Karlsruhe Institute of Technology

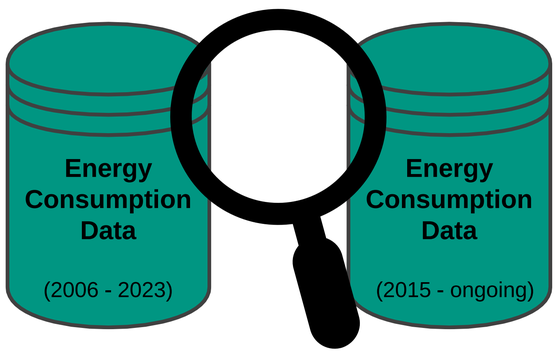
Institute for Automation and Applied Informatics (IAI)

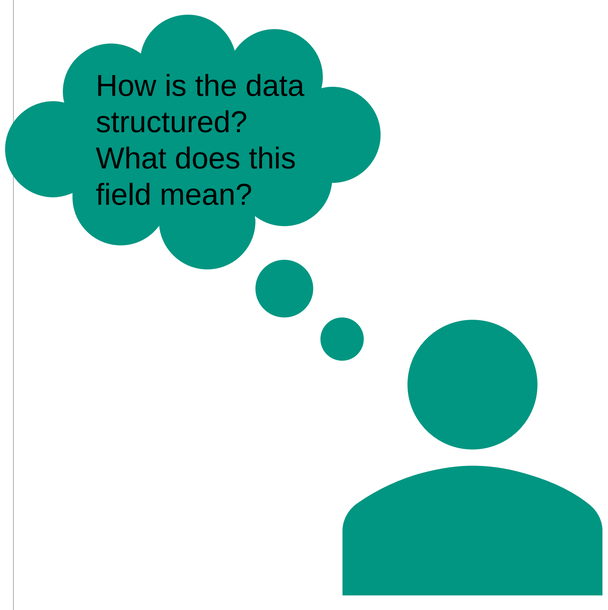
RO-Crate Time Series Exporter

for the Building Consumption Data of KIT Campus North

Jan Schweikert, Marian Turowski, Viktoria Köbe, Wolfgang Süß, Veit Hagenmeyer

1 Database Analysis





Steps taken
- Exploring available data
- Understanding the organization structure of the data
- Understand who and why the data was collected
- Check data for plausibility

Outcome
- Software artifacts
  - Data validation scripts
  - Data querying scripts

Involved Persons
- Project Team
- Data Collectors (FM)
- Stakeholders

2 Unified Domain Schema

Unified Schema

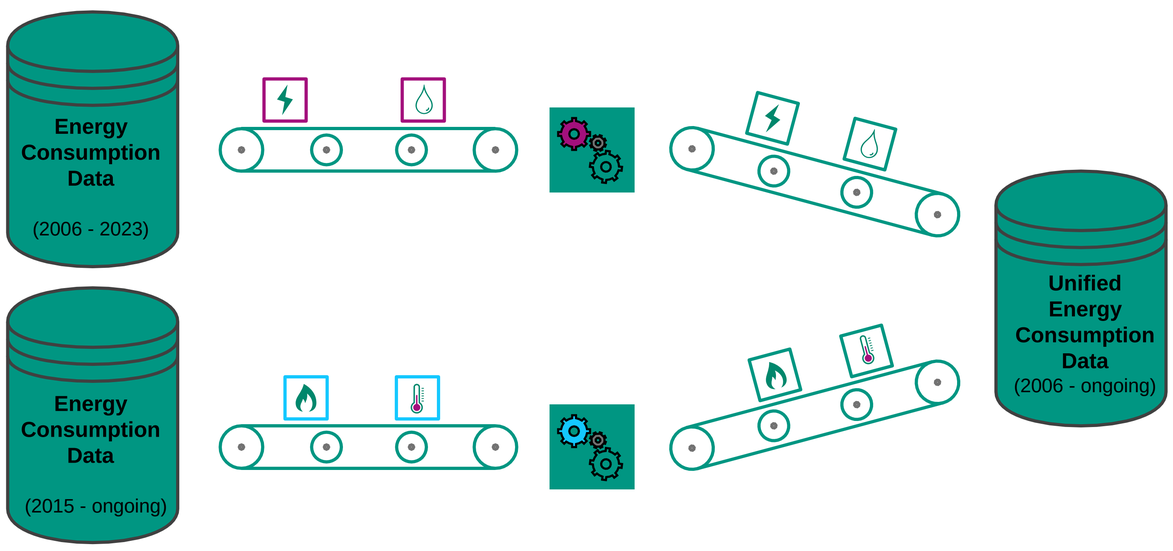
time	RFC3339
value	float
factor	float
building	string
consecutiveNumber	int
measuringType	enum
name	string
energyType	enum
quantity	QUDT Quantity
unit	QUDT Unit
granularity	ISO8601 ABNF
deactivated	boolean
locationDescription	string

Steps taken
- Identifying requirements on the data
- Agreement on shared vocabulary
- Def. a unified DB schema
- Def. the RO-Crate profile

Outcome
- Unified database schema

Involved Persons
- Project Team
- Stakeholders

3 Data Transformation

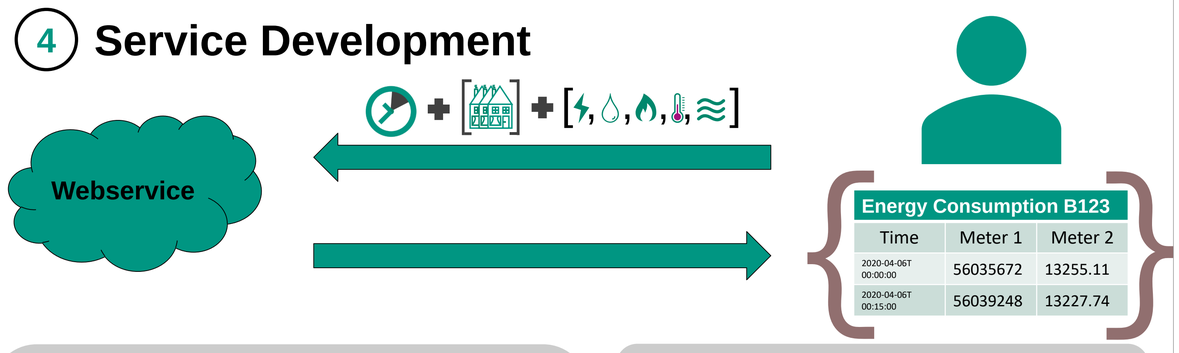


Steps taken
- Normalize field values
- Map field values to vocab terms
- Develop transformer software
- Deploy software as persistent service on our infrastructure

Outcome
- Unified database

Involved Persons
- Project Team

4 Service Development



Steps taken
- Defined REST API
- Determine programming lang
- Identify required libraries
- Program

Outcome
- Webservice

Involved Persons
- Project Team

5 Outlook

- Include weather data of requested time period (in dev.)
- Generalize the service to export data from arbitrary databases as RO-Crates
  - Exploit the database from Energy Lab 2.0
  - Automatically crawl database schemas and identify semantically unknown fields
  - Create an UI to enable researches defining the semantically unknown fields

Contact
wolfgang.suess@kit.edu

Acknowledgements
This publication was supported within the Hub Energy of the Helmholtz Metadata Collaboration <HMC>, an incubator-platform of the Helmholtz Association within the framework of the Information and Data Science strategic initiative. It was also supported by the Helmholtz Association's Initiative and Networking Fund through Helmholtz AI.