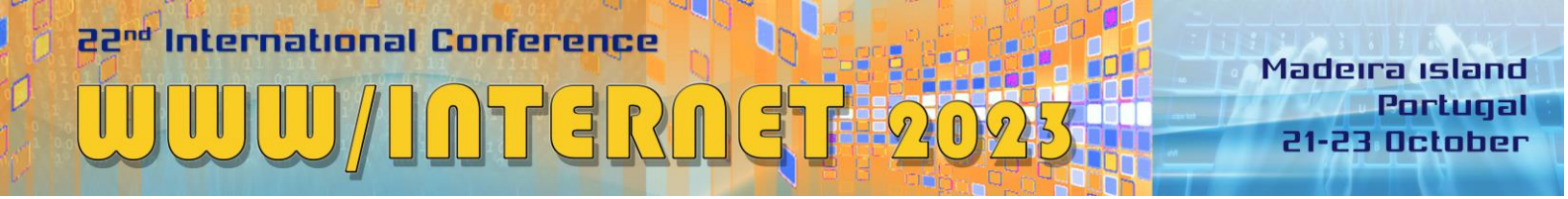


BAM
Bundesanstalt für Materialforschung und -prüfung

Sicherheit in Technik und Chemie

NATURALMSEQUERIES

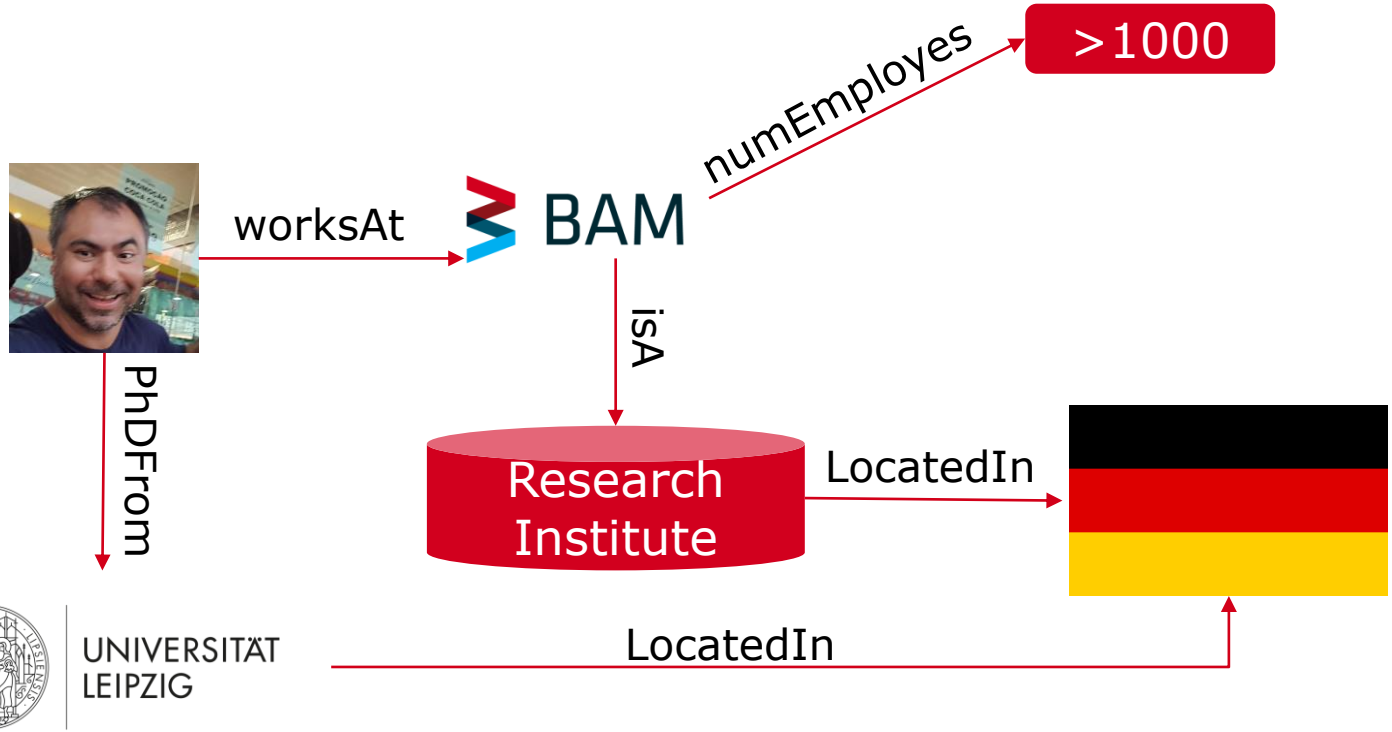
A NATURAL WAY TO QUERY MATERIALS SCIENCE ENGINEERING DATA EXPERIMENTS

Andre Valdestilhas, Thomas Hanke, Soudeh Javamasoudian, Ghezal Ahmad Jan Zia, Horst Fellenberg, Thilo Muth

Outiline

1. Preliminaries (My self, RDF, FAIR)
2. Motivation
3. Research questions (RQ)
4. Methodology (Our approach)
5. Evaluation and results – Answer RQ
6. Conclusion

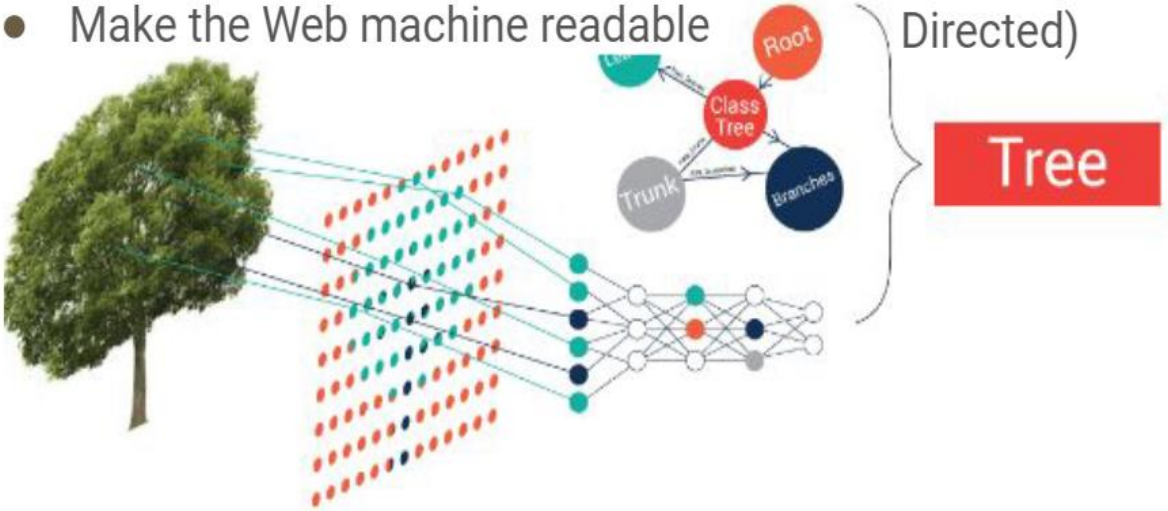
MySelf Directed Labeled Graph



Preliminaries: RDF

Semantic Web -> Resource Description Framework

- Triples (Subject – Predicate – Object)
- Make the Web machine readable
- Metadata data model
- Multigraph (Labelled, Directed)



structured Data (RDF)

\+ **shared Identifiers** (Links,URI)

\+ **query engine** (SPARQL)

= **Knowledge graph**

machine-processable navigable space

Source: https://www.ontotext.com/

Data → Knowledge – FAIR- and Open-Data

F: FINDABLE

- What data exists?
- How & where do I find the measured values?

A: ACCESSIBLE

- Is raw data & metadata accessible?
- -> Quality / value of data
- Restrictions? (Software, formats…)

I: INTEROPERABLE

- Usability beyond the originator:

-> Input & query (internal and external)

R: REUSABLE

- Value creation: Creation of new knowledge with fewer attempts or re-evaluation

How MSE domain expert obtain the Knowledge Graph

Input: Raw data from Experiment and Simulation

sample name	remark	weight	
BA Los M V-4	Kraftgeregelt 3,9 kN/s	5428,5	
BA Los M V-5	Kraftgeregelt 3,9 kN/s	5454,6	
BA Los M V-6	Kraftgeregelt 3,9 kN/s	5439	
BA Los M VI-4	Kraftgeregelt 3,9 kN/s	5354,5	
BA Los M VI-5	Kraftgeregelt 3,9 kN/s	5340,6	
BA Los M VI-6	Kraftgeregelt 3,9 kN/s	5367,6	
BA-Losert E-Modul 28d v. 04.08.14 Probe 4	Kraftgeregelt 3,9 kN/s	5342	
BA-Losert E-Modul 28d v. 04.08.14 Probe 5	Kraftgeregelt 3,9 kN/s	5404	
BA-Losert E-Modul 28d v. 04.08.14 Probe 6	Kraftgeregelt 3,9 kN/s	5314	
BA-Losert MII E-Modul 28d v. 04.08.14 Probe 4	Kraftgeregelt 3,9 kN/s	5374	
BA-Losert MII E-Modul 28d v. 04.08.14 Probe 5	Kraftgeregelt 3,9 kN/s	5365	
BA-Losert MII E-Modul 28d v. 04.08.14 Probe 6	Kraftgeregelt 3,9 kN/s	5359	
Hsken E-Modul Probe 2-1	Kraftgeregelt 3,9 kN/s	5853	
Hsken E-Modul Probe 3-1	Kraftgeregelt 3,9 kN/s	5830	
Hsken E-Modul Probe 3-3	Kraftgeregelt 3,9 kN/s	5848	

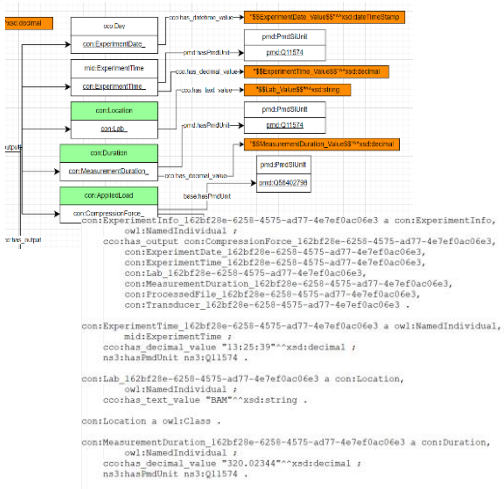
```
Masse : 5428,5
Durchmesser :   98,5
Länge : 300,2
Bediener Information Ende

Datenerfassung                                          Zeit:
Datum   Laufzeit       Ch 1 Kraft Sollwert     Ch 1 Kraft
d       s       kN      kN      mm      mm      mm      mm
41893   33,549805       -1      -1,0524206      -0,063988566
41893   34,049805       -1      -1,0790397      -0,062161684
41893   34,549805       -1      -1,0645893      -0,062426925
41893   35,049805       -1      -0,97048283     -0,061368942
41893   35,549805       -1      -0,96076912     -0,06223321
```

Metadaten and Processed data

CSV

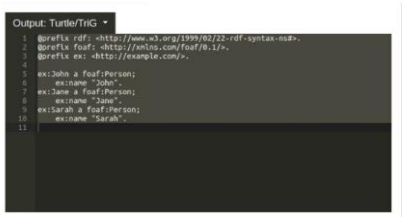
json

Mapping-Method

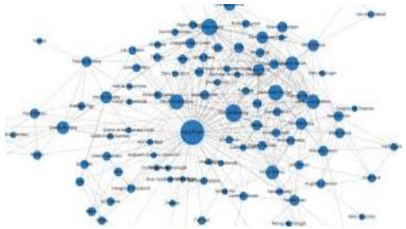


```
con:ExperimentInfo_162bf28e-6258-4575-ad77-4e7ef0ac06e3 a con:ExperimentInfo,
        owl:NamedIndividual ;
    cco:has_output con:CompressionForce_162bf28e-6258-4575-ad77-4e7ef0ac06e3,
        con:ExperimentDate_162bf28e-6258-4575-ad77-4e7ef0ac06e3,
        con:ExperimentTime_162bf28e-6258-4575-ad77-4e7ef0ac06e3,
        con:Lab_162bf28e-6258-4575-ad77-4e7ef0ac06e3,
        con:MeasurementDuration_162bf28e-6258-4575-ad77-4e7ef0ac06e3,
        con:ProcessedFile_162bf28e-6258-4575-ad77-4e7ef0ac06e3,
        con:Transducer_162bf28e-6258-4575-ad77-4e7ef0ac06e3 .

con:ExperimentTime_162bf28e-6258-4575-ad77-4e7ef0ac06e3 a owl:NamedIndividual,
        mid:ExperimentTime ;
    cco:has_decimal_value "13:25:39"^^xsd:decimal ;
    ns3:hasPmdUnit ns3:Q11574 .

con:Lab_162bf28e-6258-4575-ad77-4e7ef0ac06e3 a con:Location,
        owl:NamedIndividual ;
    cco:has_text_value "BAM"^^xsd:string .

con:Location a owl:Class .

con:MeasurementDuration_162bf28e-6258-4575-ad77-4e7ef0ac06e3 a con:Duration,
        owl:NamedIndividual ;
    cco:has_decimal_value "320.02344"^^xsd:decimal ;
    ns3:hasPmdUnit ns3:Q11574 .
```

Output: Knowledge Graph

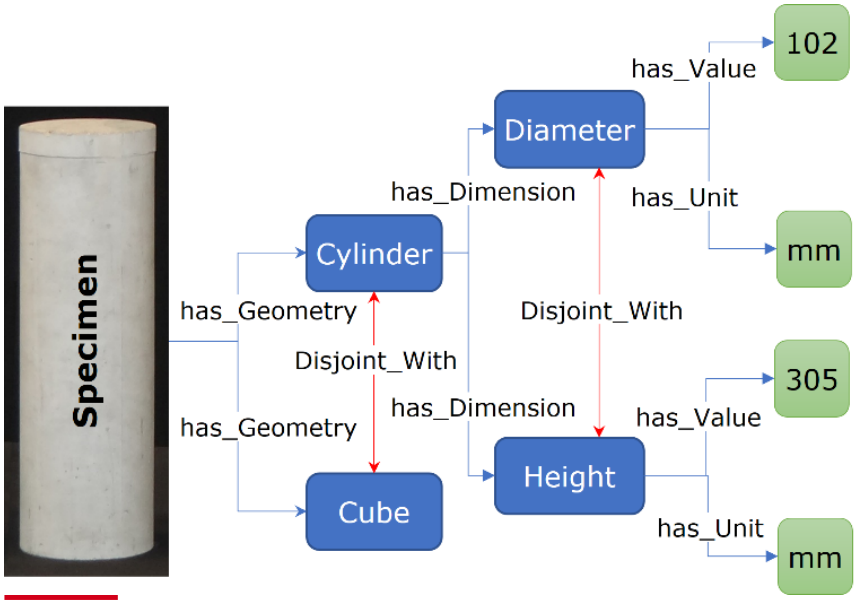


```
Output: Turtle/TriG
@prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>.
@prefix foaf: <http://xmlns.com/foaf/0.1/>.
@prefix ex: <http://example.com/>.

ex:John a foaf:Person;
    ex:name "John".
ex:Jane a foaf:Person;
    ex:name "Jane".
ex:Sarah a foaf:Person;
    ex:name "Sarah".
```

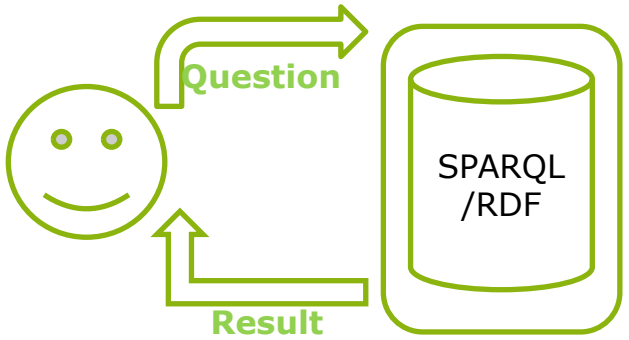


Motivation

- Material science data can be represented as **directed graphs**



- **RDF** as data model
- **Lack** of specific knowledge of **SPARQL** queries



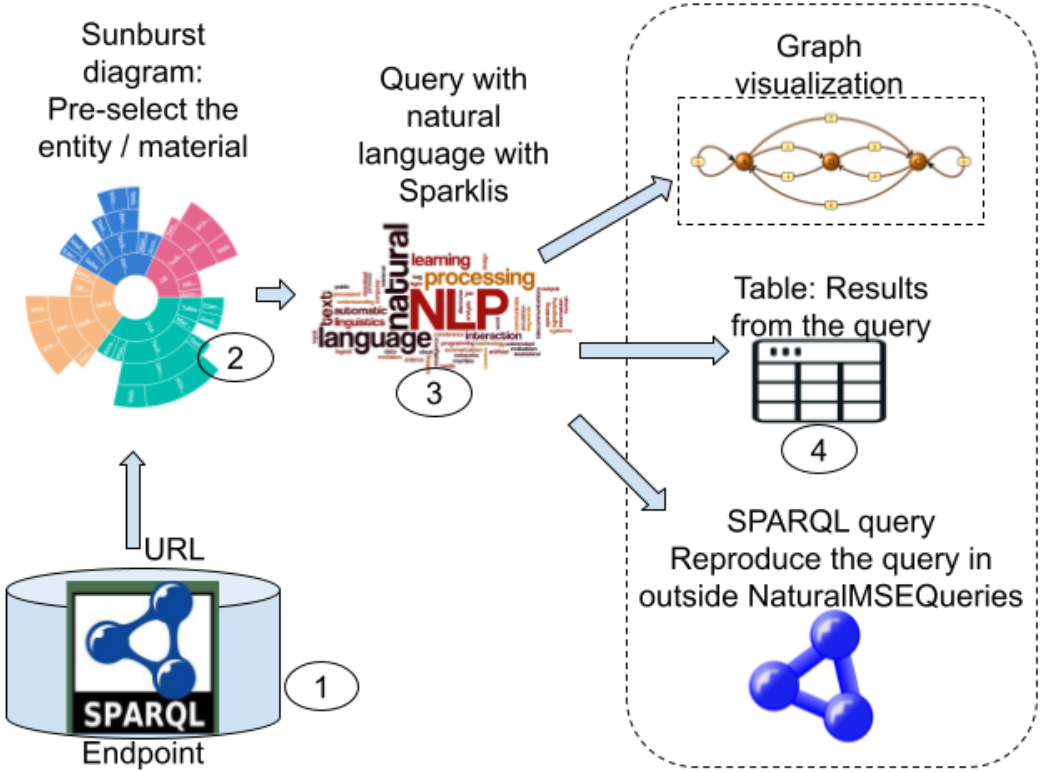
Research questions

RQ1. How to query semantic MSE data **easier** than using **SPARQL** queries?

RQ2. What is the best way to **organize Material Sciences Methods data**?

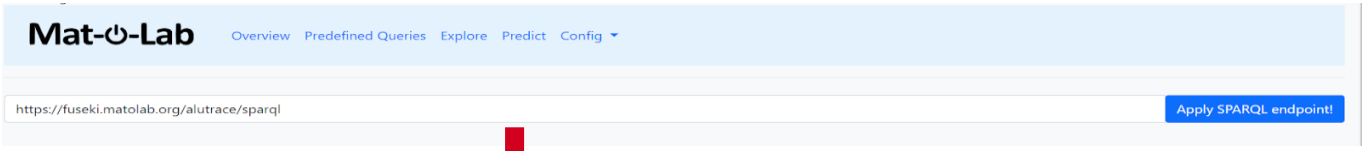
RQ3. **How much** will the framework **help** the Materials Science Engineering domain?

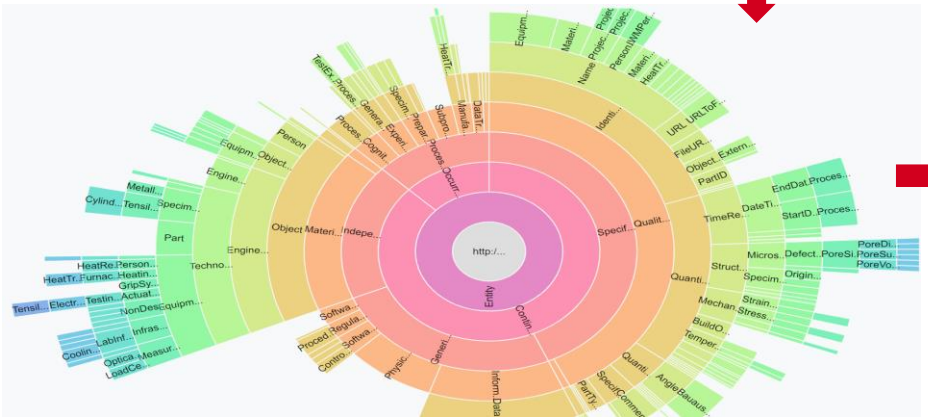
Methodology (our approach)

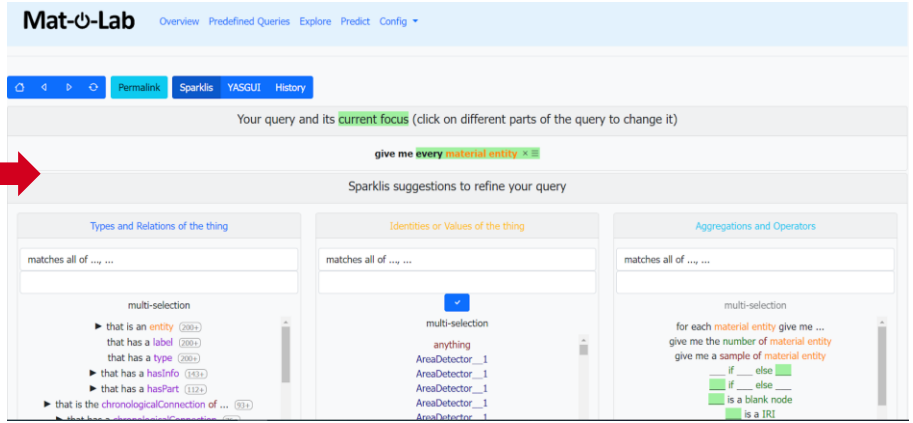


Sparklis citation: *Ferré, Sébastien. 'Sparklis: An Expressive Query Builder for SPARQL Endpoints with Guidance in Natural Language'. Semantic Web 8(3) : 405-418. IOS Press, 2017*

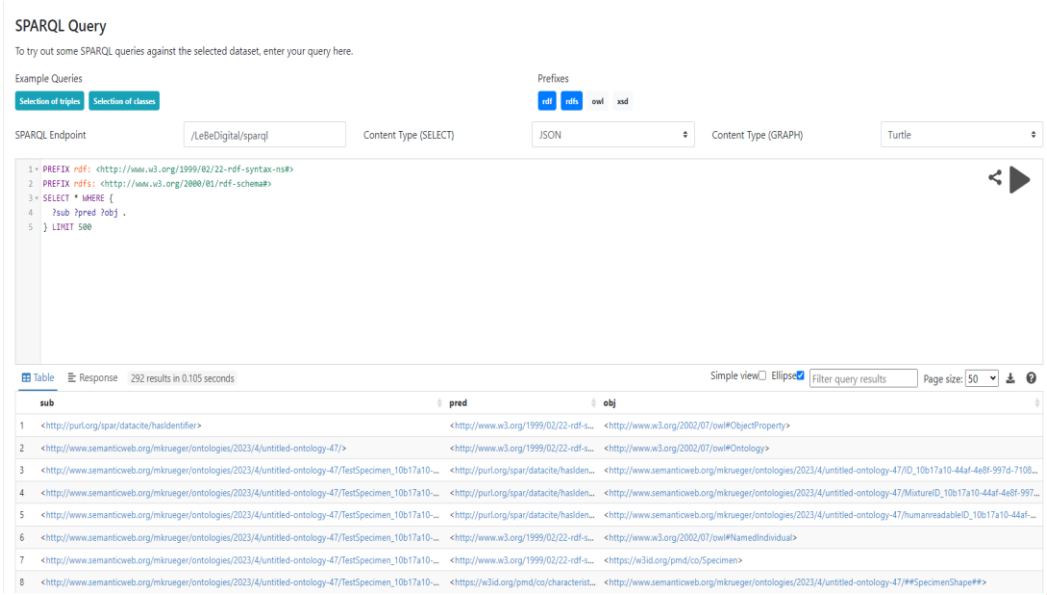
Methodology (our approach)

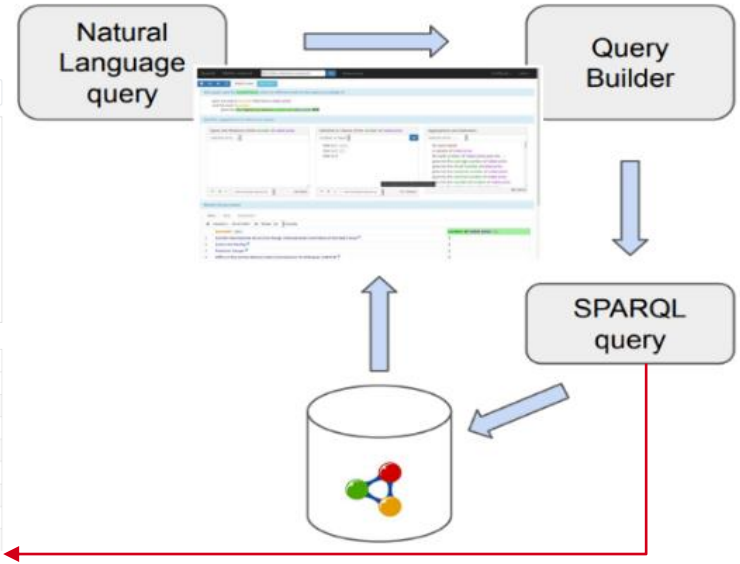






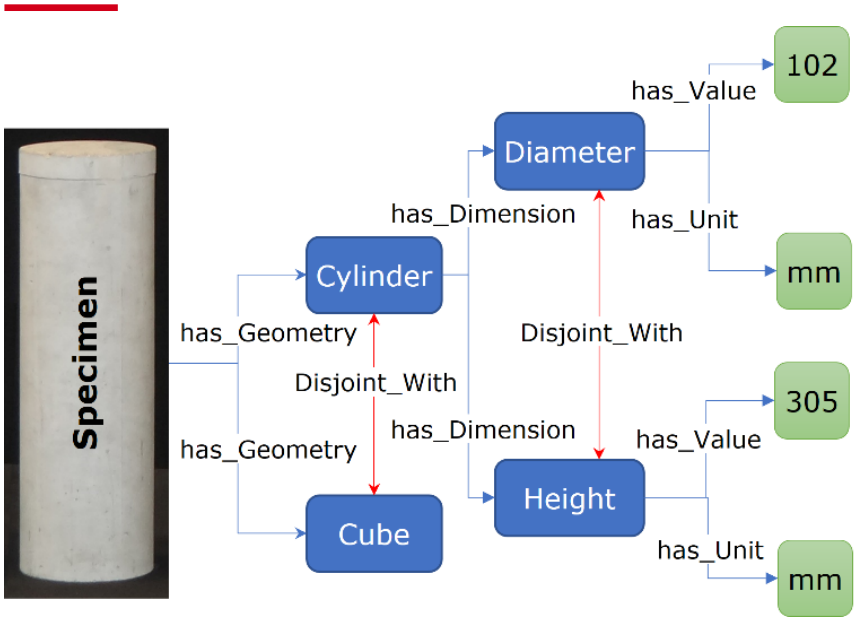
Methodology (our approach)
Querying the data

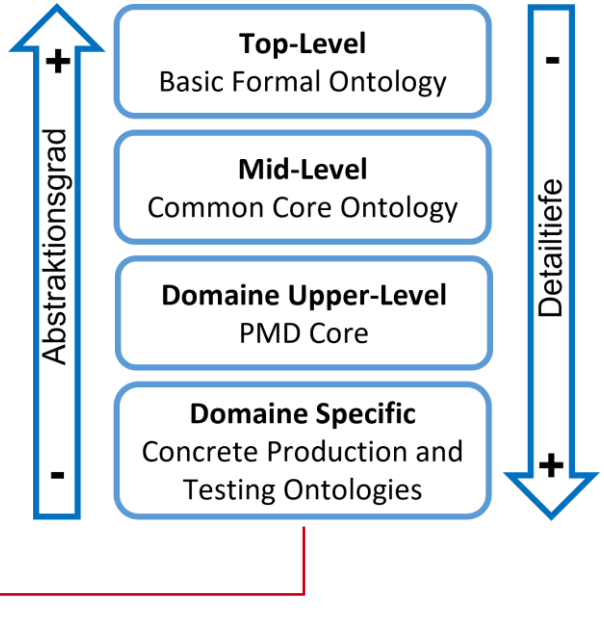




Natural Language to SPARQL query and graph visualization of the query results

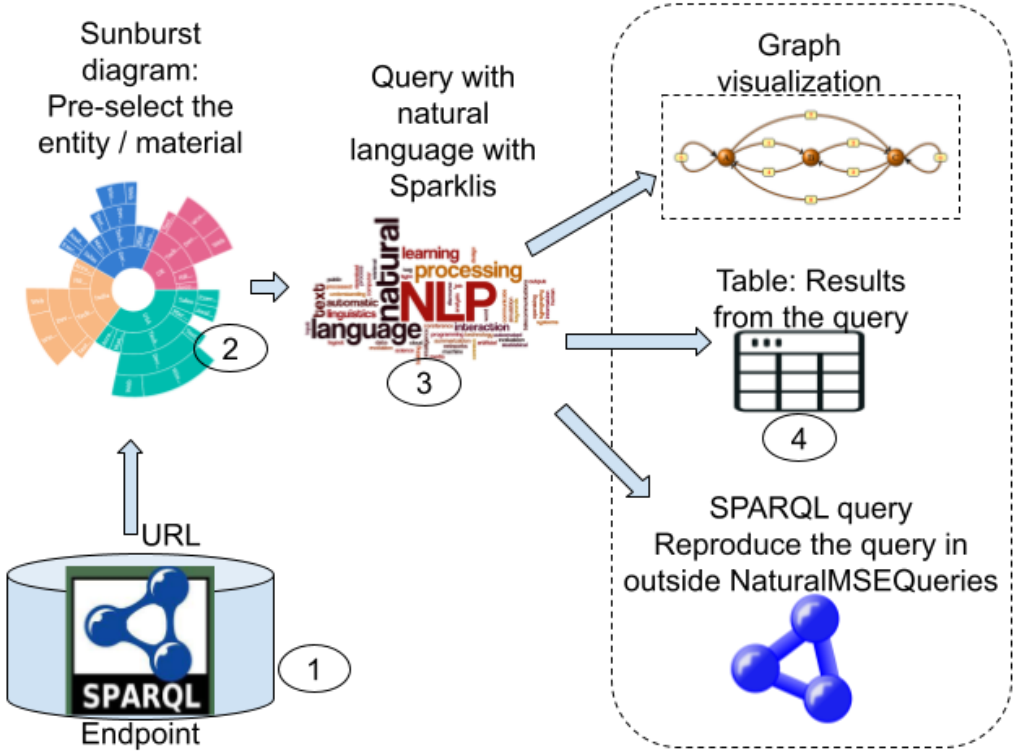
Evaluation: Domain experts creating knowledge – Ontology, querying data





- Link data across institutions (BAM – Fraunhofer – KIT, etc)
- Exploit heterogenous materials data

Approach – Research Questions



RQ1. How to query semantic MSE data **easier** than using **SPARQL** queries? (**NLP**)

RQ2. What is the best way to **organize Material Sciences Methods data**? (**RDF Knowledge Graph**)

RQ3. **How much** will the framework **help** the Materials Science Engineering domain? (**Evaluation/Usability**)

Conclusion

BAM

1 Potential for Lightweight Design

Explore how NaturalMSEQueries has successfully applied to the several projects, projects, showcasing its potential for lightweight lightweight design in materials science.

2 User-Friendly Approach

- Pioneer on the intersection between between SWT and MSE
- Understand how our approach enables enables domain experts to query materials science data more effectively, effectively, improving the overall usability and accessibility of Semantic Semantic Web technologies.

3 Future Development

NaturalMSEQueries + LLM (work in progress)

https://github.com/Mat-O-Lab/KnowledgeUI

andre.valdestilhas@bam.de