Data Strategy

Data Analytics for Smart Agriculture

Filippo Renga

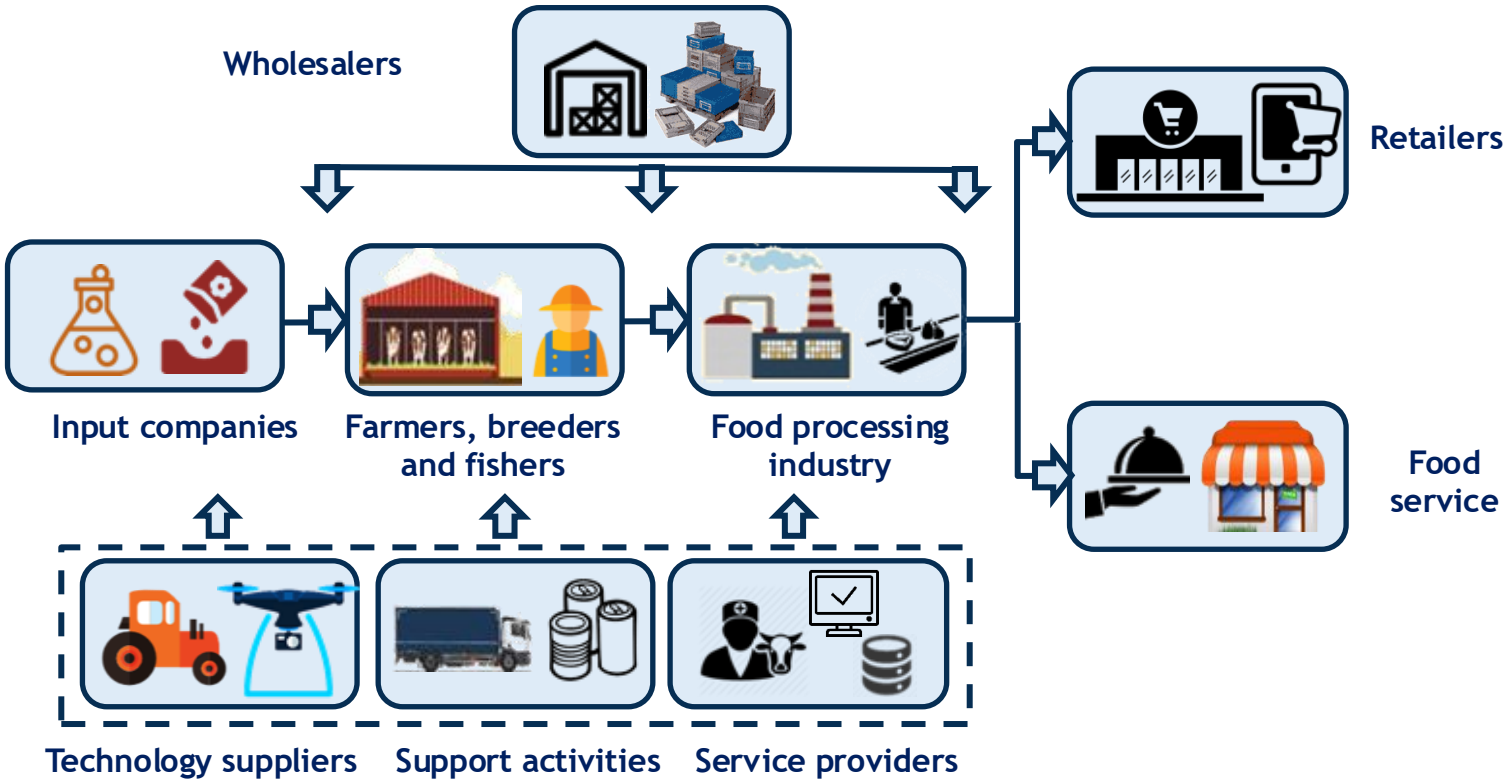
The value of Big Data Analytics

Data is the new oil. It's valuable, but if unrefined it cannot really be used. It has to be changed into gas, plastic, chemicals, etc to create a valuable entity that drives profitable activity; so must data be broken down, analyzed for it to have value.
Clive Humby, UK Mathemetician and architect of Tesco's Clubcard, 2006.

Personal data is the new oil of the internet and the new currency of the digital world.
Meglena Kuneva, European Consumer Commissioner, 2009


The Economist
Theresa May v Brussels
Ten years on: banking after the crisis
South Korea's unfinished revolution
Biology, but without the cells
JULY 6TH-12TH 2017
The world's most valuable resource
amazon
Google
Data and the new rules of competition

Actors in the agrifood supply chain



Digital Innovation in AgriFood

Farming

- Field sensors
- Sensors on machinery
- Drones
- Decision Support Systems (SSD)
- …

Transformation

- Food quality sensors
- Software ERP
- Traceability platforms
- …

Distribution

- Logistic optimization
- eCommerce
- Traceability with Blockchain
- …

Consumer

- Smart labels
- Mobile Apps
- …

101010101010101 101010101010101 101010101010101 101010101

...generating much data!

FARMING STAGE

FIELD DATA:
- Presence of infesting plants, bugs and fungal diseases
- Electrical conductivity
- Physical/biological/chemical fertility
- Various inputs costs (seeds, fertlizers, insecticide, ...)
- Type of crop
- Assessment of microbiological activity
- Supply of mineral elements
- Adaptability of plants
- Assessment of mineral fraction
- Total / Active limestone
- Measurement of nitrogen
- Vigor maps
- Prescription maps
- Irrigation
- ...

MACHINERY DATA:
- Use of machinery
- Position of the machine
- Operational data (e.g. diesel consumption)
- Machine diagnostic data
- Hours of work
- ...

EQUIPMENT DATA:
- Use of seeds, fertilizers, herbicides and water
- Working conditions

WEATHER DATA:
- Precise weather forecast
- Wind direction/strength
- Moisture (soil and air)
- Rainfall
- Temperature
- Photosynthetic efficiency index
- Solar radiation

PRODUCTION
- Origin of raw material
- Quality tests
- Operational data
- Traceability
- Certifications
- ...

LOGISTICS
- Transportation
- Transport conditions
- Routes
- Warehouse data
- ...

OPERATIONAL DATA:
- Internal / Esternal staff data
- Staff costs
- Resources / Equipment used

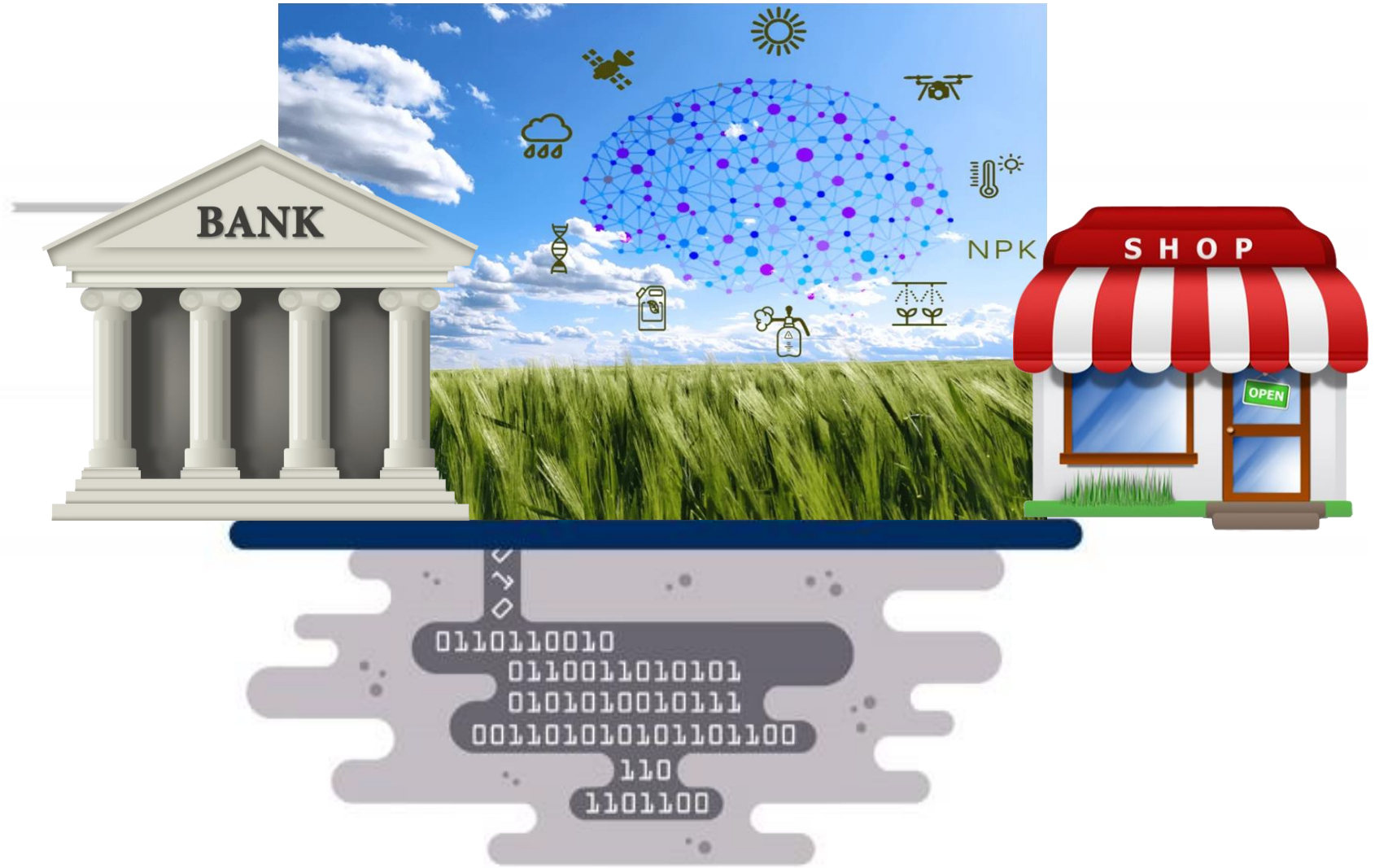
HARVEST DATA:
- Quantitative Results
- Quality
- Harvest conditions

WAREHOUSE DATA:
- Warehouse environmental data
- Product conditions
- Quantity of product in warehouse

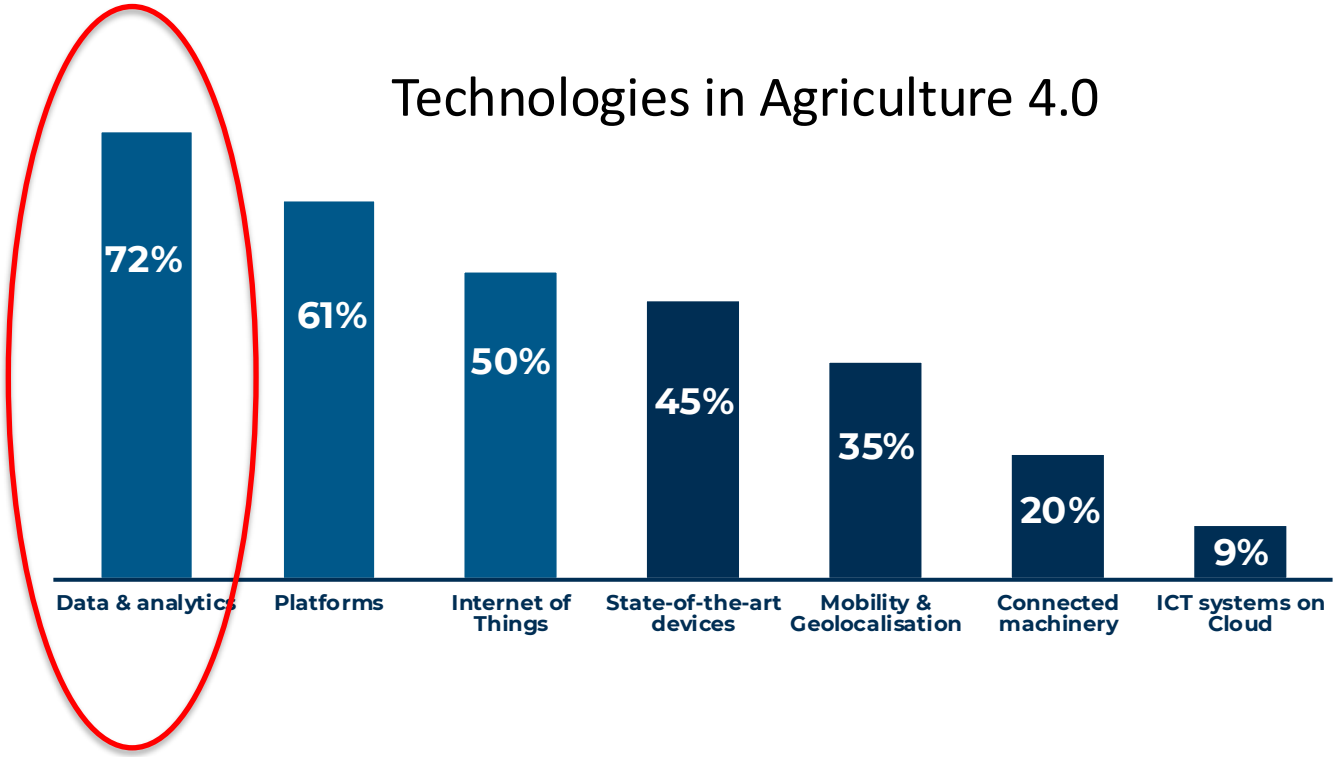
....

....

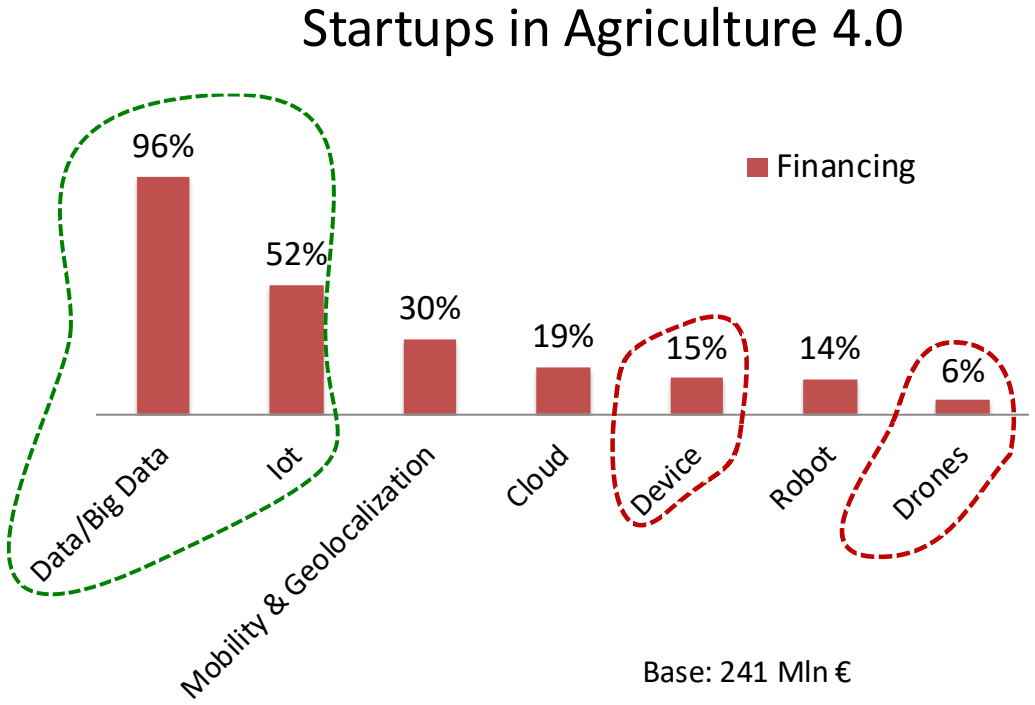
The value of Big Data Analytics



From data to information



Base: 223 Sample: 415 Agriculture 4.0 solutions in Italy
(a solution might be based on more than one technology.)



Base: 241 Mln €

Data valorization life cycle

RESEARCH AND MINING

REFINING

USE

Is it better to start with:

- *research and mining*

or

- *use cases*

?

Areas of application and types of data

RESEARCH AND MINING

REFINING

USE

Data valorization life cycle

RESEARCH AND MINING

REFINING

USE

Models to enhance data

NEW

Sale of data:

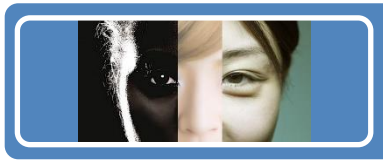
data = €

Integration of value chain:
data = new product / service

OLD

Data exchange (e.g. between agriculture and transformation):

data = opportunity of win-win business



Product / Service innovation:
data = value added to product / service

Integrated platforms allowing the interaction and the exchange of information among user groups

Data valorization life cycle

RESEARCH AND MINING

REFINING

USE

Use of data

BETTER FOOD QUALITY

30%

Food quality and traceability

30% of international case studies implemented systems based on Big Data analysis to improve the traceability and the quality of the final product

Base: 57 cases

DATA/BIG DATA PROCESSING

65%

23%

20%

Real time processing

Simulations and predictive analyses

Batch processing

Base: Mln. 241 $

Big Data areas of application

RISK MANAGEMENT

- Operational risk
- Market risk
- Credit risk
- Liquidity risk
- Reputational risk

OPERATIONS OPTIMIZATION

- Predictive maintenance
- Network infrastructure monitoring
- Performance evaluation and improvement
- Search function optimization

MARKETING AND PRODUCT DESIGN

- Churn rate prediction
- Pricing optimization
- Customer segmentation
- Cross/Up-Selling
- Demand prediction
- Marketing campaigns optimization
- Robo advisory
- PFM

FRAUD AND COMPLIANCE

- Regulatory requirements management
- Internal fraud prevention
- External fraud prevention
- …

CUSTOMER RELATIONSHIP

- After-sale monitoring
- Behaviour (in-store and online) mapping and optimization
- Customer care services optimization
- Customer interaction automation

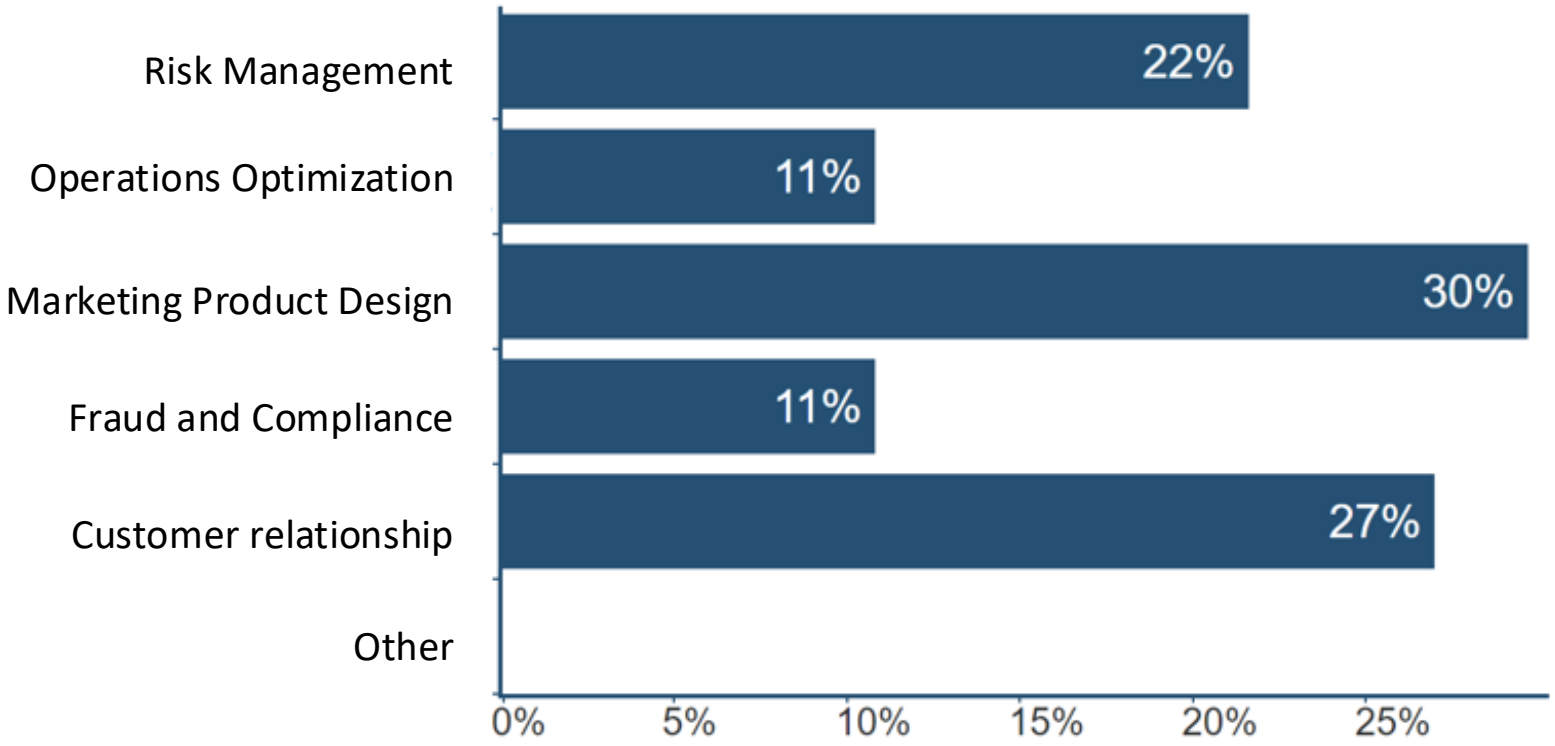
OTHER

- Market analysis for Asset Management and Trading
- HR – Talent Management and Talent Acquisition
- Competitive Intelligence

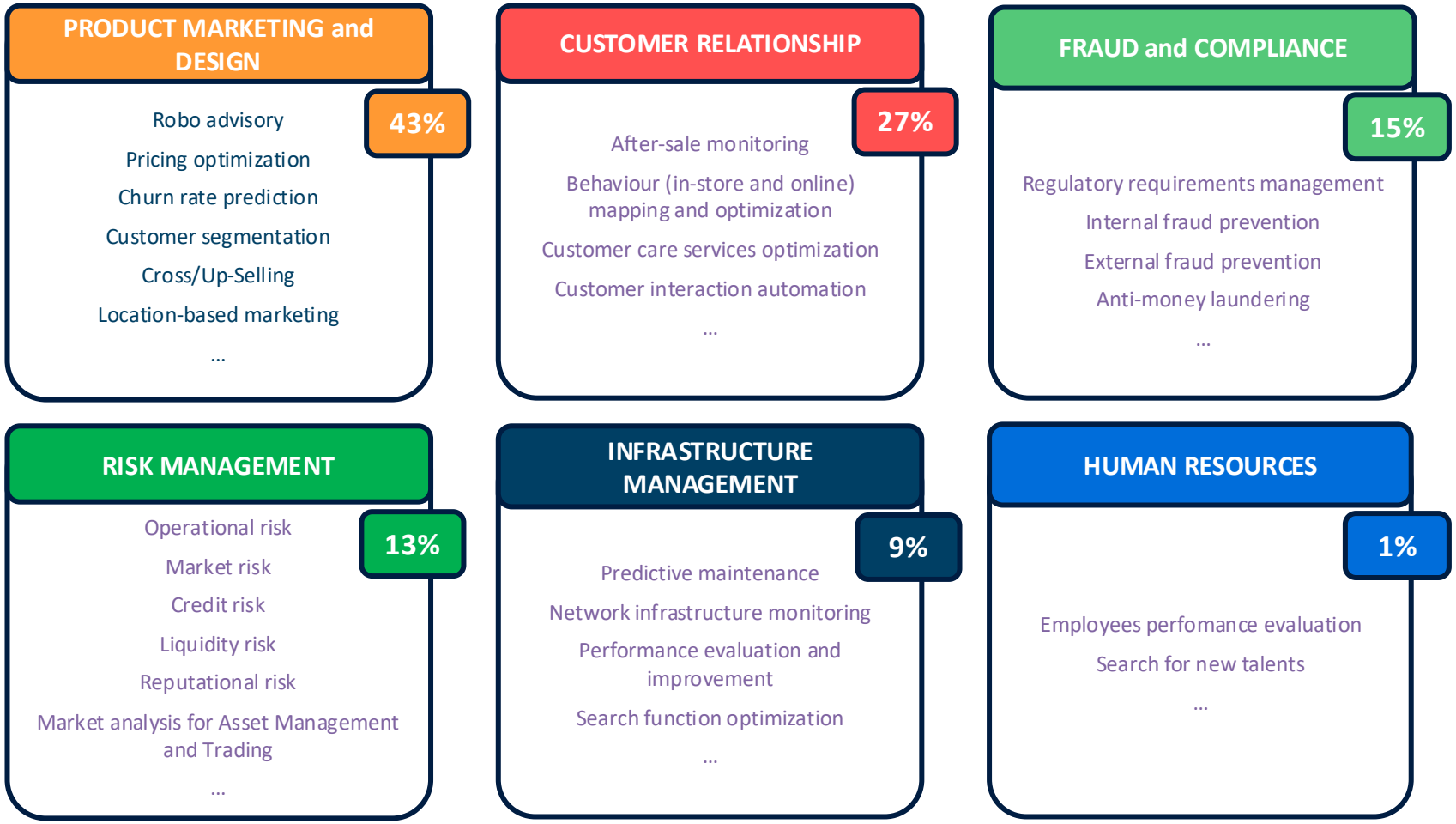
2. *Which are the priority areas of application?*
(Max 2 answers)

a) Risk Management
b) Operations optimization
c) Marketing Product Design
d) Fraud and Compliance
e) Customer relationship
f) Other

2. *Which are the priority areas of application?*
(Max 2 answers)



Areas of application and types of data



NOTE: the total amount exceeds 100% since some projects refer to more than one application area

Base: 128 Big Data Analytics international projects, January 2017

Data life cycle

RESEARCH AND MINING

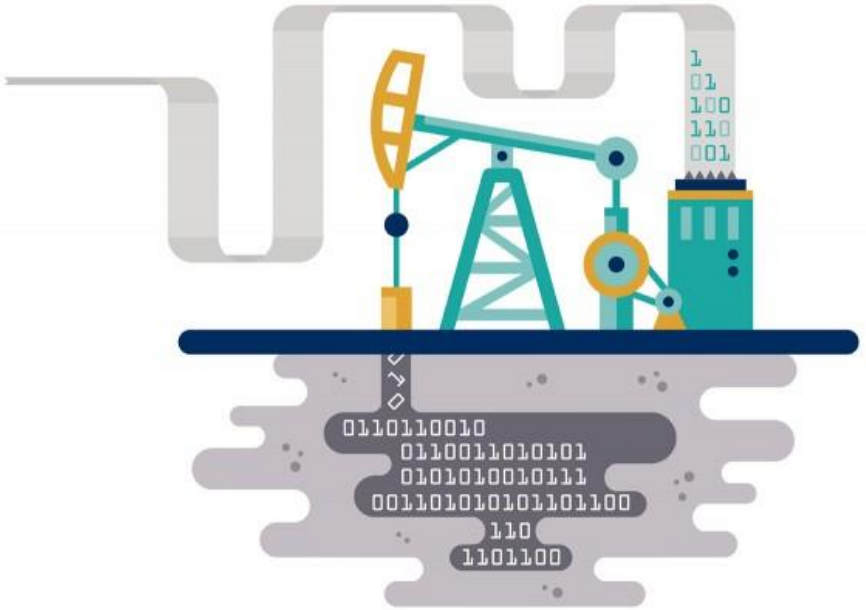
REFINING

USE

Where do we look for data?

Inside

Outside

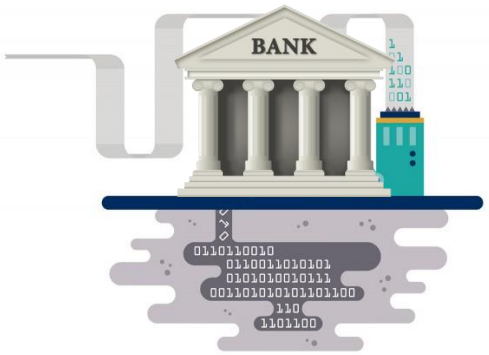


The origin of data

Internal Data:

- Data **collected** directly from the **client**
  - Personal data
  - Contact information
  - …
- Data **generated** by the **client**'s use of financial products
  - Withdrawals and deposits on the bank account
  - Credit card purchases
  - …
- Data originating from the **interaction between the client and** the bank's engagement **channels**
  - Data from web/mobile logs or online browsing
  - Customer assistance records
  - …
- Data generated internally by the **bank operators**
  - Rating (where calculated internally by the bank)
  - Report of company visits
  - …
- Data **automatically generated** by proprietary systems of the bank
  - Data from user logs
  - Data from sensors (ATMs, cameras, etc.)
  - …
- Data coming from other companies belonging to the same group
- …

The origin of



External Data:

For free ←
→ Payment

Social media

International bodies

Land Registry

Chamber of Commerce data

Central banks

News

Rating agencies

Istat

Open Data

Coalition

Reports

Central credit registers

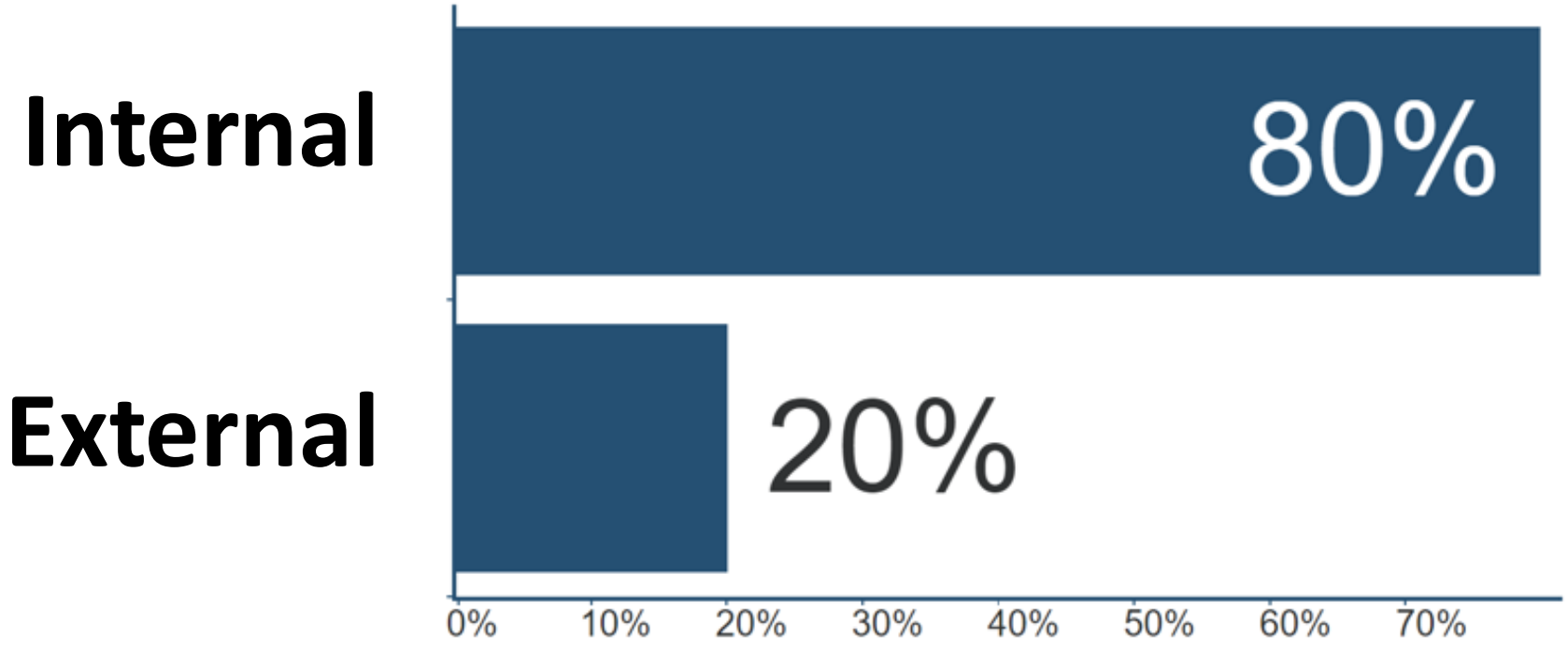
3. *Will Internal or External data generate more value?*

(Max 1 answer)

a) Internal
b) External

3. *Will Internal or External data generate more value?*

(Max 1 answer)



Areas of application and types of data

Data without any scheme that cannot be adapted to a relational database or for which there irregular or structure, bu sufficient to allow their storage in a structured mann

Image of a leaf

Video of the movement of a cow

UNSTRUCTURED DATA

STRUCTURED DATA

Temperature

GPS coordinates

Level of soil moisture

ning mainly from actional high- g systems, for anagement of ormation organised ling to predefined e schemes; these ata are easy to manipulate since they are stored in a structured manner

Data analysis speed

Batch

Near real-time

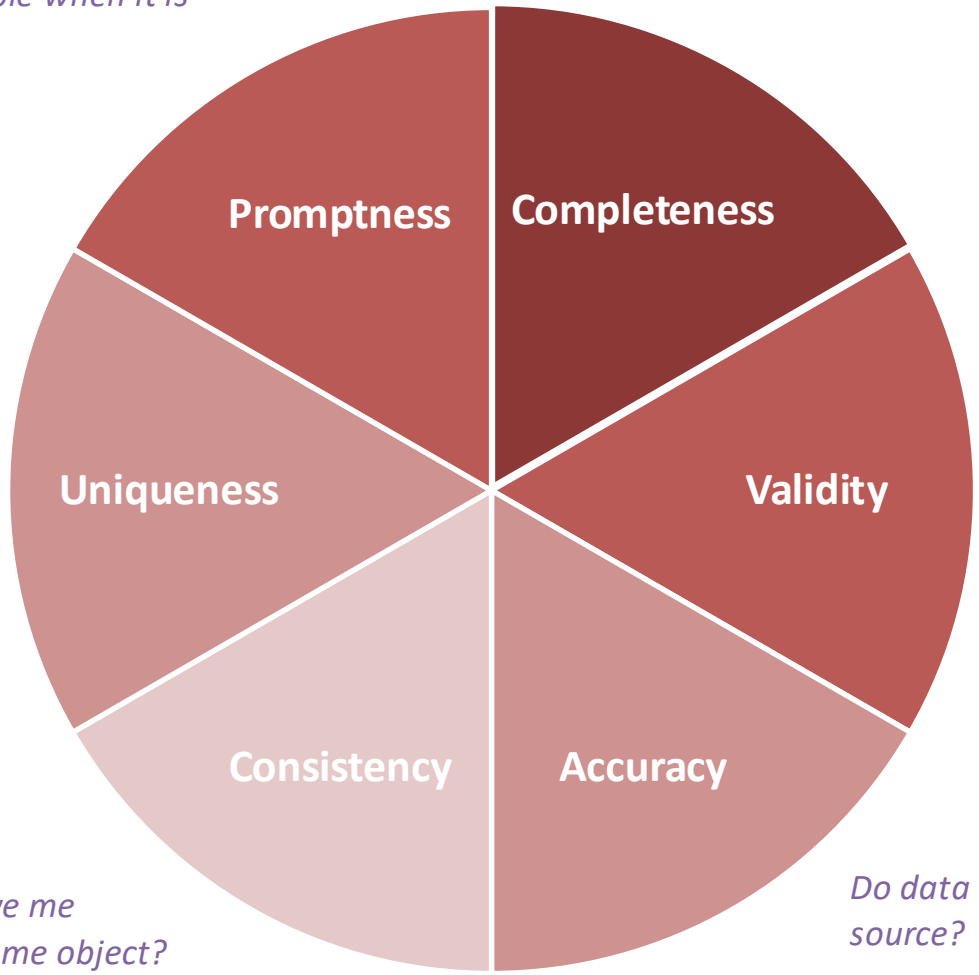
Real-time

Areas of application and types of data

Data quality

Are data available when it is needed?

Are there all the necessary data?



Is there only one version of the dataset?

Are all data within the specified domain?

Can different data give me information on the same object?

Do data reflect a verifiable source?

Source: The six primary dimensions for data quality, DAMA UK, 2013

Useful data vs useless data

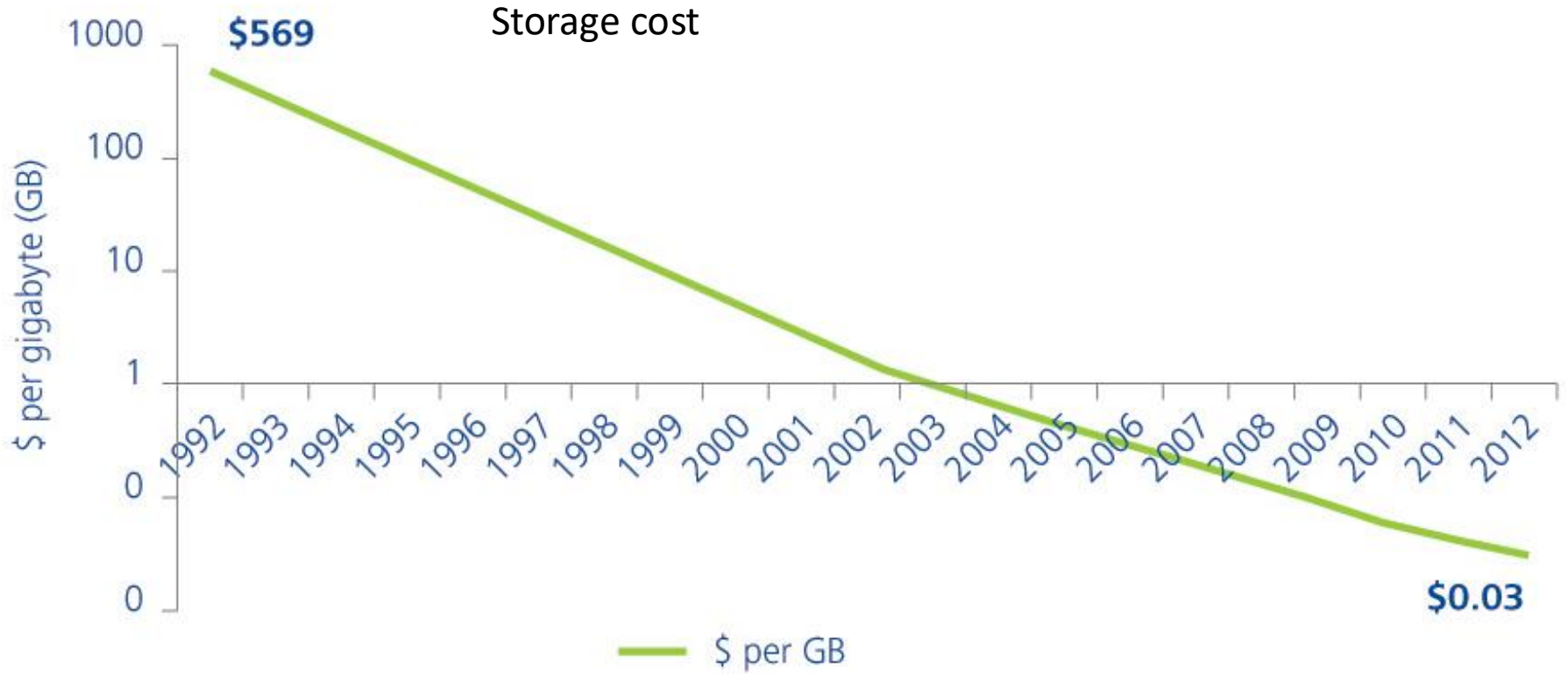
Should I collect also those data that appear to be useless at first?

How much does it cost to collect also useless data?

How can I recognize useful data?

Do useless data exist?

Data Storage cost

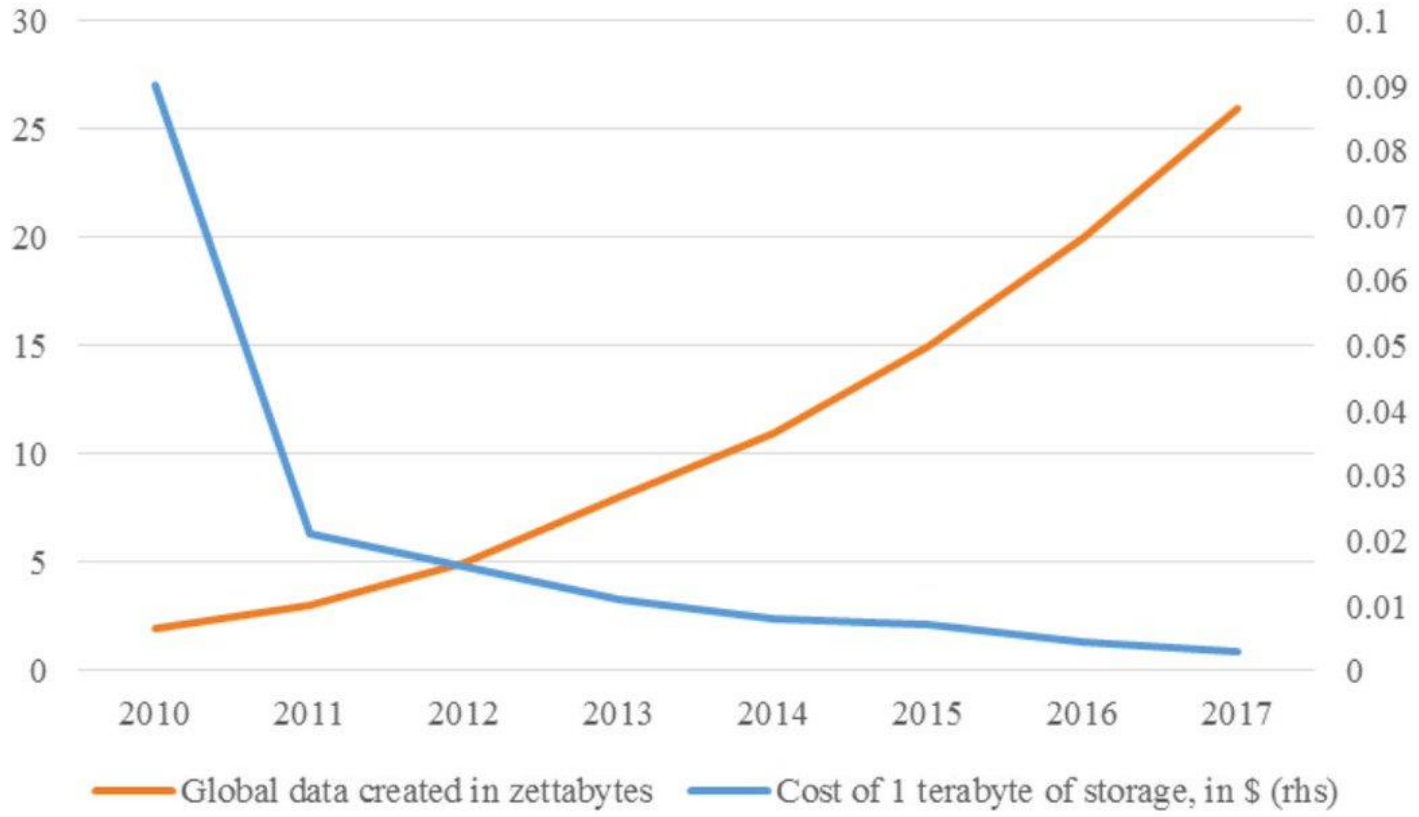


Source: Leading technology research vendor

Graphic: Deloitte University Press | DUPress.com

Data Storage cost

Figure 3: Costs of storage and global data availability, 2009-2017

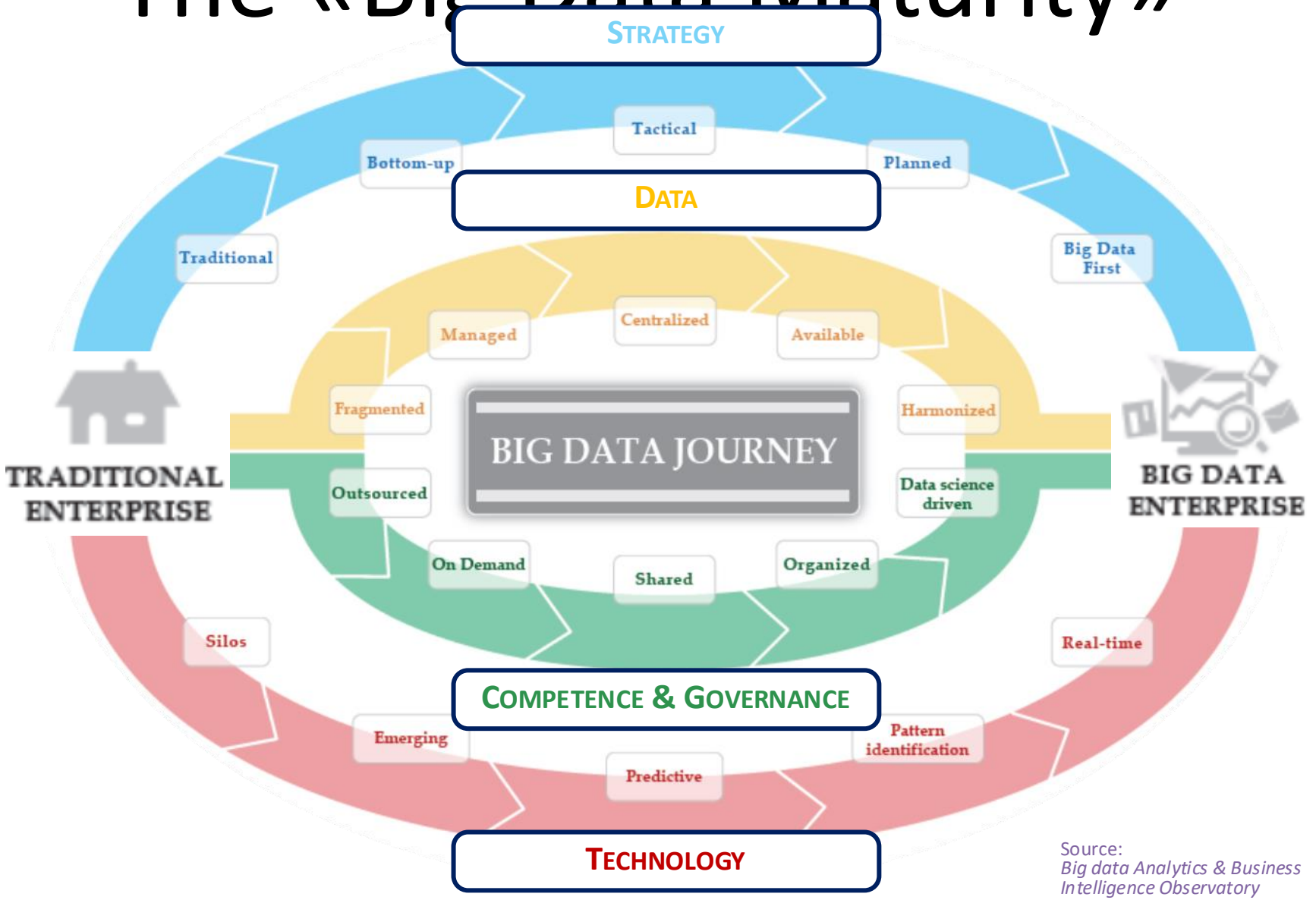


Source: Reinsel, Gantz and Rydning (2017); Klein (2017). One zettabyte is equal to one billion terabytes.

source fsb.org via @mikequindazzi

The «Big Data Journey»: the growth path

The «Big Data Maturity»



Source:
Big data Analytics & Business Intelligence Observatory

Big Data and market scenario

The «Big Data Maturity»

1

Traditional

The opportunities offered by Big Data Analytics are not yet fully understood and there isn't a defined and shared plan

2

Bottom-up

The approach to the management of the company knowledge is traditional, but there are virtuous approaches led by the vertical needs of some company functions

3

Tactical

The opportunities of Big Data Analytics are known on a Top Management level and there is an annual plan defining guidelines and priorities according to mainly temporary needs

4

Planned

There is a multiannual plan for Big Data management with a specific budget. There is an evolution roadmap in terms of technological investments and priority projects

5

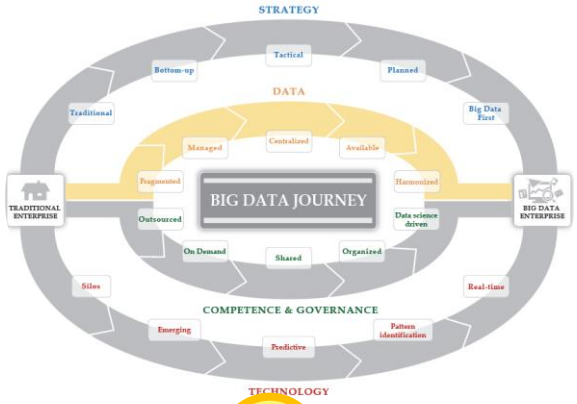
Big Data First

Big Data are part of the company's overall strategy, there is a multiannual plan and they are considered as a source of competitive advantage

Source: Big data Analytics & Business Intelligence Observatory

Big Data and market scenario

The «Big Data Maturity»

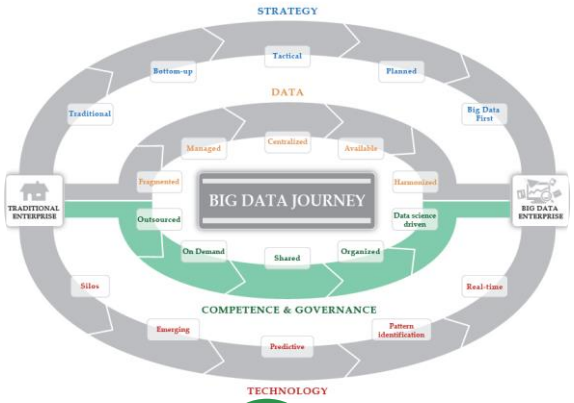


1	2	3	4	5
Fragmented	**Managed**	**Centralised**	**Available**	**Harmonized**
Data present in the company are mainly of poor quality and often provide contradictory and inconsistent information	Transactional data are available in a centralised repository. There are some policies and guidelines to manage the quality of information, the heterogeneous and non-standard degree of reliability	Data are centralised and available for the applications requiring their use. Their quality is medium-high	The reference taxonomy for data is defined in a centralised way. Data are identified uniquely within the company and they are accessible from a single point with different reading and analysis	Updated data are always available and uniquely identified for the business needs. Also unstructured data are collected

Source: Big data Analytics & Business Intelligence Observatory

Big Data and market scenario

The «Big Data Maturity»



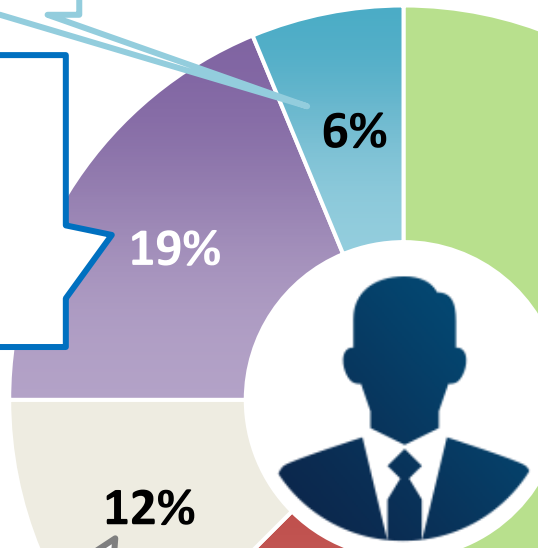
1	2	3	4	5
Outsourced	**On Demand**	**Shared**	**Organized**	**Data science**
The necessary skills to manage Big Data are largely not recognised and they are mainly present in IT Direction or they are mandated to external companies according to the temporary needs	**Skills are limited and focused on IT or in some LOB. Each business unit interacts for temporary needs**	There is a coordination plan among business units aimed at identifying needs although there isn't a structured skills development process	There is a plan aimed at creating and enhancing Big Data management skills, which include business growth paths for specific roles	There are codified figures dedicated to Big Data management and cross-functional teams with data science skills that are involved in the different business projects

Source: Big data Analytics & Business Intelligence Observatory

Roles associated to Big Data

Analysis for 59 international Banks and Financial Institutions on *93 news dealing with Big Data in 2016*

- Base: 32 professionals mentioned -



2 CEOs

Other roles
- 2 Chief Security Officers
- 1 Chief Operating Officer
- 1 HR
- 1 Chief Digital transformation
- 1 Chief Innovation Lab

Data Processing and Management
- 6 Chief Data Officers
- 1 Chief Data Governance
- 1 Chief Digital Analytics
- 1 Chief Data Analytics
- 1 Chief Data Architect
- 1 Senior Data Architect
- 1 Data Platform Engineer
- 1 Chief Analytics and Business insights
- 1 VP Data Scientist
- 1 President Analytics and Data Technology

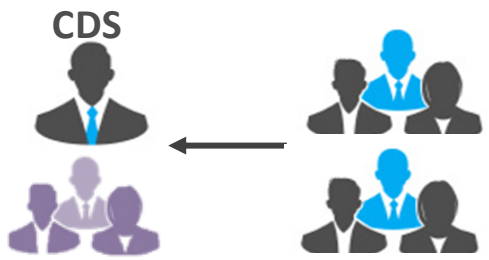
Information Systems
- 1 CIO
- 1 IT Director
- 1 Head of Information
- 1 Chief Information Developement

Marketing and Channels
- 1 Chief Strategic Media Planning
- 1 Chief Strategy e Retail Banking
- 1 Private Customer
- 1 Head of Customer Experience and Crm
- 1 Chief Multichannel Banking

Organisational approaches for Big Data management

Some possible organisational models

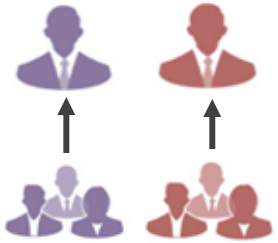
Centralised



- Widespread datascience mindset in the organisation
- Well-established management practice for datascience projects
- Possibility for datascientists to develop heterogeneous skills
- Efficient and reconfigurable structure

- Competition on sources low on data science
- Potentially critical data access
- Difficult involvement of the most traditional business units

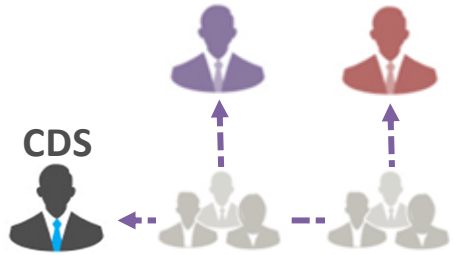
Business Driven



- Quick response time
- Knowledge of vertical business with possibility to develop new ides
- More developed data control by the business line

- Often non homogeneous and/or contrasting technological choices
- Potential partial/vertical view of data
- Datascientist with vertical view on business problems

Matrix

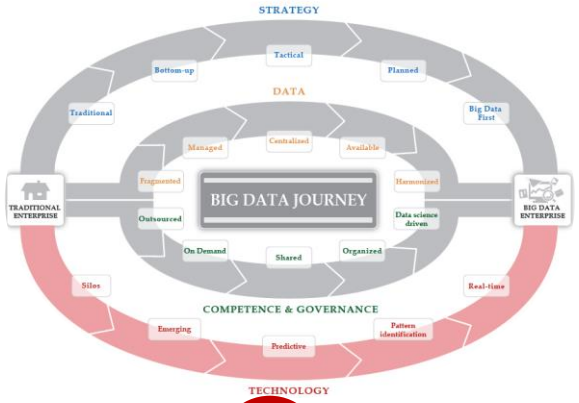


- Widespread knowledge of business problems by Data Scientists
- Harmonisation of technological choices and business analysis approaches
- Possibility of coordinated management of multifunctional projects

- Presence of double responsibility on sources
- Complexity of organisational introduction for datascientists
- Complex prioritization of projects

Big Data and market scenario

The «Big Data Maturity»



1

Silos

Data are mainly scattered in heterogeneous systems and not communicating. Technologies are outdated and sometimes they can't provide the level of detail or update required by the business

2

Emerging

There are traditional DWH systems which make the business structured data available in a cross way. Tool for data analysis is limited to some company sectors

3

Predictive

The tools present in the company allow the analysis of different data sources – mainly structured – and use technologies able to model possible future evolution scenarios in a predictive way

4

Pattern id

The technologies present in the company allow the process of large amounts of data, both structured and unstructured, in advanced data warehouses (e.g.: NoSQL Database, NewSQL Database, MPP Database, ...) and the identification of hidden patterns and correlations

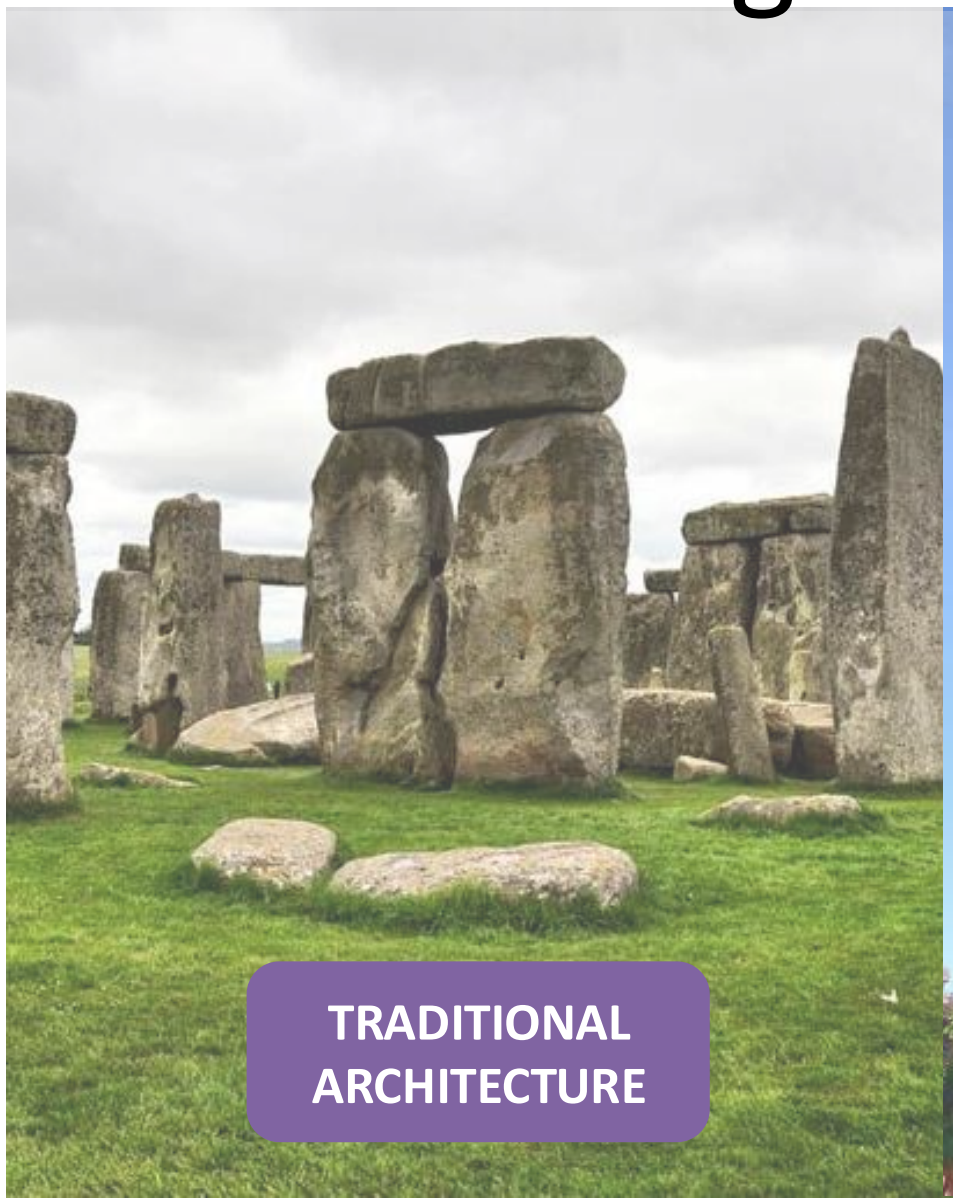
5

Real-time

The technologies present in the company can analyse large amounts of data and heterogeneous data sources. It is also possible to analyse real-time data and to provide dynamic views to line decision makers and top management

Source: Big data Analytics & Business Intelligence Observatory

Towards a Big Data architecture



References

EU Commission, 2020, «*A European strategy for data*», https://digital-strategy.ec.europa.eu/en/policies/strategy-data

EU Commission, 2022, «European Partnership under Horizon Europe Agriculture of Data», https://research-and-innovation.ec.europa.eu/document/download/a1fccc86-af53-43d4-94d2-79c54a353d0e_en?filename=ec_rtd_he-partnership-agriculture-data.pdf

Renga et al., 2022, «Valorization model for data and digital services», Ploutos-A Sustainable Innovation Framework to rebalance agri-food value chains, https://ploutos-h2020.eu/about/#1611047790490-33aa241c-29b3

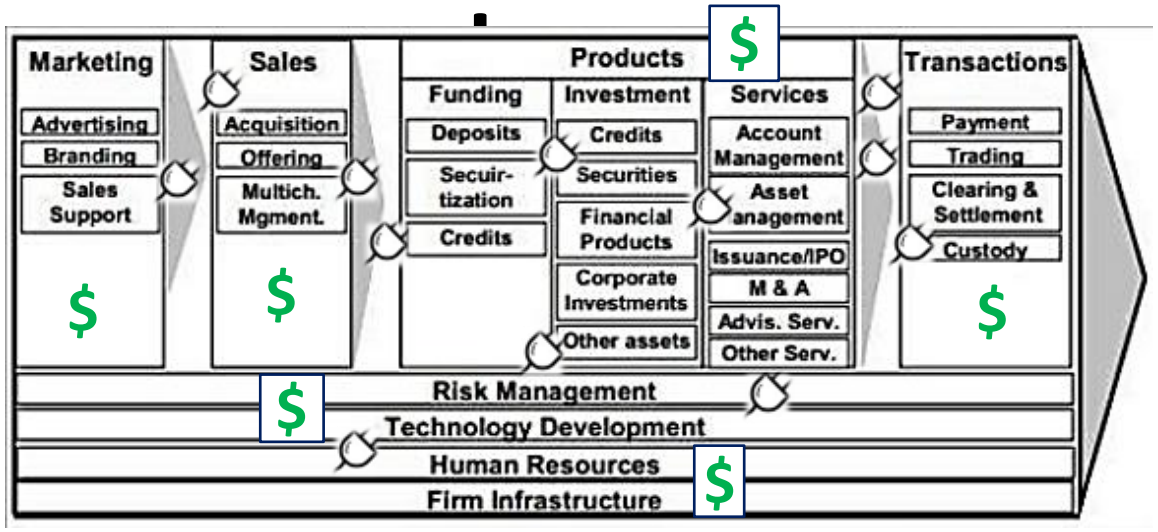
Douglas B. L., 2017, «Infonomics: How to Monetize, Manage, and Measure Information as an Asset for Competitive Advantage», Routledge

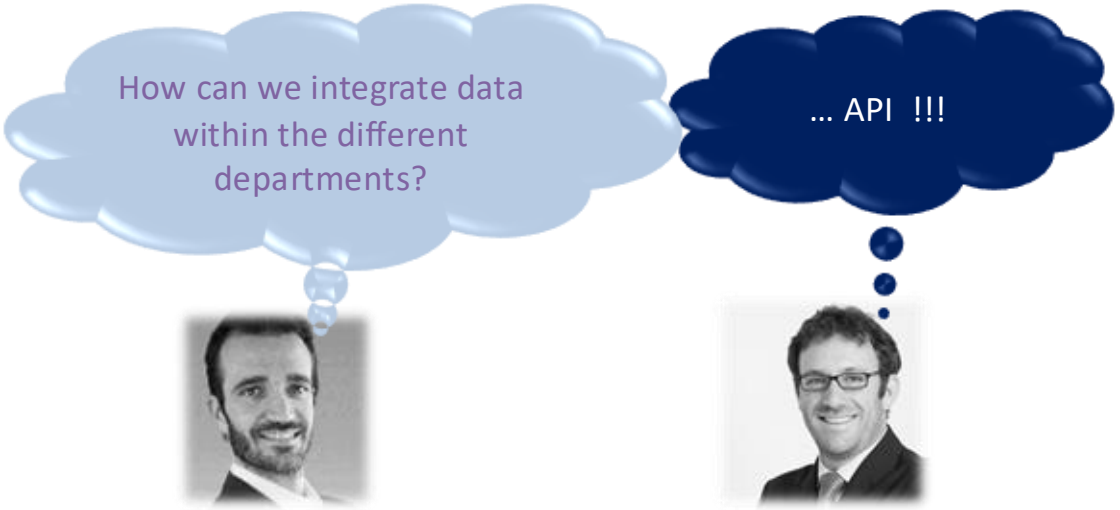
Data Management Association, 2017, «*Data Management Body of Knowledge*», Technics Publications

Wamba and et al, 2017, «Big data analytics and firm performance: Effects of dynamic capabilities», Journal of Business Research, vol. 70, pp. 356-365

Hubbard, D., 2010, «Measuring the Value of Information», in «How to Measure Anything: Finding the Value of “Intangibles“ in Business», John Wiley & Sons, Inc

All functions own relevant





Marco Bressan, Chief Data Scientist, BBVA

Application Programming Interfaces

APIs (Application Programming Interfaces) are **standards that allow software components to interact and exchange data**, particularly over the web. Put most simply, an **API is a set of instructions that allows one piece of software to interact with another**.

APIs allow different software applications to communicate with each other and exchange data directly, **without the need for human input each time**. They have become **the de facto standard for sharing data**, and have enabled organizations that hold large amounts of data to become platforms for third party innovation.

Application Programming Interfaces

Historical background

Salesforce.com releases its web-based APIs, allowing its clients to integrate the Salesforce services with the company core systems.

At the same time eBay launches the eBay Application Program Interface addressed to a small number of partners and developers

Google introduces the Google Maps APIs, allowing developers to integrate Google Maps in websites and third-party applications

Facebook releases the so-called Facebook Platform, which allows developers to develop third-party apps leveraging APIs

Netflix publicly releases its API, leading to the creation of applications and services developed by external developers

Credit Agricole releases CAStore, an online marketplace collecting ideas from the consumers for new banking applications, then allowing third-party developers to develop these apps thanks to public APIs

In 18 months the public APIs released are doubled, exceeding 10.000 units in various sectors (from telecommunications, media and finance to travel, tourism and real estate)

Application Programming Interfaces

An aspect to know and manage properly

__APIs__ are perhaps __the most critical technology__ in digital business design

Forrester Research, June 2015

In fact, a number of interviewees agreed that for most banks, __the process of choosing which technology to use, agreeing data and security standards, and getting legal sign off__ on the above would be significantly __more difficult and expensive than the actual tech implementation itself__

Open Data Institute e Fingleton Associates, September 2014