Managing Knowledge in Energy Data Spaces

Valentina Janev, Maria Esther Vidal,
Kemele Endris, Dea Pujić,

The Mihajlo Pupin Institute
University of Belgrade, Serbia
TIB-Leibniz Information for Centre
for Science and Technology, Germany

TIB LEIBNIZ INFORMATION CENTRE FOR SCIENCE AND TECHNOLOGY UNIVERSITY LIBRARY

institut MIHAJLO PUPIN

Overview

- **Policy Framework** (Energy Data Ecosystems as data-driven infrastructures)
- **Motivation: Energy Data Ecosystem**
  - The Case of Serbia
  - Actors & Challenges
- **Energy Data Space – Vision**
  - Industrial Data Space Concept
  - IMP contribution in the Energy Sector in Serbia
- **IMP Innovative Solutions for the Energy Data Ecosystem**
- **Lessons Learned & Concluding Remarks**

EU Policy Framework

- **Energy system integration** refers to the planning and operating of the energy system “as a whole”, across multiple energy carriers, infrastructures, and consumption sectors, by creating stronger links between them with the objective of delivering low-carbon, reliable and resource-efficient energy services, at the least possible cost for society.
- **Digitalisation** enables dynamic and interlinked flows of energy carriers, allow sfor more diverse markets to be connected with another, and provide the necessary data to match supply and demand at a more disaggregated level and close to real time.

institut MIHAJLO PUPIN

EU Policy Framework

- European Green Deal, December 2019
- European Strategy for Data, February 2020
- Energy System Integration Strategy, July 2020
- Governance Act, November 2020

institut MIHAJLO PUPIN

European Green Deal, December 2019

institut MIHAJLO PUPIN

- Set of policy initiatives by the European Commission with the overarching aim of making **Europe climate neutral in 2050**
- Aims to boost the efficient use of resources by moving to a clean, circular economy and stop climate change, revert biodiversity loss and cut pollution
- For the European union to reach their target of climate neutrality, one goal is to decarbonise their energy system by aiming to achieve "net-zero greenhouse gas emissions by 2050
- **Guidelines for the Implementation of the Green Agenda for the Western Balkans, October 2020**
  - Clean Energy Transition - reduce energy imports, develop renewable energy sources, strengthen regional energy security, unlock greater economic growth
  - EU's Framework Programme for Research and Innovation

https://ec.europa.eu/info/strategy/priorities-2019-2024/european-green-deal_en

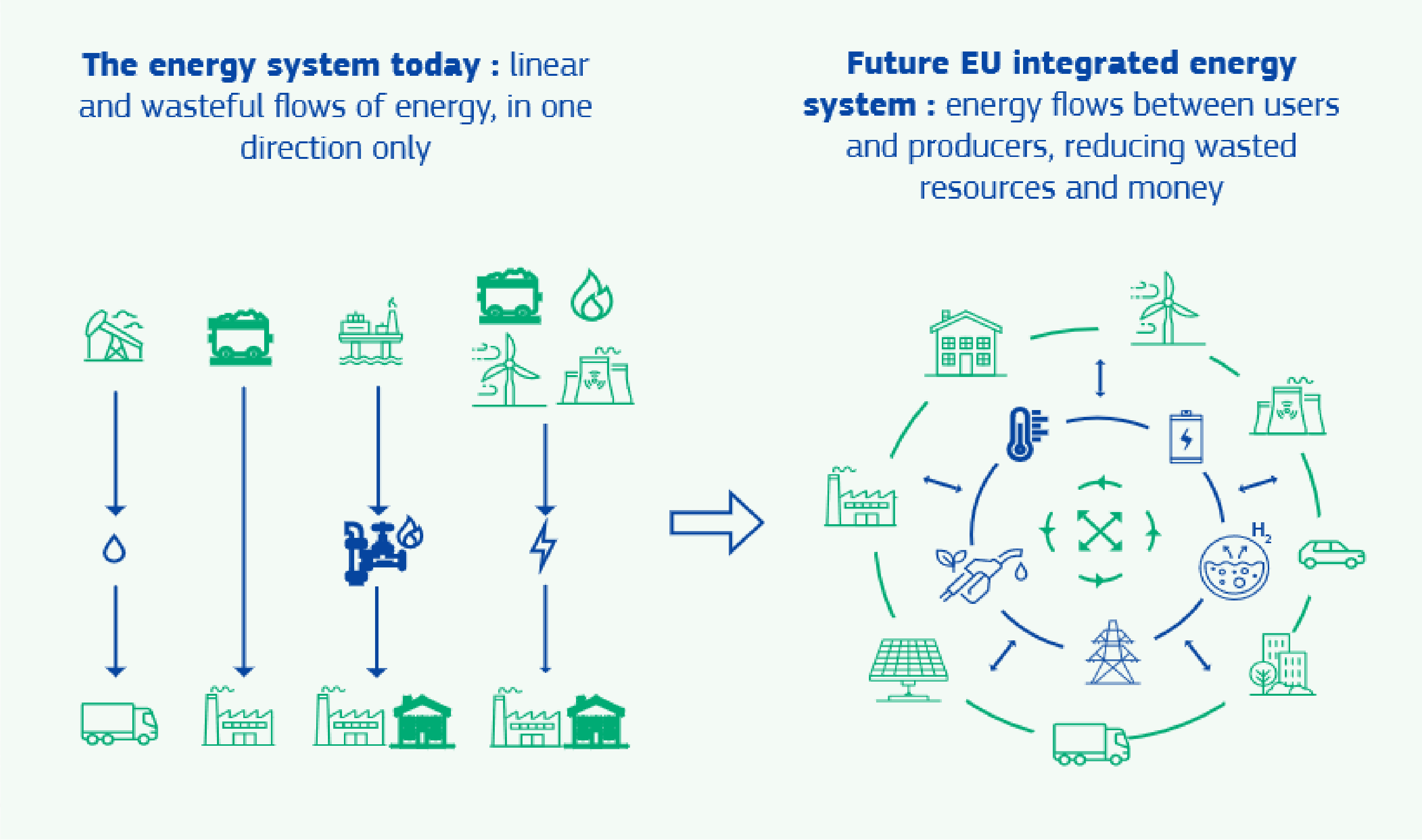
European Strategy for Data, February 2020

- aim is to create a single **European data space** where personal as well as non-personal data, including sensitive business data, are secure and businesses also have easy access to an almost infinite amount of high-quality industrial data, boosting growth and creating value, while **minimising the human carbon and environmental footprint**.

- Technical tools for data pooling and sharing
- Standards and interoperability (technical, semantic)
- Sectoral Data Governance (licensees, access rights, usage rights)
- IT capacity, including cloud storage, processing and services

Energy System Integration Strategy, July 2020

- This strategy will profoundly reform the European energy system.
- We are designing a more efficient and integrated system that links energy sources and infrastructure to support decarbonisation and build a climate neutral EU by 2050.
- It will help to build modern infrastructure, make European industry more sustainable and competitive, create jobs, and provide clean energy for citizens.



https://ec.europa.eu/energy/topics/energy-system-integration/eu-strategy-energy-system-integration_en

Data Governance Act, November 2020

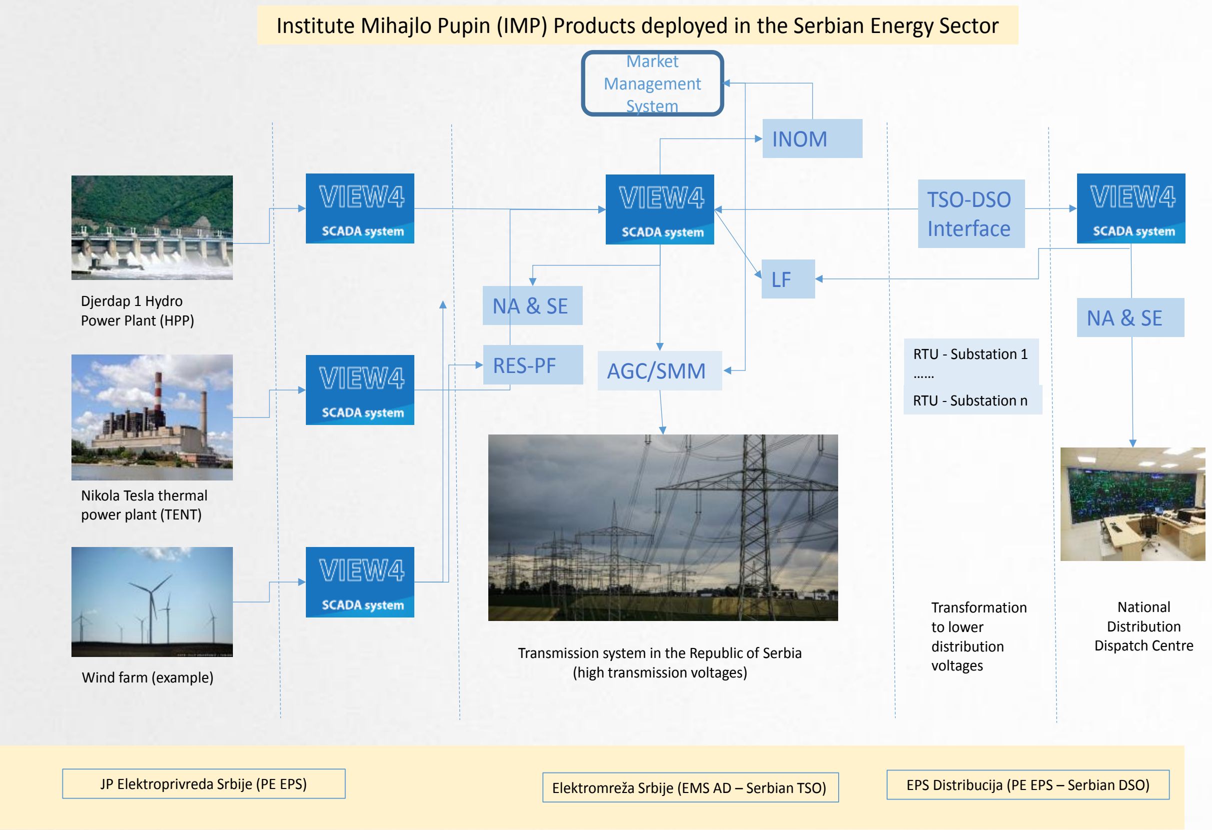
- Data Space construction is a multilayer, interdisciplinary, and has technically several complexities (cybersecurity, software architectures, interoperability, standards, etc.
- Data Governance Act defines measures to foster the availability of data for use by increasing trust in data intermediaries and by strengthening data-sharing mechanisms across the EU.
- Trustworthy data-sharing systems through four broad sets of measures:
  - Mechanisms to facilitate the reuse of certain public sector data that cannot be made available as open data.
  - Measures to ensure that **data intermediaries will function as trustworthy organisers** of data sharing or pooling within the common European data spaces.
  - Measures to make it easier for citizens and businesses to make their data available for the benefit of society.
  - Measures to **facilitate data sharing**, in particular to make it possible for data to be used **across sectors and borders**, and to enable the right data to be found for the right purpose.
- https://digital-strategy.ec.europa.eu/en/policies/data-governance

Motivation: Energy Data Ecosystem

- Actors and Challenges
- Integration and Interoperability

institut MIHAJLO PUPIN

Motivation (Example – Country level)



- **Energy management deals with monitoring and controlling the energy production, distribution and usage with different objectives including improvement of energy efficiency, increasing the flexibility and renewable generation share, and reducing the energy cost, e.g.** the Serbian Energy Value Chain
- **Challenges**
  - **Digitalization of the energy sector** Energy Management Applications are fragmented, developed against energy data silos, and data exchange is limited to few applications
  - **Big Data in the energy domain**
  - **Integration of renewable energy sources (RES)**

Acknowledgement

PLATOON - Digital PLAtform and analytical TOOls for eNergy
ARTEMIS - ARTificial Intelligence in Energy Management Innovative Services
TRINITY – TRansmission system enhancement of reg**I**o**N**al borders by means of **I**ntelligen**T** market technolog**Y**

Integration of renewable energy sources

- VIEW4 SCADA
- DSC - Digital Control Systems

- Commercial Projects, https://www.pupin.rs/en/references

Actors and Challenges

institut MIHAJLO PUPIN

ACTOR	CHALLENGES
TRADITIONAL POWER GENERATION COMPANY	Prepare to face the SmartGrid challenges soon; Use all the available information to optimize operations and extend asset useful life; Provide value to existing, often old generation facilities and be able to switch to best- in-class solutions
TRANSMISSION SYSTEM OPERATOR	Extend the portfolio of energy services that can be provided; Optimize energy sale strategies; Use all the available information to optimize operations and extend asset useful life
DISTRIBUTION SYSTEM OPERATOR	Manage hundreds/thousands of assets geographically distributed consumers; Extend asset useful life

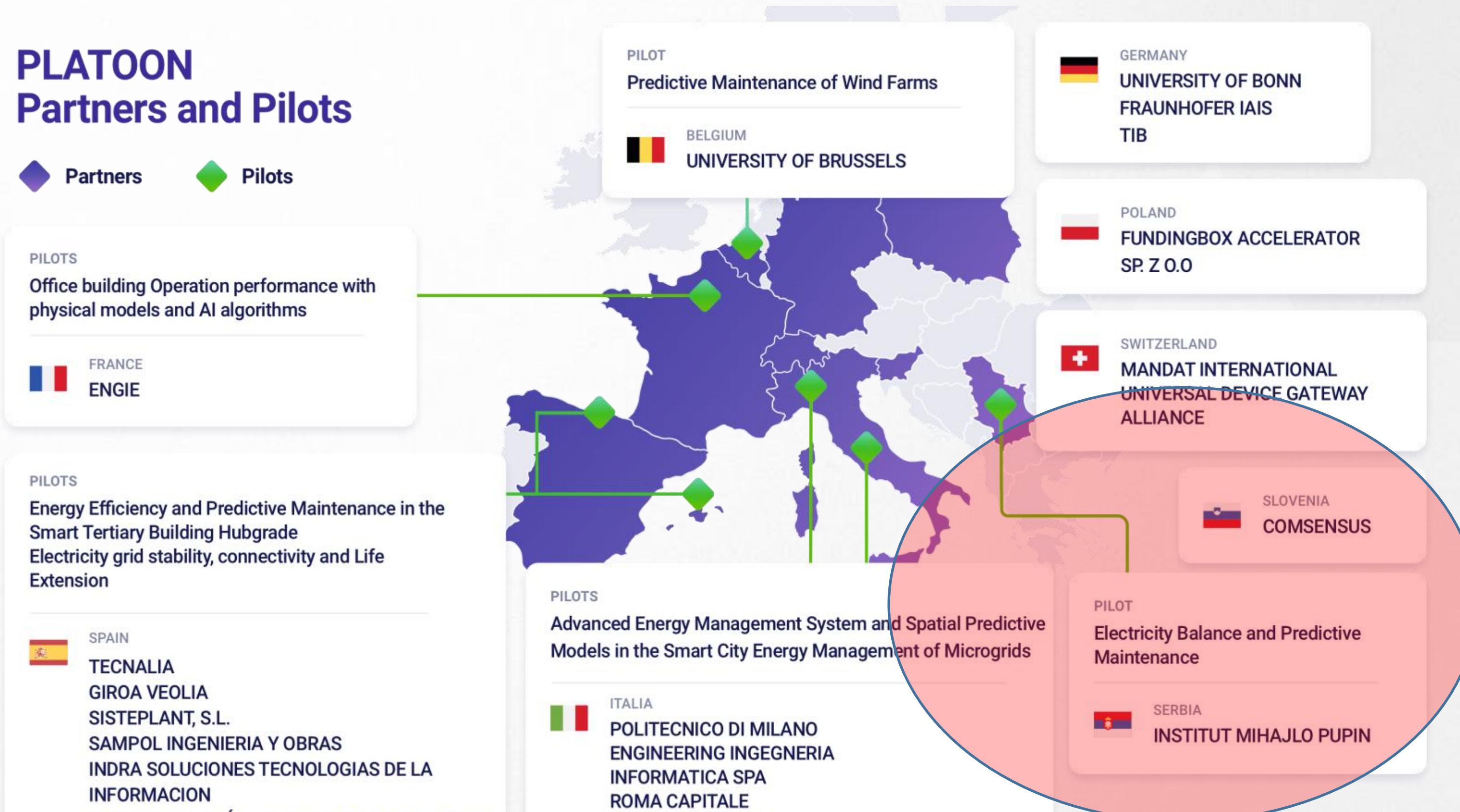
Actors and Challenges

institut MIHAJLO PUPIN

ACTOR	CHALLENGES
RENEWABLE POWER GENERATION COMPANIES	Manage geographically distributed asset fleets; Optimize energy sale strategies; Forecast producible power and optimize energy bids
ICT SUPPLIERS / TECHNOLOGY PROVIDERS	Develop specific solutions to address industry problems; Promote the use of their platforms and solutions (e.g. Institute Mihajlo Pupin SCADA System); Extend the portfolio of services to be provided
AGGREGATORS AND ENERGY SERVICES PROVIDERS	Play a role in the energy market by grouping together the interests of many individual independent producers; Extend the portfolio of energy services that can be provided ; Integrate multi-source data to benefit from opportunities and provide value to customers; Use multi-source data (weather, prices, etc.) to determine optimum consumption strategies; Optimize energy consumption, lowering costs; Extend the portfolio of energy services that can be provided

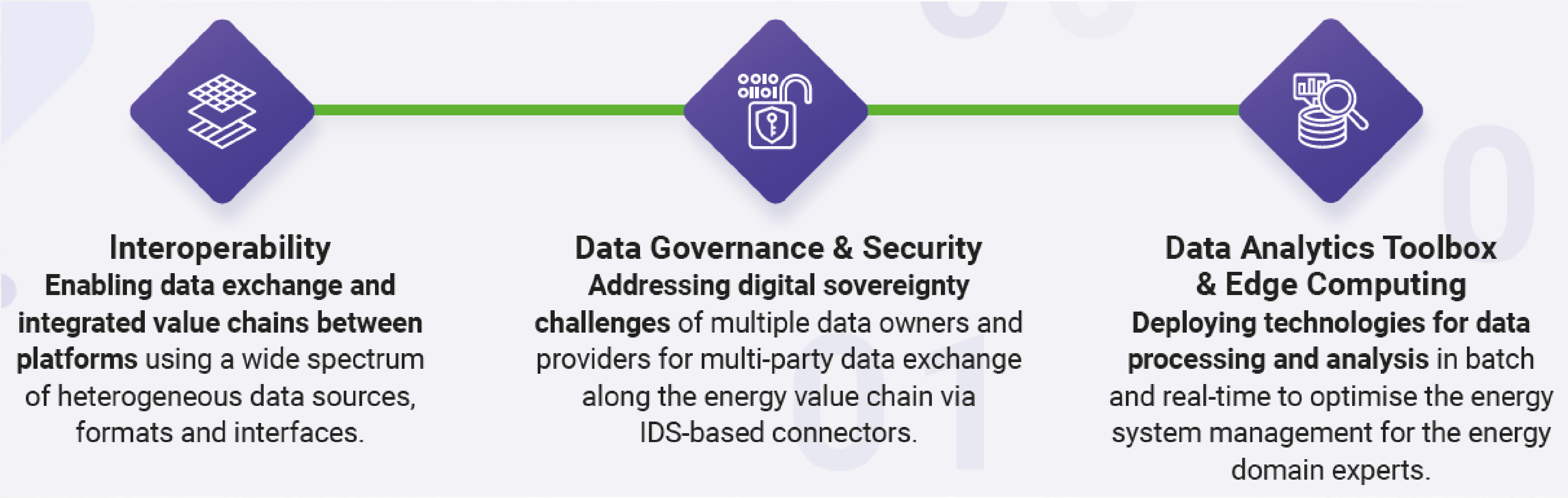
Motivation (Example – EU level)

- One of the long-term objectives of the EU is creation of common market that will eliminate trade barriers between EU Member States
- PLATOON – [Digital PLAtform and analytical TOOls for eNergy]
- Several DEs will be interconnected into a DE network
  - Data sovereignty
  - Decentralized soft architectures



Example: PLATOON Architecture for Smart Grid

institut MIHAJLO PUPIN



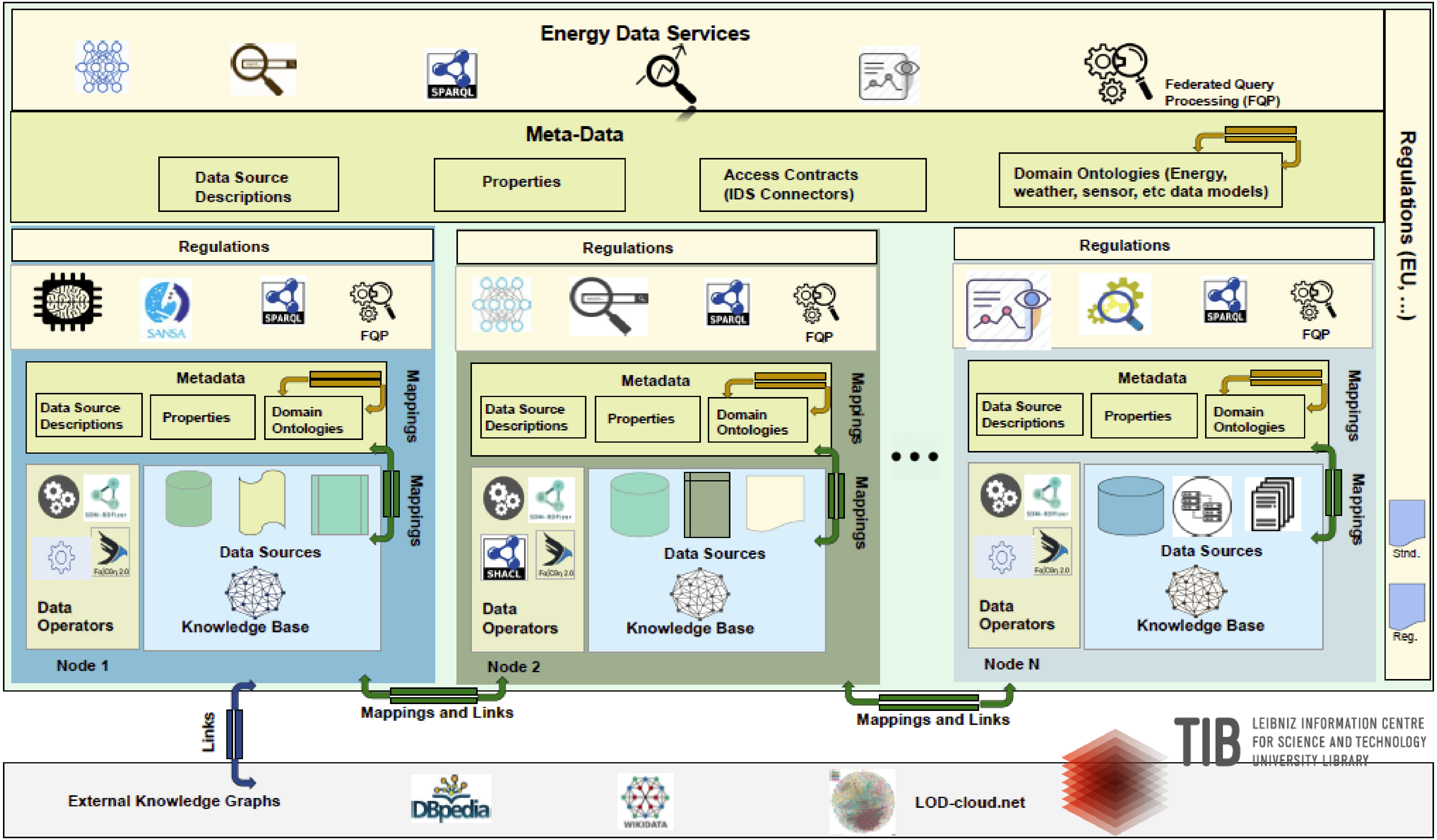
- Real-time integration and Big Data analysis needed upon the high-volume data streams from metering devices and power grid elements (e.g. switches, transformers, etc.)
- Decentralised Data Processing Architectures needed for processing multi-stream datasets of different velocity
- Variability and degree of uncertainty of power output from renewable sources increases with penetration of distributed generation (Wind / PV / Solar Power Plants) and data analytics toolbox (e.g. Accuracy of forecasting, production) and edge computing solutions are needed for optimised real-time energy system management

Energy Data Space – Vision

- Industrial Data Space Concept
- IMP role in the Energy Sector in Serbia

BDV BIG DATA VALUE ASSOCIATION

iMP institut MIHAJLO PUPIN



Energy Data Services
Federated Query Processing (FQP)
SPARQL
Meta-Data
Data Source Descriptions
Properties
Access Contracts (IDS Connectors)
Domain Ontologies (Energy, weather, sensor, etc data models)
Regulations (EU, ...)
Regulations
SANSA
FQP
Metadata
Domain Ontologies
Mappings
Data Sources
Knowledge Base
Data Operators
SHACL
Node 1
Node 2
Node N
Stnd.
Reg.
Mappings and Links
Links
External Knowledge Graphs
DBpedia
WIKIDATA
LOD-cloud.net
TIB
LEIBNIZ INFORMATION CENTRE FOR SCIENCE AND TECHNOLOGY UNIVERSITY LIBRARY

Standards for Energy Data Ecosystems

institut MIHAJLO PUPIN

- **Smart Grid Architecture Model (SGAM)** that is a product of the standardization process in the EU Mandate M/490, the work of the CEN-CENELEC-ETSI Smart Grid Coordination Group
- **THE INTERNATIONAL DATA SPACE (IDS) Information Model** Virtual data space leveraging existing standard and technologies, as well as governance models well accepted in the data economy, to facilitate secure and standardized data exchange and data linkage in a **trusted business ecosystem**

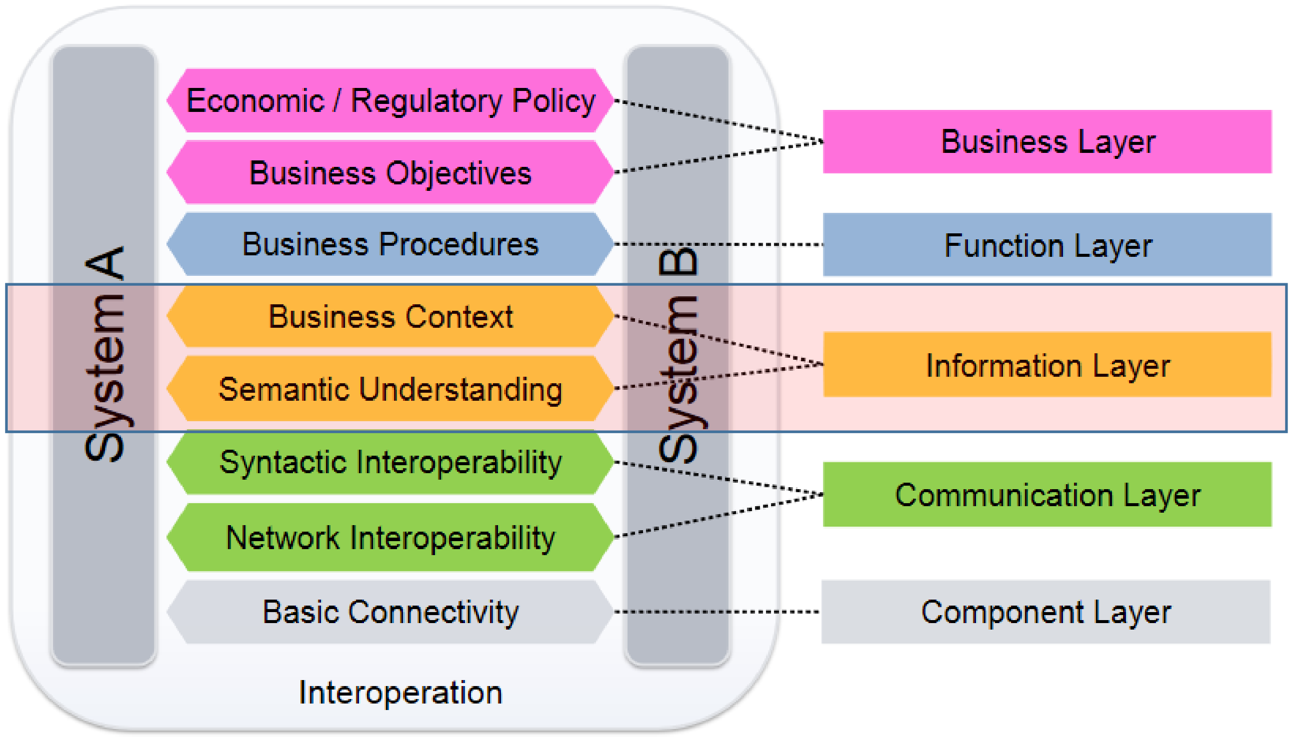


Figure 6: Grouping into interoperability layers

CEN-CENELEC-ETSI Smart Grid Coordination Group:
Smart Grid Reference Architecture

Integrated Energy Value Chains

- Data sources may have different data models, follow various data representation schemes, and contain complementary information
- New smart grids services needed for effective and scalable semantic interoperability and creation of data spaces (also supported with EU Data Strategy)
- message-based infrastructure needed to enable the communication of the different nodes and components in the energy value chain and integration in the European Energy Data Space

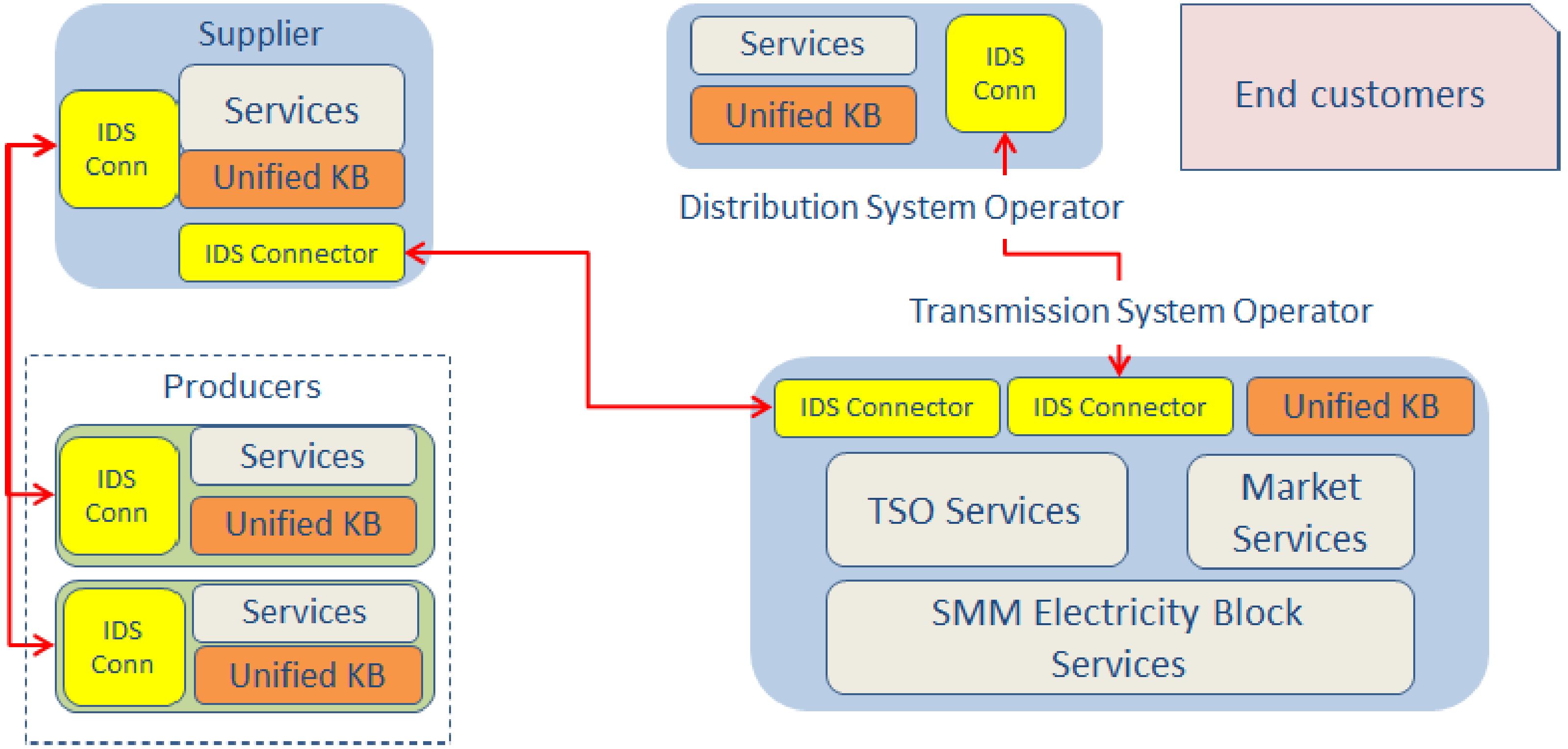


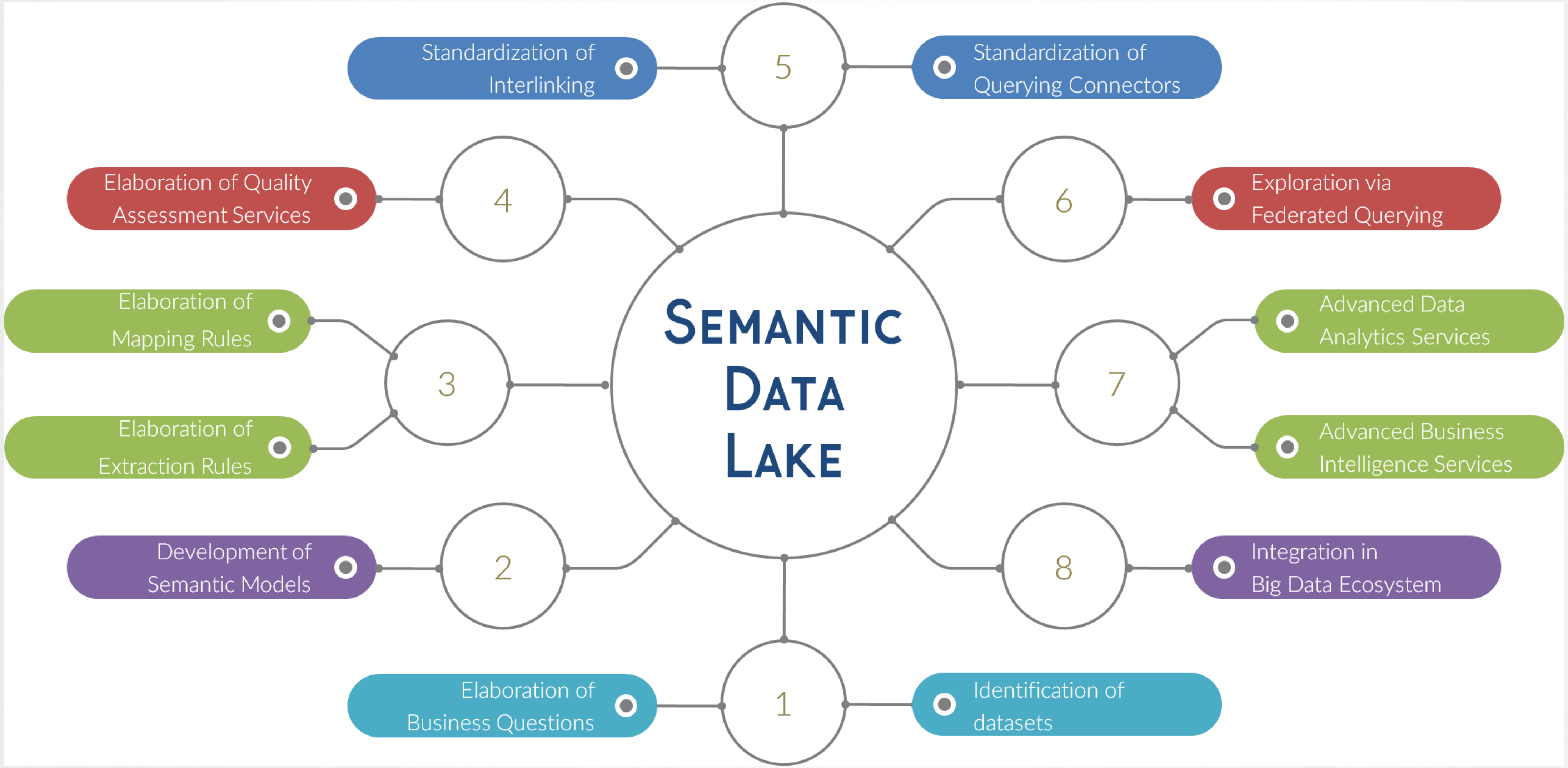
Figure 3: Multi-party data exchange based on IDS concept

APPLICATION: A USE CASE

- Semantic pipline
- Developing a Global Schema for the Energy Domain
- Unified Knowledge Graph Creation Process
- Traversing the Knowledge Graph
- Federated Query Processing

institut MIHAJLO PUPIN

Methodology



Example Scenarios

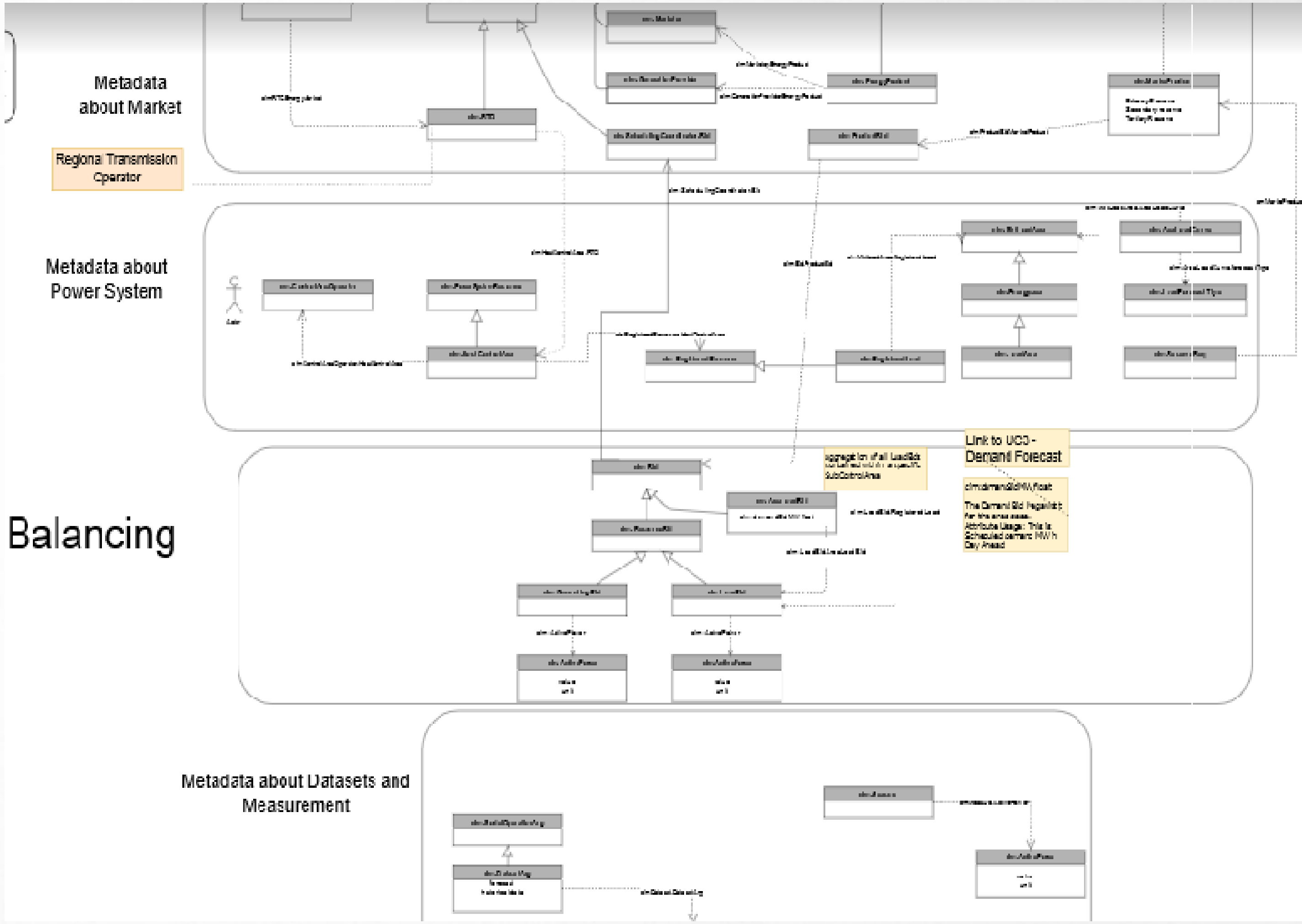
TransmissionSystem Operator (TSO):
Load/Demand forecasting. The aims of this use case are load forecasting and prediction of the load pattern. It involves accurate prediction of both magnitudes and geographical locations of electric load over the different periods of the planning horizon. Load forecasting is divided into three categories: short-term forecasts, medium-term forecasts, and long-term forecasts which is a crucial part in the electric power system planning, tariff regulation and energy trading.

Renewable Energy Sources (RES) Producer:
RES forecasting. RES allows prediction of when the transformer connecting the distribution grid to the transmission grid will be overloaded, i.e., when local wind turbine generator production will be very high. The various forecasting approaches can be classified according to the type of input (weather prediction, wind turbine generators data, historical production data). Statistically based approaches allow very short-term predictions (2 hours).

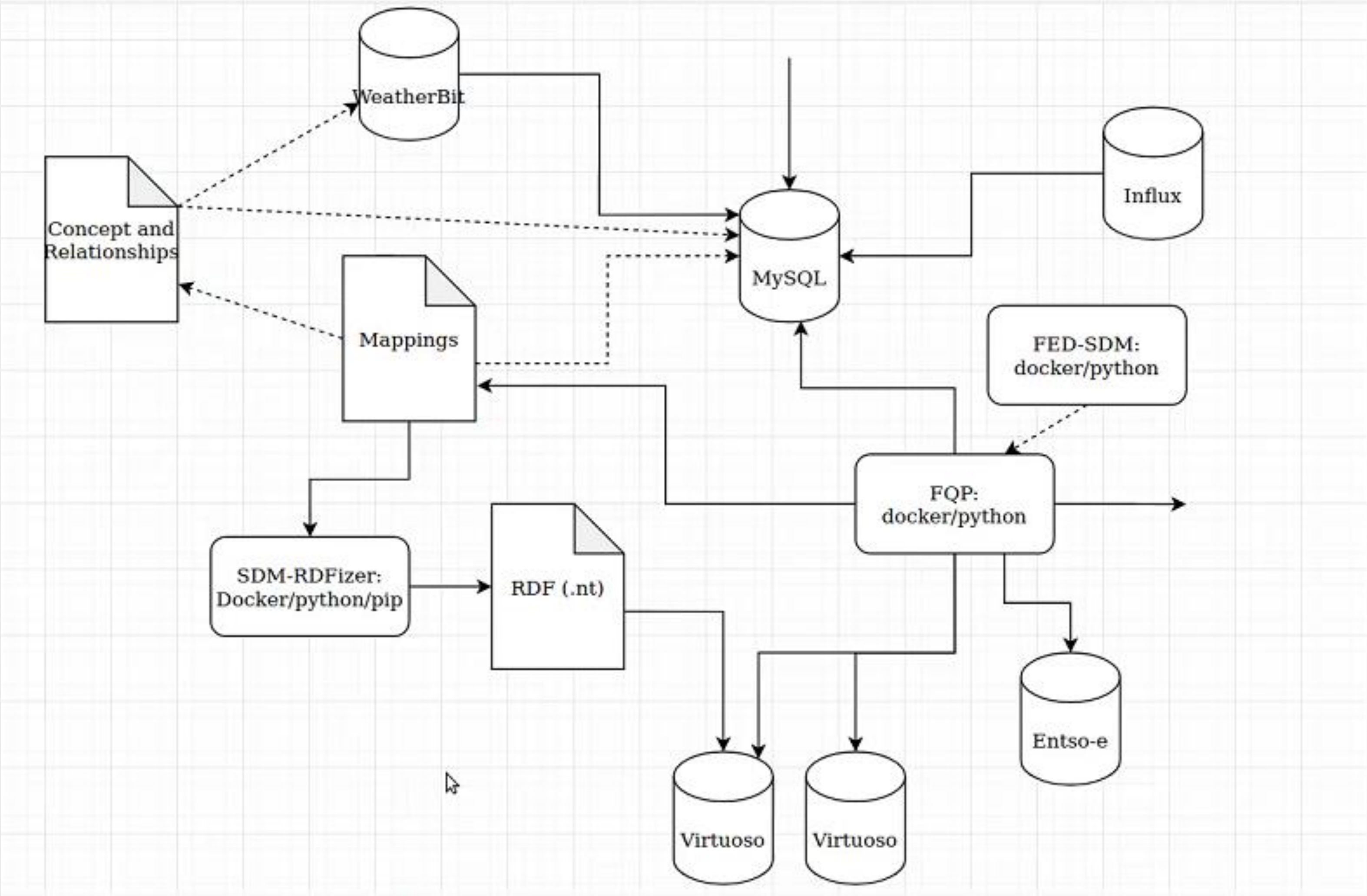
Producer:
Predictive maintenance. The aim of this use case is to provide services that able to identify rare events that could occur in power plant infrastructure due to infrastructure health problems, progressive degradation or failure.

Global Schema for the Energy Domain

CIM	**Common Information Model, https://ontology.tno.nl/IEC_CIM/**
DCMI	Dublin Core Metadata Initiative, https://dublincore.org/
SAREF	Smart Appliances REFerence ontology https://w3id.org/def/saref4city https://w3id.org/def/saref4bldg https://w3id.org/def/saref4ener/ https://w3id.org/def/saref4agri/ https://w3id.org/def/saref4envi/
IDS	Industrial Data Space (IDS) Information Model



Unified Knowledge Graph Creation Process



- Materialized Knowledge Graph Creation Process
- Virtual Knowledge Graph Creation Process

Federated Query Processing

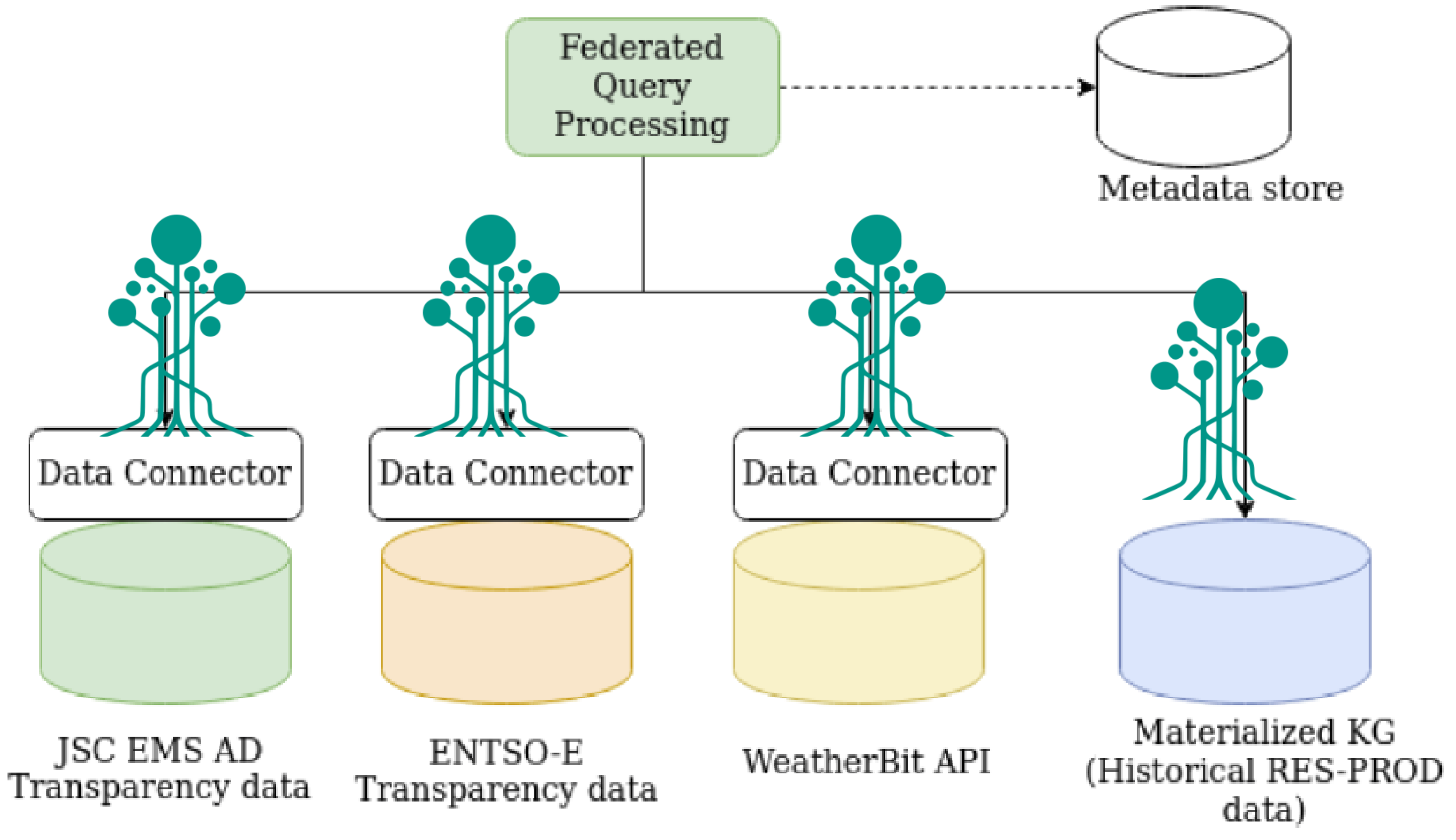


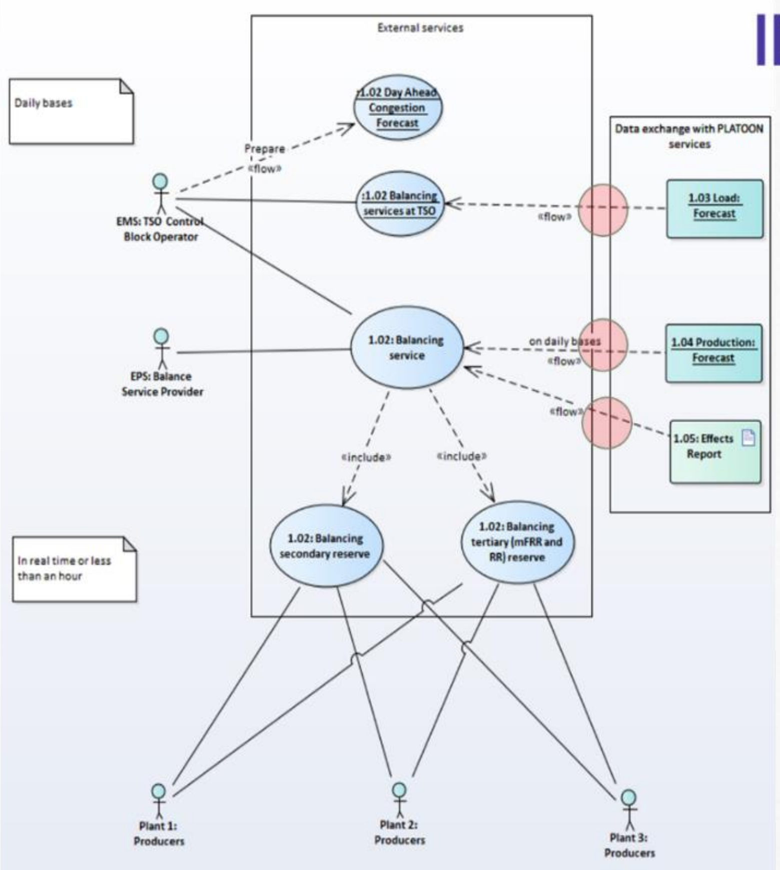
Figure 5: Federated Query Processing at TSO level

- Federated query processing system provides a unified access interface to a set of autonomous, distributed, and heterogeneous data sources.
- While distributed query processing systems have control over each data set, federated query processing engines have no control over data sets in the federation, and data providers can join or leave the federation at any time and modify their data sets independently.

https://www.dataspace-connector.io/

Deployment - Integration with existing systems

- Unified Knowledge Base
- Semantic Adapters and Federated Querying
- IDS-based connectors
- Edge services

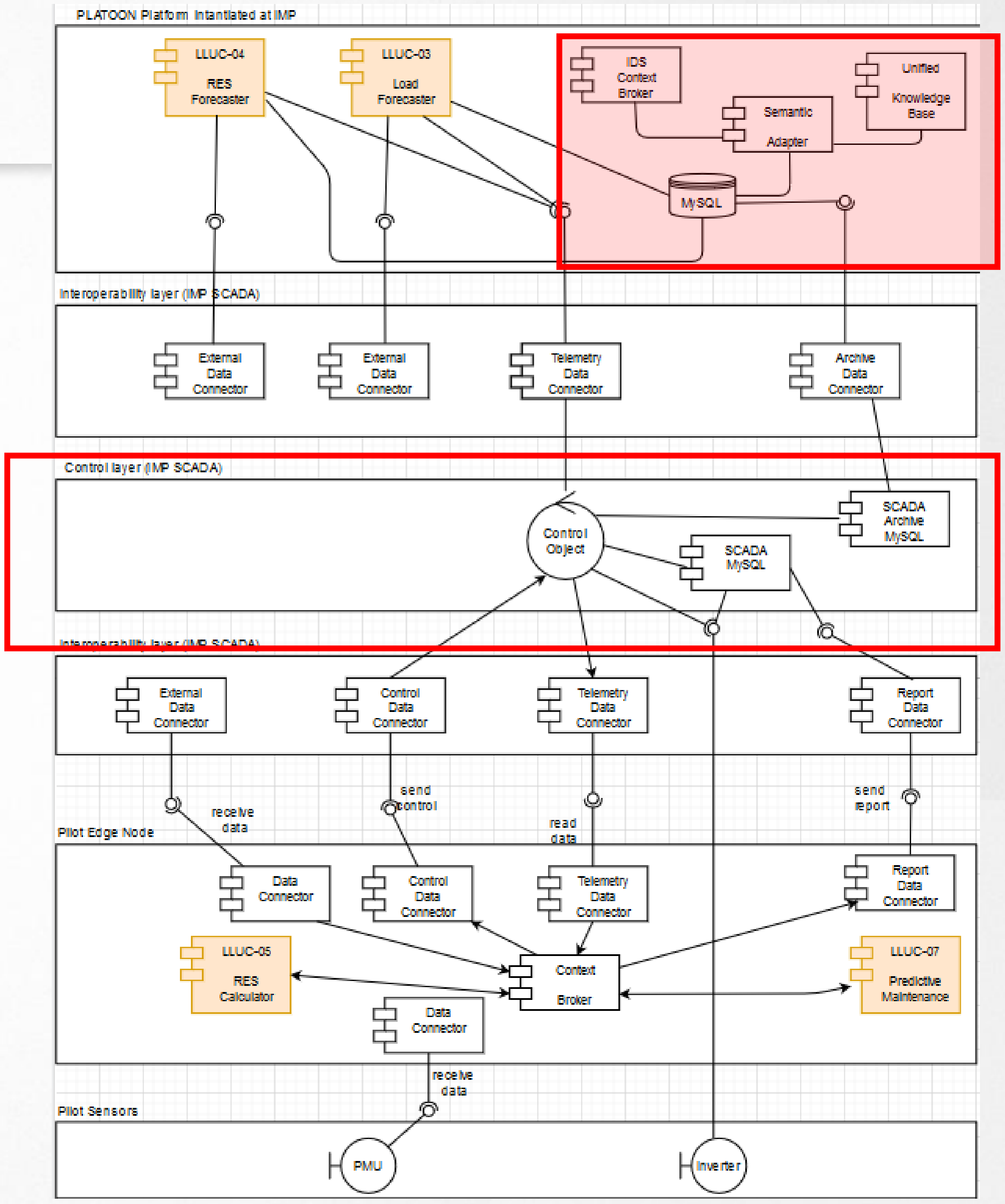


IDS Compliant Services

- Load Forecasters
- RES Forecasters
- Power Quality services for RES Effects calculation

IDS Based Architecture

- **Hybrid data and model driven local production forecasting** offering energy generation prediction based on physical models and measured data (black/grey box modelling)
- **Data-driven energy demand forecasting** providing short and long-term forecast of different load types (i.e. electricity, heating and cooling)
- **Hybrid ML-enabled energy performance assessment** delivering a measure of consumer energy efficiency by normalizing energy consumption against context-related (e.g. climate, construction type, number of inhabitants etc.) and behavior-related aspects (e.g. deviations from expected consumption), which will be used for consumer benchmarking



Concluding Remarks

Deployment of Innovations

- Large-scale penetration of renewables implies
  - Long-term energy supply sustainability
  - Decrease of Green House Gas emissions
  - Avoidance of energy distribution costs and losses

- However requires
  - Adapted power system planning and more flexible use of power plants
  - Standardized data exchange
  - Integration of services in integrated value chains
  - ...and others.

Benefits

institut MIHAJLO PUPIN

TITLE	BENEFITS
BALANCING IN CONTROL BLOCK	Improved cross-border trading
BALANCING THE GRID - DEMAND FORECASTING	Better matching of demand and supply across the energy mix; Standardize the interfacing services with the production plants
RES (WIND GENERATION) FORECASTERS	Increased reliability of supply; Increased renewable energy penetration; large-scale integration of RES
EFFECTS OF RENEWABLE ENERGY SOURCES ON THE POWER SYSTEM (DISTRIBUTION LEVEL)	Reduce the cost of Operations and improved quality of service – management of unexpected variations to voltage profile of the power system
PREDICTIVE MAINTENANCE IN RES POWER PLANTS	improved plant maintenance and optimized asset performance.

Conclusion

- Integration challenges
  - Real data in different formats
  - Granularity of observed data is not the same (from ms, s, hour, month)
- Interoperability challenges
  - Alignment of common vocabularies and ontologies
  - Selection and further development of models based on the target scenarios (e.g. for forecasting)
- Analytics Solutions and explainability challenges
  - More accurate predictive models
  - Interpretability of the Output of Predictive Models
  - Enhanced Reasoning

About Institute Mihajlo Pupin

institut MIHAJLO PUPIN

- 40 years in Process Control Systems (Thermal and Hydro PP)
- Supervisory Control and Data Acquisition Systems (SCADA) and Digital Control Systems (DCS)
- 10 years in Solar and Wind PP, Natural Gas PP, Biomass PP

99 Since 2004

- Research related to **EU climate and energy targets ("20-20-20")**
  - **10 EU projects just in the last 3 years**
    - **InBetween** - ICT enabled BEhavioral change ToWards Energy EfficieNt lifestyles; **IDEAS** - Novel building Integration Designs for increased Efficiencies in Advanced climatically tunable renewable energy Systems; **RESPOND** - Integrated demand REsponse Solution towards energy POsitive NeighbourhooDs; **REACT** - Renewable Energy for self-sustAinable island CommuniTies, …
    - **TRINITY** - TRansmission system enhancement of regIoNal borders by means of IntellIgenT market technology; PLATOON - Digital platform and analytical tools for energy, …
  - **SINERGY -** Capacity building in Smart and Innovative eNERGY management - Regional Centre of Excellence
  - **LAMBDA** – Learning, Applying, Multiplying Big Data Analytics - Regional Centre of Excellence
- Research related to Circular Economy
  - Danube-goes-Circular, **CIP/EIP** (GREEN, WEEEN, IMAGEEN), https://www.pupin.rs/en/research-and-development-projects/european-rd-projects/

Questions?

Thank you for your attention!

[Big Data Analytics Summer School, Belgrade, Serbia, June 2021]

[Doctoral Workshop 2021] **(deadline April 2021)**