Relating Big Data Business and Technical Performance Indicators

Barbara Pernici, Chiara Francalanci, Angela Geronazzo, Lucia Polidori, Stefano Ray, Leonardo Riva,
Politecnico di Milano, Italy
Arne Jørgen Berre,
SINTEF, Norway
Todor Ivanov,
University of Frankfurt, Germany

itAIS Conference, PAvia, October 12, 2018

# Building a bridge between technical and business benchmarking

DataBench

Main Activities

- Classify the main use cases of BDT by industry

- Compile and assess technical benchmarks

- Perform economic and market analysis to assess industrial needs

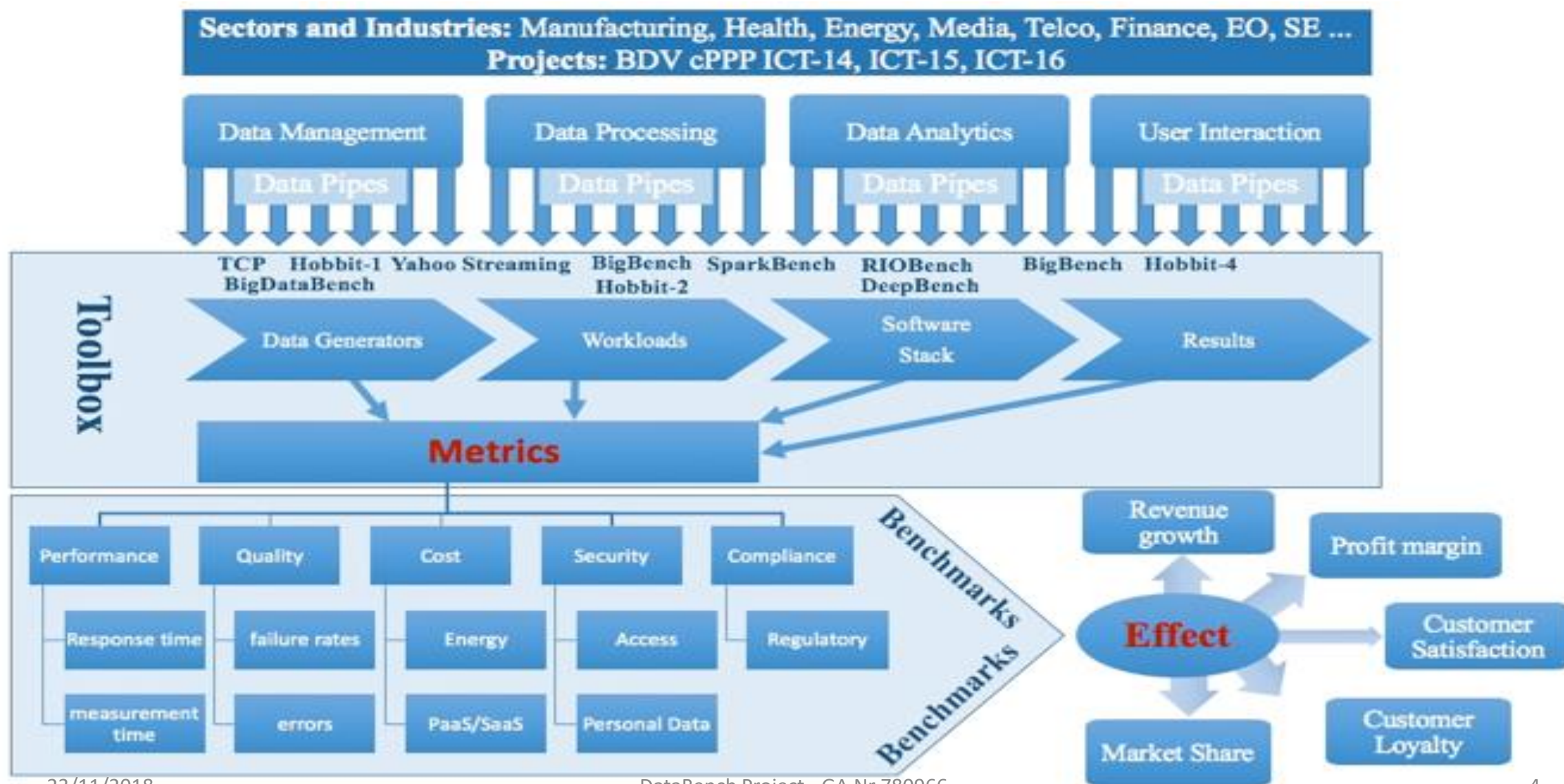- Evaluate business performance in selected use cases

Expected Results

- A conceptual framework linking technical and business benchmarks

- European industrial and performance benchmarks

- A toolbox measuring optimal benchmarking approaches

- A handbook to guide the use of benchmarks
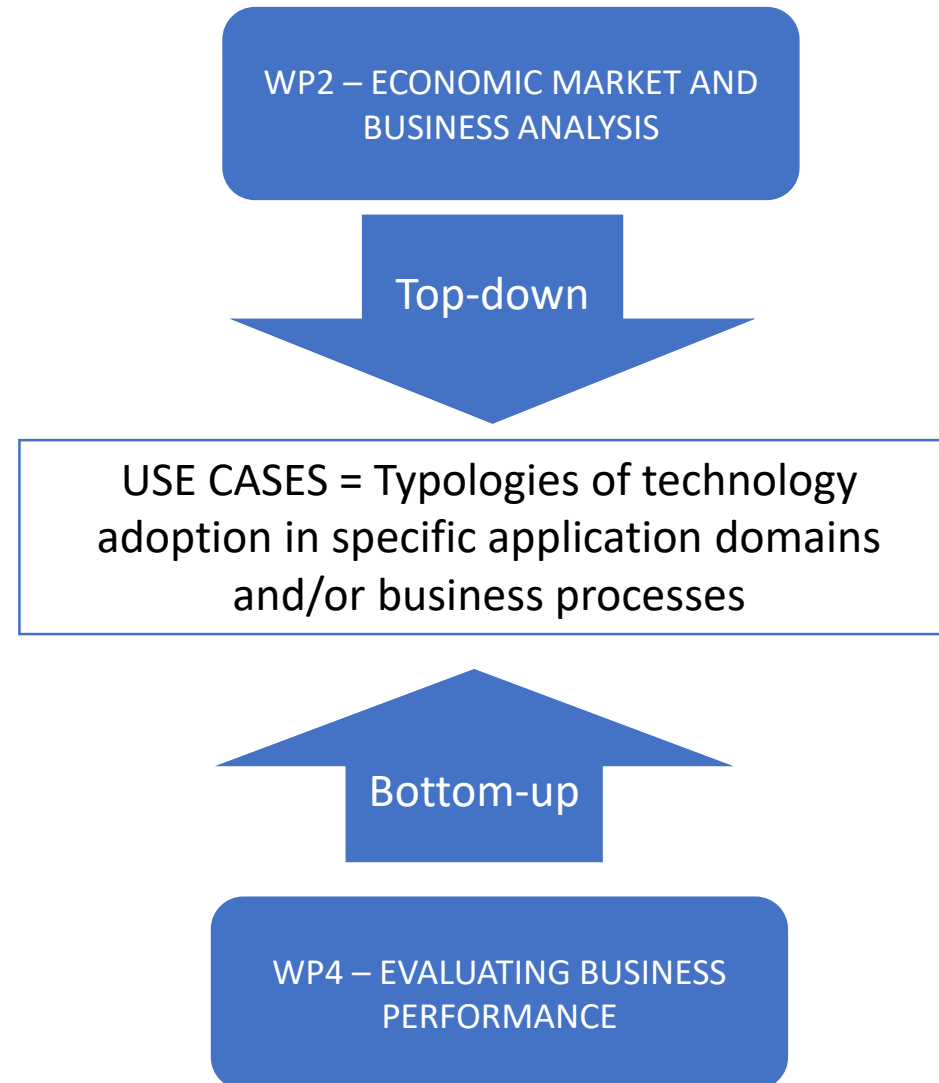
Databench Workflow

Where Magic Happens

DataBench Project - GA Nr 780966
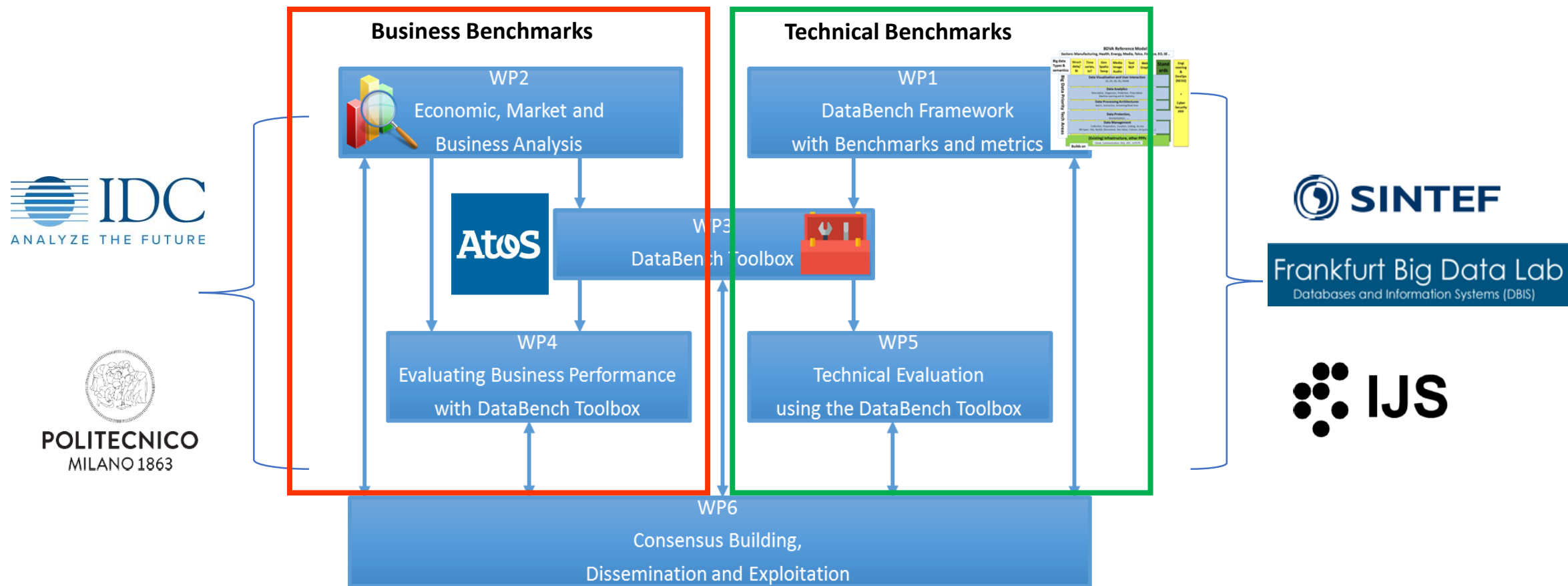
How to link technical and business benchmarking

- Focus on economic and industry analysis and the EU Big Data market

- Classify leading Big Data technologies use cases by industry

- Analyse industrial users benchmarking needs and assess their relative importance for EU economy and the main industries

- Demonstrate the scalability, European significance (high potential economic impact) and industrial relevance (responding to primary needs of users) of the benchmarks

WP2 – ECONOMIC MARKET AND BUSINESS ANALYSIS

Top-down

USE CASES = Typologies of technology adoption in specific application domains and/or business processes

Bottom-up

WP4 – EVALUATING BUSINESS PERFORMANCE

- Focus on data collection and identification of use cases to be monitored and measured
- Evaluation of business performance of specific Big Data initiatives
- Leverage Databench toolbox
- Provide the specific industrial benchmarks to WP"
- Produce the Databench Handbook, a manual supporting the application of the Databench toolbox

Results from the first year of the projects

- Modeling business indicators

- Modeling technical indicators

- Relating business and technical indicators
  - Two surveys
    - With BDVa Benchmarking group
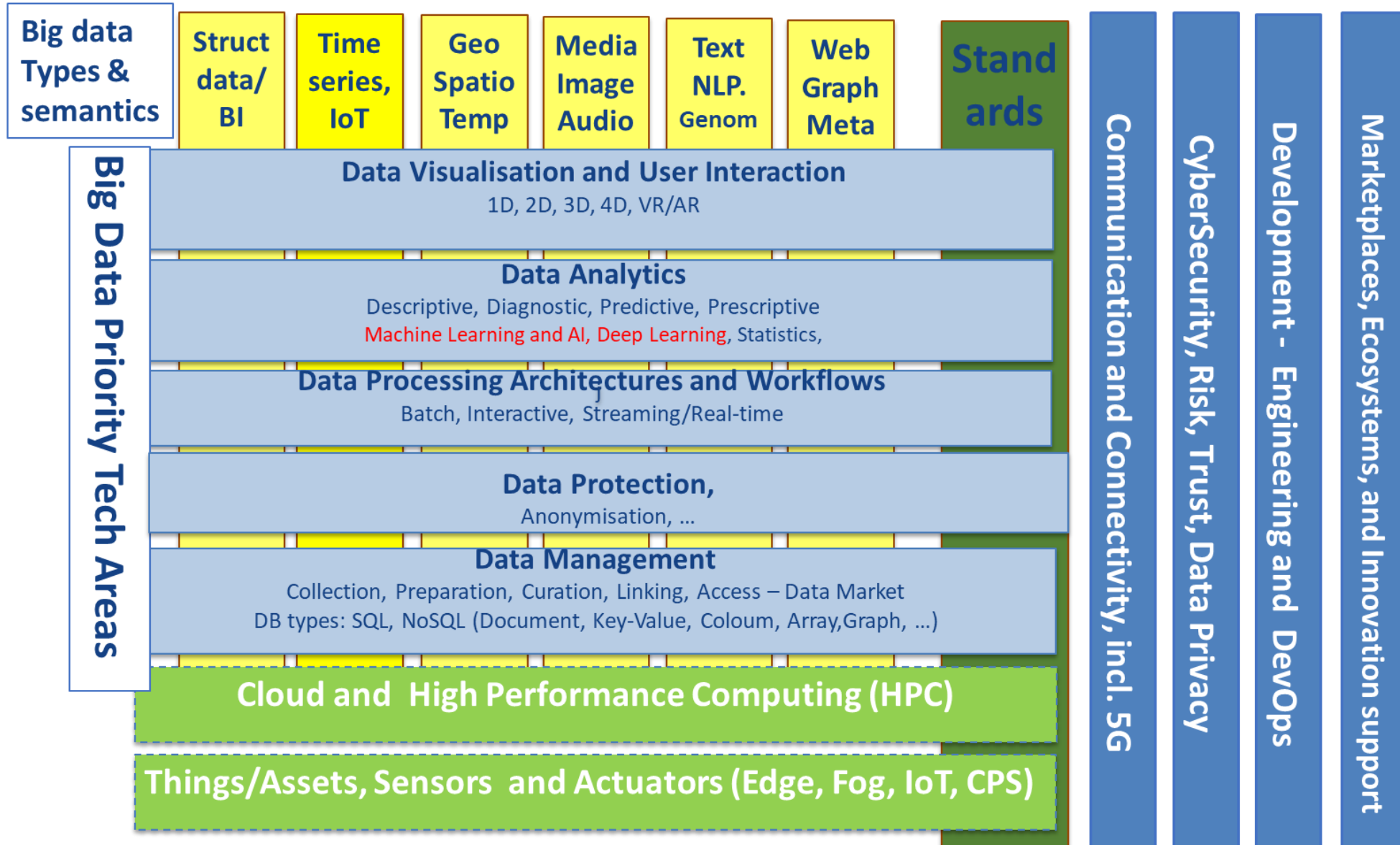   - Databench survey

Business indicators

Industry	Big Data Maturity	KPI	Scope of Big Data & Analytics	Data User	DB & Analytics Application	Size of Business	Data size	Datasource
Finance	Currently using	Cost reduction	Decision optimization task	Data Enterpreneurs	Sales	5000 or more	Gigabytes	Distributed
Manufacturing	Piloting or implementing	Time efficiency	Data driven business processes	Vendors in the ICT industry	Customer service & support	2500 to 4999	Terabytes	Centralized
Retail & Wholesale	Considering or evaluating for future use	Product/service quality	Data oriented digital transformation	User companies	IT & data operation	1000 to 2499	Petabytes	
Telecom/ Media	Not using and no plan to do so	Revenue growth			Governance risk & compliance	250 to 999	Exabytes	
Transport/ Accomodation		Customer satisfaction			Product management	50 to 249		
Utility/Oil&Gas/ Energy		Business model innovation			Marketing	10 to 49		
Professional services		Lauch of new products and/or services			Maintencance & logistics	less than 10		
Governamental/ Education					Product innovation			
Healthcare					HR & Legal			
					R&D			
					Finance			

BDVa (Big Data Value Association) Architecture

- Towards technical indicators
  - Tech areas
  - Types of data
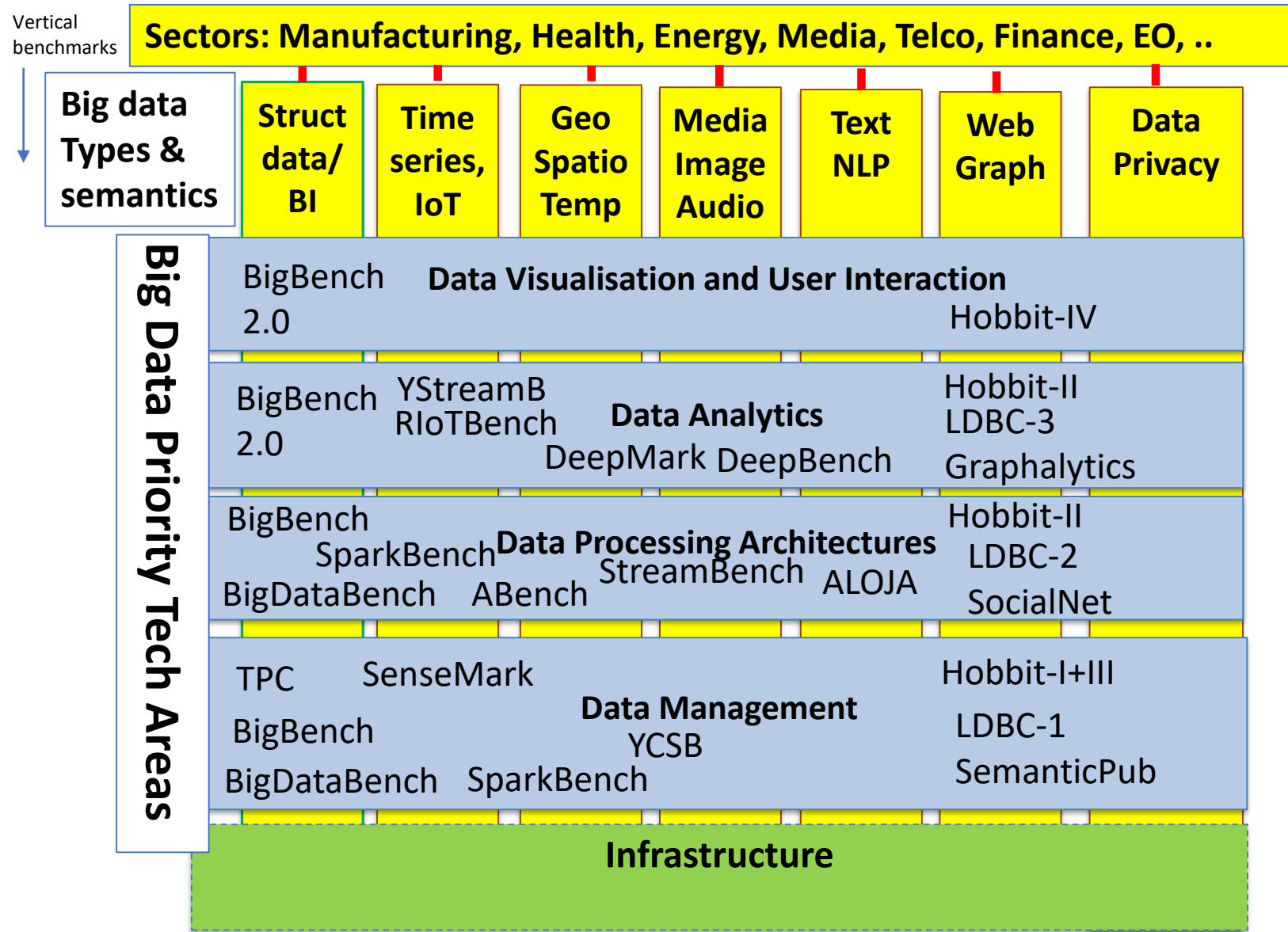
BDVA Extended (Digital Platform) Reference Model

Applications/Solutions: Manufacturing, Health, Energy, Transport, BioEco, Media, Telco, Finance, EO, SE

Big data Types & semantics

Struct data/ BI	Time series, IoT	Geo Spatio Temp	Media Image Audio	Text NLP. Genom	Web Graph Meta

Standards

Big Data Priority Tech Areas

Data Visualisation and User Interaction
1D, 2D, 3D, 4D, VR/AR

Data Analytics
Descriptive, Diagnostic, Predictive, Prescriptive
Machine Learning and AI, Deep Learning, Statistics,

Data Processing Architectures and Workflows
Batch, Interactive, Streaming/Real-time

Data Protection,
Anonymisation, …

Data Management
Collection, Preparation, Curation, Linking, Access – Data Market
DB types: SQL, NoSQL (Document, Key-Value, Coloum, Array,Graph, …)

Cloud and High Performance Computing (HPC)

Things/Assets, Sensors and Actuators (Edge, Fog, IoT, CPS)

Communication and Connectivity, incl. 5G

CyberSecurity, Risk, Trust, Data Privacy

Development - Engineering and DevOps

Marketplaces, Ecosystems, and Innovation support

Technical indicators

Metrics	Data Types	Benchmark Data Usage	Storage Type	Processing Type	Analytics Type	Architecture Patterns	Platform Features
Execution time/ Latency	Business Intelligence (Tables, Schema…)	Synthetic data	Distributed File System	Batch	Descriptive	Data Preparation	Fault-tolerance
Throughput	Graphs, Linked Data	Real data	Databases/ RDBMS	Stream	Diagnostic	Data Pipeline	Privacy
Cost	Time Series, IoT	Hybrid (mix of real and synthetic) data	NoSQL	Interactive/(near) Real-time	Predictive	Data Lake	Security
Energy consumption	Geospatial, Temporal		NewSQL/ In-Memory	Iterative/In-memory	Prescriptive	Data Warehouse	Governance
Accuracy	Text (incl. Natural Language text)		Time Series			Lambda Architecture	Data Quality
Precision	Media (Images, Audio and Video)					Kappa Architecture	Veracity
Availability						Unified Batch and Stream architecture	Variability
Durability							Data Management
CPU and Memory Utilization							Data Visualization

Vertical benchmarks

Big data Types & semantics

Sectors: Manufacturing, Health, Energy, Media, Telco, Finance, EO, ..

Struct data/ BI	Time series, IoT	Geo Spatio Temp	Media Image Audio	Text NLP	Web Graph	Data Privacy

Big Data Priority Tech Areas

BigBench 2.0   **Data Visualisation and User Interaction**   Hobbit-IV

BigBench 2.0   YStreamB RIoTBench   **Data Analytics**   DeepMark DeepBench   Hobbit-II LDBC-3 Graphalytics

BigBench   SparkBench   **Data Processing Architectures**   Hobbit-II LDBC-2 SocialNet
BigDataBench   ABench   StreamBench   ALOJA

TPC   SenseMark   **Data Management**   Hobbit-I+III
BigBench   YCSB   LDBC-1
BigDataBench   SparkBench   SemanticPub

Infrastructure

Horizontal benchmarks

Big Data Benchmarks related to the BDVA Big Data Reference model (ongoing work)

Early Results from the BDVa Questionnaire

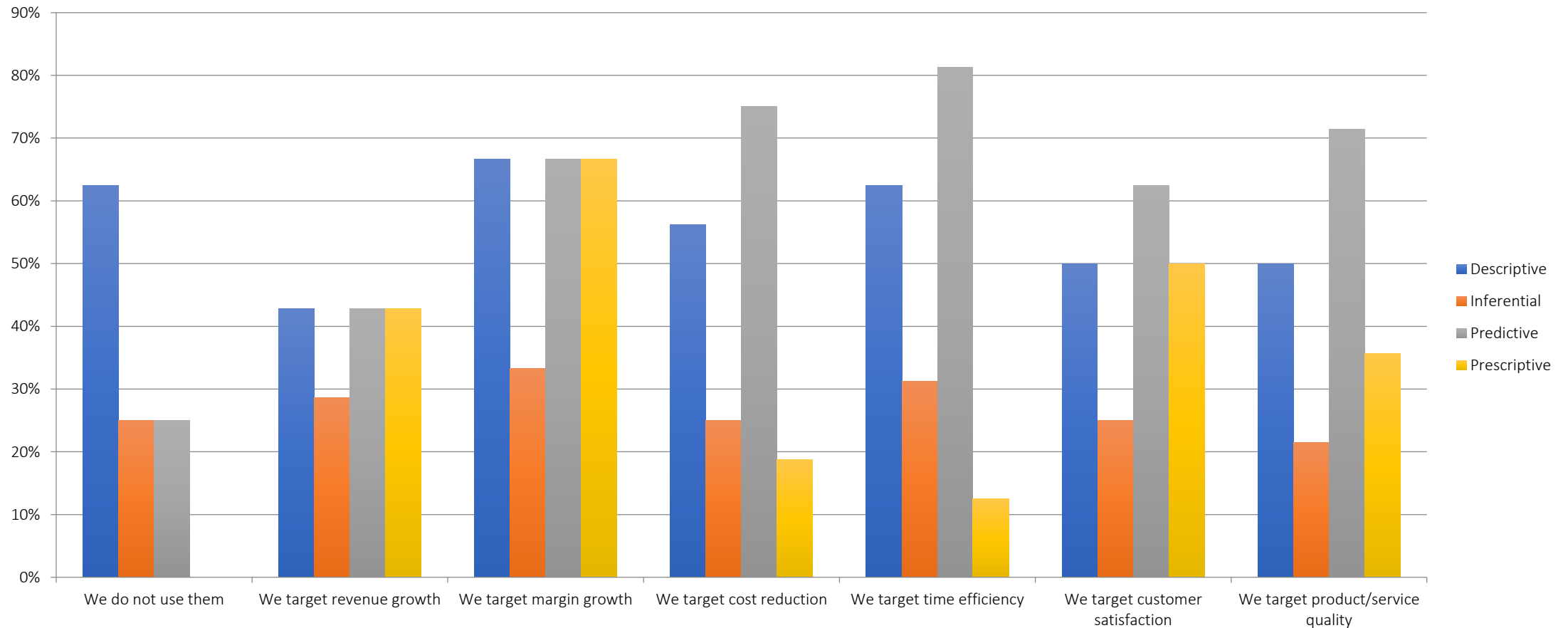# BDVa Benchmarking group (EU projects Hobbit and DataBench)

BDVa Benchmarking questionnaire

- 36 respondents
- 35 projects (mainly IC)

Relating indicators
An example: business KPI (x-axis) vs Analysis type (y-axis)

Analyzing requirements of industry

MANUFACTURING [7 respondents]

- [D10] What are your big data benchmarking goals/plans?
  - Check whether an implementation fulfils given business requirements and specifications. [86%]

- [D11] Which aspects of Big Data are you benchmarking or planning to benchmark? (ref. BDV Reference Model)
  - Data Analytics [100%]
  - Data Management [71%]
  - Data Processing [71%]

- [D12] What kind of data are you using/planning to use?
  - Real Data [86%]
  - Hybrid [86%]

- [D15] What type of Data Storage (Storage/Querying/Discovery) are you benchmarking/considering?
  - Graph Stores [67%]
  - NoSQL [67%]

- [D16] What is the most important type of Data Processing in your platform?
  - Interactive/(near) Real-time processing [71%]

- [D17] What types of data problems are you tackling?
  - Descriptive [71%]
  - Predictive [86%]

- [D18] What types of machine learning approaches do you typically use?
  - Supervised [86%]

- [D19] Which modelling techniques do you typically use?
  - Deep Learning [67%]

- [D20] What types of data are stored and processed in your system/platform? (Ref. BDV Reference Model types)
  - Time Series including IoT Data [86%]

- [D21] What are the technical key performance metrics that you (want to) measure in your system/platform/service?
  - End-to-end execution time (Runtime) [100%]
  - Throughput [67%]

- [D22] Which of the following qualitative features are important for your application/platform?
  - Fault-Tolerance [71%]

# Early Results from the Databench Business users Survey

Survey

- DataBench Survey, IDC, Interim results, 401 interviews
- Aiming at 800 (on going) + case studies
- October 2018
- 11 EU countries (7.7% in Italy)
- Final results to be presented at the European Big Data Value Forum and in the DataBench report due in December 2018

Users recognize the relevance of business benchmarking...
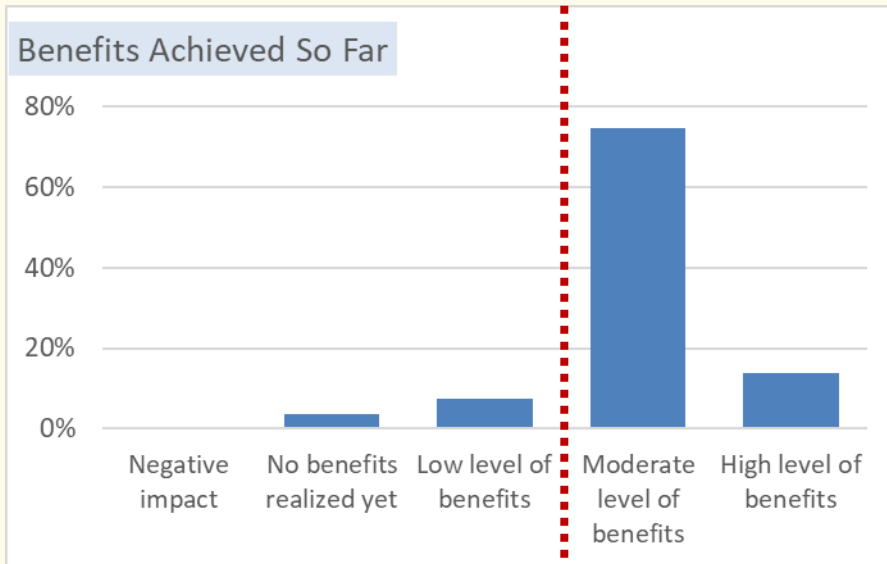
Respondents by Type of Use of BDA
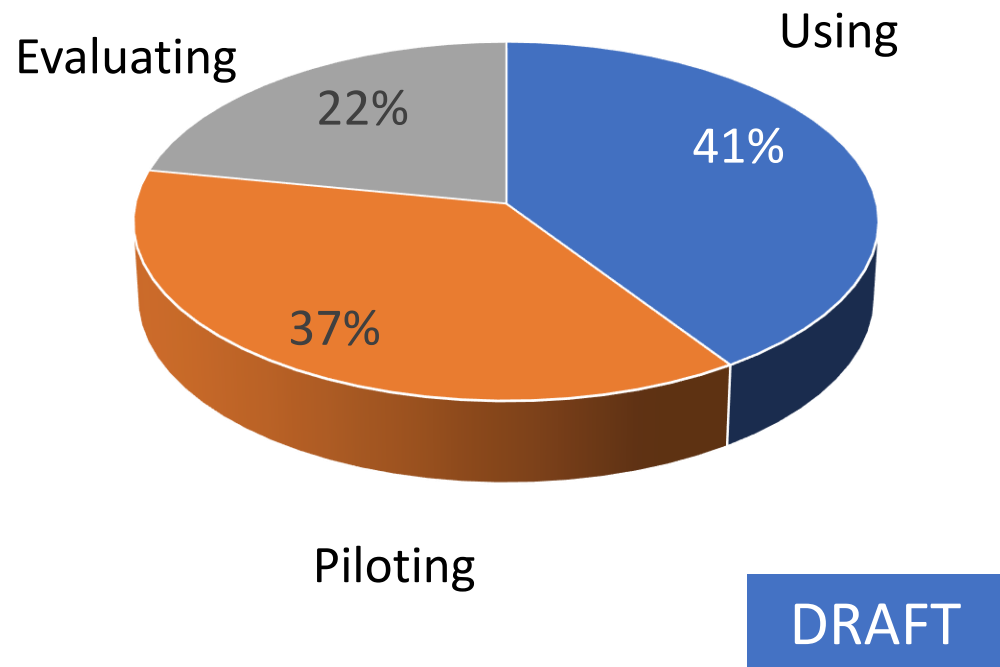
How important is benchmarking the business impacts of BDA?
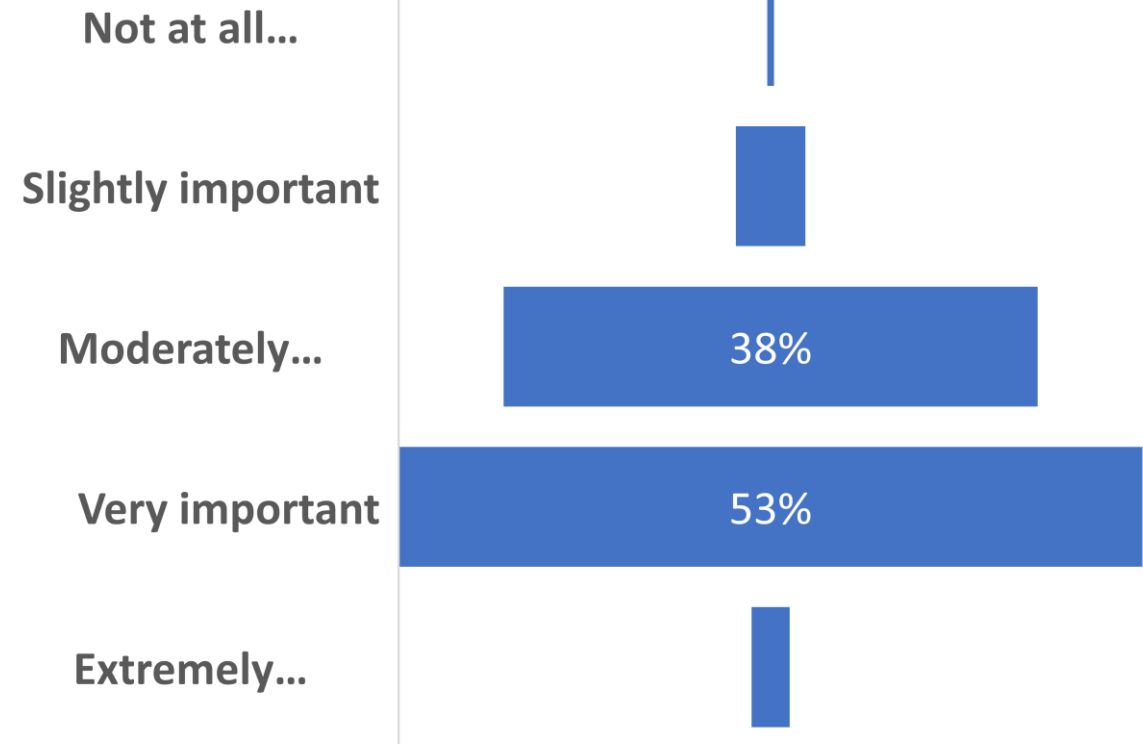
Source: Databench Survey, IDC, Interim results, 401 interviews, October 2018
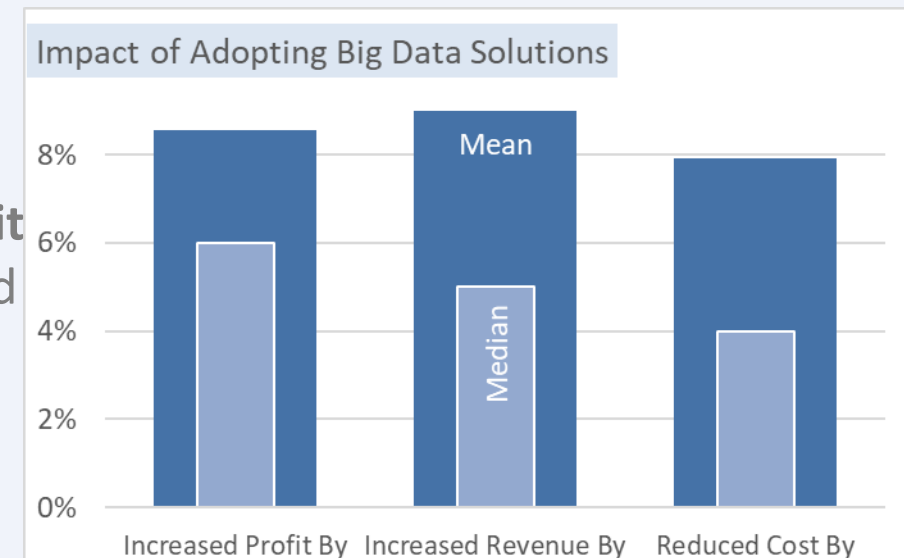
Big Data is Worth the Investment

Benefits Achieved So Far
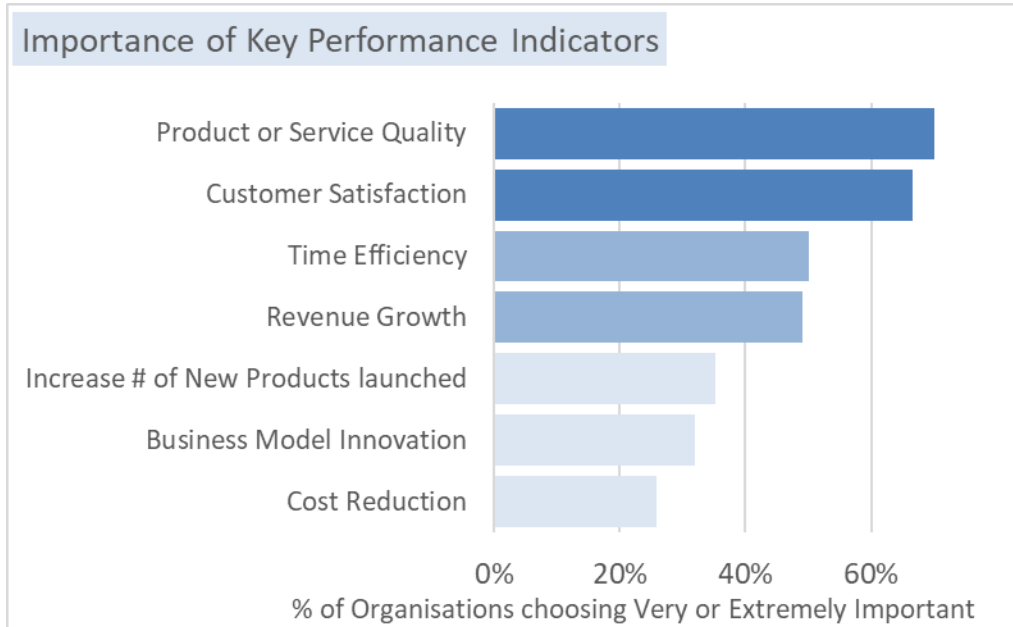

Nearly **90%** of businesses saw **moderate** or **high** levels of benefit in their Big Data implementation

DRAFT

Impact of Adopting Big Data Solutions

Mean

Median

Increased Profit By   Increased Revenue By   Reduced Cost By

Adopting Big Data Solutions increased **profit** and **revenue** by more than 8%, and reduced **cost** by nearly 8%

IDC | ANALYZE THE FUTURE

DataBench

Big Data implementation focus

Importance of Key Performance Indicators

Product or Service Quality
Customer Satisfaction
Time Efficiency
Revenue Growth
Increase # of New Products launched
Business Model Innovation
Cost Reduction
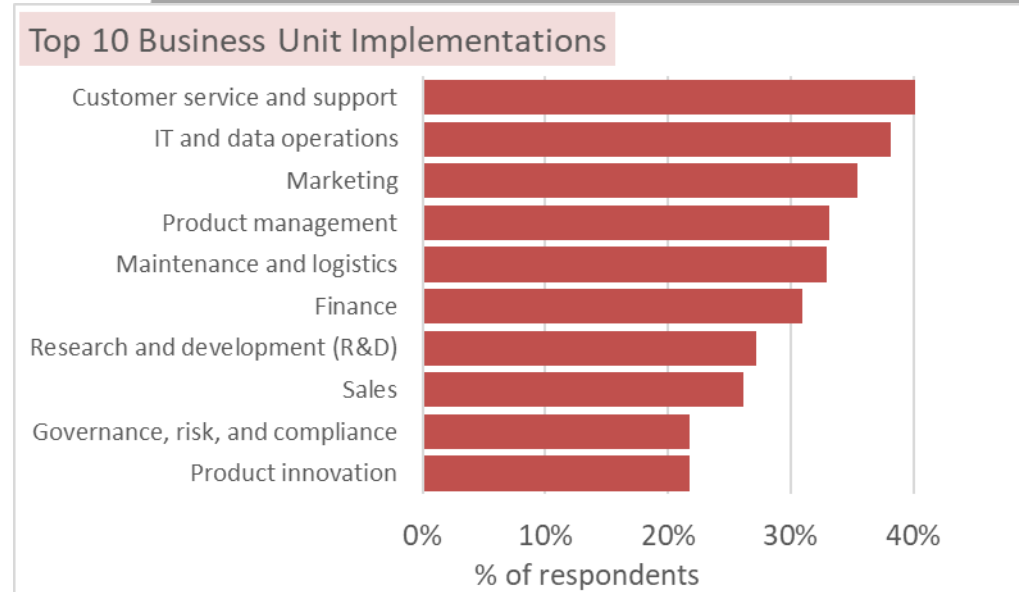
0%   20%   40%   60%
% of Organisations choosing Very or Extremely Important

Quality and **Customers** are the two most important KPI's

DRAFT

But implementation is balanced across **all** business units

Top 10 Business Unit Implementations

Customer service and support
IT and data operations
Marketing
Product management
Maintenance and logistics
Finance
Research and development (R&D)
Sales
Governance, risk, and compliance
Product innovation

0%   10%   20%   30%   40%
% of respondents

Overall, Big Data preference is for **growth** – with **new products and markets** – rather than improve efficiency and save cost

DataBench

Big Data – Key Use Cases

Top 20 Use Cases by # Respondents

- Risk exposure assessment
- New product development
- Price optimization
- Regulatory intelligence
- Automated Customer Service
- Customer profiling, targeting, and...
- Product & Service Recommendation...
- Fraud prevention and detection
- Supply chain optimization
- Predictive Maintenance
- Inventory and service parts optimization
- Customer scoring and/or churn mitigation
- Quality management investigation
- Asset management
- Smart warehousing
- Intelligent Fulfillment
- Connected vehicles optimization
- Network analytics and optimization
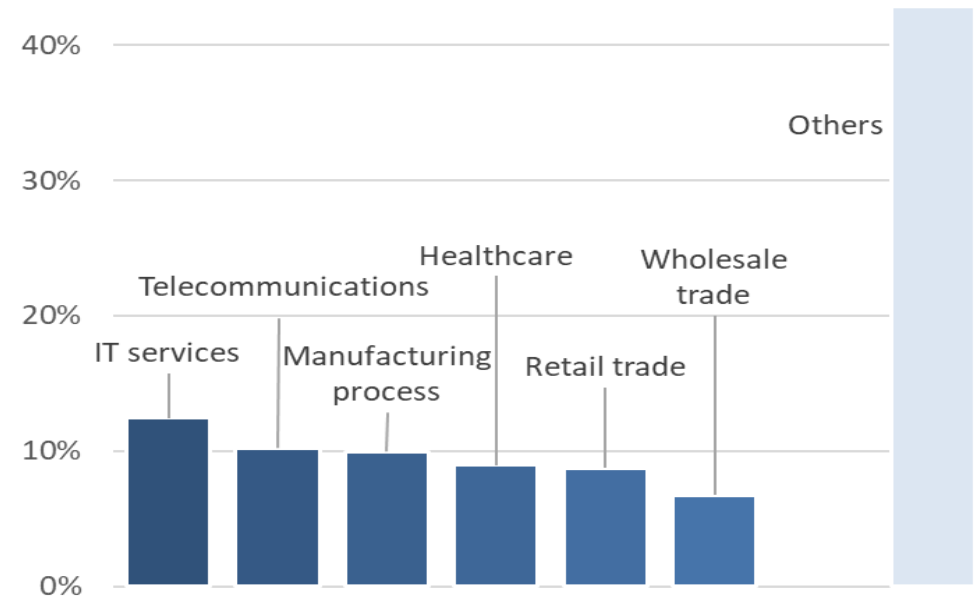- Energy consumption analysis and...
- Field service optimization

using or Evaluating

Respondents

Final results to be presented at the European Big Data Value Forum and in the Databench report due in December 2018

DRAFT

Respondent share by industry

IT services — Telecommunications — Manufacturing process — Healthcare — Retail trade — Wholesale trade — Others

© IDC

Goals of DataBench

- Provide methodologies and tools to help assess and maximise the business benefits of BDT adoption

- Provide criteria for the selection of the most appropriate BDTs solutions

- Provide benchmarks of European and industrial significance

- Provide a questionnaire tool comparing your choices and your KPIs with your peers

Interested in participating?

- Expression of interest to become a case study and monitoring your Big Data KPIs

- Answer a survey on your Big Data experiences

Contacts

✉ info@databench.eu

🐦 @DataBench_eu

f DataBench

in DataBench Project

DataBench

▶ DataBench Project

Evidence Based Big Data Benchmarking to Improve Business Performance

barbara.pernici@polimi.it

rstevens@idc.com