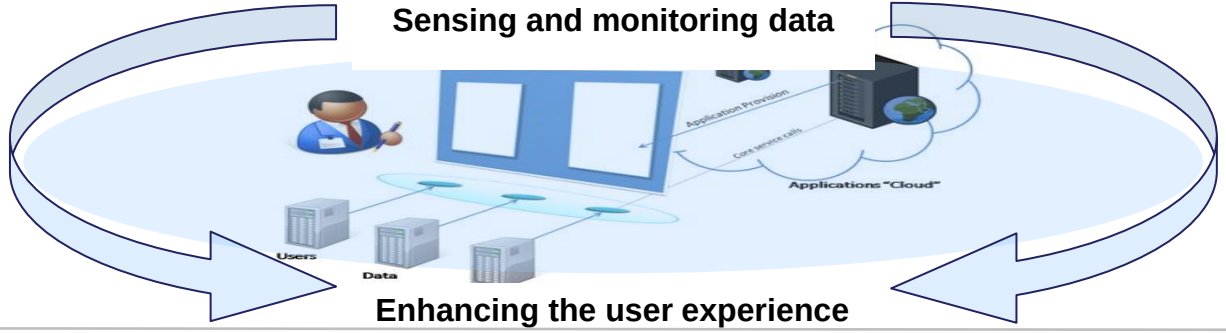
[6]

The Always-On Society generates true Big Data

Connected devices

Sensing and monitoring data

Providing connected Apps



Extracting meaningful information

Enhancing the user experience

Connected Solutions

Data science will create new professions

Large scale

Jeremy
Data Regulatory Officer

Melissa
On-line Data Business Strategist

Public

Private

Debora
Consultant Data Concessions

Ron
Web-data Entrepreneur

Small and medium scale

Every profession becomes a Data Science Profession

Health Analytics

On the Case at Mount Sinai, It's Dr. Data

By STEVE LOHR MARCH 7, 2015

Jeffrey Hammerbacher uses his finance and tech experience to understand diseases.
Sam Hodgson for The New York Times

Jeffrey Hammerbacher is a number cruncher — a Harvard math major who went from a job as a Wall Street quant to a key role at Facebook to a founder of a successful data start-up.

But five years ago, he was given a diagnosis of bipolar disorder, a crisis that fueled in him a fierce curiosity in medicine — about how the body and brain work and why they sometimes fail. The more he read and talked to experts, the more he became convinced that medicine needed people like him: skilled practitioners of data science who could guide scientific discovery and decision-making.

Now Mr. Hammerbacher, 32, is on the faculty of the Icahn School of Medicine at Mount Sinai, despite the fact that he has no academic training in medicine or biology. He is there because the school has begun an ambitious, well-funded initiative to apply data science to medicine.

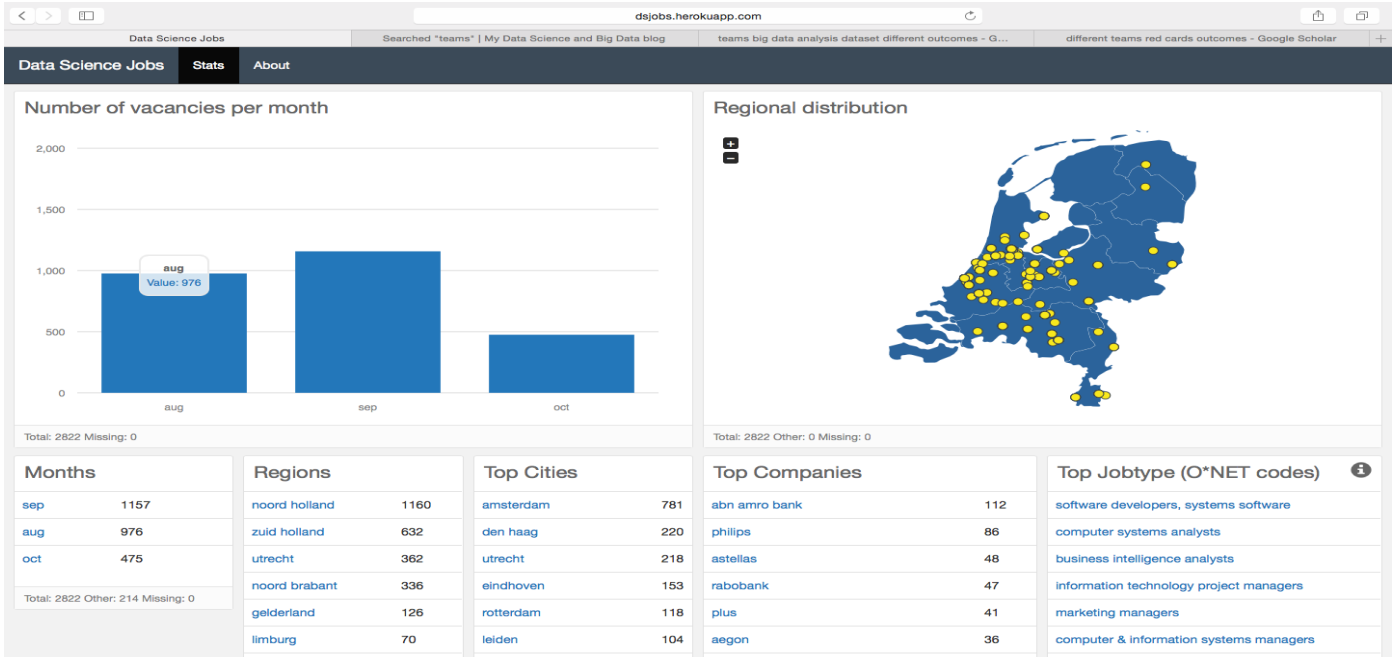
Email

Marketing Analytics

HR Analytics

Accounting & Finance Analytics

Explosive demand of “data scientists” in the Netherlands



Data Science is the next big thing

Data science seeks to use all relevant, often complex and hybrid data to effectively tell a story that can be easily understood by non-experts

Data Science is also the next big risk

Big data can and should bring greater safety, economic opportunity, and convenience to all people.

There is discrimination in delivery of these ads.

"Big data" allows for a new level of specificity in underwriting, changing how risk is allocated.

A hiring preference against workers who live far away may be accurate—they may really average shorter tenure in the job—but is it fair?

There are few laws governing the data brokerage industry.

A lack of high-quality, individualized financial data can exclude a person from the mainstream financial system.

Despite their importance, the algorithms used by E-Verify are not disclosed to the public.

Body-worn cameras are poised to help boost accountability for law enforcement and citizens.

http://bigdata.fairness.io/wp-content/uploads/2014/11/Civil_Rights_Big_Data_and_Our_Algorithmic-Future_v1.1.pdf

Data Science has two words

Data

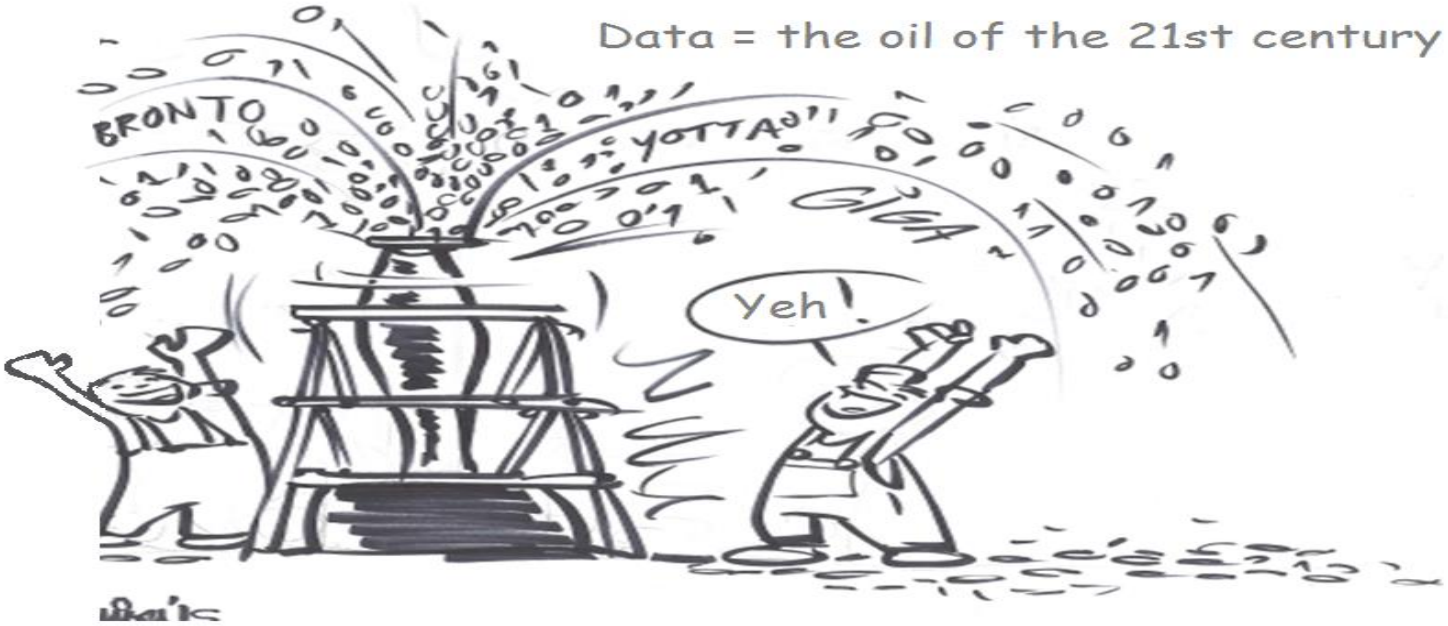
Science

….we need both to subtract value from data


Does the data always reveal the truth?


I'M SCARED TO DEATH
BIG DATA DOESN'T MEAN
BIG WISDOM.
ux
design
research
hci
usability
IA
ixd

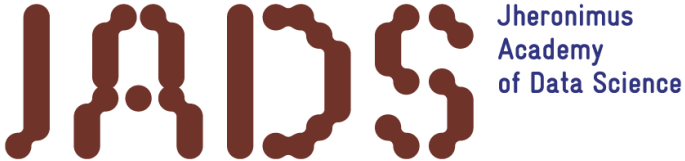
The vision is clear: data is the new oil water



[16]


Because there
is no Plan B

Cool Examples



Big data and gene-environment interactions


Setting up Big Registries to find new answers to complex questions
wiseGEEK

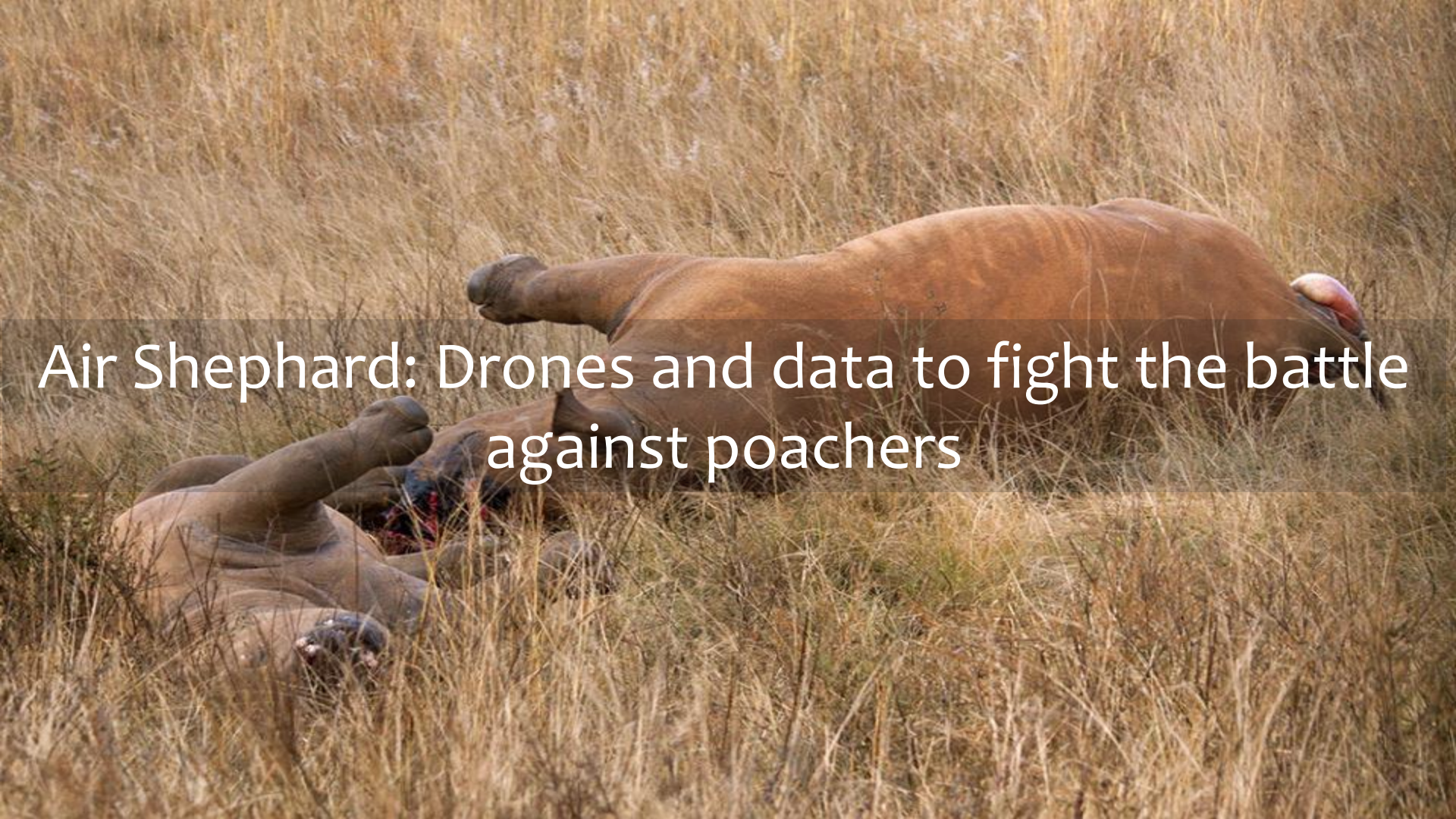
GENOME ANALYTICS MADE SIMPLER

Your I-phone has become a biomedical research device

HEALTH

The biological passport as the way to find cheaters


SMART MAINTENANCE USING PREDICTIVE ANALYTICS
(e.g. GE, BMW & DAF)

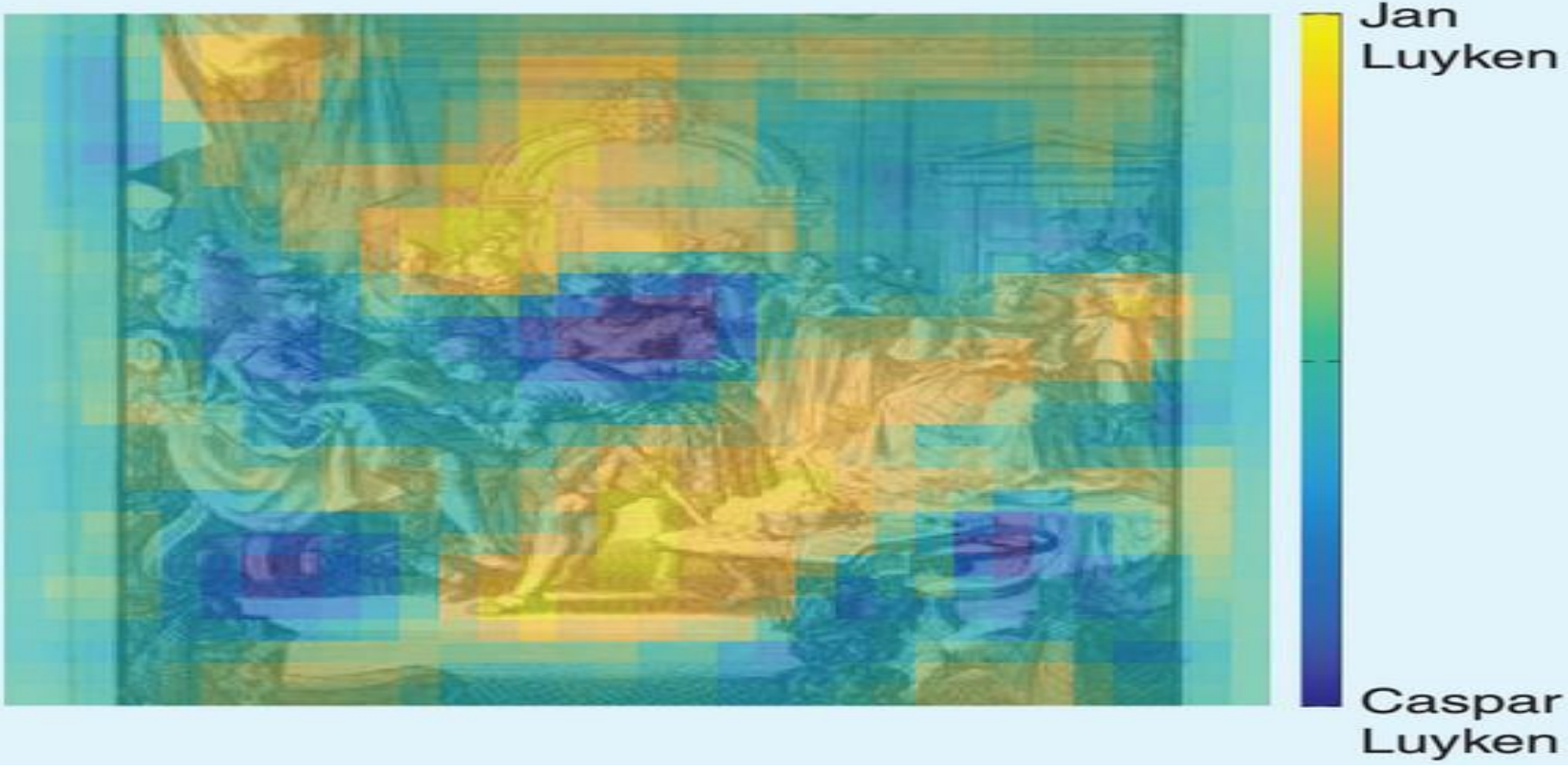


Air Shephard: Drones and data to fight the battle against poachers

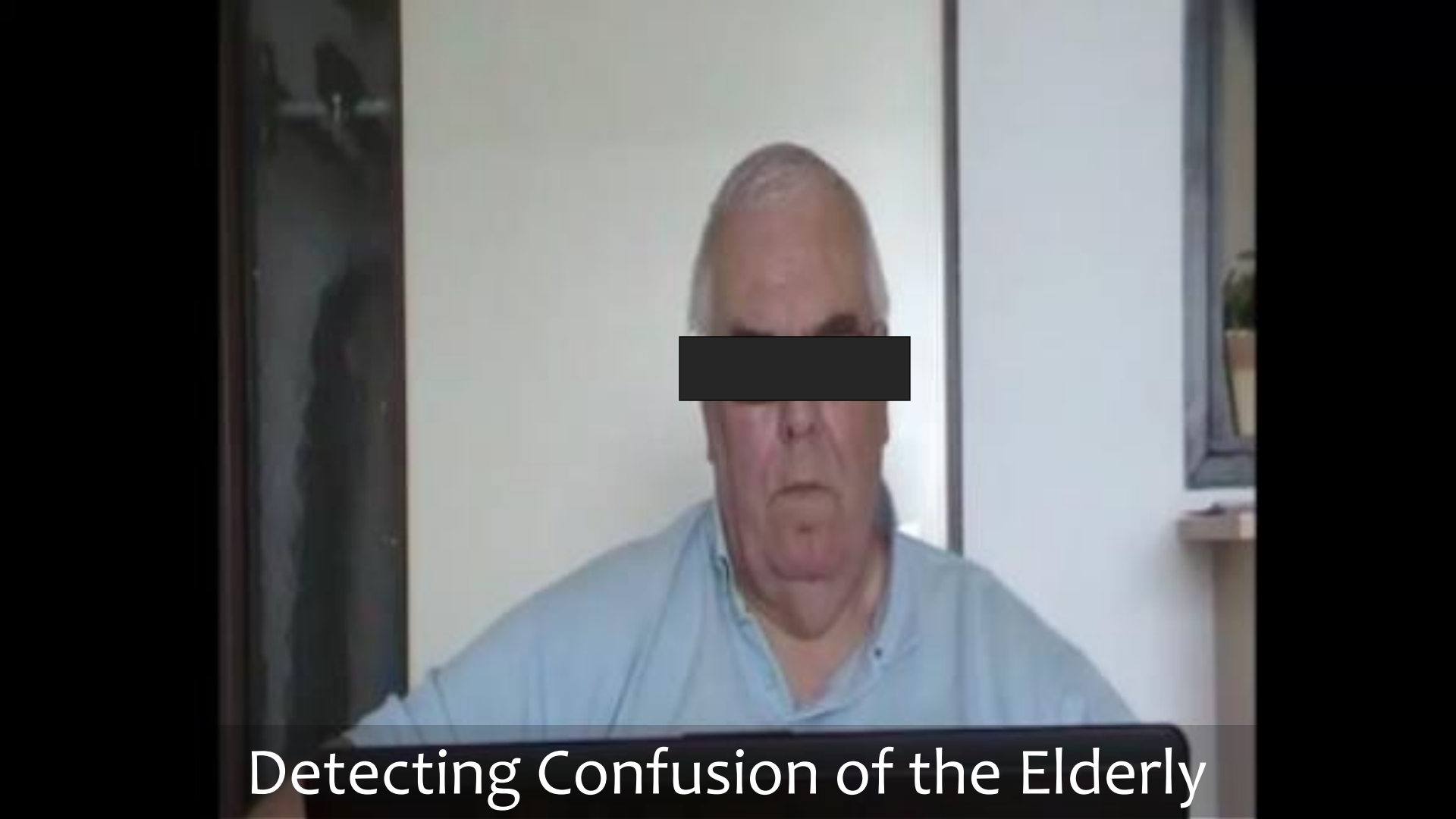

PREDICTING DEFORESTATION


FROM SEED PRODUCER TO BIG DATA COMPANY

[FIG5] An image of the *Overdracht der Nederlande, aan de Infante Isabella* (1697–1699) by Jan Luyken.



[FIG6] **Visualization of how characteristic each image region is for the artists Jan and Caspar Luyken. The yellow regions are characteristic of Jan Luyken, whereas the blue regions are characteristic of Caspar Luyken.**


Detecting Confusion of the Elderly

Does Spotify Increase Variety?

Design Targeted Marketing Actions


More Products than Persons?

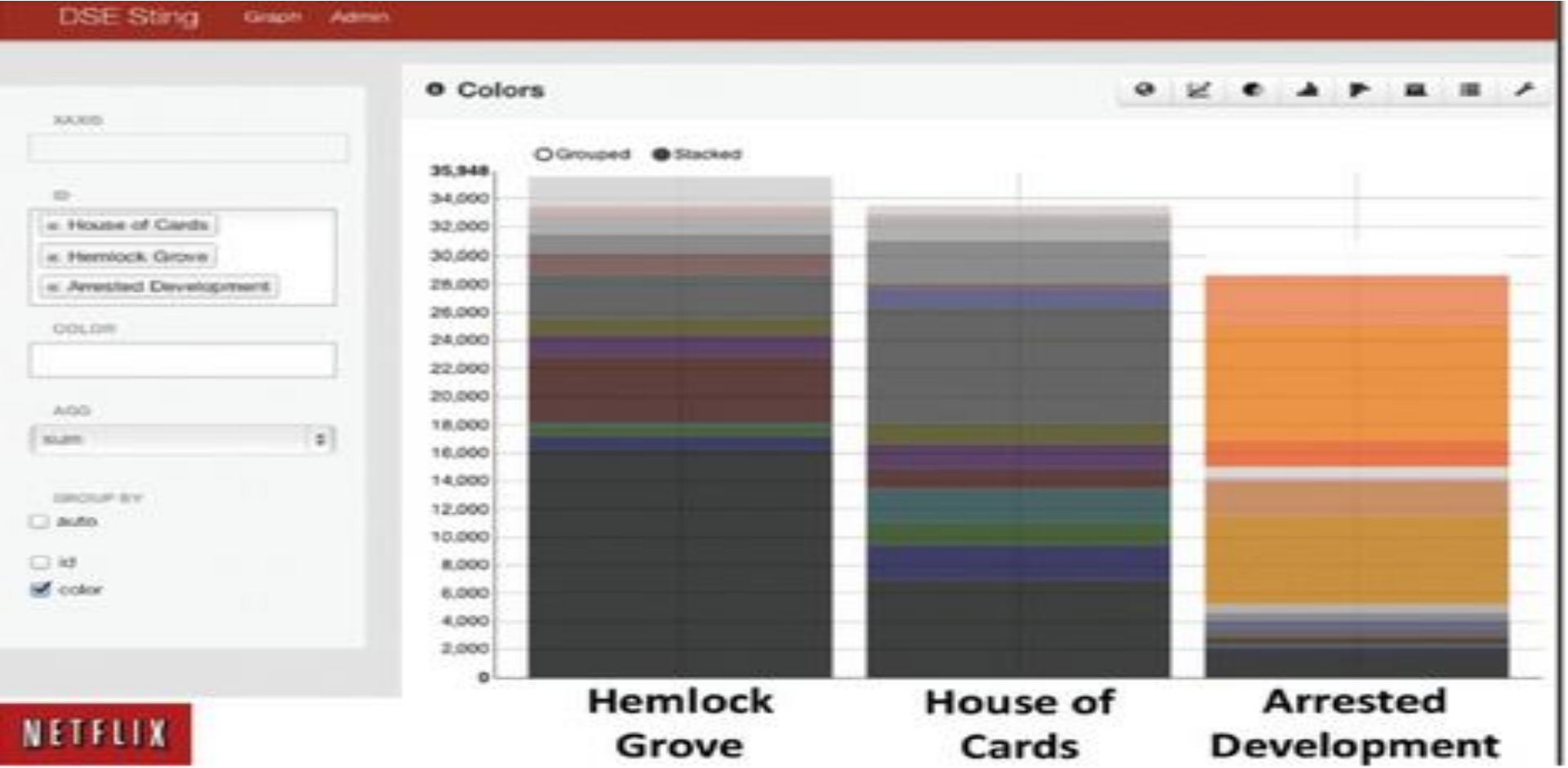


Figure 3.2 Detailed Color Comparison of *Hemlock Grove*, *House of Cards*, and *Arrested Development*
Source: Netflix Technology Blog (techblog.netflix.com)

Pricing for advertising space on billboards, benches and busses based on actual # views

route
outdoor media analytics

HR Analytics: Overhaulin’ HR


How does big data affect competitive processes?

Computational Legal Studies™

Tag Archives: *Judicial Decision Making*

Obamacare's Chances Of Survival Have Improved (But It Is Still Very Close)

Posted on March 6, 2015 by Daniel Martin Katz

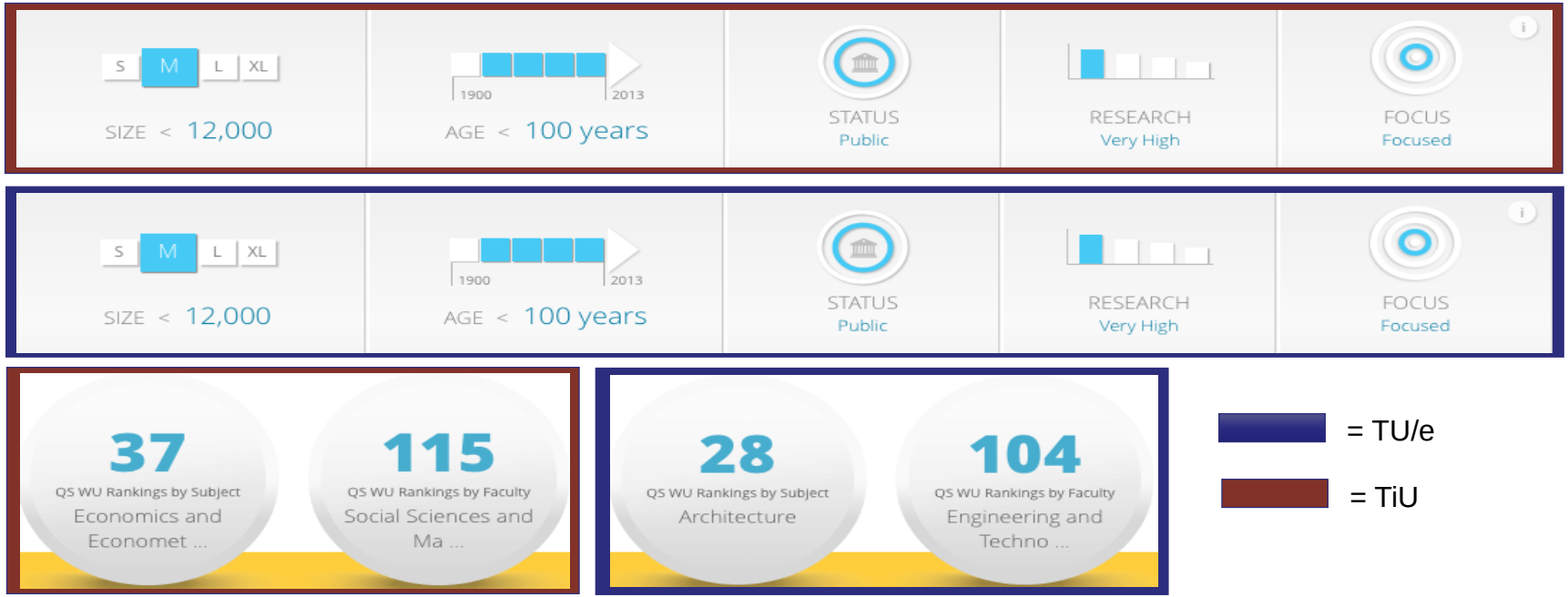
Posted in Uncategorized | Tagged artificial intelligence and law, computational legal studies, fantasy scotus, Judicial Decision Making, obamacare, quantitative legal prediction, supreme court, wisdom of crowds | Leave a comment

Earthquake Haiti (Jan 2010)

Who we are

JADS
Jheronimus
Academy
of Data Science

Two focused, comparable, yet distinct top universities

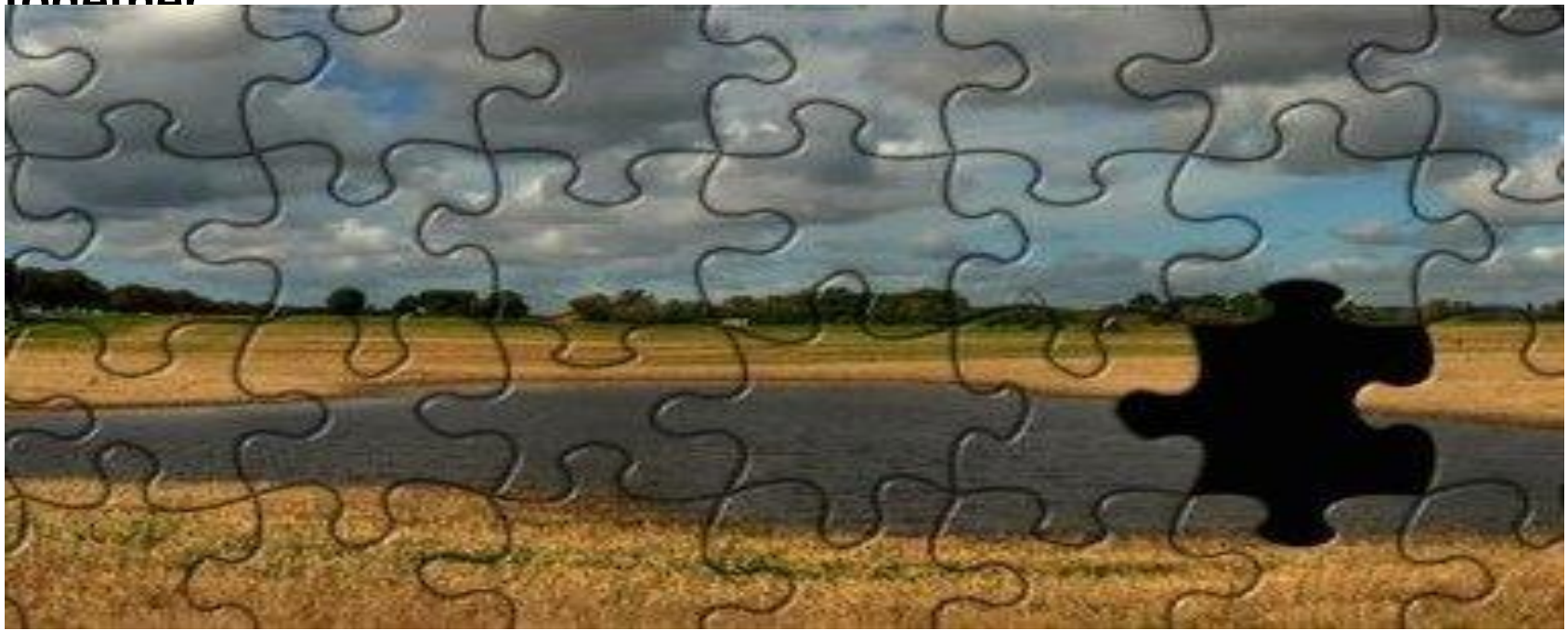


Are Business & IT finally meeting up?

JADS: What, Why & How

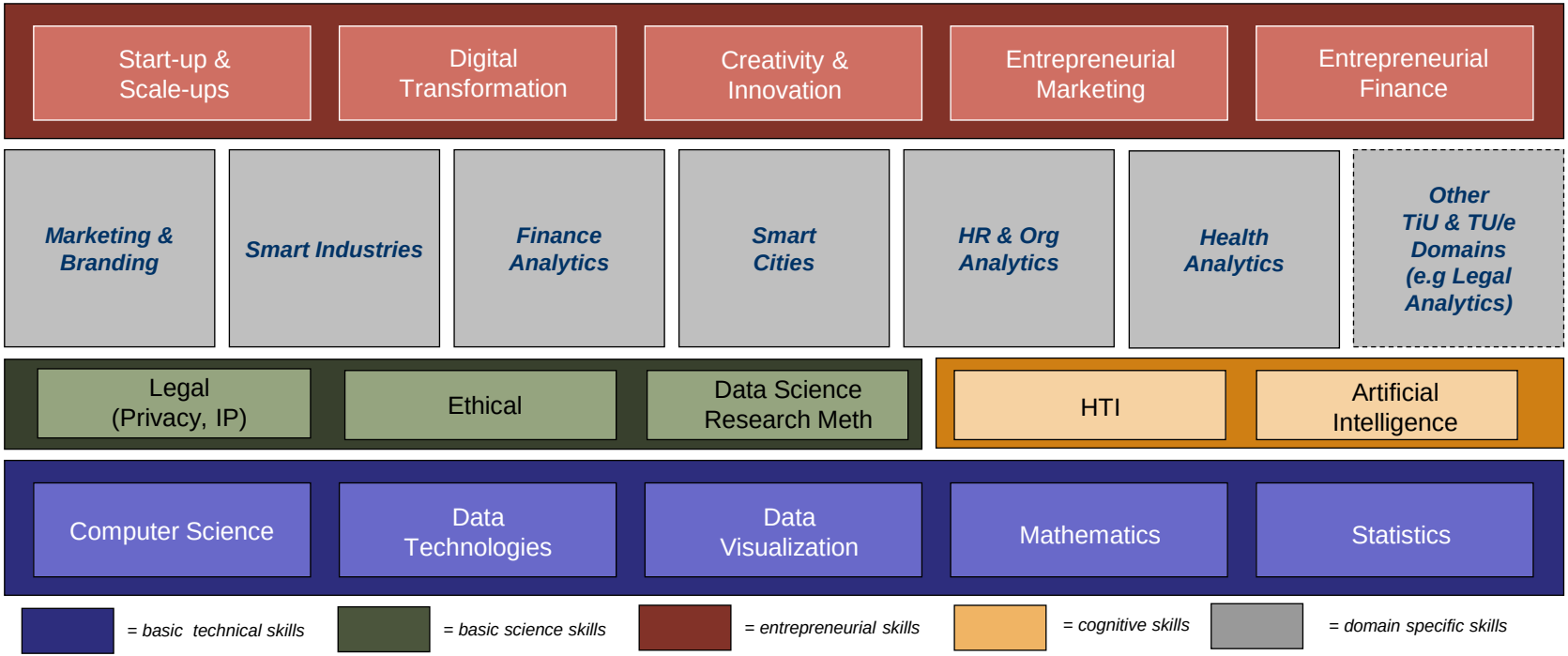
Jheronimus
Academy
of Data Science

WHY JADS? Bringing Technology and Sociology/Business together



Key Ingredients of Data Science

Our Capability Framework



JADS: close collaboration of 2 universities on 3 areas within data science

education

JADS provides a number of data science programs at undergraduate, graduate and post-graduate level to educate the data scientist of the future

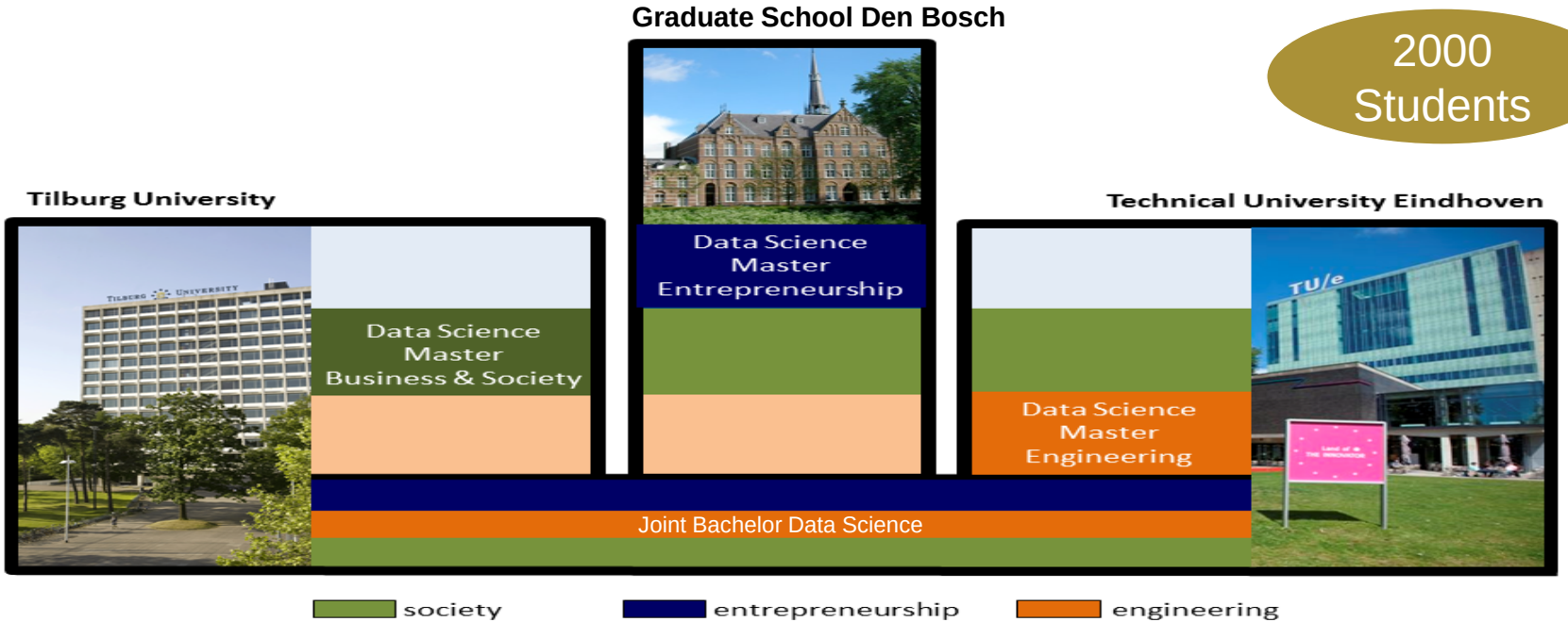
research

Offering Data science research in three locations. Research programs that enhance the relevance of academic research to business and society.

business

JADS offers great opportunities for organizations to become part of the data science ecosystem in 's-Hertogenbosch.

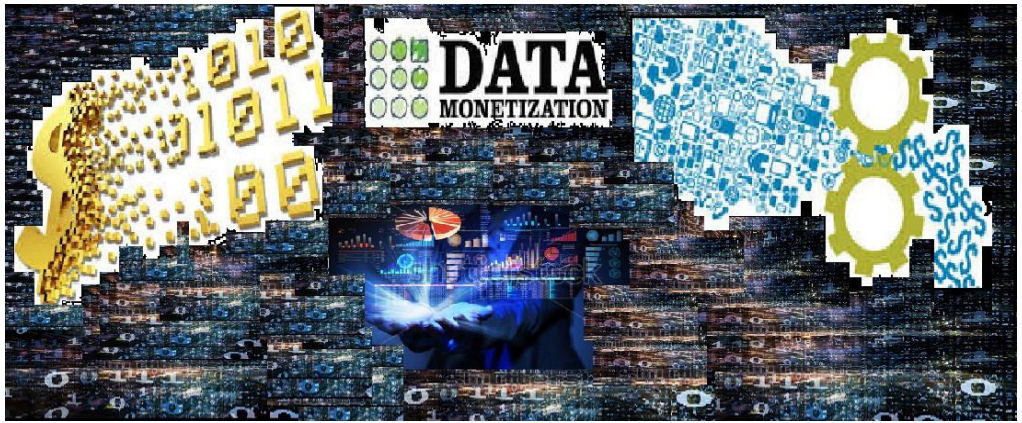
Education: nine data science programs offered at three locations



[48]

Research: we will have one generic research theme …

Data Innovation & Data Entrepreneurship



Transforming data into value through analysis and visualization

[49]

… and six dedicated research themes

Personal care and wellbeing

Monitoring non-intrusively biomarkers and body functions and feeding back information.

Productivity and maintenance

Monitoring performance remotely and enabling long-distance control and maintenance.

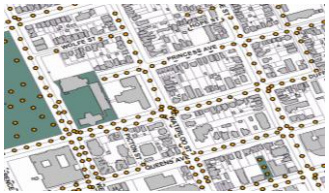
Branding and marketing

Analyzing customer behavior and mapping needs on business propositions.

Finance and insurance

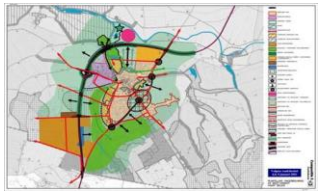
Developing data driven options for innovative financial and insurance related business improvements.

City management and control



Enabling adaptive public space support by urban service management and commissioning.

Agro, food, and life science



Providing decision support by correlating information from different life-science related sources.

Five unique elements of our proposition

Size: 2000 students in various programs with 3 research centers

Novelty: educational programs are newly designed using DS design criteria

Multi-disciplinary: Mixing business and IT to drive innovation

Ambition & Quality: world class staff & students to make an impact

Relevance: Intense collaboration with industry in education and research

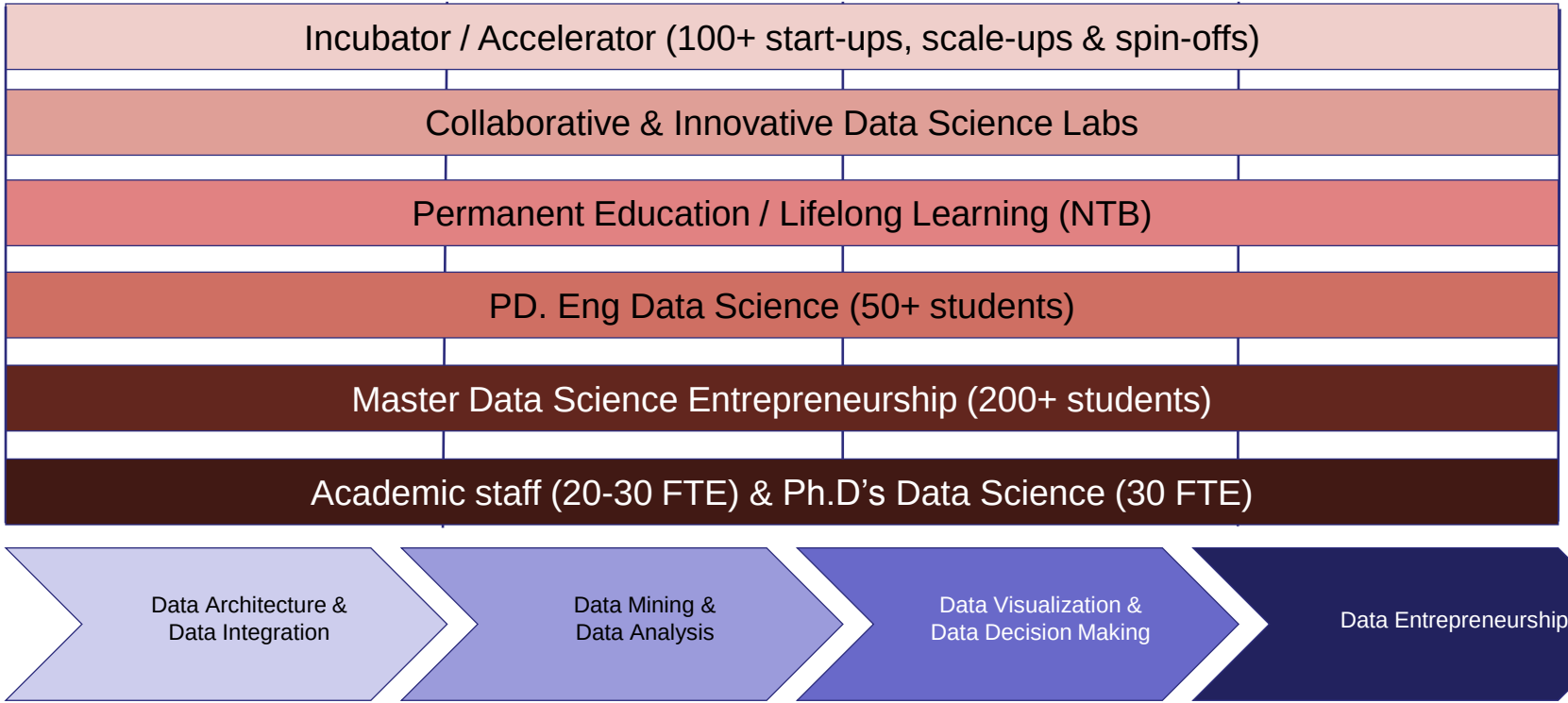
Mariënburg: the hotspot of our initiative

Jheronimus Academy of Data Science

Combining Data Science & Entrepreneurship

Jheronimus Academy of Data Science

Functions in our ‘hotspot’: ‘s-Hertogenbosch

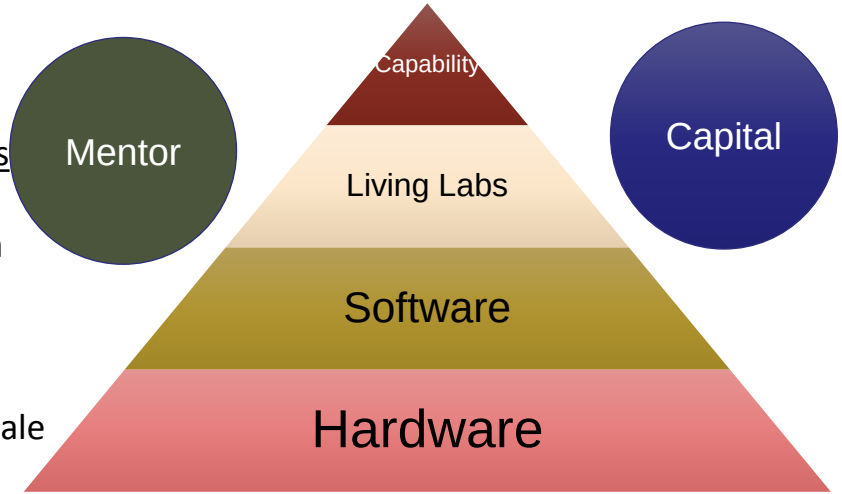


The place where we will collaborate intensively with our partners

Our offering to start-ups, scale-ups and spin-offs (and students)

We offer a unique proposition to start-ups (incubator), scale-ups (accelerator) and spin-offs to take shop in and around Mariënburg. This offer has several layers.

1. **Hardware** – state of the art facilities.
2. **Software** a scientific cloud (e.g. R, Python, Matlab, Mathematica, Hadoop) and commercial applications (e.g. SAS, Tableau) on premises.
3. **Living Labs** with proprietary datasets and open data
4. **Data science capability** by having established researchers open to discussion and dedicated scientific programmers who do data science.
5. **Financial capital** available for students, start-ups, scale
6. **Mentoring network** of 100+ data science entrepreneurs, data scientists, advisors (IP)



Unlocking the Value of Data

Entrepreneurship = the art & science of spotting and grabbing opportunities

To really change the world one needs to understand both business & IT

elastic

Booking.com

UBER

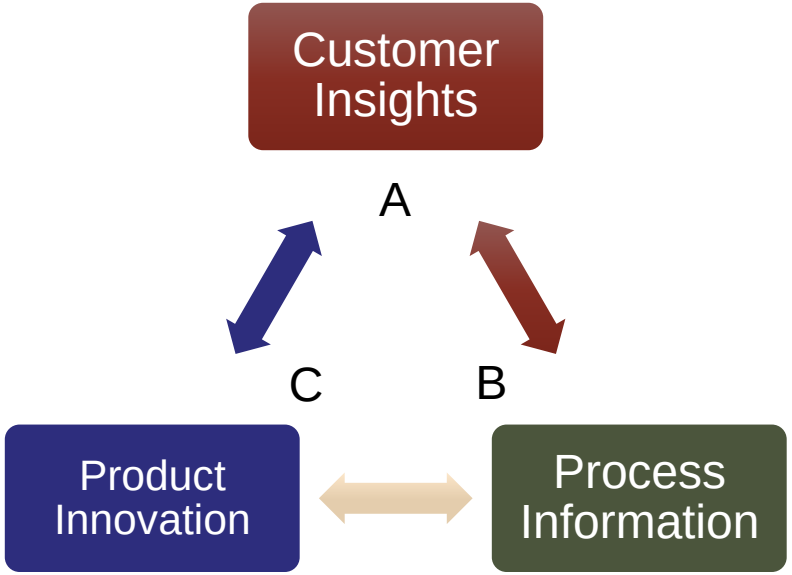
airbnb

Unlocking Value out of Data

- The triple pathway to unlocking Data Value
- Volume, Variety, Velocity, Veracity and Value
- The Process of Data Science
- The Organization of Data Science
- The Business Models of Data Science
- The difficulties of building a Data Science Capability

The Triple Pathway to Data Value

- There are various ways to extract value from Data:
  - Improved customer relations
  - Effective and efficient processes
  - New products and services
- The current focus is mainly on improving customer understanding



The FOUR V's of Big Data

From traffic patterns and music downloads to web history and medical records, data is recorded, stored, and analyzed to enable the technology and services that the world relies on every day. But what exactly is big data, and how can these massive amounts of data be used?

As a leader in the sector, IBM data scientists break big data into four dimensions: **Volume, Velocity, Variety and Veracity**

Depending on the industry and organization, big data encompasses information from multiple internal and external sources such as transactions, social media, enterprise content, sensors and mobile devices. Companies can leverage data to adapt their products and services to better meet customer needs, optimize operations and infrastructure, and find new sources of revenue.

By 2015

4.4 MILLION IT JOBS

will be created globally to support big data, with 1.9 million in the United States

Volume

SCALE OF DATA

40 ZETTABYTES

[43 TRILLION GIGABYTES]

of data will be created by 2020, an increase of 300 times from 2005

It's estimated that

2.5 QUINTILLION BYTES

[2.3 TRILLION GIGABYTES]

of data are created each day

6 BILLION PEOPLE

have cell phones

WORLD POPULATION: 7 BILLION

Most companies in the U.S. have at least

100 TERABYTES

[100,000 GIGABYTES]

of data stored

Variety

DIFFERENT FORMS OF DATA

As of 2011, the global size of data in healthcare was estimated to be

150 EXABYTES

[161 BILLION GIGABYTES]

30 BILLION PIECES OF CONTENT

are shared on Facebook every month

By 2014, it's anticipated there will be

420 MILLION WEARABLE, WIRELESS HEALTH MONITORS

4 BILLION+ HOURS OF VIDEO

are watched on YouTube each month

400 MILLION TWEETS

are sent per day by about 200 million monthly active users

Velocity

ANALYSIS OF STREAMING DATA

The New York Stock Exchange captures

1 TB OF TRADE INFORMATION

during each trading session

Modern cars have close to

100 SENSORS

that monitor items such as fuel level and tire pressure

By 2016, it is projected there will be

18.9 BILLION NETWORK CONNECTIONS

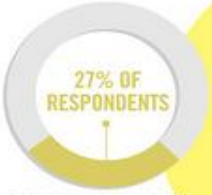
– almost 2.5 connections per person on earth

Veracity

UNCERTAINTY OF DATA

1 IN 3 BUSINESS LEADERS

don't trust the information they use to make decisions



in one survey were unsure of how much of their data was inaccurate

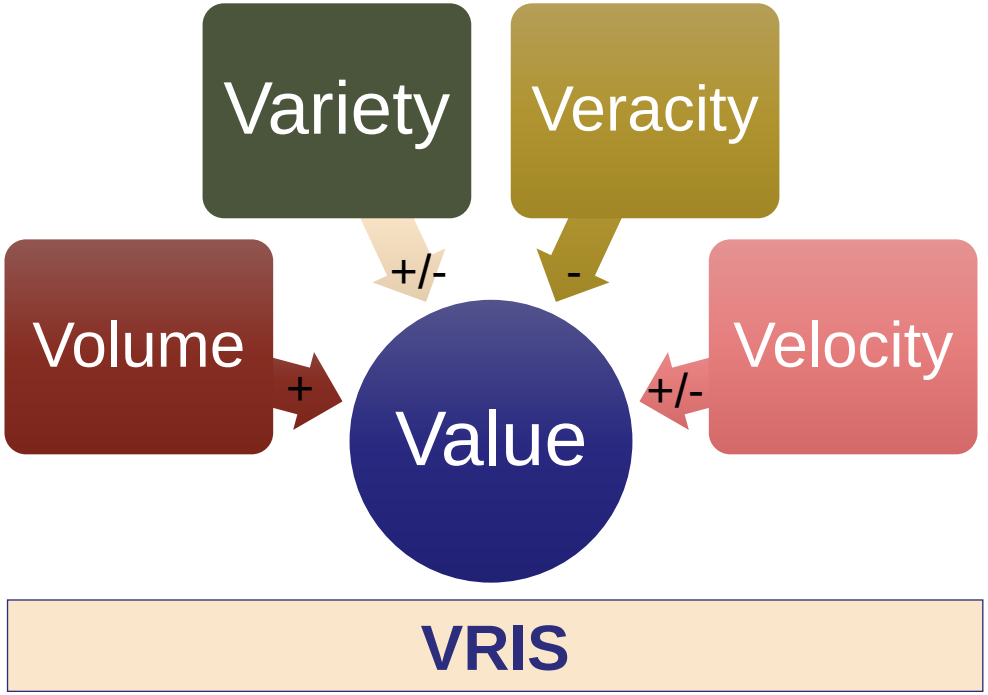
Poor data quality costs the US economy around

$3.1 TRILLION A YEAR

Sources: McKinsey Global Institute, Twitter, Cisco, Gartner, EMC, SAS, IBM, MEPTEC, QAS

IBM

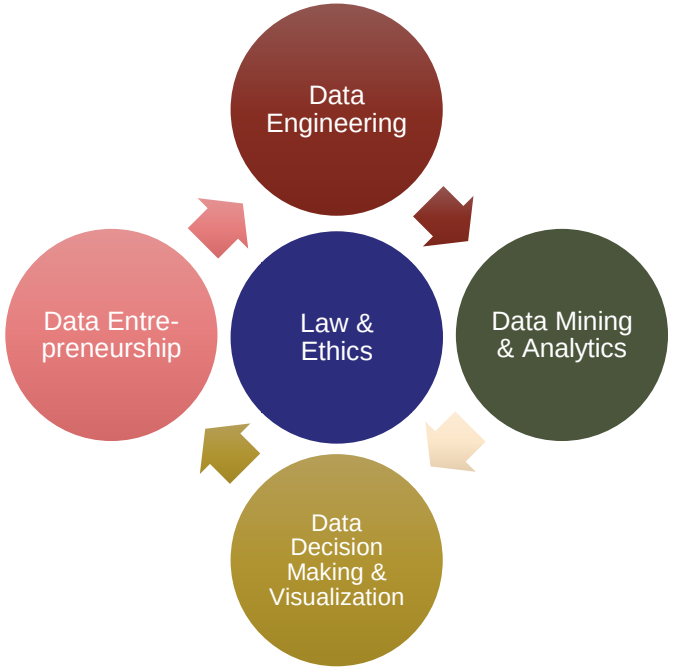
The Potential Value of Big Data is determined by it's uniqueness



[63]

The High-Level Process of Data Science

- Data value can be created in every phase of the data value process
- Data value can be destroyed in every phase of the data value process
- Without potential value, there is no way value can be created



[64]

The Mature Data Science Organization

A mature data science organization…

- …democratizes all data and data access.
- …uses Agile for everything and leverages DataOps (i.e., DevOps for Data Product Development).
- …leverages the crowd and works collaboratively with businesses (i.e., data champions, hackathons, etc.).
- …follows rigorous scientific methodology (i.e., measured, experimental, disciplined, iterative, refining hypotheses as needed).
- …attracts and retains diverse participants, and grants them freedom to explore.
- …relentlessly asks the right questions, and constantly searches for the next one.
- …celebrates a fast-fail collaborative culture.
- …shows insights through illustrations and tells stories.
- …builds proof of value, not proof of concepts.
- …personifies data science as a way of doing things, not a thing to do.

[65]

Data Science is not Business Intelligence

Data science als aanvullende discipline op bestaande analytics.

Bron (aangepast): http://blog.revolutionanalytics.com/2013/05/statistics-vs-data-science-vs-bi.html

Project Aristotle: Google's question on how to build teams

Google Myths

" Building the best team means combining the best people"

'It's better to put introverts together"

"Teams are more effective when everyone is friends away from work"

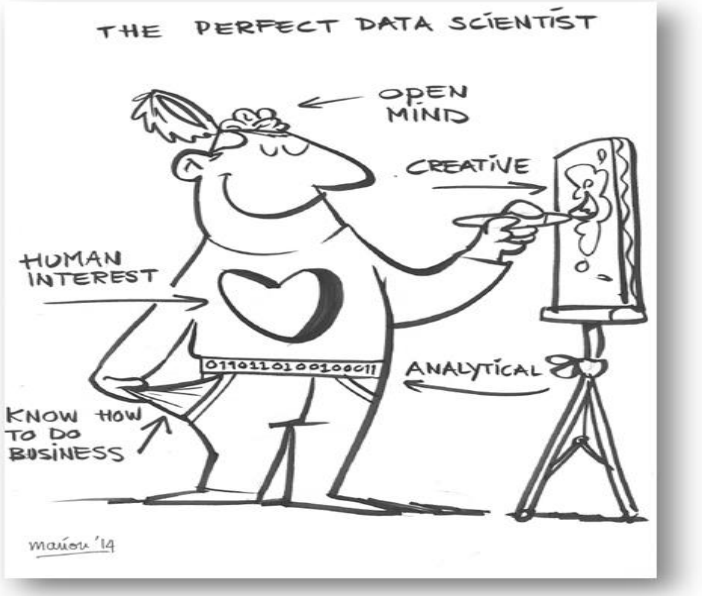
[67]

It’s [redacted] very much about human capital

We need T-shaped people

Data scientist have knowledge of

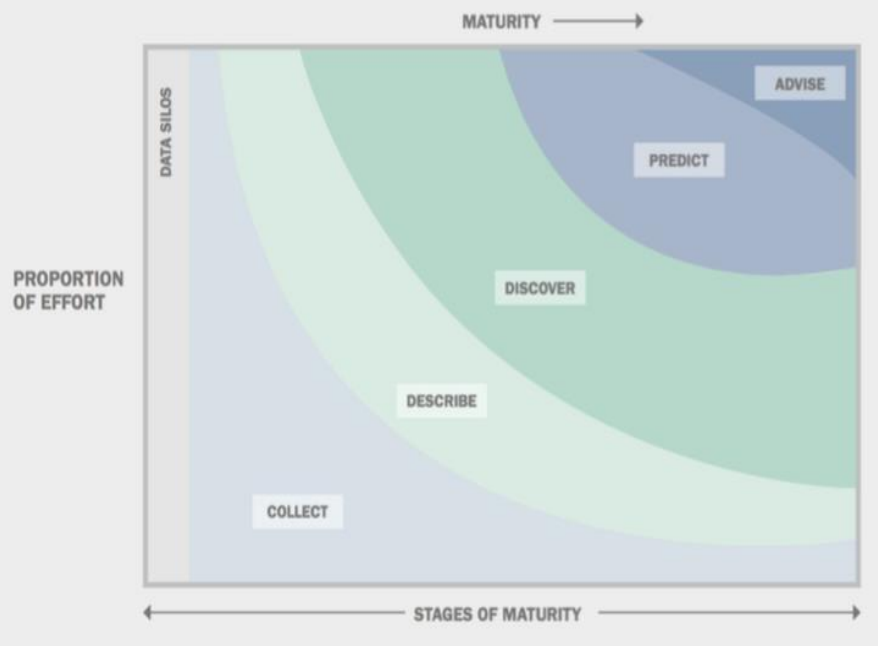
- Engineering
- Business and society
- Entrepreneurship



Data Science Maturity is generally low

A Data Driven Organization Has:

- Data Technology
- Data Processes
- Data Governance
- Data Skills & Data Culture

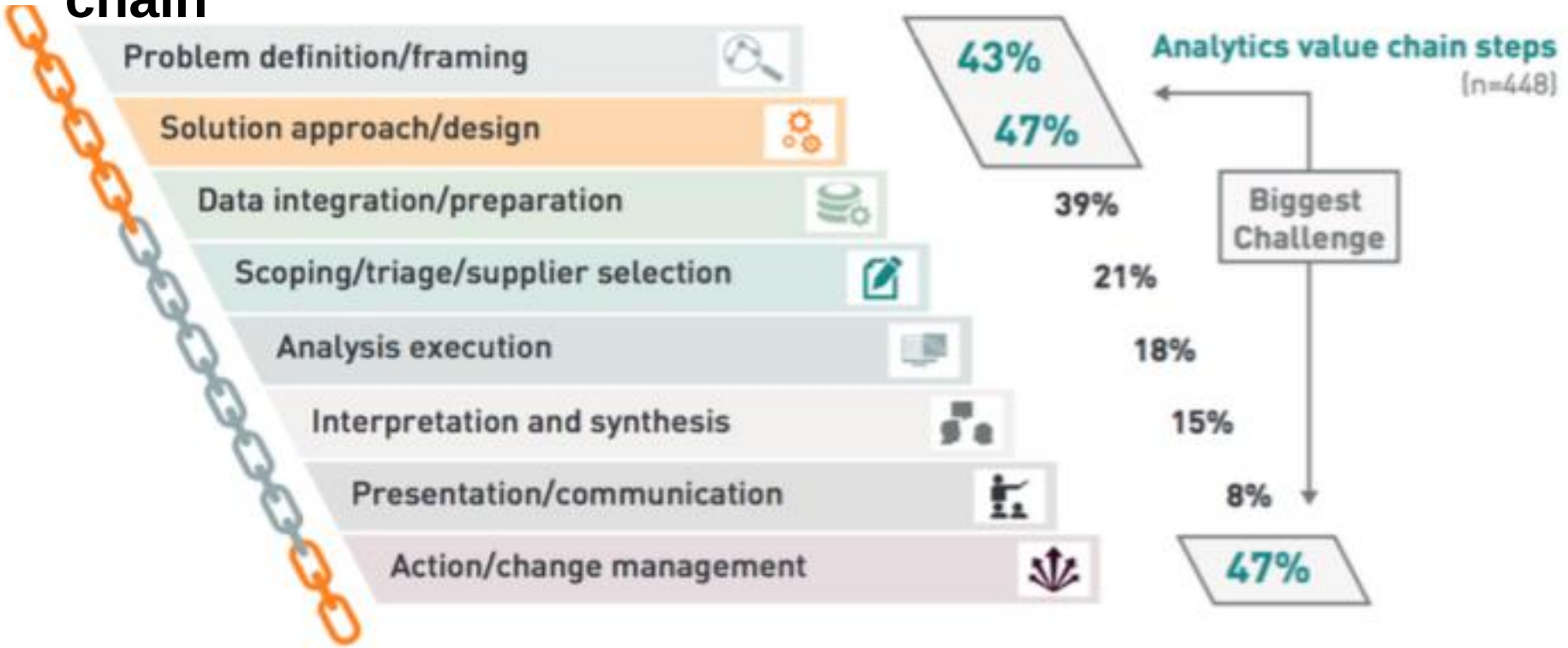


[69]

Key Aspects of Your Data Science Business Model

- Who is the data customer?
- What is the product or service?
- What is your pricing model?
- What is the price?
- Where is your added value?

Problems at the end and beginning of the value chain



INDUSTRY CHANGE AHEAD

A contemporary fairytale

"ONCE UPON A TIME IN DEN BOSCH"

... A DREAM BECAME TRUE

APPENDIX