Digital Twins: a New Perspective on Cyber Physical Systems

SARA VINCO

Politecnico di Torino, Italy

sara.vinco@polito.it

Politecnico di Torino

A Digital What?

Digital twin

wikipedia.org

From Wikipedia, the free encyclopedia

A **digital twin** is a real-time virtual representation of a real-world physical system or process (a *physical twin*) that serves as the indistinguishable digital counterpart of it for practical purposes, such as system simulation, integration, testing, monitoring, and maintenance.[1] [2] [3] Though the concept originated earlier, the first practical definition of a digital twin originated from NASA in an attempt to improve physical-model simulation of spacecraft in 2010.[4] Digital twins are the result of continual improvement in the creation of product design and engineering activities. Product drawings and engineering specifications have progressed from handmade drafting to computer-aided drafting/computer-aided design to model-based systems engineering.

WIKIPEDIA
The Free Encyclopedia

Does this sound familiar?

WIKIPEDIA
The Free Encyclopedia

Digital twin

wikipedia.org

From Wikipedia, the free encyclopedia

A **digital twin** is a real-time virtual representation of a real-world physical system or process (a *physical twin*) that serves as the indistinguishable digital counterpart of it for practical purposes, such as system simulation, integration, testing, monitoring, and maintenance.[1] [2] [3] Though the concept originated earlier, the first practical definition of a di
attempt to improve physical-model simul
twins are the result of continual improver
and engineering activities. Product drawi
have progressed from handmade draftin
aided design to model-based systems er

www.cpsschool.eu

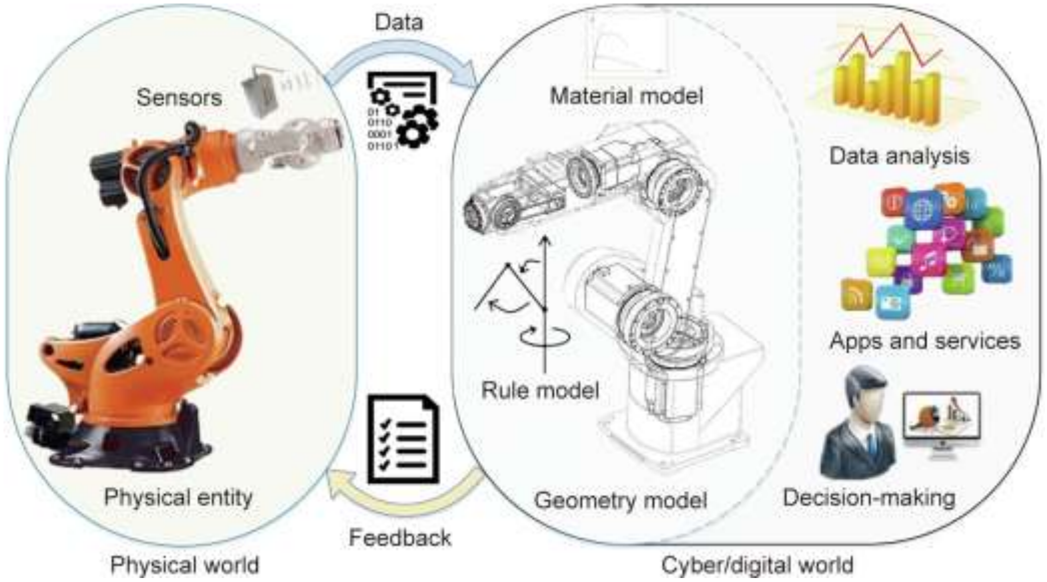
Cyber-physical systems are complex and autonomous ensembles of different components that directly cooperate to offer smart and adaptive functionalities. They are composed of sensors, actuators, and processing components that are deeply entangled: they constantly exchange information and actively interact with the external environment. So they are increasingly used in a variety of applications with a growing market. CPSs will potentially bring about significant social benefits; nevertheless, there is no such thing as a free lunch and several new challenges and trade-offs must be faced in their design, especially when the CPS should adapt to the changing environment or heal itself. Indeed, different nodes could cooperate to enrich the offered functionality.

Does this sound familiar?

CPSs and DTs share the same essential concepts and look at the same reality

- Intensive cyber-physical connection, real-time interaction, research for «smart» reaction to evolving conditions
- So why two definitions?
  Almost simultaneous: CPS in 2006, DT in 2003

Shift in perspective!



CPSs vs. DTs

Cyber Physical System

- Sensors and actuators as main modules
- Control commands generated based on predefined rules or models, that activate actuators to adapt the system to changes
- May control more than one process/object simultaneously

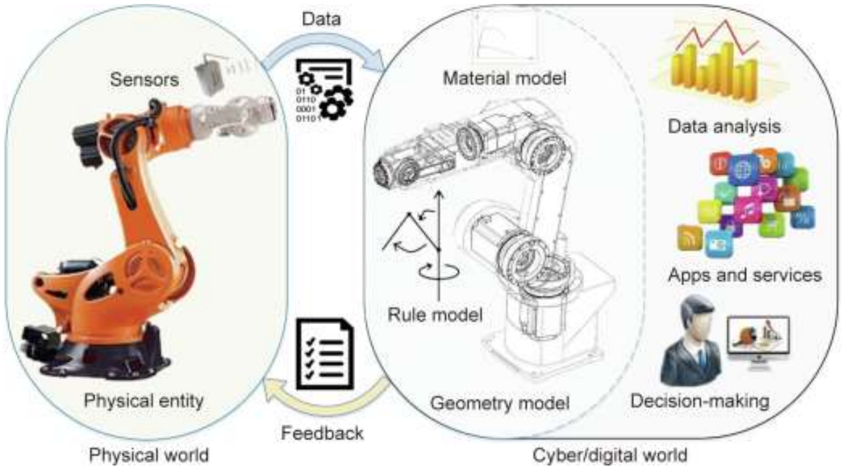
Digital Twin

- Goal is to provide a physical and functional description of an object
- Virtual models and physical process/object co-evolve throughout the whole lifetime
- Sensors as data generators
- Actuators to react to any predicted malfunction or decision derived by the cyber part

CPSs vs. DTs

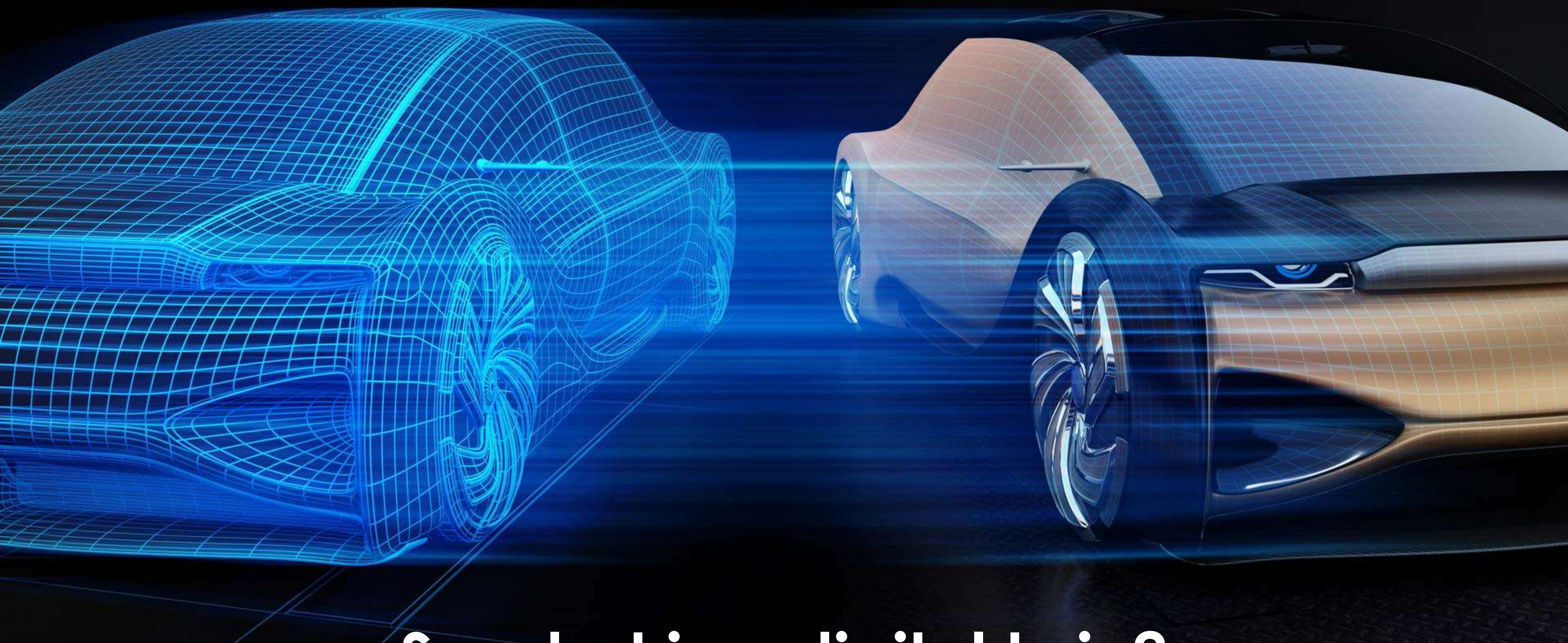
Cyber Physical System

- *Science perspective*: focus on control problems and feedback loop



Digital Twin

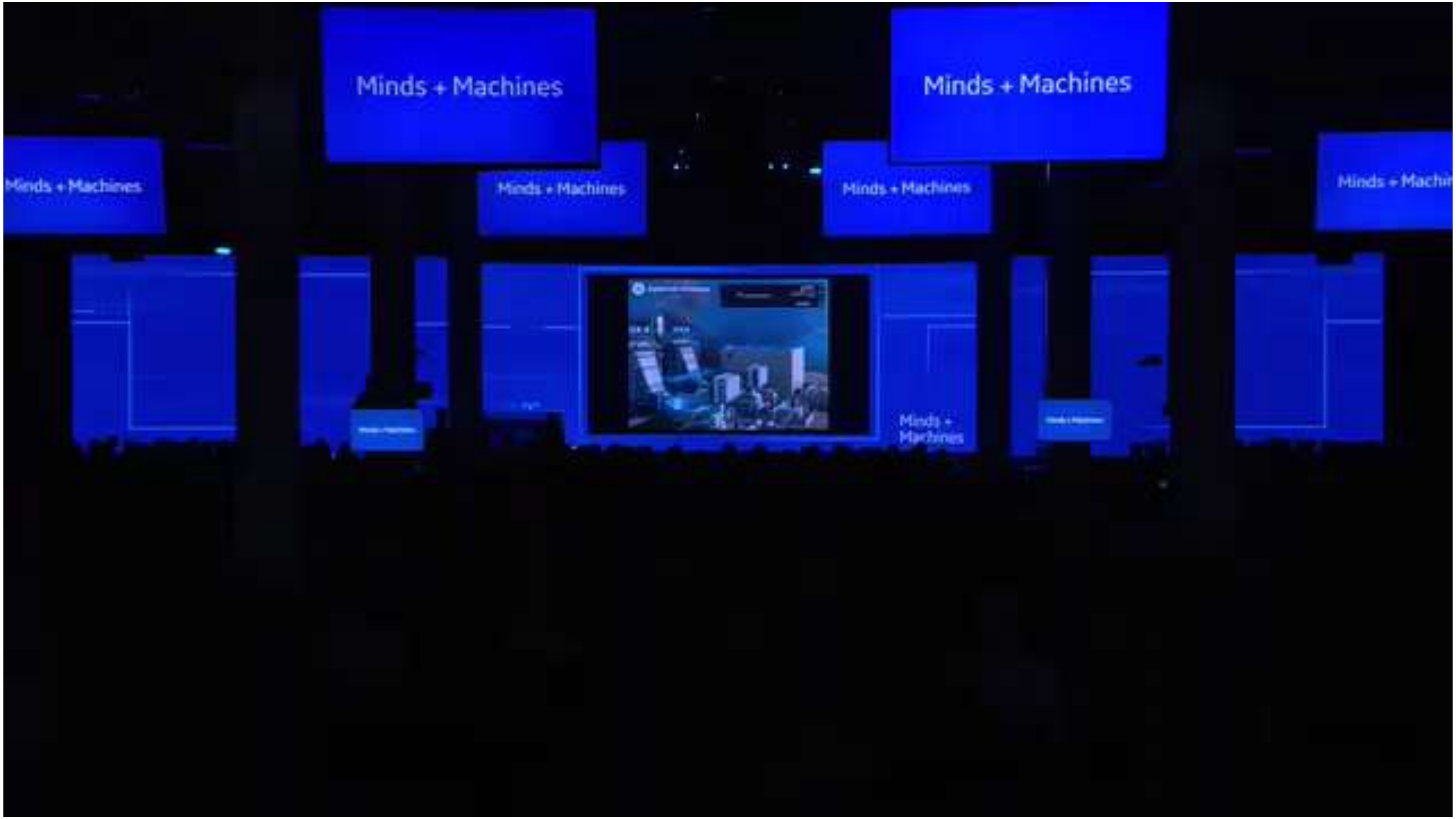
- *Engineering and manufacturing perspective*
- Exploits data analysis of collected data (sensed, historical, etc.) to enable accurate predictions, rational decisions, informed production
- Data and data-centered models as core elements



So what is a digital twin?

So what is a Digital Twin?

www.youtube.com/watch?v=2dCz3oL2rTw



So what is a Digital Twin?

Virtual representation that interacts with the physical object throughout its lifecycle to provide intelligence for prediction, evaluation, optimization, etc.

5 main ingredients:

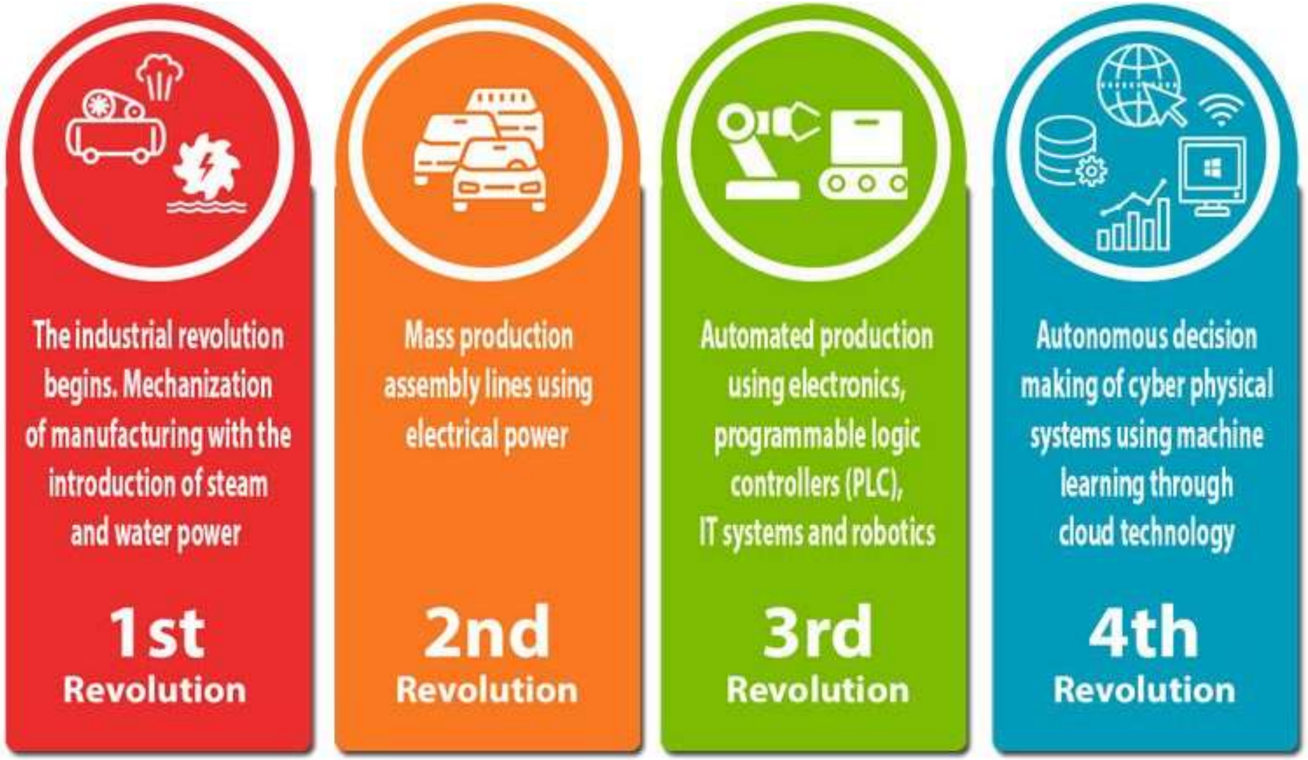
- Physical space plus virtual space
- Their connection for virtual-physical interaction
- Data from virtual and physical domains used for comprehensive information capture
- Functions for unified management and on-demand usage:
  - Detection, judgment, prediction

THE CONTEXT

INDUSTRY 4.0 AND ITS ENABLING TECHNOLOGIES

Industry 4.0

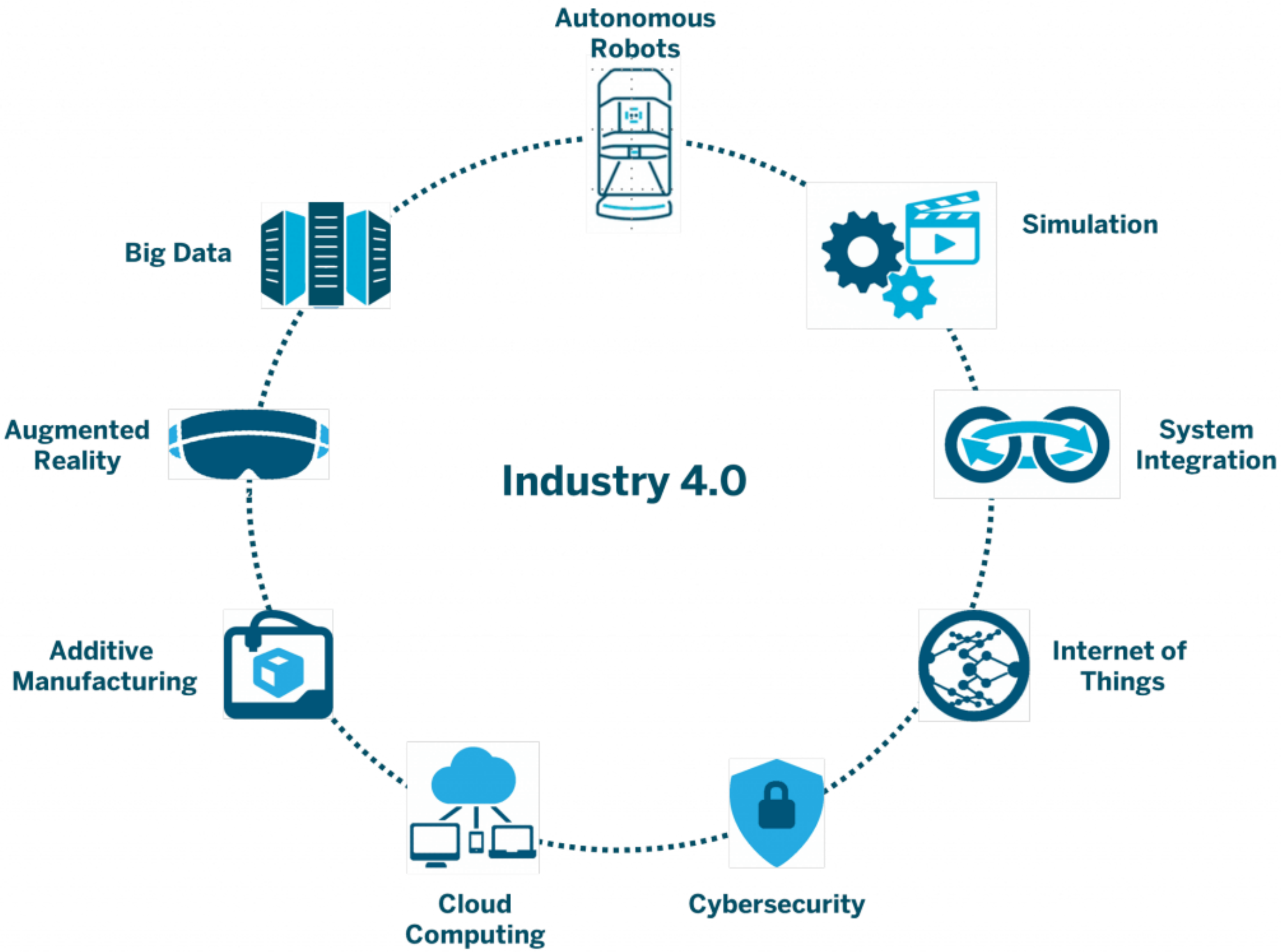
Industry 4.0 has been defined as *"a name for the current trend of automation and data exchange in manufacturing technologies, including cyber-physical systems, the Internet of things, cloud computing and cognitive computing and creating the smart factory"*



Industry 4.0 and Digital Twins

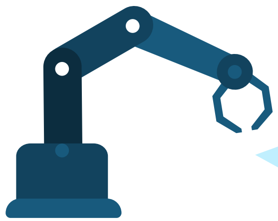
Different ingredients, i.e., enabling technologies

- Even more automation than in the 3rd industrial revolution
- Bridge the physical and digital world through Industrial IoT
- Shift from a central industrial control system to one where smart products define the production steps
- Closed-loop data models and control systems
- Personalization/customization of products



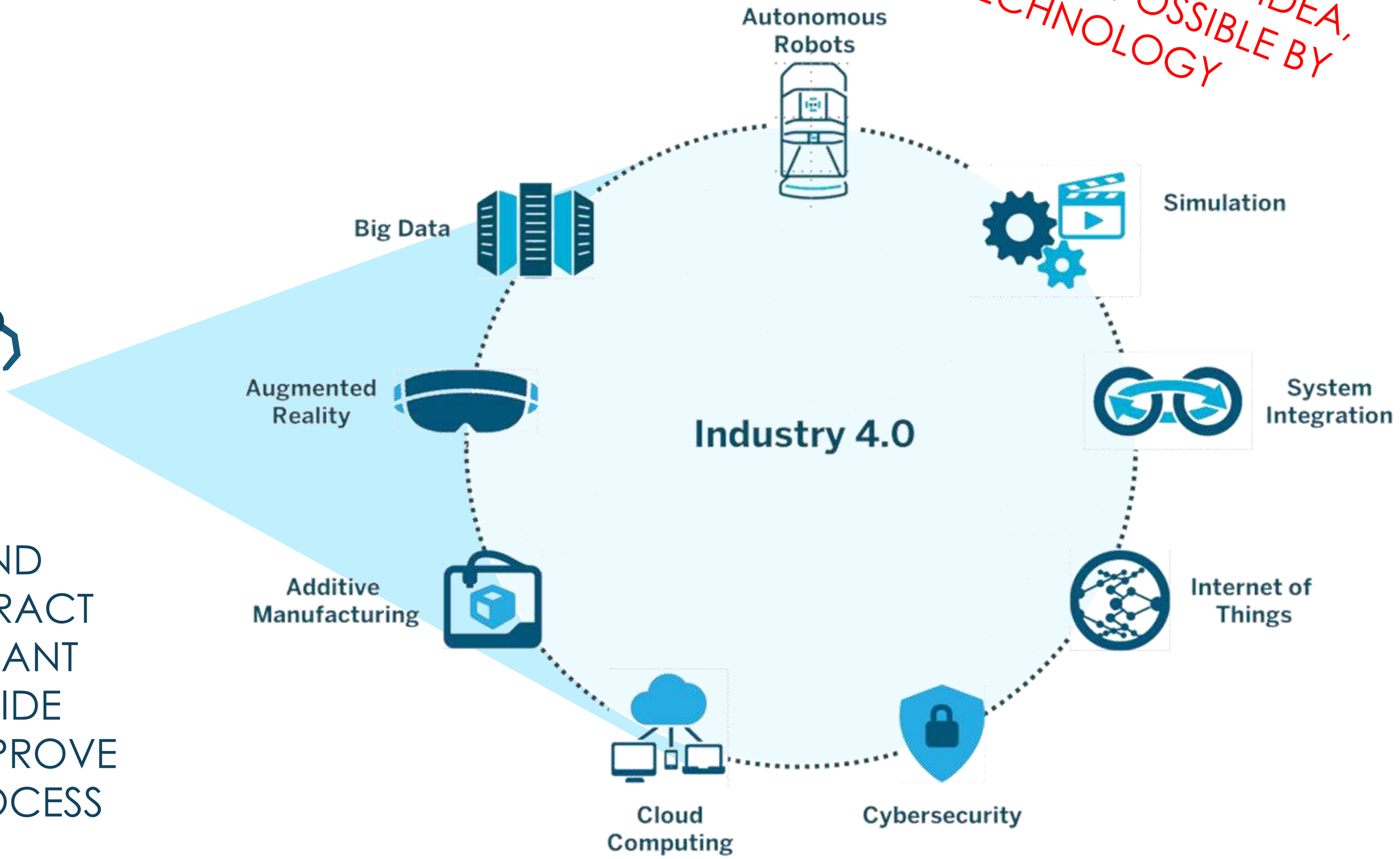
Industry 4.0 and Digital Twins

ONCE A VISIONARY IDEA, NOW MADE POSSIBLE BY TECHNOLOGY



DIGITAL TWINS

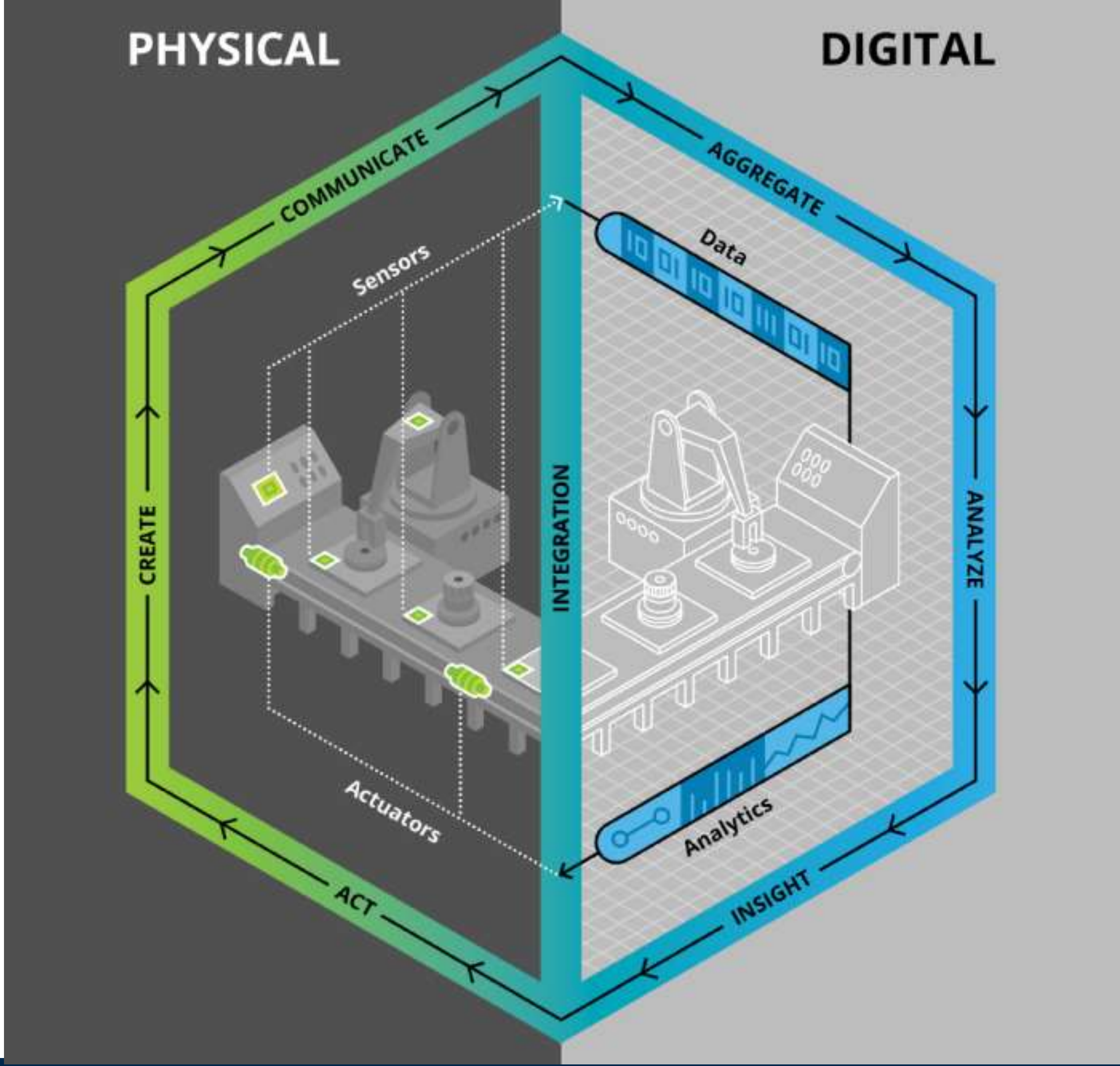
VIRTUAL MODEL AND PHYSICAL PLANT INTERACT THROUGHOUT THE PLANT LIFECYCLE TO PROVIDE INTELLIGENCE AND IMPROVE THE PRODUCTION PROCESS



Physical and Virtual: the Digital Twin

Digital twin as a practical application of technologies for seamless integration of physical and virtual

- For each physical object there exists a virtual mirror model able to analyze, evaluate, optimize, predict, etc.
- The two parts interact with each other and remain synchronous in closed loop
- Data from both physical and virtual can be fused to generate more comprehensive information



https://www2.deloitte.com/cn/en/pages/consumer-industrial-products/articles/industry-4-0-and-the-digital-twin.html

DIGITAL TWINS

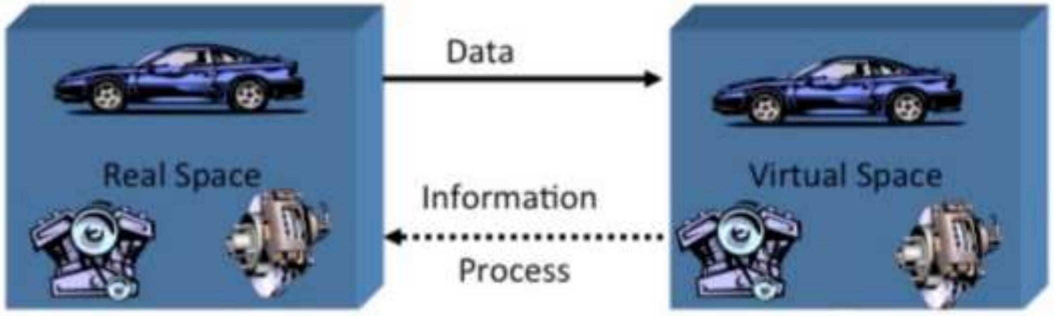
DEFINITIONS, PARTS AND ENABLING TECHNOLOGIES

2003 2005 2010 2012 2014 2017 2018

First definition

Given by Michael Grieves, University of Michigan

Presentation to industry for the formation of a Product Lifecycle Management (PLM) center



KEY CONCEPT: INFORMATION MIRRORING MODEL

- Real space + virtual space
- Information flow between virtual space and real space
- During all lifespan: creation, manufacture, operation, and disposal

M. Grives, Conceptual Ideal for PLM, 2002

2003 2005 2010 2012 2014 2017 2018

Definition by NASA

“Integrated, multi-physics, multi-scale, probabilistic simulation of a vehicle or system that uses the best physical models, sensor updates, fleet history, etc. to mirror the life of its flying twin”

- Concept first applied in the 1970s during the Apollo 13 program
- Rapidly account for changes to the vehicle while exposed to the extreme conditions in space
- NASA found they could no longer base corrective decisions on the original model
- The actual module had undergone significant changes
- The original model needed to be updated to more closely mirror the current state of the module!

2003 2005 2010 2012 2014 2017 2018

Digital twin as key technology

For future vehicles by NASA and US Air Force for structural health management

Multidisciplinary physics-based methodology

Whitepaper of digital twin

Application to more domains, including automotive, oil and gas, healthcare and medicine

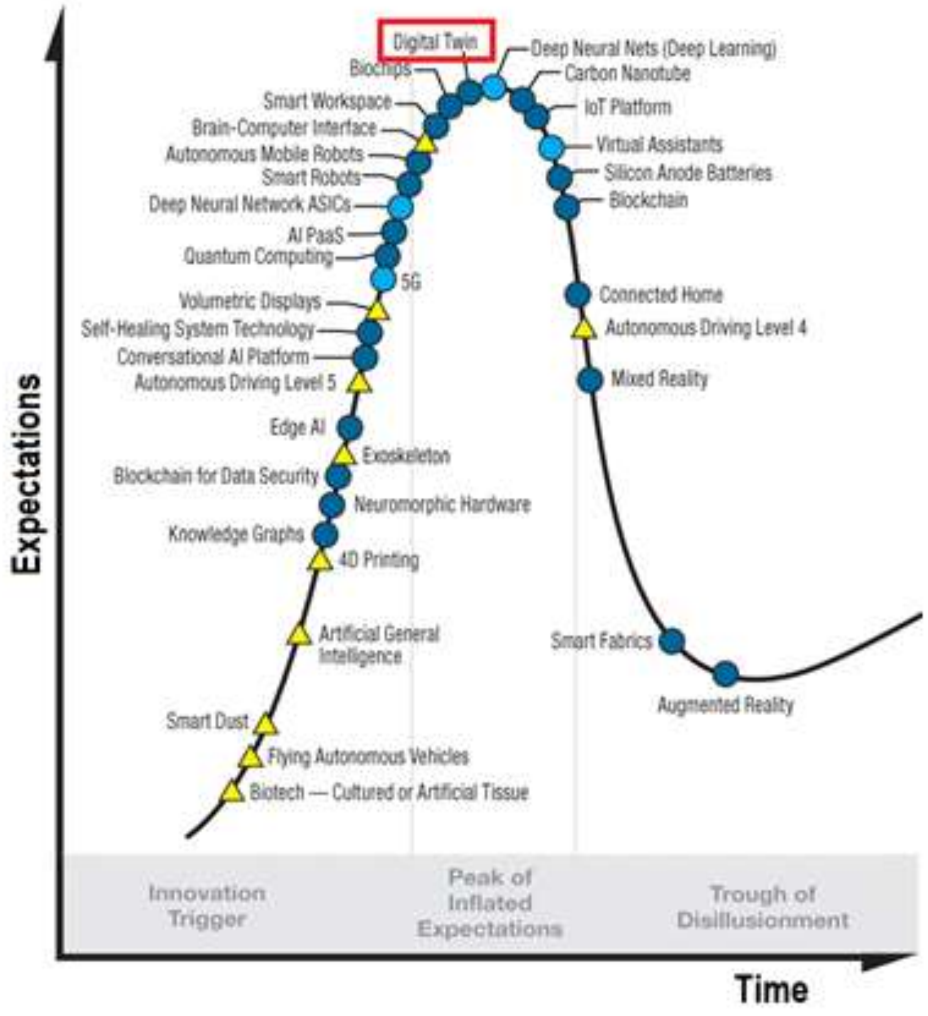
First conference

Exploration of key technologies, mechanisms, implementation methods

Definition of digital twin-driven design

2003 2005 2010 2012 2014 2017 2018

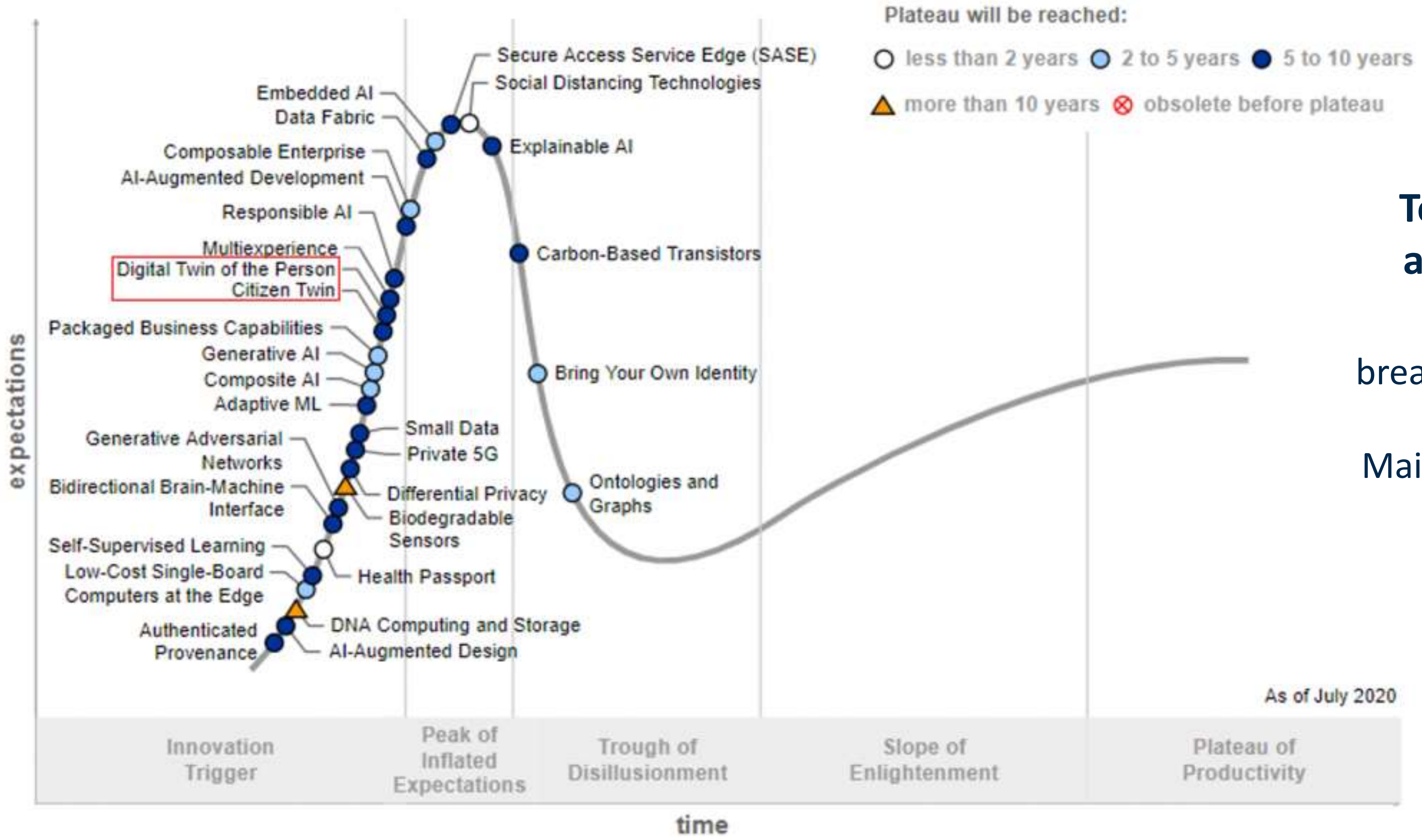
Hype Cycle for Emerging Technologies, 2018



Digital twin as top trend

Gartner: digital twin as one of the top 10 most promising technological trends

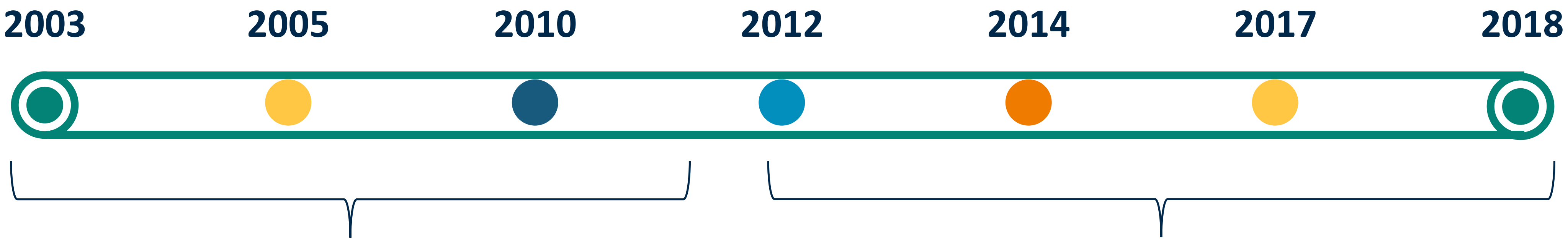
2003 2005 2010 2012 2014 2017 2018



Today: digital twin as available technology

Potential technology breakthrough, early proof-of-concept stories

Mainstream adoption in 5 to 10 years

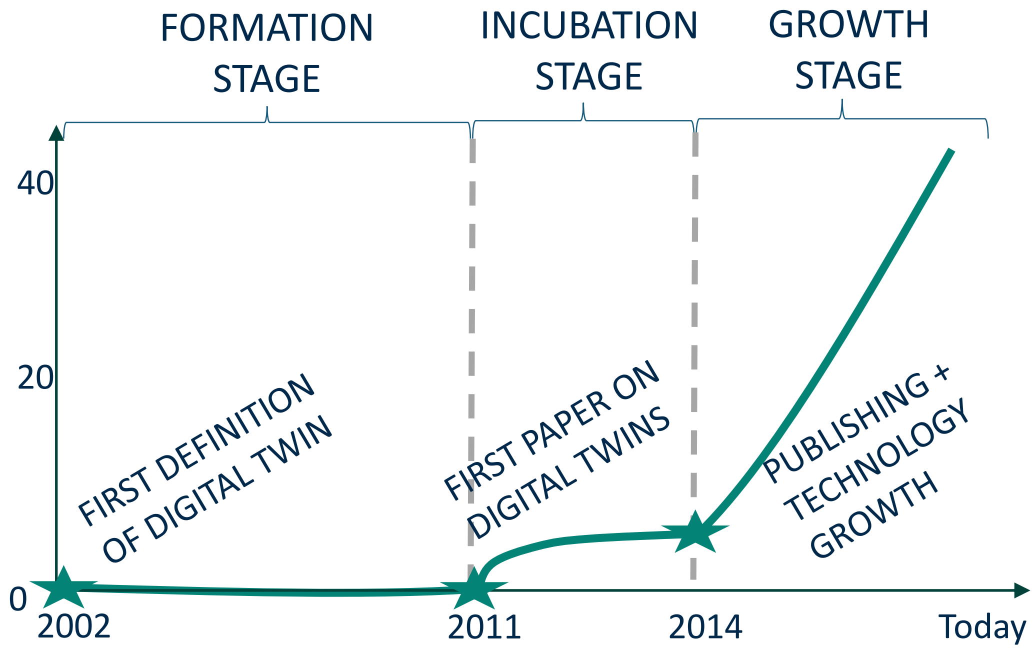


Technology gap

From 2003 to 2011 technology was immature to support practically viable digital twins

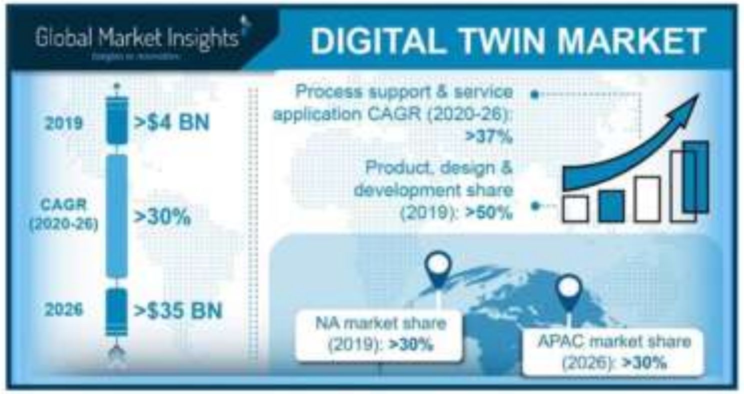
Rapid growth of research and development

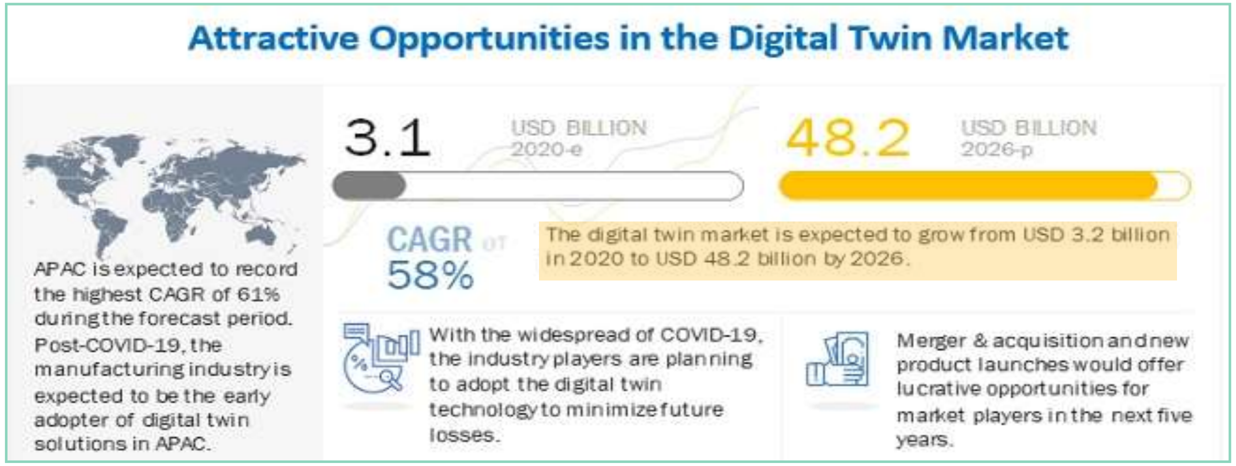
Triggered by technology growth, e.g., cloud computing, IoT, big data, sensor technologies, …
Enabling technology is crucial to allow digital twins!

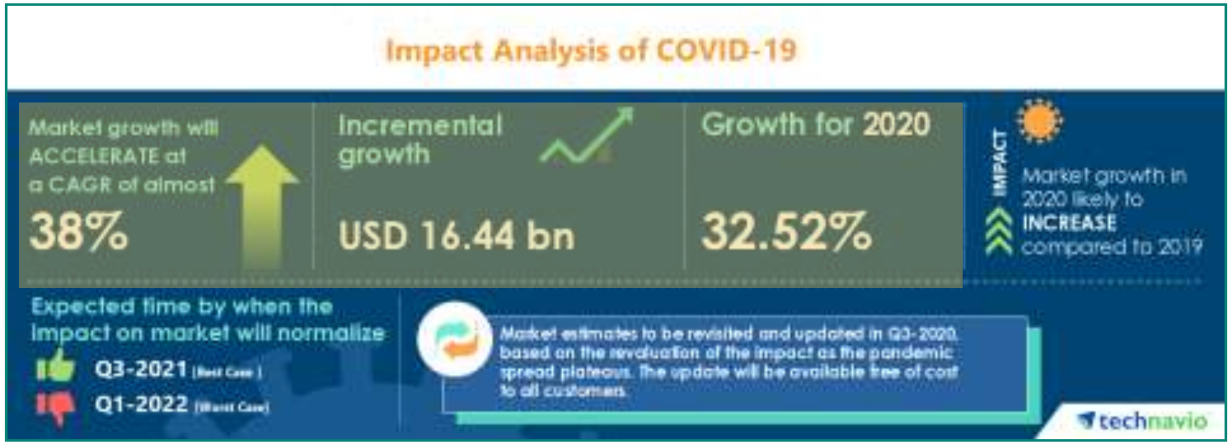


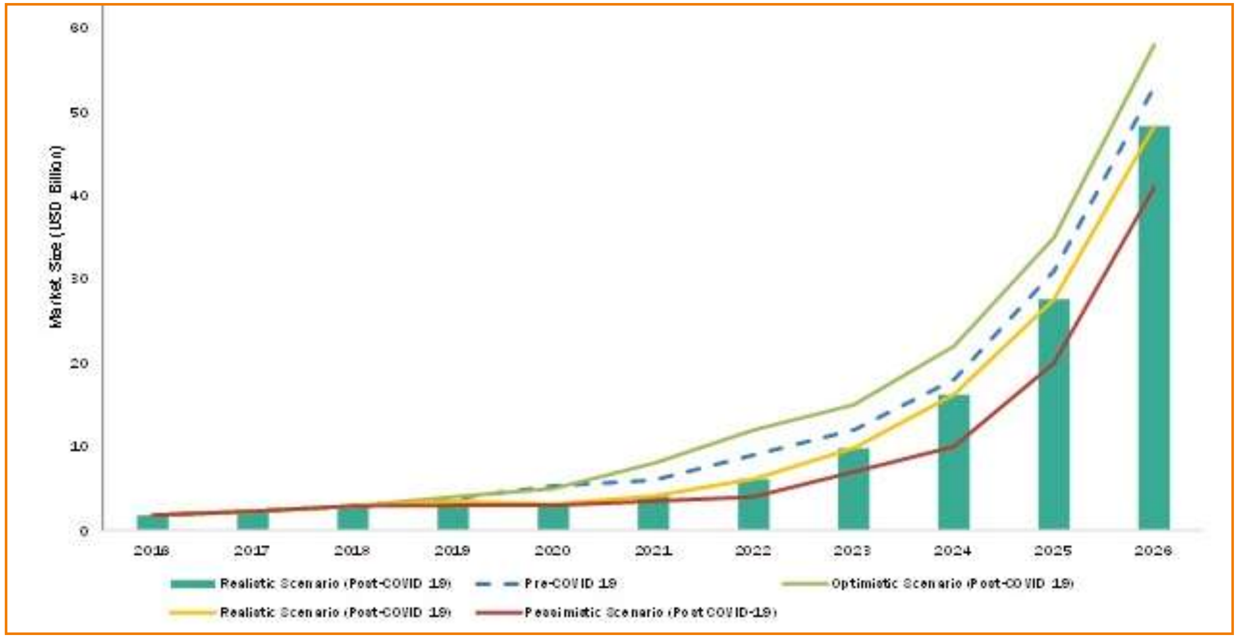
Industry Trends

Digital Twin Market size exceeded USD 4 billion in 2019 and is estimated to grow at a CAGR of over 30% between 2020 and 2026. The rising internet penetration coupled with the proliferation of smartphones and the advent of advanced technologies such as VR, AR, AI, machine language, deep learning, and blockchain is driving the market growth. This technology is helping to reduce the time to market, increase operational efficiency, and improve product lifecycle for various end-use segments such as healthcare, automotive, and aerospace.





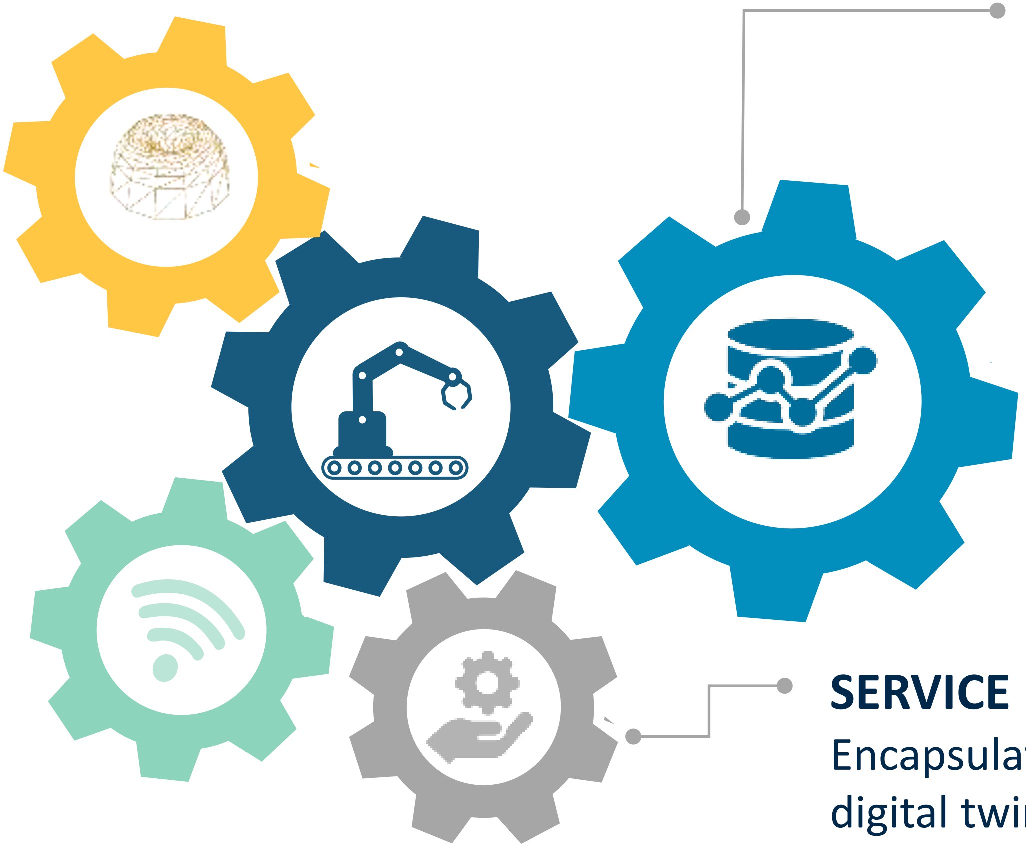




Parts of a Digital Twin

MODEL

Digital companion is made of a set of models
Reproduce with high fidelity the properties, behaviors and rules of the physical object
Operate autonomously in the virtual space
Ability to predict problems on physical side
Validate performance before system completion



DATA

Physical plus virtual data provides intelligence, e.g., digital model built from sensor data, decision based on simulated data, and operations based on predefined models

CONNECTIONS

Enable closed-loop interaction between digital and physical elements

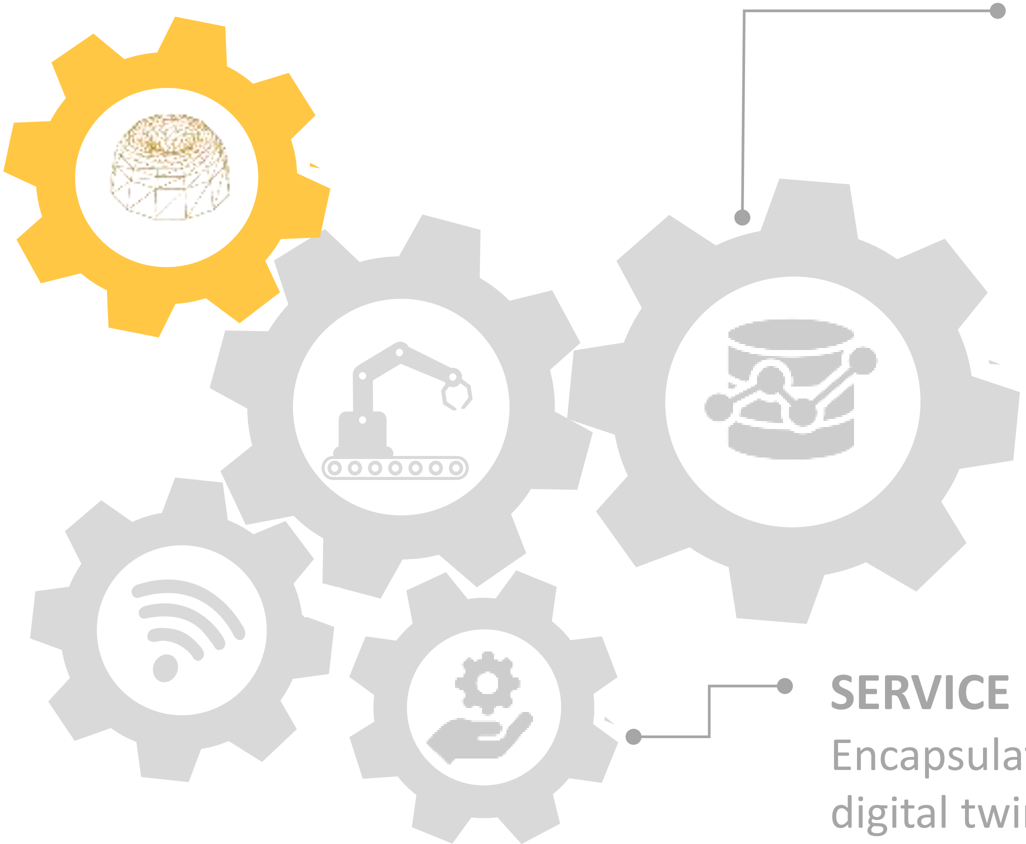
SERVICE

Encapsulate functions of the digital twin into services of easy and convenient usage

Parts of a Digital Twin

MODEL

Digital companion is made of a set of models
Reproduce with high fidelity the properties, behaviors and rules of the physical object
Operate autonomously in the virtual space
Ability to predict problems on physical side
Validate performance before system completion



DATA

Physical plus virtual data provides intelligence, e.g., digital model built from sensor data, decision based on simulated data, and operations based on predefined models

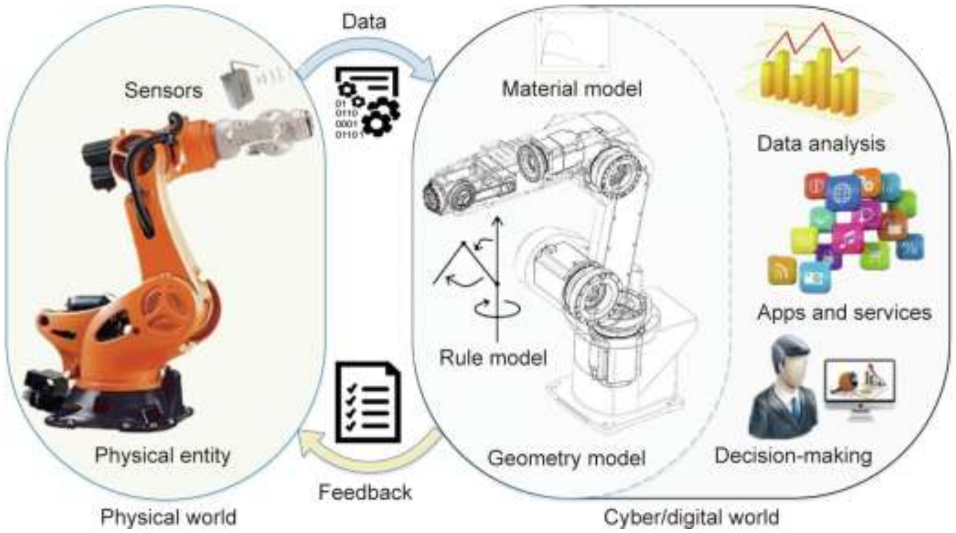
CONNECTIONS

Enable closed-loop interaction between digital and physical elements

SERVICE

Encapsulate functions of the digital twin into services of easy and convenient usage

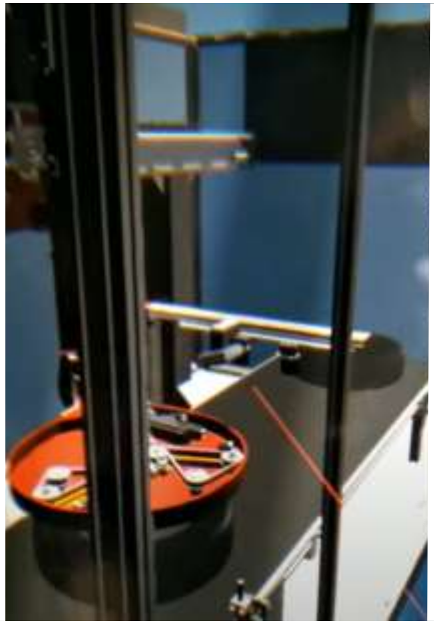
Modeling the Virtual Entity



Physical entity exists physically and can complete missions in space and produce outputs

Virtual entity:

- Made of a set of models to describe the physical entity from different perspectives (e.g., geometrical dimensions, physical properties)
- No standard modeling flow
- Strictly dependent on application, level of detail, expertise of the designer, etc.
- How and what do I want to reproduce the physical aspects?



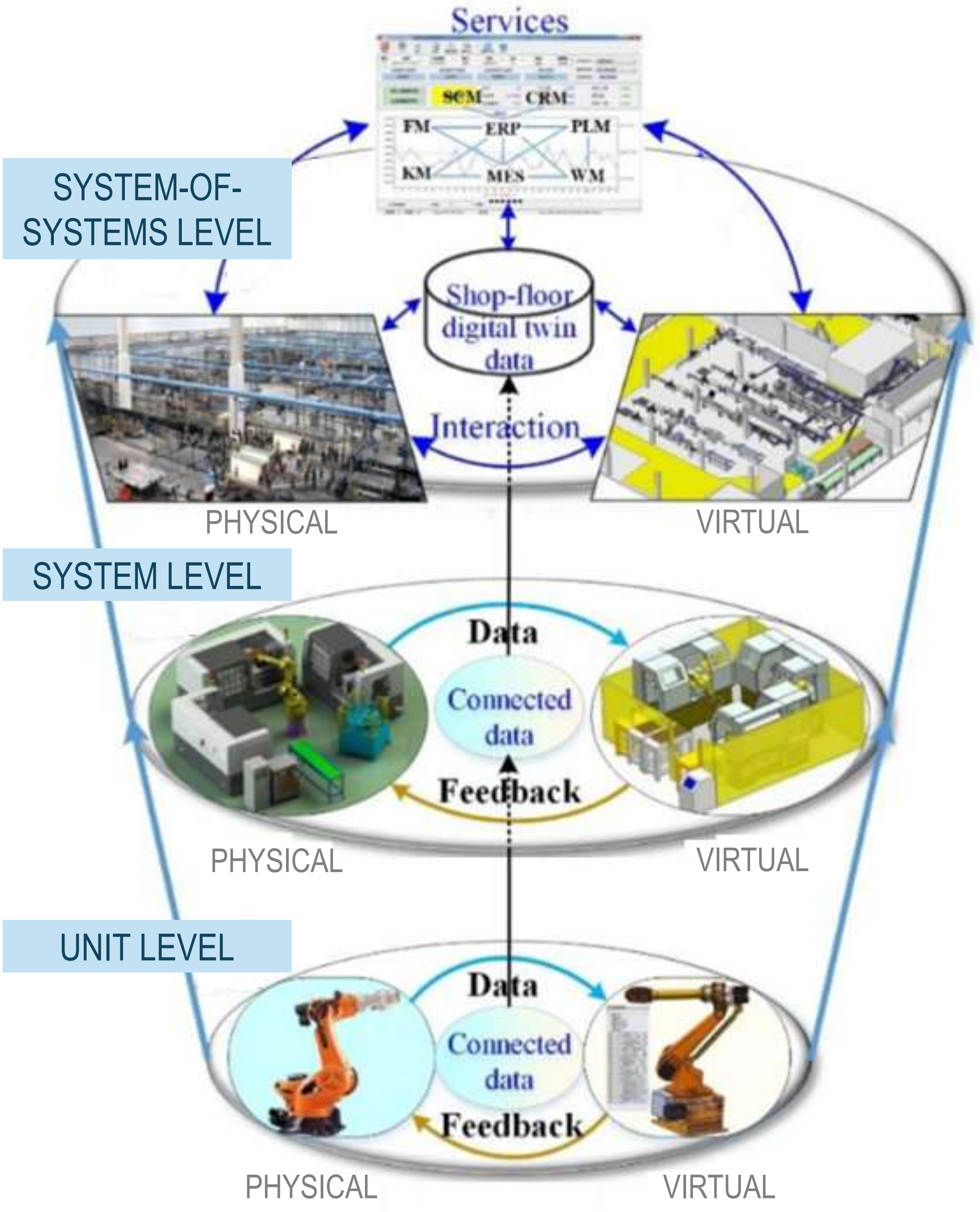
https://www.youtube.com/watch?v=s4Ztp_d_q54

Modeling the Virtual Entity

No generic model

Different types of models, depending on the scope and on the application:

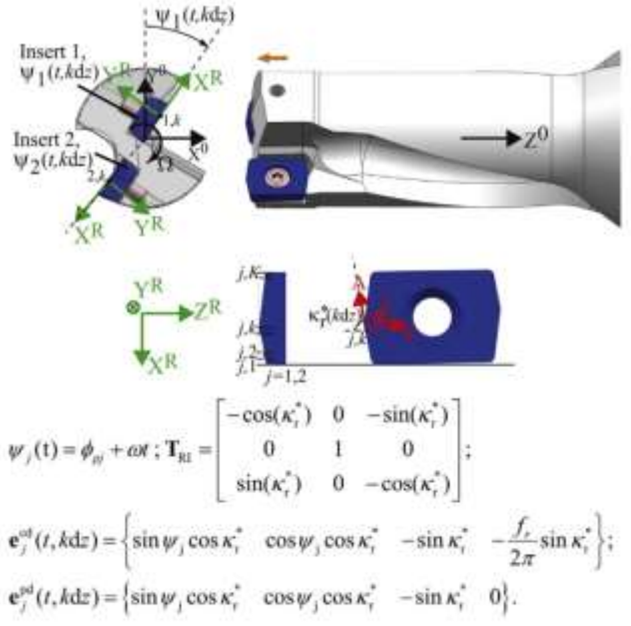
- *Unit level:* equipment and production line, basic closed loop
- *System level:* emphasis on interconnection of components and on collaboration, geographically concentrated
- *System of Systems level:* enterprise-wide integration during entire product life-cycle, multiple production lines, supply chain

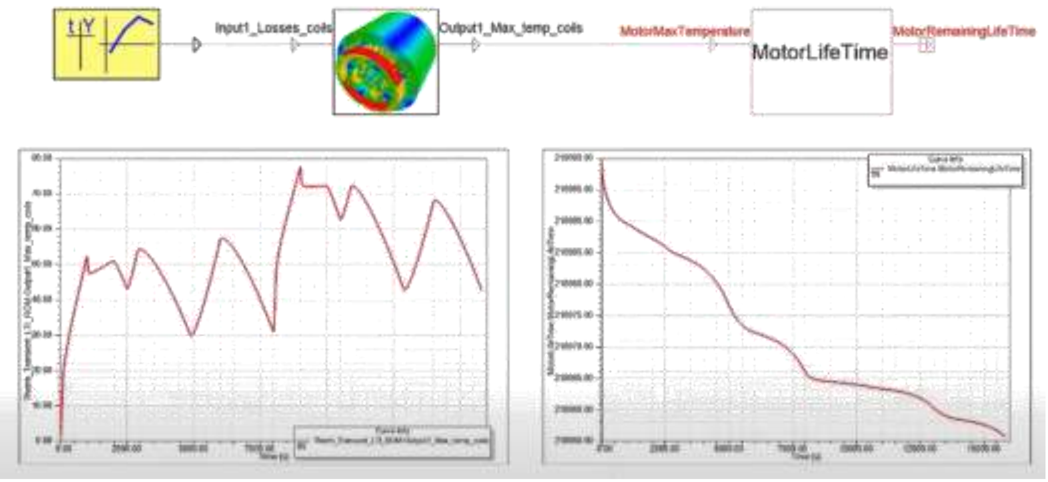


Modeling the Virtual Entity

Sound mathematical models of multi-physical physical processes

- Multiphysics simulation, e.g., fluid, structure, aero-acustic, thermodynamics
- Integrate existing models by adding the closed loop with the physical counterpart – no need to invent new ones
- Include 3D solid modeling to describe the geometric appearance of the physical entity
- Limited to the design phase of man made objects/systems
  - Prone to numerical instability and too computationally demanding
  - Underutilized unless their computational efficiency is improved



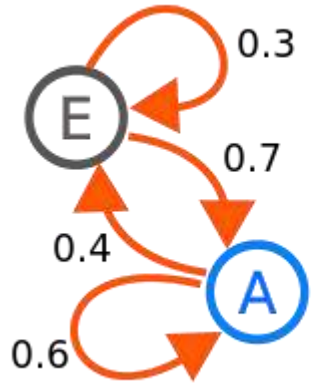


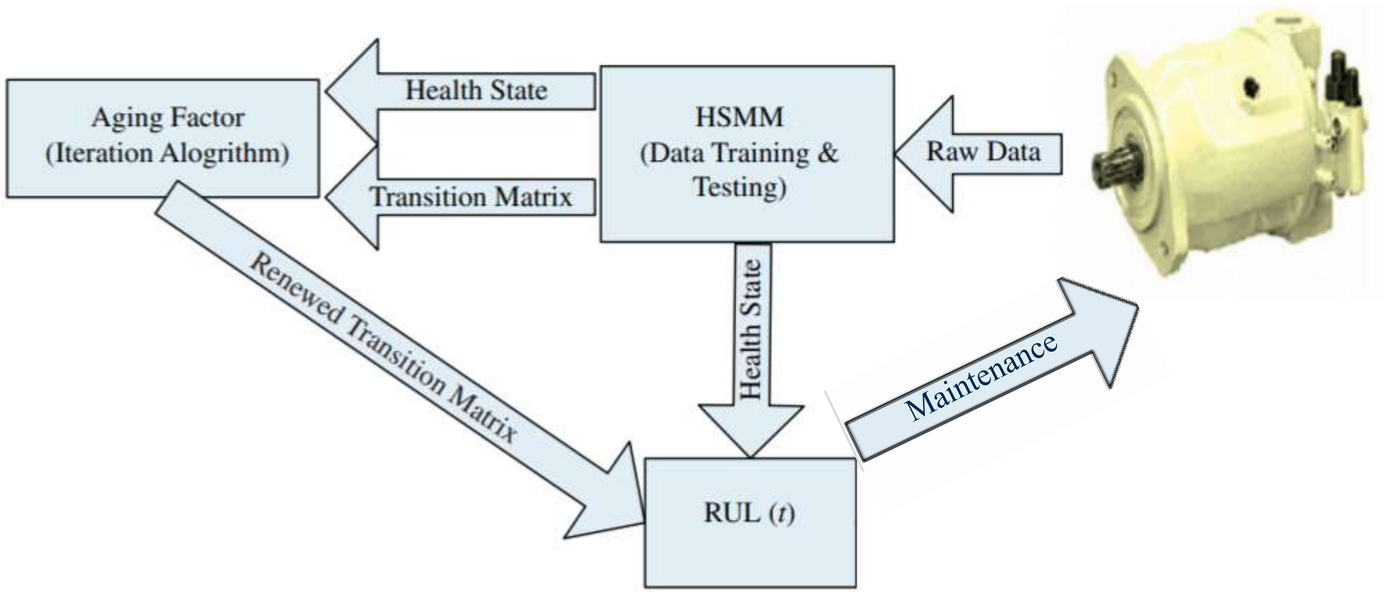
Y. Altintas, P. Kersting, D. Biermann, E. Budak, B. Denkena, I. Lazoglu, Virtual process systems for part machining operations, CIRP Annals, Volume 63, Issue 2, 2014, Pages 585-605.

Modeling the Virtual Entity

Statistical models like Markov chains and Bayesian networks

- System as sequence of possible events in which the probability of each event depends only on the state attained in the previous event



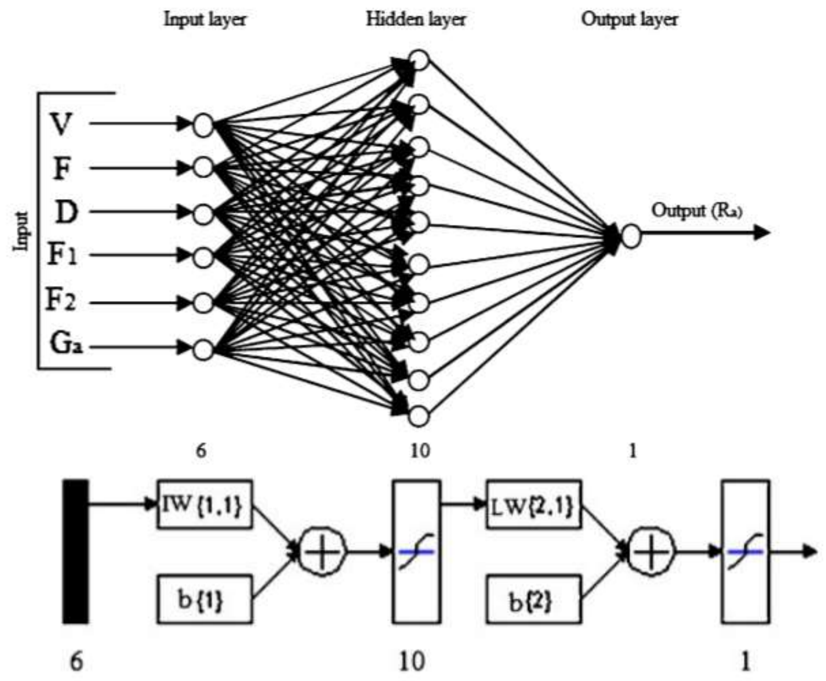


Y. Altintas, P. Kersting, D. Biermann, E. Budak, B. Denkena, I. Lazoglu, Virtual process systems for part machining operations, CIRP Annals, Volume 63, Issue 2, 2014, Pages 585-605.

Modeling the Virtual Entity

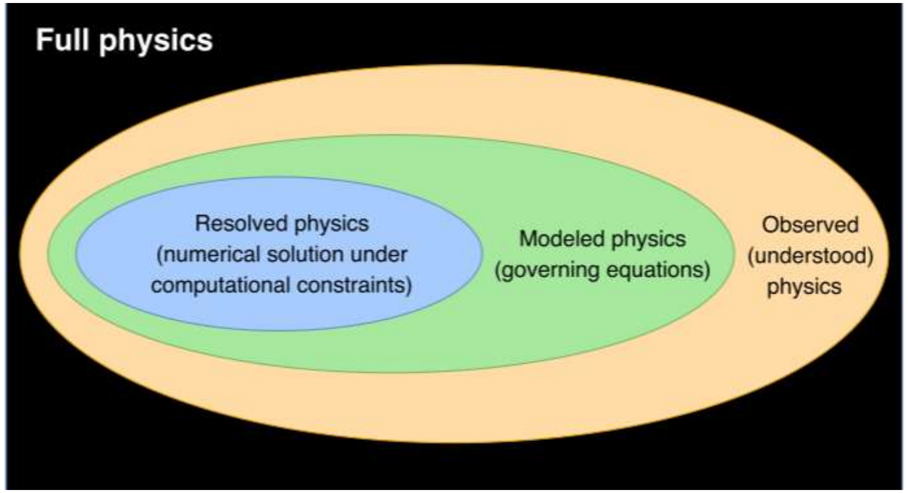
Data-driven modeling

- Use machine learning libraries and high performance architectures to infer models from data
  - Collects real-time data and historical data for model training, model verifying, and model updating
- Exploit available data (e.g., sensor data) to infer a model automatically
  - User does not need detailed knowledge of the physical phenomenon (that may be too complex to model mathematically)
  - Model keeps on improving as more data is collected
- Most advanced learning approaches are too complex to be interpretable
  - Not acceptable in critical contexts (e.g., healthcare)

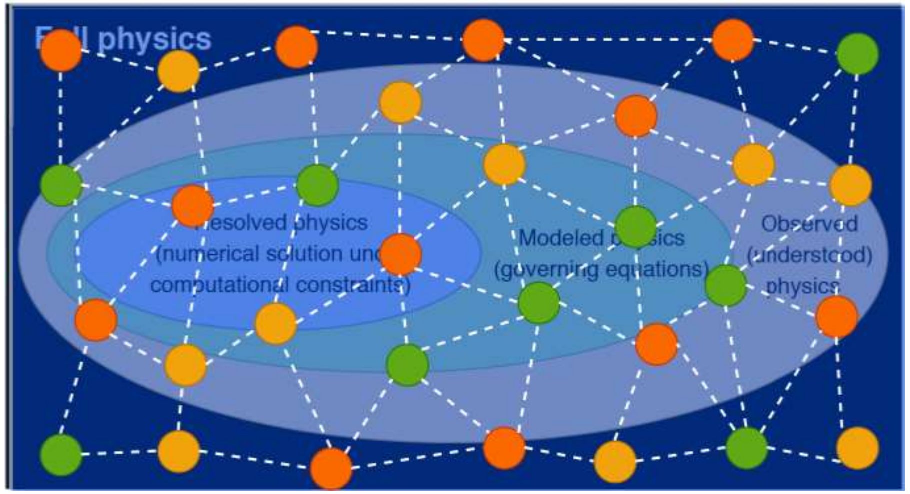


Subbiah, Palani. Prediction of surface roughness in CNC end milling by machine vision system using artificial neural network based on 2D Fourier transform. Int. Journal of Advanced Manufacturing Technology
A. Rasheed, O. San and T. Kvamsdal, "Digital Twin: Values, Challenges and Enablers From a Modeling Perspective," in IEEE Access, vol. 8, pp. 21980-22012, 2020, doi: 10.1109/ACCESS.2020.2970143.

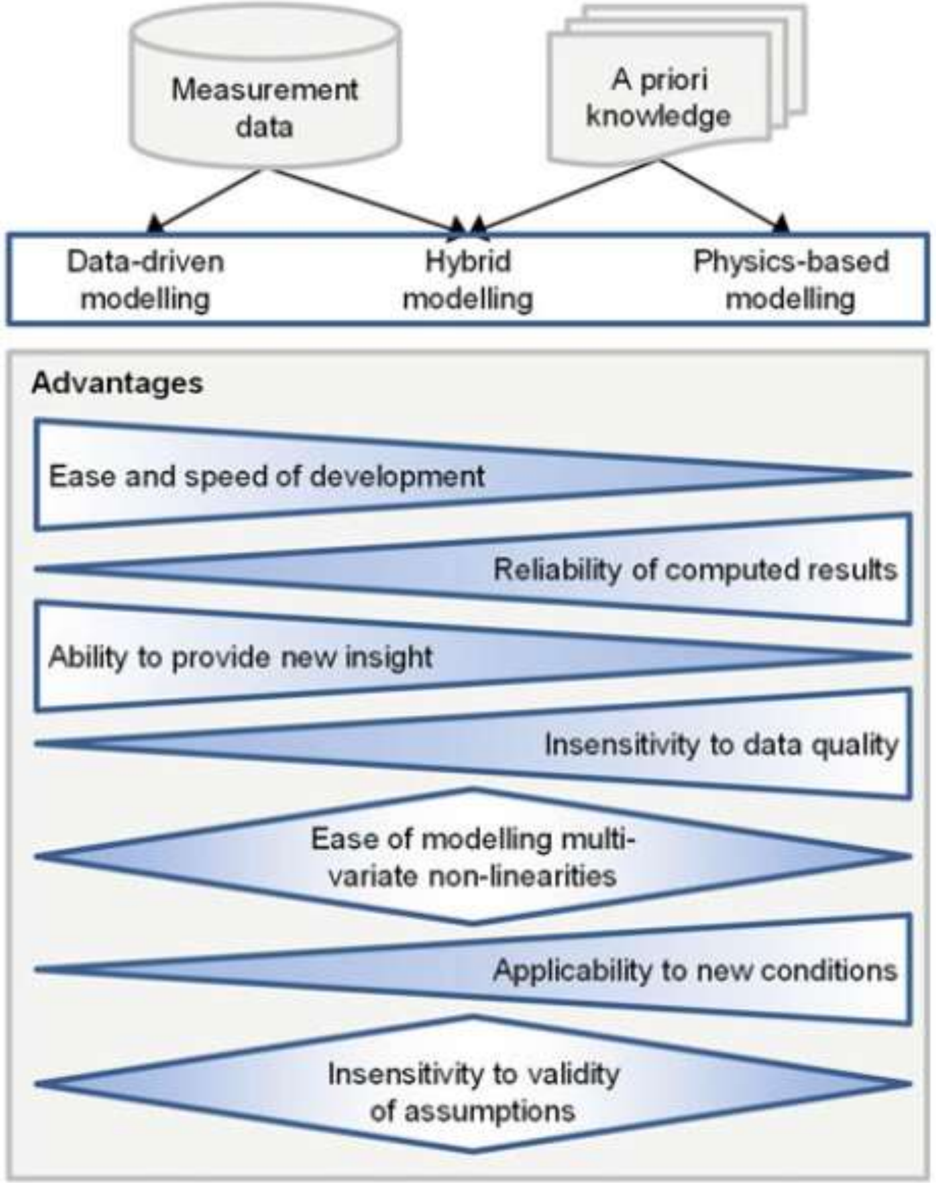
Modeling the Virtual Entity



A-PRIORI KNOWLEDGE



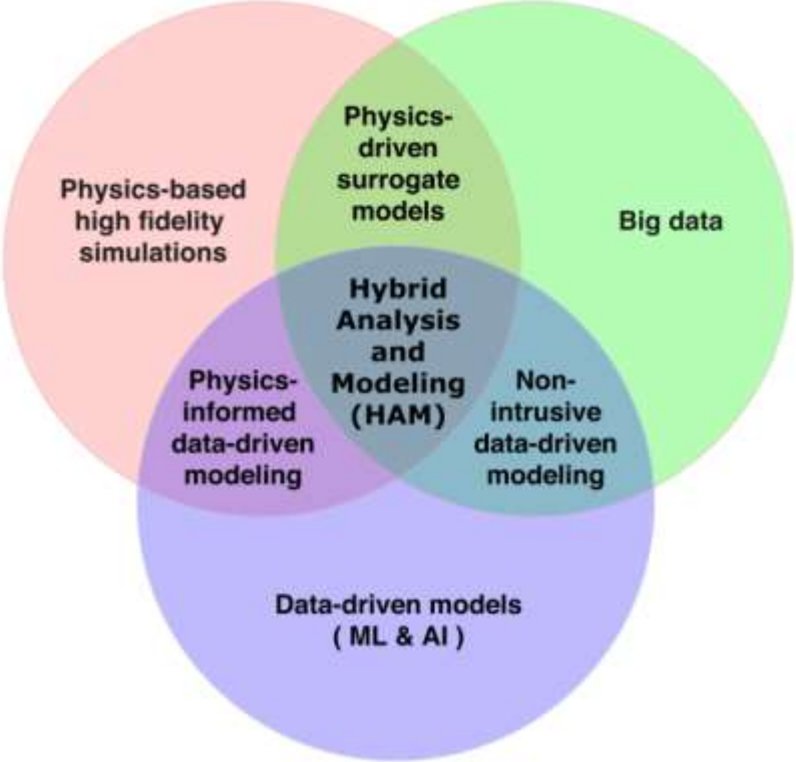
MEASUREMENT DATA

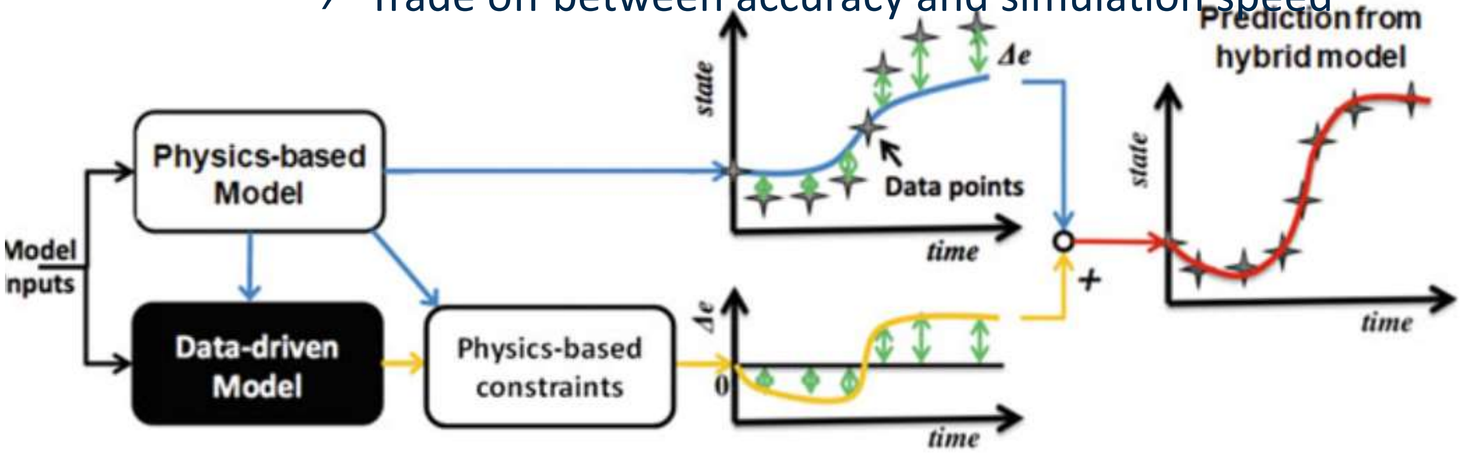


Modeling the Virtual Entity

Possible future direction: Hybrid Analysis and Modeling

- Combine physics-based models with data-driven intelligent models
- Remove shortfalls of both approaches
  - Interpretability, robust foundation of the former
  - Accuracy, efficiency, automatic pattern capabilities of the latter
- Fitting models to data, model order reduction, replacement of equations with ML, physics-informed ML, …
  - Trade off between accuracy and simulation speed





Modeling the Virtual Entity

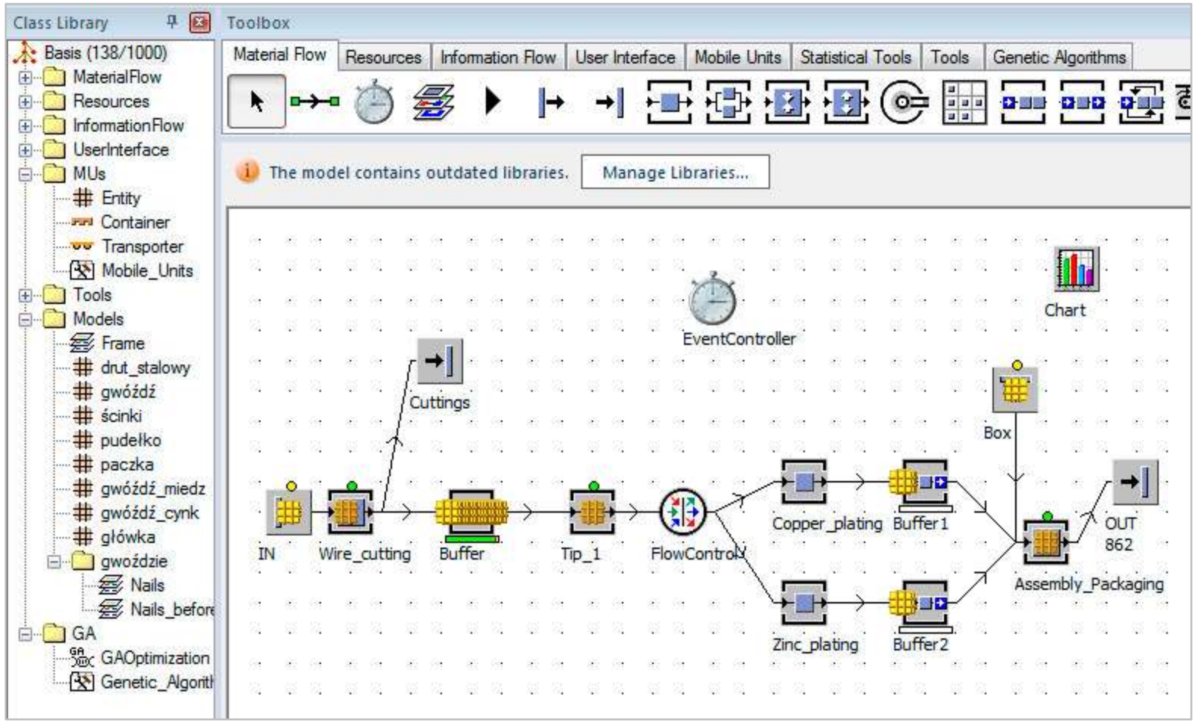
It may be necessary to enlarge the scope from single component to production line…

- E.g., line and space optimization, occupancy simulation and optimization, supply chain monitoring
- Merge models of different aspects of the shop-floor, e.g., geometry, multi-physics, behaviors and rules, etc.
  - Correlation, mapping and fusion of different models with mutual influence
- Useful simulation tools that can be connected in closed loop with the physical production line like Plant Simulation
  - Currently used for static digital twins
  - Extend the simulation tool with communication to and from the physical line

Modeling the Virtual Entity

System level (production process or plant)

MONITORING AND OPTIMIZATION OF PART FLOW, ENERGY CONSUMPTION, LINE ARCHITECTURE

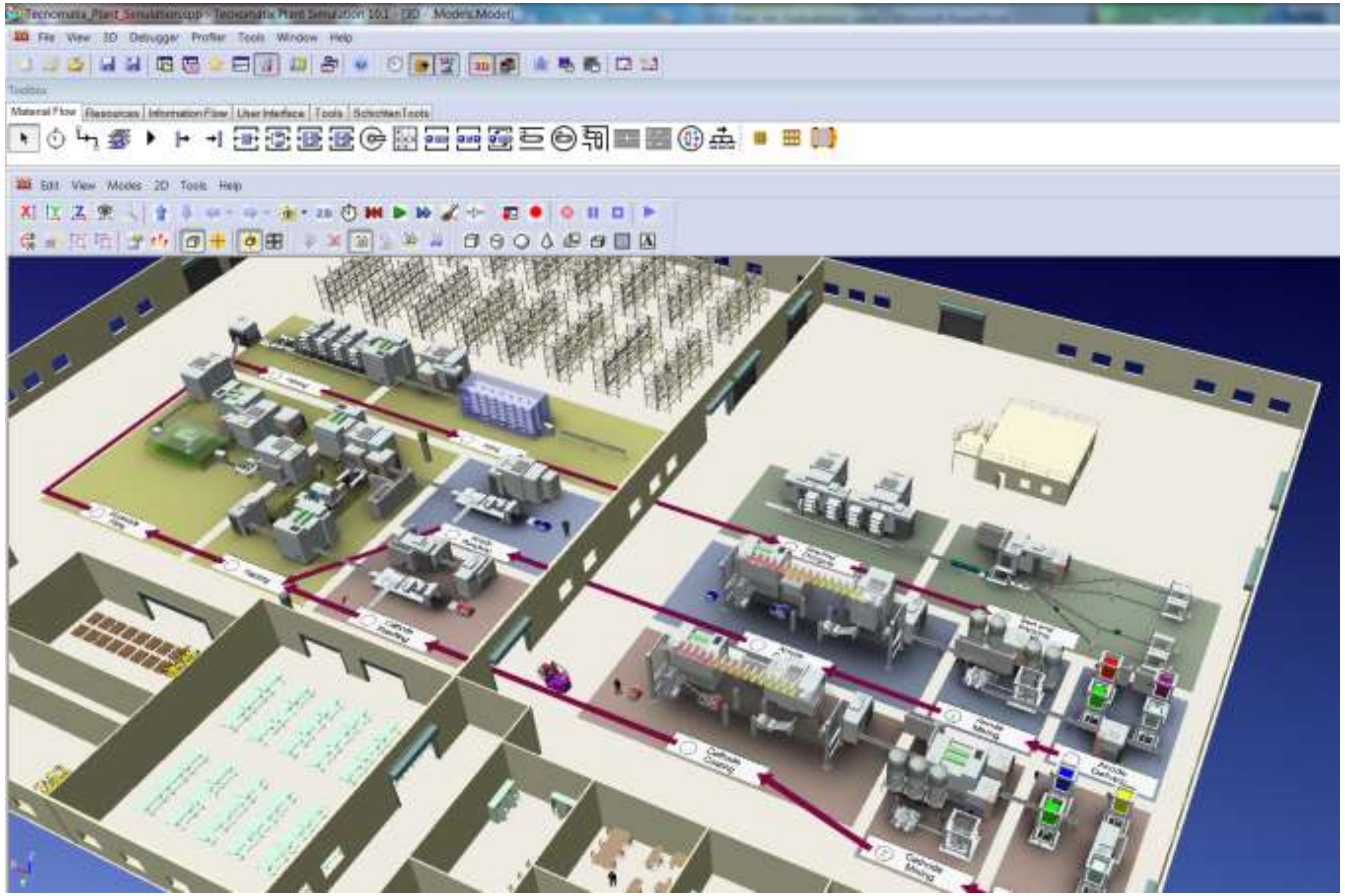


MONITORING AND OPTIMIZATION OF PRODUCTION LINE

Modeling the Virtual Entity

System-of-Systems level

- Enterprise-wide view
- Less detailed model of single processes and equipment
- Global view of production, supply chain, etc.



ENTERPRISE VIEW

Modeling the Virtual Entity

Models do not have to be new per se…

- Can use existing models at any level of detail
- Exploit the expertise and the available tools

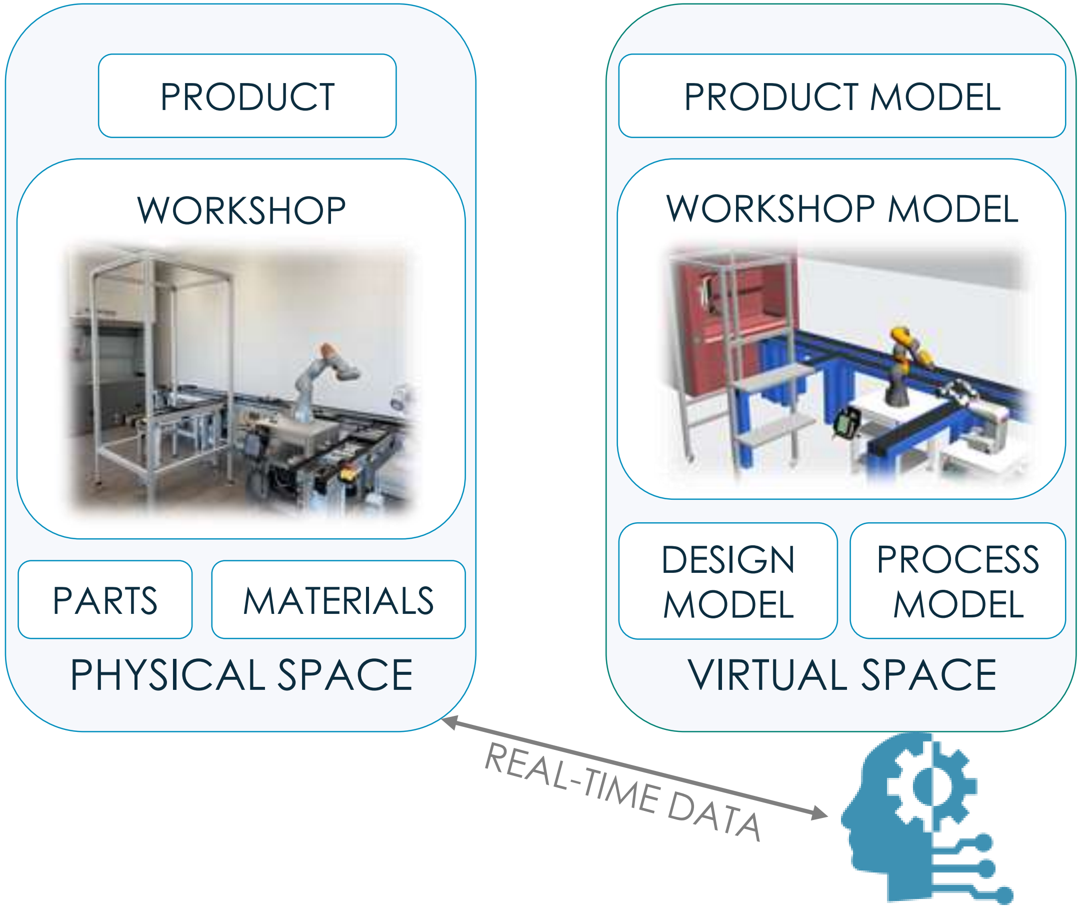
Models and simulations must be enriched with communication with the physical entity and possibility to react to evolving conditions of the plant

- E.g., parameters are not constant and user defined but rather inputs (or elaborated from inputs) coming from the physical plant

Enabling technology: artificial intelligence (AI)

Digital twins must be able to run analytics in real time or faster, provide a high degree of prediction accuracy and integrate data from a collection of disparate and often incompatible sources

- Meeting these goals lies beyond the reach of traditional design and simulation technologies
- Artificial intelligence to create models based on observed behavior and historical data rather than just the design information



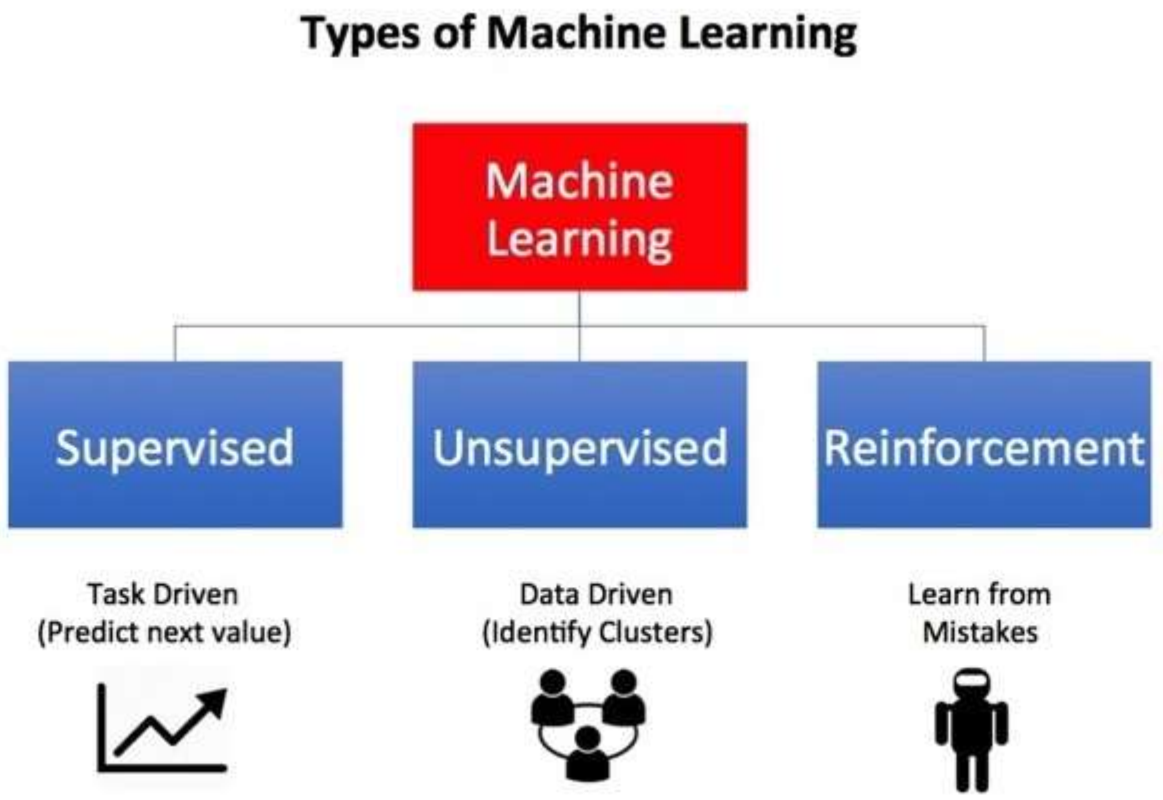
Enabling technology: artificial intelligence (AI)

…and in particular Machine Learning

- Ability of IT systems to independently find solutions to problems by recognizing patterns in data

Key idea: use real-time data and historic data for model training, model verifying, and model updating

- Real potential exploited now that massively parallel architectures allow efficient computation
- Computers are outperforming humans in many (even creative) tasks



Artificial Intelligence and Digital Twins

The digital twin continuously collects real-time data from the physical production line, to utilize real-time and historic data for model training, model verifying, and model updating

› Model is iteratively trained and optimized based on continuously updated and accumulated data

› Adapt to the continuous changes in the real factory environment

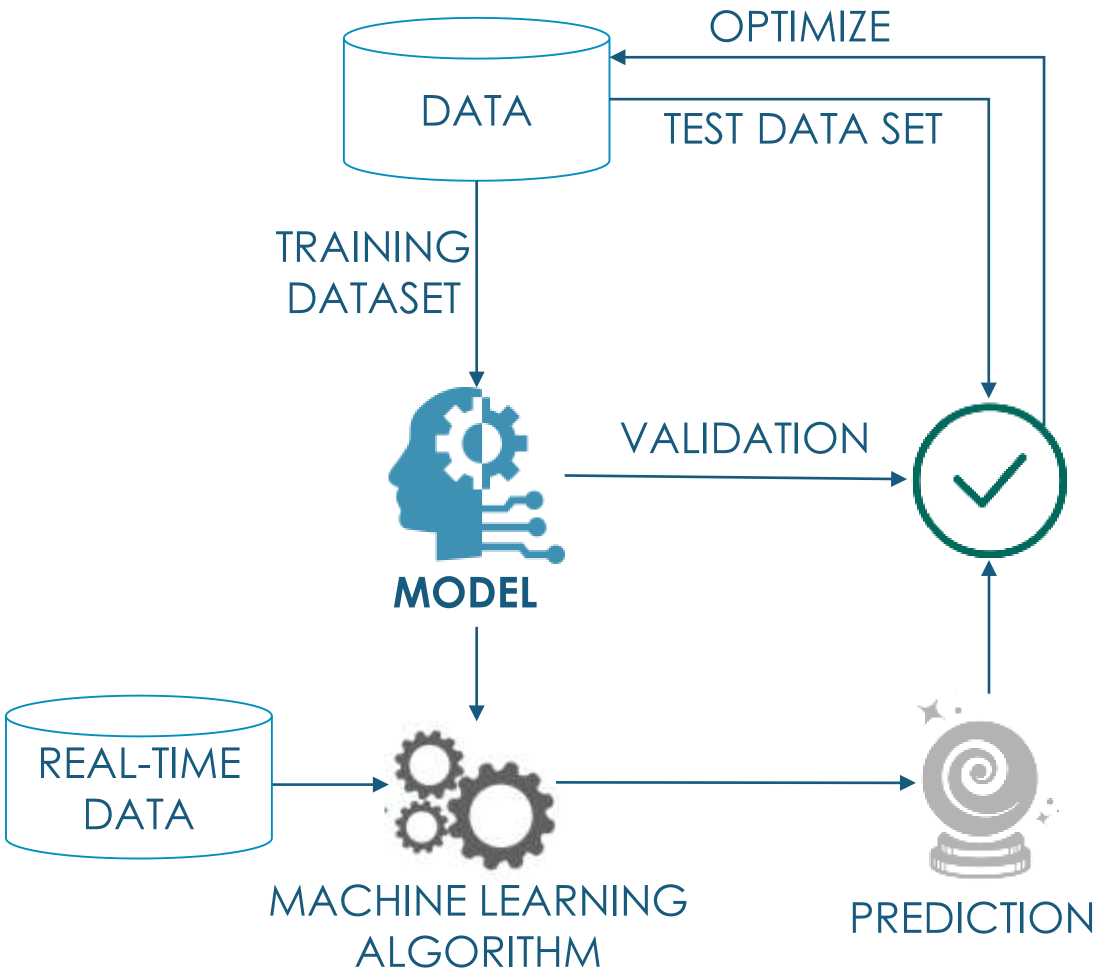
Still ongoing research

› AI systems to make predictions when data is abundant and the processes being evaluated are relatively simple

› <u>Lack of data is a critical issue</u> – data collection campaigns are necessary

Artificial Intelligence and Digital Twins

1. Data collection and cleaning
   - Data might be incomplete, inconsistent or contain errors and missing value
   - Requires construction of a proper dataset
   - Include potential defects and failure data, usually not available
2. Data is used to train the model
   - Training data set used to learn specific parameters over the training period that will minimize prediction error
   - Try different machine learning algorithms, e.g., random forest, neural networks
3. Evaluation on the test data set



Benefits of Artificial Intelligence with Digital Twins

Benefits:

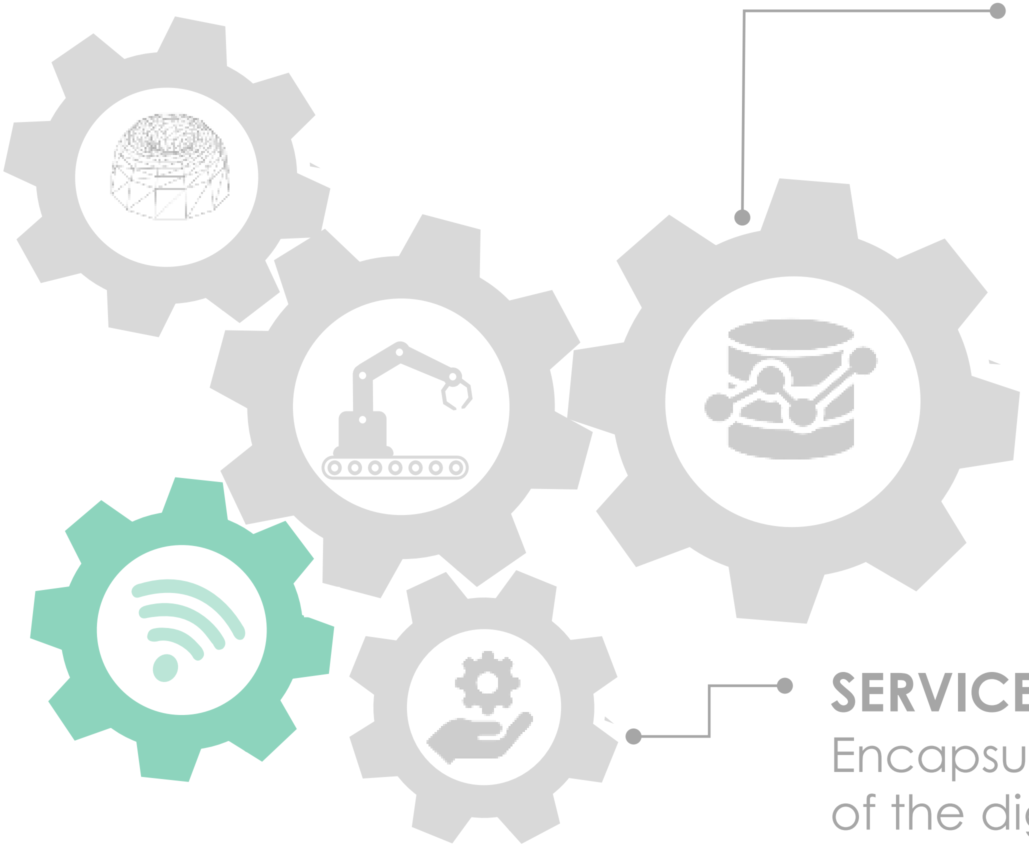
- Accurate and agile production control
- Help to reduce the cost of inefficient production, improve the economic benefits and enhance sustainability

Limitations:

- Few investigations on the theory, approach, process and guidelines of implementation of machine learning and digital twin in manufacturing
- Still ongoing research, but for sure a promising direction!

Parts of a Digital Twin

MODEL
Digital companion is made of a set of models
Reproduce with high fidelity the properties, behaviors and rules of the physical object
Operate autonomously in the virtual space
Ability to predict problems on physical side
Validate performance before system completion



DATA
Physical plus virtual data provides intelligence, e.g., digital model built from sensor data, decision based on simulated data, and operations based on predefined models

CONNECTIONS
Enable closed-loop interaction between digital and physical elements

SERVICE
Encapsulate functions of the digital twin into services of easy and convenient usage

Enabling Connection

It is critical to build a connection between the virtual entity and the physical entity for data exchange

- Backbone of the digital twin

Involves a number of technologies

- Communication protocol analysis
  - Real time communication mechanisms
  - Wireless communication, but not exclusively (e.g., RFID, Ethernet)
- Protocol standardization
  - Crucial to enable interoperability and extensibility
  - Stable interaction and cooperation
- Middleware technology
- Devices for access to communication

Internet of Things (IoT)

Industrial IoT employs a network of sensors to collect critical production data and uses cloud software to turn this data into valuable insights about the efficiency of manufacturing operations

- Visibility into shop floor and field operations
- Visibility into the manufacturing supply chain
- Visibility into remote and outsourced operations

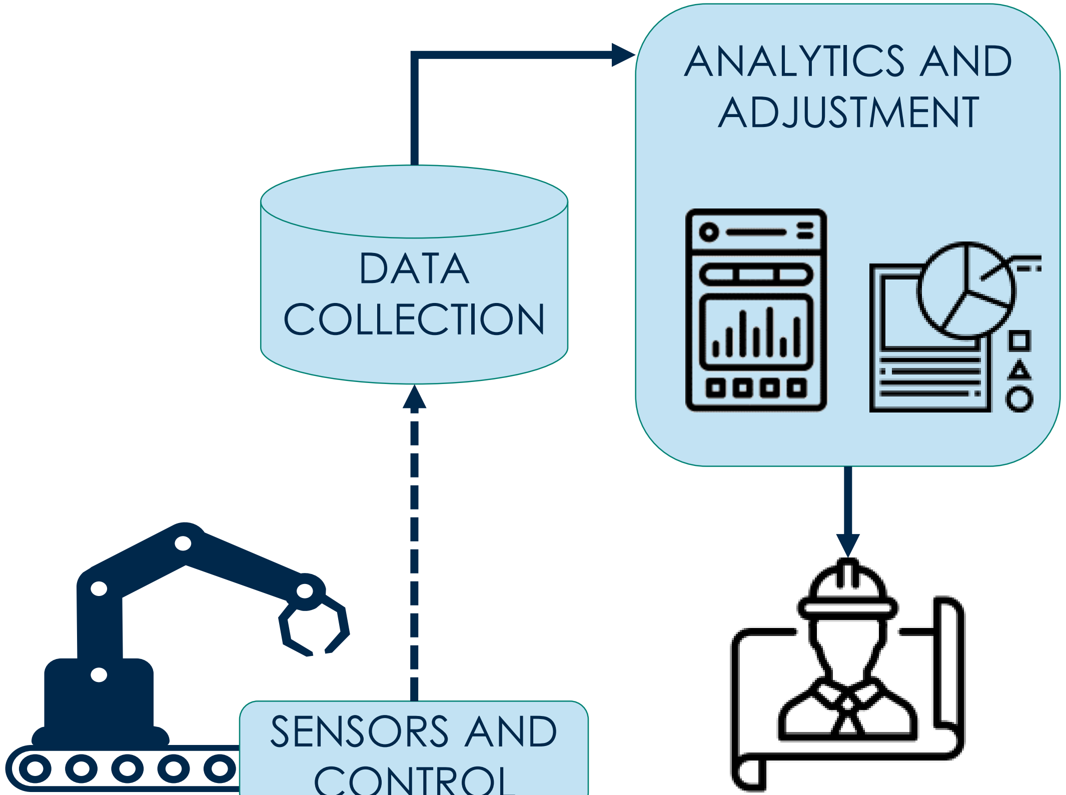
Worldwide numbers:

- 62% of surveyed manufacturing enterprises are already executing digital transformation pilots
- 86% of manufacturers have already adopted IIoT solutions and 84% of them find IIoT extremely effective
- IoT applications in manufacturing are expected to generate $1.2 to $3.7 trillion of economic value annually by 2025

Internet of Things (IoT)

Mainly adopted in industrial contexts for:

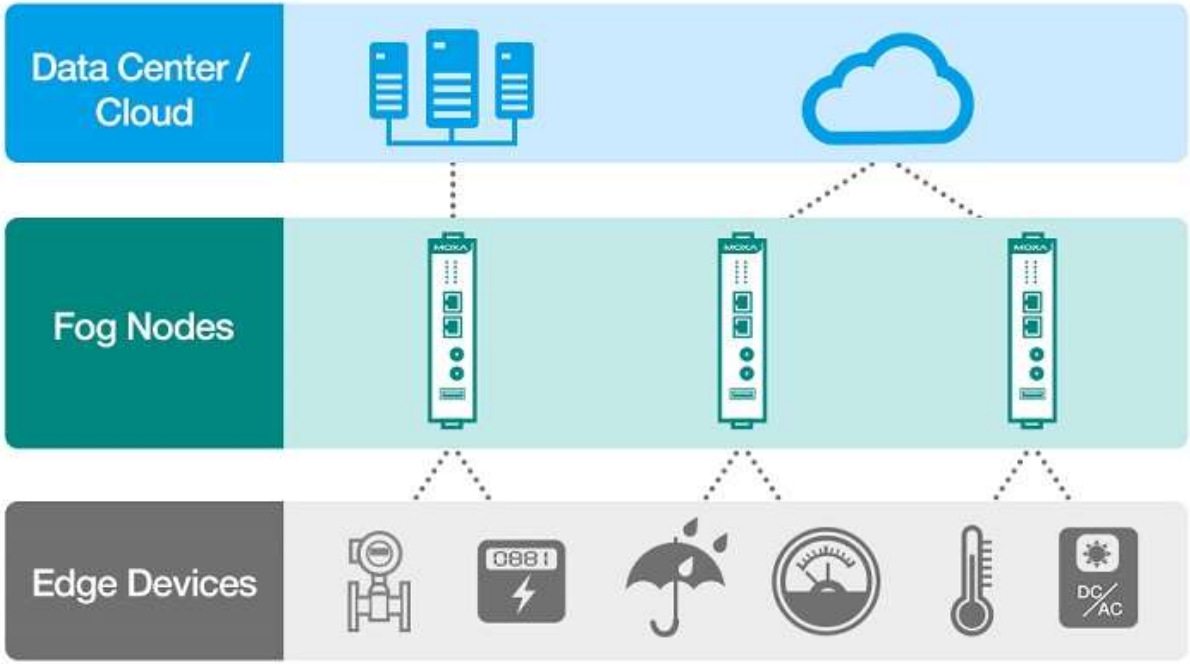
- Cost reduction
  - Optimized asset and inventory management, reduced machine downtime
- Shorter time-to-market
  - Faster and more efficient manufacturing and supply chain operations
- Mass customization
  - Source of real-time data required for thoughtful forecasting, shop floor scheduling and routing
- Improved safety
  - Addresses safety problems in potentially hazardous environments



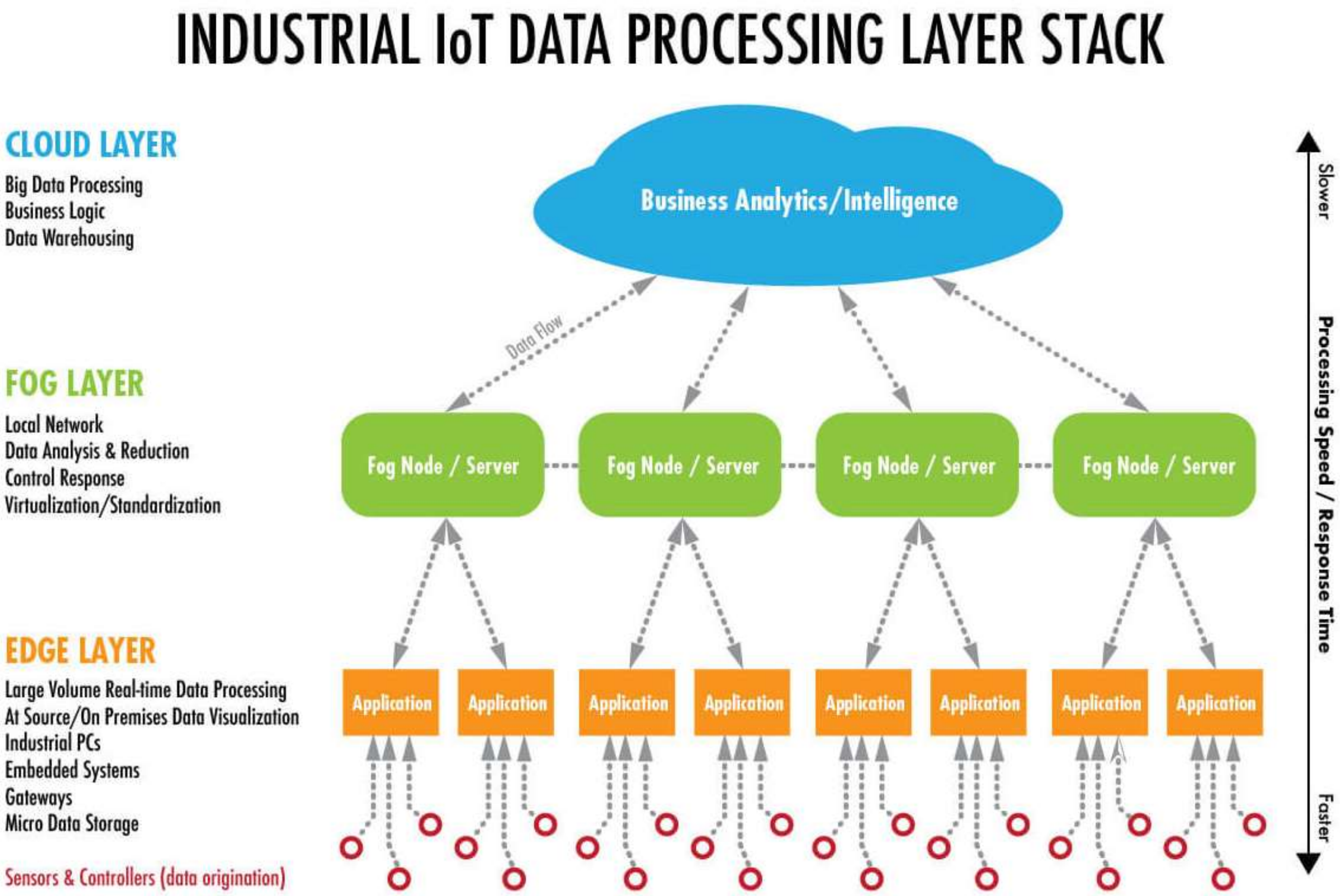
Internet of Things (IoT)

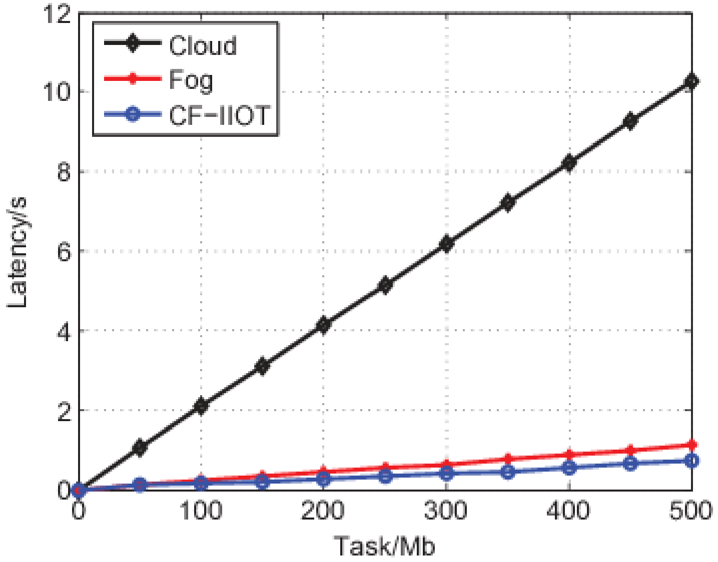
Three different data technologies depending on where data is stored and processed

- Edge computing
  - Computing to the edge of the network (closer to data sources)
  - Allow real-time sensing and actuation and agile connectivity with few computing resources
  - Enhance security and minimize data transfers
- Cloud computing
  - Data stored and processed on servers
  - Ubiquitous convenient on-demand access to shared resources
  - High computing and storage capabilities
  - Security issues: public/private/hybrid
- Fog computing
  - Mix of the two approaches
  - Portions of data stored locally

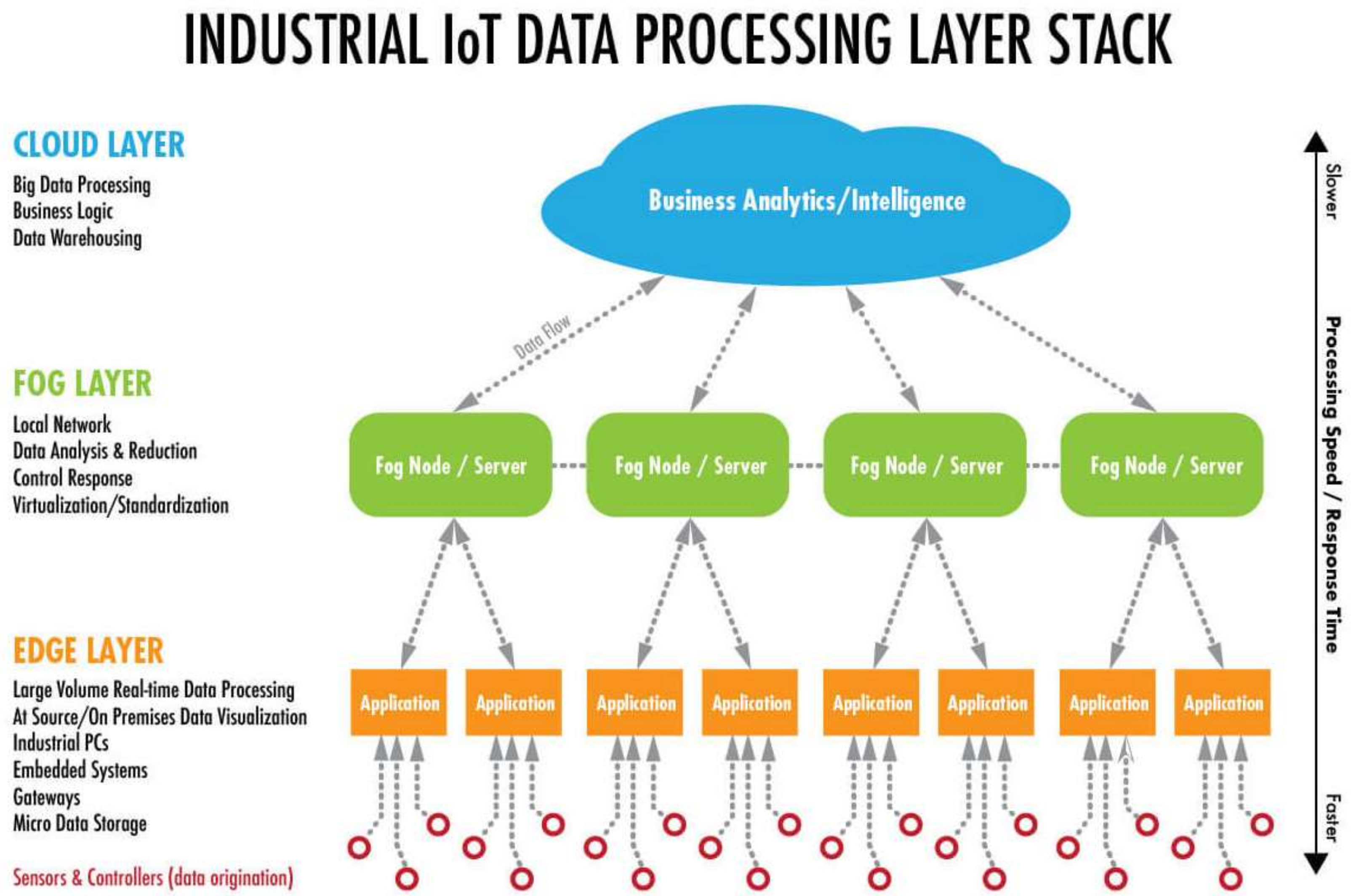


Internet of Things (IoT)





Internet of Things (IoT)



SYSTEM-OF-SYSTEMS LEVEL:
ENTERPRISE-WIDE PERSPECTIVE
MASSIVE DATA
CONECTION TO ERP AND MES

UNIT LEVEL:
BASIC ATTRIBUTES + REAL TIME SENSOR DATA WITH LOW LATENCY
LOW COMPUTING POWER

IIoT Protocols Classification

Category	Protocol/Standard/Bus
IT	REST/HTTP/TLS
	WebSocket
	JMS
	DDS
	XMPP
IoT	Wi-Fi
	4G/LTE
	5G
	MQTT
	CoAP
	AMQP
	Bluetooth
	ZigBee
	RFID/NFC
	LoRa
	NB-IoT
	WirelessHart
	Z-Wave
OT	PROFINET
	Niagara AX
	CAN
	Siemens S7
	EtherNet/IP
	Hart/IP
	BACnet
	Modbus
	OPC-DA
	OPC-UA
	DNP3/IEC-62351
	ANSI C12.22
	IEC-61850/IEC-62351
	IE-60870-5-104/IEC-62351
	IEC-60870-6-503/IEC-62351

INFORMATION TECHNOLOGIES

MANAGES FLOWS OF DIGITAL INFORMATION (HTTP, REST)

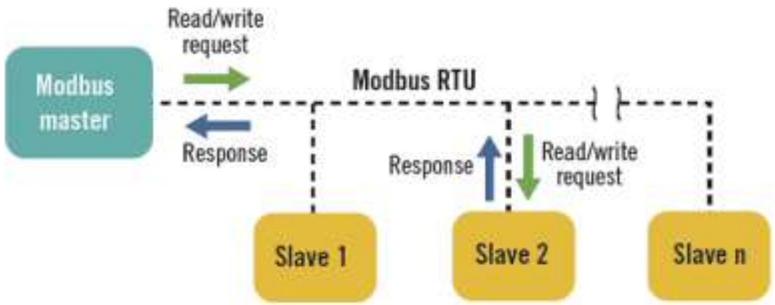
INTERNET OF THINGS

PHYSICAL OBJECTS THAT ARE EMBEDDED WITH SENSORS, PROCESSING ABILITY, SOFTARE, AND TECHNOLOGIES THAT CONNECT AND EXCHANGE DATA WITH OTHER DEVICES AND SYSTEMS OVER COMMUNICATION NETWORKS

OPERATIONAL TECHNOLOGIES

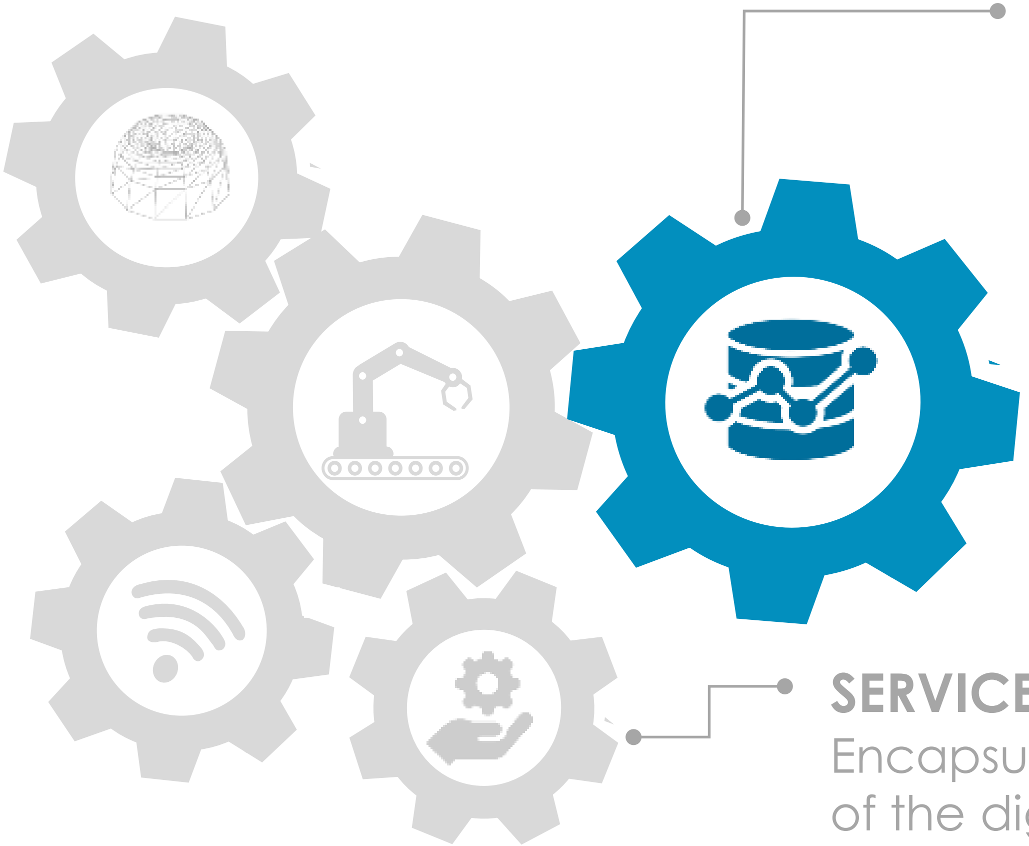
COMPUTING AND COMMUNICATION SYSTEMS TO MANAGE, MONITOR AND CONTROL INDUSTRIAL PHYSICAL DEVICES AND PROCESSES

E.G., MODBUS OPEN STANDARD USED TO COMMUNICATE OVER SERIAL OR ETHERNET WITH PLC, HMI, I/O DEVICES AND SENSORS
MASTER-SLAVE APPROACH
NOT PROTECTED AGAINST CYBERSECURITY ATTACKS OR INTERNAL SECURITY BREACHES



Parts of a Digital Twin

MODEL
Digital companion is made of a set of models
Reproduce with high fidelity the properties, behaviors and rules of the physical object
Operate autonomously in the virtual space
Ability to predict problems on physical side
Validate performance before system completion



DATA
Physical plus virtual data provides intelligence, e.g., digital model built from sensor data, decision based on simulated data, and operations based on predefined models

CONNECTIONS
Enable closed-loop interaction between digital and physical elements

SERVICE
Encapsulate functions of the digital twin into services of easy and convenient usage

Data Management

The first critical aspect is data generation

- Embed the physical entity with sensors
  - Choose the relevant information to be monitored
  - Sensor layout optimization to reduce redundant information, improve utilization rate of sensors, and keep stability of collected data
- Choose communication technology
  - RFID to track and identify entities
  - Wireless sensor network to generate data autonomously for each physical entity, then process and collect data at a central location
- Consider adopting soft sensors to estimate physical entity parameters that are difficult to measure in hardware

Data Management

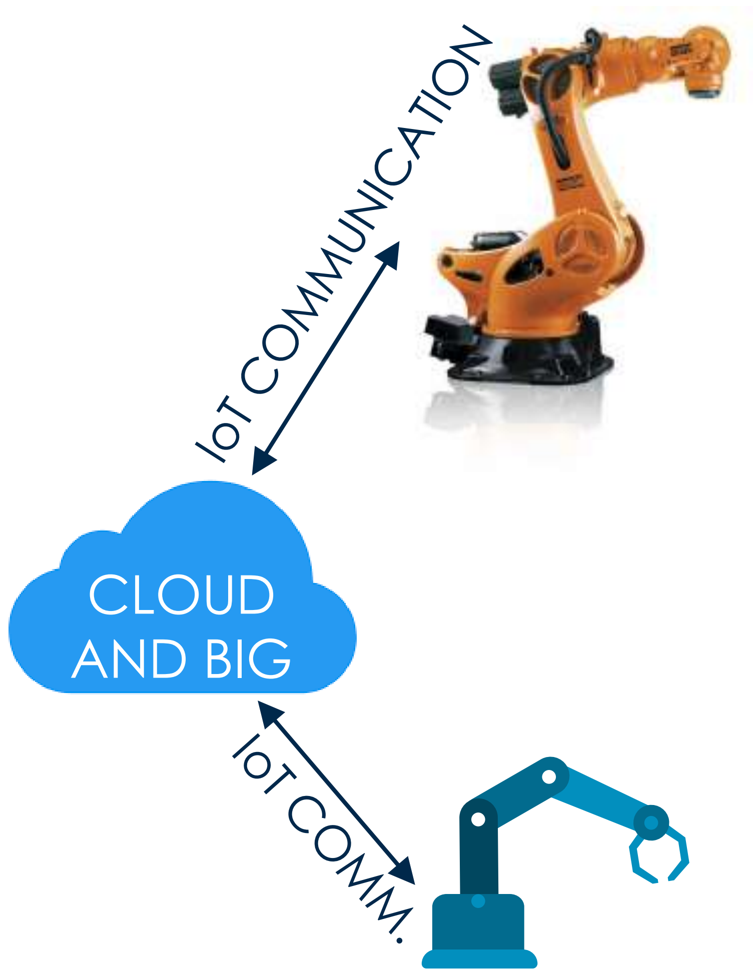
Digital twins must handle a massive amount of data

- Must process a data from a variety of channels: machine, physical environment, virtual space, historical database, etc.

It is necessary to perform data preprocessing to clean data and prepare it for advanced data analysis methods

Standardization is (one of) the keys

- Enable and ease interoperability and extensibility
- Ease data management and cleaning
- Allows to automate conversion between different (standard) formats



Data Management

Data collection

- Issues due to data volume, variety, velocity and heterogeneity
- Deal with dirty data, e.g., duplicated, incorrect, incomplete or delayed

Data pre-processing

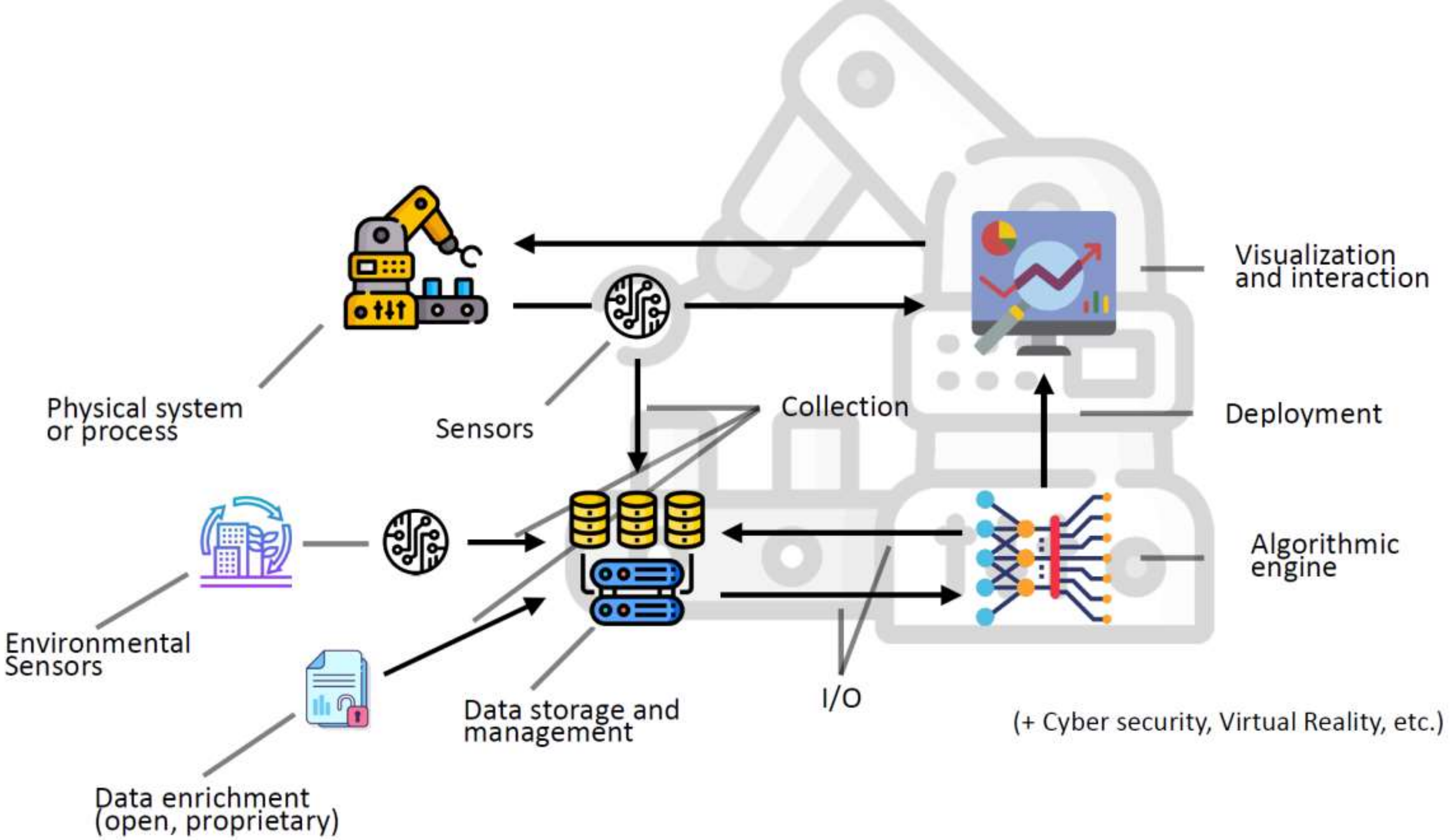
- Cleaning (remove outliers and repeated data) and aggregation (e.g., over time windows)
- Feature extraction and selection

Data elaboration

- Derive understandable and meaningful information from data
- Clustering: discover groups of data with similar behaviors
  - May require manual labeling from a human expert
- Mining: discover patterns in large sets of data (e.g., ML)

Data security

- Confidential and critical data for the company


Visualization
and interaction
Physical system
or process
Sensors
Collection
Deployment
Environmental
Sensors
Algorithmic
engine
Data storage and
management
I/O
(+ Cyber security, Virtual Reality, etc.)
Data enrichment
(open, proprietary)

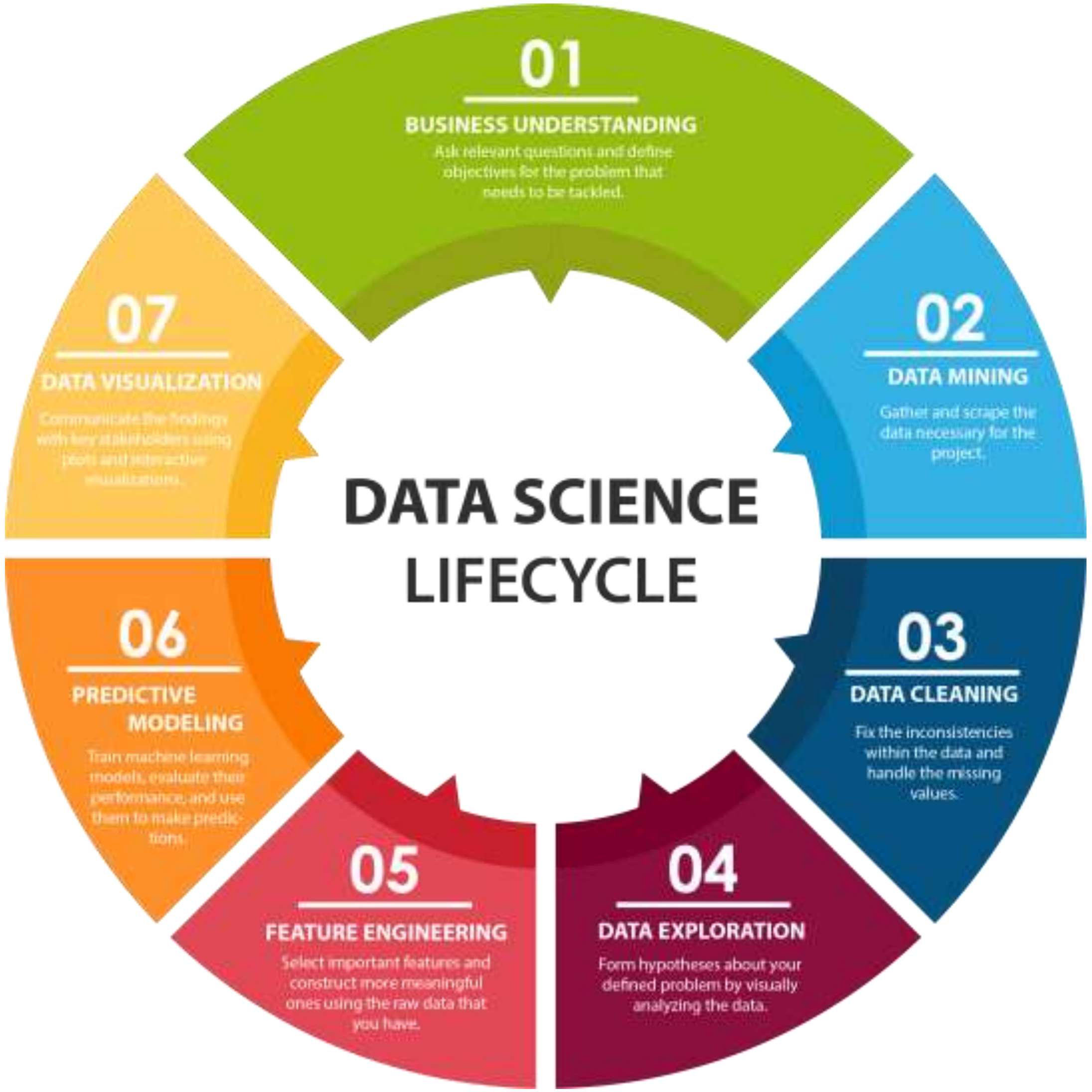
Big Data and Digital Twins

Data driven means circular and iterative

- Start with simple and robust models deeply rooted in the business expertise, and build more complex and refined models step by step

Data driven means that for most complex, real-world scenarios we cannot anticipate exactly the result of a single analysis

- We must follow the scientific method and be prepared to reject hypotheses, but we can also discover new and useful aspects of the process



Time series databases

Database

› Organized collection of data stored and accessed electronically from a computer system

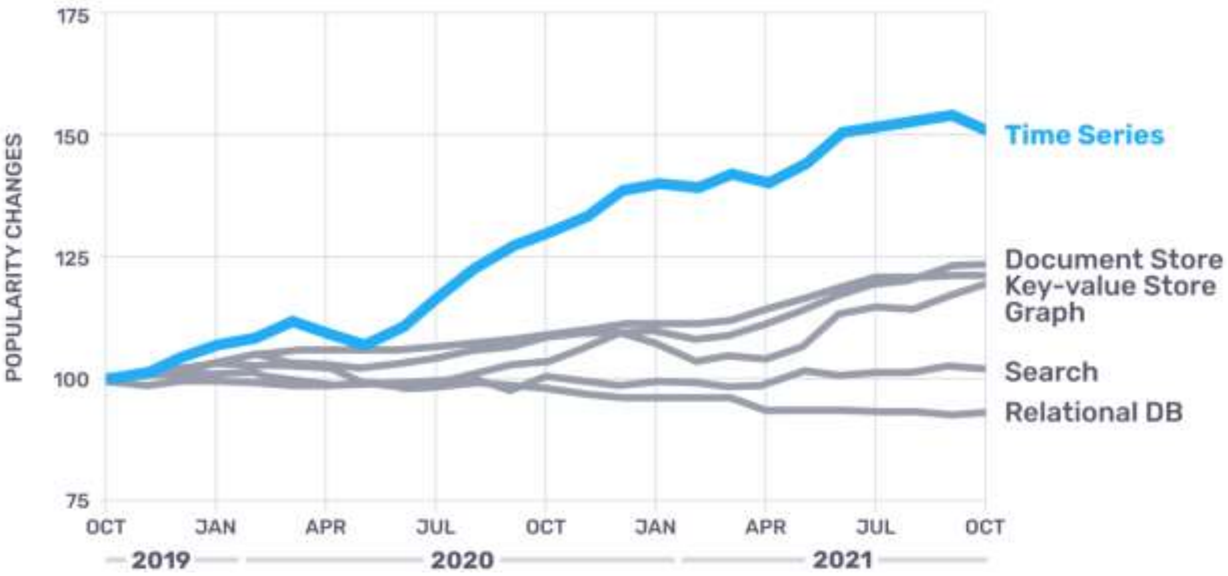
Time series database

› Optimized for time-stamped or time series data: measurements or events that are tracked, monitored, down-sampled, and aggregated over time

› Offer time-stamp data storage and compression, lifecycle management and summarization, ability to handle scans of many records, and time series aware queries

  › Keep high precision data around for a short period of time

  › Request a summary of data over a large time period (e.g., percentile increase)



Time series databases

Cassandra

- Can be used for time series too
- Offers optimizations for storing and retrieving temporal data, such as its reordering according to the timestamp key
- Favours retrieval tasks based on time ranges

MongoDB

- Document-oriented

InfluxDB

- Defined specifically to handle Time Series
- Data is physically ordered by time
- Ability to define, for a given data group, some retention policies, i.e. specific rules to manage old data (e.g., deleting all the data older than three months or replacing them with some aggregated values)
- Possibility to define continuous queries, i.e. a query tool able to work continuously on a stream of input data, rather than in a batch fashion

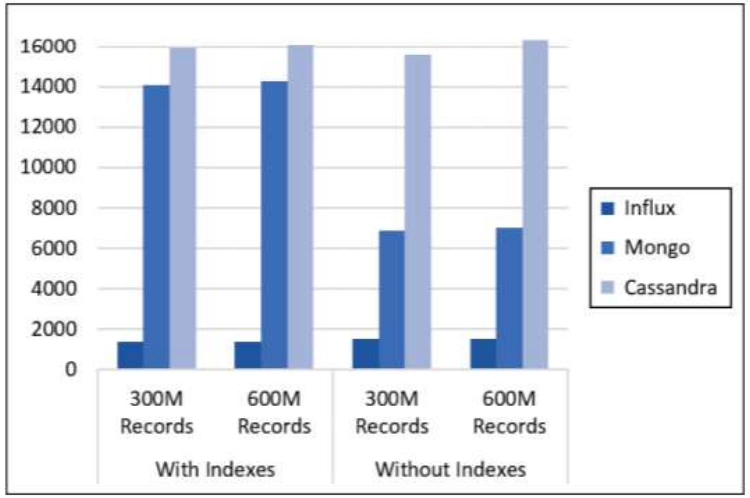


Fig. 1. Time, in seconds, to import 300M and 600M records for each DBMS, with and without indexes.

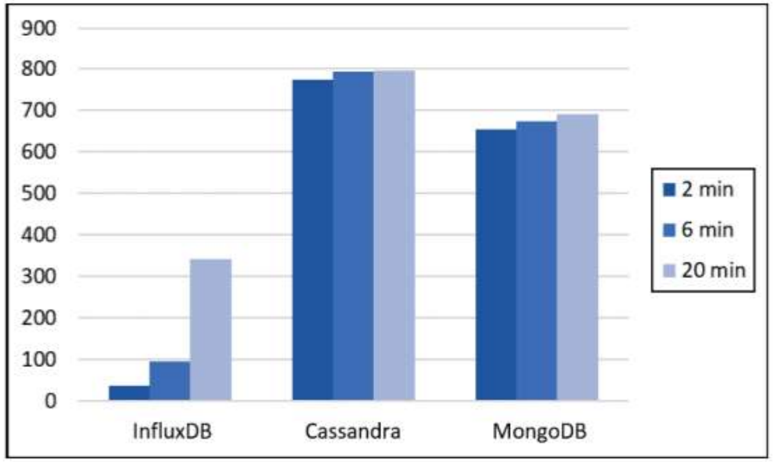
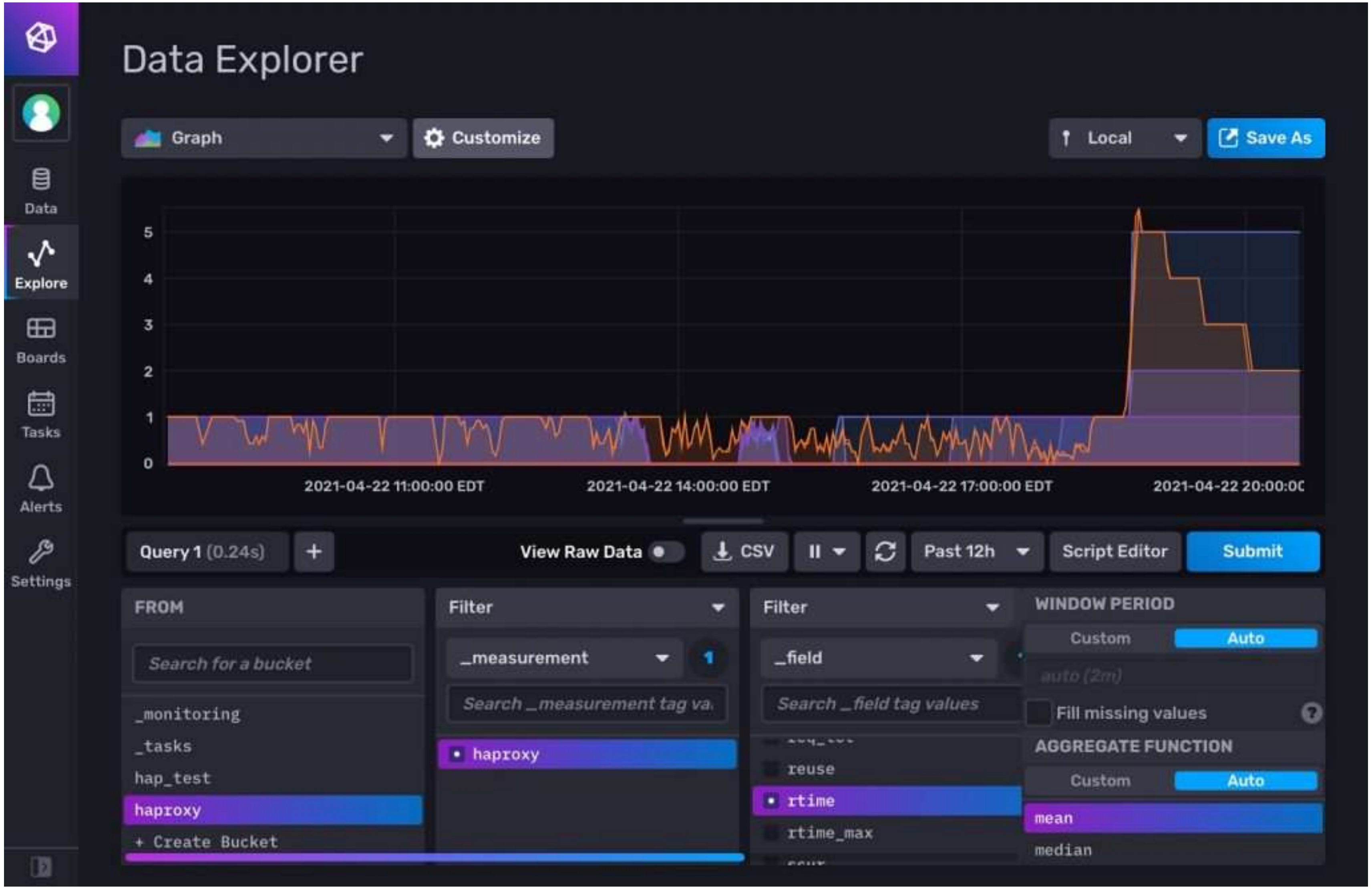


Fig. 2. Average Execution Time of Queries with Time-based Filtering, in Seconds, for three different amounts of data.

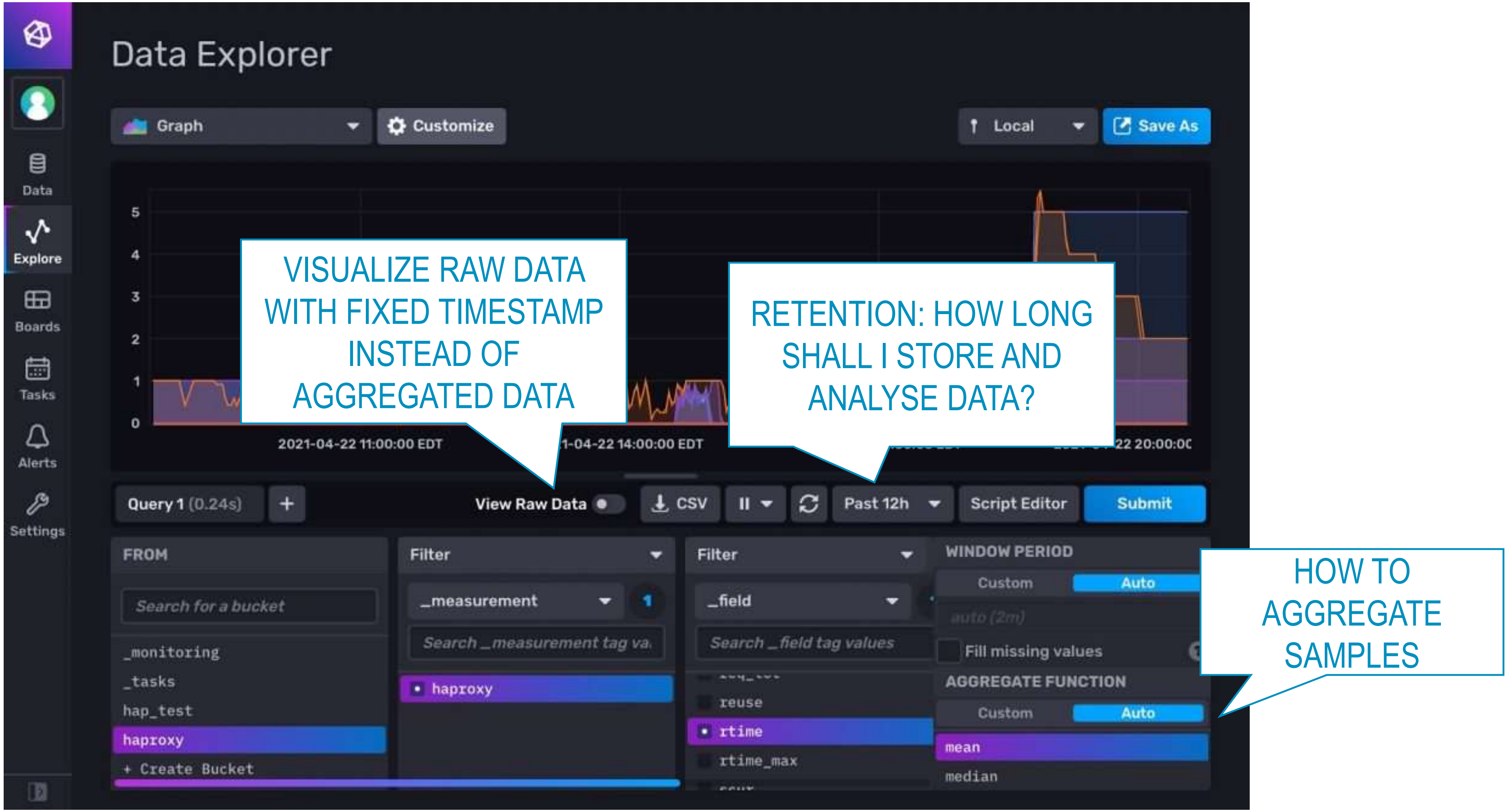
S. Di Martino, L. Fiadone, A. Peron, A. Riccabone and V. N. Vitale, "Industrial Internet of Things: Persistence for Time Series with NoSQL Databases," IEEE WETICE, 2019, pp. 340-345

InfluxDB



Data Explorer
Graph
Customize
Local
Save As
Data
Explore
Boards
Tasks
Alerts
Settings
5
4
3
2
1
0
2021-04-22 11:00:00 EDT
2021-04-22 14:00:00 EDT
2021-04-22 17:00:00 EDT
Query 1 (0.24s)
View Raw Data
CSV
Past 12h
Script Editor
Submit
FROM
Search for a bucket
_monitoring
_tasks
hap_test
haproxy
+ Create Bucket
Filter
_measurement
Search _measurement tag va
haproxy
Filter
_field
Search _field tag values
reuse
rtime
rtime_max
WINDOW PERIOD
Custom
Auto
auto (2m)
Fill missing values
AGGREGATE FUNCTION
Custom
Auto
mean
median

InfluxDB


Data Explorer
Graph
Customize
Local
Save As
VISUALIZE RAW DATA WITH FIXED TIMESTAMP INSTEAD OF AGGREGATED DATA
RETENTION: HOW LONG SHALL I STORE AND ANALYSE DATA?
HOW TO AGGREGATE SAMPLES
2021-04-22 11:00:00 EDT
Query 1 (0.24s)
View Raw Data
CSV
Past 12h
Script Editor
Submit
FROM
Search for a bucket
_monitoring
_tasks
hap_test
haproxy
+ Create Bucket
Filter
_measurement
Search _measurement tag va
haproxy
Filter
_field
Search _field tag values
reuse
rtime
rtime_max
WINDOW PERIOD
Custom
Auto
auto (2m)
Fill missing values
AGGREGATE FUNCTION
Custom
Auto
mean
median
Data
Explore
Boards
Tasks
Alerts
Settings

Parts of a Digital Twin

MODEL

Digital companion is made of a set of models

Reproduce with high fidelity the properties, behaviors and rules of the physical object

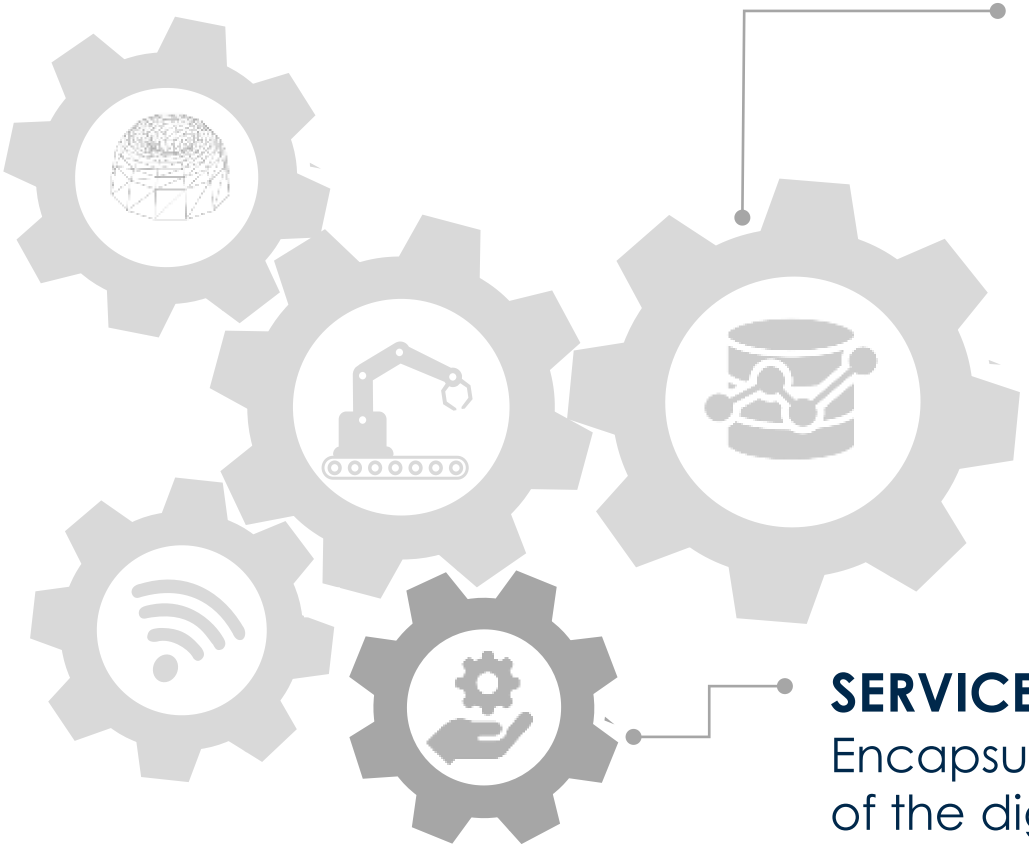
Operate autonomously in the virtual space

Ability to predict problems on physical side

Validate performance before system completion

CONNECTIONS

Enable closed-loop interaction between digital and physical elements



DATA

Physical plus virtual data provides intelligence, e.g., digital model built from sensor data, decision based on simulated data, and operations based on predefined models

SERVICE

Encapsulate functions of the digital twin into services of easy and convenient usage

Service Modeling

Define services, i.e., functions with friendly interface for easy and on-demand usage

- What do I want to report to the user? And how?
- How to make information readable?
- Support to decision making

Enable convenient access to digital twin data and functionality

- Bridge between the manufacturing physical space and virtual space

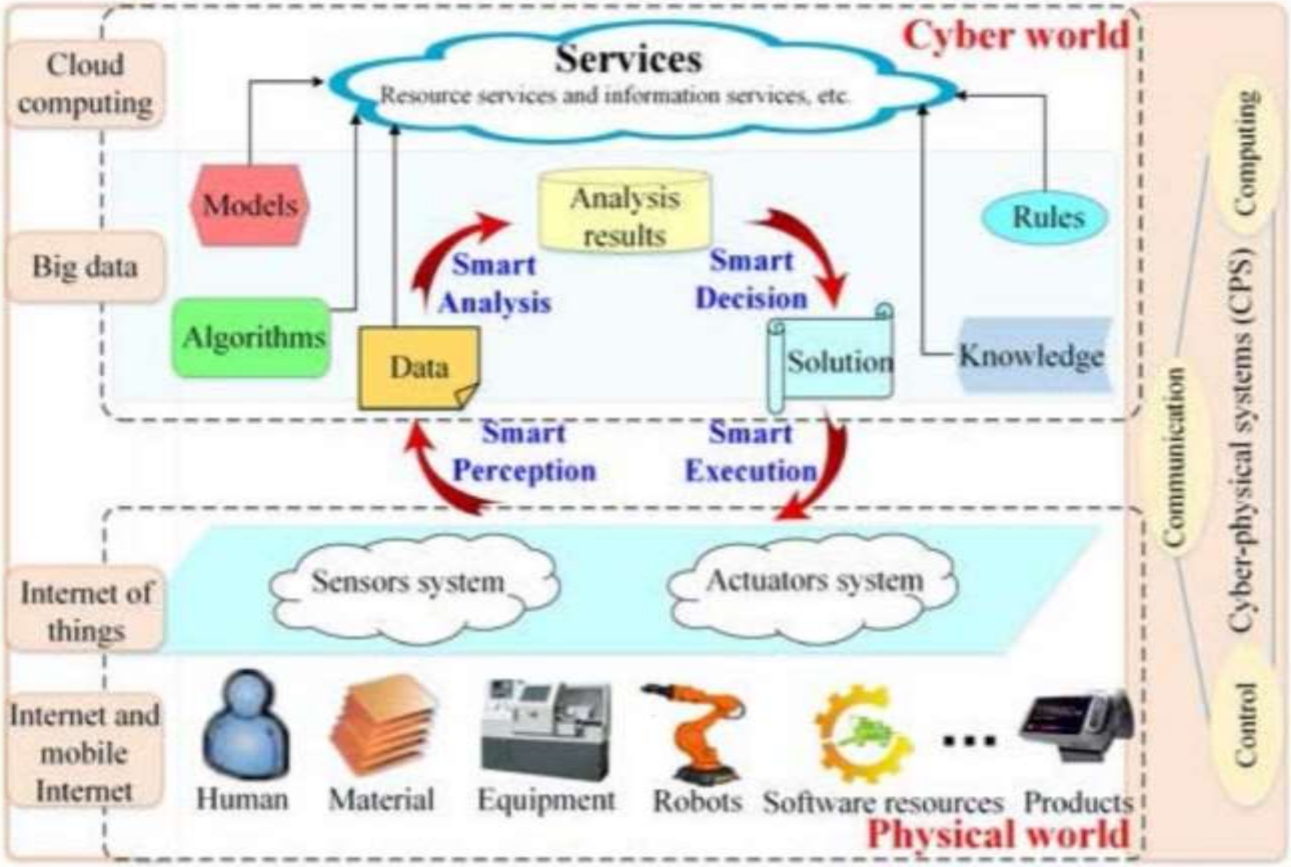
Services as a mean to add value to the product

- Manufacturing as a service, on demand access to digital twin and related services
- Enable resource sharing and cooperation

Service Modeling

Example of services:

- Resource management (models, connections, data)
- Visualization and encapsulation of services provided by the digital twin
- Service evaluation (e.g., time, cost, reliability, energy efficiency)
- Human-machine interface and training
- Quality and health records



Service Modeling

Complex and dynamic smart decision-making problems

- E.g., energy consumption management, precise control, predictive maintenance

Services are used to build a bridge between the manufacturing physical space and virtual space

- E.g., production scheduling, product quality management, equipment health management
- Fusion and collaboration of services through dynamic service invocation, scheduling and combination

Service Modeling

All manufacturing resources and capabilities in product lifecycle can be virtualized and encapsulated in services

- Resource services
  - Developed from equipment, materials, hardware resources
  - Can encapsulate dimensions of the digital twin
- Information services
  - Software resources, such as professional software, experience models, domain knowledge, algorithms
- Extended to capabilities such as design, simulation, production and maintenance
  - Prediction models, maintenance scheduling, visualization applications
  - Augmented reality and 3D modeling and simulation tools

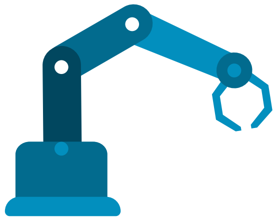
Service Modeling

Physical entity services

- Description and status: name, ID, location
- Quality of service: time, cost, reliability, satisfaction
- Capacities: precision, size, process
- Real-time status: overload, idle, in maintenance

Virtual entity services

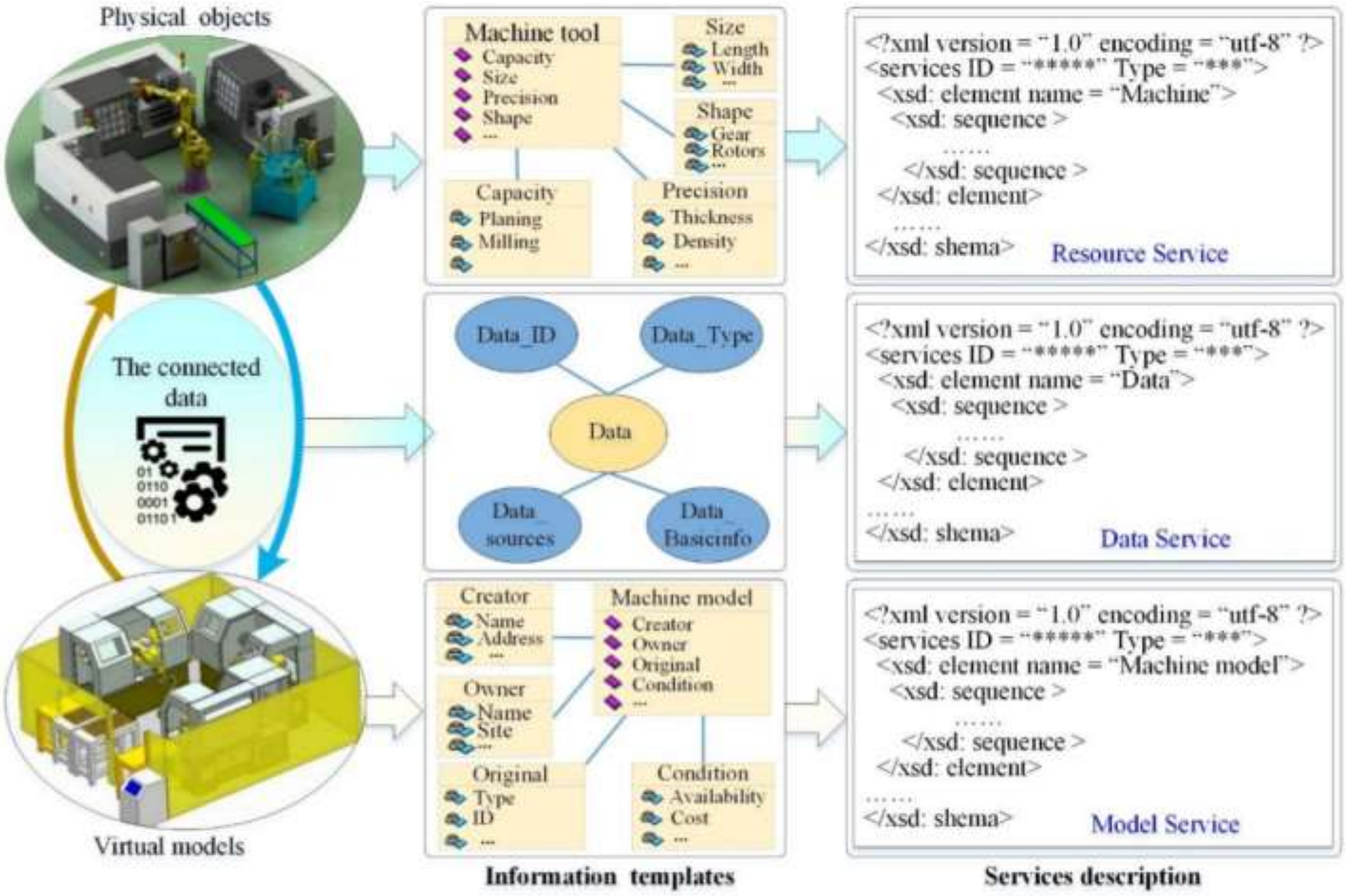
- Can be shared and accessed by multiple users at the same time
- Corresponding physical entity, creator, owner, online site
- Quality of service of the virtual entity: cost, reliability, functions, etc.



Service Modeling

Data services

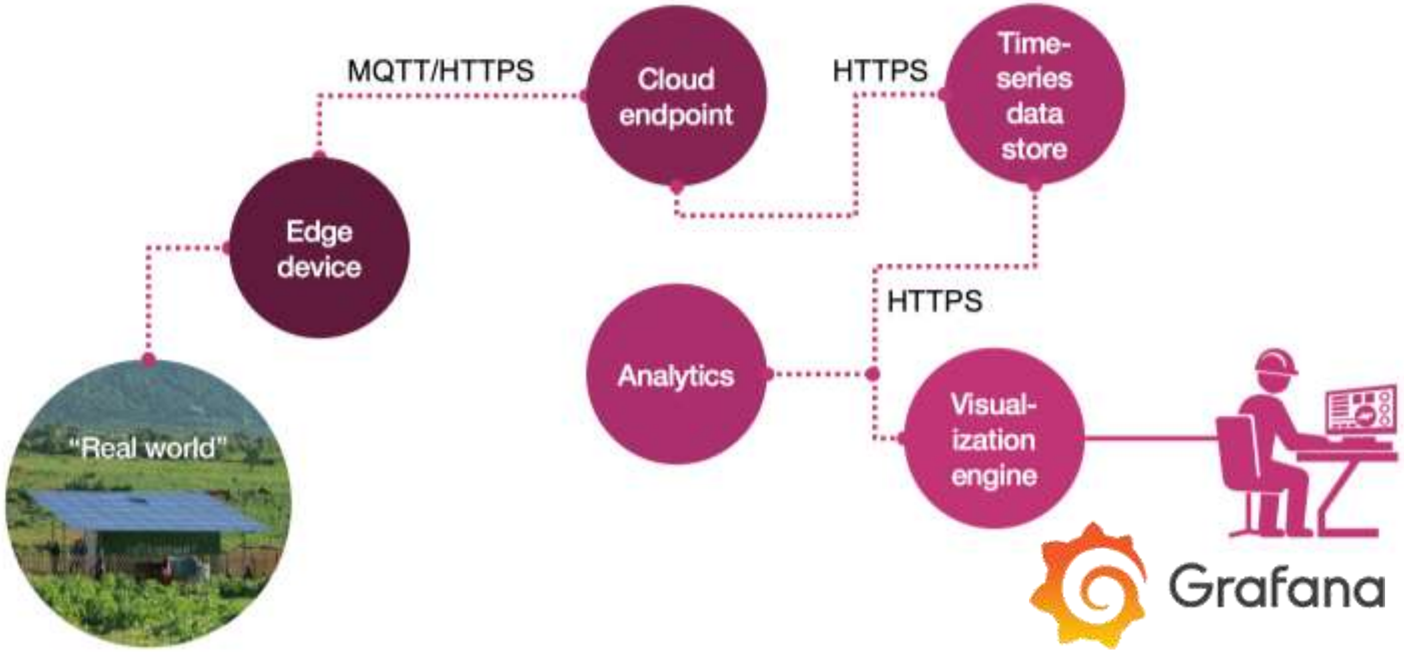
- Allow simplified access to data through unified templates
  - Data provider, sources, type and description
  - Encapsulate information relative to different entities in a chosen language, e.g., XML



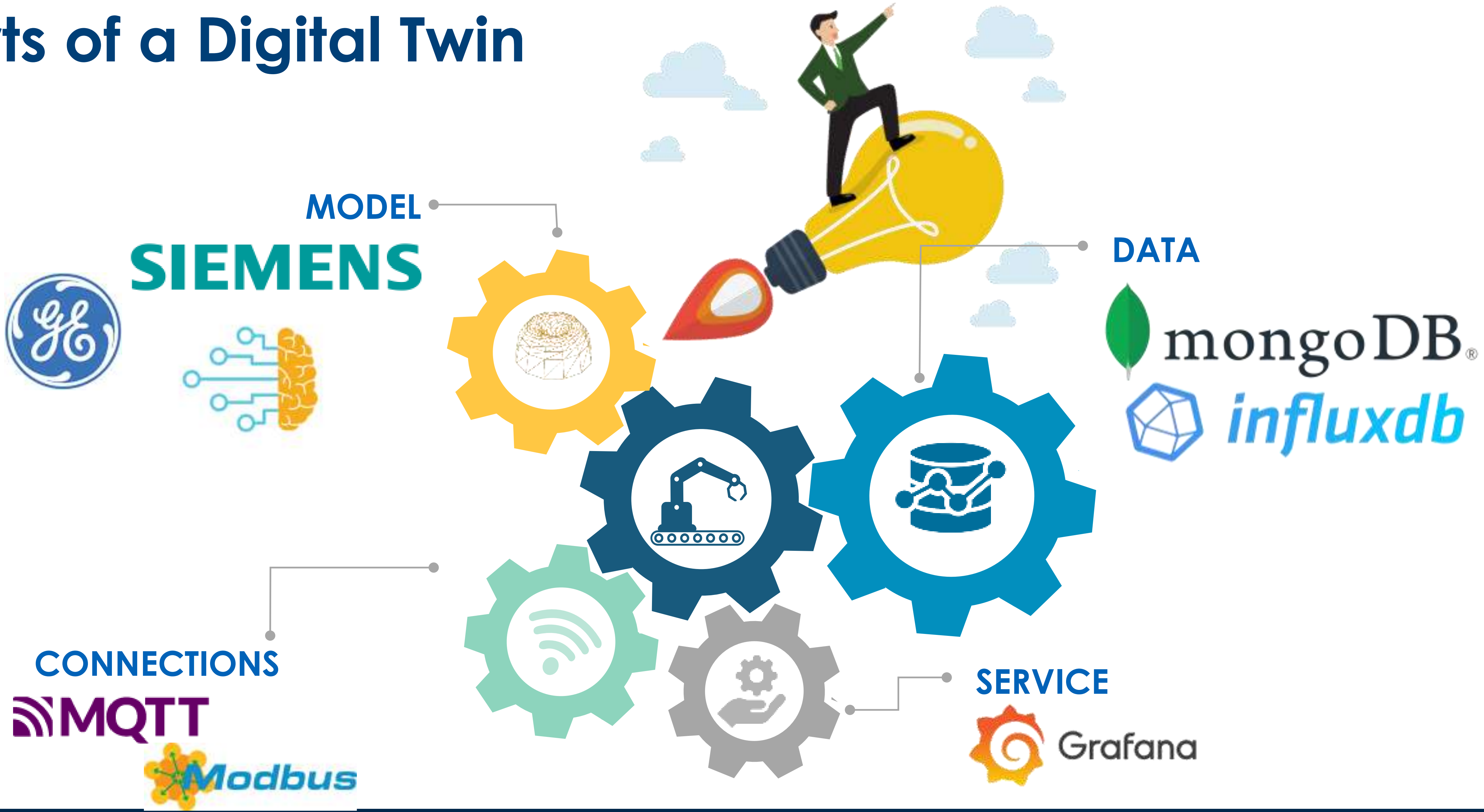
Grafana

Open source observability stack that allows you to monitor and analyze metrics, logs and traces

- Query, visualize, alert on and understand data
- Create dashboards to show data evolution, statistics, etc.



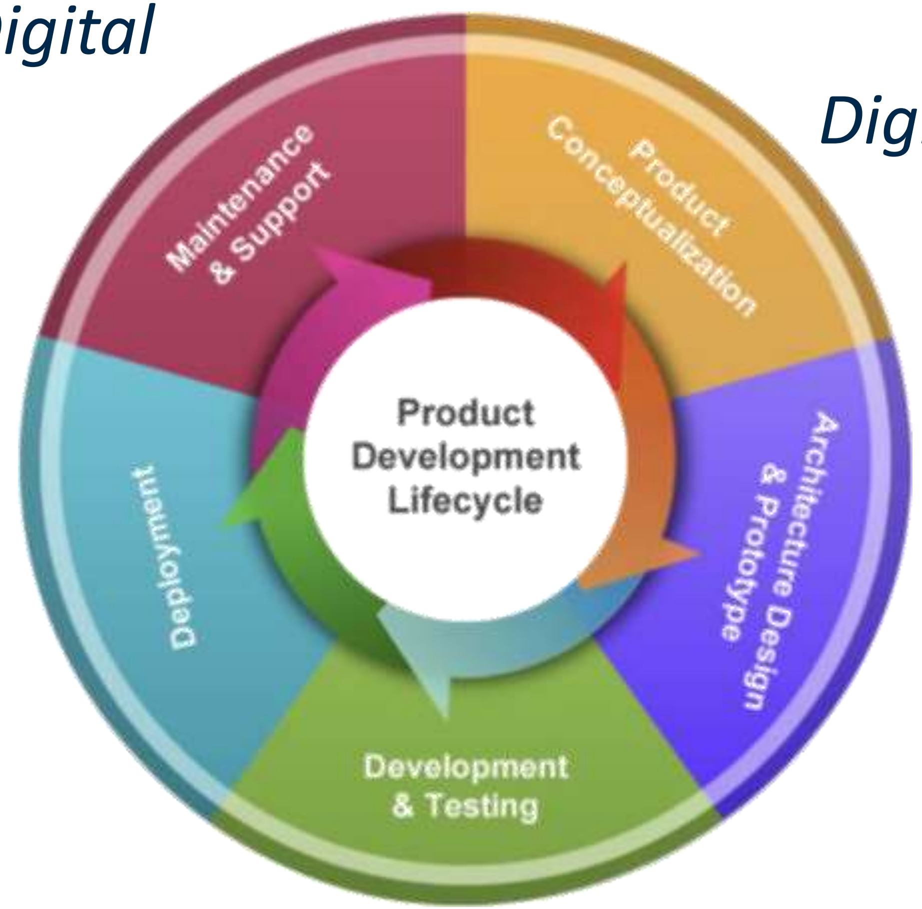
Parts of a Digital Twin



APPLICATIONS IN PRODUCT LIFECYCLE

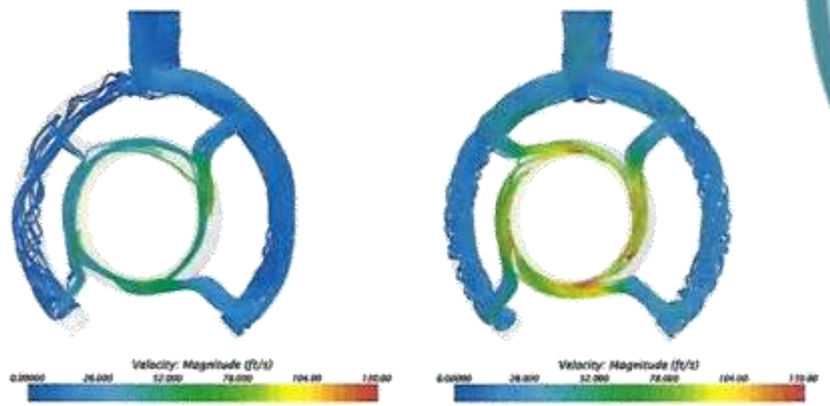
Digital Twins in the Production Lifecycle

At which stage are Digital Twins useful?



When and how are Digital Twins employed?

Digital Twins in the Production Lifecycle



DESIGN PHASE

Digital twin-based methodologies

- More responsive, efficient, and informed design of products
- Physics-based modeling
- Conceptual design and pre-verification
- Prediction of performance and quality at early stages of design
- Deal with uncertainty and lack of physical data to eliminate potential failures

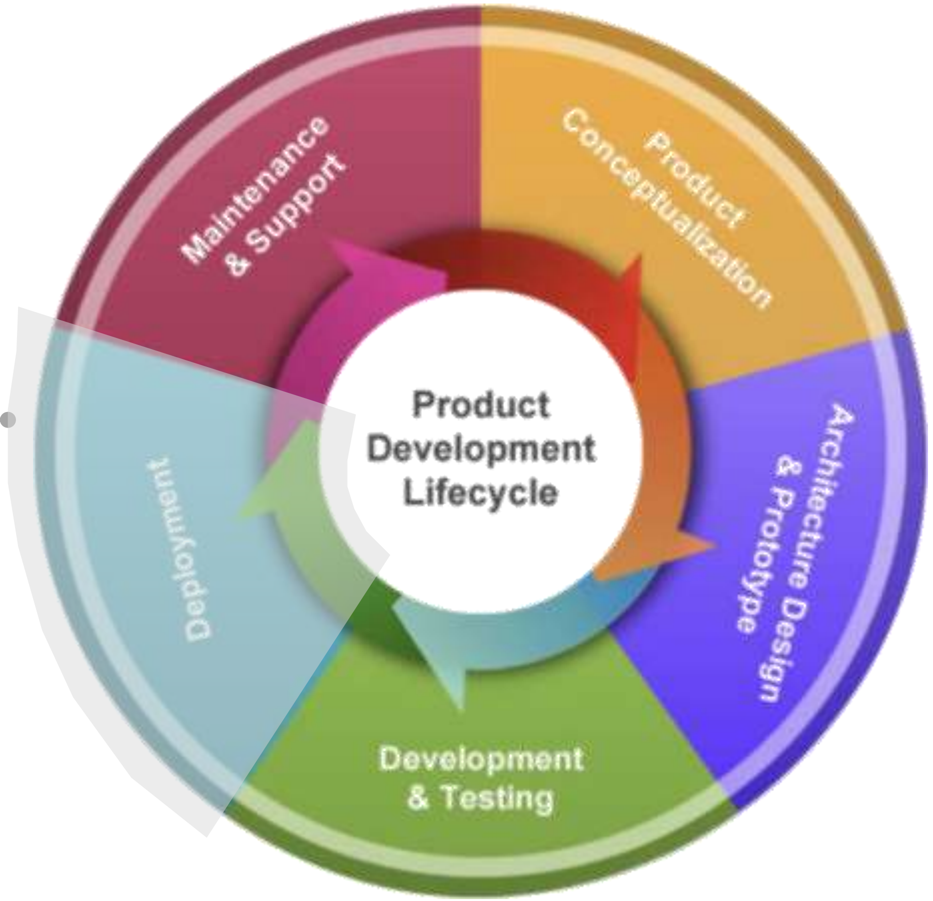
Stephen Ferguson, Edward Bennett, Artem Ivashchenko, Digital twin tackles design challenges, World Pumps, Volume 2017, Issue 4, 2017, Pages 26-28.

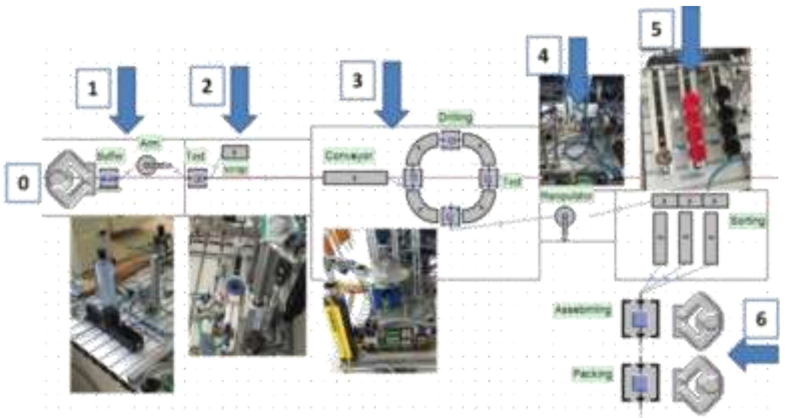
Digital Twins in the Production Lifecycle

PRODUCTION PHASE

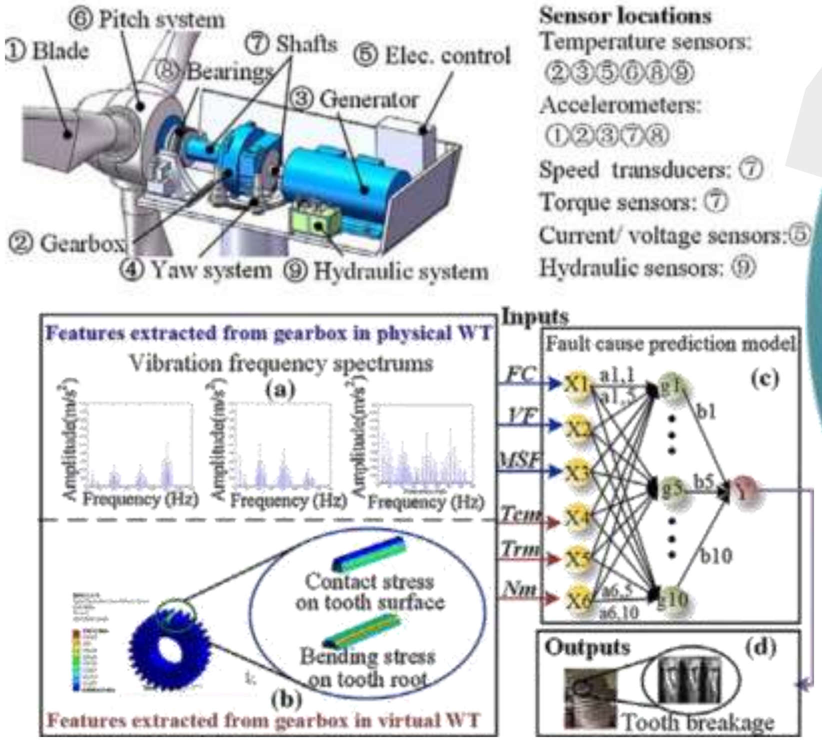
Improve the production process: virtual space mirrors physical space in timely manner and controls physical behaviors in real time

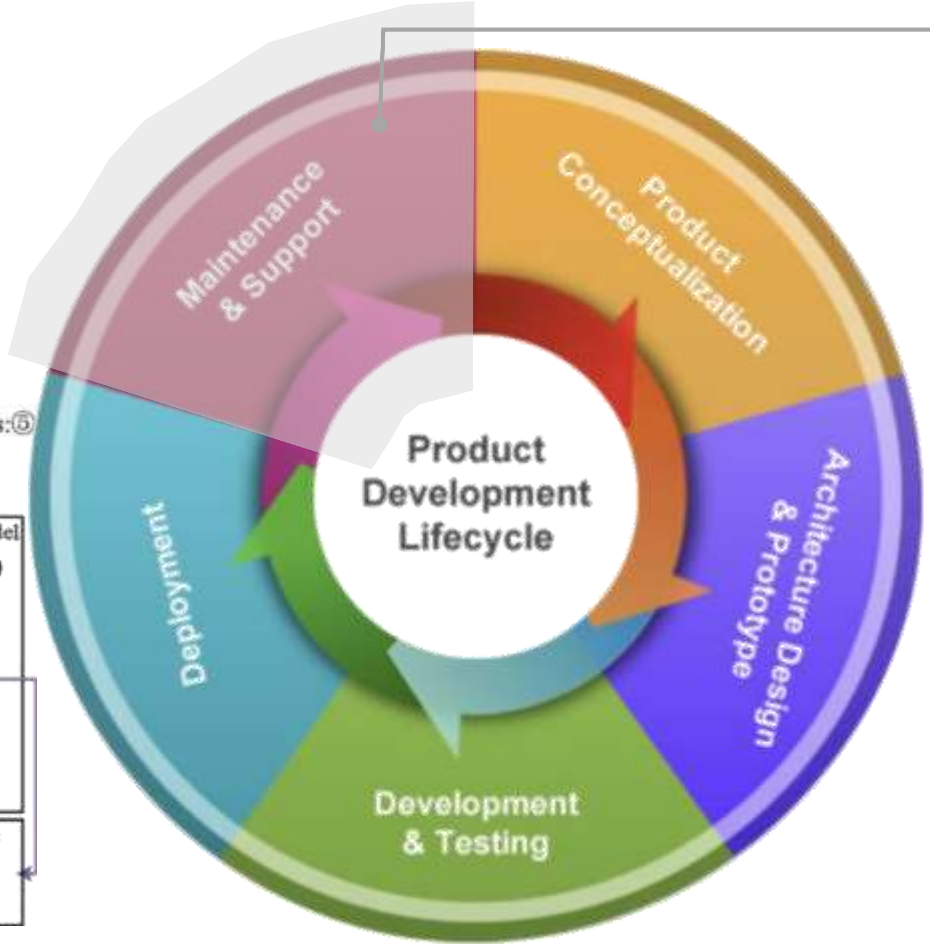
- Production becomes more reliable, flexible and predictable
- Visualize and update real time status
- Accurate forecasting
- Optimize energy consumption and throughput, make timely adjustments
- Prolong machine lifetime





Digital Twins in the Production Lifecycle





MAINTENANCE PHASE

Combine sensor and historical data for fault prediction

- Multi-physics, multi-scale modeling
- Machine learning
- Integrate models and simulated data to generate valuable information for efficiency, accuracy and maintenance
- Increase reliability and prolong machine lifetime

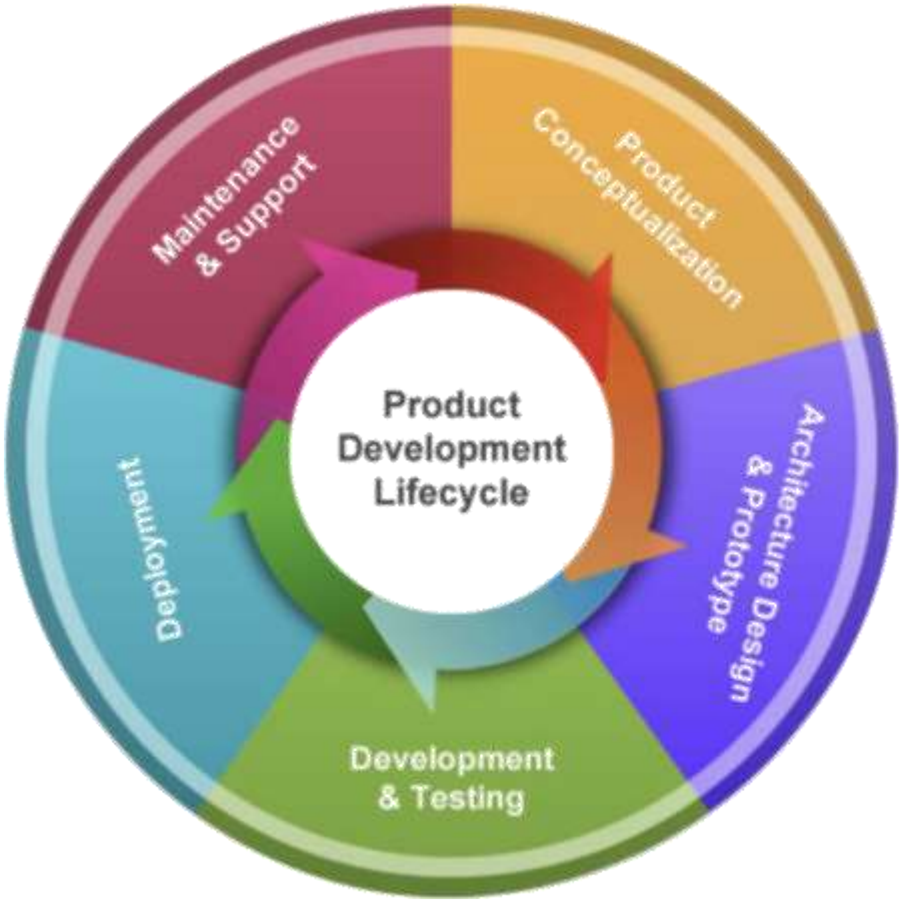
Fei Tao, Meng Zhang, Yushan Liu, A.Y.C. Nee, Digital twin driven prognostics and health management for complex equipment, CIRP Annals, Volume 67, Issue 1, 2018, Pages 169-172

Digital Twins in the Production Lifecycle

ACROSS ALL PHASES

The digital twin can integrate data from the various stages seamlessly

- Design digital twin adjusted during service for maintenance
- Support across product lifetime
- Secure good geometrical quality through statistical variation simulation (design), inspection (pre-production) and root cause identification (production)
- Historical lifecycle experience used in design to make accurate simulations of variations, detect faults and their cause

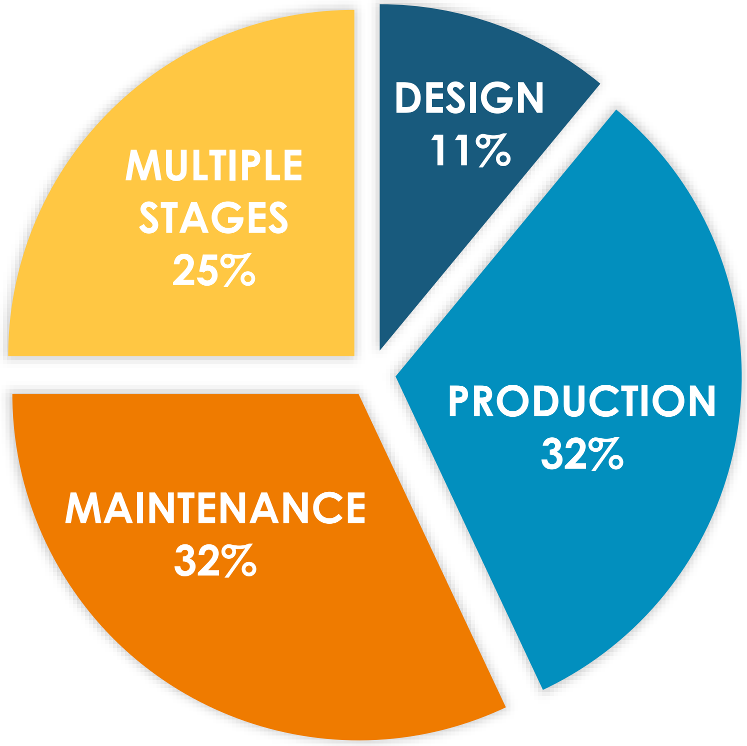


Fei Tao, Meng Zhang, Yushan Liu, A.Y.C. Nee, Digital twin driven prognostics and health management for complex equipment, CIRP Annals, Volume 67, Issue 1, 2018, Pages 169-172

Digital Twins in the Production Lifecycle

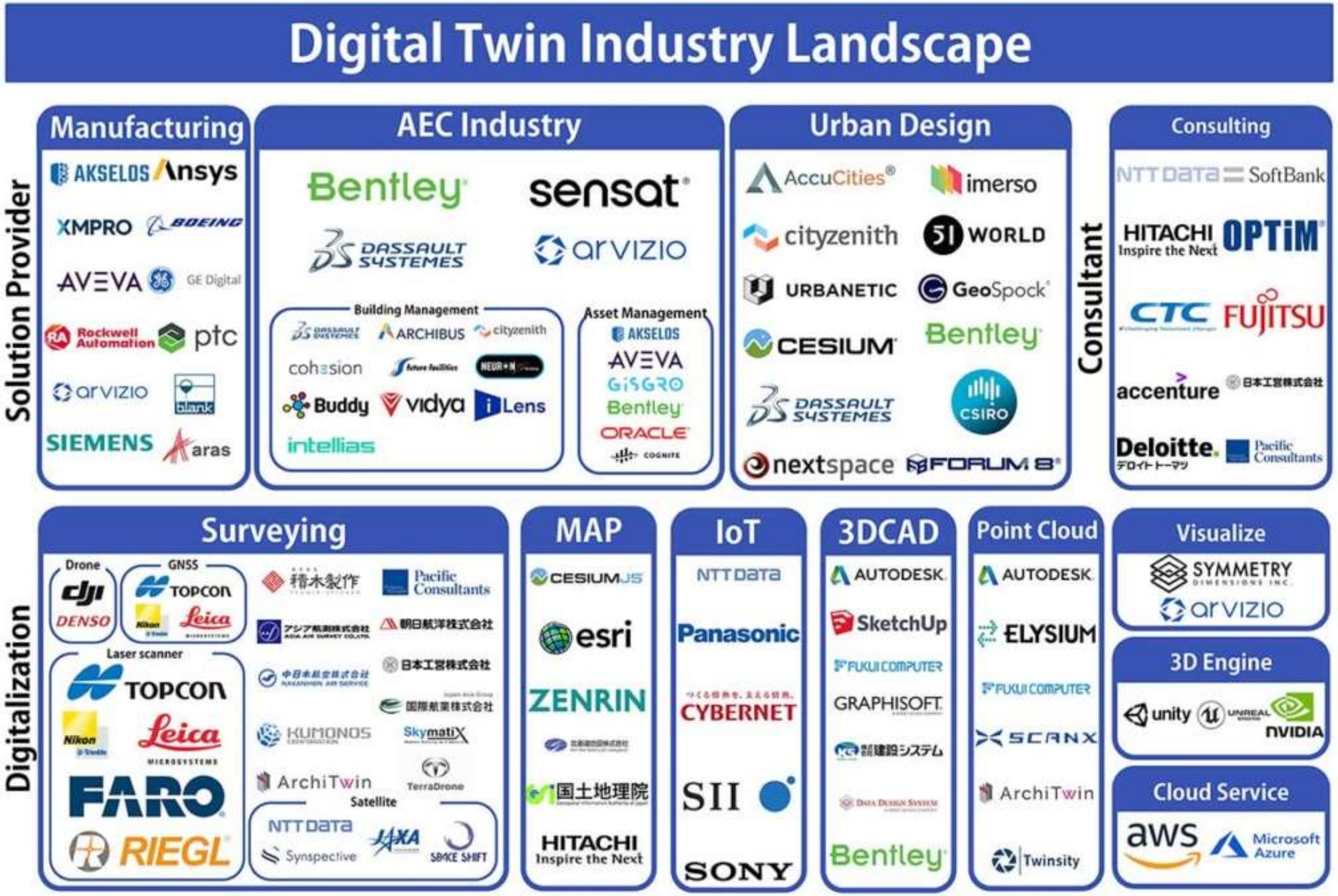
Production and maintenance are the most popular fields of application

- Maintenance was the first field of application
  - Lots of attention, time and resources
  - Still limited application
- Production to realize smart operations
  - Great potential in real-time control and optimization and for accurate prediction
  - Future stage of application
- Covering all stages requires high data integration effort



INDUSTRIAL APPLICATIONS

Current landscape



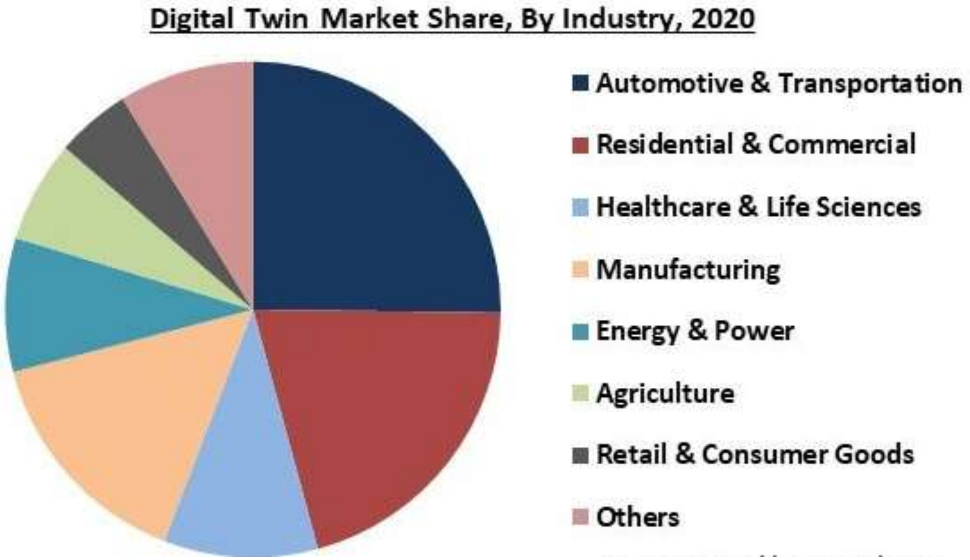
https://www.grandviewresearch.com/industry-analysis/digital-twin-market

Future market

Despite of the challenges, digital twins have very promising prospects

› Different estimations but all agree on a growth of Revenue Growth Rate (CAGR) higher than 40%

Report Attribute	Details
Market size value in 2020	USD 5 Billion
Market size forecast in 2027	USD 63.5 Billion
Base Year	2020
Historical Period	2017 to 2019
Forecast Period	2021 to 2027
Revenue Growth Rate	CAGR of 41.7% from 2021 to 2027



Industrial applications

Despite of the challenges, digital twins are already applied in industry in a number of contexts

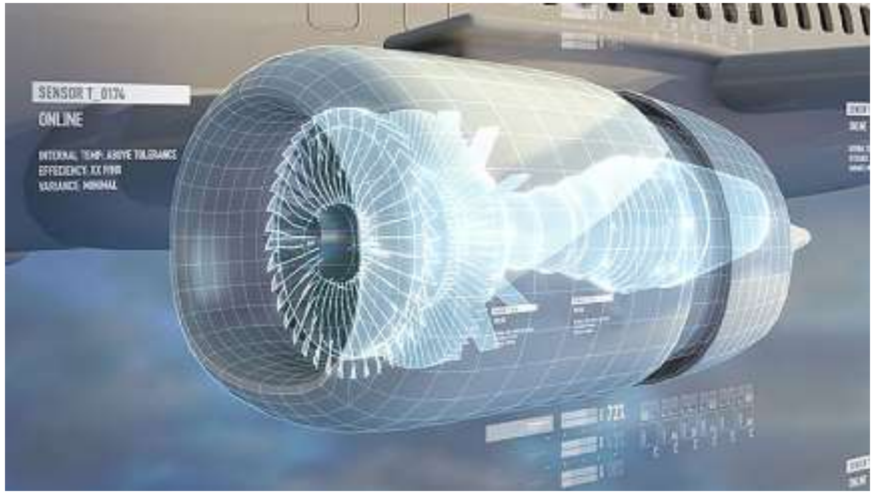
Digital twin in Aerospace

Earliest industrial domain adopting digital twins

- Large companies combined digital twin technology with their business, including maintenance, production, assembly, safety and security management
- Airbus, Boeing, General Electric Company, …

Mature application scenario

- Good amount of research
- Can perform real time prediction and high fidelity validation in virtual space
- Important to improve reliability, reduce accidents and conserve resources



Digital twin in Electric Power Generation

Digital wind farm to collect real-time data, optimize maintenance of each turbine to increase annual energy production

- Mix and match different turbine configurations
- Collect and analyze data from the real-life version
- Increase efficiency (average loss -10% to -33%, +20% efficiency)

Real-time data management

- Acquire and manage critical data to view and control turbines in real time
- Protection from digital threats

https://youtu.be/OZ-N3h40IPc
https://www.ge.com/renewableenergy/wind-energy/onshore-wind/digital-wind-farm

Digital twin in Automotive

Car are becoming more complex, with higher requirements for high precision testing and maintenance

- Make repair scenarios for trucks and locomotives
  - Collect real time data to find out why the breakout happened, e.g., health conditions of parts or changes in relevant variables
  - Find how the breakout happened and reproduce engine or vehicle model under specific conditions
  - Reduce downtime cost and duration
  - Optimize fuel efficiency
- Taycan Porsche is provided with a digital twin
  - Recommend optimal time and required scope of service
  - Customize service intervals and allow servicing for specific components, based on how the customer uses their vehicle

https://newsroom.porsche.com/en_US/2021/technology/porsche-digital-chassis-twin-26323.html

Digital twin in Oil and Gas

Remote areas with severe conditions with many issues in process management

- Unplanned downtime leads to loss of revenue and profitability up to $1 trillion per year

Emergence of digital twins brings opportunity to improve maintenance and operation

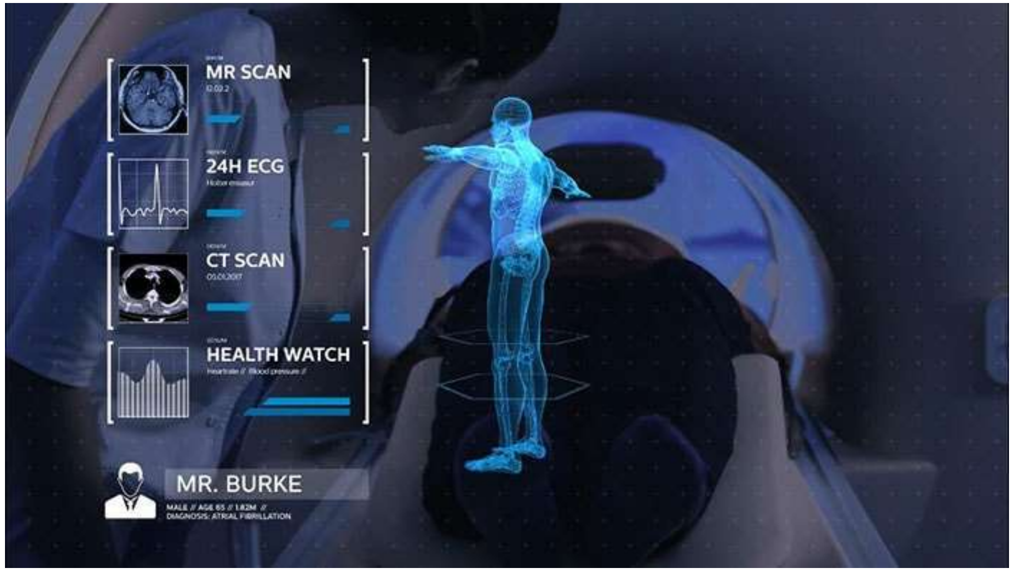
- Help operating under risks
- Equipment installation and maintenance
- Identify and manage changes in design
- Efficient means for data collection, visualization and state monitoring

https://youtu.be/OZ-N3h40IPc
https://www.ge.com/renewableenergy/wind-energy/onshore-wind/digital-wind-farm

Digital twin in Healthcare and Medicine

Personalized medicine

- Models of a human organ accounting for blood flow, mechanics and electricity
- … or even of a whole human!
- Possible applications:
  - Discover undeveloped illnesses to improve patient care
  - Experiment with treatments and develop innovative ways to deal with difficult illnesses/understand how different individuals would react
  - Improve preparation for surgeries
  - In the future may also help physicians optimize the performance of patient-specific treatment plans



https://youtu.be/H6JzPCbyVSM
https://www.challenge.org/insights/digital-twin-in-healthcare/

Digital twin in Maritime and Shipping

Digital twins attracted attention for the creation of «virtual sister ship»

- Support data visualization and analytics
  - High speed real time sampling
  - Compared with data from the model to detect difference or performance degradation
- Optimization of performance and communication
- Ease system integration and quality assurance
  - Predictive analytics of critical equipment
  - Solve problems before they actually impact on ship operation
- Systematic framework to produce information and high quality reports
- Training personnel

Digital twin in City Managment

cityzenith

Widely known example is Singapore:

- Enable virtual experiments and large scale simulations to optimize
  - Way to improving accessibility, simulating emergency situations
  - E.g., heating or cooling to optimize consumption
  - Intelligence in operation, analysis, and prediction
- Long term planning and decision making
  - Scale from a single building to a portfolio of hundreds and even thousands of building

https://youtu.be/QnLyy0owGL0
https://www.smartcitylab.com/blog/digital-transformation/singapore-experiments-with-its-digital-twin-to-improve-city-life

Digital twin in Agriculture

Potential to revolutionize agriculture

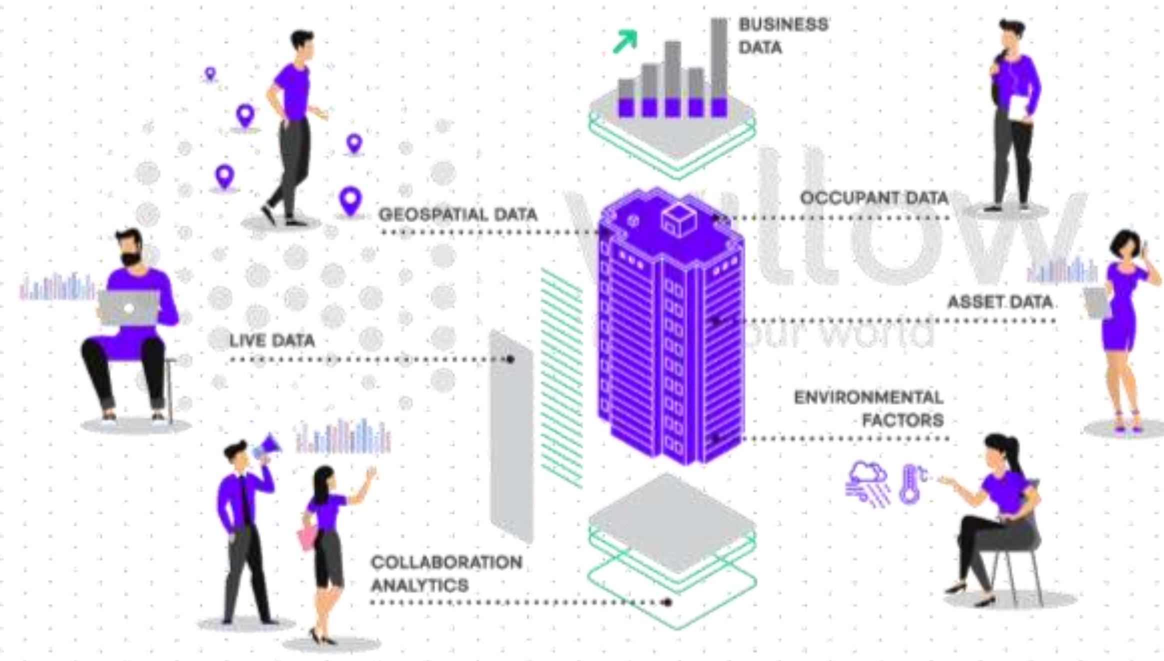
- Remote monitoring of growing cultures and cattle
- Recording and identifying pests and diseases
- Monitor and optimize management of silos of livestock farms
- Track machinery fleet in real time
- Accelerate farm supply, production, harvest, packaging and distribution

Will be an indispensable technology for the agriculture of the future

Digital twin in Construction

Help in all phases, from design and construction to operation

- Access and analyze real time data and historical data plus documents
- Support decision making to make effective and timely changes during design phase
- Centralized construction data collection and management, handle all activities and documents
- At handover, make available models and information in one place
- Optimize building operation and occupation, monitor resources and waste, prevent hazards



Digital twin in Security and Emergency

Digital twins can be adopted for security and emergency measures, to achieve lower costs and better performance and guarantees

- Data from the lifecycle can be used to generate valuable information for handling with machine failures, and for predicting and applying maintenance before the occurrence of failures
  - As soon as an anomaly occurs, it is then possible to digitally replay exactly what happened to the equipment and perform diagnostics, assess the gravity of the situation, and identify root causes
- Replace or complement physical inspections in hazardous environments
  - Visibility and predictability without embarking on a costly and risky physical exploration
  - Reduce the risk of unnecessary shutdowns or damages due to errors in judgment
- Security of equipment and staff can be guaranteed

EXAMPLE OF DIGITAL TWIN OF A ROBOT ARM

Why a robot arm

Widely used in manufacturing scenarios

- Exposed to physical stress
- May be harmful to nearby humans

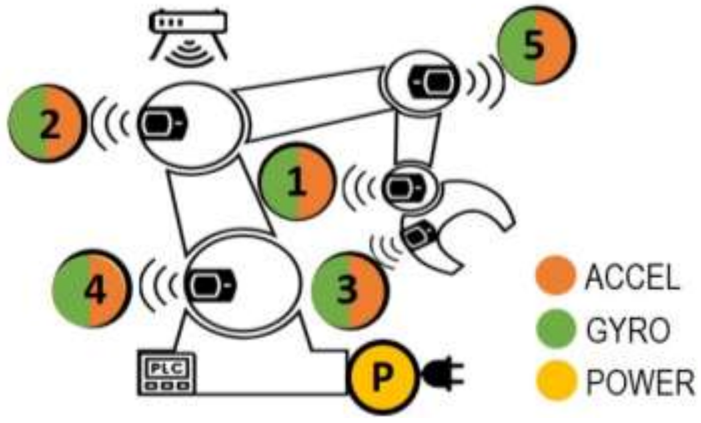
Desirable to have a way to monitor its behavior to intercept anomalous behaviors or even predict future failures

- How can we approach its modeling?
- Goal: not reproducing its movement but rather apply anomaly detection

Data collection from the robot arm

Setup of the monitoring infrastructure

- 11 sensors overall
- 5 gyroscopes and 5 accelerometers to detect movement (positions *1* to *5*)
- 1 power sensor to detect power consumption (position *P*)
- Attached to the existing “non digital” arm
- Overall 56 different signals



Setup of the network infrastructure

- Sensors natively communicate via Modbus
- Data sent to a gateway and forwarded to an InfluxDB database

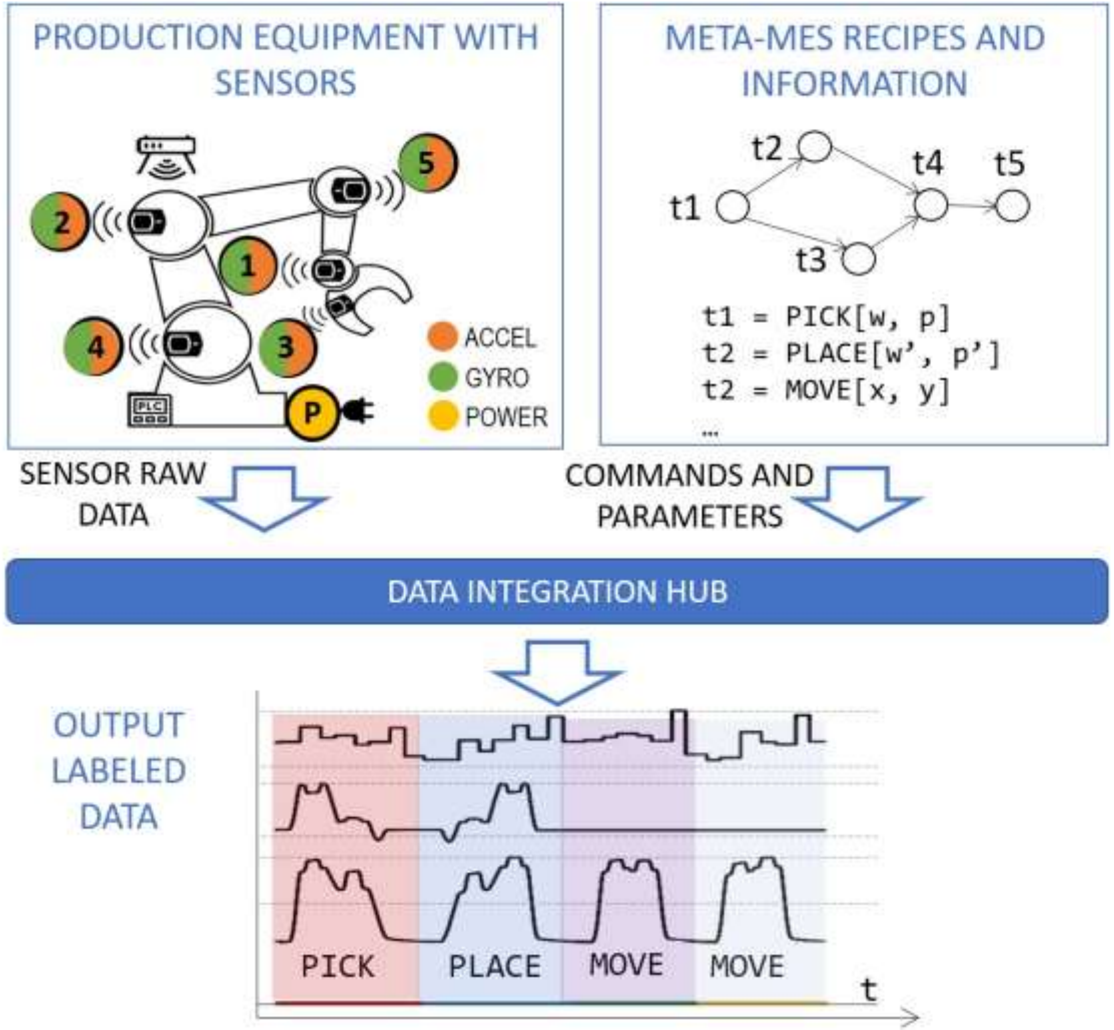
Data collection infrastructure

To have a clear view of the arm conditions it is important to consider simultaneously

- Sensor data traces
- What the arm is doing/supposed to do

How to merge such information?

- Usually commands are labels applied by hand (costly, errors, small datasets)
- Alternative: apply automatic data fusion and labeling
  - Merge information from different sources
  - Handle different types/formats and time scales
  - Automatic rich dataset



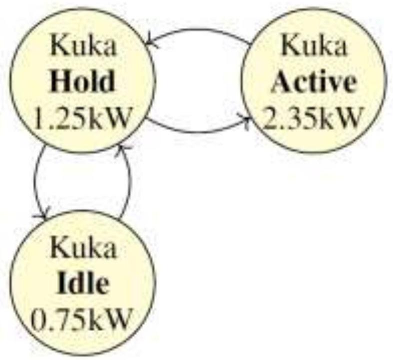
Incremental models of the digital twin

Manually applied labels

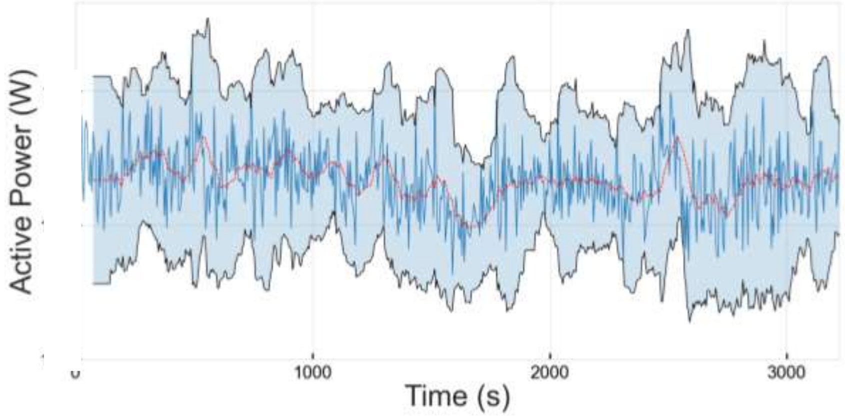
› Few samples available with limited annotations (e.g., only commands with no parameters) – limited visibility on arm operating conditions

› Different kind of models with different levels of accuracy

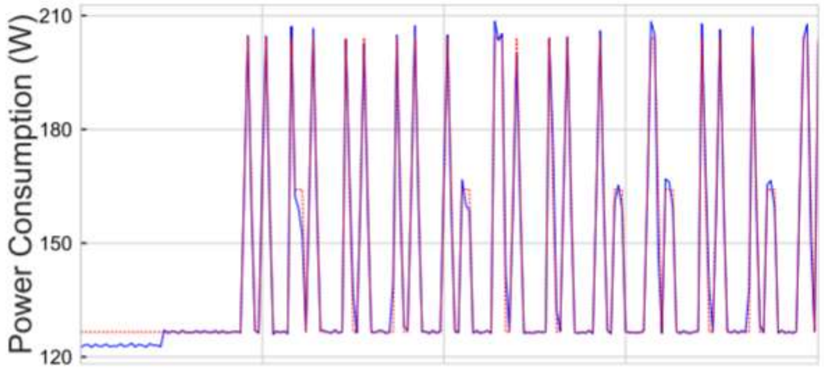
› Focus on power consumption



POWER STATE MACHINE: VERY ABSTRACT, TYPICAL CONSUMPTION



MOVING AVERAGE BASED ON MEAN AND STANDARD DEVIATION FOR OPERATING MODE

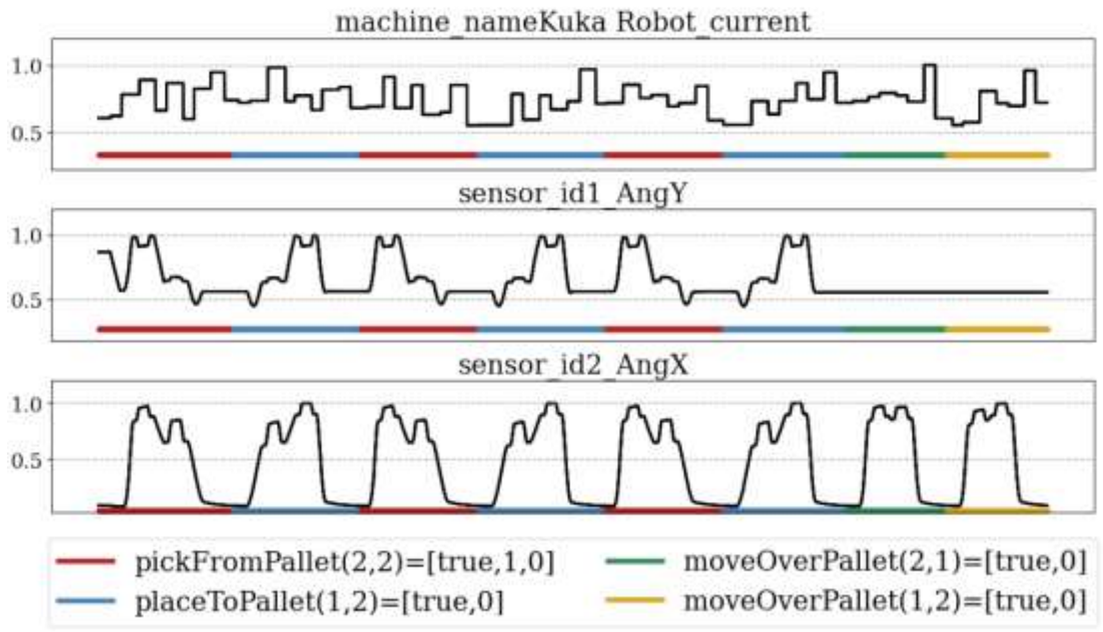


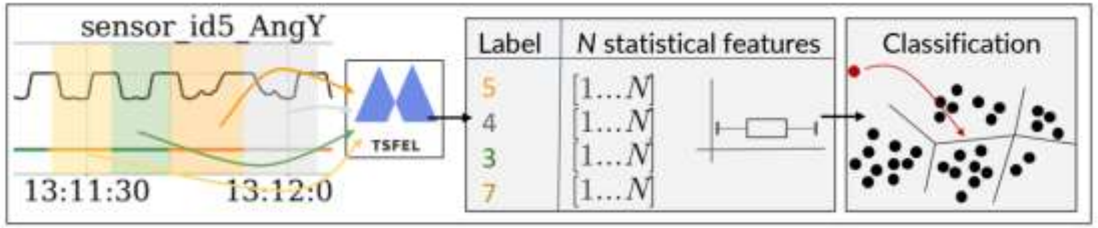
NEURAL NETWORK MODEL: RELATIONSHIP BETWEEN OPERATING MODE AND CONSUMPION

More complete model

Automatically labelled data

- Longer acquisitions (e.g., hours of operation with 1s time step for sensors)
- Slice each signal based on annotated operating mode
- Statistical feature extraction and feature selection to identify most relevant aspects
  - Accelerometers 2, 4, and 5 plus power sensors are the most informative data sources
- Resulting features (from 1980 to 50) used to construct a classifier

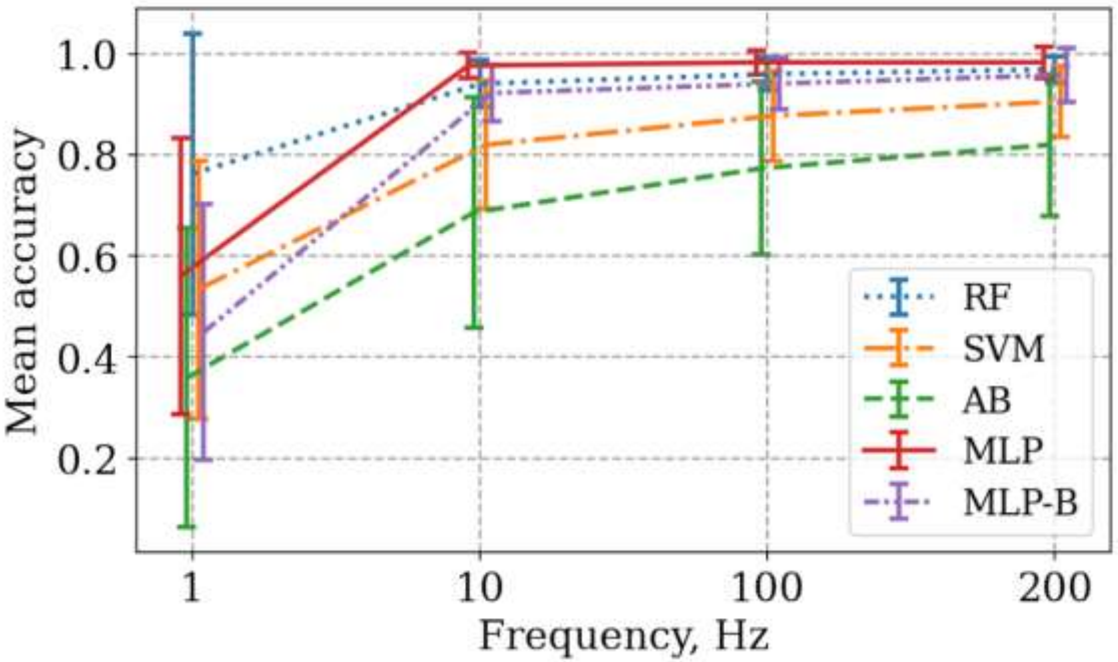




More complete model

Evaluation of different classifiers and different sampling rates of sensor data

- Increasing accuracy at increasing sampling frequency
- Multilayer Perceptron (MLP) is the one that provides the highest and most balanced mean accuracy values (97.6±2.5% at 10 Hz)
- Bayesian (MLP-B) also provides a measure of confidence



Anomaly detection

Use the classifiers to detect anomalous behaviors

- Action predicted is different from the one specified by the MES → ANOMALY!
- Action predicted with low confidence → POTENTIAL ANOMALY!
- 79% anomalies detected correctly

Interesting: we do not need to train the algorithms on specific anomalies

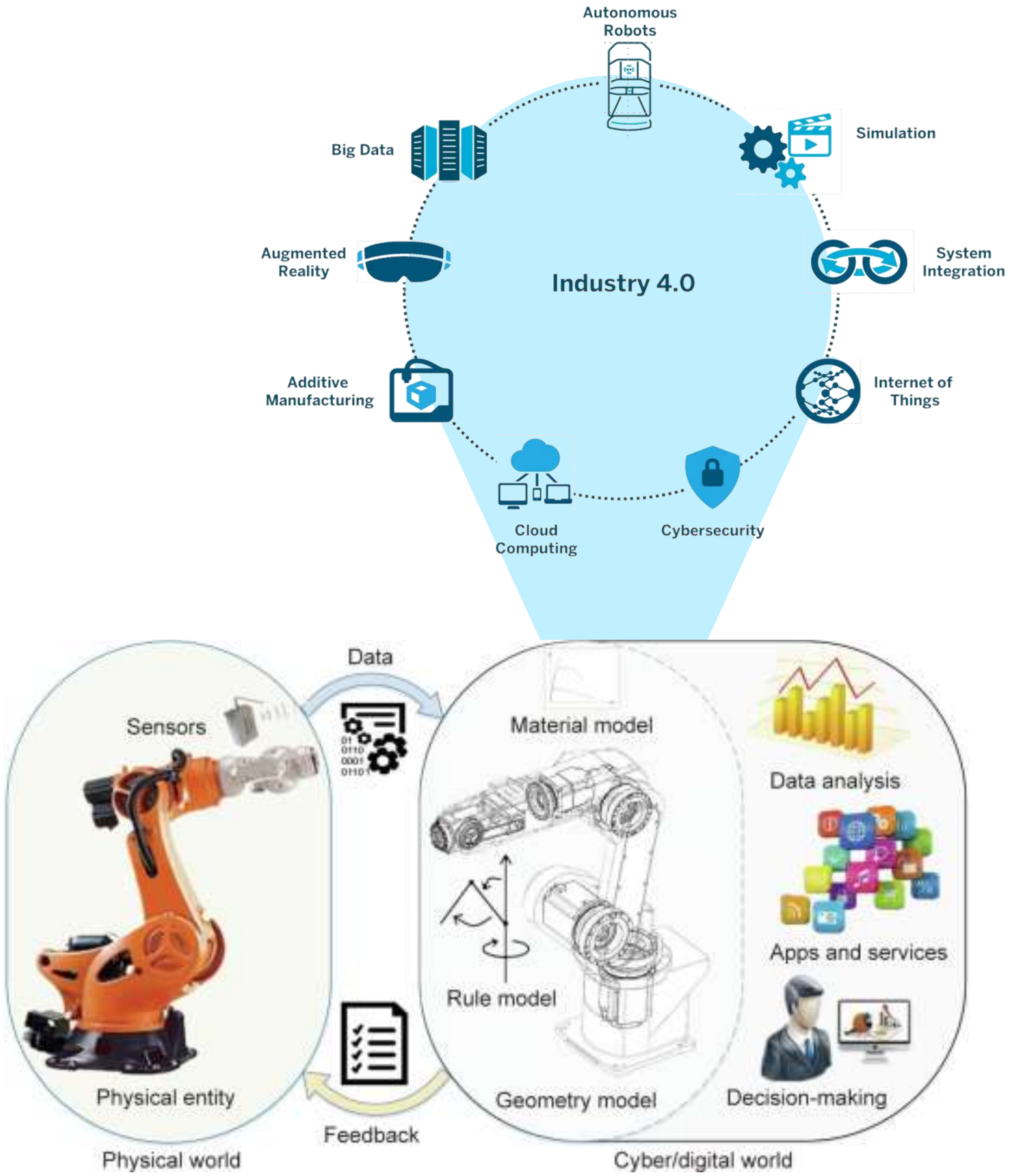
- Anomalies are not known at the beginning of equipment lifetime
- Not all typical anomalous behaviors may be known

Next step: cloud/edge partitioning

Wrap up!

Digital twins are CPS seen from a different perspective (engineering/manufacturing)

- Focus on data and its management
  - Wide exploitation of data-driven methods as physical entities are complex or difficult to model otherwise
- Benefit from related technological research
  - Span across all modern technologies, from IoT and cloud to cybersecurity and data analysis
- Strong pressure from the market



THANK YOU!

sara.vinco@polito.it