DATA MANAGEMENT, FROM ASSET DEFINITIONS TO OPERATORS NOTES

Professor Jerker Delsing
Luleå University of Technology
Sweden

LULEÅ UNIVERSITY OF TECHNOLOGY

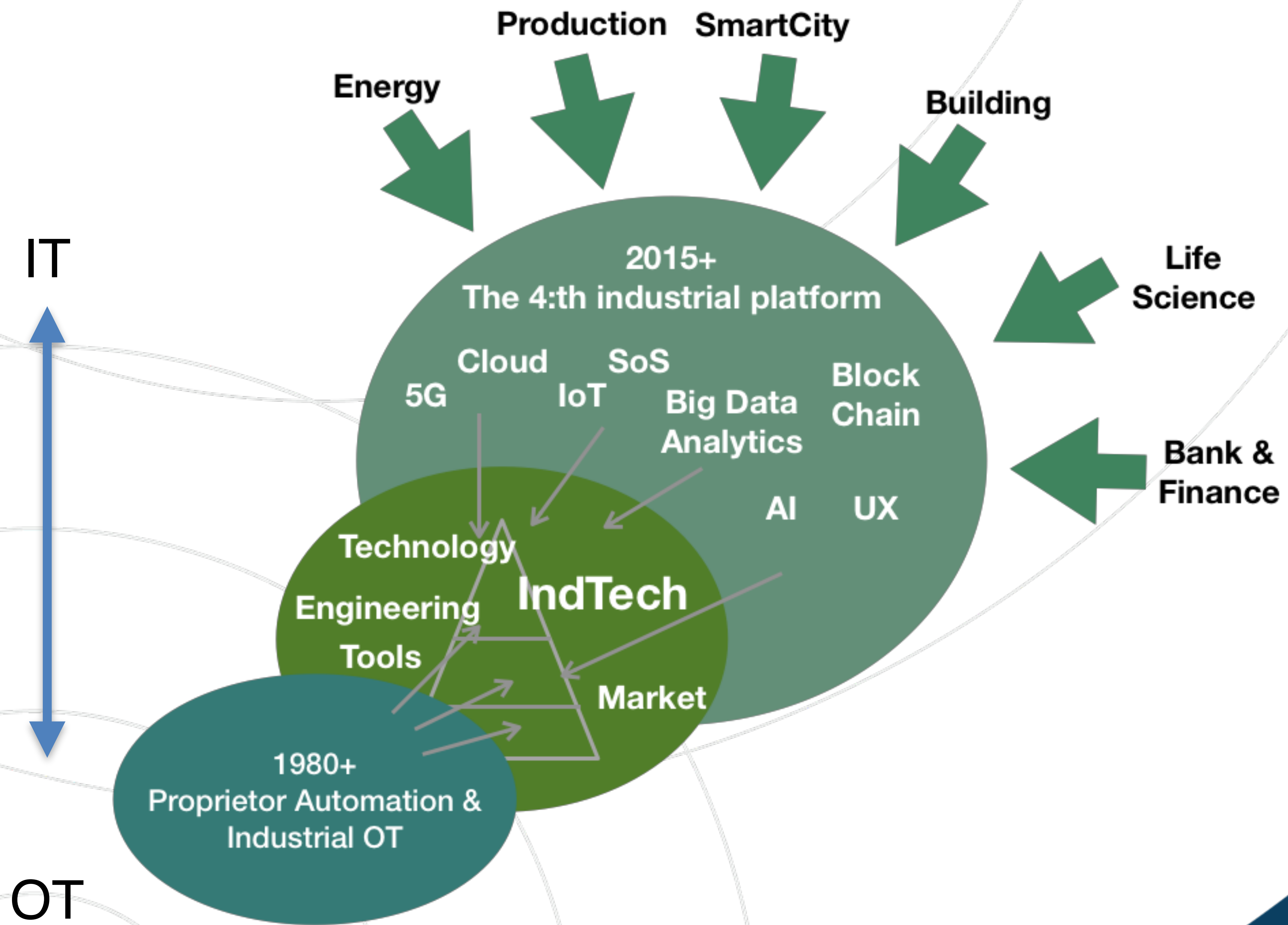
HOW TO ACQUIRE THE FOR AI NECESSARY DATA

Professor Jerker Delsing
Luleå University of Technology
Sweden

LULEÅ UNIVERSITY OF TECHNOLOGY

Our data environmemt

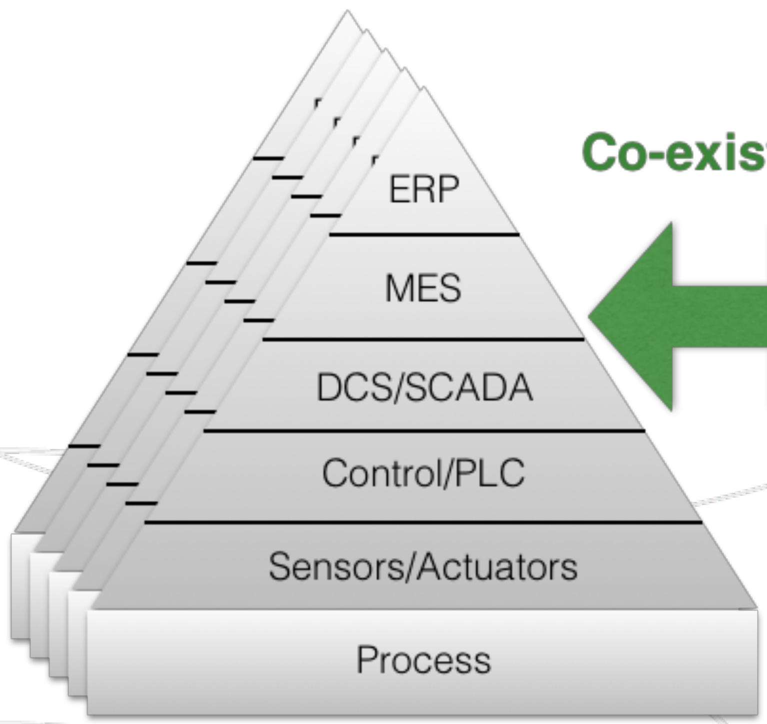
OT meets IT


Production
SmartCity
Energy
Building
Life Science
Bank & Finance
IT
OT
2015+
The 4:th industrial platform
Cloud
SoS
5G
IoT
Big Data Analytics
Block Chain
AI
UX
Technology
Engineering
Tools
IndTech
Market
1980+
Proprietor Automation & Industrial OT

Complex production automation environment

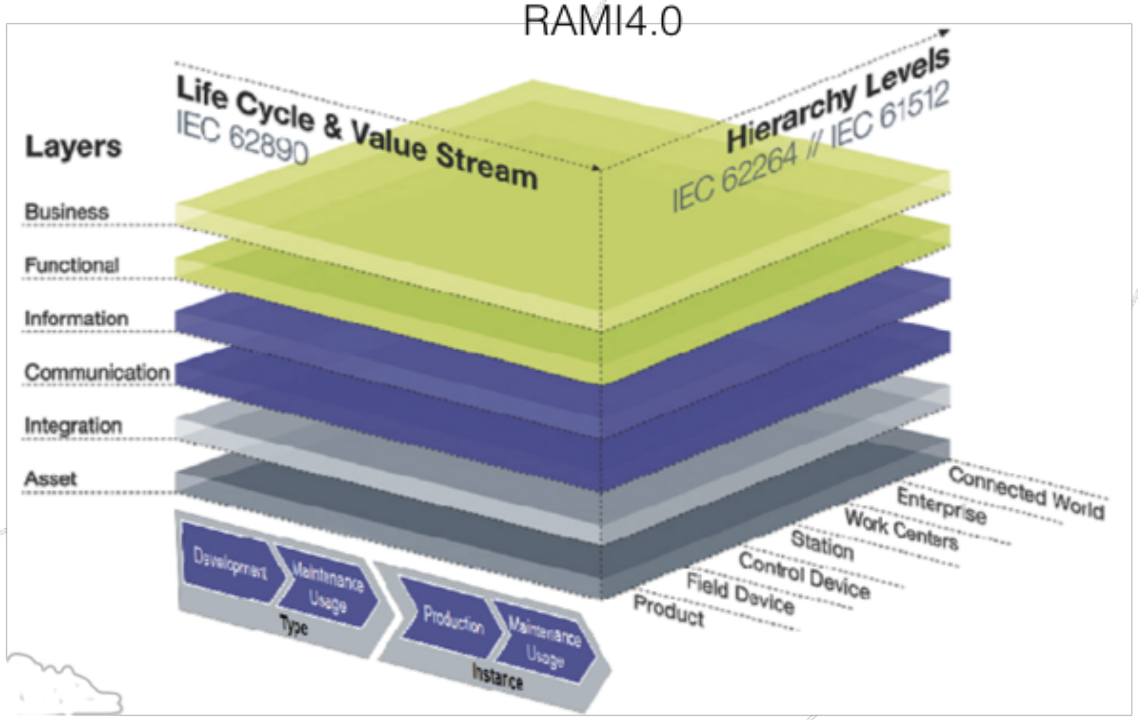
Legacy ISA 95

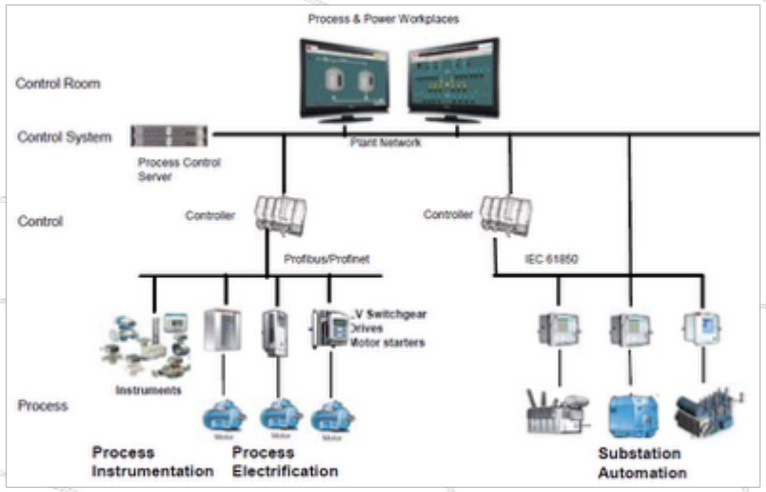
ISA-95



Co-existens

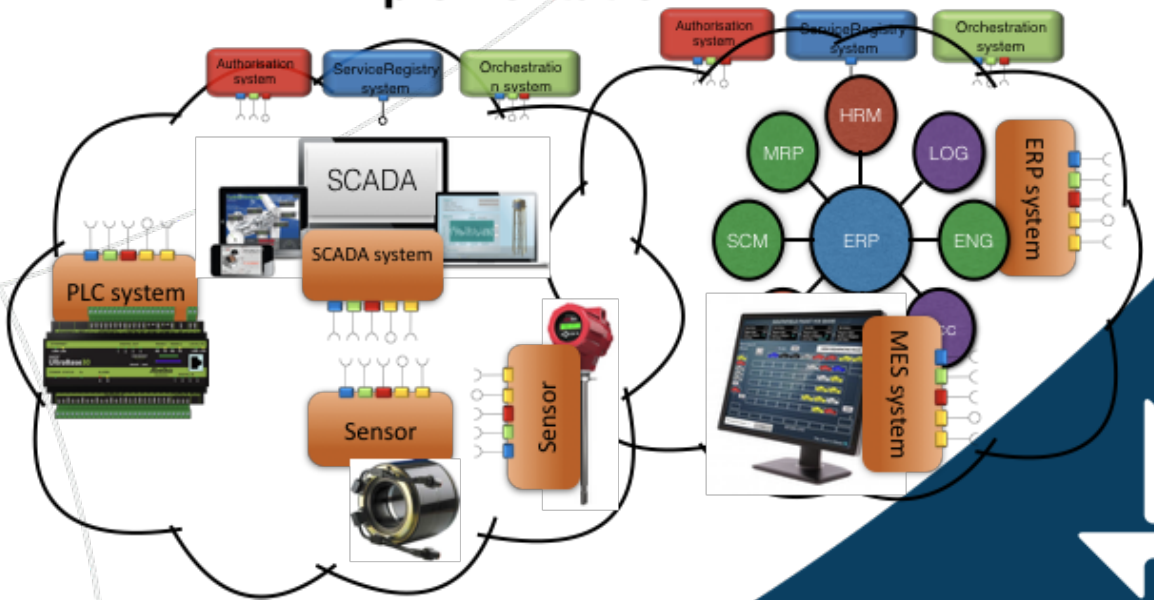
RAMI4.0





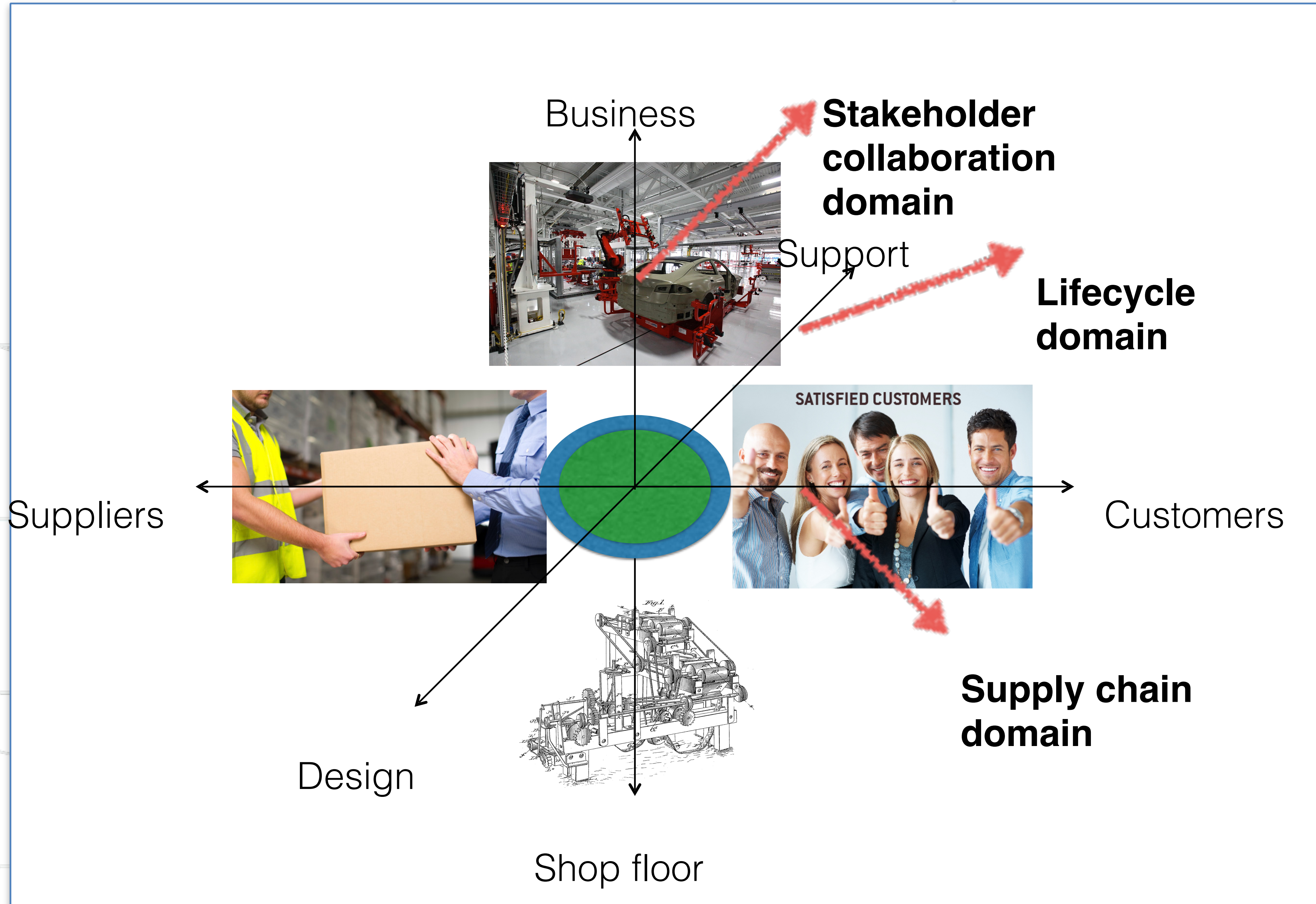
Industrie4.0

Local automation cloud implementation



ARROWHEAD TOOLS

From enterprise to multi stakeholder operation



ARROWHEAD TOOLS

Multi-stakeholder environment
Value networks

Data ownership

Access to data

Trusted logs of actions

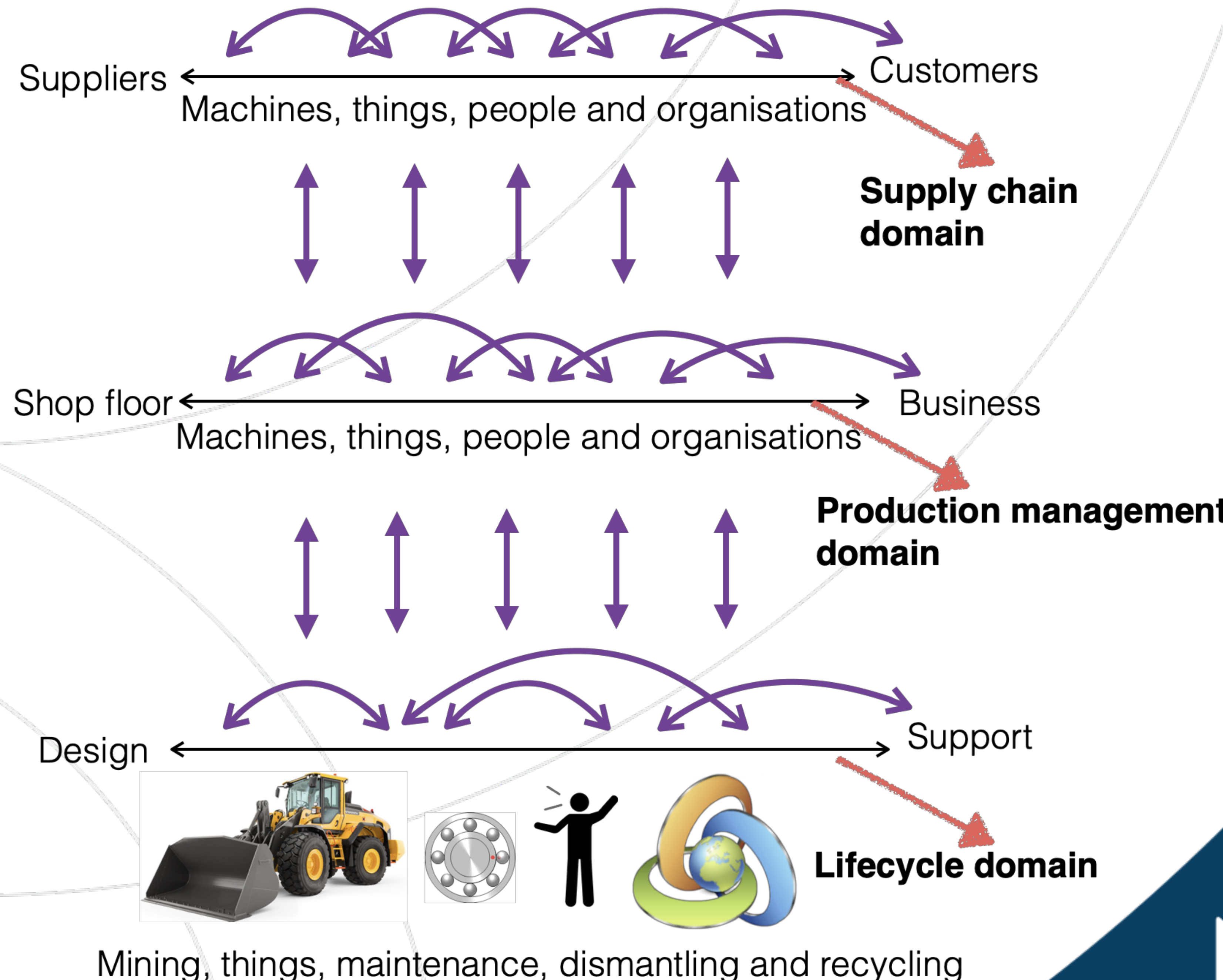
M2M business

Real time monetisation

…

…

…



ARROWHEAD TOOLS

AI and data

- Most AI is today applied on one parameter data sets
- Our facilities have a large set of parameters for which data is collected
- AI have the potential to reduce complexity of multi modal data sets
  - Providing meaningful understanding of our operations, production and products

Where are the AI data bottlenecks

- Access to data?
  - Legally
  - Technically
- Data quality?
- Data documentation?
- Understanding of data?
  - Context is necessary!!
- **The multitude of data models developed over decades and decades to come**

Access to Data from Production

- Production data is currently regarded as key company assets
- Production data exists with many stakeholders in production
- Secure and efficient data sharing is key to many operations
  - Protection for non agreed use of shared data is important
  - No generally accepted solutions!!
    - Early technology initiatives from IDS and GAIA-X

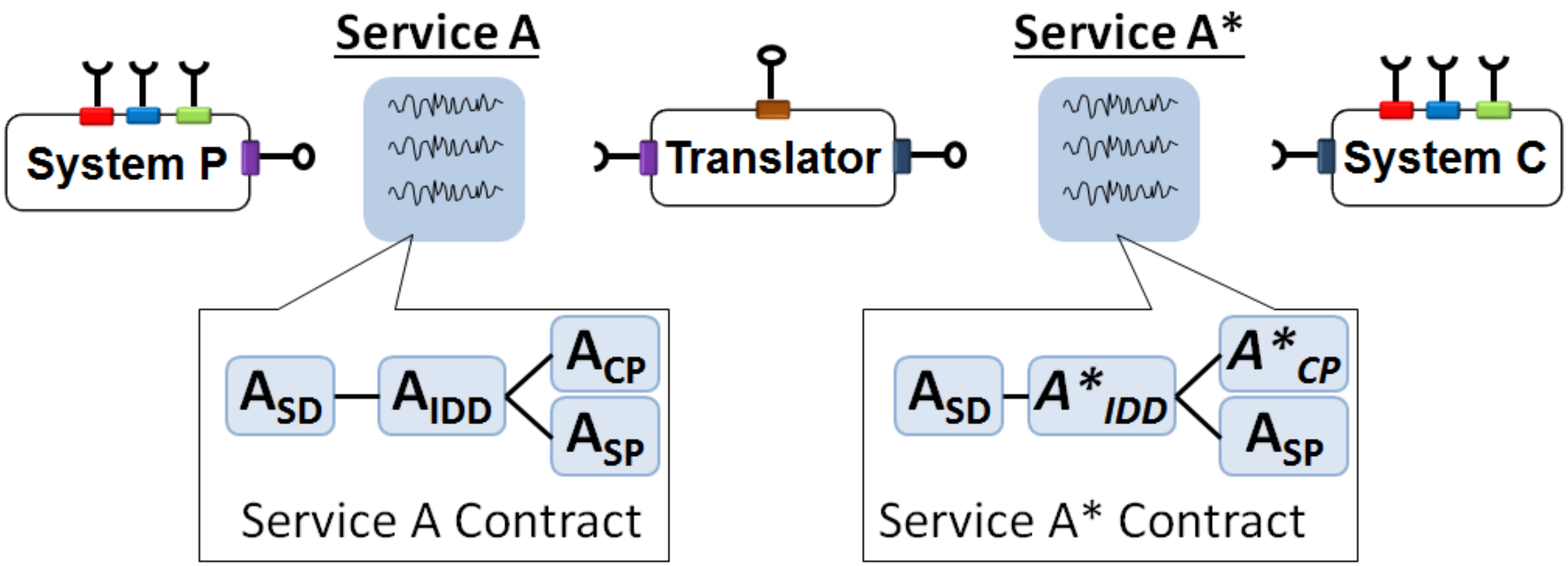
Data sources

- Machines - time series of data
  - The protocol and encoding problem
  - The data model problem
- Context
  - The asset data model problem
- Humans - operators, maintenance, engineering
  - The non digital data problem
    - Digitalisation and understanding of human notes
    - AI based recognition and natural language capturing

Protocol and data encoding problem

Translation between different protocols, and encodings

- e.g. HTTP, CoAP, MQTT, OPC-UA, Modbus…
- e.g. JSON, XML, CBOR, ….



H. Derhamy, J. Eliasson and J. Delsing, "IoT Interoperability—On-Demand and Low Latency Transparent Multiprotocol Translator," in IEEE Internet of Things Journal, vol. 4, no. 5, pp. 1754-1763, Oct. 2017.

ARROWHEAD TOOLS

Machine data model problem

Machine A message

[
{“n”: “OO_temp_sensor”,
“t”: 318350,
“u”: “K”,
“v”: 294.05}
]

Machine B message

[
{“bn”: “temp_sensor”, “bt”: 318350},
{“u”: “Cel”, “v”: 20.9},
{“u”: “Lon”, “v”: “1”},
{“u”: “Lat”, “v”: “-1”}
]

Same standard
Same ontology
Same data
Do not look the same!!

Interoperable????

Machine data model standards

Sensor data

- SenML (RFC 8428) developed by OMA
- SensorML (OGC standard)

Addresses same type of data but are not interoperable

Asset data model standards

Asset descriptions

- ISO 15926
- ISO 10303 (AP 223)
- Asset administration shell DIN ….
- “Old” asset standards, non machine readable
- Loosely documented “standards”

Addresses same type of asset data but are not interoperable

Data model standards and evolution

Domain standards including data models

- Domain does not necessary talk to each other

Standard life times

- New standards are created
- Standards are updated - 5-15 years

Technology life times and update and upgrade cycles

- Asset mechanical lifetime - 20-100 years
- Automation/IT HW lifetime - 10 years
- Automation/IT **SW lifetime - ….. months to a few years?!**

Data model lifetime

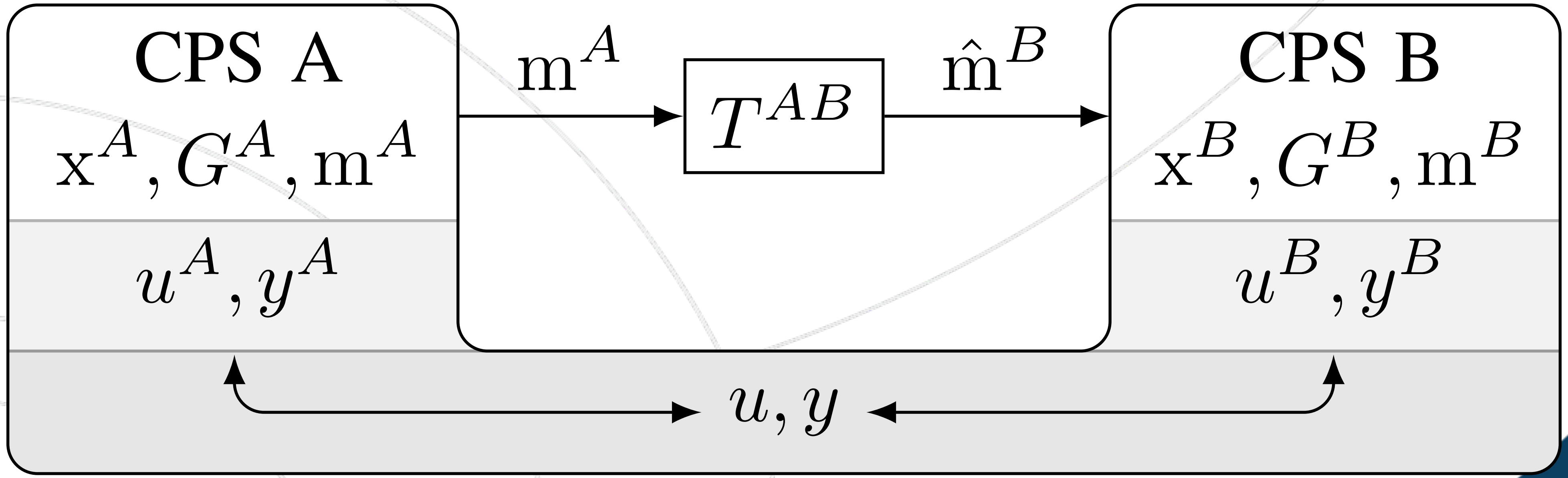
- Semantics/ontologies - 5-10 years
- New once invented every hour

Data interoperability initiatives
Production asset data interoperability initiatives
Paper&Pulp, Oil&Gas

Data model interoperability

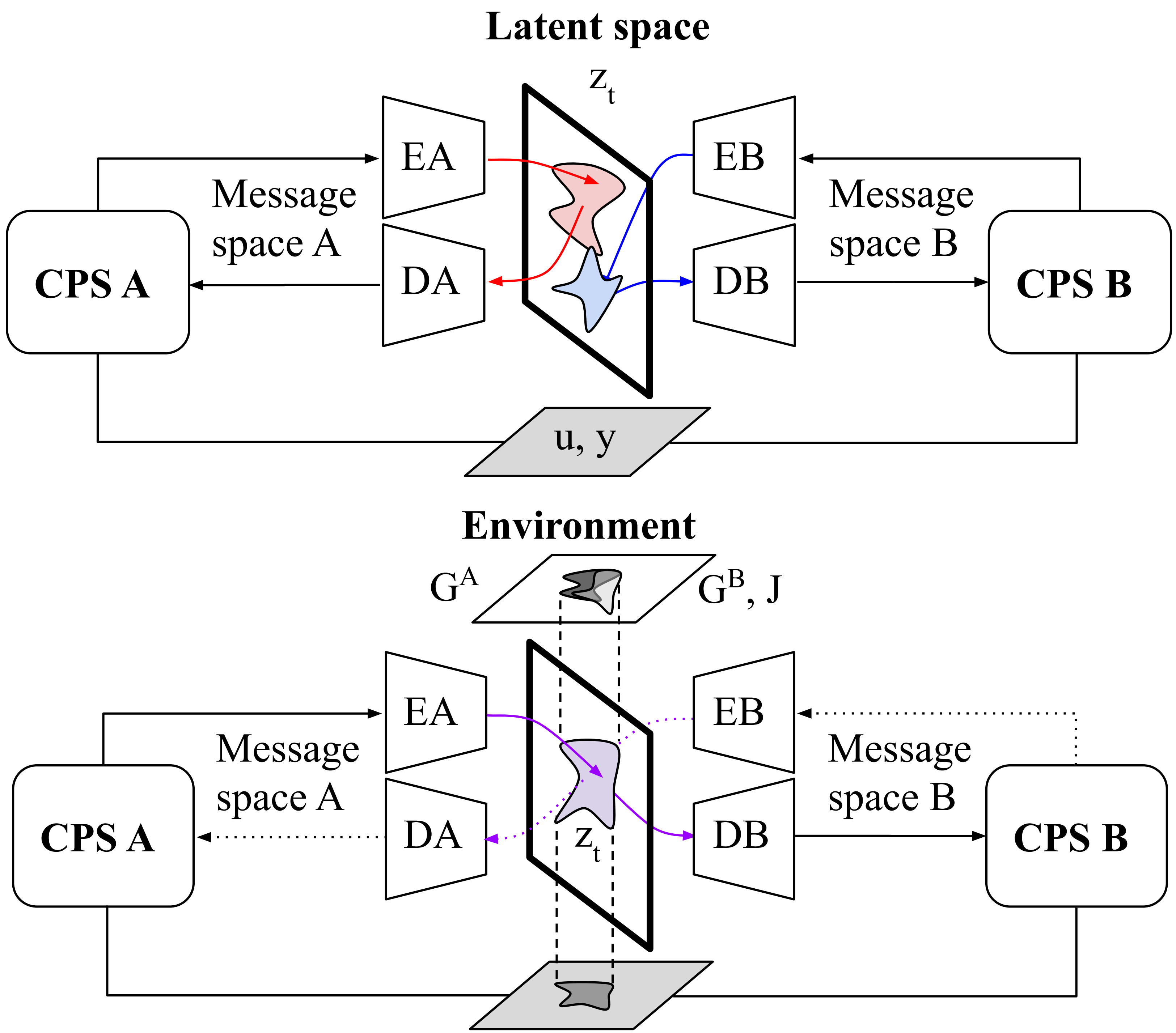
Machine learning approach

- Model of communicating cyber-physical systems (CPS) with different data representations and semantic definitions that interact in a physical environment (gray) and service-oriented architecture (white) via messages m translated by a function T^{AB}



J. Nilsson, F. Sandin and J. Delsing, "Interoperability and machine-to-machine translation model with mappings to machine learning tasks," 2019 IEEE 17th International Conference on Industrial Informatics (INDIN), Helsinki, Finland, 2019, pp. 284-289.

Data semantics translation approach



Time domain data

Context data Availability??

Nilsson, J. (2019). System of Systems Interoperability Machine Learning Model (Licentiate dissertation). Luleå University of Technology. Retrieved from http://urn.kb.se/resolve?urn=urn:nbn:se:ltu:diva-76229

Data semantics interoperability - early results

Semantics translation results

TABLE I: Tested models and their results.

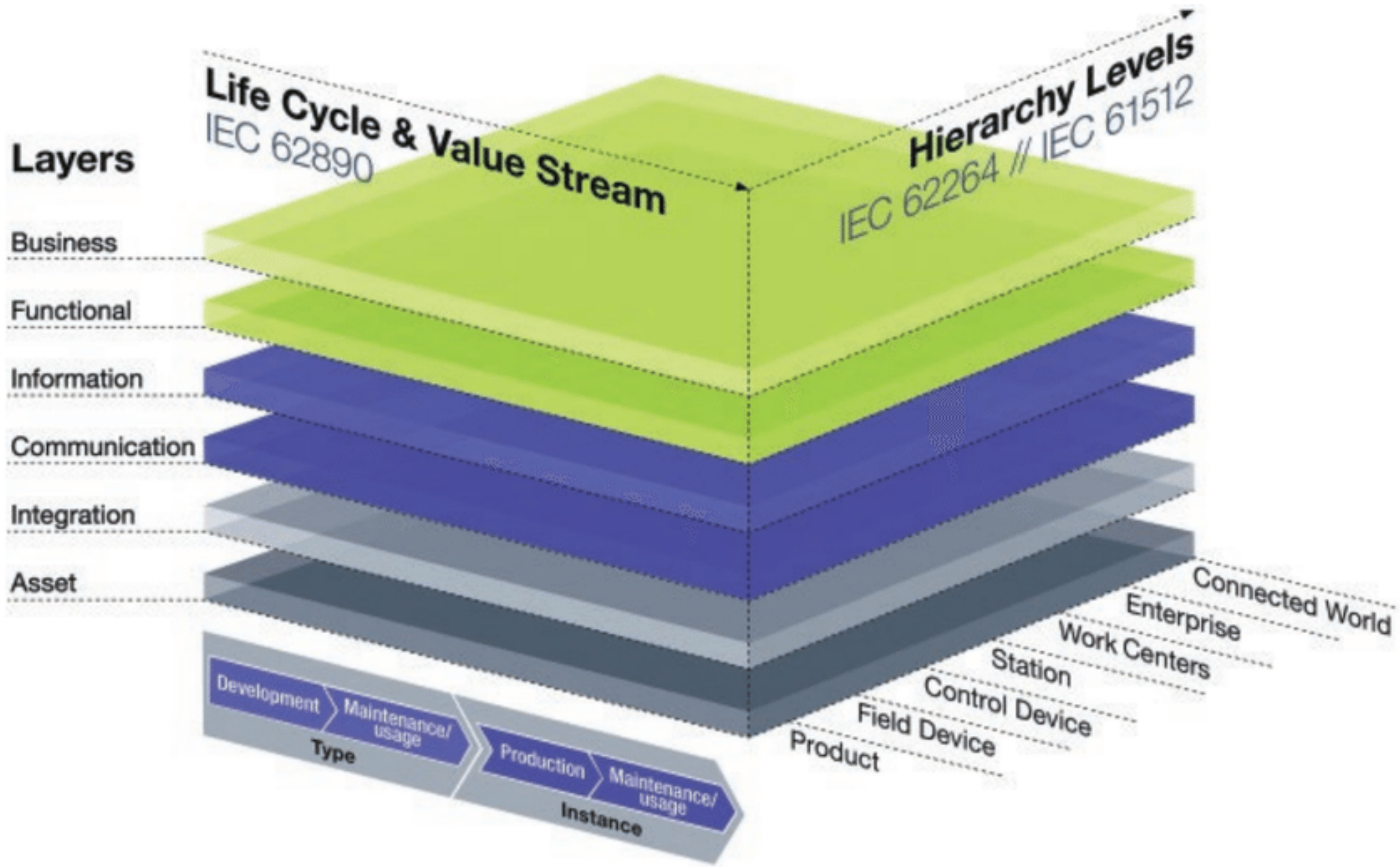
Model	Kind	Strategy	Size	Accuracy		Error	
				Max	Mean	Min	Mean
0	non-shared	2	1-layer	0.70	0.44	0.57	4.0
1	non-shared	2	2-layer	0.73	0.38	0.50	4.9
2	non-shared	1	1-layer	0.66	0.39	0.48	6.7
3	non-shared	1	2-layer	0.74	0.34	0.71	12.0
4	shared	2	2-layer	0.70	0.34	0.54	15.0
5	shared	3	2-layer	0.75	0.41	0.43	2.7
6	shared	1	2-layer	0.69	0.33	0.53	12.0
7	supervised	–	1-layer	1.0	1.0	0.16	0.17
8	supervised	–	2-layer	1.0	0.99	0.16	0.19

How to manage extraction of data for AI from process automation and industrial digitalisation?

ARROWHEAD TOOLS

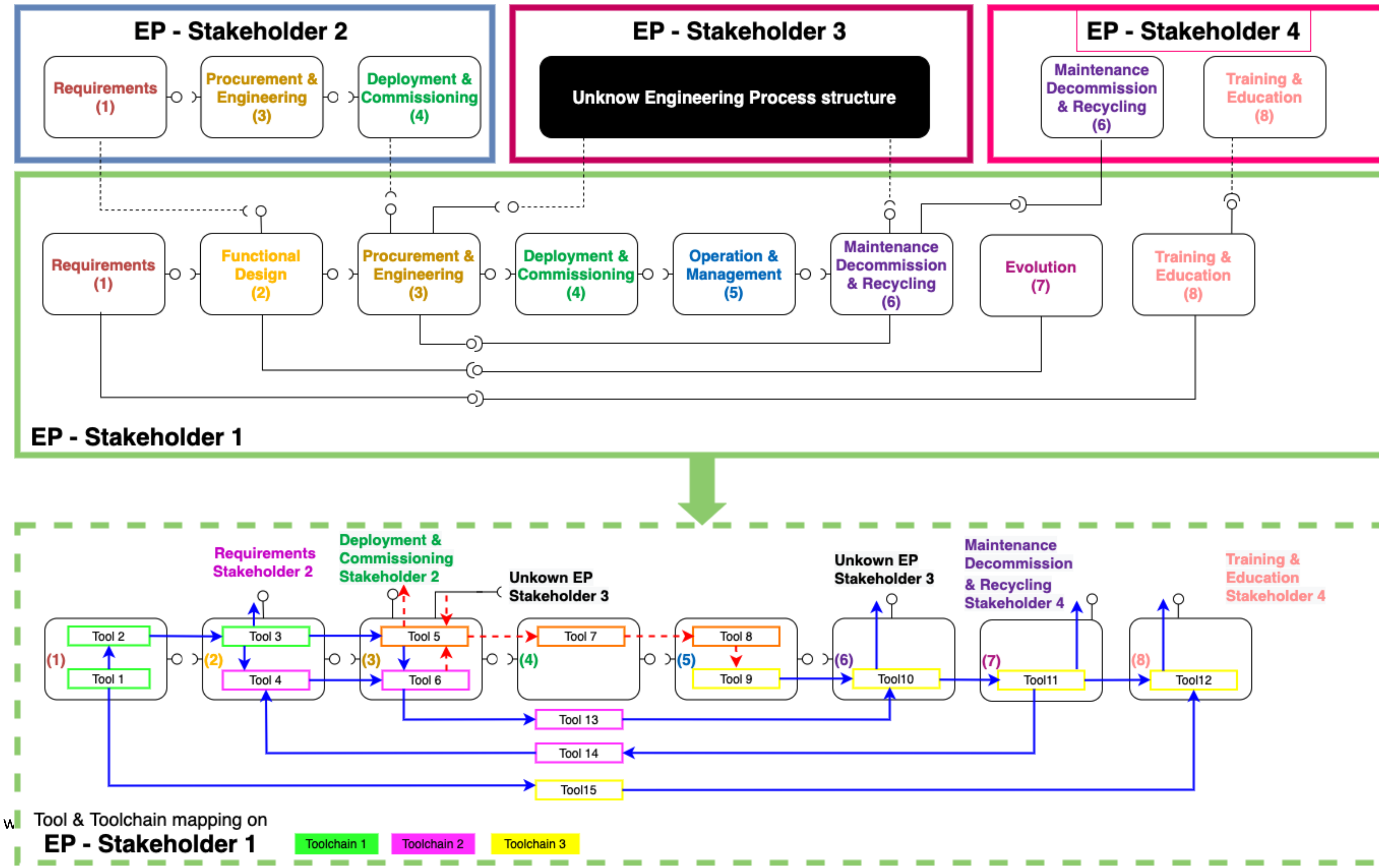
Data generation in multi stakeholder environment

- Data on the engineering process
- Operational data
- Business data
- Lifecycle data
- Stakeholder data
- Context data
- Asset data
- Maintenance data
- Evolution data



ARROWHEAD TOOLS

Engineering process need to consider data capturing in multi stakeholder environment



ARROWHEAD TOOLS

Model supported engineering

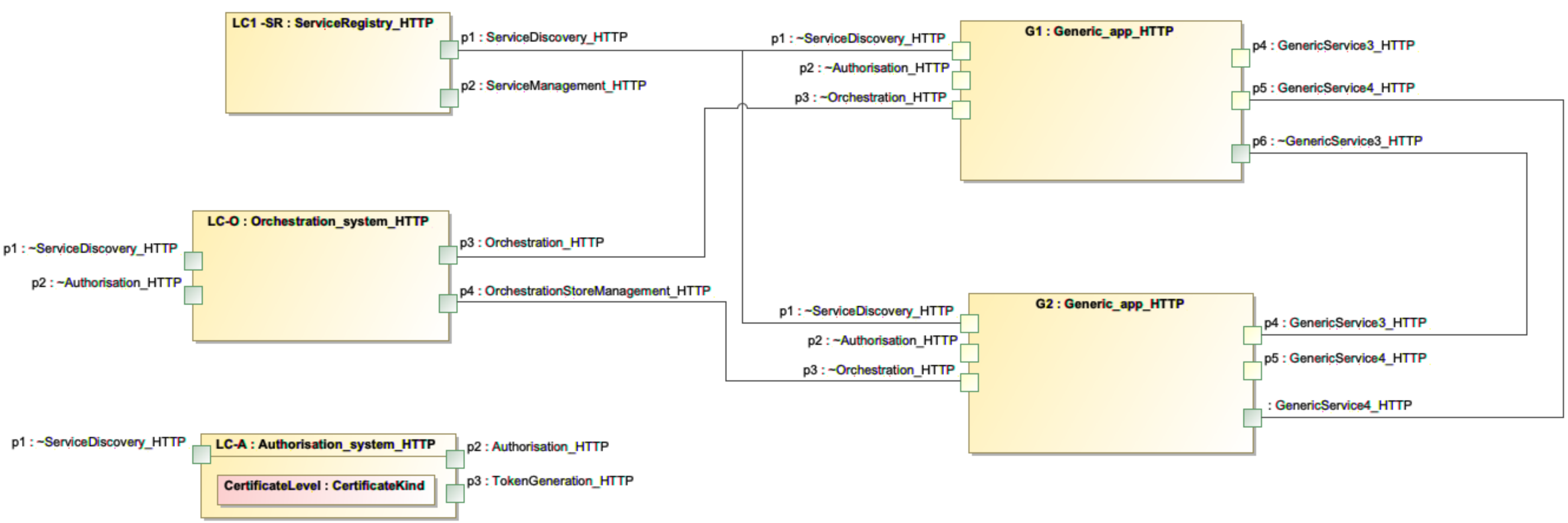
Modelling of solutions - SysML

Machine supported engineering

Automated SW generation

Data extraction supported by modelling guidance

- In the engineering process
- Operations
- …….



In conclusion

- AI in production depends on data availability and accesability
- Complex problems need multimodal data sets
- Data interoperability is fundamental
- Engineering processes need to address data generation
- Operational Strategic Data Management

This research work has been funded by the European Commission, through the European H2020 research and innovation programme, ECSEL Joint Undertaking, and National Funding Authorities from 18 involved countries under the research project Arrowhead Tools with Grant Agreement no. 826452.

ARROWHEAD TOOLS

QUESTIONS

jerker.delsing@ltu.se

LULEÅ UNIVERSITY OF TECHNOLOGY