Industrial Analytics

Symposium and Summer School On Service-Oriented Computing
June 17-23, 2019 in Crete, Greece

Dr.-Ing. Christoph Gröger
Lead Architect Data Analytics
Robert Bosch GmbH
Christoph.Groeger@de.bosch.com

BOSCH

Agenda

1. Introduction
2. Concepts
3. Use Cases
4. Processes & Techniques
5. Platforms & Architectures
6. Challenges

BOSCH

Introduction
Bosch: Business Sectors and Key Figures*

Bosch Group

€ 78.5 billion euros in sales

~410,000 associates

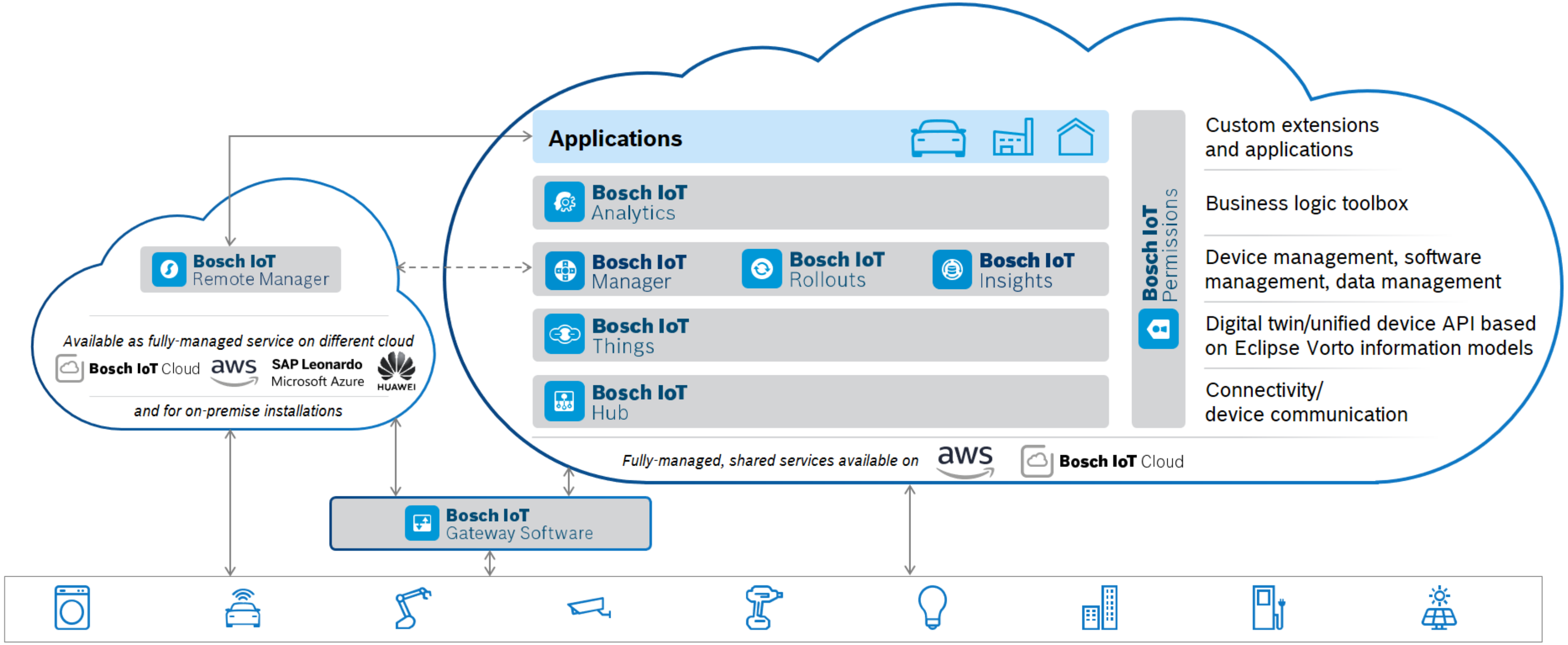
Mobility Solutions	▶ One of the world's leading providers of mobility solutions
Industrial Technology	▶ Leading in drive and control technology, packaging, and process technology
Energy and Building Technology	▶ One of the leading manufacturers of security and communication technology ▶ Leading manufacturer of energy-efficient heating products and hot-water solutions
Consumer Goods	▶ Leading supplier of power tools and accessories ▶ Leading supplier of household appliances

* As of 12.18

* Figures adjusted for extraordinary effects resulting from changes in the consolidated group and methodological changes and depreciation and amortization resulting from purchase price allocation.

BOSCH

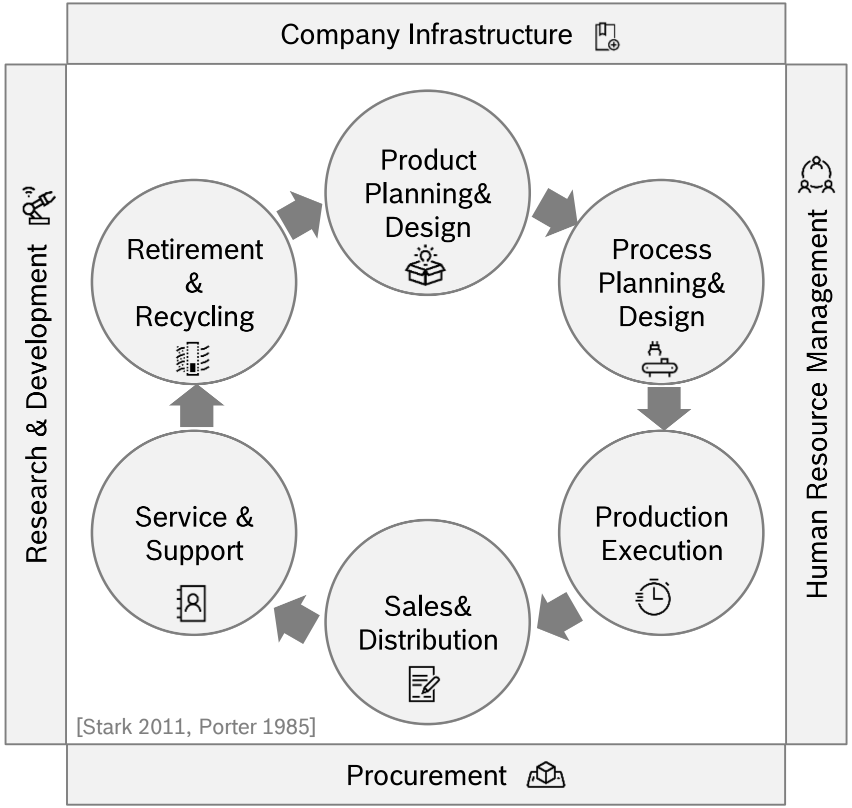
Introduction
Bosch: IoT Suite



[http://www.bosch-iot-suite.com]

Concepts

Terminology: Industrial Analytics



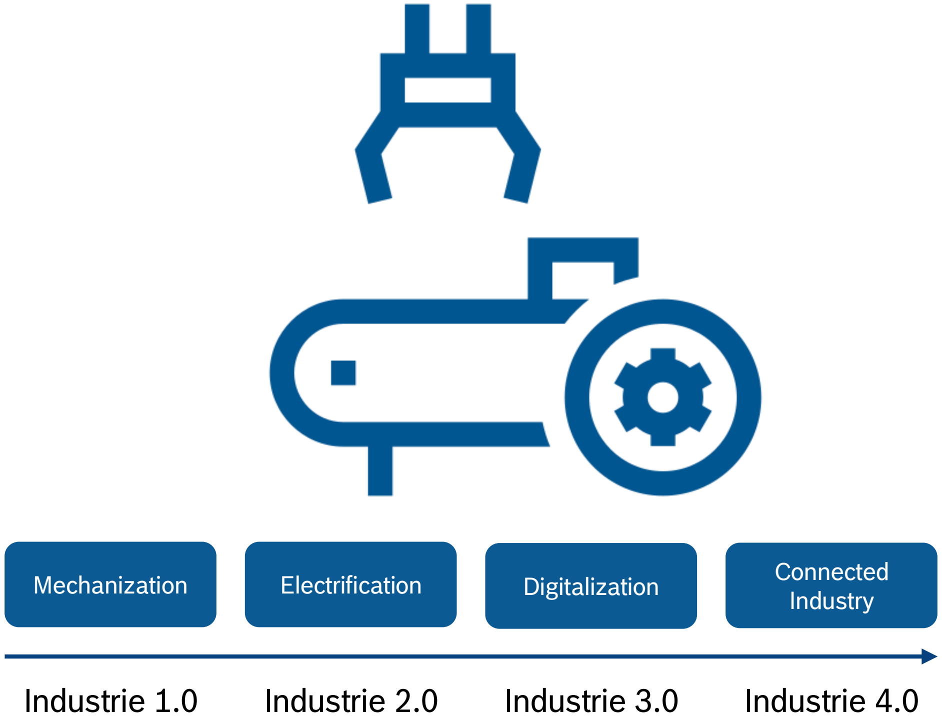
Defining “Industrial Analytics”

- Data analytics for industrial applications
- That is, industrial enterprises and industrial value chains as application domains of data analytics
- Sometimes also called “Industrial Intelligence” or “Industrie 4.0 Analytics”

Note: Industrial Analytics refers to the entire industrial value chain, not only single phases such as production.

[Gröger 2018, Harper et al. 2015]

Concepts
Terminology: Industrie 4.0

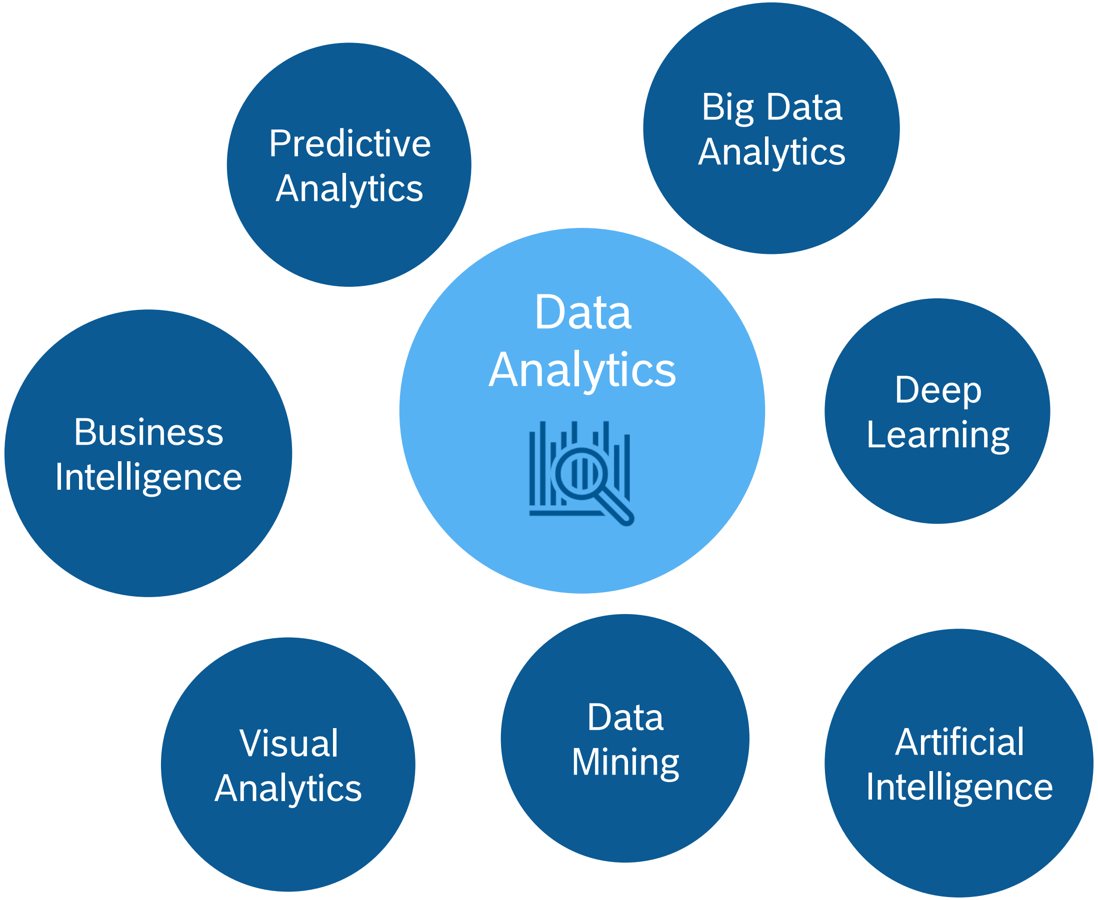


Defining "Industrie 4.0"

- Next generation of industrial value generation based on the comprehensive use of internet-of-things (IoT) technology and cyber-physical systems
- Aims at the complete digital interconnection of all processes and objects across the industrial value chain

[Bauernhansl 2014]

Concepts
Terminology: Data Analytics



Defining “data analytics”

Broader sense

- Analysis of data for decision support and knowledge extraction
- Then, business intelligence refers to the use of data analytics in enterprises

Narrower sense

- Synonym for “advanced analytics”
- Collective term for all analysis techniques beyond classical reporting and OLAP, especially data mining, visual analytics, text analytics and more

Note: Data analytics are applied across both small data and big data.

[Runkler 2012, Kemper et al. 2010]

Concepts
Drivers for Industrial Analytics

Novel Data Processing and Analytics Technologies

- Hadoop: HDFS, Map&Reduce, Spark, Impala, ...
- NoSQL: MongoDB, Cassandra, Neo4j, ...
- Deep Learning: TensorFlow, Keras, Caffe, ...
- ...

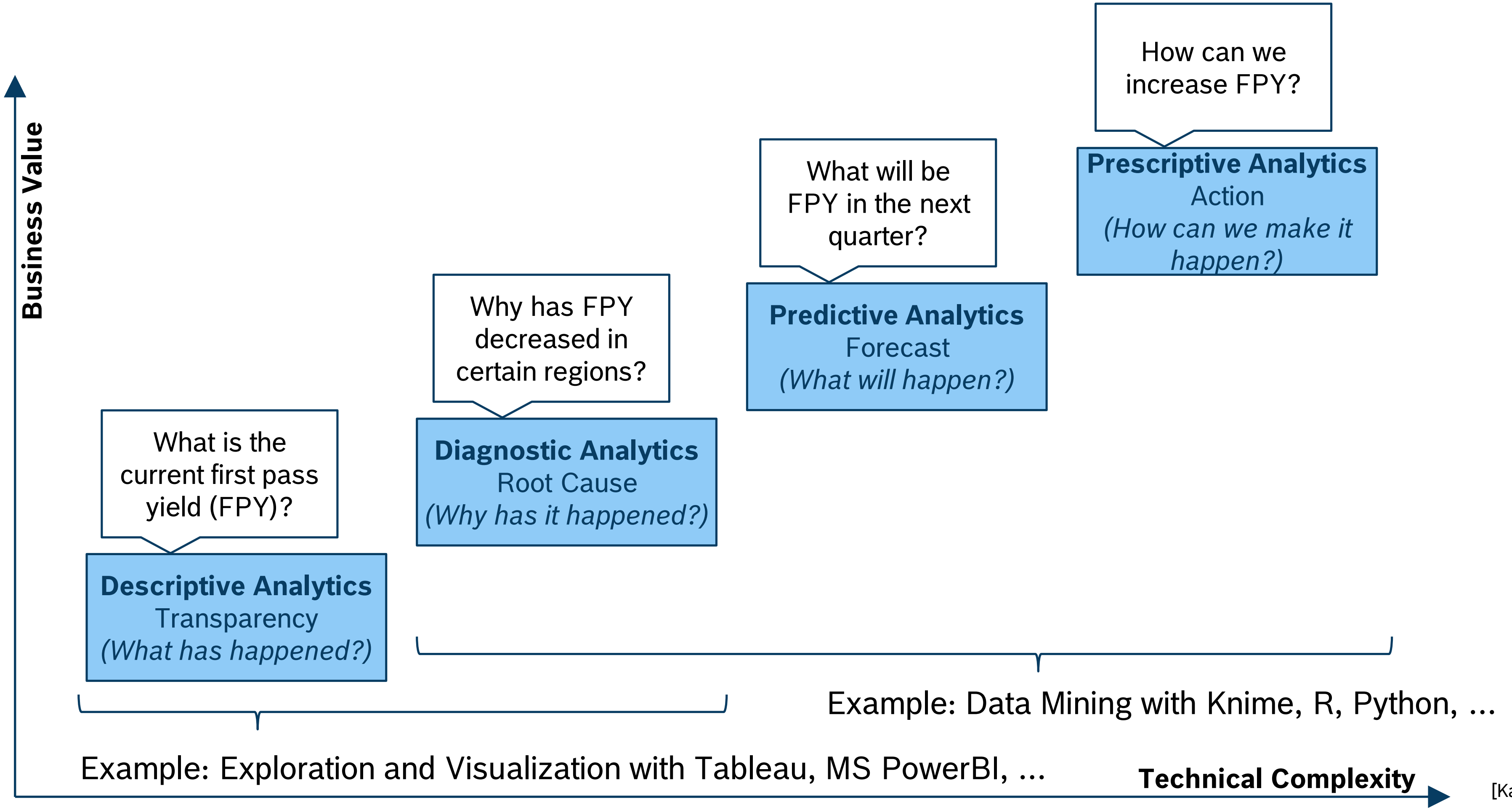
Cloud Computing

- Massively scalable storage and compute power in pay-as-you-go mode
- Complex analytics tools and frameworks automatically provisioned and always up-to-date
- Separation of storage and compute for higher flexibility, e.g. different Hadoop tool stacks on the same data
- ...

Industrie 4.0 and Digitalization of Industrial Value Chain

- Digital twins across product life cycle
- Manufacturing execution systems
- Sensors on the shop floor
- ...

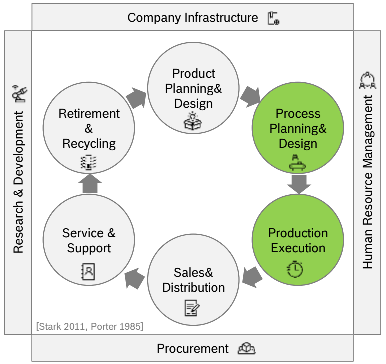
Concepts
Types of Data Analytics



[Kart et al. 2013, Gröger 2018]

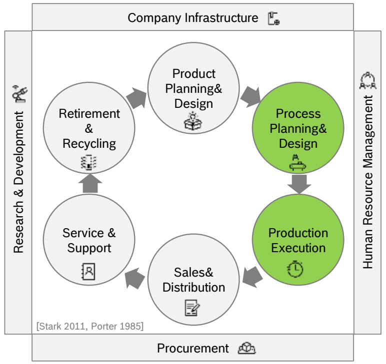
Use Cases
Manufacturing Analytics at Global Scale (I)

- Improving quality, reducing scrap, improving first pass yield
  - Single Process Optimization
    - Techniques: Visual Analytics, Data Mining
    - Data Sources: Machine Log Data, Sensor Data, MES Data
    - Challenges: Data Integration, Data Quality
  - Value Stream Optimization
    - Techniques: Visual Analytics, Data Mining
    - Data Sources: Measurement Data, Machine Log Data, Sensor Data, MES Data
    - Challenges: Data Integration, Data Quality



Use Cases
Manufacturing Analytics at Global Scale (II)

- Improving uptime, reducing cost, optimizing resources
  - Predictive Maintenance
    - Concept: Two types of predictive maintenance
      - Avoid unplanned downtime due to machine failures
      - Avoid unnecessary maintenance on quality related parts (tools like nozzles, stencils, molding tools)
    - Techniques: Data Mining
    - Challenges: Documentation of changes on the machines are often incomplete, manual and non-machine-readable
  - Customer Demand Forecast
    - Concept: Optimize Midterm planning accuracy
    - Techniques: Data Mining
    - Challenges: Extracting ERP data without overloading source systems

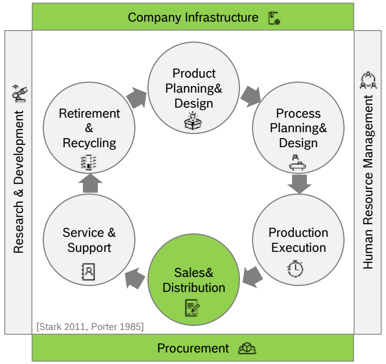


BOSCH

Use Cases
End-to-End Process Mining

- Improving process transparency, performance and quality end-to-end
  - Techniques : Process Mining (uses specialized data mining algorithms on event logs to automatically reconstruct the process flow and identify trends, patterns and bottlenecks)
  - Data Sources: Process events from ERP systems
  - Challenges: Extracting ERP data without overloading source systems, data integration across ERP systems

Process	Focus Topics
Order-to-Cash	• Reduction of manual order settlement • Elimination of delivery blocks • Automation of backlog reporting
Purchase-to-Pay	• Automation rate • Manual interventions • Price changes rates
IT Ticket Handling	• Time to resolution • Ping pong tickets • Multi hop tickets

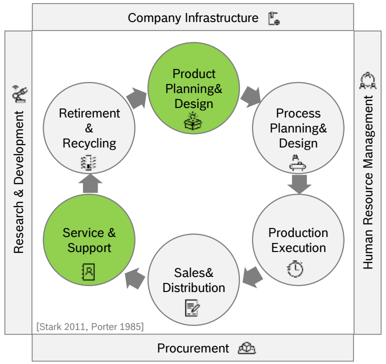


CI/DAP4.2 - Dr.-Ing. Christoph Gröger | 2019-06-18

Use Cases
Engineering in the Loop

- Improving product design based on real-world product usage data
  - Techniques: Visual Analytics
  - Data Sources: Simulation data from engineering, master data, sensor data on product usage
  - Challenges: Data Acquisition, Data Integration

“Me as an Engineer, I want to evaluate an as-is-design with field data to identify a possible product redesign.”



Processes & Techniques
Process Model CRISP-DM

Cross-Industry Standard Process for Data Mining

Competencies	CRISP-DM Steps
domain knowledge, project management	Business Understanding
domain knowledge, IT systems knowledge	Data Understanding
SW engineering, data engineering	Data Preparation
algorithm knowledge	Modeling
algorithm knowledge, domain knowledge	Evaluation
domain knowledge, IT systems knowledge	Deployment

Key role: (Citizen) data scientist

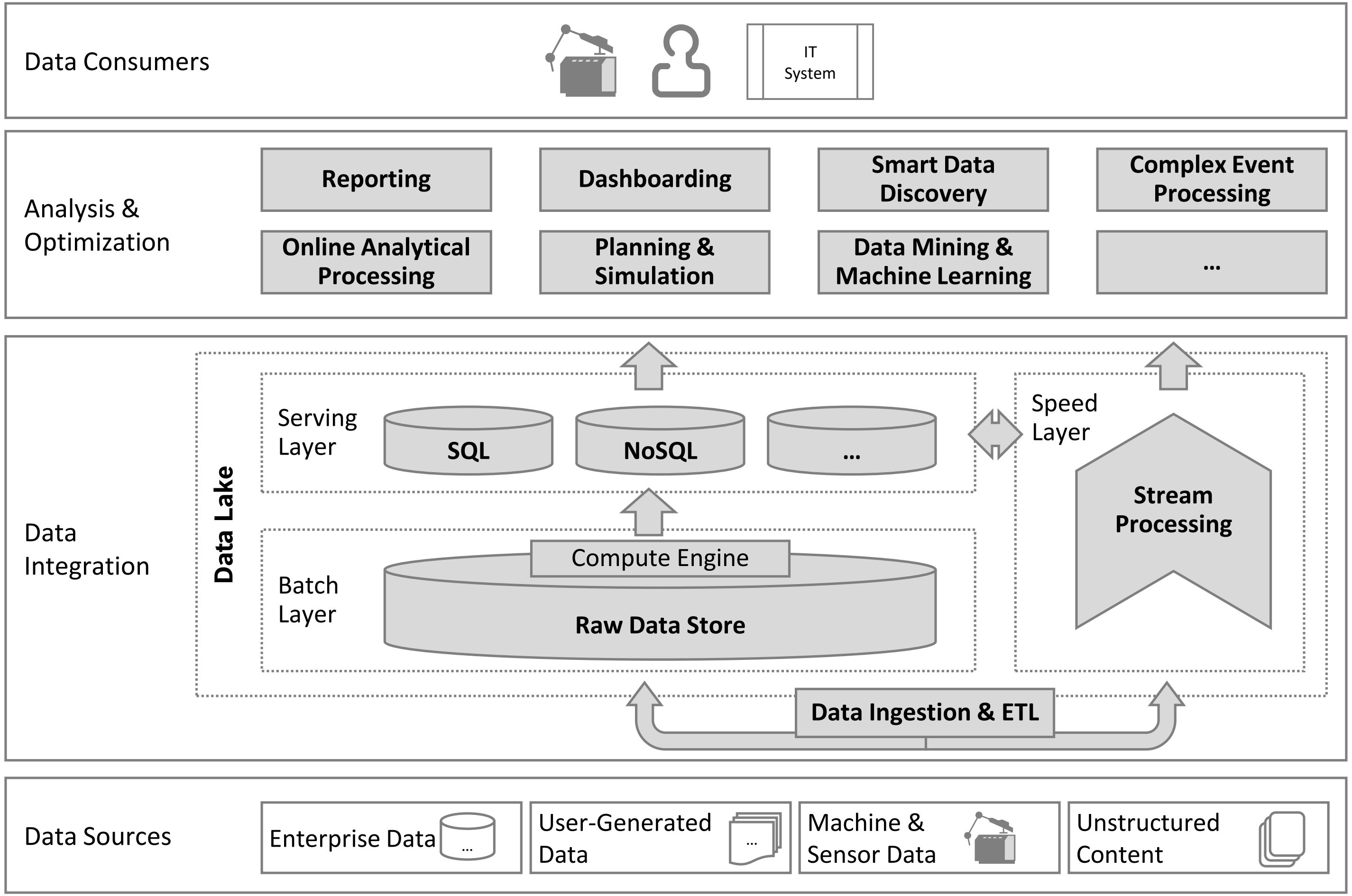
60-80% of effort

CRISP-DM: Cross-Industry Standard Process for Data Mining

- Iterative
- Experimental
- Agile

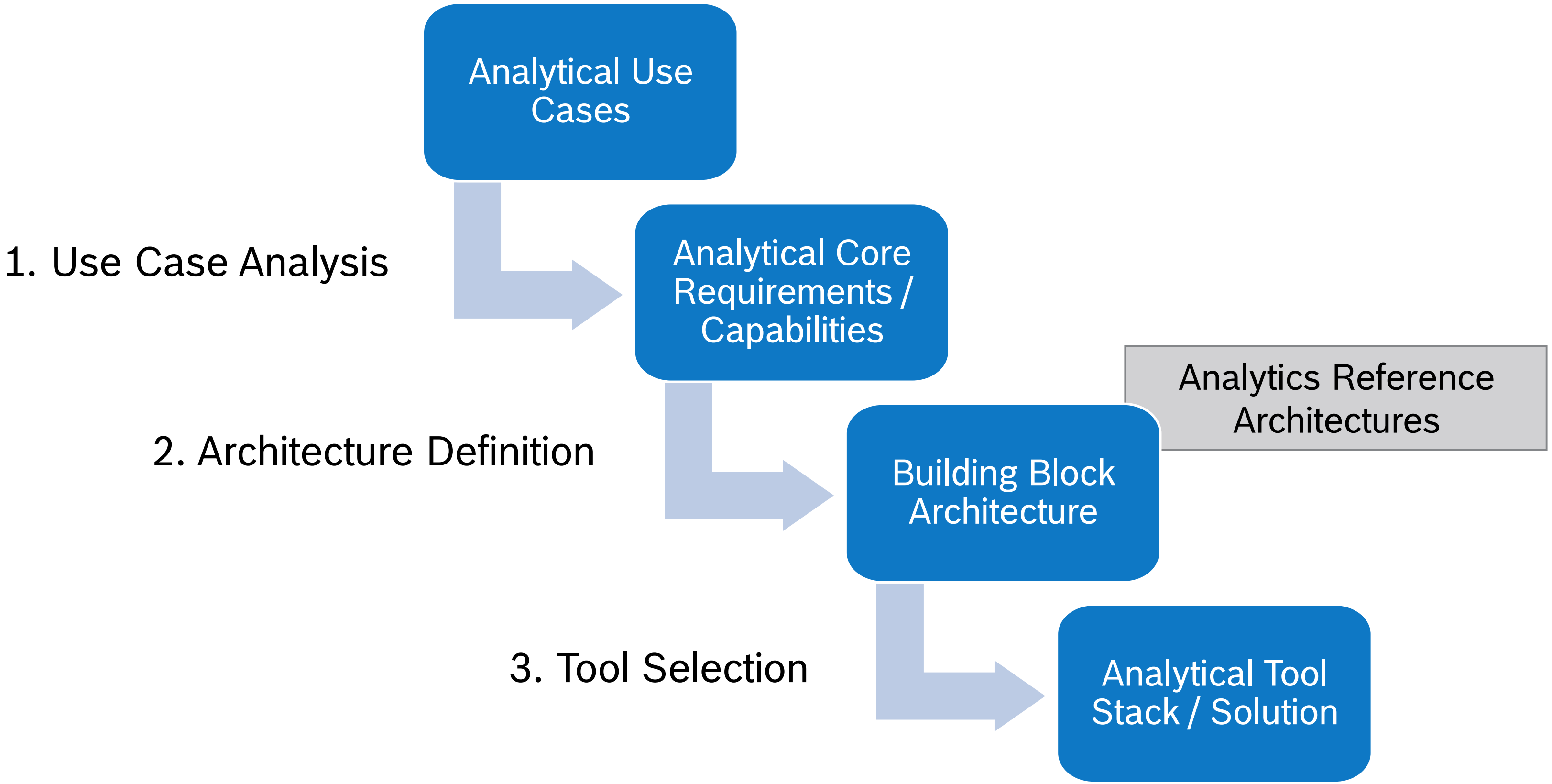
[Han et al. 2012]

Platforms & Architectures
Analytical Platforms: Sample Data Lake Architecture



[Gröger 2018, Gröger et al. 2019]

Platforms & Architectures
Architecture Methodology



Challenges
Key Challenges and Research Directions

Architecture

Developing standardized and reusable analytical services across products, processes and factories

➢ **The analytical operating system**

People

Empowering business domain specialists to do advanced analytics

➢ **The citizen data scientist**

Organization

Governing a federated data lake landscape

➢ **Holistic analytics governance**

[Gröger 2018]

BOSCH

References

Baars, Kemper (2015): *Integration von Big Data-Komponenten in die Business Intelligence,* Zeitschrift für Controlling 27

Bauernhansl (2014): *Die Vierte Industrielle Revolution – Der Weg in ein wertschaffendes Produktionsparadigma*, in: Industrie 4.0 in Produktion, Automatisierung und Logistik, Springer

Bose (2009): *Advanced Analytics,* Industrial Management & Data Systems 109(2)

Gartner (2016): *Select the Right Architecture Model for Your Modern BI and Analytics Platform,* Gartner Report

Gluchowski, Gabriel, Dittmar (2008): *Management Support Systeme und Business Intelligence*, Springer

Gröger (2015): *Advanced Manufacturing Analytics*, Josef Eul Verlag

Gröger (2018): *Building an Industry 4.0 Analytics Platform*, Datenbank-Spektrum 18(1)

Gröger, Hoos (2019): *Ganzheitliches Metadatenmanagement im Data Lake*, Proc. of BTW 2019 Conference

Han, Kamber, Pei (2012): *Data mining. Concepts and techniques*, Morgan Kaufmann

Harper, Zheng, Jacobs et al. (2015): *Industrial Analytics Pipelines*, Proc. of IEEE BigDataService 2015 Conference

Kart , Linden, Schulte (2013): *Extend your portfolio of analytics capabilities*. Gartner research note G00254653

Kemper, Baars, Mehanna (2010): *Business Intelligence*, Vieweg+Teubner

Kromer (2015): *Modern hybrid big data warehouse architectures*, Business Intelligence Journal 19(4)

Marz, Warren (2015): *Big Data*, Manning Publications

Mathis (2017): *Data Lakes*, Datenbank-Spektrum 17(3)

Porter (1985): *Competitive Advantage*, The Free Press

Runkler (2012): *Data Analytics*, Springer

Russom (2011): *Big Data Analytics*, TDWI Report

Stark (2011): *Product Lifecycle Management*, Springer

Terrizzano, Schwarz, Roth, Colino (2015): *Data Wrangling: The Challenging Journey from the Wild to the Lake*, Proc. of CIDR 2015 Conference

Zheng, Peltsverger (2015): *Web Analytics Overview*, in: Encyclopedia of Information Science and Technology, IGI Global