TALOS ERA CHAIR IN ARTIFICIAL INTELLIGENCE FOR HUMANITIES AND SOCIAL SCIENCES

ΠΑΝΕΠΙΣΤΗΜΙΟ ΚΡΗΤΗΣ
UNIVERSITY OF CRETE

What is Ontology?

Prof Christophe Roche

ERA Chair Holder - University of Crete

https://christophe-roche.fr/

May 2025

Horizon ERA Chair TALOS AI4SSH Project funded by the European Commission
Grant Agreement n° 101087269, https://cordis.europa.eu/project/id/101087269

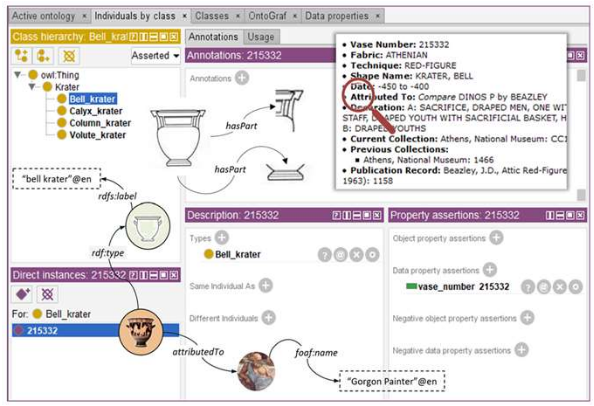
Funded by the European Union

This MOOC is an introduction to ontology in the sense of Knowledge Engineering. Used to represent, structure and share knowledge in a domain, ontologies are at the heart of IT applications that require a semantic dimension.

What is Ontology?

Contents

1. **Introduction**
2. **Definition**
3. **Examples**
4. **Theories of Concept**
5. **Representation Languages**
6. **Environments**

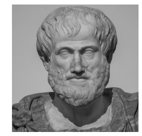


This MOOC is divided into six parts. We will begin by reviewing the origins of this discipline to arrive at its definition in the sense of knowledge engineering. Three examples will illustrate the importance of this concept in different fields. We will then briefly discuss different theories of concept, which define as many approaches to knowledge modelling. Finally, we will list the families of languages used to represent ontologies and conclude with two ontology-building environments—one based on the notion of class, and the other on the notion of concept.

1. Introduction

1 **“The branch of metaphysics dealing with the nature of being”** OXFORD UNIVERSITY PRESS

“The science of being as being… independently of its particular determinations”

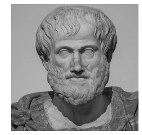


Ontology has a long history, with roots in philosophy, and more precisely in metaphysics. It first appeared with the Greek philosophers, notably Aristotle. It is defined as the science of being as being—that is, what is essential in things, independently of the states they may be in. The term comes from the Greek *ontos* (being) and *logos* (science, language, discourse).

1. Introduction

1

“The branch of metaphysics dealing with the nature of being” OXFORD UNIVERSITY PRESS



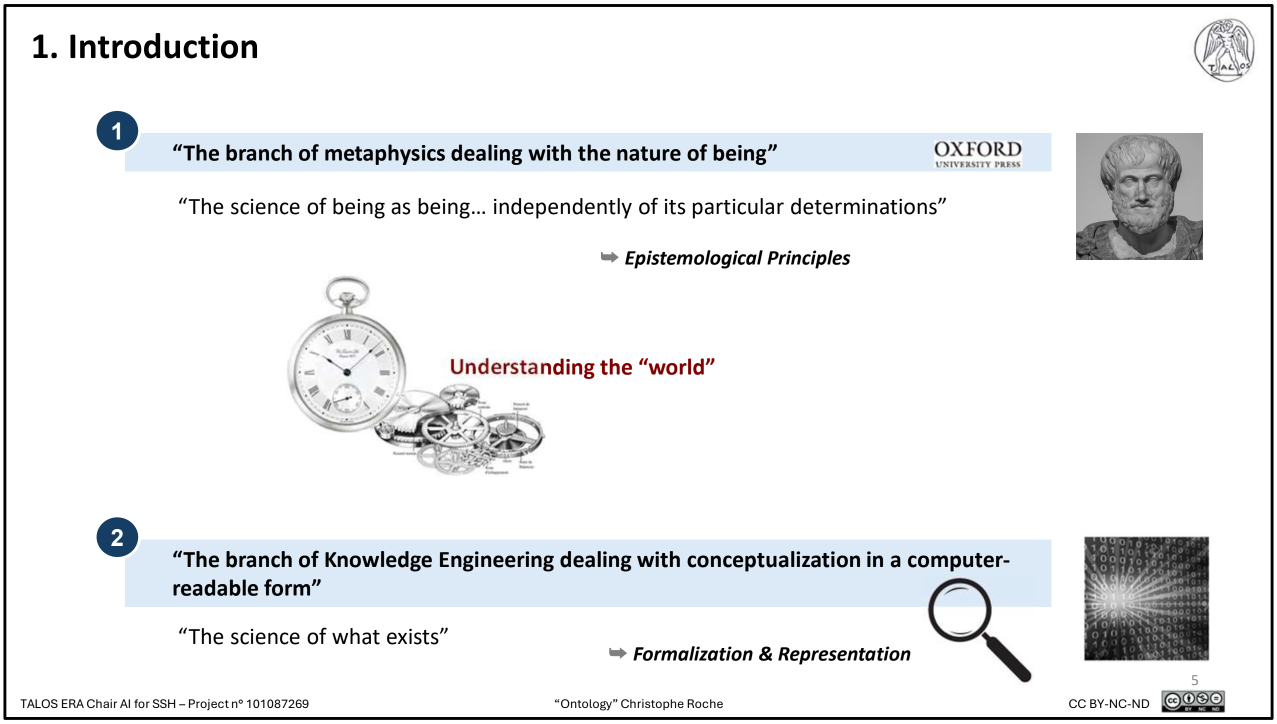
“The science of being as being… independently of its particular determinations”

2

“The branch of Knowledge Engineering dealing with conceptualization in a computer-readable form”

“The science of what exists”

With the advent of computer science, artificial intelligence, and the semantic web, the term “ontology” has acquired a new meaning. Ontology now refers to that branch of knowledge engineering focused on the conceptualisation of a domain in a form that a computer can read and understand for information processing.

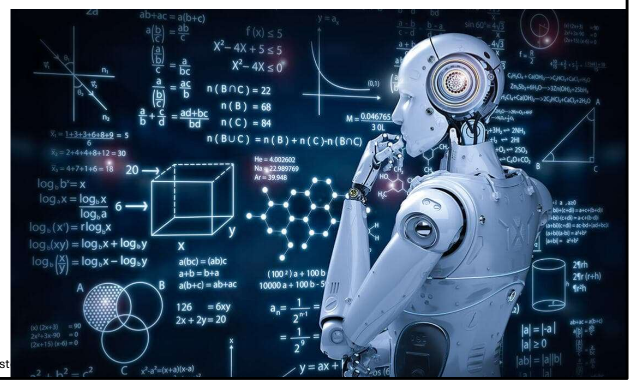


The borrowing of the word "ontology" from philosophy by knowledge engineering is justified by a shared concern: understanding the world and the objects that populate it. The means, however, differ. Metaphysics focuses on epistemological principles (*epistēmē* meaning knowledge in Greek), while knowledge engineering is ultimately concerned with the formalization and representation of ontologies for computational purposes.

1. Introduction

Knowledge Engineering

- ✓ How to represent & organise the objects which populate the world in a computer-readable form?
- ✓ Everything that can be represented exists



TALOS ERA Chair AI for SSH – Project nº 101087269

“Ontology” Christ

The ontology of knowledge engineering aims to answer the following question: “How can the objects that populate a reality be represented and organised in a form that can be understood by a computer?” Bearing in mind that for artificial intelligence, anything that can be represented exists.

2. Definition

- **Set of Concept and Relationship Definitions**

What is an Ontology ? Short answer: **An ontology is a specification of a conceptualization**.

In the context of knowledge sharing, I use the term ontology to mean a *specification of a conceptualization*. That is, an ontology is a description (like a formal specification of a program) of the concepts and relationships that can exist for an agent or a community of agents. This definition is consistent with the usage of ontology as set-of-concept-definitions, but more general. *Tom Gruber*

There are different definitions of ontology depending on the conceptual approach and whether or not a linguistic dimension is taken into account.

An ontology can be defined as a specification of a conceptualization—that is, the definition of concepts and relationships of a domain using a language that a computer can understand.

2. Definition

- **Set of Concept and Relationship Definitions**

What is an Ontology ? Short answer: **An ontology is a specification of a conceptualization**.

In the context of knowledge sharing, I use the term ontology to mean a *specification of a conceptualization*. That is, an ontology is a description (like a formal specification of a program) of the concepts and relationships that can exist for an agent or a community of agents. This definition is consistent with the usage of ontology as set-of-concept-definitions, but more general. *Tom Gruber*

- **Vocabulary of Terms**

"An [explicit] ontology may take a variety of forms, but necessarily it will include **a vocabulary of terms and some specification of their meaning** (i.e., definitions)" *M.Ushold & M.Gruninger. Knowledge Engineering Review, Vol.11, n°2, June1996*

W3C "There is no clear division between what is referred to as "vocabularies" and "ontologies"."

OWL 2 Web Ontology Language Document Overview (Second Edition)
W3C Recommendation 11 December 2012

"Ontologies are formalized vocabularies of terms, often covering a specific domain and shared by a community of users. They specify the definitions of terms by describing their relationships with other terms in the ontology. "

It can also be seen as a formally defined vocabulary. Care must be taken not to confuse terms and concepts, linguistic relations between terms with conceptual relations between concepts. The formal definition of terms is in fact the formal definition of the concepts denoted by those terms, to which a natural language translation of the formal definition may be associated.

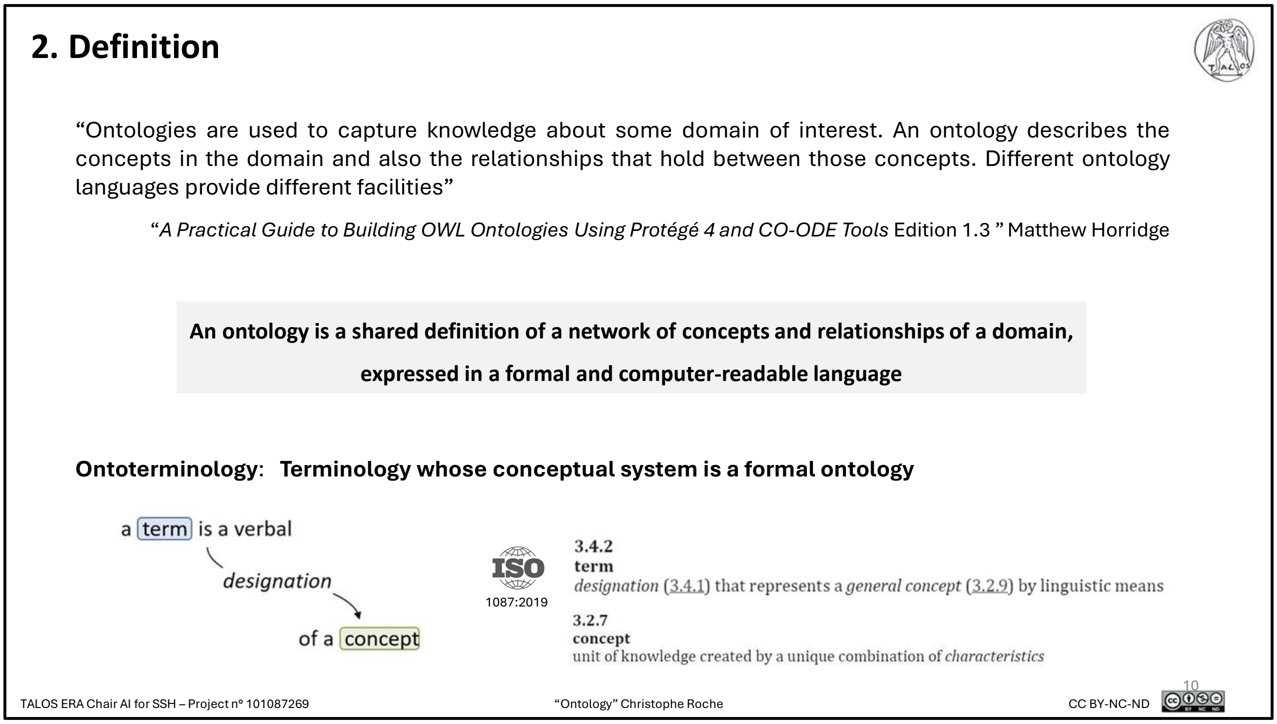
2. Definition

"Ontologies are used to capture knowledge about some domain of interest. An ontology describes the concepts in the domain and also the relationships that hold between those concepts. Different ontology languages provide different facilities"

"*A Practical Guide to Building OWL Ontologies Using Protégé 4 and CO-ODE Tools* Edition 1.3 " Matthew Horridge

An ontology is a shared definition of a network of concepts and relationships of a domain, expressed in a formal and computer-readable language

This is why we prefer definitions that emphasise the concepts of the domain and their relationships, which can be summarised by saying that: "An ontology is a shared definition of a network of concepts and relationships of a domain, expressed in a formal and computer-readable language".

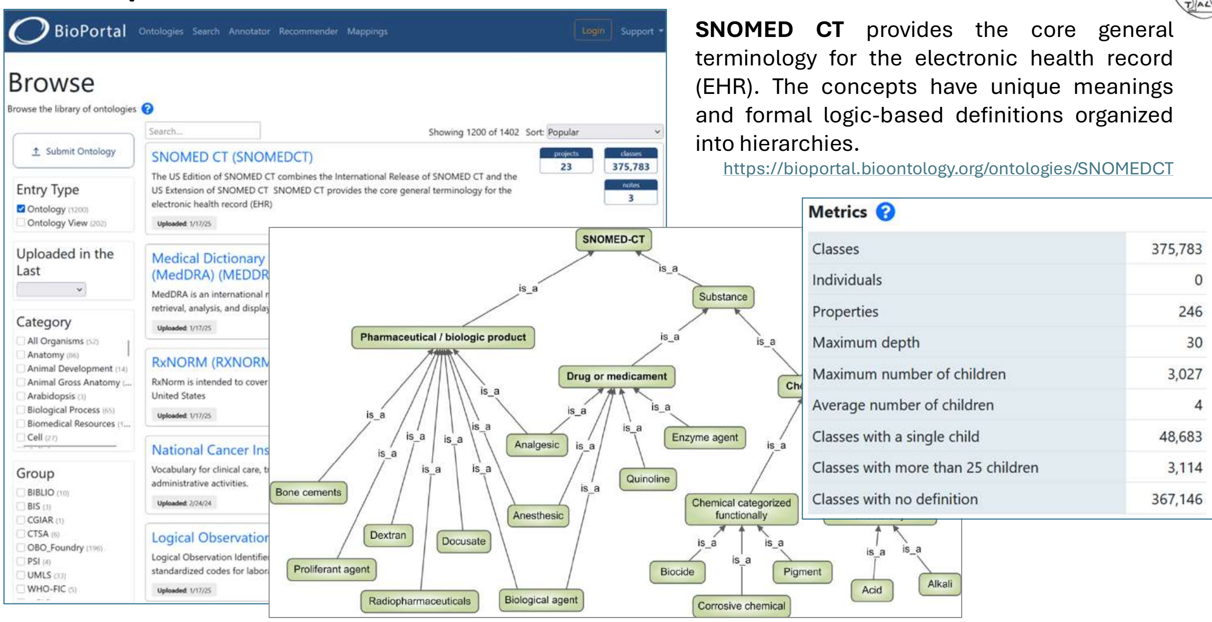


If we wish to take the linguistic dimension into account, it is preferable to use the notion of ontoterminology, a terminology whose conceptual system is a formal ontology, which also has the advantage of following the ISO definitions of term and concept (ISO 1087:2019).

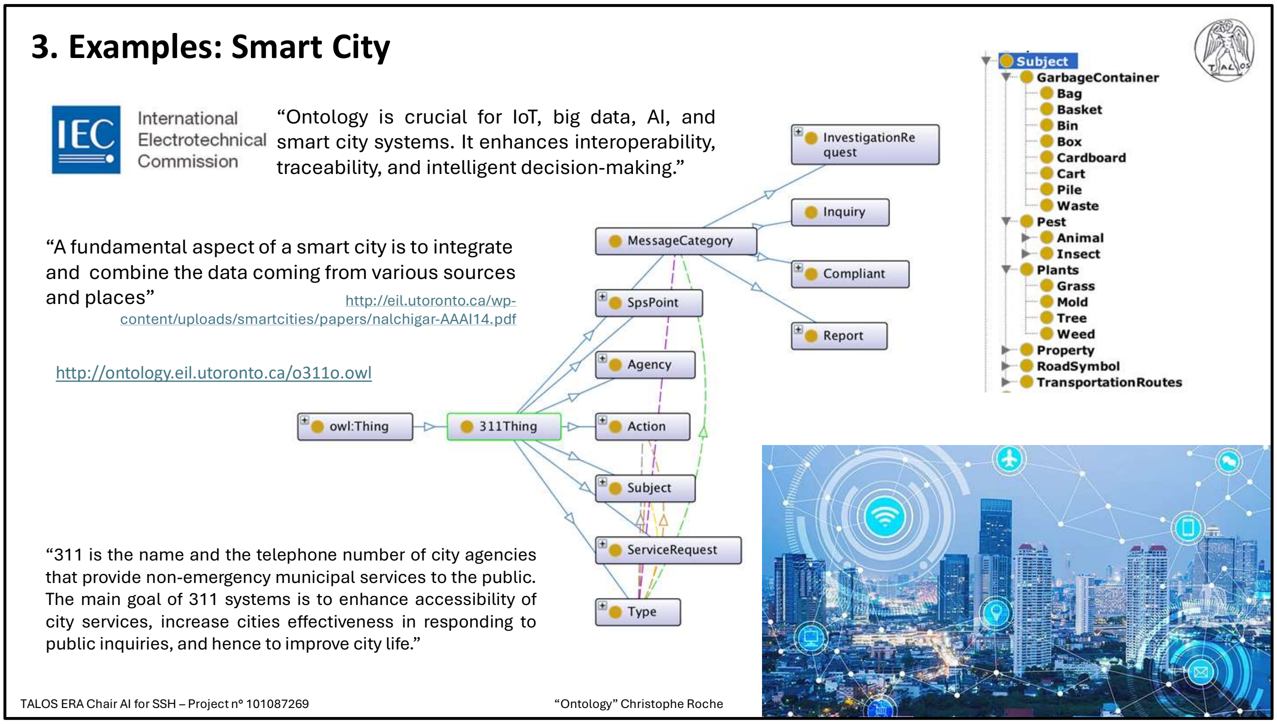
3. Examples: Medicine

SNOMED CT provides the core general terminology for the electronic health record (EHR). The concepts have unique meanings and formal logic-based definitions organized into hierarchies.

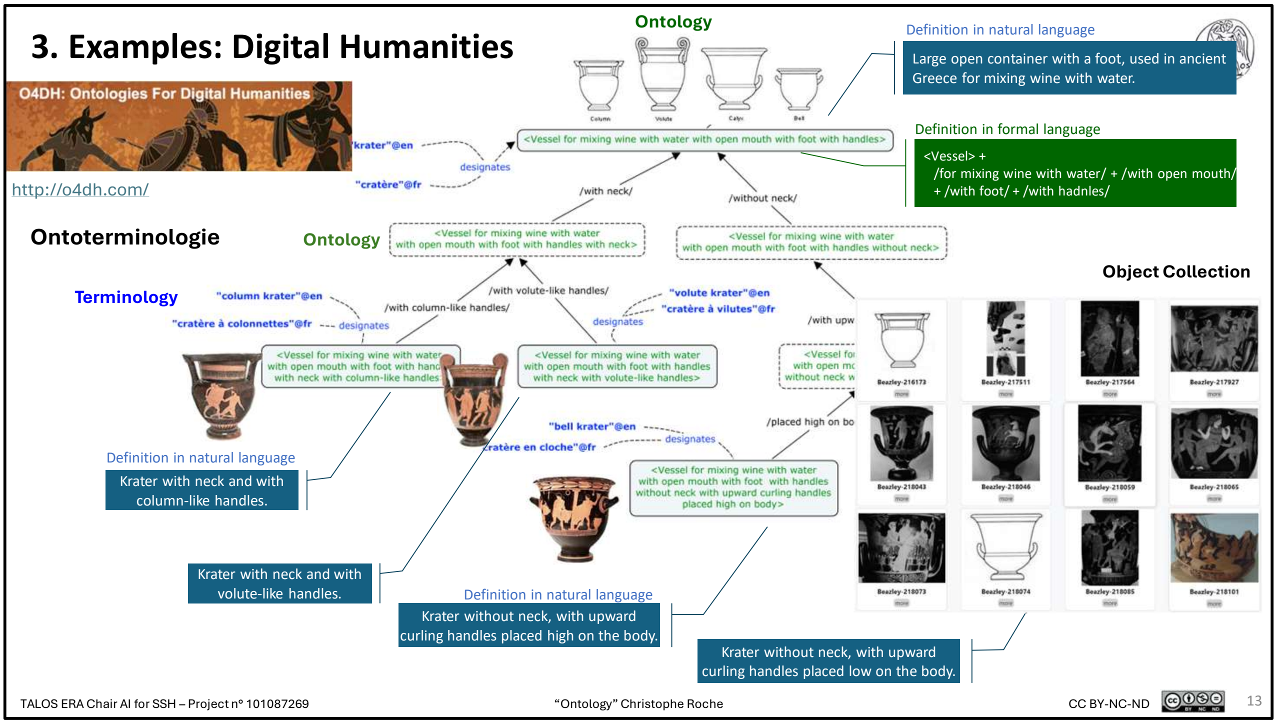
https://bioportal.bioontology.org/ontologies/SNOMEDCT



Ontologies are central to many applications. For example, in medicine, the BioPortal website provides access to several ontologies, including SNOMED CT, which defines the clinical terms used in medical records. These terms designate concepts defined on a logical basis and organised in a hierarchy.



A second example is the Smart City. Ontologies play a fundamental role in the interoperability between systems and the integration of data from different sources. For example, the "311 ontology" defines concepts for managing public services.

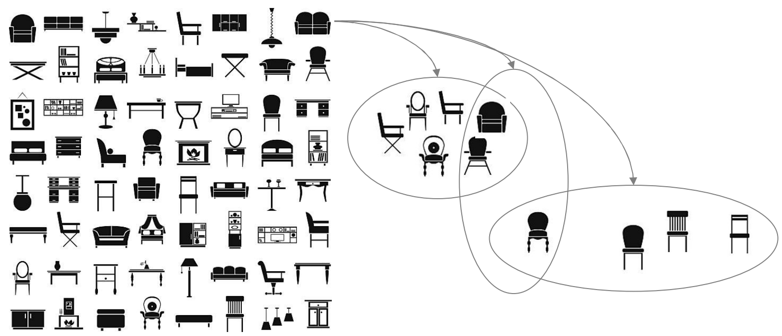


A final example is Digital Humanities. Ontologies enable artefacts to be organised and named, making them easier to manage. The use of ontoterminologies make it possible to clearly distinguish between the linguistic and conceptual dimensions of terminology, facilitating the construction of multilingual electronic dictionaries and the management of object collections.

4. Theories of Concept

<u>Concept</u>: Unit of knowledge about.... a plurality of things

✓ Understanding the "World"
✓ Organizing the objects



....verifying a same property

The notion of concept is central to ontology. Ontology aims to understand the world and organize the objects which populate it. A concept is therefore a unit of knowledge about a plurality of things that verify the same property. Objects are structured into sets, called classes, which can overlap.

4. Theories of Concept

■ **Epistemological Principles**

■ **« Nature » (essence) of thing**: as I *conceive* things

- **essential characteristic**

■ **« Description» of thing**: as I *perceive* things

- **descriptive characteristic (attribute)**
- **relation** between objects

There are different ways of approaching the conceptualization of a domain. One may focus on the nature of things as they are conceived, that is, their essential characteristics. A characteristic is essential if removed from the thing, the thing is no longer what it is. For example, /with a backrest/ is an essential characteristic of the concept of chair. One can be interested in the description of things as they are perceived, that is, their descriptive characteristics, also called attributes, such as colour. One can also consider relationships between objects. These are all epistemological principles that can be combined.

4. Theories of Concept

■ **Knowledge about a plurality of things**

■ **Concept**: unique combination of essential characteristics

<Chair> ::= /for one person/ + /with feet/ + /with back/ + /without arms/

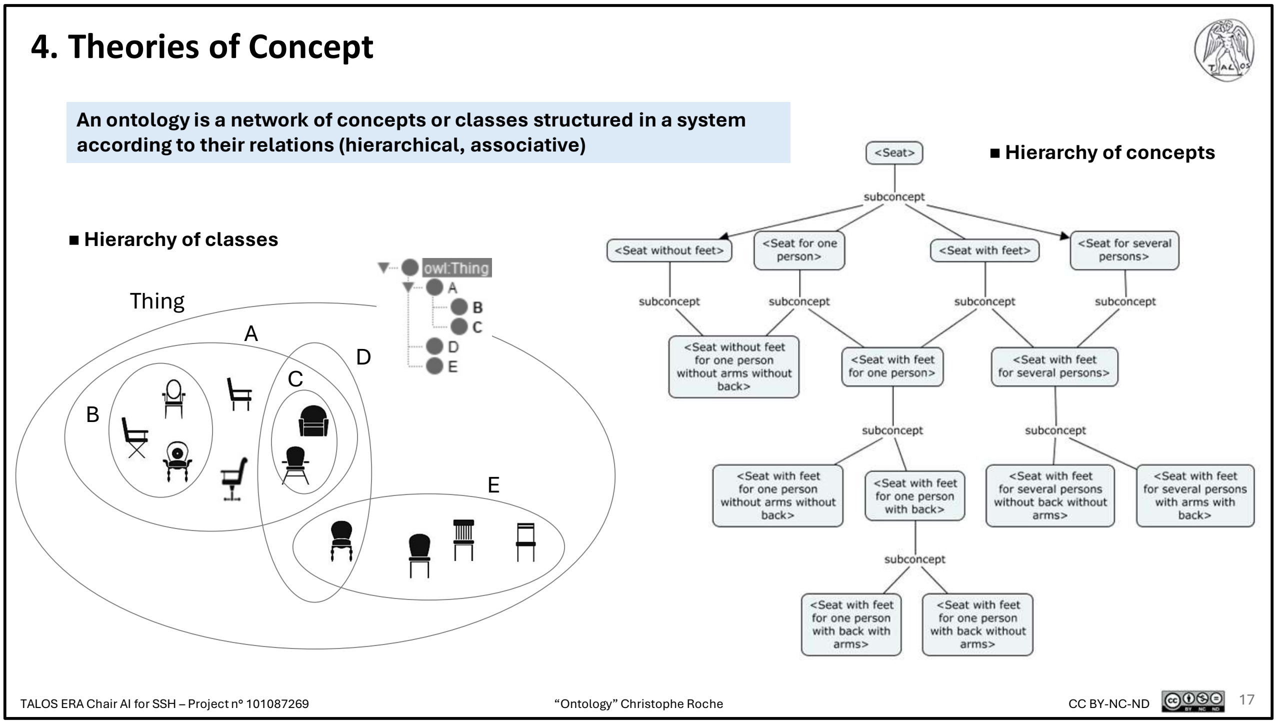
■ **Class**: set of objects verifying a same property

Parisian ::= { x / Person(x) ∧ livesIn (x, Paris) }

Things can be defined not according to their "nature" but through their relationships

A concept can thus be defined as a combination of essential characteristics, emphasizing the nature of things. For example, the concept of a chair is defined by the essential characteristics: /for a person/, /with legs/, /with backrest/, /without arms/.

The notion of class, on the other hand, focuses on individuals that verify the same property such as a relationship to a specific object. For example, the class of Parisians groups together all the objects that are instances of the Person concept living in Paris, it means linked by the relation *lives-in* to the object Paris, which is an instance of the City concept. In this way, it is possible to define a class according to the relationships between its elements, and not according to their nature.



An ontology is a network of concepts and/or classes structured into a system according to their hierarchical and associative relationships. Classes are hierarchically organized by set inclusion—for instance, the class of chairs is a subclass of the class of seats. Concepts are structured according to a relationship of inheritance of the essential characteristics that define them.

5. Representation Languages

- **The expression of knowledge is limited to the well-formed formulas of the formal language**

- **The formal languages are not equivalent**

  The Sapir-Whorf's hypothesis is true for all languages

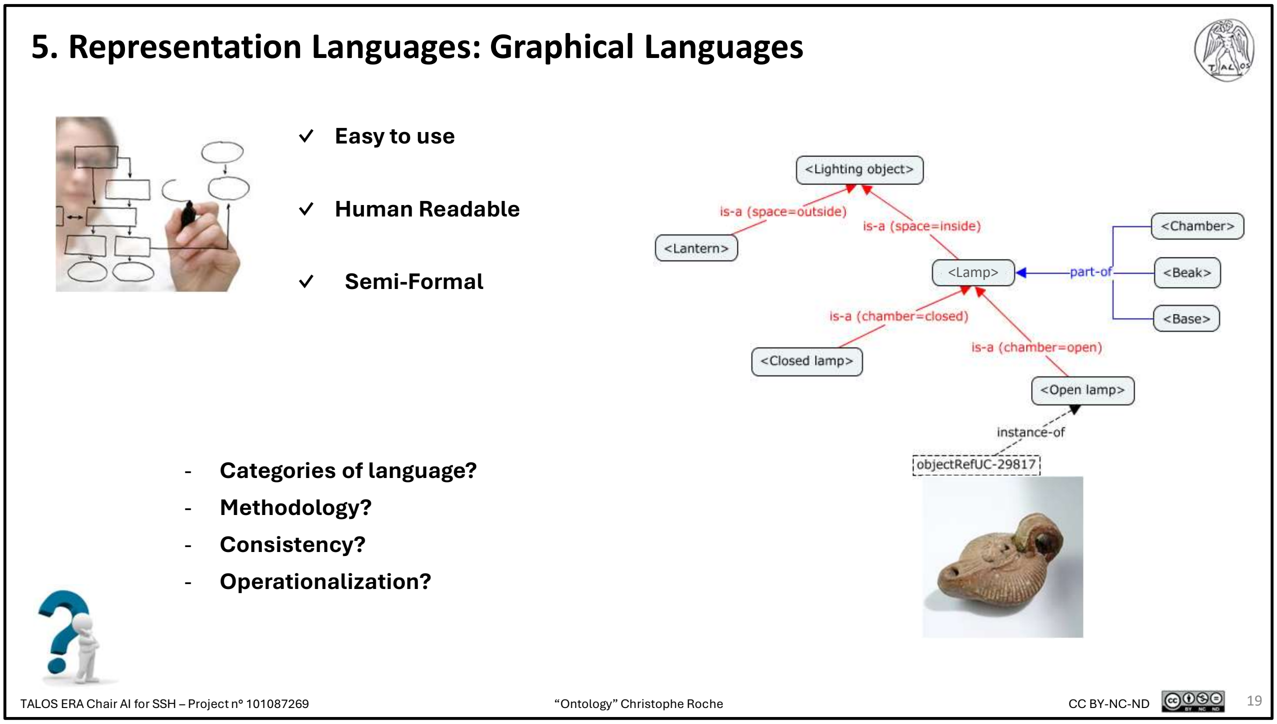
  - ✓ Domain of application
  - ✓ Power of expression (categories of thought supported by the categories of the language)
  - ✓ Logical Properties
  - ✓ Operationalization

TALOS ERA Chair AI for SSH – Project n° 101087269 "Ontology" Christophe Roche 18

The choice of the representation language is crucial. It depends, above all, on the field of application, which may require a specific approach and therefore a specific language.

It is important to remember that the representation language, and therefore the epistemological principles it conveys, shape our vision of the world: *we only see what we can express*. The result will not be the same depending on whether the modeling is done with classes or concepts, both of which can be combined.

Choosing a language based on logical properties is also essential, as it allows verification of certain properties, such as the consistency of the ontology (it means the absence of contradictions).



Graphical representation languages, which are easy to understand and use, are particularly helpful at the beginning of a project to identify the main concepts and their relationships. They can also be useful at the end of a project to illustrate parts of the ontology. However, it is difficult to build a complete ontology using these languages. Moreover, they are semi-formal and do not always incorporate logical properties, such as preventing the creation of loops for hierarchical relationships.

5. Representation Languages: AI Languages

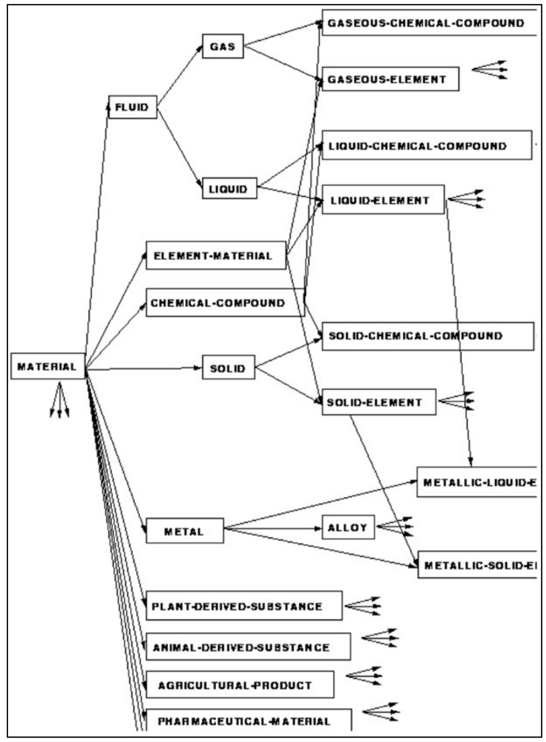
Schema (Frame) - Minsky

Define the **object structure**

A **class** is defined as a set of **slots** with **values**

```
(defun-class lamp
    (is-a 'lighting-object)
    (space 'inside)
    (has-part 'chamber 'beak 'base))
```

Clear, powerful, readable both by human and computer

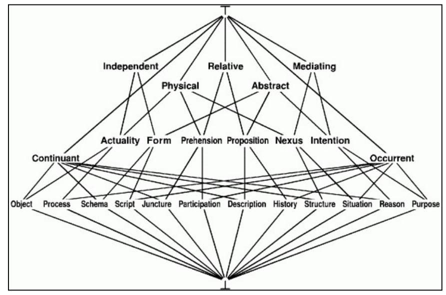


AI-based languages, such as schemas or frames, offer many advantages. They are understandable by both humans and machines. They are based on the notion of class, defined as a set of attribute–value pairs.

5. Representation Languages: Logical Languages

Syntax and Semantics:

- ✓ Clear
- ✓ Precise
- ✓ Formally specified



A concept (category) is an unary predicate.

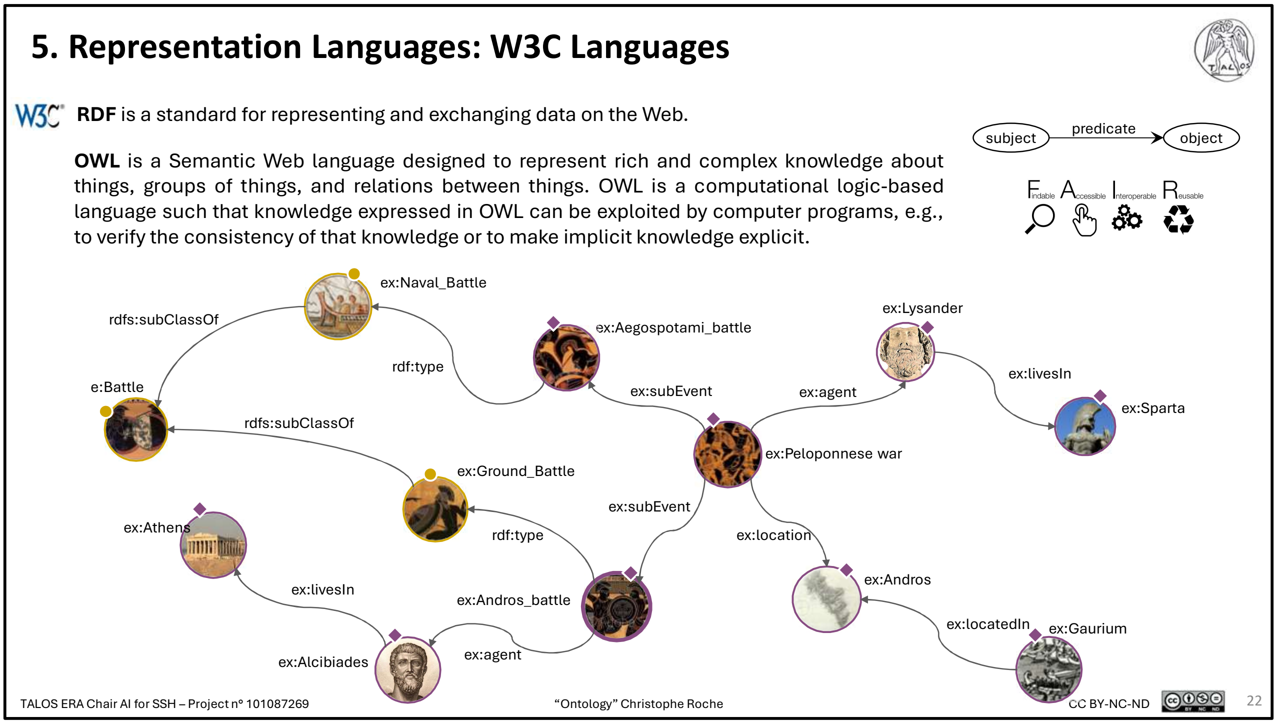
Form(x) ::= Independant(x) $\wedge$ Abstract(x)

➡ **Properties of Axiomatic System**

Definitions are:

- Objective
- Coherent
- Precise
- Consensual
- Reusable
- Sharable
- Readable (for an expert)

Logical languages, based on precise syntax and semantics, offer mechanisms for validation and inference. They have become indispensable. Most environments provide user interfaces to help write logical expressions.



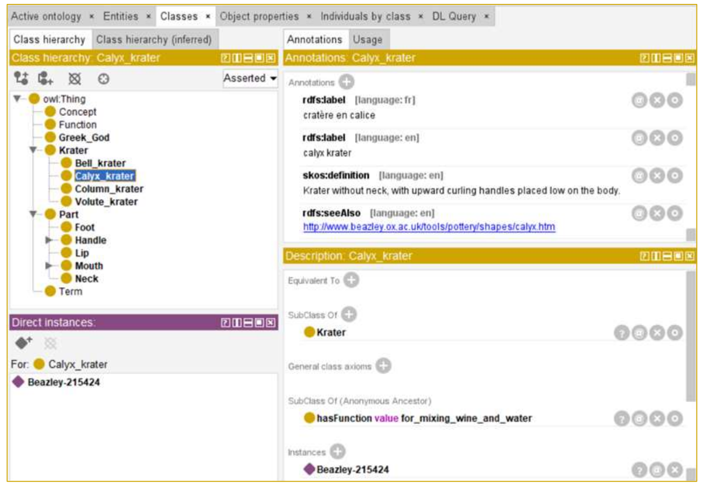
Ontologies are knowledge graphs. Using W3C standards built around RDF, such as OWL, enables ontologies to be represented as RDF graphs, making it easier to share and link them for the Semantic Web and Linked and Open Data. The goal is to offer data sets that respect the principles of FAIR Data, so that data is ‘easy to find, accessible, interoperable and reusable’.

6. Environments: Protégé

https://protege.stanford.edu/

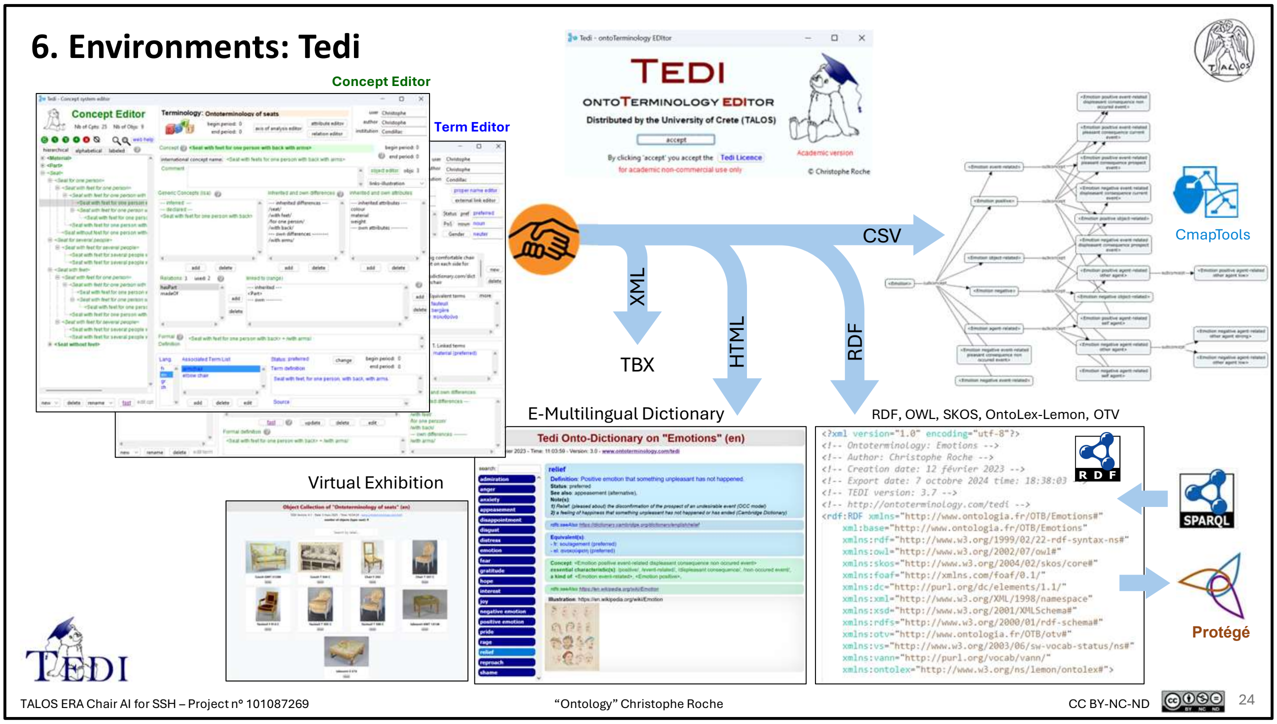
protégé

- ✓ Free environment
- ✓ Large Community
- ✓ Definition based on relations between objects (Class)
- ✓ Description Logic Reasoners
- ✓ W3C Standards compliant



- How to represent essential characteristics (Concept)?
- How to represent the linguistic dimension?
- How to take into account the way of thinking of Experts?

Let's conclude this MOOC by presenting two environments. Protégé is the most widely used ontology-building environment. It is free open-source software developed by Stanford University. It is supported by a large user community. The notion of class, based on relationships between individuals, replaces that of concept. Protégé relies on first-order logic, allowing reasoning engines to check ontology consistency. It supports W3C formats such as OWL, enabling ontologies to be exported as RDF knowledge graphs. However, essential characteristics, the linguistic dimension (which requires explicit term representation), and taking into account the way experts think remain problems that are difficult to tackle in Protégé.



Tedi is an ontoterminology building environment distributed free of charge by the University of Crete as part of the European TALOS project for academic, research and teaching purposes only, to the exclusion of all commercial applications. An ontoterminology is a terminology whose conceptual system is a formal ontology. Concepts are defined by combinations of essential characteristics, following the terminology principles of the ISO standards (1087 and 704). Tedi integrates several methodologies for building ontologies and terminologies while verifying logical properties. Ontoterminologies are exported as RDF graphs including OWL, SKOS, and OntoLex-Lemon, as well as multilingual electronic dictionaries and object collections.