An ontology for units of measures across history, standards, and scientific and technology domains

Oskar Andersson*, Huanyu Li, Patrick Lambrix, Rickard Armiento

* Materials Design and Informatics, Theoretical Physics,
Department of Physics Chemestry and Biology, Linköping University, Sweden

LINKÖPING UNIVERSITY

Physical units and databases

- The world's science relies on correct use of units of measure.
- Lost satellite rover due to assumption of metric units when software data output were in English units.
- Data collected today are commonly stored in databases. Units are crucial to communicate quantities in this data.

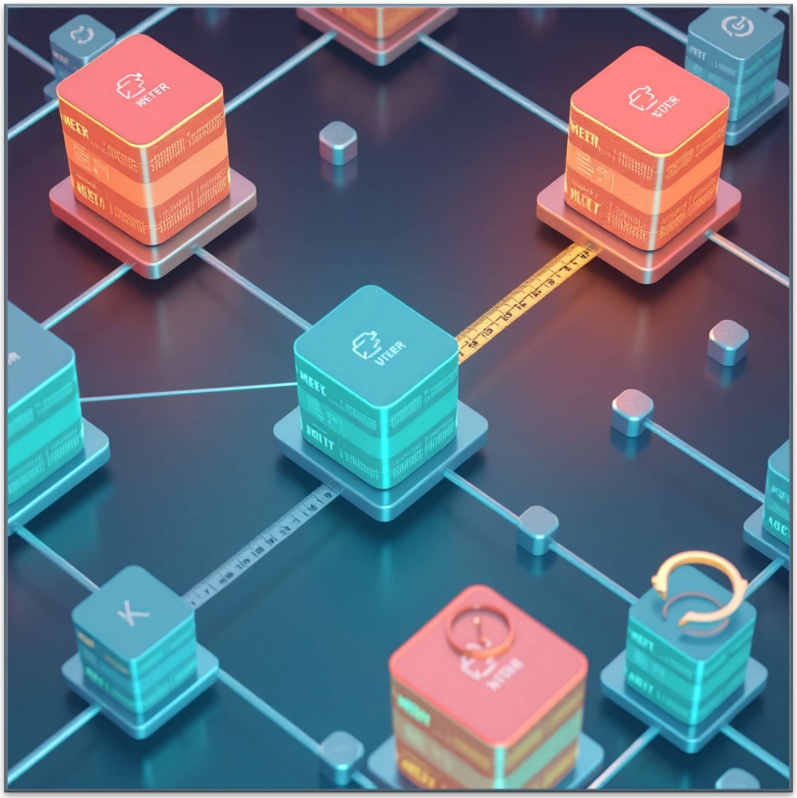


Image generated with Flux AI

Unit definitions and impact of its change

- Unit definitions, e.g., metre:
  - 1960: Defined in terms of wavelengths in vacuum of the radiation of the krypton 86 atom.
  - 1983: Redefined in terms of the speed of light.
- Meaning of stored quantity changes when unit changes
  - $L = 2.589321933608542278 \cdot 10^{-9}$ m (assume [m 1960])
  - $L = 2.589321932352753296 \cdot 10^{-9}$ [m 1983] †
  - Even more complex for other units, e.g., electron volt.
- Database workers need to know which [m] is used.
  - Find in code base? Through publication year?
- Compound unit expressions: $Pa = kg \cdot m^{-1} \cdot s^{-2}$.
- Reinterpretation when transferring between standards, e.g., SI based systems forces conversion of CGS based data.

LINKÖPING UNIVERSITY

†. 1 [1960 m] = 1650763.73 wavelengths of krypton 86. Wavelength of Krypton $605780210.3 \cdot 10^{-15}$ [1983 m] => 1 [1960 m] = 0.999999999515012419 [1983 m]

Imprecise accuracy of Krypton 86 wavelength measurements => Meaning of difference is unknown

Approaches to units of measure ontologies

- Exists ontologies to describe materials data stored in various databases expressed in different units.
  - Materials Design Ontology (MDO)[1].
  - Using existing units of measure ontologies.
- Two approaches to craft such unit ontologies:
  1. Ontology with as precisely and scientifically model quantities and units. (Most are here.)
  2. Precisely describe information present in existing databases.
- 2) enables database integration and clear communication of published data

1. P. Lambrix, R. Armiento, H. Li, O. Hartig, M.A.N. Pour, Y. Li, 2022

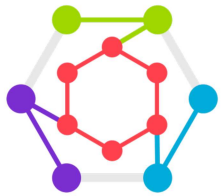
Existing unit ontologies

- QUDT – *"architecture for the conceptual representation of quantities, quantity kinds, units, dimensions, and data types."*
  - SI based through dimensions/quantities. Compound expressions as separate concepts, e.g., unit:KiloGM-PER-PA-SEC-M.
- UCUM – *"code system intended to include all units of measures being contemporarily used in international science, engineering, and business."*
  - SI based through fundamental design of unitsystems closely tied to SI unitsystem. Symbols are standardized to one single unit.
- And others…
- None of the surveyed ontologies support inclusion of changes in unit definitions over time.

OPTIMADE

Open Databases Integration for Materials Design

- Community driven effort for common API for databases.
- OPTIMADE's latest release: standardization work of property definitions.
- Part of property definitions is physical unit definitions useful for materials science.
- Our work is based on these physical unit definitions.



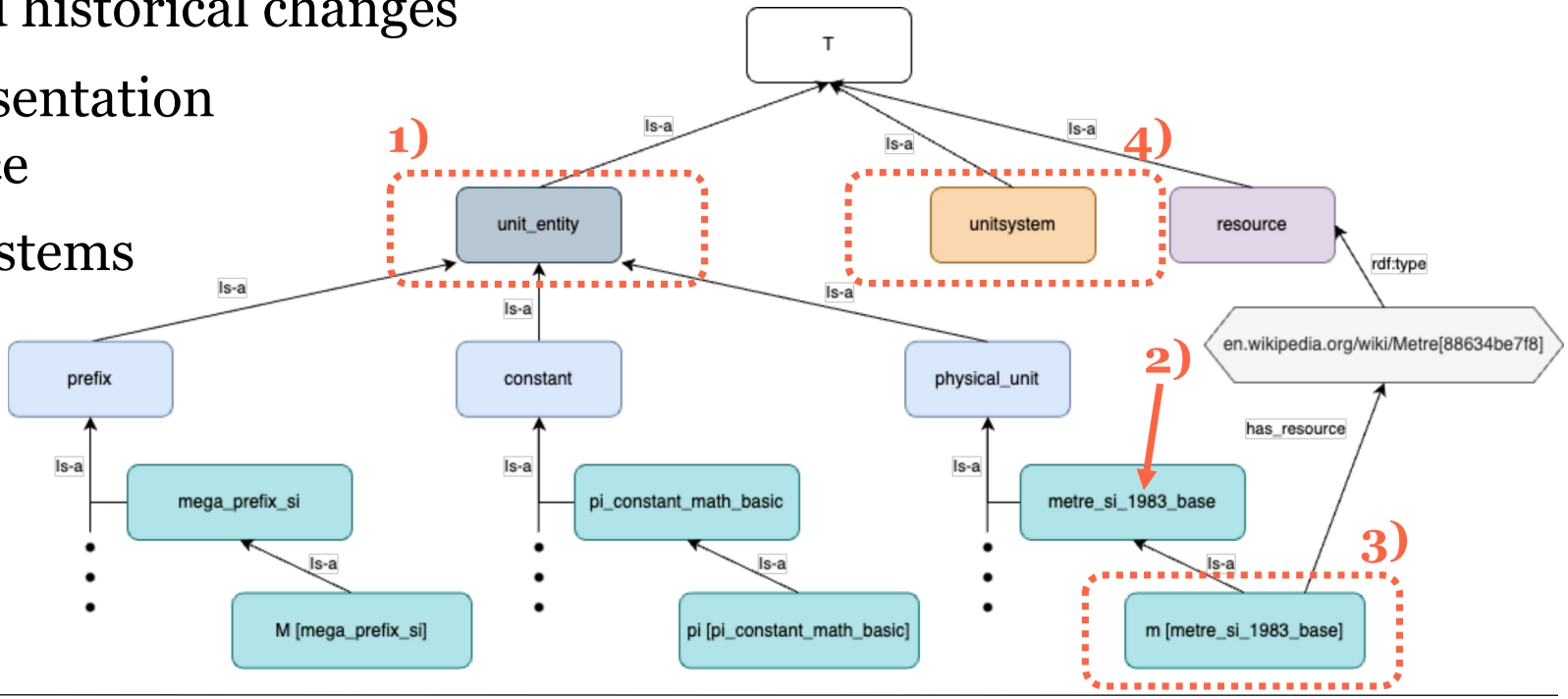

```
{
    "$id": "https://schemas.optimade.org/defs/v1.2/units/si/general/atomicmassunit",
    "$schema": "https://schemas.optimade.org/meta/v1.2/optimade/physical_unit_definition.json",
    "title": "atomic mass unit",
    "symbol": "u",
    "display-symbol": "u",
    "alternate-symbols": [
        "dalton",
        "Da"
    ],
    "description": "A unit of mass representing 1/12 of the mass of a free carbon 12 atom (i.e., a typical value of the
    "compatibility": [
        "https://schemas.optimade.org/defs/v1.2/units/si/1970/accepted/atomicmassunit",
        "https://schemas.optimade.org/defs/v1.2/units/si/1998/accepted/dalton",
        "https://schemas.optimade.org/defs/v1.2/units/si/general/dalton"
    ],

    ],
    "approximate-relations": [
        {
            "base-units": [
                {
                    "symbol": "kg",
                    "id": "https://schemas.optimade.org/defs/v1.2/units/si/general/kilogram"
                }
            ],
            "base-units-expression": "kg",
            "scale": {
                "value": 1.6605390666e-27,
                "standard_uncertainty": 5e-37
            }
        }
    ],
    "x-optimade-definition": {
        "label": "atomicmassunit_si_general",
        "kind": "unit",
        "format": "1.2",
        "version": "1.2.0",
        "name": "atomicmassunit"
    }
}
```

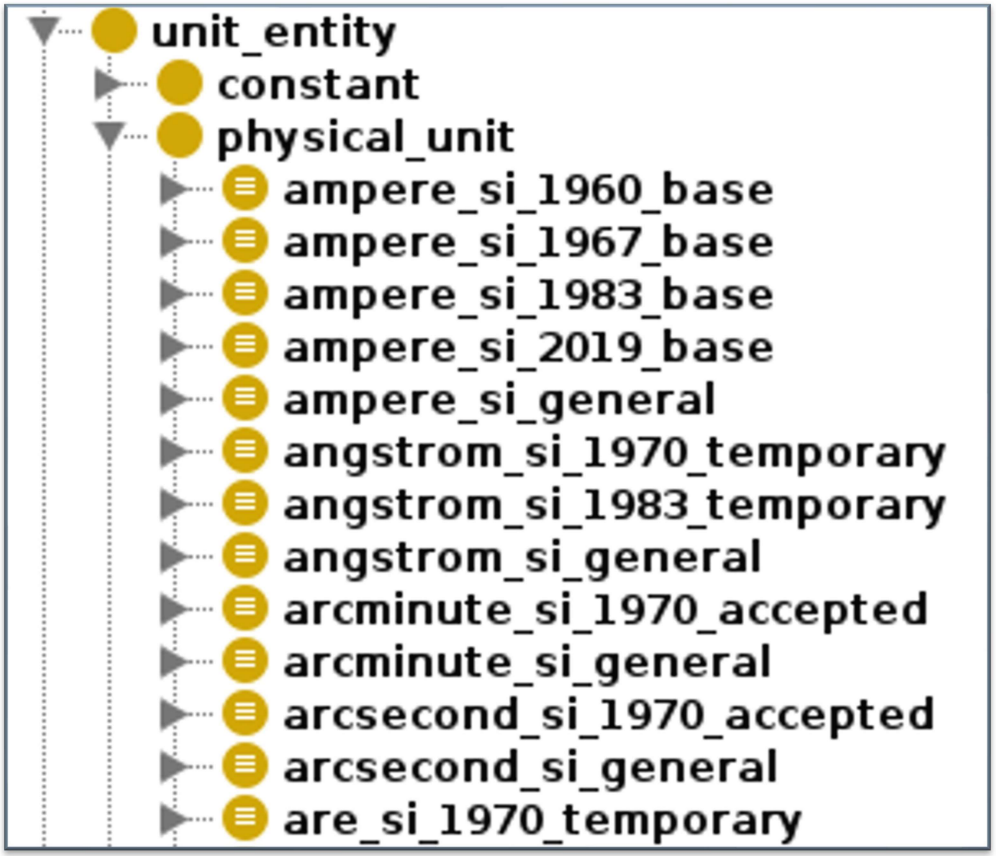
Ontology design – overview

1. Includes unit entities
2. Unit definition and historical changes
3. Rich symbol representation through inheritance
4. Customized unitsystems
5. Standard agnostic (not bound to SI)

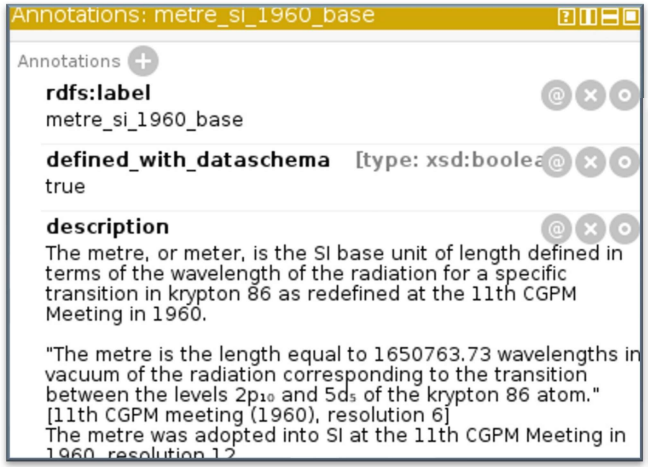
5) No dimensions or categories



Ontology design – historical definitions

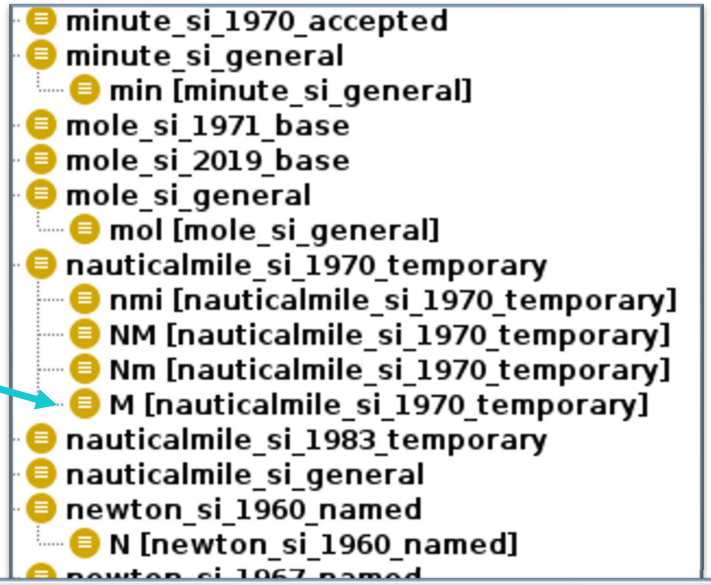


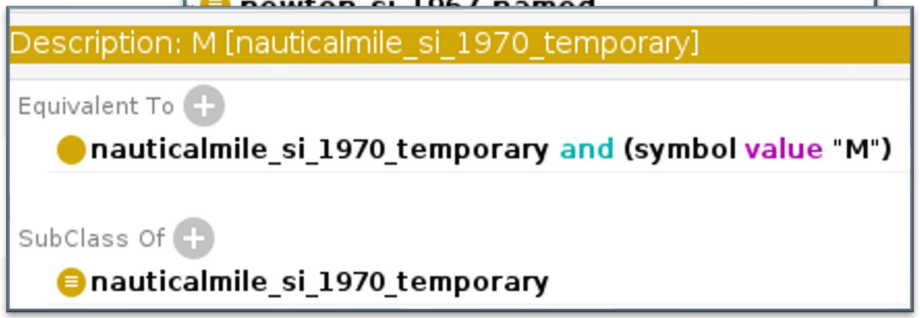
- Each historical definition as separate concepts.
- Allows: Preservation of exact historical definition and original meaning.



Ontology design – symbols

- Reference to exact definition through symbol concepts.
- Allows:
  - Same symbol for different units.
  - Different symbols for same units
- Symbol variations:

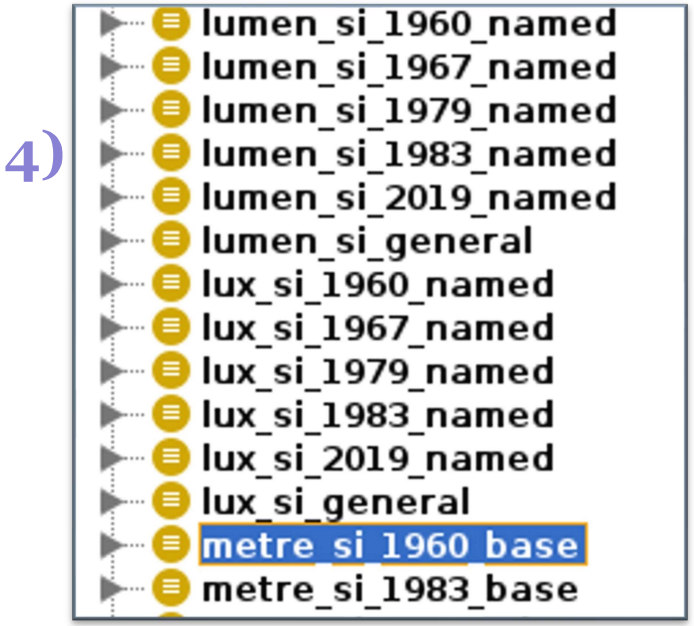




- Compound expressions:
  - String expression, base units are linked to symbol concept.

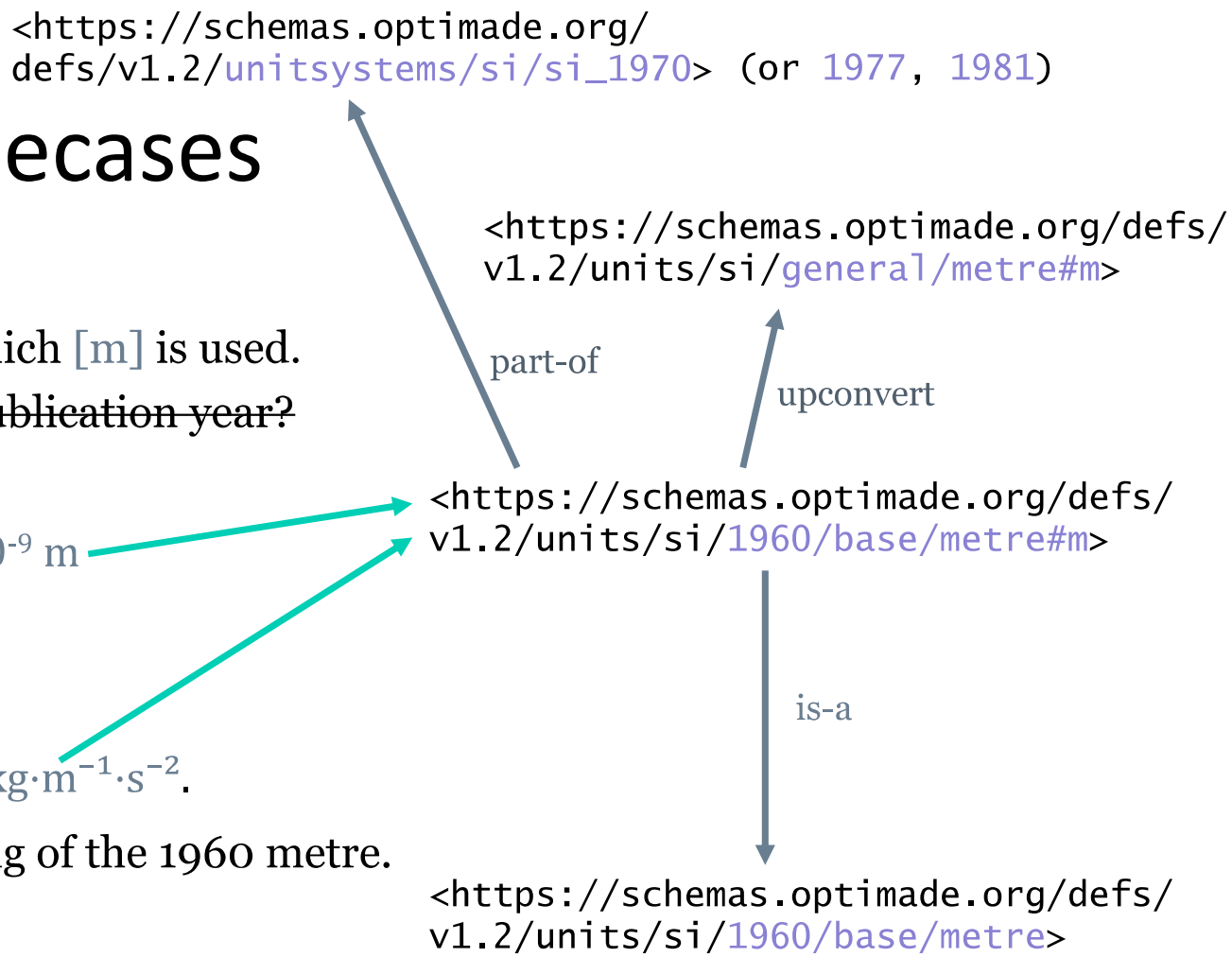
Versioning and IRIs

1. Retain IRIs from OPTIMADE standard unit definitions, includes OPTIMADE semantic versioning.
2. Changes on OPTIMADE level is contained in concepts IRI.
3. Changes on ontology design is contained in ontology version IRI.
4. Unit redefinition by standard organization yields new concept.



Application and usecases

- Publishing data.
- Database workers need to know which [m] is used.
  - ~~Find in code base? Through publication year?~~
- Which [m] is used? Link to IRI
  - L = 2.589321933608542278·10^{-9} m
- Context with unitsystem.
- Semantic context, e.g., JSON-LD.
- Compound unit expressions: Pa = kg·m^{-1}·s^{-2}.
- Unit concept holds original meaning of the 1960 metre.
- Can find related definitions.



Summary and outlook

- A units of measure ontology capable of:
  - Preservation of original historical unit definition
  - Flexible usage of symbols useful in science and technology
  - Support for units across all standards
- Outlook:
  - More unit definitions.
  - Include dimensions in a unitsystem-agnostic way.
  - Extend work to full property definitions (includes more data formats and units).
- Collaborators: Huanyu Li, Patrick Lambrix, Rickard Armiento.
- Contact: oskar.andersson@liu.se




LINKÖPING UNIVERSITY

SeRC Swedish e-Science Research Centre

Knut and Alice Wallenberg Foundation

Vetenskapsrådet