Distributed Data Infrastructures in Energy Systems Analysis

Carsten Hoyer-Klick

Johannes Frey, Ulrich Frey, Anastasis Giannousakis, Tobias Hecking, Sebastian Hellmann, Christian Hofmann, Ludwig Hülk, Sophie Jentzsch, Michaja Pehl, Marion Schroedter-Homscheidt, Vera Sehn

Supported by:

Federal Ministry for Economic Affairs and Energy

on the basis of a decision by the German Bundestag

Project Number 03EI1005A

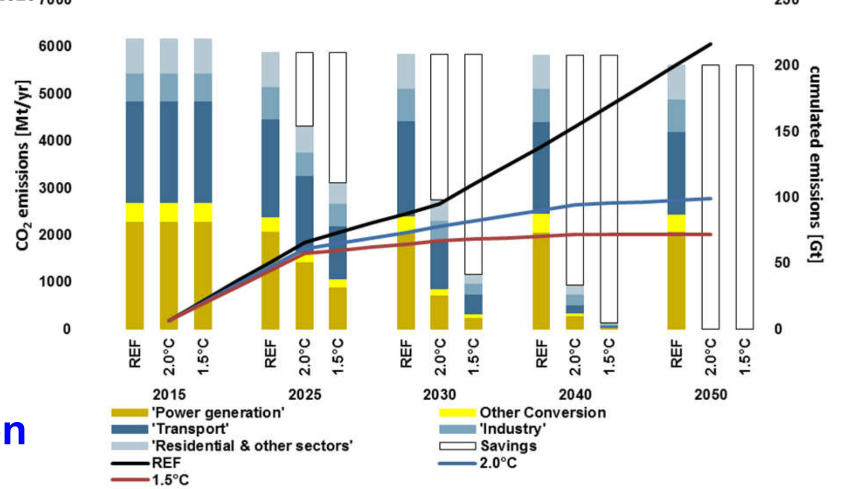
DLR

LOD-GEOSS

Outline

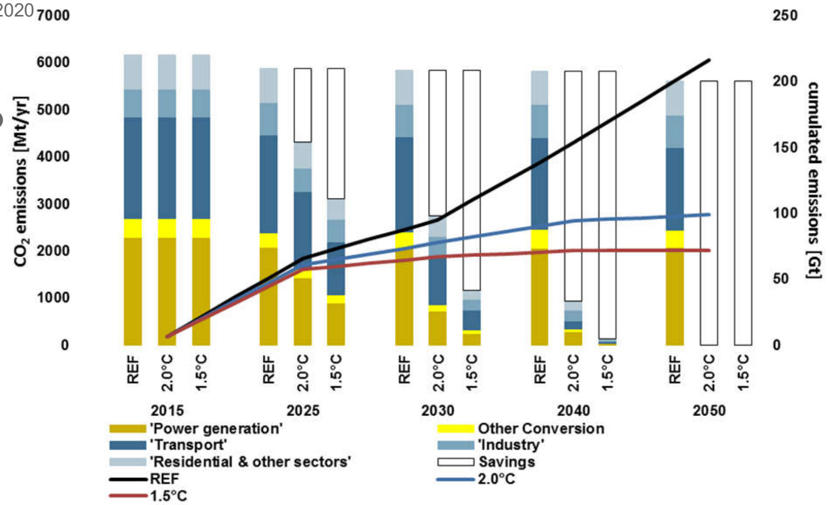
- The domain of energy systems analysis
  - Scenario development
  - Energy systems modelling
- The struggle with data in energy systems analysis
- The Global Earth Observation System of Systems (GEOSS) as an example for a distributed data infrastructure
- Using the data bus for handling of data and metadata in energy systems analysis

What is an Energy Scenario?



- **An image of the temporal development of an energy system** on local, regional, national, continental or global scale
  - taking into account all relevant **energy demand and generation**
  - taking into account all relevant **energy carriers**
  - based on **socio-economic drivers**, **energy intensities** and assumptions on **efficiency gains**
  - based on **technology development pathways** for generation and demand
  - Simulation of **if-then relations** with different temporal horizons (mostly 2020, 2030, 2050, 2100)
  - **Scenarios are not a prognosis!** They are a possible consistent pathway into the future.

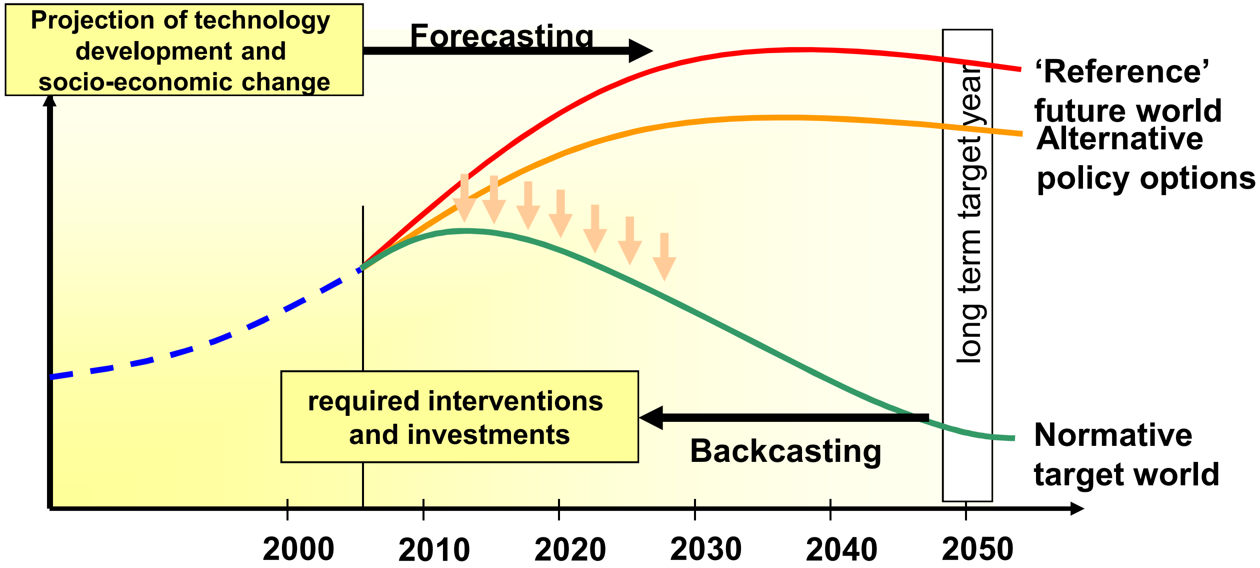
Which Role to Energy Scenarios Play in Politics?



- Compare different **options for future developments**
- Describe the **impact of policy decisions** (support schemes, targets, infrastructure decisions) on the energy system
- Describe **innovation processes** and **investment decisions**
- Basis for **policy support** and consulting and for **political arguments**
- Basis to **design the structural change** to a **sustainable energy system**
- Questions: What do we expect? What do we want to achieve? How can we achieve this? What impacts will be associated with it?

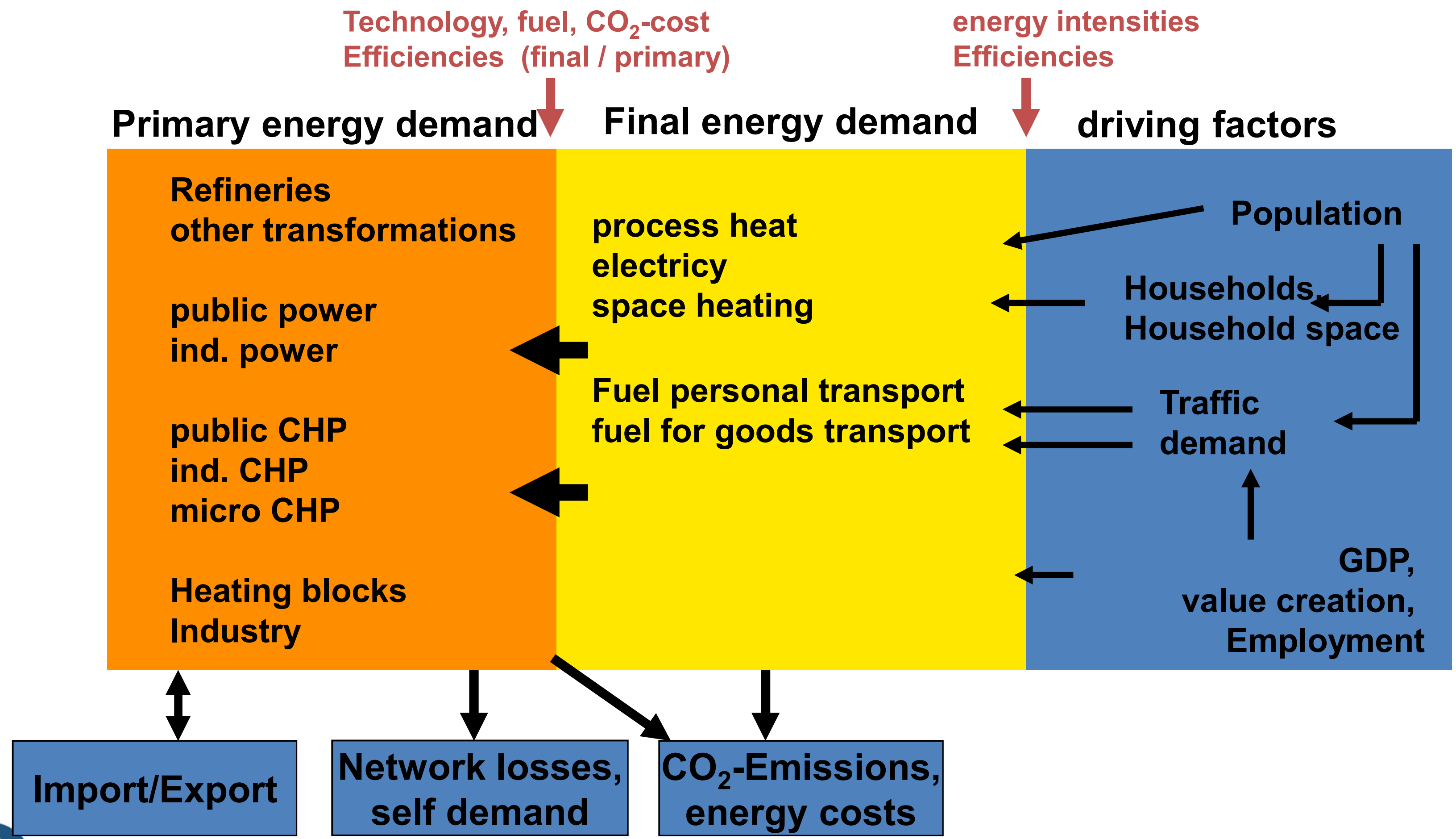
Methodology – Scenario Logic

- long term scenarios starting from **normative targets** (backcasting)
- needed developments based on **RE potentials, plausible market developments, sustainable and robust strategies**
- model calibration with **energy balances. Population** & **GDP** development etc. as drivers of demand
- **review process, validation (dynamic modelling), analysis of economic effects**



Structure of the Energy System in Models

- Energy system models are **data driven** to a very large extend
- Data input is **very heterogeneous** and from a **large variety of different data sources**
- **Data transparency** is getting increased attention in the public debate with (should) be based on results

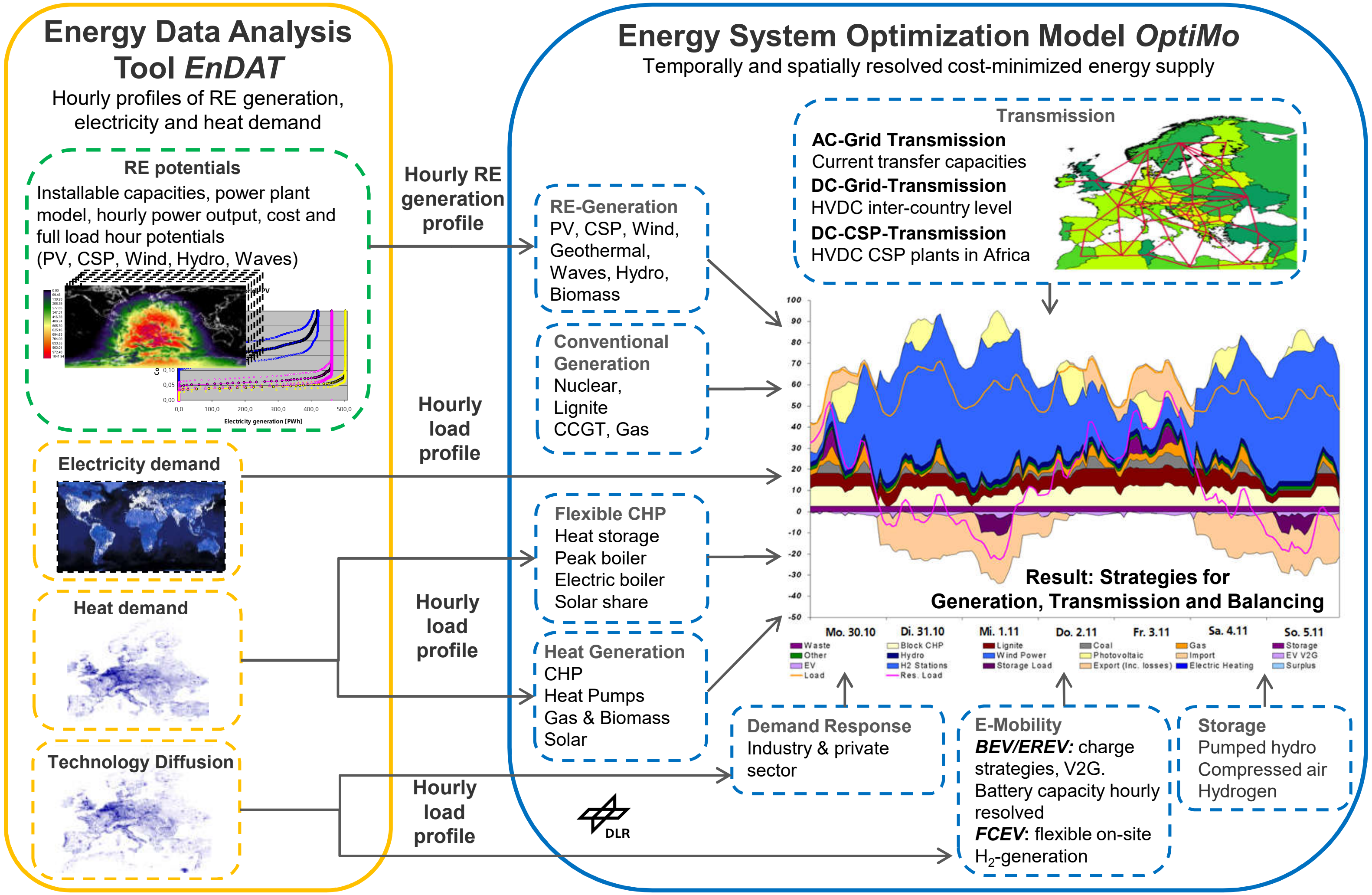


Project Number 03EI1005A

LOD-GEOSS

High resolution modelling

- Integration of **fluctuating renewable** energy sources needs **high resolution modelling**
- Typical modelling approaches need a lot of **time series data**, e.g.
  - Meteorology
  - Power demand
  - Power exchange data



Project Number 03EI1005A

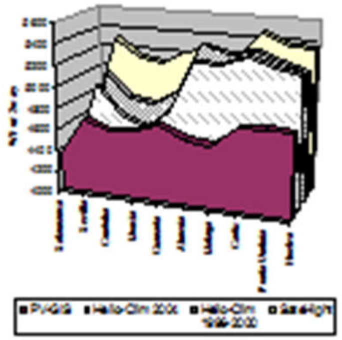
LOD-GEOSS

Always the Struggle with the Data

- Data availability and quality is determining the results of energy systems analysis
- High effort in data research, data bases of very different quality
- Different semantics, different definitions.
- Many different data formats
- Unknown or unsuitable licenses
- Update of data is difficult, time lag between data research and publication

Déjà-Vu : Motivation for EU Project MESOR (Management and Exploitation of Solar Resource Knowledge – mesor.org)

Guidance and Access to Data



- Many sources for solar resource knowledge are available
- Every source has its own access mechanism and data format
- Quality of the sources is often not well known
- Results are difficult to compare

There is quite a number of data sources, but this creates uncertainty of the results, especially if they do not agree

Similar Problems in Earth Observation around the turn of the Century

- Many providers of Earth Observation data
- Every provider with his own access mechanisms
- Data is hard to find and to access
- Solutions:
  - Founding of GEO (Group on Earth Observation) as a G8 initiative in 2005.
  - GEOSS (Global Earth Observation Systems of Systems) as an architecture to network Earth Observation around the Globe.
  - Data and Architecture Committee to define interoperability and communication standards
  - Use of data standards of the Open Geospatial Consortium (OGC) for geographical data and webservices.

Some Principles of GEOSS

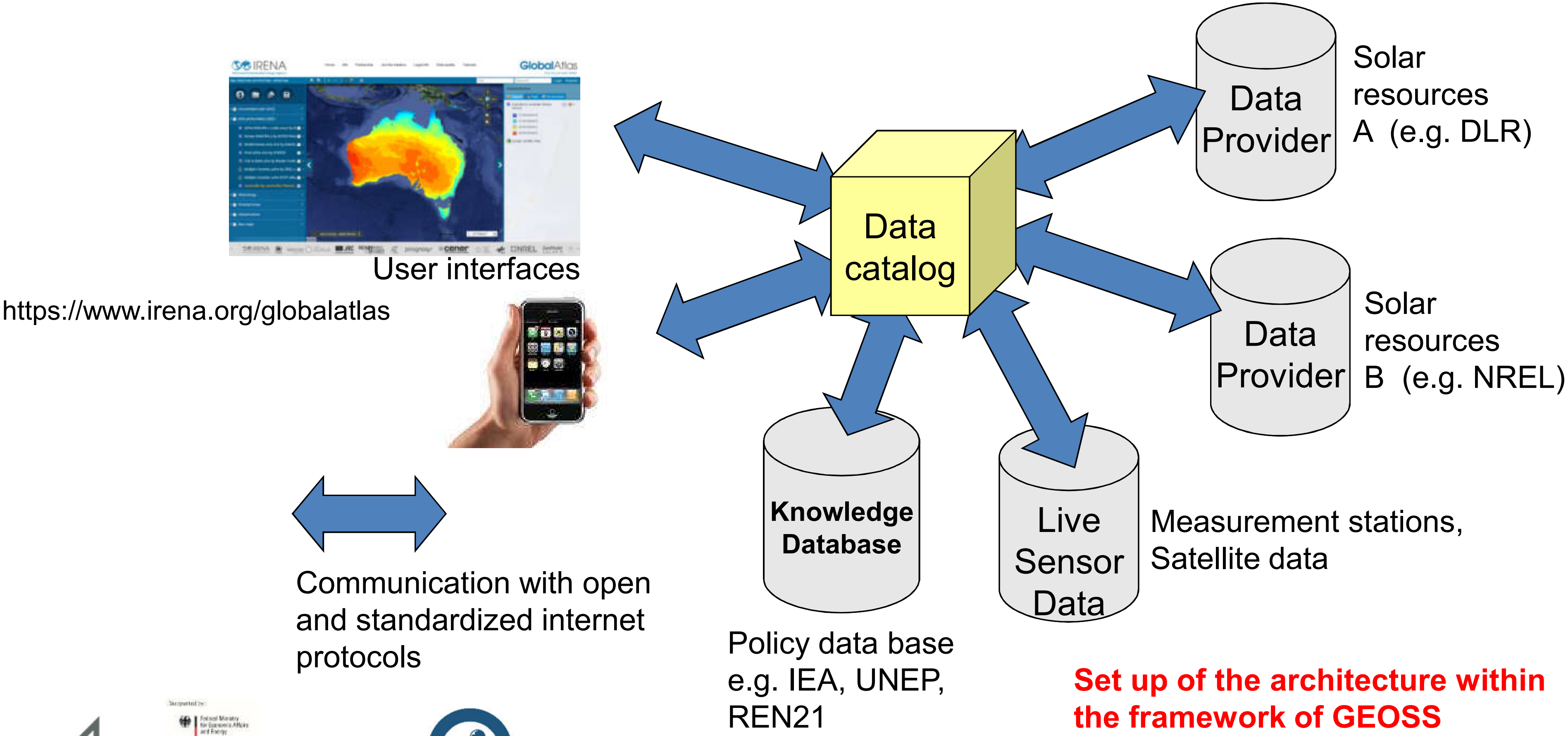
- Decentralized provision of data
  - Every provider stays responsible of maintenance of his data
  - Every provider keeps the rights to his data

- Central hierarchical catalog(s) for the discovery of data, metadata is collected in the catalogs.

- Standardized description of interfaces, e.g.
  (WSDL – Web-service Description Language)

- OGC (Open geospatial consortium) interfaces as standard for geo data.

GEOSS Example: The Global Atlas of Renewable Energies of IRENA
Open Architecture – Collaborative Information Systems



Set up of the architecture within the framework of GEOSS

Learning from Linked Open Data and Open Data Initiatives

- Distributed and networked Data bases :
  - Linked Open Data
  - dbpedia/Wikipedia
- Integrated Assessment Models
  - SSPs (Share Socioeconomic Pathsways)
  - IPCC Scenario data bases
- Open Data Initiatives in Energy Systems Analysis
  - Open Energy Platform (OEP)
  - Szenario DB
  - Open Power Systems Data Base (OPSD)
  - OpenEI (US initiative of NREL and others)

reegle information gateway for renewable energy and energy efficiency

OEP

ipcc INTERGOVERNMENTAL PANEL ON climate change WMO UNEP

OpenEI Open Energy Information

What do we need

- Unified description of data:
  - Glossary
  - Definitions

  } Common Ontologies
- Proper licensing of data
- A flexible expandable networked data infrastructure
- Interfaces to GEOSS (meteorological and geographical data)
- Interfaces to models

REINER LEMOINE INSTITUT

DLR Deutsches Zentrum für Luft- und Raumfahrt

InfAI Institute for Applied Informatics

JÜLICH Forschungszentrum

Universität Stuttgart IER Institut für Energiewirtschaft und rationelle Energieanwendung

PIK Potsdam Institute for Climate Impact Research

DLR Project Number 03EI1005A LOD-GEOSS

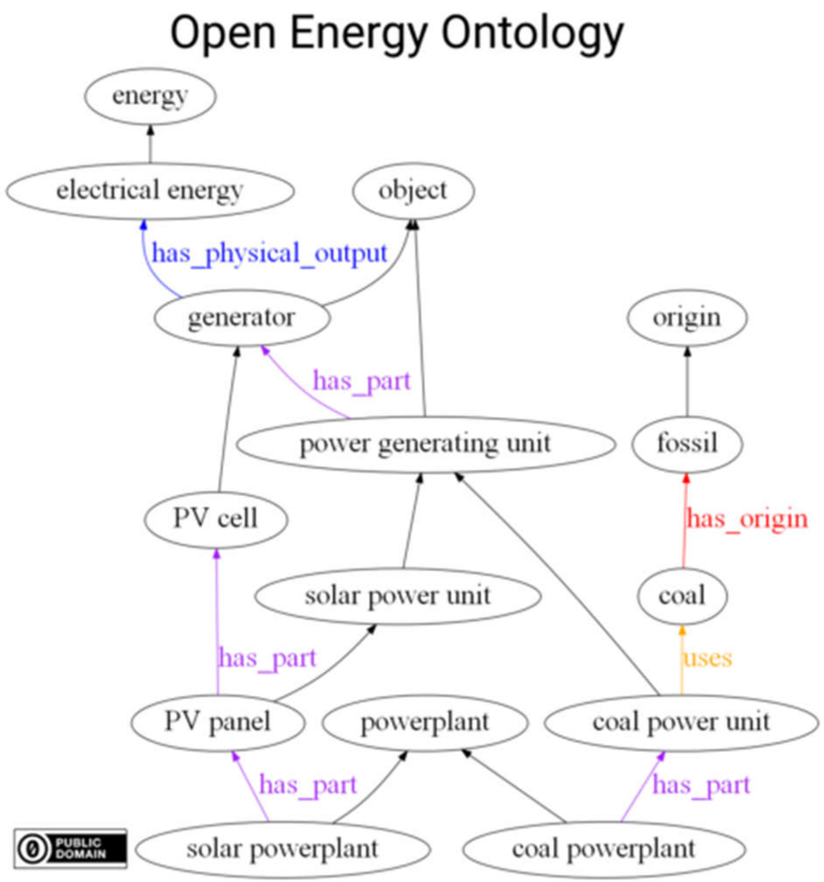
The Open Energy Ontology (OEO)

What is it good for?

- Standardizing terminology in the energy domain
- Data annotation and data integration
- Templates for data capture
- Visualization
- Text and data mining
- Semantic similarity analysis

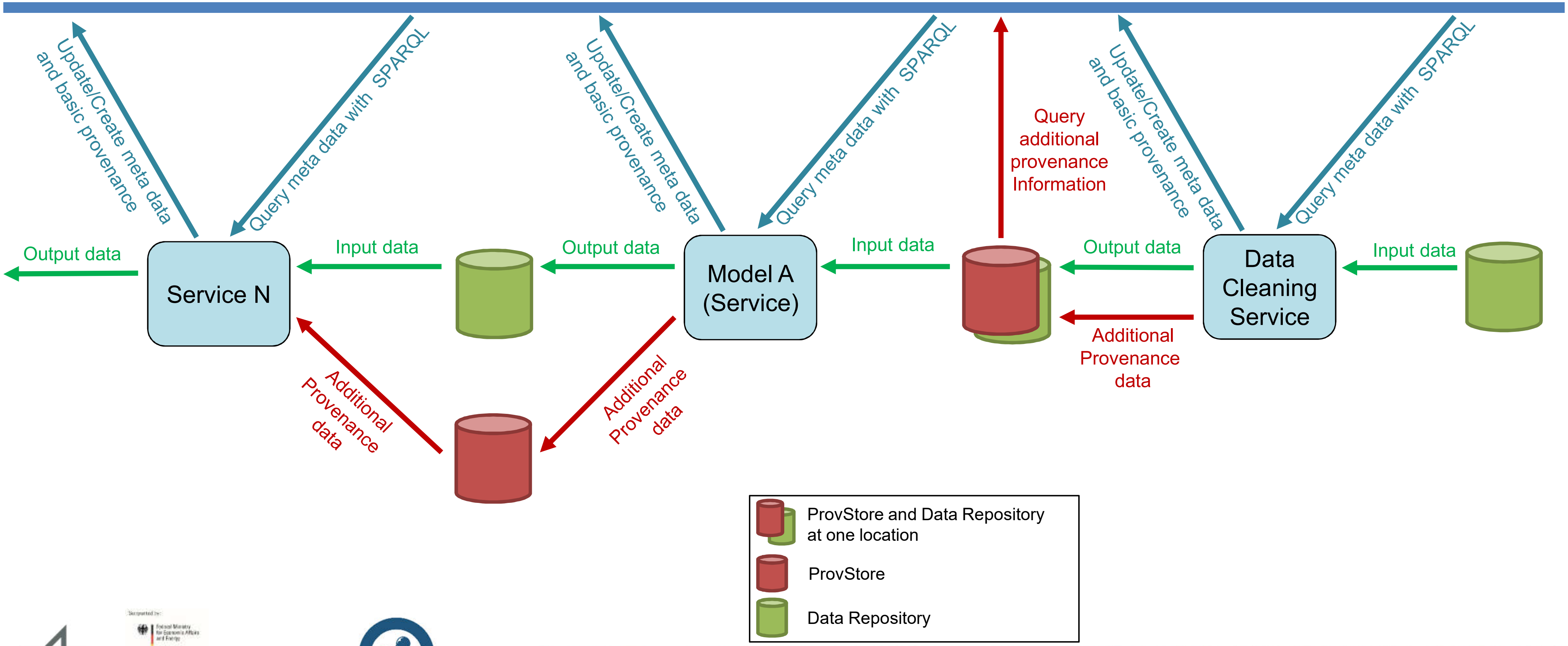
Main Characteristics

- Manchester OWL syntax
- Uses the BFO as top level ontology
- Openly licensed (CC0)
- Version 1.1.0 released, monthly release cycles



https://github.com/OpenEnergyPlatform/ontology

Databus aggregates meta data (including basic provenance) via external maven repositories.
This includes locations of data and provenance.



Project Number 03EI1005A

Databus concept

1. The Databus collects and handles all meta data (catalogue)
2. A service can query information about available data from the Databus. It contains:
   - Meta data about the data set, including basic provenance
   - Location of the data set
   - Location of additional provenance

3. The service retrieves the data from a maven repository
4. The service generates a new data set based on the input (optionally recording detailed provenance)
5. The service generates a new meta data description

6. The data set is hosted in a maven repository
7. The meta data to the data is hosted in a maven repository
8. The Databus is notified about the location of a new dataset and its meta data

REPEAT

Contribution to the FAIR Principles

- **F**inable
  - Development and setup of the data bus as metadata catalog for data in energy systems analysis
- **A**ccessible
  - Descriptions of the data formats and interfaces to data bases, best practice buides, links to the data sources in the metadata
- **I**nteroperable
  - Enhancement of on Open Energy Ontology as a common data language, open data format descriptions
- **R**eusable
  - Improvement of data licenses, here with a special focus on GEOSS data, provenance information with the data

Project Number 03EI1005A

Contact us

German Aerospace Center
Energy Systems Analysis

Carsten Hoyer-Klick

Email: carsten.hoyer-klick@dlr.de

License