



DATA SCIENCE AT DNB

WORLD BANK - FINSAC CONFERENCE ON FINTECH

IMAN VAN LELYVELD

22 MAY 2019

Agenda

Process

How did it go? What have we accomplished so far?

Technology

Challenges & Solutions

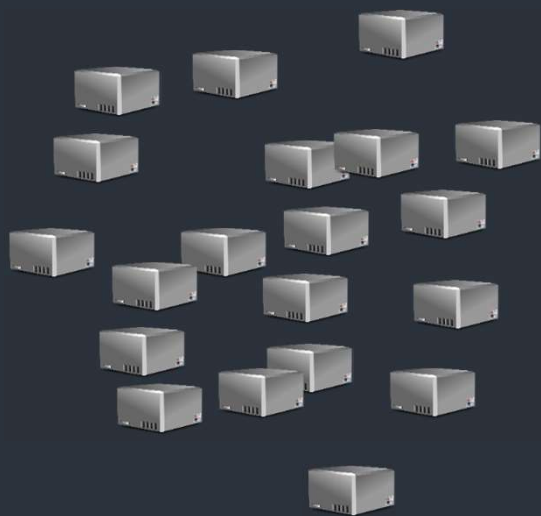
Showcase

Explanation of PoCs: what are the new possibilities?

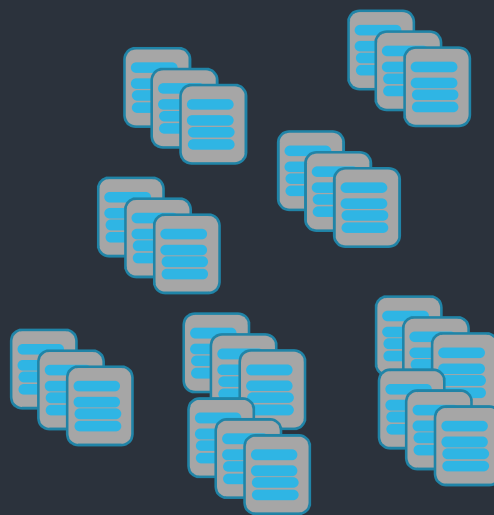
What's next?

What have we learned?

A dot on the horizon ...



IT and tools

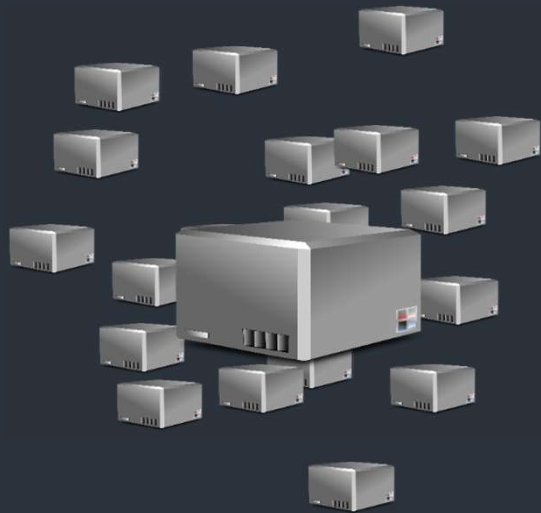


Data



People

A dot on the horizon ...



IT and tools



Data



People

The dream ...



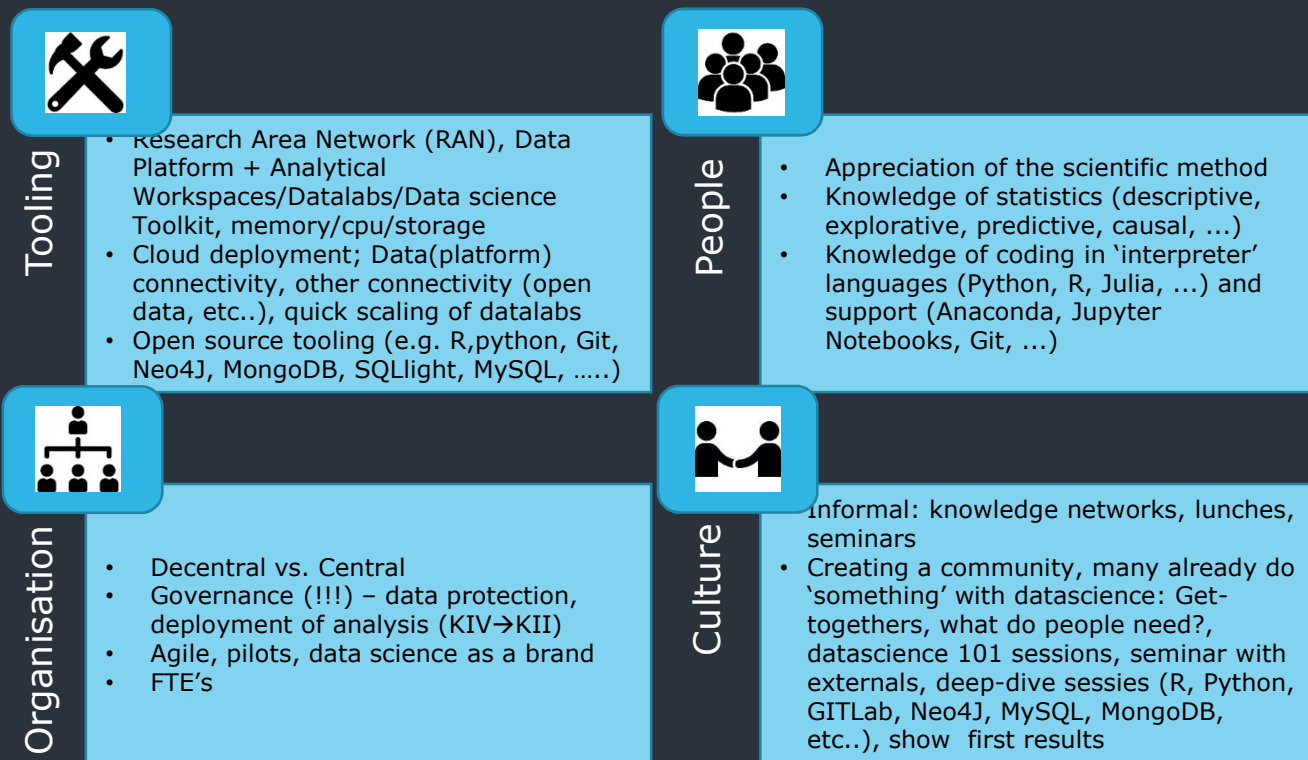
... and reality



DeNederlandscheBank

EUROSISTEEM

Which elements do we need?



Process: venturing down unbeaten paths...

- Means and projects
- Legal
 - tension between experiments and a complete contract
- Technology



Agenda

Process

How did it go? What have we accomplished so far?

Technology

Challenges & Solutions

Showcase

Explanation of PoCs: what are the new possibilities?

What's next?

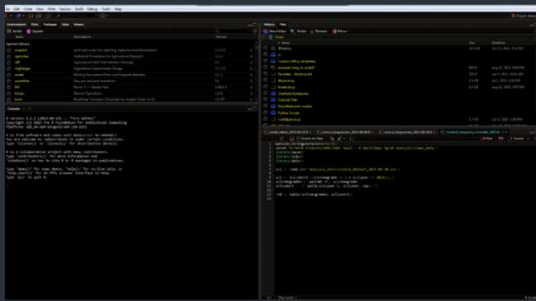
What have we learned?

Modern
Flexible
Scalable

Secure
Manageable
Traceable

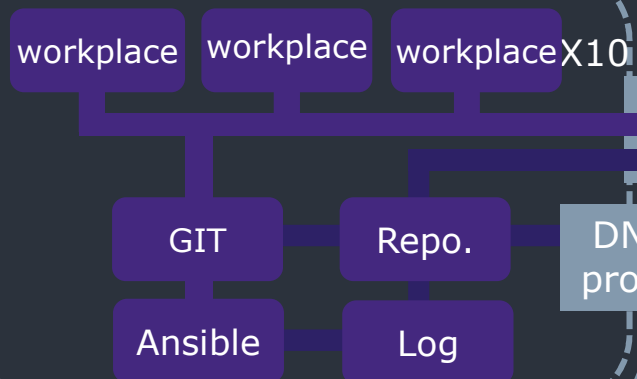
DeNederlandscheBank

EUROSYSTEEM



DNB

RAN High Secure

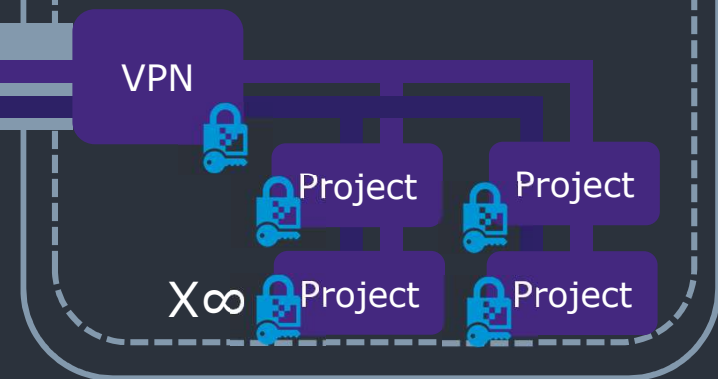


Internet

VPN Tunnel
SSH Tunnel
HTTPS

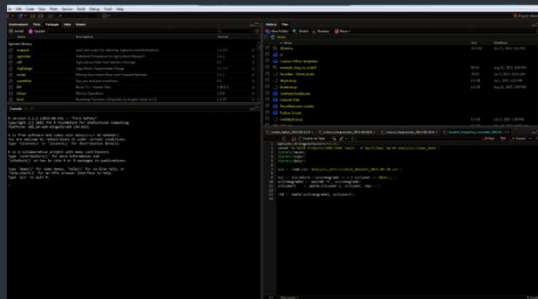
Surf Sara

DNB Net.



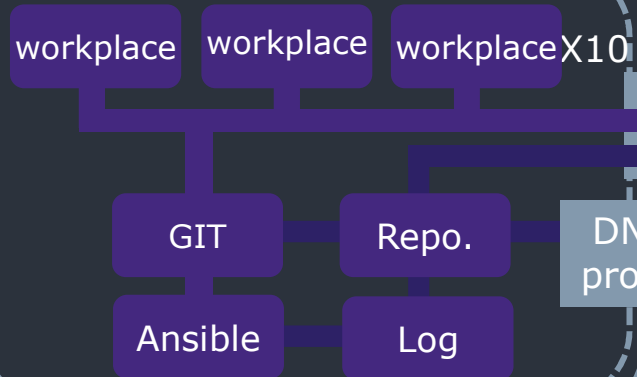
DeNederlandscheBank

EUROSYSTEEM



DNB

RAN High Secure



Internet

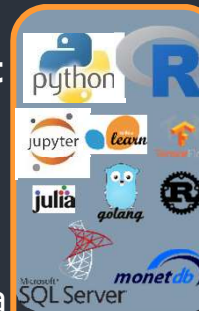
VPN Tunnel
SSH Tunnel
HTTPS

Surf Sara

DNB Net.

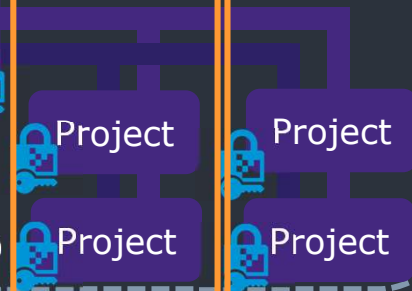


X ∞



Per VM
- 100TB
- 80cores
- 512GB

X.. VM's
per project



DeNederlandscheBank

EUROSYSTEEM

Agenda

Process

How did it go? What have we accomplished so far?

Technology

Challenges & Solutions

Showcase

Explanation of PoCs: what are the new possibilities?

What's next?

What have we learned?

PoC 1: Cred
 PoC 2: CCP
 PoC 3: CDS
 PoC 4: IRS
 PoC 5: Patter
 PoC 6: Resid
 PoC 7: Ana
 PoC 8: Futu

SharePoint

BLADEREN PAGINA

Home Groepen Taken

Overview

Over deze site

Information

Knowledge Documents (DNB-UNRESTRICTED)

Knowledge Documents (DNB-RESTRICTED)

Blog

Internal Contacts

Calendar

Onze projecten

Links

Notitieblok

Wiki

Recent

Over Ons

SharePoint

BLADEREN ITEMS LIJST

Home Groepen Taken

Overview

Over deze site

Information

Knowledge Documents (DNB-UNRESTRICTED)

Knowledge Documents (DNB-RESTRICTED)

Blog

Internal Contacts

Calendar

Onze projecten

Links

Notitieblok

Wiki

Recent

Over Ons

PoC 5: Pattern recognition in Solvency II reports

Pattern recognition in Solvency II reports

Topic:

In this project we use machine learning algorithms to detect rules and patterns within Solvency II supervisory reports and between the data. The Solvency II supervisory reports cover a wide range of subjects and contain financial information of insurance undertakings on a quarterly basis. This information plays an important role in our day-to-day insurance supervision. In this project we want to automatically identify points of interests contained in these reports, such as erratic behavior of specific insurance undertakings, or adverse developments. By defining a smart set of features based on the supervisory reports, we want to automatically identify non-obvious relations ('rules') between the financial condition of insurance undertakings and market developments. We want to identify outliers, i.e. developments in one insurance undertaking that behave differently than in other insurance undertakings, and we might be able to do this based on identified patterns and general financial market data.

Why is it important?

- Automatic detection of points of interest and data errors could increase the efficiency of insurance supervision;
- By applying machine learning algorithms we can detect obvious patterns but also non-obvious relations and points of interest, i.e. patterns in the data;
- The approach taken, models and the experience gained can be applied for other (supervisory) reports (pension funds and banking).

Key questions:

- What are the patterns in Solvency II reports in combination with financial market data?
- To what extent can abnormalities in Solvency II be detected using pattern recognition?
- To what extent can Solvency II reports be predicted in combination with identified patterns and financial market data?
- What is the performance of different algorithms with the above mentioned questions?

Titel : PoC 5: Pattern recognition in Solvency II reports (1)


Titel : PoC 6: Residential Real Estate (1)

Titel : PoC 7: AnaCredit (1)

Titel : PoC 8: future securities statistics (1)

objecten TB Training Open Source Lunches

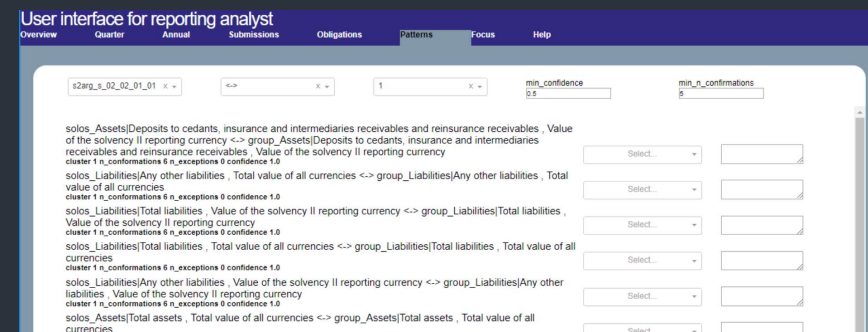
BLOGS



Word Datapreneur
 We zijn bezig met het Word Datapreneur leertraject. Wat is dit eigenlijk?

PoC 5: Pattern recognition in Solvency II reports

- Plausibility of group reporting in Solvency II
 - Sizable reports (> 4.000 dimensions per group)
 - Domain knowledge required
 - Specifying DQ checks labour intensive
- Applying machine learning algorithms
 - Checking causal relations within reports
 - Comparing group reporting and solo reporting
- On going work
 - Further DQ checks
 - Linking with other financial data
 - Other algorithms (# dimensions >> # groups)
 - User-interface for visualising and feedback



Agenda

Process

How did it go? What have we accomplished so far?

Technology

Challenges & Solutions

Showcase

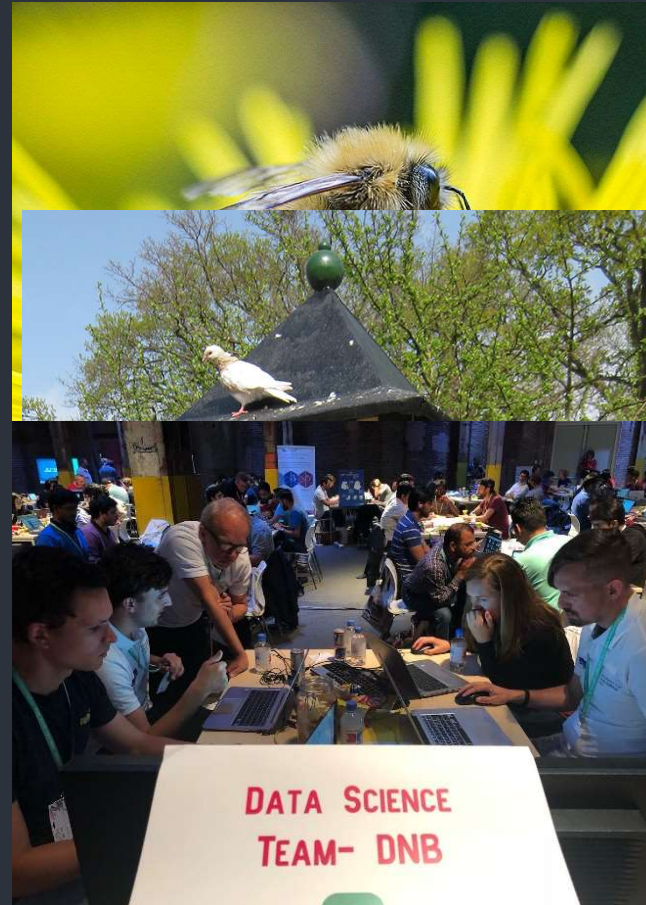
Explanation of PoCs: what are the new possibilities?

What's next?

What have we learned?

Lessons?

- Cross-pollination
 - Lelione
 - STAT ML
- The hard core
- Hackathon



Community

- Python & R lunches
- A whiff of Data Science
 - Hands on case study with Jupyter Notebook for DNB board and management
- Manifest
 - What is responsible data science?
- Datapreneur
 - 22 participants from across the whole bank starting with open source and tackling their own business problems
 - Python, R, GIT, Agile, co-coding
- Training
 - Data Science 101, Joint with DNB Academy