ONTOLOGY AND WHY CONTEXT MATTERS

Ahmad Zainul Ihsan[1] and Stefan Sandfeld[1,2]

1 Forschungszentrum Jülich, Institute of Advanced Simulations – Materials Data Science and Informatics (IAS-9), Germany
2 RWTH Aachen University, Faculty 5, Germany

MSE Day, Hereon Geesthacht, 14.11.2023

Mitglied der Helmholtz-Gemeinschaft

JÜLICH Forschungszentrum

MOTIVATION

LANGUAGE AND COMMUNICATION

- Language can be used to **represent knowledge**.
- What is Language?
  - System of conventional **spoken**, **manual**, or **written symbols** combined to **convey meaning** [1]**.**
  - Language is used by human beings to **express** themselves.
- **Communication** is one of the functions of language

- Successful communication depends on:
  - Information has to be correctly **encoded** via **syntax**.
  - The meaning (**semantics**) of the encoded information must be interpreted correctly (**understanding**).
  - The **understanding or decoding** process is influenced by the **context** of sender and receiver

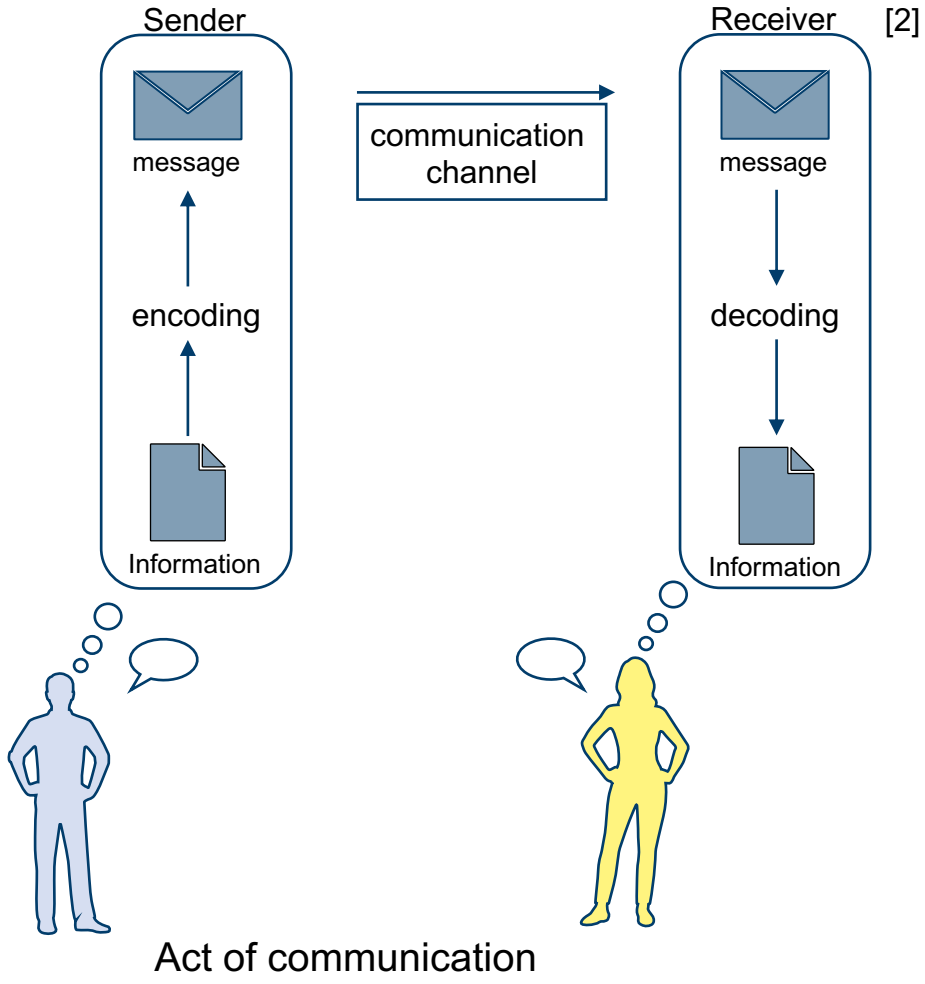
[1] https://www.britannica.com/topic/language

MOTIVATION

LANGUAGE AND COMMUNICATION

- Language can be used to **represent knowledge**.
- What is Language?
  - System of conventional **spoken**, **manual**, or **written symbols** combined to **convey meaning** [1]**.**
  - Language is used by human beings to **express** themselves.
- **Communication** is one of the functions of language

- Successful communication depends on:
  - Information has to be correctly **encoded** via **syntax**.
  - The meaning (**semantics**) of the encoded information must be interpreted correctly (**understanding**).
  - The **understanding or decoding** process is influenced by the **context** of sender and receiver



[1] https://www.britannica.com/topic/language

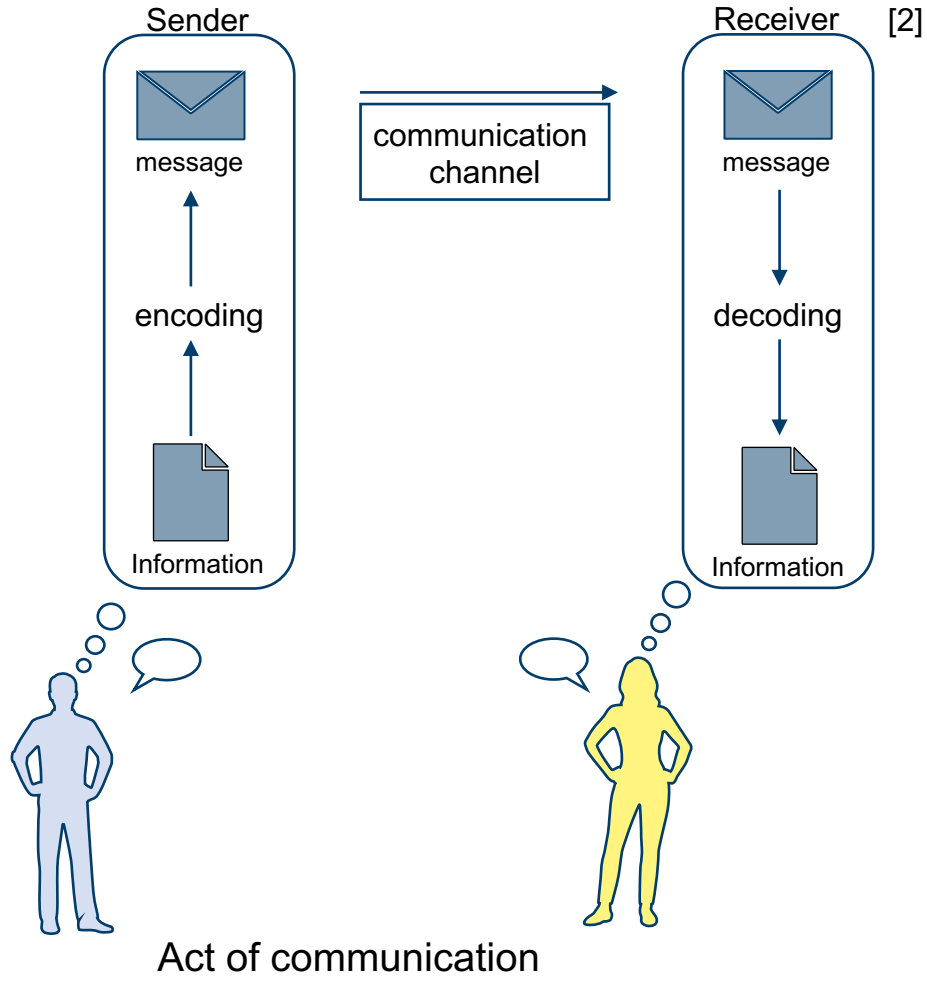
[2] Adapted from Knowledge Graphs Course by Prof. Harald Sack

MOTIVATION

LANGUAGE AND COMMUNICATION

- Language can be used to **represent knowledge**.
- What is Language?
  - System of conventional **spoken**, **manual**, or **written symbols** combined to **convey meaning** [1].
  - Language is used by human beings to **express** themselves.
- **Communication** is one of the functions of language

- Successful communication depends on:
  - Information has to be correctly **encoded** via **syntax**.
  - The meaning of the encoded information must be interpreted correctly (**semantically correct**) -> **understanding**.
  - The **understanding or decoding** process is influenced by the **context.**

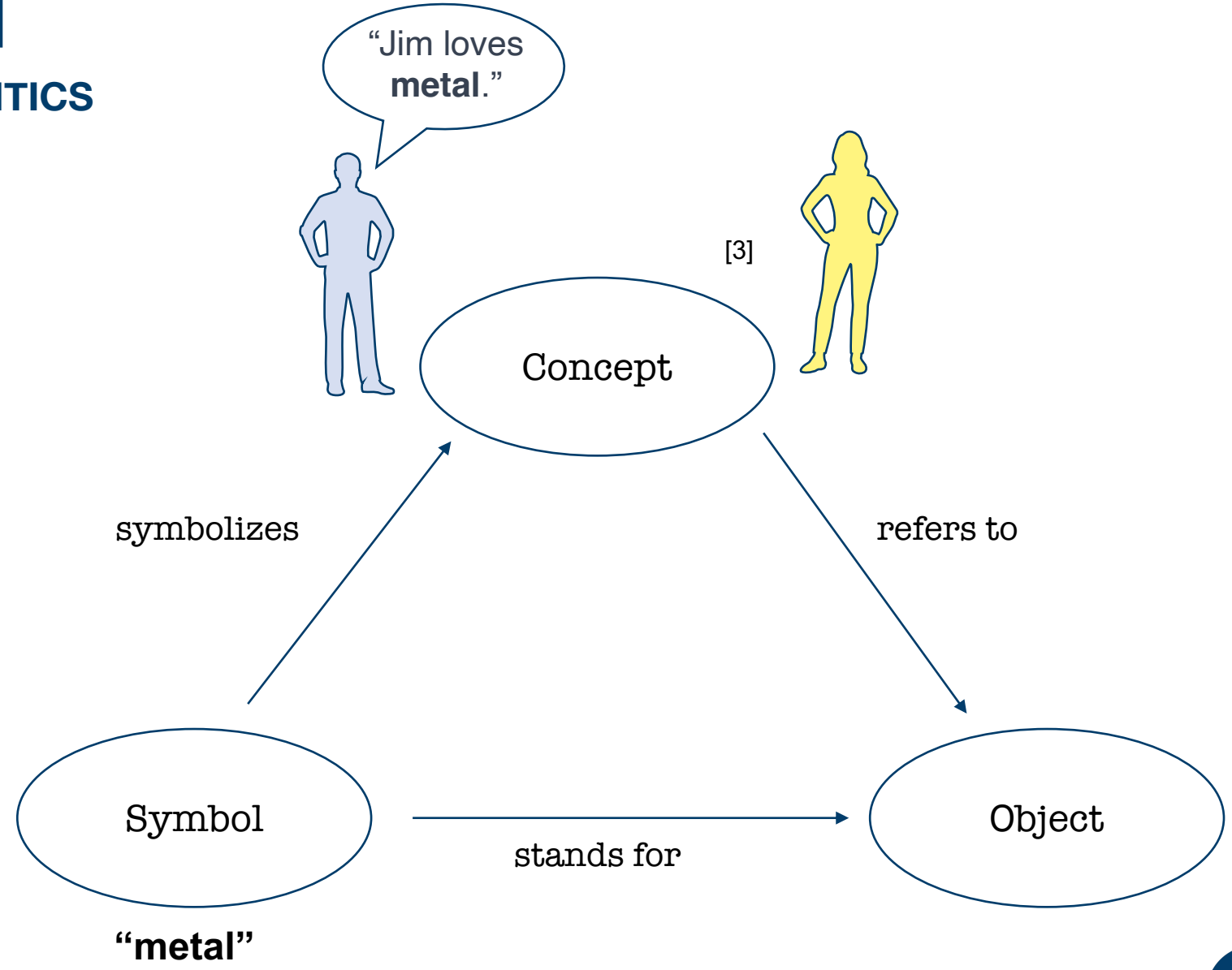


[1] https://www.britannica.com/topic/language

[2] Adapted from Knowledge Graphs Course by Prof. Harald Sack

MOTIVATION

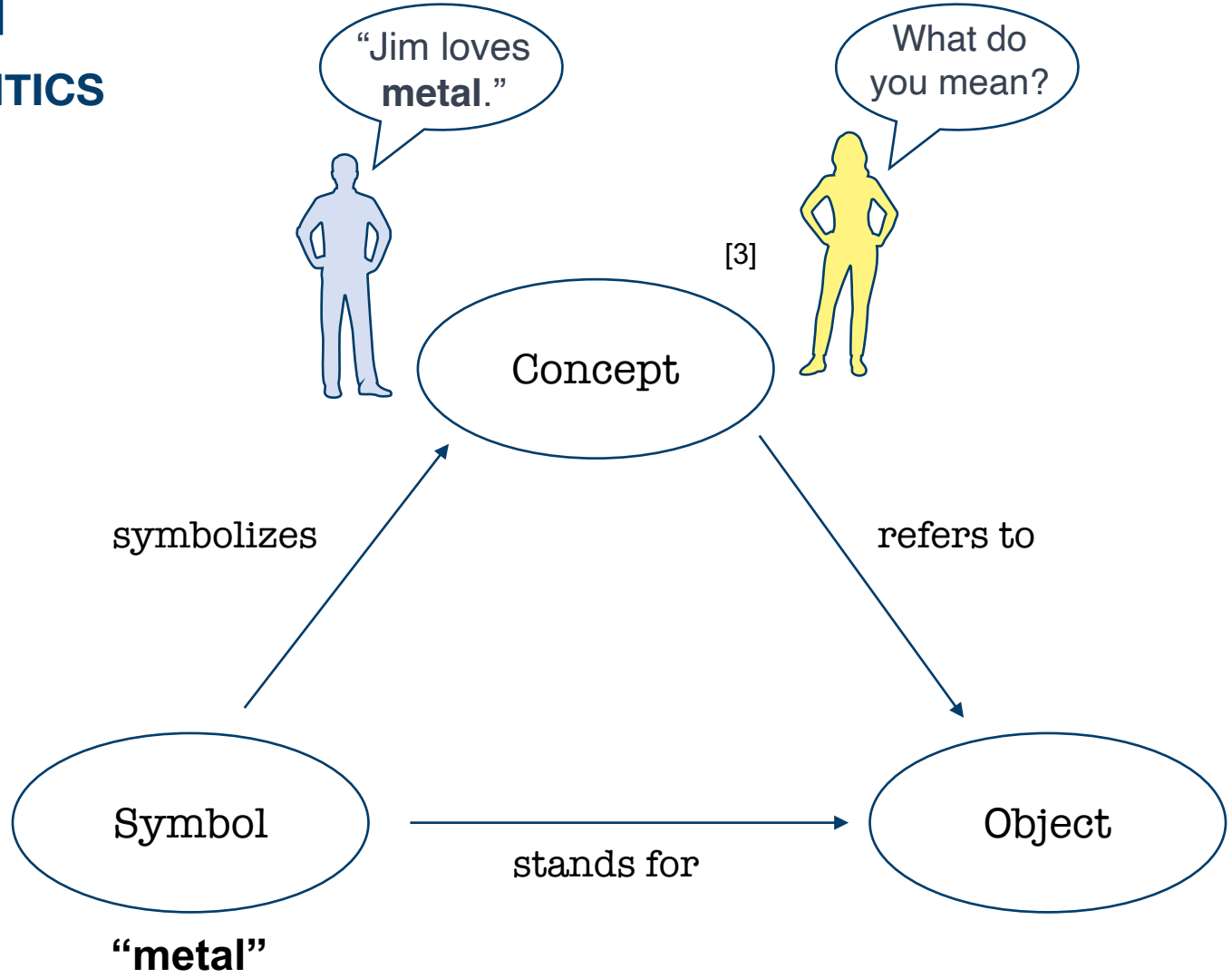
CONTEXT AND SEMANTICS



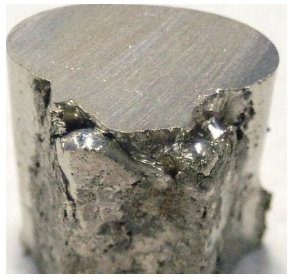
[2] C. K. Ogden and I. A. Richards (1923) The Meaning of Meaning

MOTIVATION

CONTEXT AND SEMANTICS



Music(Metal)



[4]

Material(Metal)

[5]

Software(Metal)

[3] C. K. Ogden and I. A. Richards (1923) The Meaning of Meaning

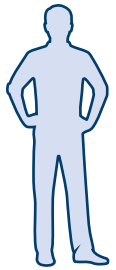
[4] https://en.wikipedia.org/wiki/Nickel

[5] https://developer.apple.com/metal/

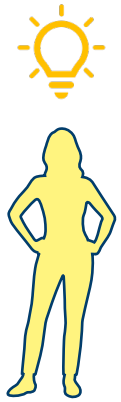
MOTIVATION

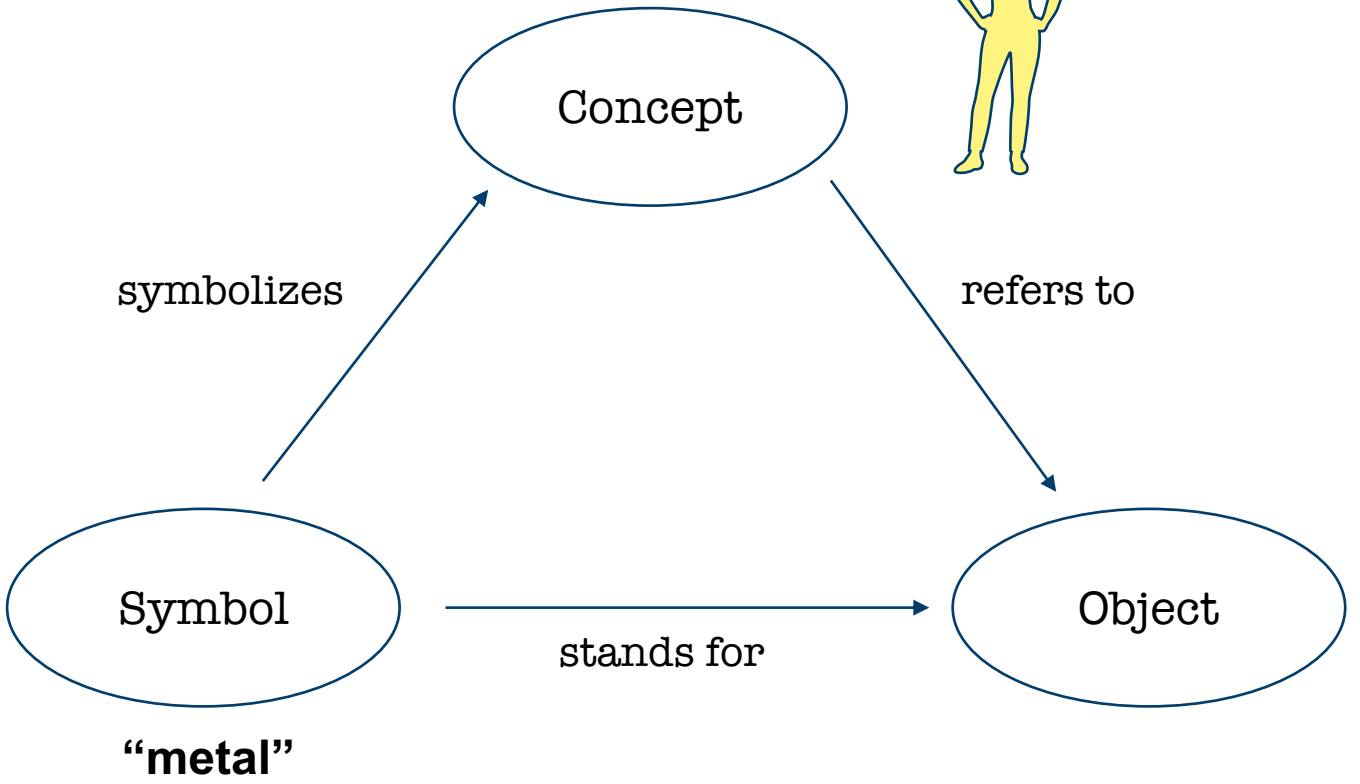
CONTEXT AND SEMANTICS

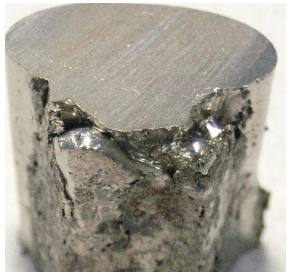
"I mean, he is a metalworker and crafts intricate designs using the medium of **metal.**"



"…can be crafted, something malleable, expertise is needed…"



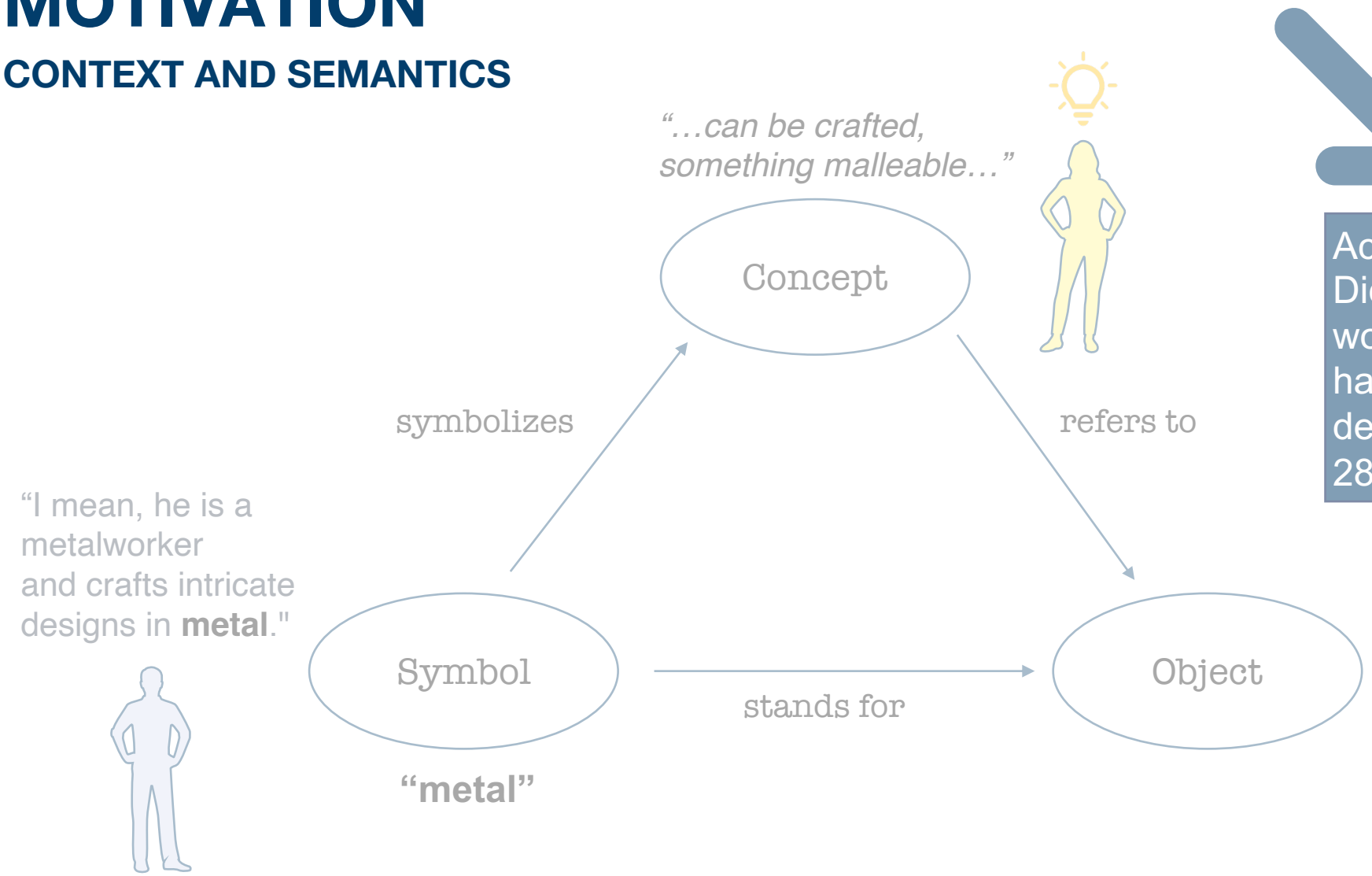




Metal Material

MOTIVATION

CONTEXT AND SEMANTICS



Metal Material

According to the Oxford Dictionary, the 500 most used words in the English language have at least 14.070 different definitions. This is an average of 28 meanings per word [6].

[6] http://inmyownterms.com/mysmartterms/mysmarterms-5-the-semantic-triangle-words-dont-mean-people-mean/

MOTIVATION

FORMAL KNOWLEDGE REPRESENTATION

- Formal knowledge representation:
  - It is a field of AI.
  - It unambiguously captures the **semantics** of **concepts**, **properties**, **relationships**, and **entities.**
  - These semantics relates to specific **knowledge of domains**, i.e., fields of interest or areas of concern.
- **Machines (computers)** must be able to **understand** knowledge representations

MOTIVATION

FORMAL KNOWLEDGE REPRESENTATION

- Formal knowledge representation:
  - It is a field of AI.
  - It unambiguously captures the **semantics** of **concepts**, **properties**, **relationships**, and **entities.**
  - These semantics relates to specific **knowledge of domains**, i.e., fields of interest or areas of concern.
- **Machines (computers)** must be able to **understand** knowledge representations

FAIR

To be interoperable…

I1. (meta)data use a **formal**, accessible, shared, and broadly applicable **language for knowledge representation.**

I2. (meta)data use vocabularies that follow FAIR principles.

I3. (meta)data include qualified references to other (meta)data.

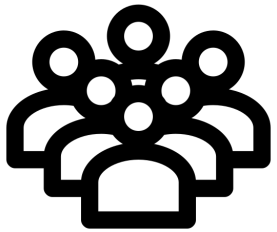
ONTOLOGY

WHAT IS ONTOLOGY?

An ontology is an **explicit, formal specification of a shared conceptualization** [7].

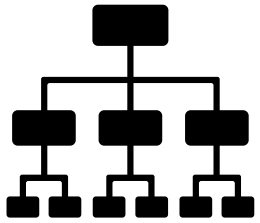
Shared

Reach consensus (shared understanding) about ontology.



Conceptualization

Abstract model about specific domain.

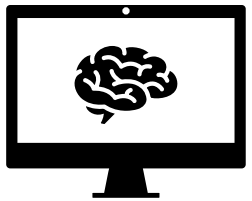


Explicit

Meaning of all concepts must be defined.

Formal

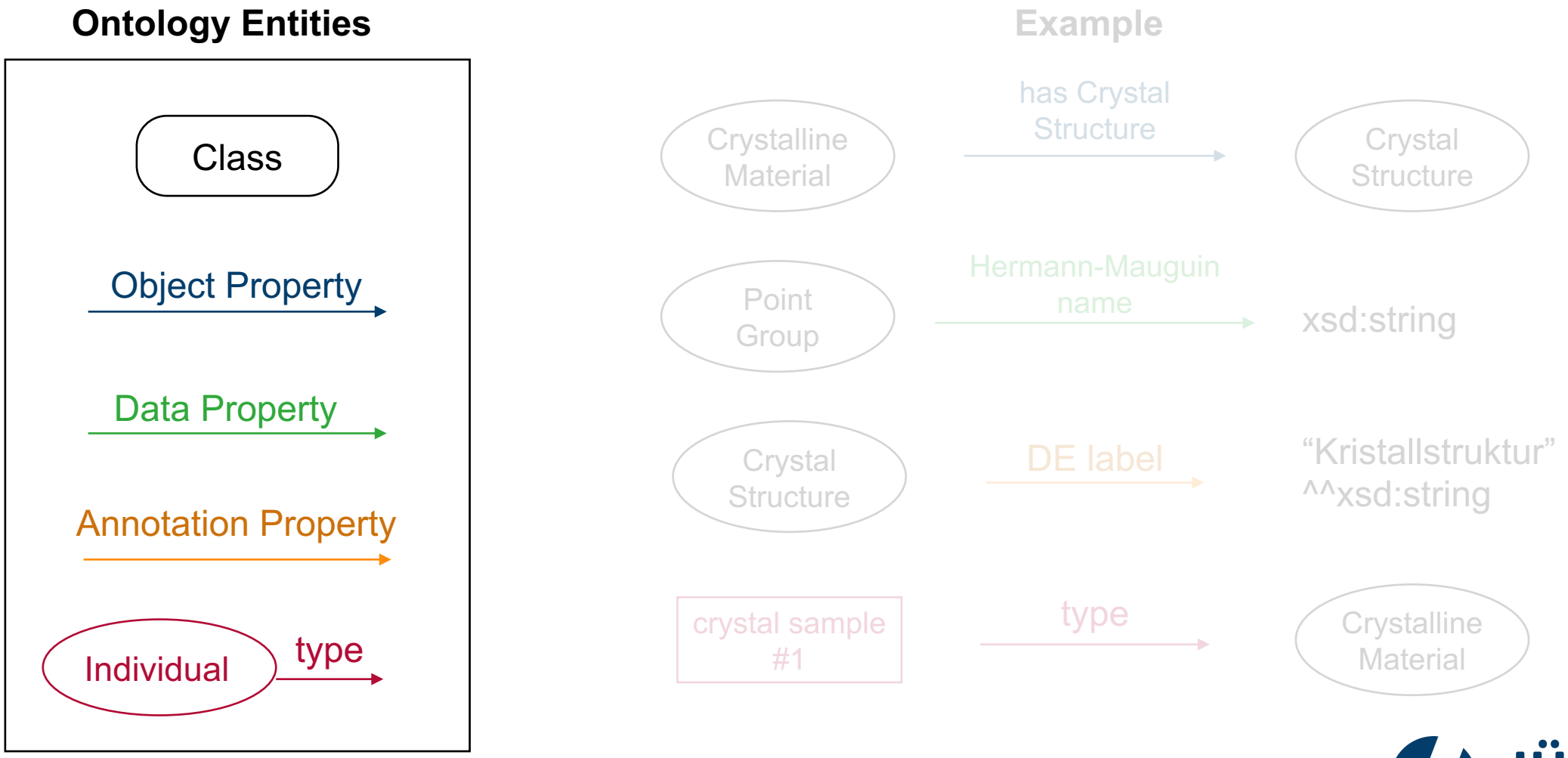
Machine understandable.



[7] Thomas R. Gruber: A Translation Approach to Portable Ontology Specifications. Knowledge Acquisition, 5(2):199-220, 1993.

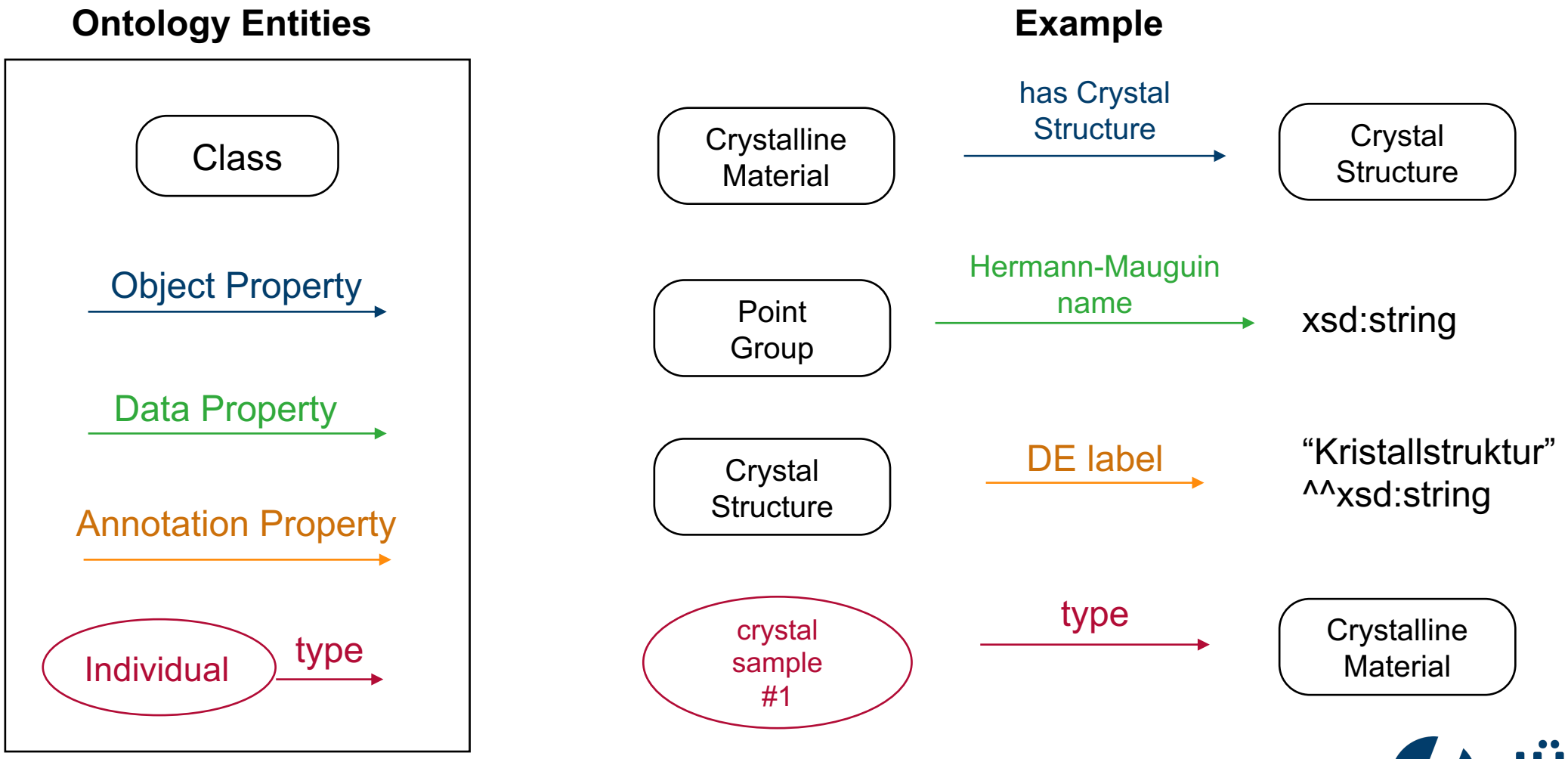
ONTOLOGY

HOW TO REPRESENT ONTOLOGY (1)



ONTOLOGY

HOW TO REPRESENT ONTOLOGY (2)



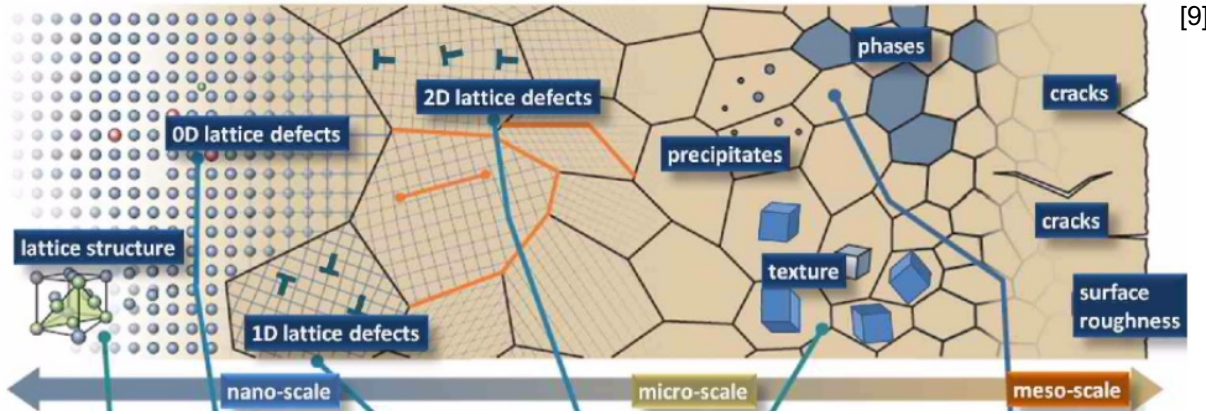
ONTOLOGY APPLICATIONS
CRYSTALLOGRAPHIC DEFECT ONTOLOGY

- **Crystal Defect Ontology** (CDO): harmonized understanding & description of crystal defect across
  - scales (nano/micro/meso),
  - dimensions (0D-3D), and
  - methods (simulations & experiments)

Crystallographic Defect Ontology

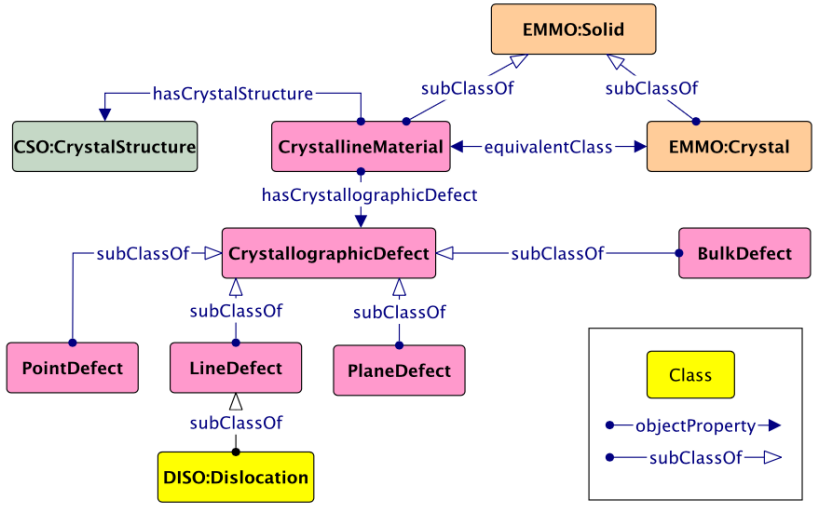
Dislocation Ontology

Planar and Point Defect Ontology



[9]

Defects in crystalline materials

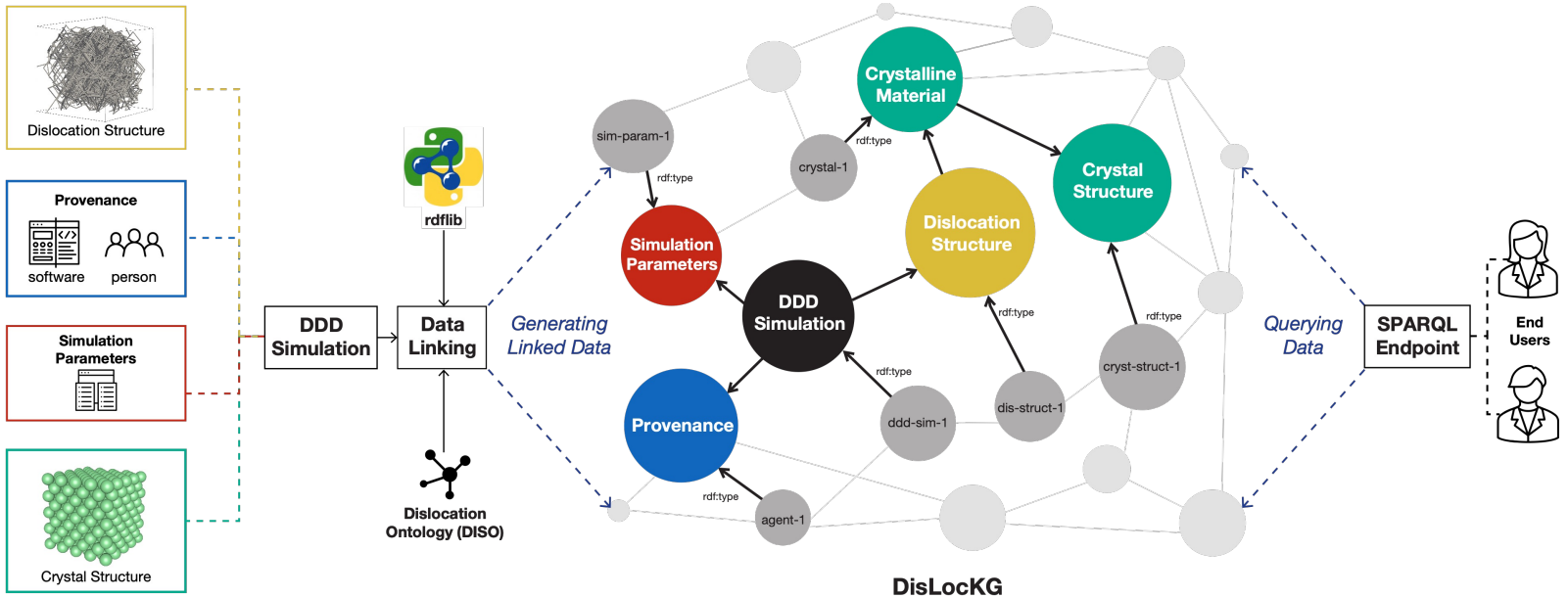


Crystallographic Defect Ontology
AZ Ihsan, AA Guzman, S Fathalla, V Hofmann, S Sandfeld

[8] https://dl.acm.org/doi/abs/10.1145/3555776.3578739
[9] http://www.dierk-raabe.com/multiscale-modeling/

ONTOLOGY APPLICATIONS
MODELING DISLOCATION DYNAMICS DATA

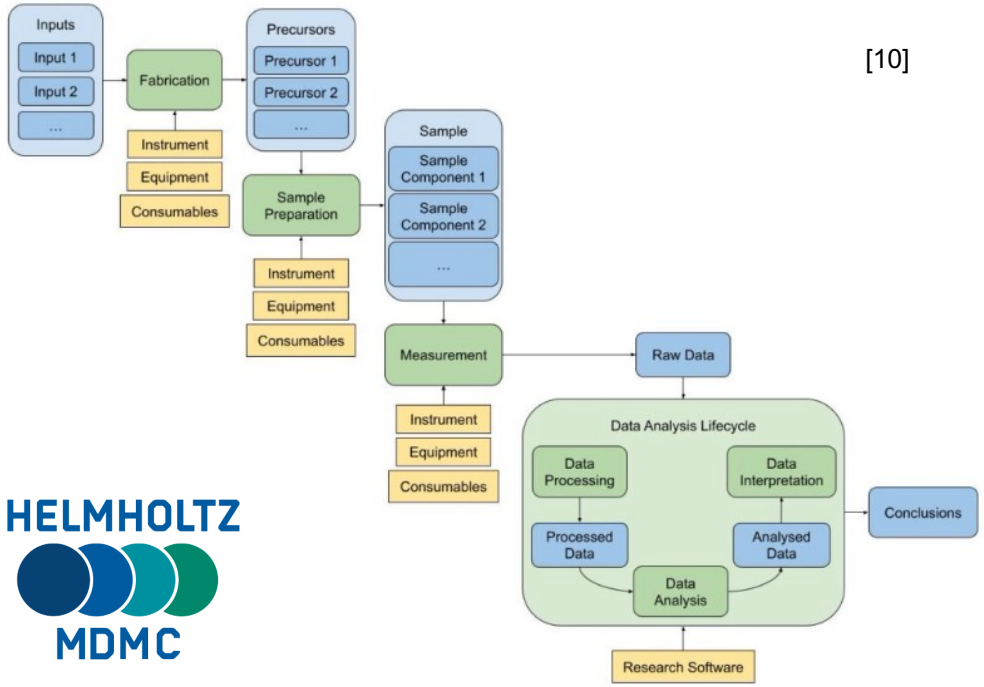
- Dislocation dynamic (DD) simulation
  - We use Dislocation Ontology (DISO) to annotate dislocation dynamics data (data linking)
  - Dislocation dynamics data consists of: structure, crystal structure, simulation parameters, and simulation provenance.
  - We generate dislocation simulation linked data as knowledge graph, **Dislocation Knowledge Graph (DisLocKG)**
  - Via its SPARQL endpoint, one can query dislocation simulation data



DDD simulation data as linked data.
Modeling Dislocation Dynamics Data Using Semantic Web Technologies,
AZ Ihsan, S Fathalla, S Sandfeld, ArXiv pre-print https://arxiv.org/abs/2309.06930

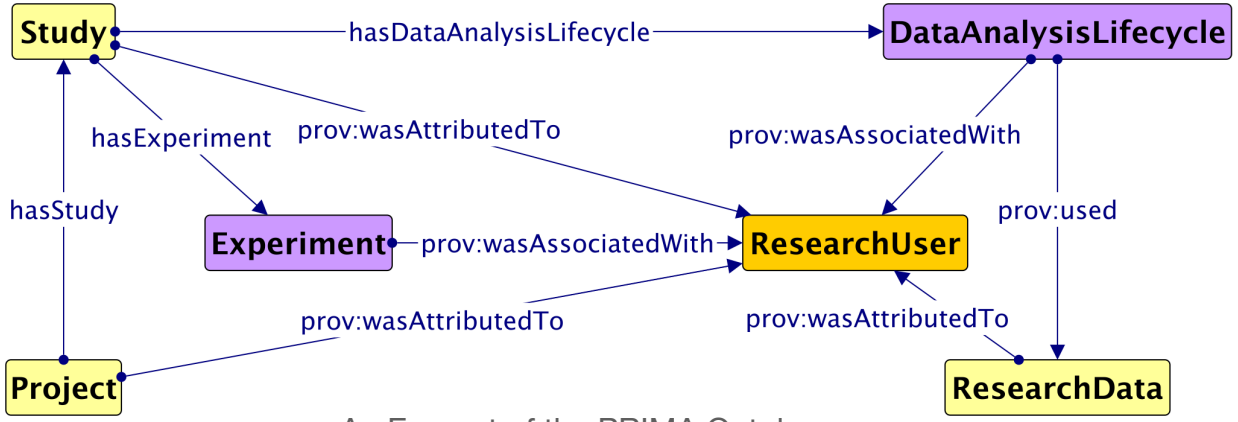
ONTOLOGY APPLICATIONS
PROVENANCE ONTOLOGY FOR EXPERIMENTAL WORKFLOW

- The Provenance Information for Materials Science (PRIMA) Ontology:
  - Formalize the experimental workflow into an ontology
  - Used to annotate the experimental data and to increase the reusability by capturing provenance information.



[10]

Basic steps experimental workflow



An Excerpt of the PRIMA Ontology
AZ. Ihsan, M. Jalali, R. Aversa, M. Panighel, E. Osmenaj.

PRIMA

Collaboration through use cases

nffa.eu nanoscience foundries & fine analysis

herbie hereon

[10] https://jl-mdmc-helmholtz.de/mdmc-activities/metadata-working-group/metadata-wg-topics/experimental-workflow/

JÜLICH Forschungszentrum

TAKE-HOME MESSAGE

- Understanding communication involves an encoding (syntax) - decoding (semantics) process.
- Context influences the interpretation of meaning in understanding.
- Formal knowledge representation overcomes the problem of language in knowledge representation.
- Ontology is a kind of formal knowledge representations.

ACKNOWLEDGEMENTS

HMC HELMHOLTZ Metadata Collaboration
HELMHOLTZ MDMC
erc European Research Council Established by the European Commission
NFDI MatWerk