



Just do it

Data-Driven Risk & Compliance

Pressure Cooker

2016

Willem van Asperen & Michiel Krol (Michiel.krol@paconsulting.com)



Thanks for joining us...



Data Driven Innovator

- 18 years Financial Services Industry
- Actuarial background and Solvency
- Leads PA Consulting Group's Risk & Compliance Team in the Benelux



Data Scientist

- Background in Artificial Intelligence
- Leads PA Consulting Group's Data Science Team in the Benelux
- Speaks 14 programming languages

A red rectangular warning sign with rounded corners and a white border. It features two silver screws at the top and bottom center. The word "WARNING!" is written in large, bold, white, 3D-style capital letters at the top. Below it, the text "AIM OF THIS PRESSURE COOKER:" is written in white, all-caps, sans-serif font. At the bottom, the phrase "UNDERSTANDING HOW TO DRIVE RISK&COMPLIANCE VALUE FROM DATA" is written in white, italicized, all-caps, sans-serif font.

WARNING!

AIM OF THIS PRESSURE COOKER:

*UNDERSTANDING HOW TO DRIVE
RISK&COMPLIANCE VALUE FROM DATA*

This Pressure Cooker is about Entrepreneurship in (Y)Our daily reality...

1	The new normal: Customer online + Connected
2	New (distributed) computing methods + Sensors
	(Customer) Interaction Logging is the new oil

Warming-up: What is a killer case?



We have created an approach to accelerate Data Driven Innovations

rockstart.
We love startups.

Rock start is our inspiration. Killer cases are our start-ups

rockstart.

We love killer cases.

The birth of a killer case



The birth of a killer case

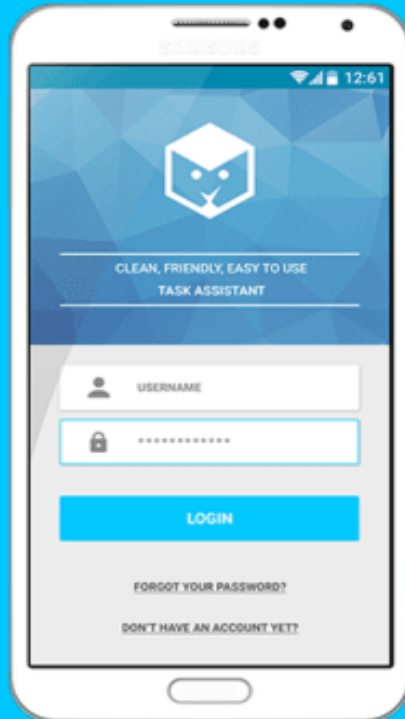


The birth of a killer case

Define
(Portfolio Management)

Deliver
(Insight project delivery)

Deploy
(Benefit realisation)



MARKO, YOUR SMART TASK BUDDY

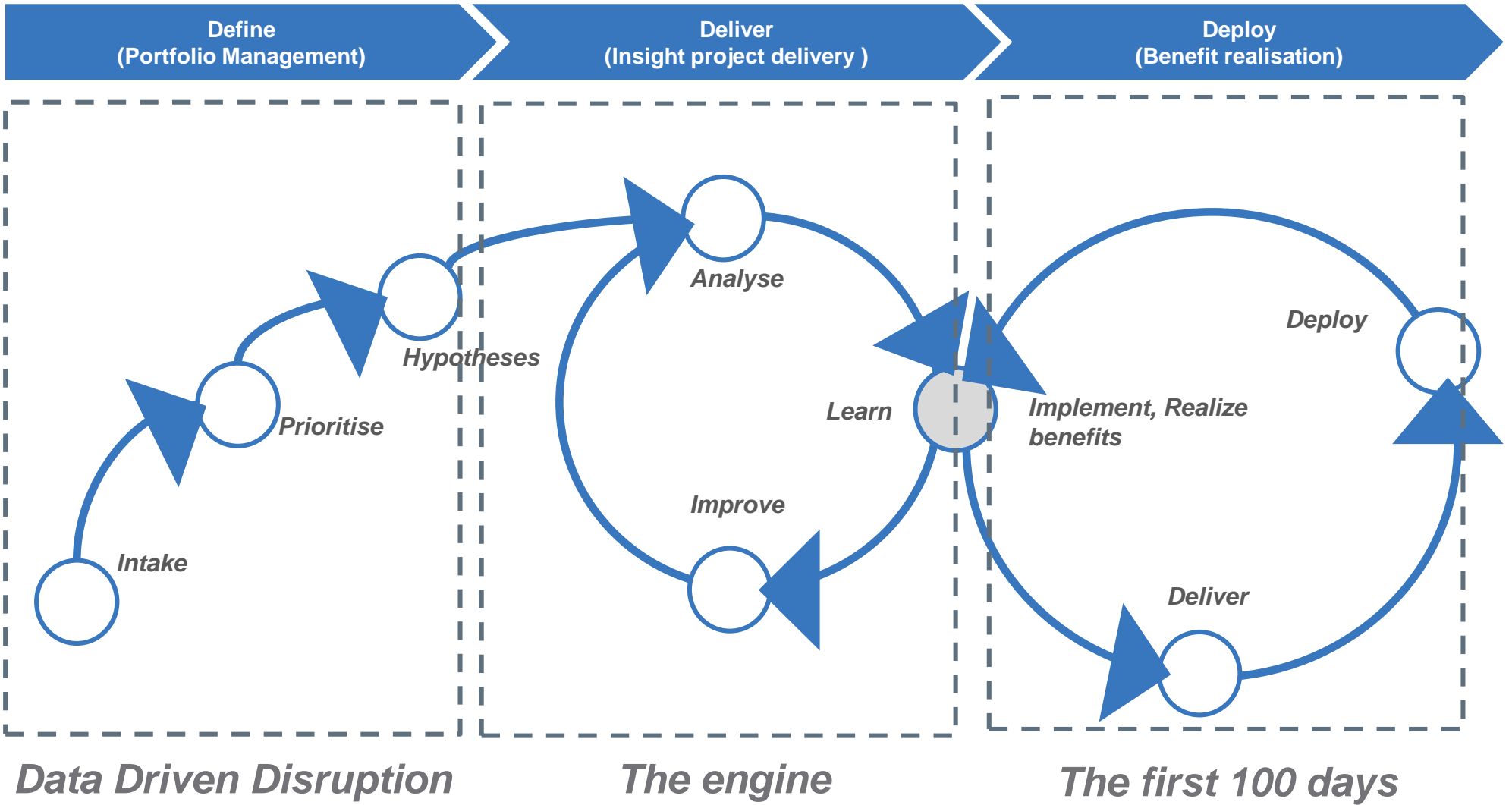
is part of

rockstart.

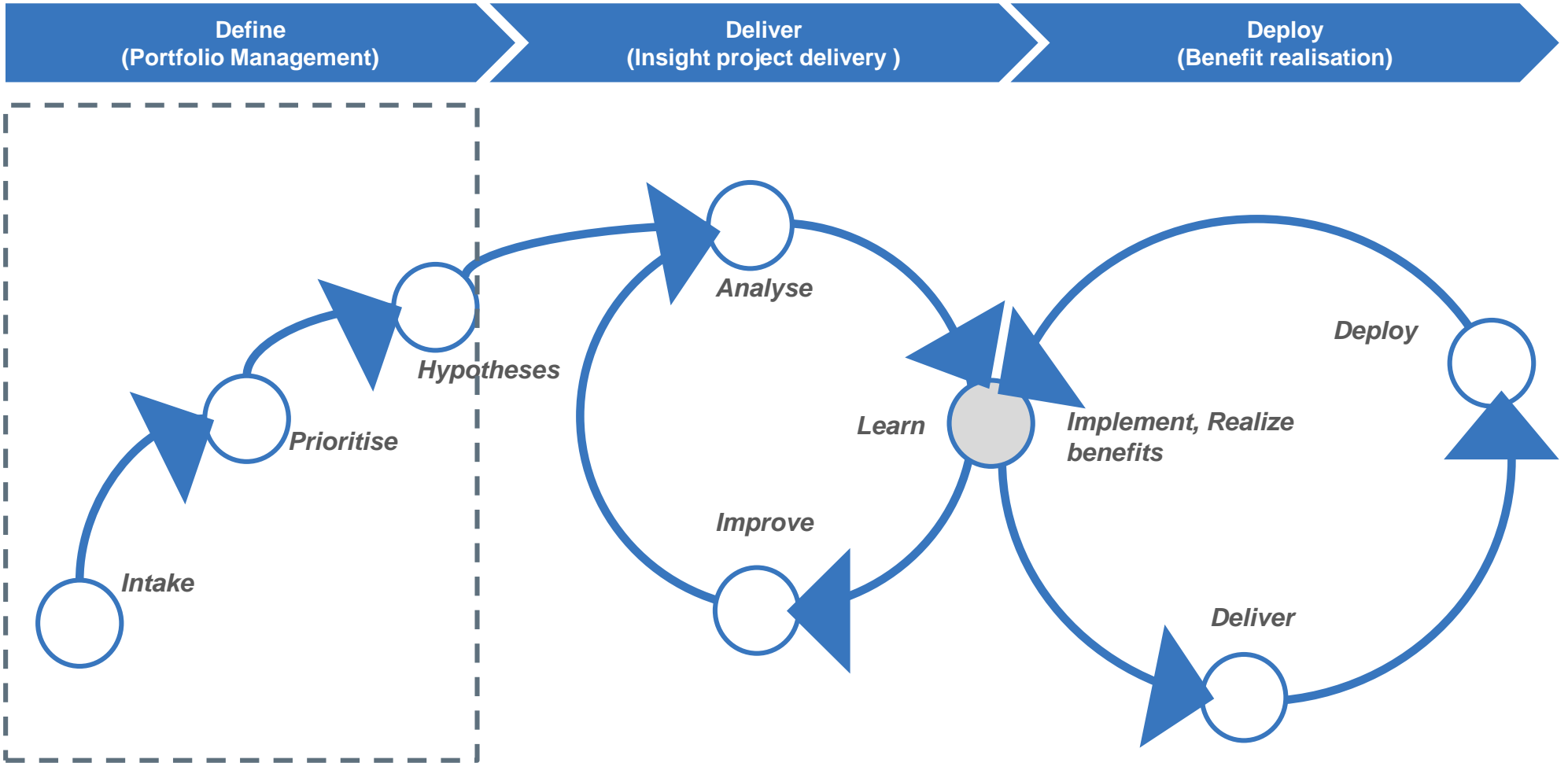
ACCELERATOR

www.marko.rocks

The birth of a killer case, explained in three Pressure Cooker Blocks



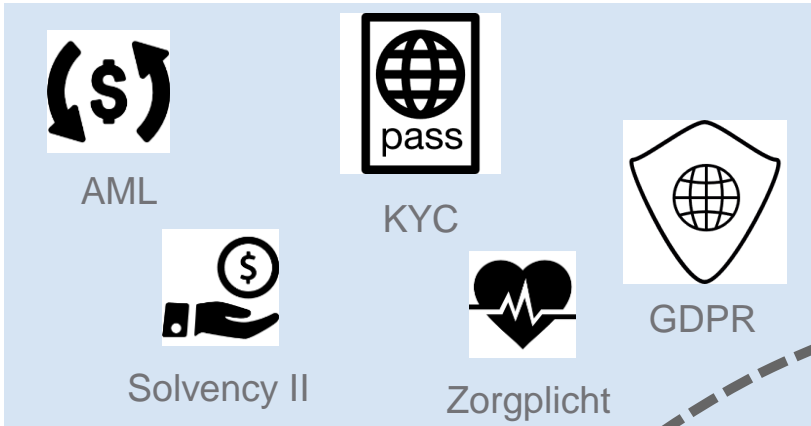
The birth of a killer case, explained in three Pressure Cooker Blocks



Data Driven Disruption

Data as a key driver of Risk & Compliance impact in a Digital World

Many regulatory policies



CHALLENGE!

Few regulatory policies

Traditional Risk & Compliance

Sufficient budget and FTE



Pressure on budget and FTE

Data as a key driver of Risk & Compliance impact in a Digital World

Many regulatory policies

AML
 Solvency II
 KYC
 Zorgplicht
 GDPR

Smart Risk & Compliance



Few regulatory policies

Traditional Risk & Compliance

Achmea schrapt 4000 banen
 Verzekeringconcern Achmea schrapt de komende 3 jaar ongeveer 4000 banen. Dat heeft het concern, de grootste verzekeraar van Nederland, woensdag aangekondigd, waarvan

[NU.nl](#) > [Economie](#) > [Ondernemen](#)

Foto: ANP

n Duin, in opdracht van de provincie. Het project kost 250 miljoen en wordt uitgevoerd door de aannemers Zilveren Kruis.

Rabobank schrapt in komende jaren 9.000 banen

Sufficient budget and FTE

Pressure on budget and FTE

Advanced Data Analytics help large banks to be compliant with Know Your Customer (KYC) regulations and reduce compliance costs.

Situation

- Enhanced Due Diligence regulations
- Verification

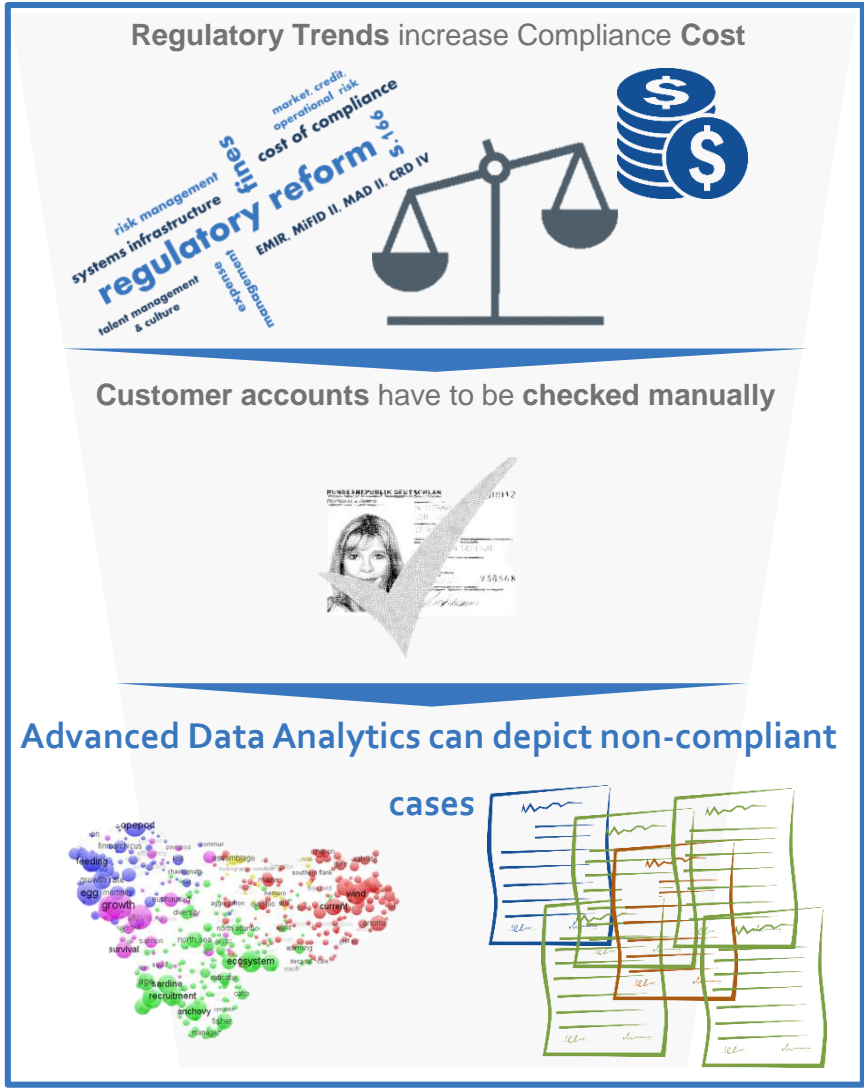
Challenge

- Unstructured files
- Spot checks
- Time-consuming
- Cannot ensure 100% compliancy.

Efficient compliancy while reducing manual labour?

Answer

- Advanced data analytics
- Machine learning, predictive analytics
- Probability



European Data Protection Regulations – enlarged responsibilities for financial institutions concerning Personal Data

Situation

- Data Protection Regulations
- Protection of personal data
- The penalty of 5% of worldwide revenue

Challenge

- Contracts are scattered and unstructured
- Time intensive and sampling
- Not continuous

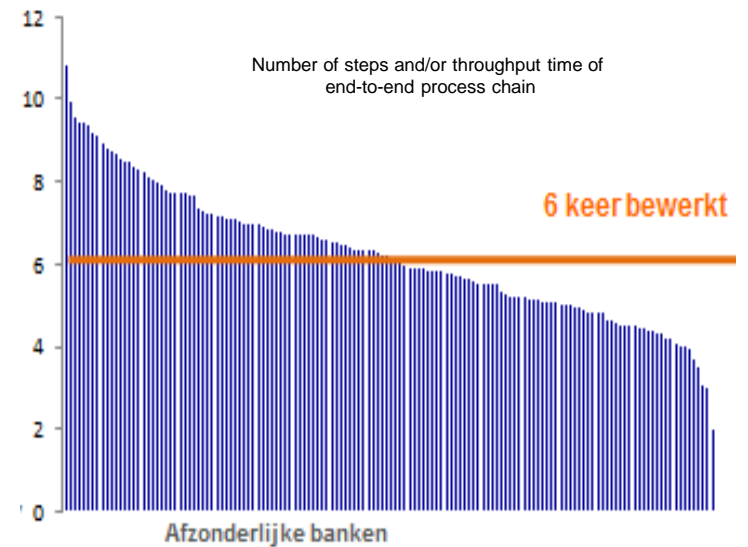
How can numerous files be reviewed in a short period of time to ensure compliancy?

Answer

- Advanced analytics, text mining, natural language processing
- Data driven compliance officers and data scientists
- 6-8 weeks



Bottleneck remover | reduce the process steps and save up to 20% costs



Who is this?



The Palantir of

Infrastructure (Palantir products):

- Palantir Gotham: Used to integrate, secure, manage and analyze enterprise data.
- Palantir Metropolis : Used to integrate, analyze, enrich and model any kind of quantitative data

Killer cases (Palantir solutions)

- | | |
|--|--|
| <ul style="list-style-type: none"> • Anti Fraud • Capital Markets • Case Management • Crisis Response • Cyber Security • Defense • Disaster Preparedness • Disease Response • Healthcare Delivery | <ul style="list-style-type: none"> • Insider Threat • Insurance Analytics • Intelligence • Law Enforcement • Legal Intelligence • Palantir Verus • Pharma R&D • Trader Oversight • Custom Solutions |
|--|--|

.... Banking (Killer cases)



Infrastructure (Palantir products):

- Palantir Gotham: Used to integrate, secure, manage and analyze enterprise data.
- Palantir Metropolis : Used to integrate, analyze, enrich and model any kind of quantitative data

Killer cases (Palantir solutions)

- | | |
|--|--|
| <ul style="list-style-type: none">• Anti Fraud• Capital Markets• Case Management• Crisis Response• Cyber Security• Defense• Disaster Preparedness• Disease Response• Healthcare Delivery | <ul style="list-style-type: none">• Insider Threat• Insurance Analytics• Intelligence• Law Enforcement• Legal Intelligence• Palantir Verus• Pharma R&D• Trader Oversight• Custom Solutions |
|--|--|

.... Your business

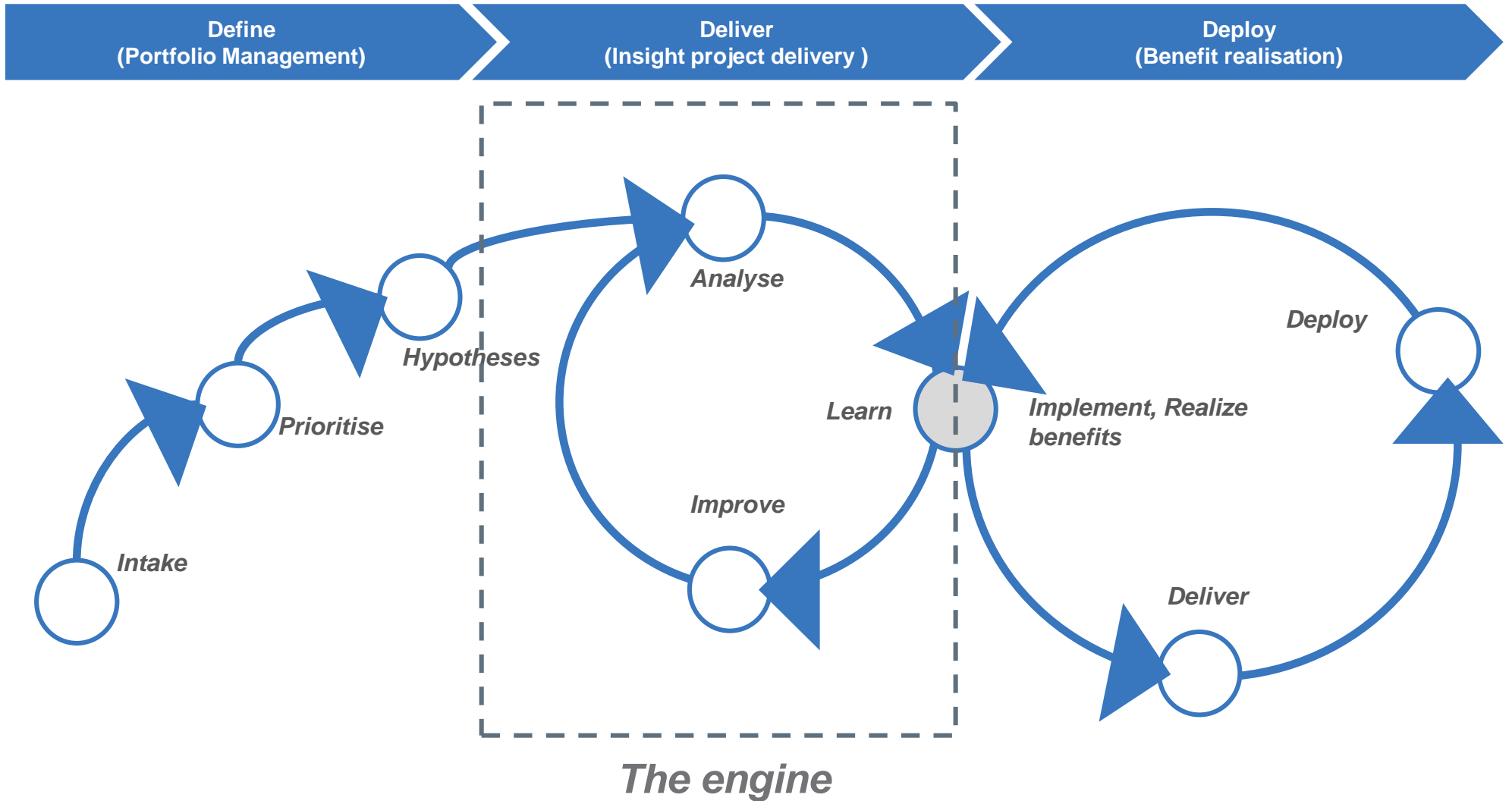


Pressure Cooker Workshop: Idea generator

- We split in groups
- Brainstorm potential killer cases for Risk & Compliance
 1. Where do we see **risk of mal-compliance**? Which regulations? What makes it difficult?
 2. Where do we see **efficiency potential**? Which processes? What makes it difficult?
 3. How can we make more **impact**? Early warning indicators, prioritization?
- Pitch your top-3 ideas to the group



The birth of a killer case, explained in three Pressure Cooker Blocks

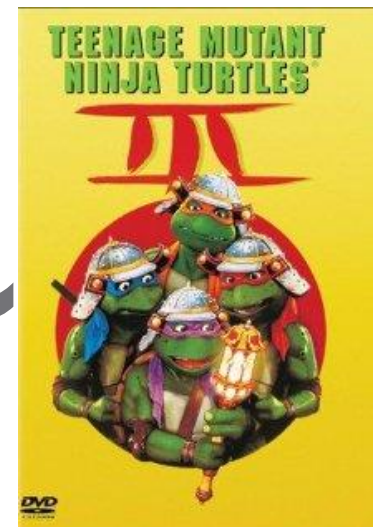
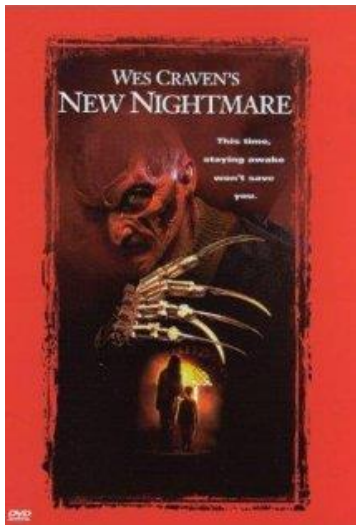


Methods developed in the 70s can now be applied to large data sets

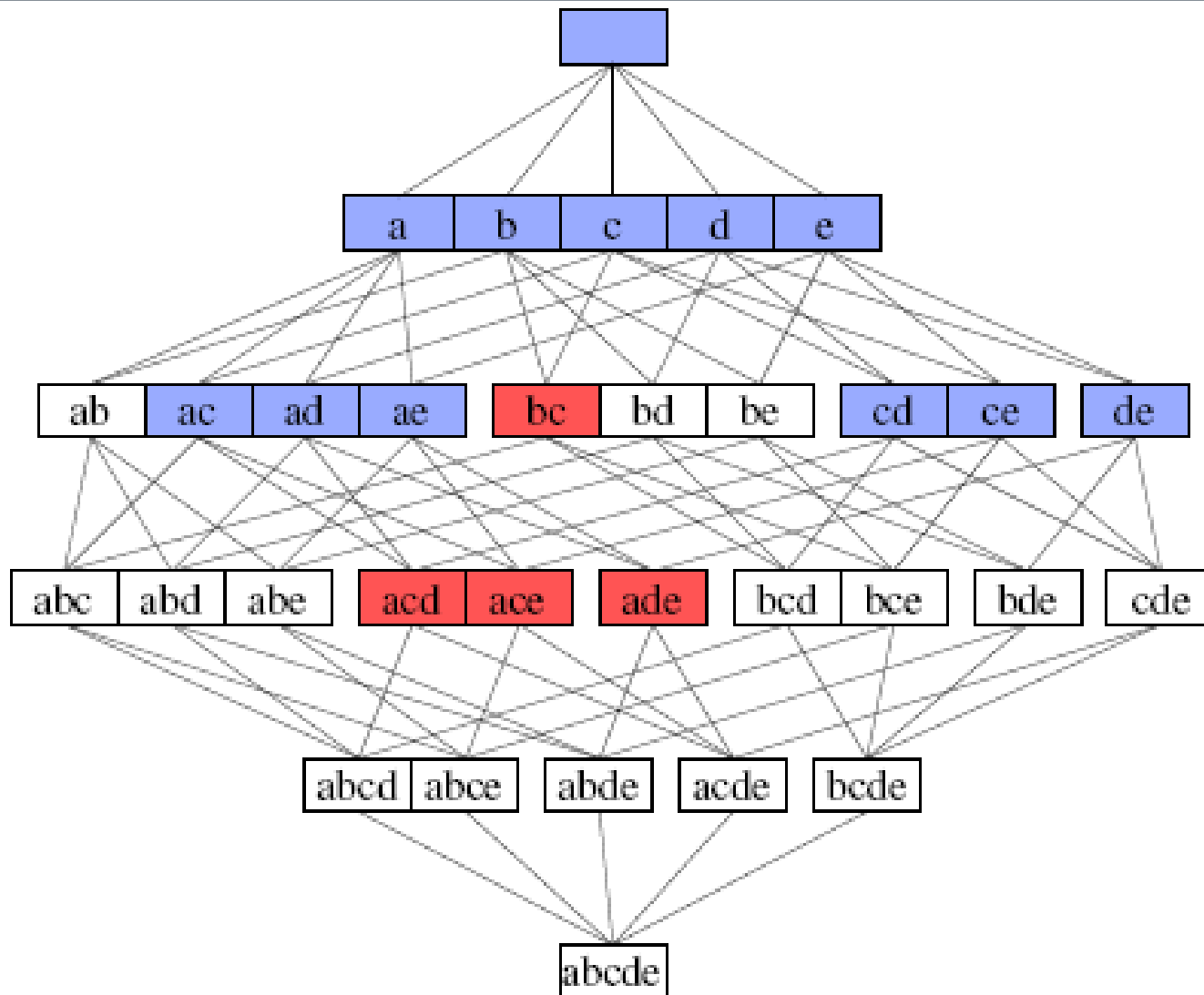
- Pearson Correlation Coefficient

$$r = r_{xy} = \frac{n \sum x_i y_i - \sum x_i \sum y_i}{\sqrt{n \sum x_i^2 - (\sum x_i)^2} \sqrt{n \sum y_i^2 - (\sum y_i)^2}}$$

- Example: which movies are scored high together?



Amazon uses Frequent Item Sets to analyze “shopping baskets”



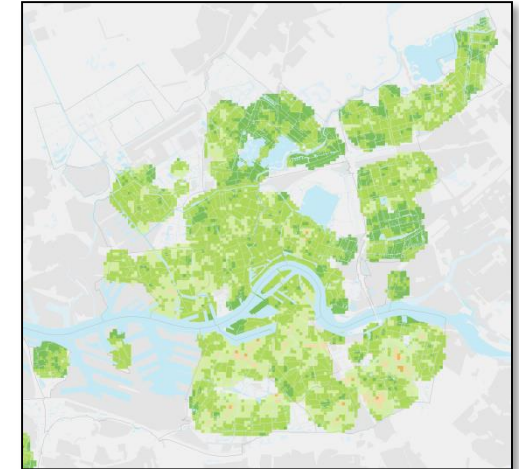
Municipality of Rotterdam use Decision Trees to spot potential fraud

```

huwdat > 20 Nov 1999: FALSE (2416/240.5)
huwdat <= 20 Nov 1999:
:...C.leeftijd > 84: TRUE (171.1/29.9)
  C.leeftijd <= 84:
:...MIN_BWJR_PND <= 1968:
    :...MIN_BWJR_PND > 1924: FALSE (1475.4/308.5)
    :   MIN_BWJR_PND <= 1924:
    :     :...C.leeftijd <= 55: TRUE (595.5/271.7)
    :     :   C.leeftijd > 55: FALSE (367.2/84.1)
MIN_BWJR_PND > 1968:
:...buurt in {1 ... 98}: FALSE (972.9/311.5)
  buurt in {12 ... 99}:
    :...BELBVEILIG2013 <= 1.8277: TRUE (1078.9/375.5)
    :   BELBVEILIG2013 > 1.8277: FALSE (293/120.1)
  
```

1

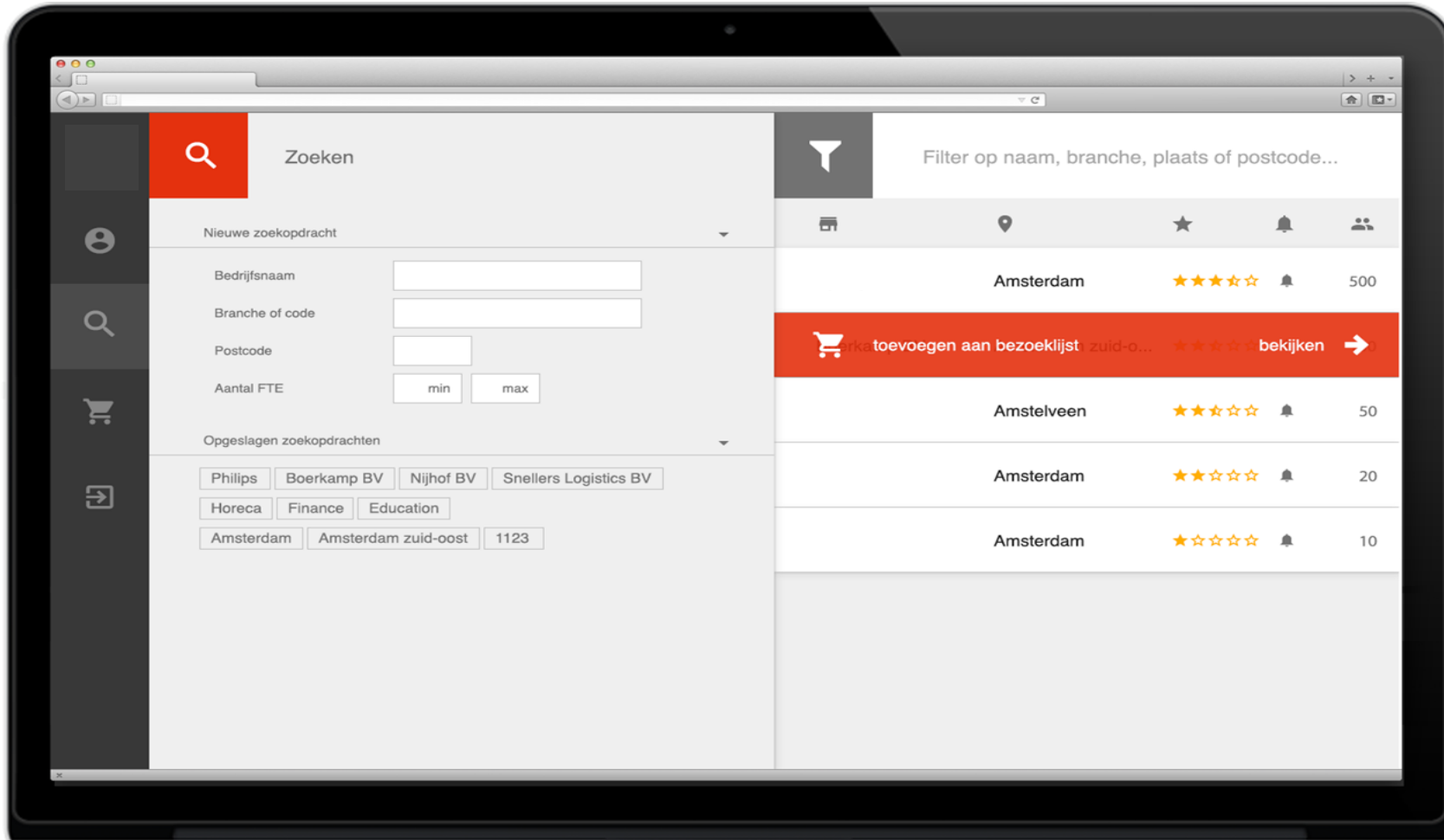
2



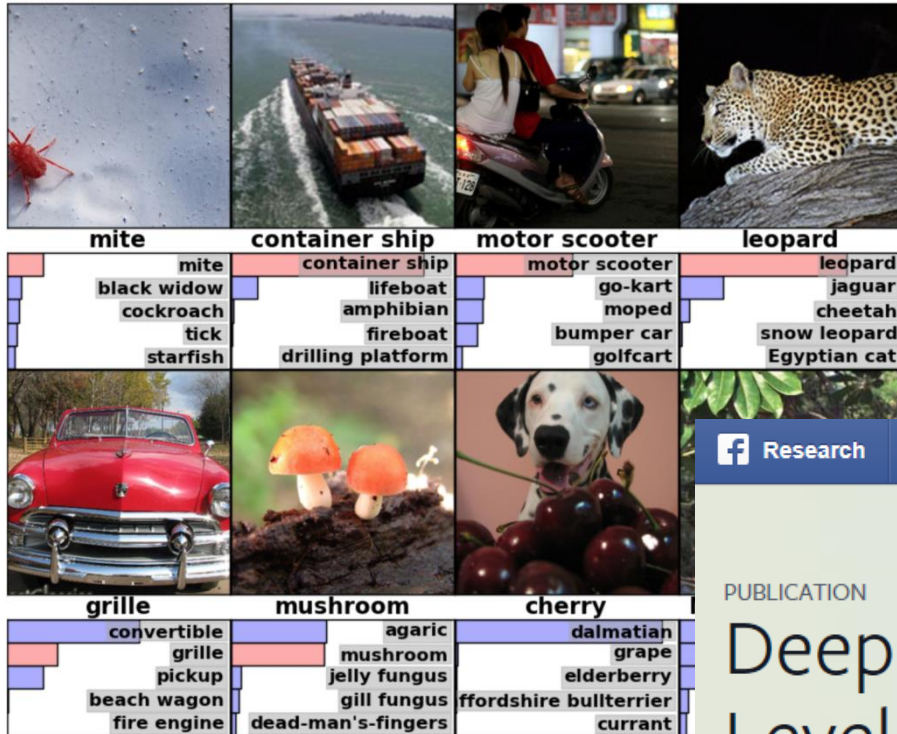
Determining factors for “married, not living at the same address” appear to be:

1. Age and age of the home
2. The sense of safety in the local neighborhood

A Temp Agency uses Random Forrest to spot propensity to buy



University of Munich use Deep Learning to classify random images to support patients with disabilities – Facebook is researching DeepFace



Facebook navigation: Research, Our Research, Programs, Publications, Events, Blog

PUBLICATION

DeepFace: Closing the Gap to Human-Level Performance in Face Verification

Yaniv Taigman, Ming Yang, Marc'Aurelio Ranzato, Lior Wolf
 Conference on Computer Vision and Pattern Recognition (CVPR) · 24 juni 2014

Facebook interaction: Vind ik leuk, Delen, 222

Over the last seven years, the way of data analytics has fundamentally changed

How we used to do it	How we are now doing it
Closed Source	Open Source
Scale Up	Scale Out
Schema on Write	Schema on Read
Bring Data to experts	Bring experts to the Data
Bring expert knowledge to the data	Bring data to the experts
Statistics	Machine Learning
Innovation led by corporations	Innovation led by the crowd
Build to last	Build what is required NOW

“Hadoop was built in the Google distributed file system (2003) and a Google paper on Map-Reduce (2004)”

Conclusion

This is how Netflix can predict the titles that are most likely to be of interest to you

Scores per title of individual users

	A	B	C	D	E	F	G
user 101	3	2	4		3	4	3
user 102		3		3	3		5
user 103	3	5			3		
user 104	4	3	5	5	1		4
user 105	5		2		5		3
user 106	1	4		3	5	5	3
user 107	3	3	5	5	2	4	
user 108			2			3	3
user 109	3	5		3		4	2
user 110			3	3			
user 111	4		1	5		2	3
user 112	3		5	1			
user 113		2		5		5	3
user 114	5	4			4		
user 115	2					4	3
user 116	4		3		3	3	3
user 117	5			4		3	5
user 118		5	4	1	5		
user 119	3	3			1	5	
user 120							

Correlations between the scores of individual titles

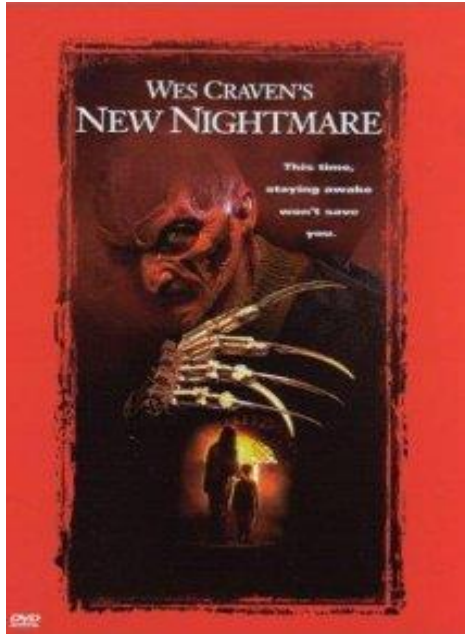
	A	B	C	D	E	F	G
A	1.0	-0.1	-0.3	0.5	-0.2	-0.3	0.6
B	-0.1	1.0	-0.3	-0.3	0.6	-0.3	-0.1
C	-0.3	-0.3	1.0	-0.3	-0.3	0.7	-0.2
D	0.5	-0.3	-0.3	1.0	-0.3	-0.2	0.6
E	-0.2	0.6	-0.3	-0.3	1.0	-0.4	-0.2
F	-0.3	-0.3	0.7	-0.3	-0.4	1.0	-0.1
G	0.6	-0.1	-0.2	0.6	-0.2	-0.1	1.0

Scores of title F are strongly correlated with those of title C



User 104 will therefore probably like title F

$$r = r_{xy} = \frac{n \sum x_i y_i - \sum x_i \sum y_i}{\sqrt{n \sum x_i^2 - (\sum x_i)^2} \sqrt{n \sum y_i^2 - (\sum y_i)^2}}$$

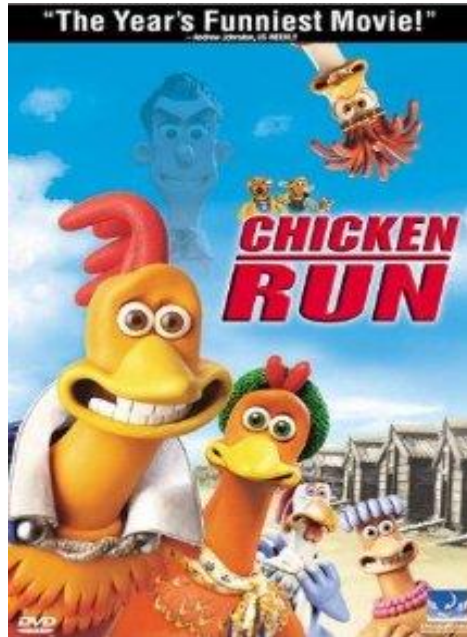


A demonic force has chosen Freddy Krueger as its portal to the real world. Can Heather play the part of Nancy one last time and trap the evil trying to enter our world?

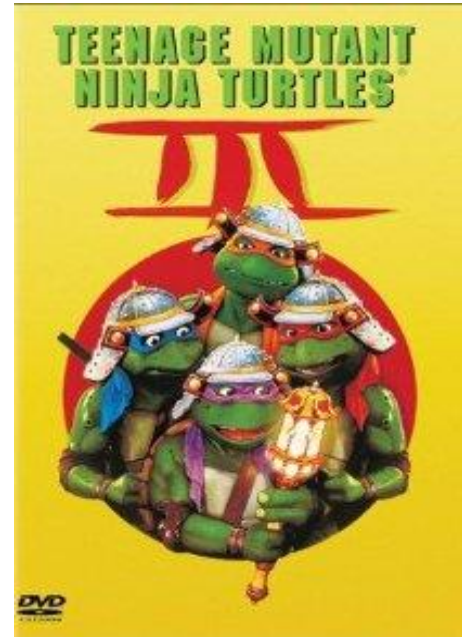


Chucky, the doll possessed by a serial killer, returns for revenge against Andy, the young boy that defeated him and has since become adult.

$$r = r_{xy} = \frac{n \sum x_i y_i - \sum x_i \sum y_i}{\sqrt{n \sum x_i^2 - (\sum x_i)^2} \sqrt{n \sum y_i^2 - (\sum y_i)^2}}$$

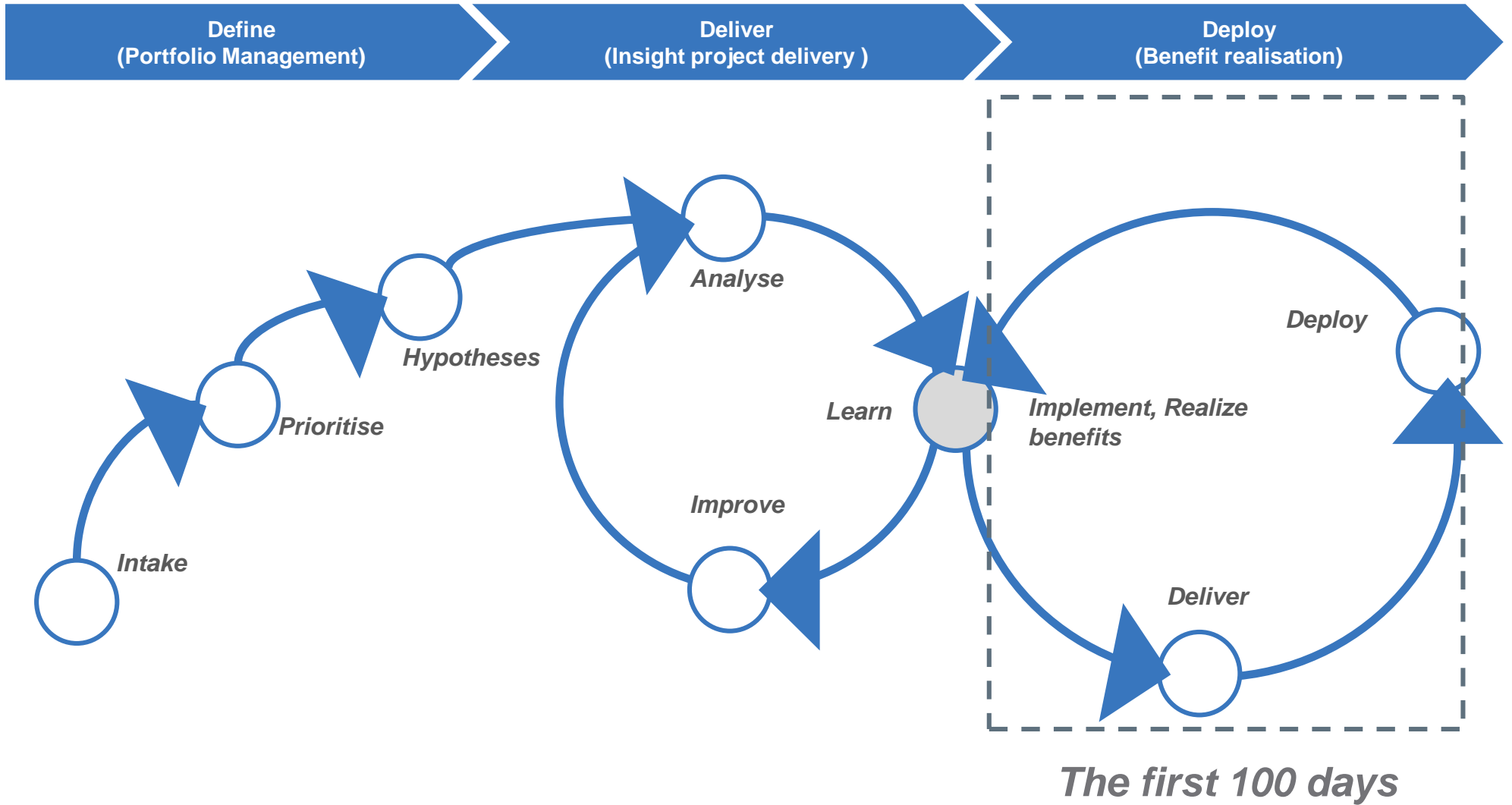


When a bird "flies" into a chicken farm, the fellow chickens see him as an opportunity to escape their evil owners.



The turtles find themselves transported back in time to ancient Japan.

The birth of a killer case, explained in three Pressure Cooker Blocks



Did we meet your expectations? What are your take aways?



How we help to accelerate Data Driven Innovation...



1 Data Driven Innovation Inspiration session



2 Data-Driven Transformation maturity scan

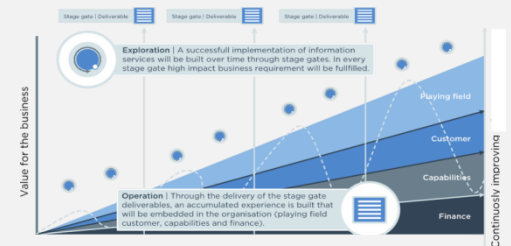
Maturity scan		Traditional Business	Data Driven Business
Enablers	Organization	Competenties	Getrainde data sciences en data-driven leadership capability
		Processen	Centraal geprioritiseerde focus op: data-driven projecten
Technology		Platform	Centraal gefunde intelligence hub
		Analysa tools / methoden	Real-time big data platform
Data		Data logistiek	Toepassing prescriptive analytics
		Governance / Process	Quick connect connecties met interne en externe bronnen
		Kwaliteit	Ingerichte data eigenaarschap en gebruiksprocessen
	Rijkheid		Validatie en uitwisseling ruwe en verrijkte data
			Taps naar ~50 interne en externe bronnen

3 Killer case projects



4 Sandbox platform

5 Digital transformation support



6 Data-driven transformation serious game

