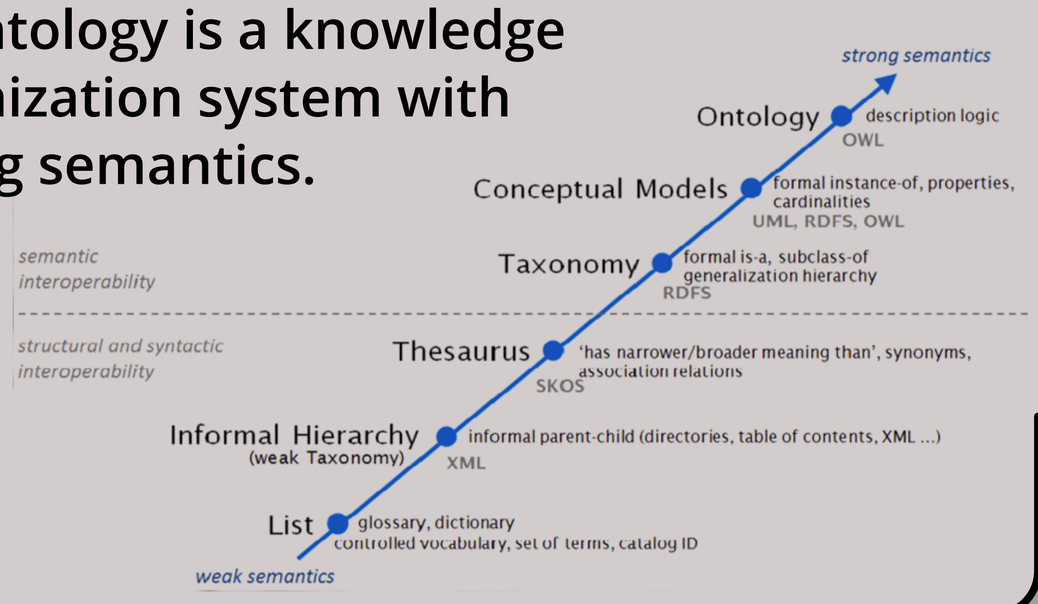
Ontologies in Computational Materials Science

FRITZ-HABER-INSTITUT
MAX-PLANCK-GESELLSCHAFT

Maja-Olivia Lenz, Luca M. Ghiringhelli, Carsten Baldauf, Matthias Scheffler
Email: lenz@fhi-berlin.mpg.de

NOMAD
NOVEL MATERIALS DISCOVERY

An ontology is a knowledge organization system with strong semantics.

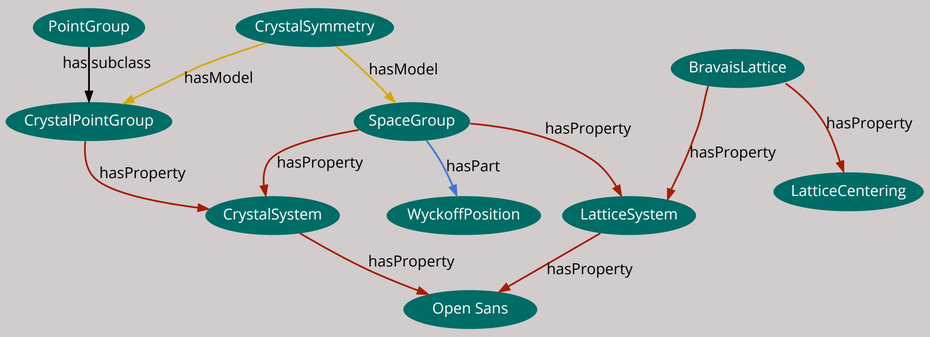


"An ontology is a formal, explicit specification of a shared conceptualization."

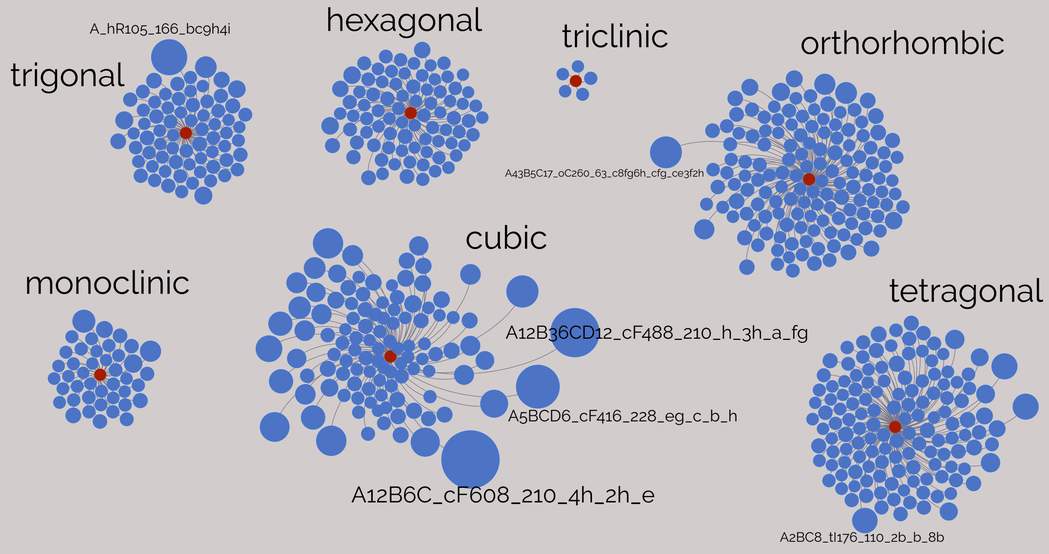
- formal → Machine-readable
- explicit specification → Concepts, properties, relations, functions, constraints, axioms are explicitly defined
- shared → Consensual knowledge
- conceptualization → Abstract model and simplified view of some phenomenon in the world that we want to represent

The Crystal Structure Ontology

Semantically represent crystal structures and their symmetries
Extract for crystal symmetry at class level



Ontology can be instantiated with AFLOW Library of Crystallographic Prototypes
Querying this knowledge graph allows visualization as network:
Prototypes (blue) size-scaled by number of atoms clustered in crystal systems (red)



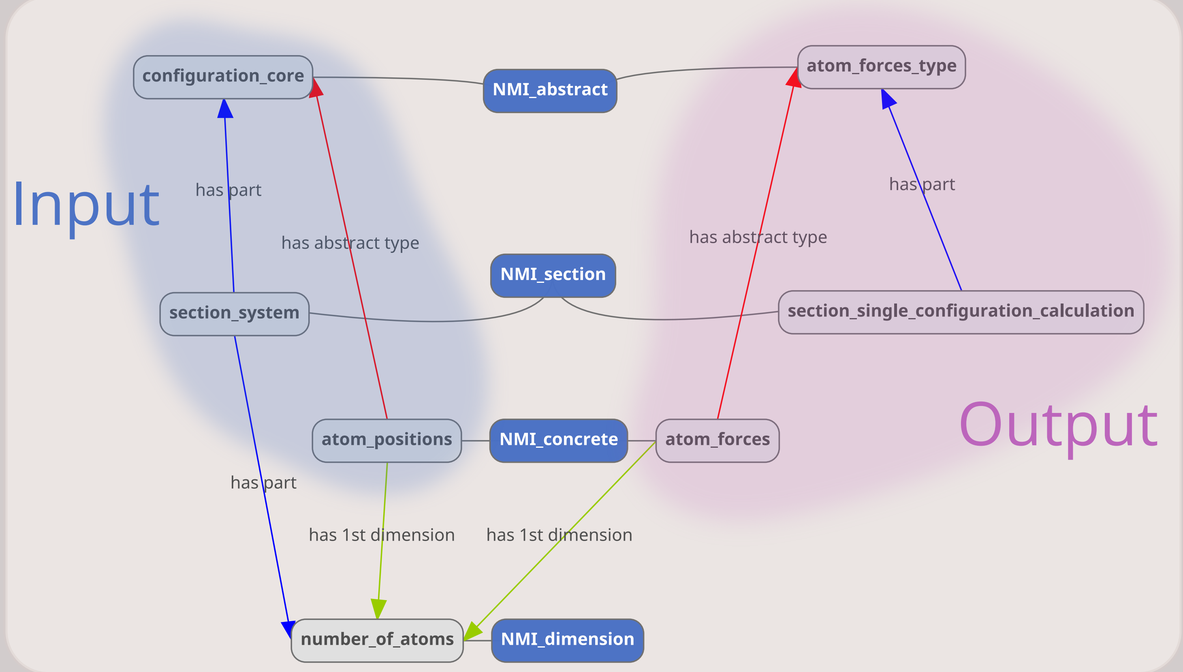
The NOMAD Meta Info

The NOMAD Meta Info is the meta data scheme for data in the NOMAD Archive – the largest database for normalized data in computational materials science.

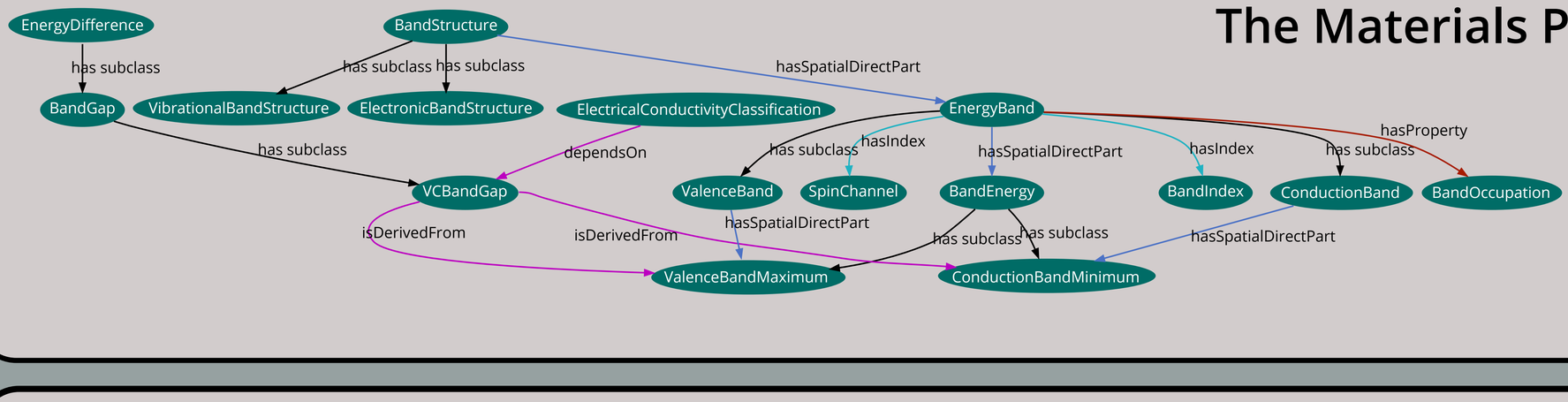
There are currently four types of meta data that structure the data and assign relations between them:

Concrete values are the labels to the values (strings, scalars, vectors, ...) parsed by parsers.
Sections represent different parts of a computer simulation.
Abstract types are mata data for meta data, they describe the type fo data that is labeled by a Concrete Value or a Section
Dimensions classify some meta data terms as integers that define the lengths of a dimension of a non-scalar Concrete Value

The NOMAD Meta Info contains 5 types of relations. 4 are shown as arrows.
The 5th is the has reference which points from section to section.
⟶ Can be represented ontologically and linked to more semantic ontologies



The Materials Properties Ontology

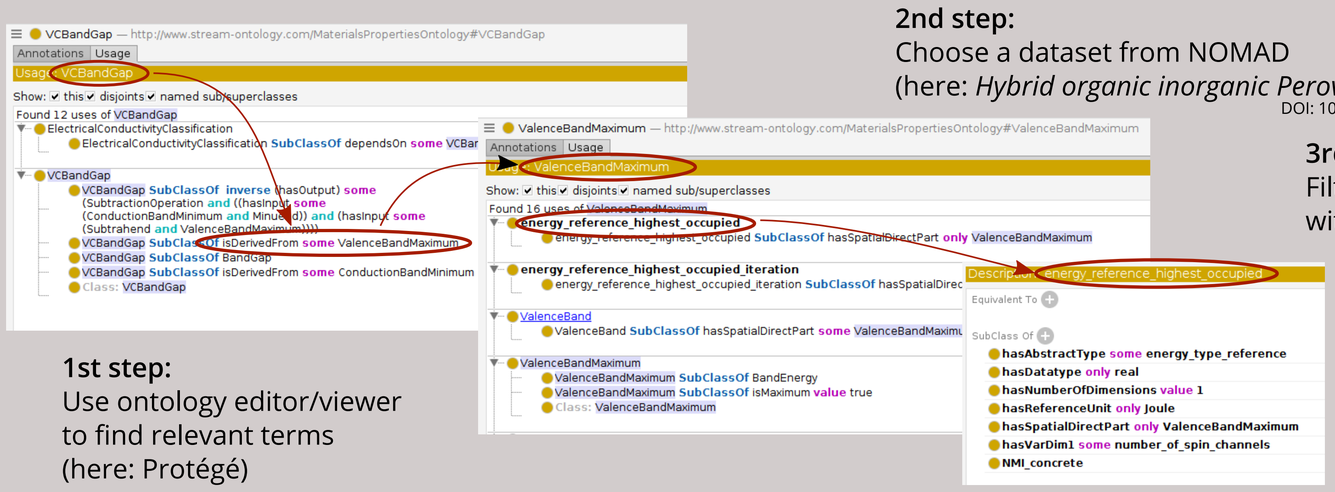


Important properties occuring in computational materials science:

Scalar:
- Total energy
- Band gap

Non-scalar:
- Atomic forces
- Stress
- Band structure

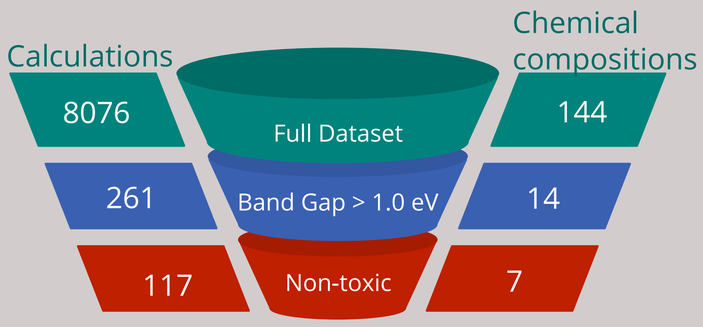
Application: Search for a better solcar cell material



1st step:
Use ontology editor/viewer to find relevant terms
(here: Protégé)

2nd step:
Choose a dataset from NOMAD
(here: *Hybrid organic inorganic Perovskites*)
DOI: 10.17172/NOMAD/2017.03.15-1

3rd step:
Filter for materials with band gap > 1.0 eV *
⟶ reduces number of materials in dataset by factor 10



4th step:
Filter out all materials containing elements that cause some intoxicating effect according to Wikidata

*Shockley Queisser criterion for maximum efficiency of solar cell: Band gap ~ 1.3 eV