Introducing MAMBO

Materials And Molecules Basic Ontology

Fabio Le Piane

fabio.lepiane@ismn.cnr.it

CNR-ISMN / University of Bologna / Italy

June 7, 2021

DAIMON
www.daimoners.eu

Introduction

- Advanced materials and their applications have become a key field of research
- Need for systematic and efficient methods for organizing knowledge in the field and conduct computational (multiscale modelling) or experimental investigations
- Moreover, the recent developments of data-driven technologies led to significant progress in most strategic fields, but they need a very systematic and organized approach to data collection and management

For these reasons, organizing knowledge in a systematic fashion is quickly becoming an undelayable necessity

DAIMON

Related works

- EMMO: a multidisciplinary effort to develop a standard representational ontology for applied sciences
- ChEBI: a dictionary of molecular entities focused on ‘small’ chemical compounds
- MDO: an ontology for materials design field, representing the domain knowledge specifically related to solid-state physics and computational materials science
- OSMO: namely, the Ontology for Simulation, Modelling, and Optimization

MAMBO also aims at connecting with pre-existing materials databases:

- OPTIMADE
- NOMAD

Application scenarios

MAMBO is not intended to be a foundational ontology for the whole materials science. Instead, it aims to fill one gap left from the previous ontologies: molecular materials and related systems.

An operational ontology for this sub-domain is still lacking, with the aforementioned ontologies focusing on different aspects of materials science (like MDO does with crystalline structures)

In alternative, one could rely on EMMO, which is a huge projects with the aim to be a comprehensive knowledge base for materials science (molecular materials included) but with no specific operational goal or sub-domain

On the other hand, MAMBO aims to be an ontology tailored for helping researchers in molecular materials development and application to accelerate their work and easing the sharing of the results

Application scenarios

- Retrieving structured information

- Supporting the development of new, complex workflows for modelling systems based on molecular materials
  - Modelling complex computational workflows for specific problems related to materials science
  - Organize the process of using data obtained by simulations for data-driven techniques in order to realize predictive models (property predictions, designing new materials, …)

- Integrating data generated via computational simulations and experimental workflows

Development process

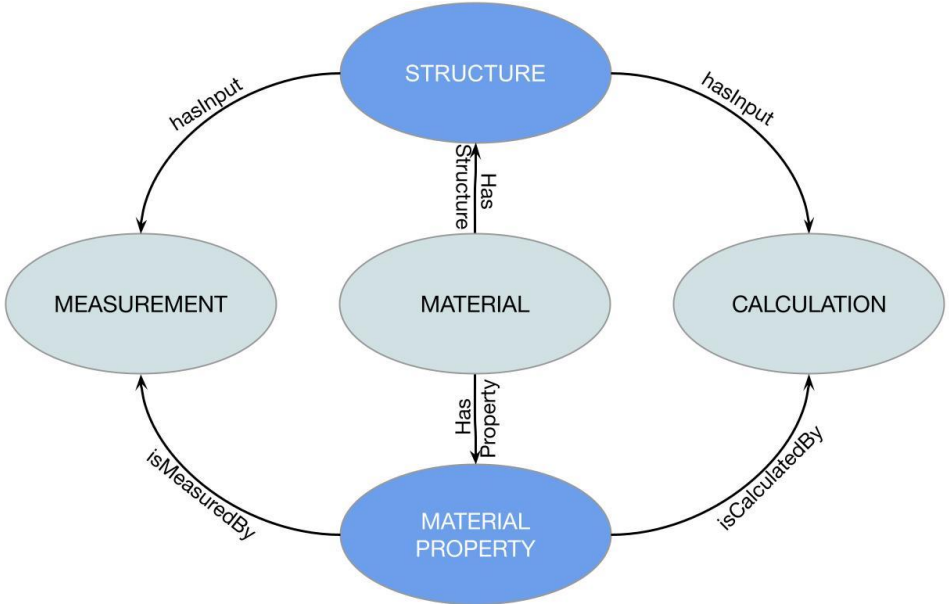
The process started with meetings with domain experts:

- A set of competency questions
- A set of tasks to organize
- A set of use cases

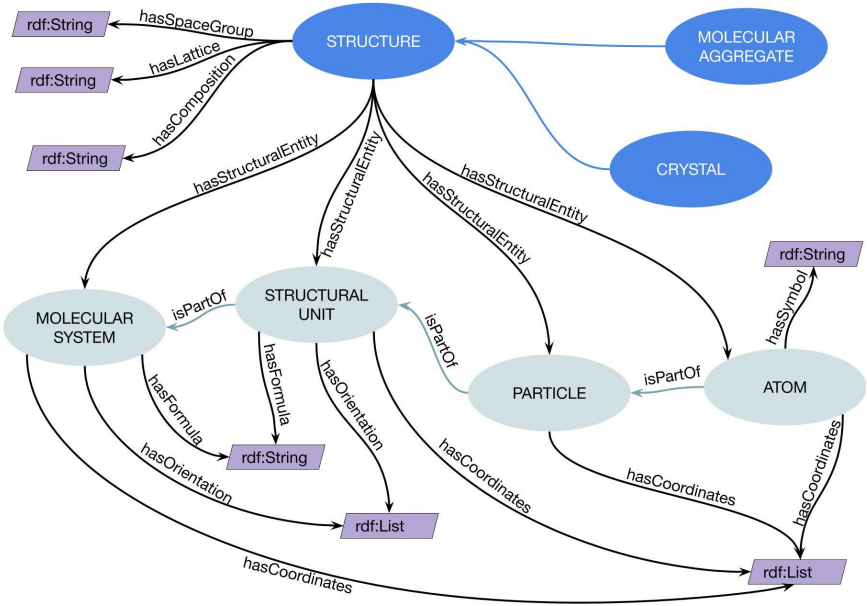
Due to the peculiar nature of the typical development approaches pursued in the considered application area, we modelled the main concepts of the ontology associating them to specific problem solving methods (PSMs)

Core concepts

- The central concept is the one of (molecular) material
- A material is described by the collection of its structural characteristics and its chemico-physical properties
- A Structure or a Property may be obtained via a empirical experiment (Measurement) or a computational simulation (Calculation)
- All of these concepts are present in MAMBO as classes



The Structure class



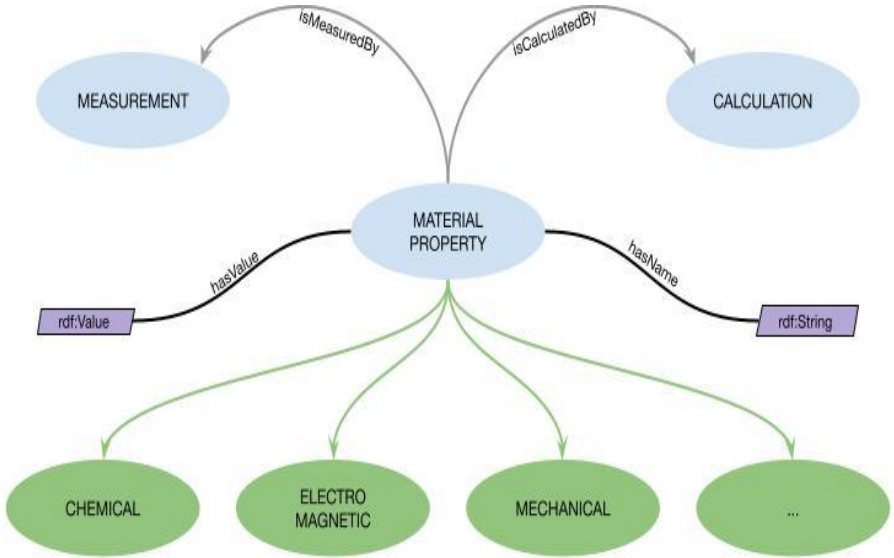
- A Structure is composed by one or more structural entities
- It has many sub-classes, two of which are Molecular Aggregate (peculiar of MAMBO) and Crystal (which could serve as an integration point with MDO)
- It can be described with many characteristics (Has it got a space group? Or a lattice? Which is its composition?)

DAIMON

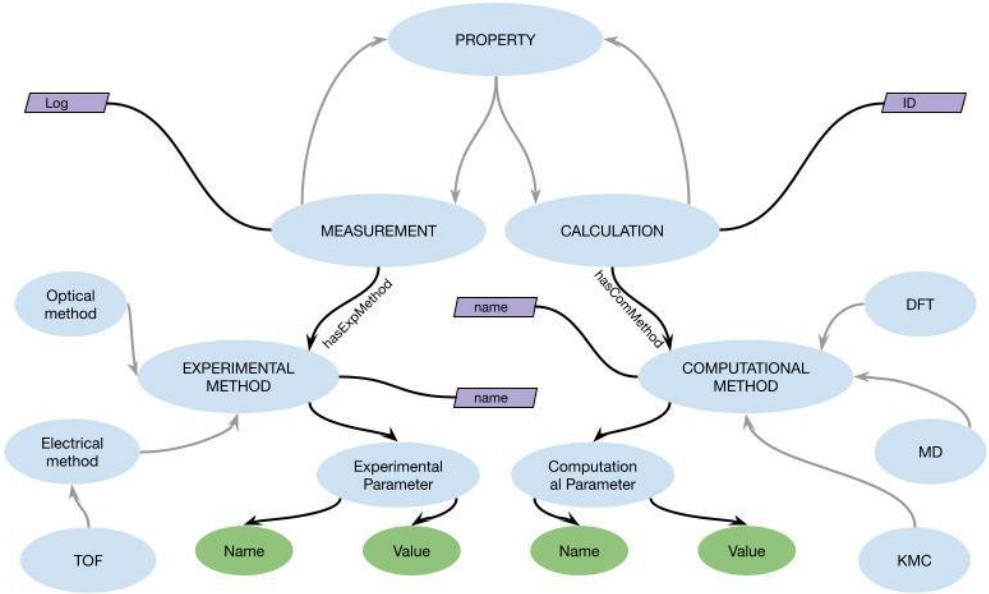
The Property class

- A property is summarised as its value and its name and properties are divided into different types/categories

- A property can be determined with a measurement or a calculation



Measurement - Calculation



- Both have a corresponding Method class, which collects the different, related methods and techniques (optical methods, DFT, MD, …)
- At the same way, both collects the parameters of their respective methods

- A similar relationship will be developed with the Structure class

Future steps

- Fix naming convention
- Assess effectiveness of different implementation strategies
- Instantiation
- Extension to specialized domains
- Design a database for molecular materials

DAIMON