# **IBM Connect**

WIECZOREK Jacques





- 1 Goal
- 2 NRB Group
- 3 A digital context
- 4 Advanced analytics & Big Data
- 5 Big Data own experience
- 6 Risks & Statistics
- 7 Conclusions



# > The Goal

**Concret set of activities related to Big Data** 





Who we are...





## Some figures













Turnover

146,7 m€

FTE

544

Finance & Insurance

Public & Social, (Reg. & Fed.)

**Utilities** 

Industry

23,4 m€

149

Health care

26,6 m€

188

Public & Social (Prov. & Loc) Brussels & South 30,0 m€ 4,9 m€

201

Public & Social (Prov. & Loc) North 1,7 m€

17

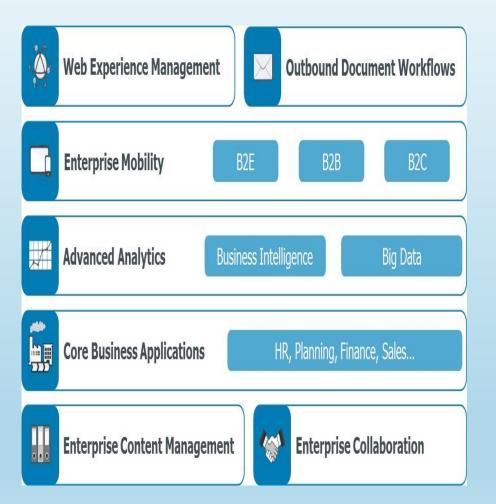
B2B Mobile & Web Development

2014 - turnover : 222,3 millions € - FTE : 1100 employees



### Solutions

#### Covering all the needs of your organization





**Mainframe** 

IBM z/OS: 7 500+ MIPS,

800+ million transactions/year

> Bull GCOS8

**Distributed systems** 

Windows, Linux, UNIX and AS/400

Storage & Back up

Storage capacity: > 1 PB (Petabyte) disk space

Automated backup management

**Workplace services** 

12 000+ work stations (including VoIP)

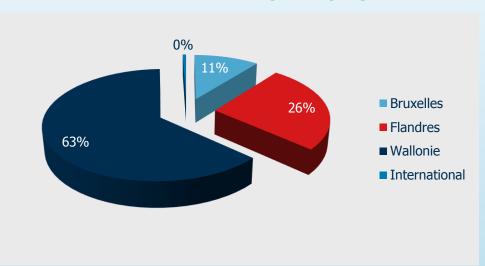
Network

Dedicated inter data center optical fiber connectivity Multiple broadband suppliersin place with dedicated fibers

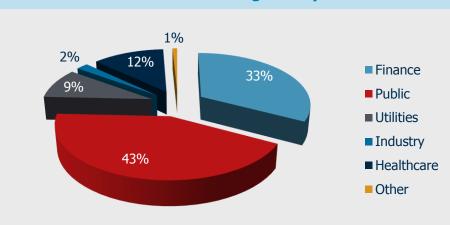


#### Revenues breakdowns

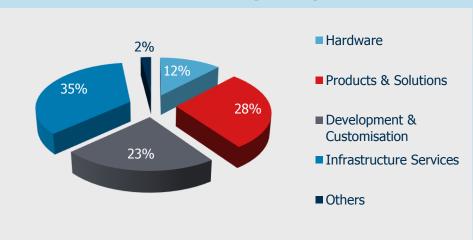
#### **Distribution of sales figures by region**



#### **Distribution of sales figures by sector**



#### **Distribution of sales figures by services**





Our main customers (NRB)

Finance & Insurance



delta lloyd









Public & Social Fed. & Reg.















**Utilities** 

















Industry





















Our main customers (xperthis, Civadis, Cevi & logins)

Health care











Public & Social Loc. South



Public & Social Loc. North















# A digital context





# Digital transformation: we are all learning





## Customer experience is the new strategic battefield

#### A product/ a service experience

## **Use of Technology** is exploding

- ☐ Moore Law
- □ Improvement of Data access through expansion of physical network
- ☐ Web 2.0 & participative collaboration



### **Power of the Mass**



- □ Always connected
- **Always sharing**
- □ Always aware
- □ UGC
- □ yesterday<**sharing**<today</p>

- More access
- More influence



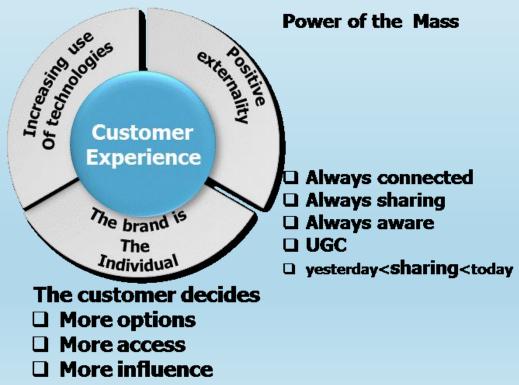
## Scope of the customer experience

We carry our consumers habits and expectations into any workplace!



# Use of Technology is exploding

- Moore Law
- ☐ Improvement of Data access through expansion of physical network
- Web 2.0 & participative collaboration



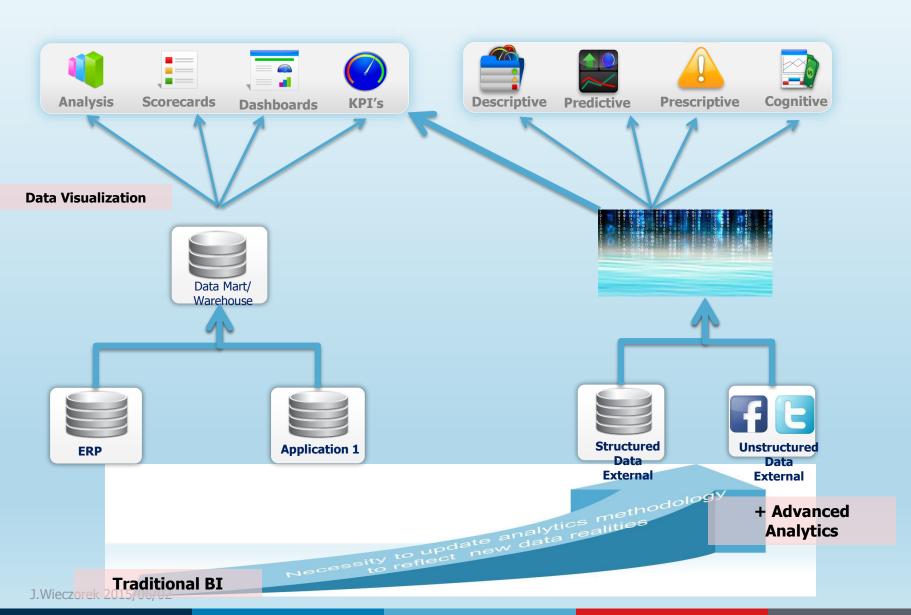


# Advanced Analytics & Big Data





## BI but Advanced Analytics looming on the horizon





### <u>Vision of the Walloon region: sustaining the Marshall plan through Big Data</u>

Getting the 6 competitiveness Clusters ahead competition

Master Global Information

- Local Data
- Open Data

Securing Data Privacy for some specific sectors Healthcare & Industries

- Private Cloud
- Local Storage
  - 2 Data Centers

Retaining Human Capital

- High potential jobs
  - Data Scientist
  - Business Scenario Modeler (BSM)

Turning R&D outcomes into production

- Research Centers
- Universities
- IT Companies

Expectation from NRB:
To drive the process as major ICT company

## **NRB's Big Data Hub**



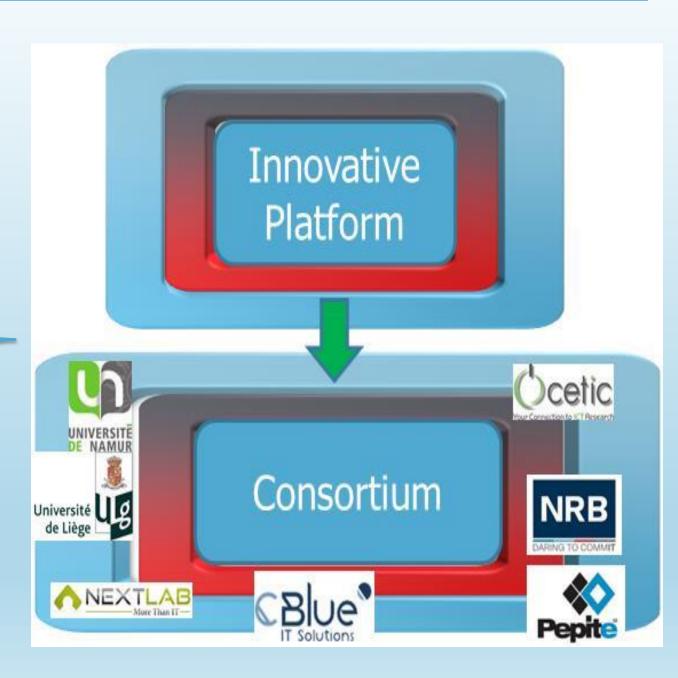














## Some specific added values

Added Value through Big Data

**Increase Cash Flow** 

**Asset Efficiency** 

Governance

**Customer Insight** 

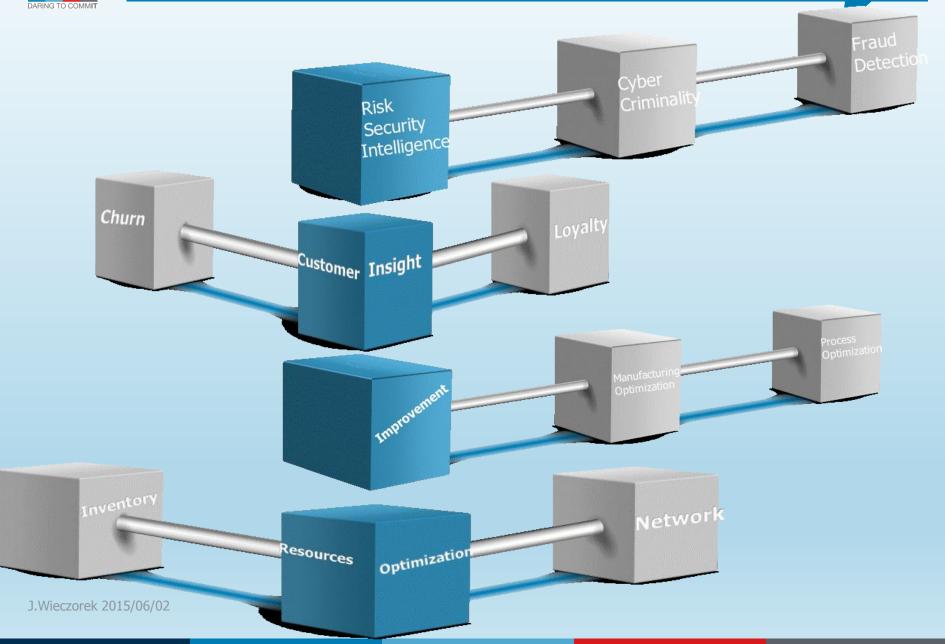
Resources Optimization Process
Productivity
Improvement

Risk Security Intelligence

J.Wieczorek 2015/06/02

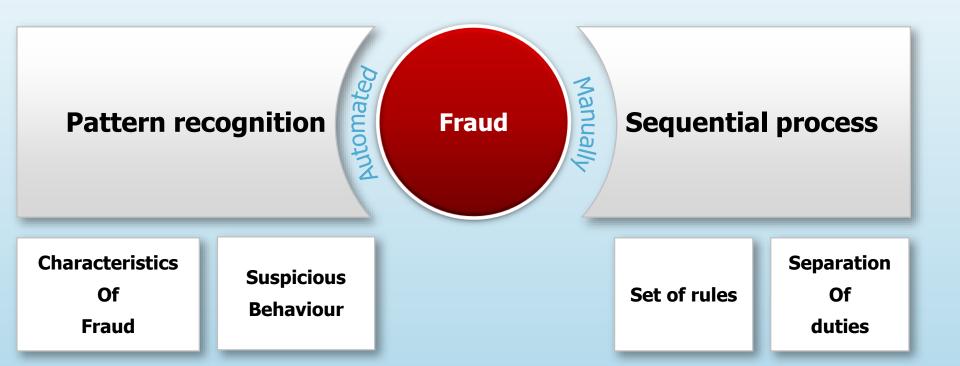


# Detailed added value through Big Data





## How to detect and prevent fraud?



Governance and fraud Risk management



### The context of the Fraud

#### **Fraudster**

- Average offender ( random and /or occasional dishonest behavior
- Professional fraudster: organized / group crime offender

Files Fraudster Fraud Reimbursement Fraud Ring Internal Data

### **Availability problem**

- Obtaining real data to legal and competitive reasons
- Alternative :
  - Create synthetic data which matches closely to actual data ( data generation)

Once-Off
Occasional
Fraud is:
Seasonal
Regular

## Services Portfolio

## **Consultancy**

- Scope definition ( choice of the service analytics )
- Business experts
  - Business scenario modeller

# **Design Implementation**

- Design
- Customization of an existing
- Analytics packages :
  - predictive, prescriptive, descriptive, diagnostic,...
- Research & Experimentation of new algorithm or new analytics package

### Run

- Deployment of an analytics package or algorithm
  - Execution
  - Maintenance



## A powerfull end to end service

#### Consultancy

- service analytics )
- Business experts
  - Business scenario
- All stakeholders

#### Design **Implementation**

- Analytics packages
  - predictive, prescriptive, descriptive, diagnostic,...
- Research & Experimentation of new algorithm or new analytics package
- All stakeholders

#### Run

- analytics package or

  - Maintenance

Powerfull Infrastructure

#### **Primary data center**







**Dedicated Hadoop** cloud

#### Secondary data center





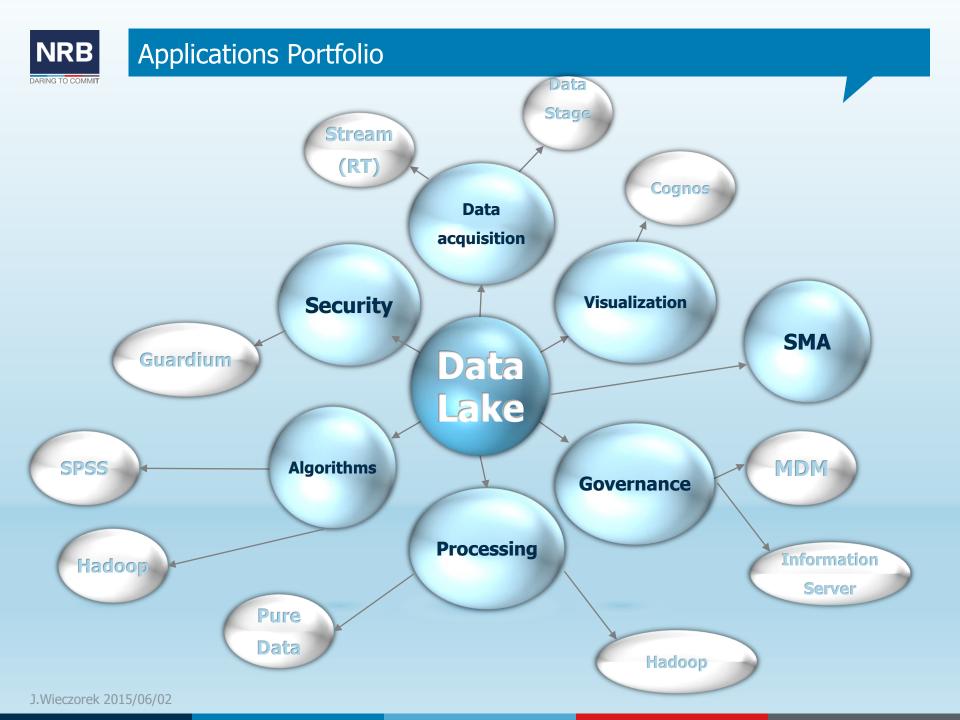


**PureData** 4 Analytics NRB shared cloud



Hadoop

- **Computation cloud based on dedicated Vblock**
- **Two dedicated infrastructures for Hadoop service**
- **Shared infrastructure in secondary data center**





## The value of having a Social Media Strategy

Tailoring product and services to specific needs





## The goal is to increase interactions with customers



- Understand attitudes, opinions and evolving trends;
- □ Correct structural claims faster than competitors;
- Create customized campaigns;
- ☐ Understand what customers like but also don't like;
- ☐ Can be an early warning system when customers are turning against;
- Managing reputational risks.

Product satisfaction



# Big Data own experience



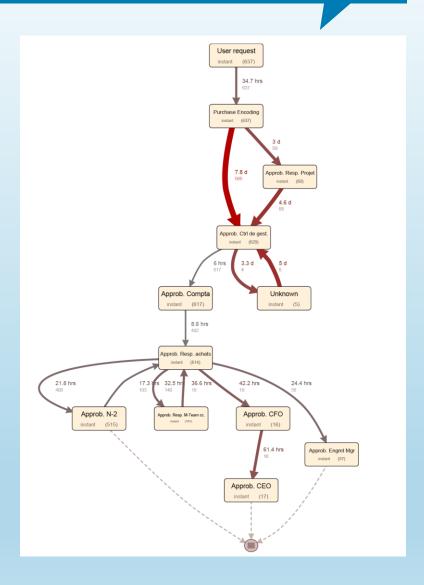


### **Process mining**

> Based on purchase requests logs

DA 🔻	Commande 🔻 Date	<b>→</b> 1 Stratégie	Ressource	e Activity
18002613	4570020797 19/12/2013 (	00:00	3 P06333	User Request
18002618	4570019402 20/12/2013 (	00:00	1 P06677	User Request
18002620	4570019405 23/12/2013 (	00:00	1 P06069	User Request
18002621	4570019406 23/12/2013 (	00:00	1 P06069	User Request
18002637	4570019397 30/12/2013 (	00:00	1 P16002	User Request
18002640	4570019396 02/01/2014 0	00:00	1 P05086	User Request
18002640	4570019396 02/01/2014 0	00:00	1 Purchase	Purchase Encoding
18002637	4570019397 03/01/2014 0	00:00	1 Purchase	Purchase Encoding
18002618	4570019402 03/01/2014 0	00:00	1 Purchase	Purchase Encoding
18002620	4570019405 03/01/2014 0	00:00	1 Purchase	Purchase Encoding
18002621	4570019406 06/01/2014 0	00:00	1 Purchase	Purchase Encoding
18002646	4570019422 07/01/2014 0	00:00	1 P06677	User Request
18002647	4570019423 07/01/2014 0	00:00	1 P06677	User Request
18002645	4570019522 07/01/2014 0	00:00	1 P16029	User Request
18002637	4570019397 07/01/2014	10:26	1 P06252	Approb. Ctrl de gest.
18002618	4570019402 07/01/2014	10:27	1 P06252	Approb. Ctrl de gest.
18002640	4570019396 07/01/2014	10:31	1 P06252	Approb. Ctrl de gest.
18002620	4570019405 07/01/2014	10:32	1 P06252	Approb. Ctrl de gest.
18002621	4570019406 07/01/2014	10:32	1 P06252	Approb. Ctrl de gest.
18002637	4570019397 07/01/2014	11:11	1 USER_PO	Approb. Compta
18002618	4570019402 07/01/2014 :	11:12	1 USER_PO	Approb. Compta
18002640	4570019396 07/01/2014 :	11:14	1 USER_PO	Approb. Compta
18002620	4570019405 07/01/2014 :	11:14	1 USER_PO	Approb. Compta
18002621	4570019406 07/01/2014 :	11:14	1 USER_PO	Approb. Compta
18002637	4570019397 07/01/2014	13:06	1 USER_PO	Approb. Resp. achats
18002618	4570019402 07/01/2014	13:08	1 USER_PO	Approb. Resp. achats

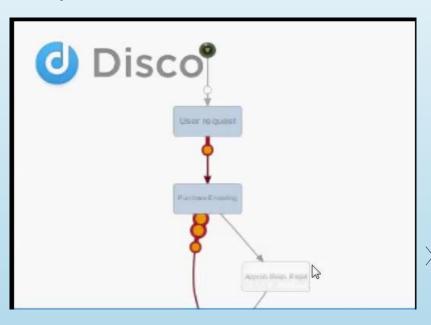
- Highlight steps and paths followed by the request in order to validate the conformity relative to the defined process
- Compute minimum, maximum and average time to process a step / activity

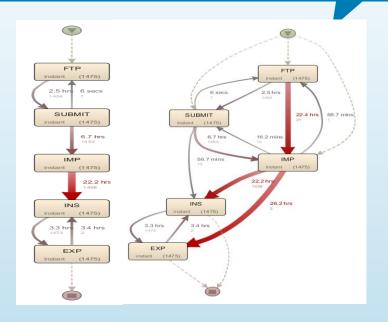




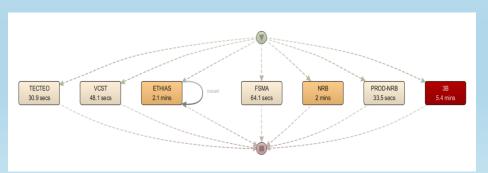
### **Process mining**

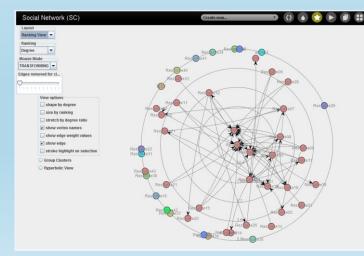
Allow to go deeper in details and show steps and paths little used and bottlenecks





> Show interactions between stakeholders







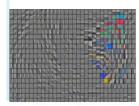
## Some of our competencies



# Stanford

Machine Learning

Learn about the most effective machine learning techniques, and gain practice implementing them and getting them to work for yourself.



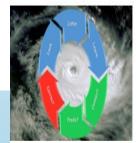


Neural Networks for Machine Learning

Learn about artificial neural networks and how they're being used for machine learning, as applied to speech and object recognition, image segmentation, modeling language and human



This course will use social network analys both its theory and computational tools, to make sense of the social and information networks that have been fueled and rendered accessible by the internet.





Web Intelligence and Big Data

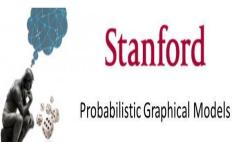
This course is about building 'web-intelligence' applications exploiting big data sources arising social media, mobile devices and sensors, using new big-data platforms based on the 'map-reduce' parallel programming



COLUMBIA UNIVERSITY

Big Data in Education

Education is increasingly occurring online or in educational software, resulting in an explosion of data that can be used to improve educational effectiveness and support basic research on learning. In this course, you will learn how and when to use key methods for educational data mining and learning analytic on this data.



In this class, you will learn learn algorithms for using a PGM to reach conclusions about the world from limited and noisy evidence, and for making good decisions under uncertainty.



# Risks & Statistics





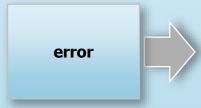
# Category of risks

**Sporadic** 

**Recurring** 

**Statistics** 

**Cognitive** 





# Category of risks

**Sporadic** 

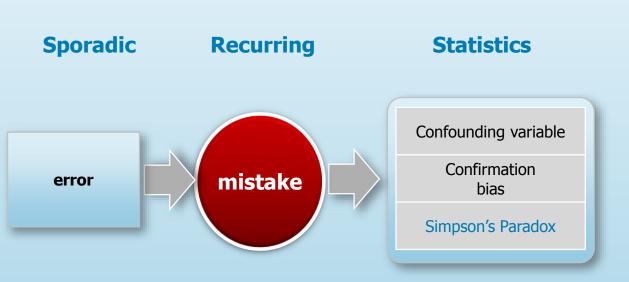
**Recurring** 

**Statistics** 

**Cognitive** 



# Risks





# Simpson's Paradox





# Simpson's Paradox

Homer







Same name, different first names, differents fates...



## Emergency Services: global indicator of the performances

Statistical indicators: Survival rate / Period: 10 days

Hospital	Total	Survivors	Deaths	Survival Rate
Hospital A	1000	800	200	80 %
Hospital B	1000	900	100	90 %

**Additional Information** 

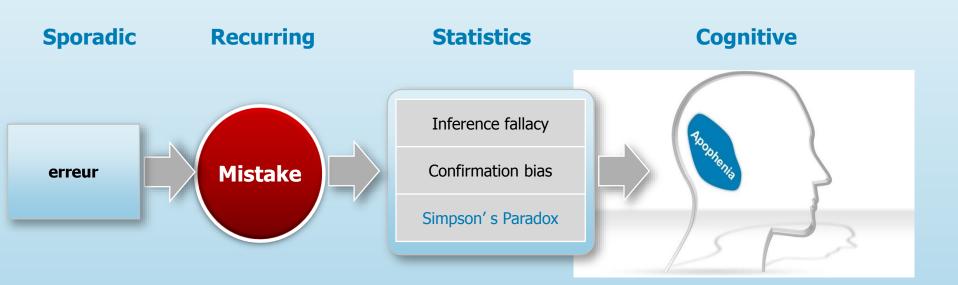
**Identical criteria for:** 

**Hospital A** 

**Hospital B** 



# Apophenia



## listen to your intuition

Theater ticket + parking ticket = 1,1€

The Theater ticket costs one euro more than the parking ticket

How much does the ? Parking ticket cost?

The Parking ticket

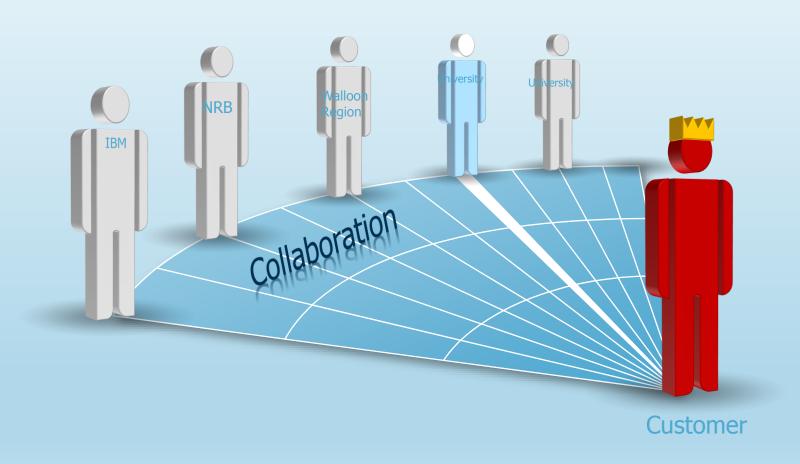


# Conclusions





## Principle of creative cooperation in developing solutions





### Discover

#### Discover

#### According to Business Maturity

#### **Business Understanding**

#### **Strategy & Positionning**

Global Strategy

Positionning of Products/Services

Scorecards – KPI's

Digital Strategy

Touchpoints for

Customers/Prospects

Search, ecommerce, Social Media,

Websites, ...

**Direct Competitors** 

Disruptive Elements

Disruptive Models

Disruptive Technologies

Disruptive Behaviors

#### Determine Potential Use Case(s)

Objectives

Priorities

Costs, ROI, constraints, Value to

The Client

Scope

New Data Sources

#### Determine High Level Feasibility

Data Mining Goals & Success Criteria Resources Availability

nesources Availability

Time to Deliver

#### Use Case Understanding

#### **Assess Business Maturity**

#### Determine

**Use Case Objectives** 

Value to the Client

Business Success Criteria

#### **Assess Situation**

Inventory of Resources Requirements,

Assumptions, and

Constraints

Risk and Contingencies Terminology

Costs and Benefits

#### **Refine Data Mining Goals**

Data Mining Goals Data Mining Success Criteria

#### Produce Project Plan

Approach

Deliverables

Schedule

(Data) Risk Mitigation (Privacy, etc.)

Stakeholders to Involve Initial Assessment of

Tools and Techniques

#### **Business Understanding**

1 to 5 days Workshop(s) 2 to 10 days Analysis & Consolidation

1 Senior Consultant 1 Business Scenario Modeller

#### Use Case Understanding

1 to 2 days Workshop(s) 2 to 5 days Analysis & Consolidation

1 Senior Consultant 1 Business Scenario Modeller