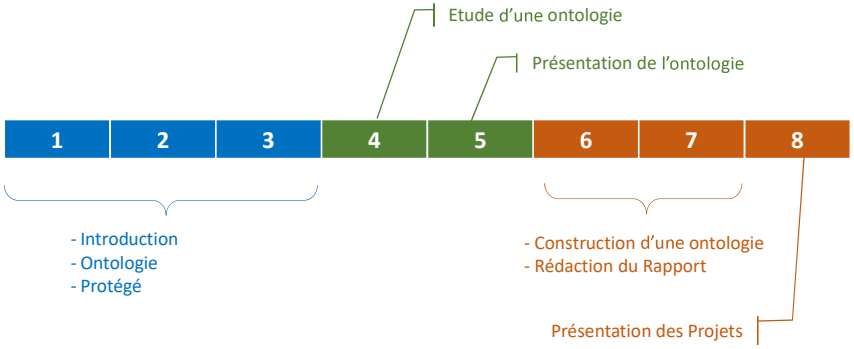
STIC Info – Master 2

Info 911: Knowledge Engineering

Web des Données (Données Ouvertes et Liées)

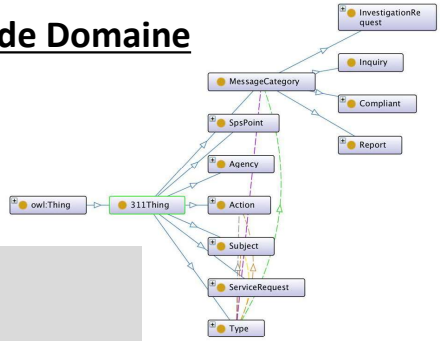
Info 911

Ingénierie des Connaissances

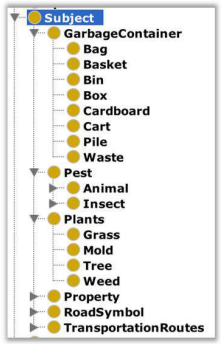


Construction d'une Ontologie de Domaine à l'aide de Protégé

Protégé

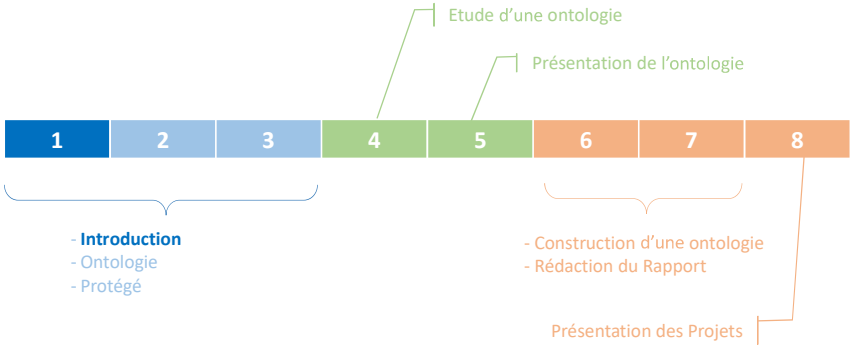


1) **Présentation du domaine choisi**
2) **Questions de Compétences (>= 5)**
3) **Vocabulaires utilisés**
4) **Description de l'ontologie de domaine :**
   1) Hiérarchie des concepts
   2) Définition des concepts (restrictions de propriétés)
   3) Peuplement de l'ontologie (s'il y a lieu)
   4) Vérification : Raisonneur(s)
   5) Définition de règles (s'il y a lieu)
5) **Evaluation**
6) **Requêtes SPARQL (Questions de Compétences)**
7) **Rapport**



Info 911

Ingénierie des Connaissances



Introduction: the Web

The **World Wide Web** (WWW), commonly known as the Web, is **an information system** where documents and other web resources are identified by Uniform Resource Locators (**URLs**, such as https://www.example.com/), which may be interlinked by hypertext, and are **accessible over the Internet**. The resources of the WWW may be accessed by users by a software application called a web browser.

Web 1.0 was an early stage of the conceptual evolution of the World Wide Web, centered around a top-down approach to the use of the web and its user interface. […] users could **only view webpages** but not contribute to the content of the webpages.

A **Web 2.0** site may allow users to **interact and collaborate** with each other in a social media dialogue as creators of user-generated content in a virtual community, in contrast to websites where people are limited to the passive viewing of content. Examples of Web 2.0 include **social networking** sites, blogs, wikis, folksonomies, video sharing sites, hosted services, web applications, and mashups

Definitions of **Web 3.0** vary greatly. Some believe its most important features are the **Semantic Web** and personalization.

Introduction: History

- Tim Berners-Lee, Consultant at CERN

- *1980 : program of note taking with creation of links (hypertext)*
  mainframes
- *1989 : hypertext for the management of information*
  links extend (reference) to other web addresses
  data and links are decentralized on the internet => (error 404)

- End of 1990 : 1st server and 1er navigator named World Wide Web tested by means of an internet connection (NeXT machine)

- 1993 : web technology from CERN becomes free of charge and free of copyright

- 1994 : Creation of the World Wide Web Consortium (W3C)

- *Founding members : INRIA, MIT, Keio University*
- *publication of web standards*

Web Architecture

✓ **3 fundamental notions** (original web)

1. **URI** (Universal Resource Identifier)
   - Unique identifier permitting to reference a resource on a network

   URL (Universal Resource Locator)
   - Point of access to a representation of a resource on the Web

2. **HTTP Protocol** (HyperText Transfer Protocol)
   - Protocol of client-server communication developed for the Web

3. **HTML** (HyperText Markup Language)
   - Markup language of hypertexts dedicated to the representation and the interchange of web pages.
   - HTML is derived from SGML (Standard Generalized Markup Language), language of description with tags and an ISO standard since 1986 (ISO 8879:1986)

Web Architecture

✓ **Textual Interchange format of structured data**

XML (Extensible Markup Language)

'extensible markup language', computer langage informatique of generic markup derived from SGML; it is an interchange format of structured content (tree structure) between heterogeneous information systems (interoperability).

⇒ **Web of structured data**

Link the data ?

Web Architecture

✓ **RDF (Resource Description Framework)**

RDF is a linked data interchange format for the Web.
Data are represented in the form of graphs (set of triplets 'subject-predicate-object').

✓ **SPARQL (SPARQL Protocol And RDF Query Language)**

Query language permitting to search, add, modify and delete RDF data on the Internet.

⇒ **Web of data (open) linked**

Data management ?

⇒ Metadata (structure, meaning…)

⇒ **Semantic Web**

Web Architecture

✓ **RDFS (RDF Schema)**

Extensible language of knowledge representation :

- *Classes*
- *Sub-classes*
- *Properties (domain et co-domain)*

✓ **OWL (Ontology Web Language)**

- langage of representation of ontologies on the basis of RDF and RDFS
- Based on description logics

✓ **RIF (Rule Interchange Format)**

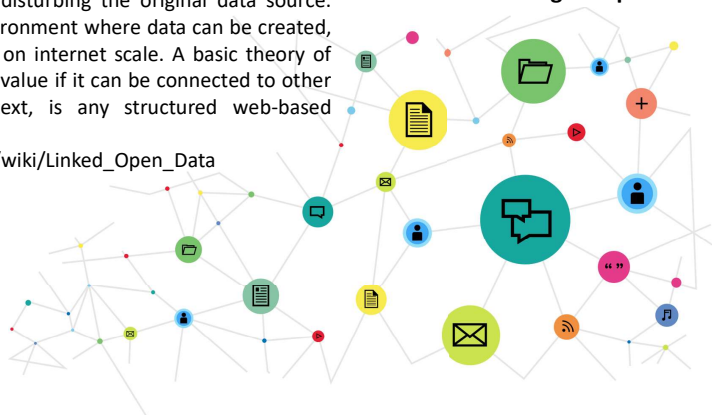
Interchange Format of inference rules (If condition Then action)

⇒ **Semantic Web**

Linked and Open Data

"LOD is often thought of as a virtual data cloud where anyone can access any data they are authorized to see and may also add to any data without disturbing the original data source. This provides an open environment where data can be created, connected and consumed on internet scale. A basic theory of LOD is that data has more value if it can be connected to other data. Data, in this context, is any structured web-based information."
https://www.w3.org/egov/wiki/Linked_Open_Data

Knowledge Graph



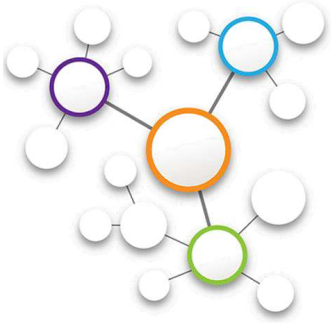
"In computing, linked data (often capitalized as Linked Data) is structured data which is interlinked with other data so it becomes more useful through semantic queries. It builds upon standard Web technologies such as HTTP, RDF and URIs, but rather than using them to serve web pages only for human readers, it extends them to share information in a way that can be read automatically by computers. Part of the vision of linked data is for the Internet to become a global database."
https://en.wikipedia.org/wiki/Linked_data

Web des données

https://fr.wikipedia.org/wiki/Web_des_donn%C3%A9es

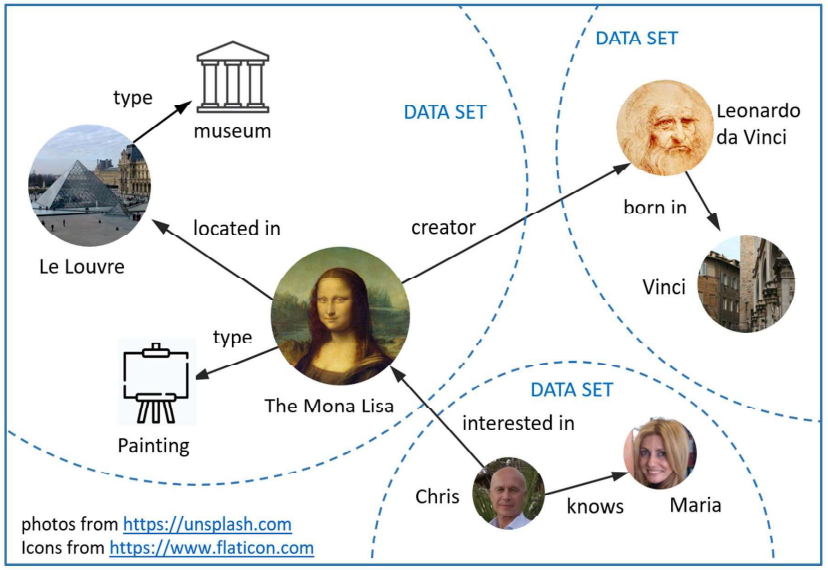
Le **Web des données** (*linked data*, en anglais) est une initiative du W3C (Consortium World Wide Web) visant à favoriser la publication de données structurées sur le Web, non pas sous la forme de silos de données isolés les uns des autres, mais en les reliant entre elles pour constituer un réseau global d'informations.

Il s'appuie sur les standards du Web, tels que HTTP et URI - mais plutôt qu'utiliser ces standards uniquement pour faciliter la navigation par les êtres humains, le Web des données les étend pour partager l'information également entre machines. Cela permet d'interroger automatiquement les données, quels que soient leurs lieux de stockage, et sans avoir à les dupliquer[1].



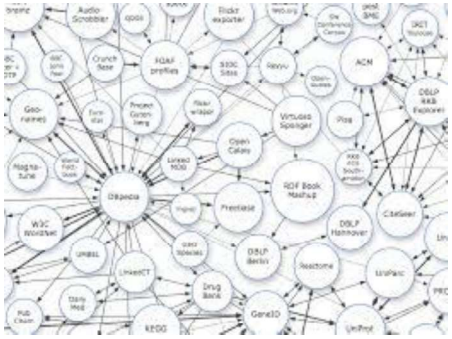
Knowledge Graph

http://o4dh.com/dh-ch-en



Web Architecture

Web of Open and Linked Data (RDF)



Knowledge Graph (RDF)

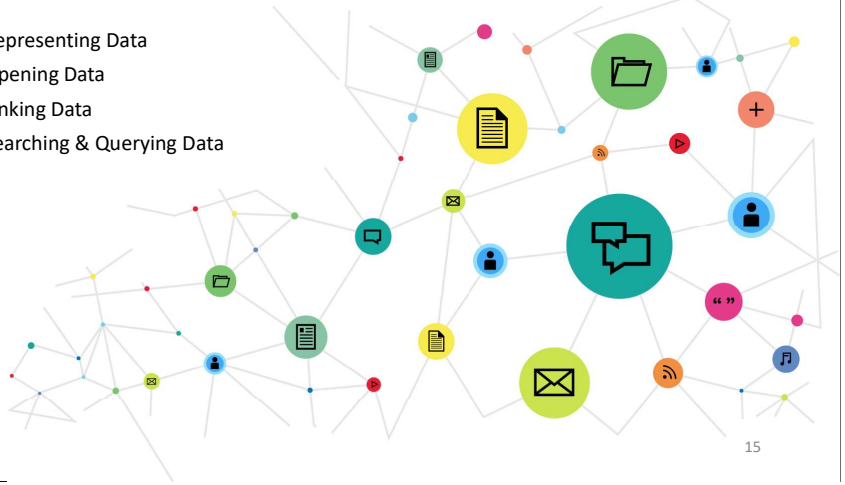
- **No relational schema**
- **Anybody can say anything on any subject (URI)**

Knowledge Graph

A **knowledge graph** is a **special kind of database** which stores **knowledge in a machine-readable form** and provides a means for information to be collected, organised, shared, searched and utilised.

Main Issues:

- Representing Data
- Opening Data
- Linking Data
- Searching & Querying Data



Resource Description Framework (RDF)

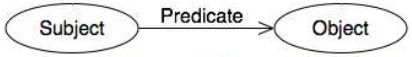
✓ A Graph of Data

Resource Description Framework (RDF) is a model of graphs dedicated to describing in a formal fashion the Web resources and their metadata, in such a way so as to permit the automatic treatment of such descriptions.

RDF is a standard model for data interchange on the Web. RDF has features that facilitate **data merging** even if the underlying schemas differ, and it specifically supports the evolution of schemas over time without requiring all the data consumers to be changed.

RDF extends the linking structure of the Web to use **URIs to name the relationship between things as well as the two ends of the link** (this is usually referred to as a "triple"). Using this simple model, it allows structured and semi-structured **data to be mixed, exposed, and shared** across different applications.

This linking structure forms a **directed, labeled graph**, where the edges represent the named link between two resources, represented by the graph nodes. This graph view is the easiest possible mental model for RDF and is often used in easy-to-understand visual explanations



An RDF graph with two nodes (Subject and Object) and a triple connecting them (Predicate)

RDF Vocabulary

Vocabulary[1] RDF :

« **An RDF vocabulary is a collection of IRIs intended for use in RDF graphs**. »

An RDF triple consists of three components:

- the subject, which is an IRI or a blank node
- **the predicate, which is an IRI**
- the object, which is an IRI, a literal or a blank node

IRI : Internationalized Resource Identifier

The Internationalized Resource Identifier (IRI) – is an internet protocol standard which extends the ASCII characters subset of the Uniform Resource Identifier (URI) protocol. While URIs are limited to a subset of the ASCII character set, IRIs may contain characters from the Universal Character Set (Unicode/ISO 10646), including Chinese, Japanese kanji, Korean, and Cyrillic characters.

[1] : *Ontology = Vocabulary whose terms are formally defined*

RDFS Vocabulary

Vocabulary RDF Schema 1.1 (www.w3.org/TR/rdf-schema/)

RDF Schema provides a data-modelling vocabulary for RDF data.
RDF Schema is an extension of the basic RDF vocabulary.

RDF Schema (**Resource Description Framework Schema**, variously abbreviated as **RDFS**, **RDF(S)**, **RDF-S**, or **RDF/S**) is a set of classes with certain properties using the RDF extensible knowledge representation data model, providing basic elements for the description of ontologies, otherwise called RDF vocabularies, intended to structure RDF resources. These resources can be saved in a triplestore to reach them with the query language SPARQL.

rdf:Property

rdfs:domain

rdfs:range

RDF & RDFS VOCABULARIES

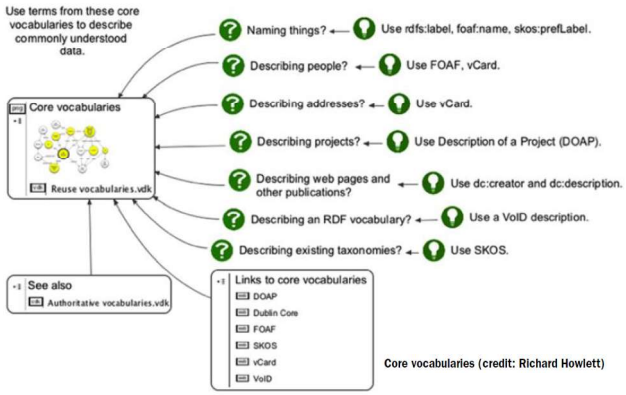
Property name	comment	domain	range
rdf:type	The subject is an instance of a class.	rdfs:Resource	rdfs:Class
rdfs:subClassOf	The subject is a subclass of a class.	rdfs:Class	rdfs:Class
rdfs:subPropertyOf	The subject is a subproperty of a property.	rdf:Property	rdf:Property
rdfs:domain	A domain of the subject property.	rdf:Property	rdfs:Class
rdfs:range	A range of the subject property.	rdf:Property	rdfs:Class
rdfs:label	A human-readable name for the subject.	rdfs:Resource	rdfs:Literal
rdfs:comment	A description of the subject resource.	rdfs:Resource	rdfs:Literal
rdfs:member	A member of the subject resource.	rdfs:Resource	rdfs:Resource
rdf:first	The first item in the subject RDF list.	rdf:List	rdfs:Resource
rdf:rest	The rest of the subject RDF list after the first item.	rdf:List	rdf:List
rdfs:seeAlso	Further information about the subject resource.	rdfs:Resource	rdfs:Resource
rdfs:isDefinedBy	The definition of the subject resource.	rdfs:Resource	rdfs:Resource
rdf:value	Idiomatic property used for structured values.	rdfs:Resource	rdfs:Resource
rdf:subject	The subject of the subject RDF statement.	rdf:Statement	rdfs:Resource
rdf:predicate	The predicate of the subject RDF statement.	rdf:Statement	rdfs:Resource
rdf:object	The object of the subject RDF statement.	rdf:Statement	rdfs:Resource

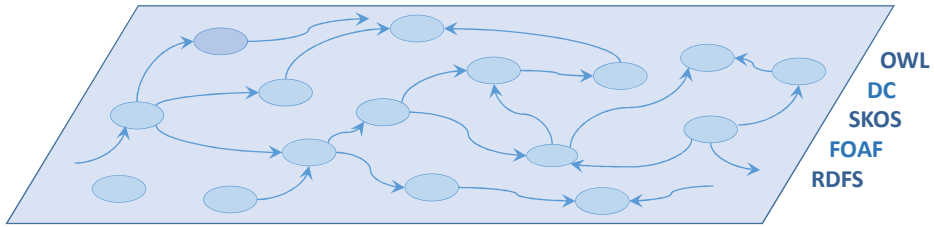
Class name	comment
rdfs:Resource	The class resource, everything.
rdfs:Literal	The class of literal values, e.g. textual strings and integers.
rdf:langString	The class of language-tagged string literal values.
rdf:HTML	The class of HTML literal values.
rdf:XMLLiteral	The class of XML literal values.
rdfs:Class	The class of classes.
rdf:Property	The class of RDF properties.
rdfs:Datatype	The class of RDF datatypes.
rdf:Statement	The class of RDF statements.
rdf:Bag	The class of unordered containers.
rdf:Seq	The class of ordered containers.
rdf:Alt	The class of containers of alternatives.
rdfs:Container	The class of RDF containers.
rdfs:ContainerMembershipProperty	The class of container membership properties, rdf:_1, rdf:_2, ..., all of which are sub-properties of 'member'.
rdf:List	The class of RDF Lists.

RDF Graph

RDF (family of RDF standards) is a framework for representing information in the web:

- **RDF**
- **RDFS**
- **FOAF**
- **DC**
- **SKOS**
- **OWL**
- **Etc.**



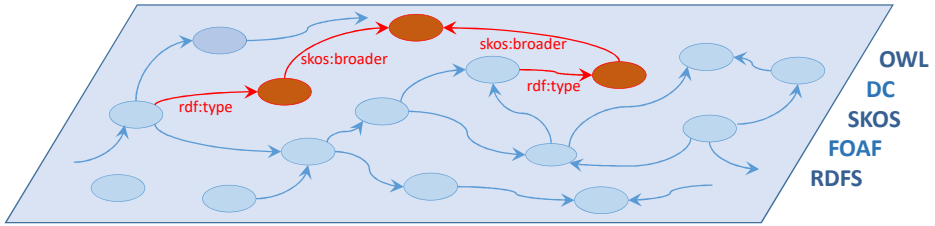


Knowledge Graph

How to structure data?

Semantic Level

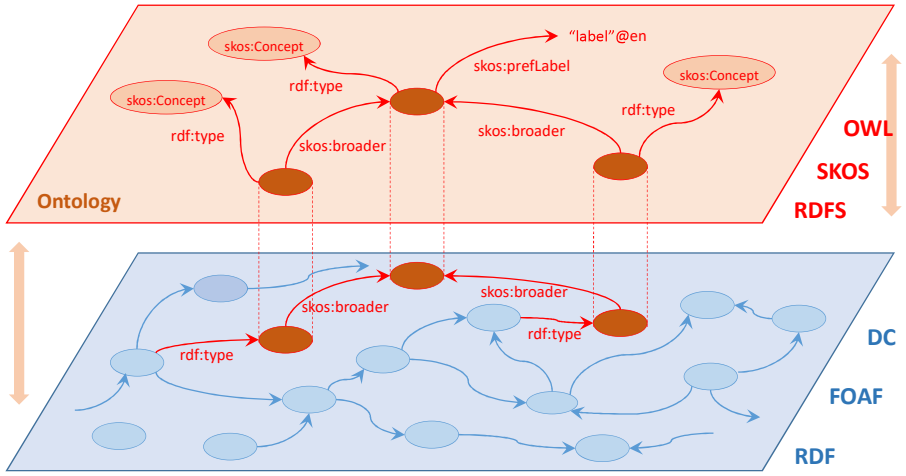
- Organization of Resources into classes
- Defining properties for predicates (semantics)



Knowledge Graph

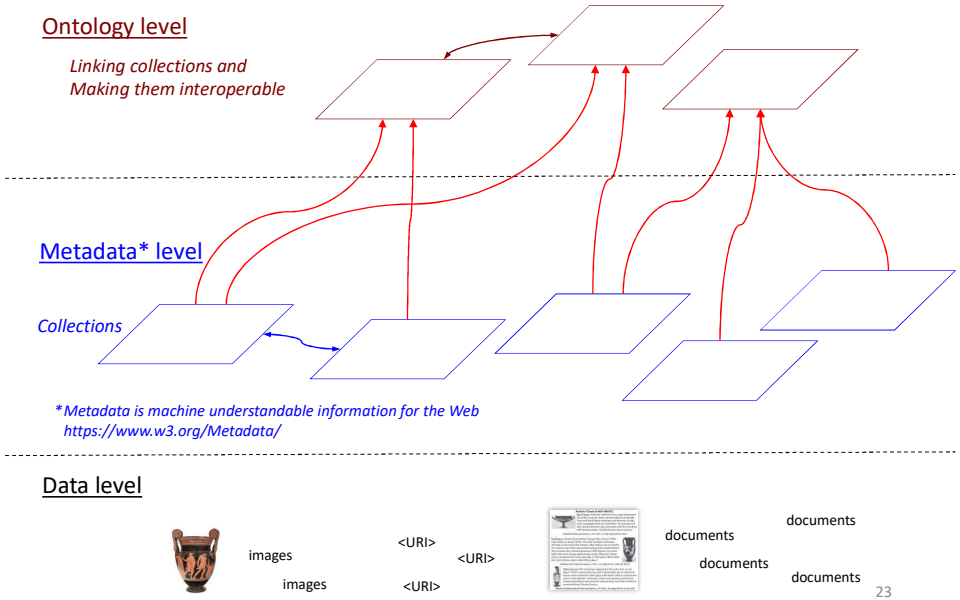
How to structure data?

Semantic Level: Ontology



Knowledge Graph

Data, Metadata & Ontology Levels



Semantic Web

https://www.w3.org/standards/semanticweb/

In addition to the classic "Web of documents" W3C is helping to build a technology stack to support a "**Web of data**," the sort of data you find in databases.

The ultimate goal of the Web of data is to enable **computers** to do more useful work and to develop systems that can support trusted interactions over the network.

The term "Semantic Web" refers to W3C's vision of the Web of linked data. Semantic Web technologies enable people to create data stores on the Web, build **vocabularies**, and write rules for handling data.

Linked data are empowered by technologies such as RDF, SPARQL, OWL, and SKOS.

Semantic Web

https://en.wikipedia.org/wiki/Semantic_Web

"**The ultimate goal of Semantic Web is to make the machine to understand the Internet data**. To enable the encoding of semantics with the data, well-known technologies are RDF(Resource Description Framework) and OWL(Web Ontology Language). These technologies formally represent the meaning involved in information".

"The Semantic Web is an extension of the World Wide Web through standards by the World Wide Web Consortium (W3C). The standards promote common data formats and exchange protocols on the Web, most fundamentally the Resource Description Framework (RDF). According to the W3C, "**The Semantic Web provides a common framework that allows data to be shared and reused across application, enterprise, and community boundaries**". The Semantic Web is therefore regarded as an integrator across different content, information applications and systems."

"The term was coined by Tim Berners-Lee for a web of data (or data web) that can be processed by machines—that is, one in which much of the meaning is machine-readable."

Semantic Web

Linked Data

The Semantic Web is a Web of data — of dates and titles and part numbers and chemical properties and any other data one might conceive of. **RDF** provides the foundation for publishing and linking your data. Various technologies allow you to embed data in documents (RDFa, GRDDL) or expose what you have in SQL databases, or make it available as RDF files.

Vocabulary

At times it may be important or valuable to organize data. Using **OWL** (to build vocabularies, or "**ontologies**") and **SKOS** (for designing knowledge organization systems) it is possible to enrich data with additional meaning, which allows more people (and more machines) to do more with the data.

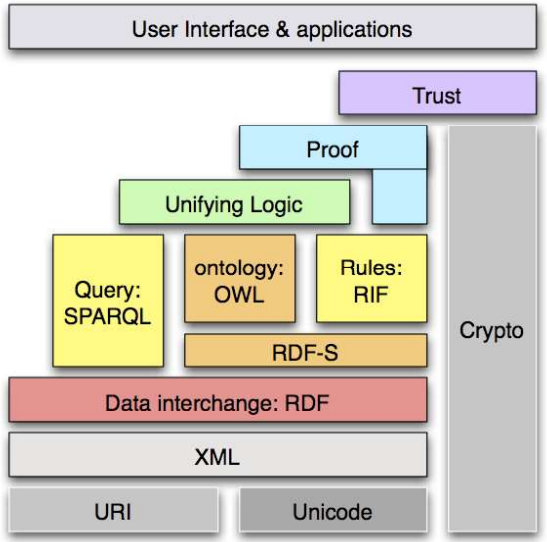
Query

Query languages go hand-in-hand with databases. If the Semantic Web is viewed as a global database, then it is easy to understand why one would need a query language for that data. **SPARQL** is the query language for the Semantic Web.

Inference

Near the top of the Semantic Web stack one finds inference — reasoning over data through rules. W3C work on rules, primarily through **RIF** and OWL, is focused on translating between rule languages and exchanging rules among different systems.

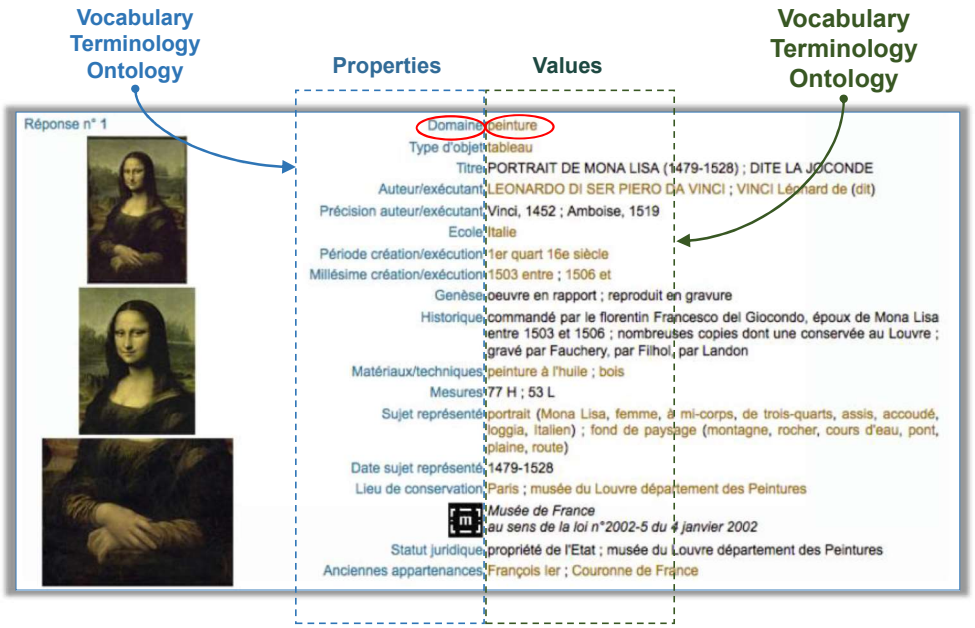
Web Architecture



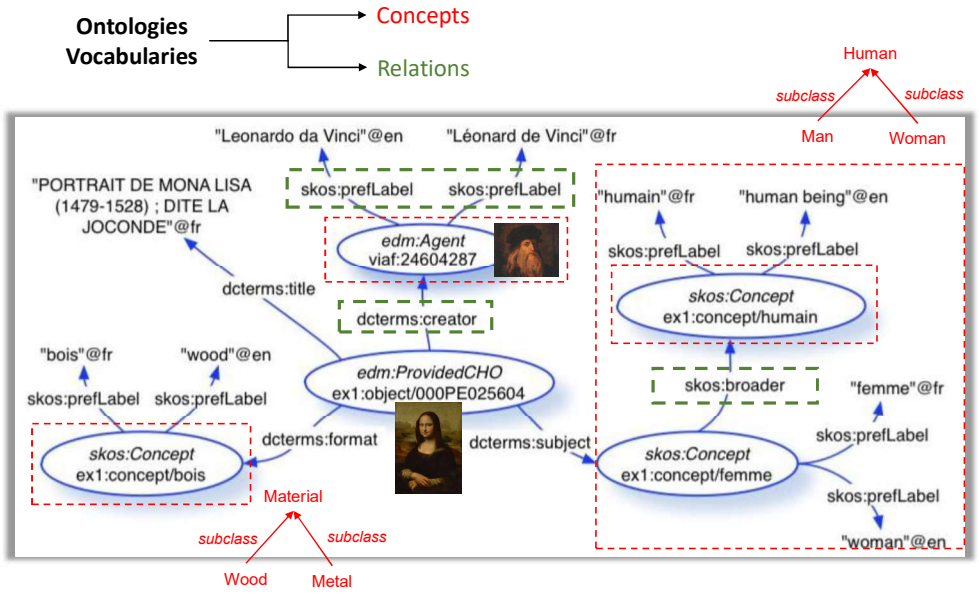
Semantic web standards (Semantic Web Layer)

Vocabularies – Ontologies – Terminologies?

Vocabularies – Ontologies – Terminologies?



Vocabularies – Ontologies – Terminologies?



Vocabularies – Ontologies – Terminologies?

https://www.w3.org/standards/semanticweb/ontology

What is a Vocabulary?

On the Semantic Web, **vocabularies** define the **concepts** and **relationships** (also referred to as "terms") used to describe and represent an area of concern.

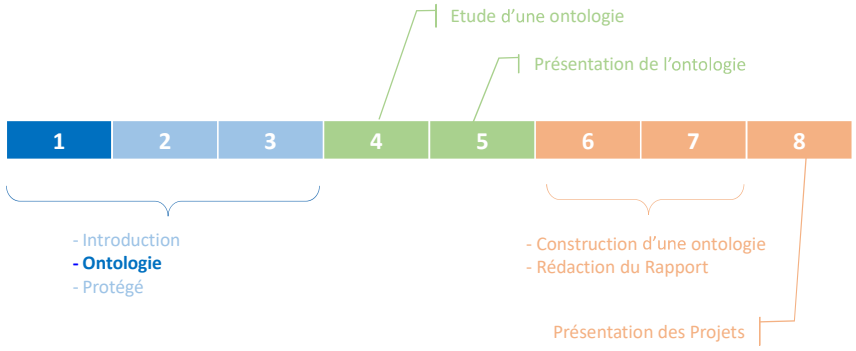
Vocabularies are used to classify the terms that can be used in a particular application, characterize possible relationships, and define possible constraints on using those terms.

In practice, vocabularies can be very complex (with several thousands of terms) or very simple (describing one or two concepts only).

There is no clear division between what is referred to as "vocabularies" and "ontologies". The trend is to use the word "ontology" for more complex, and possibly quite formal collection of terms, whereas "vocabulary" is used when such strict formalism is not necessarily used or only in a very loose sense. Vocabularies are the basic building blocks for inference techniques on the Semantic Web.

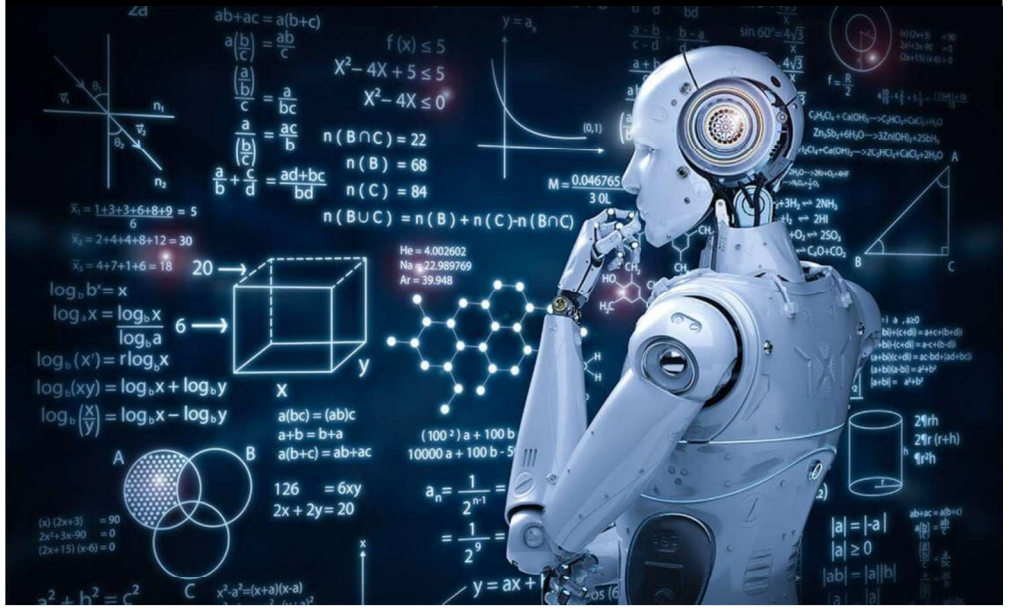
Info 911

Ingénierie des Connaissances

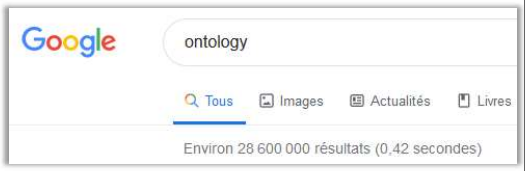


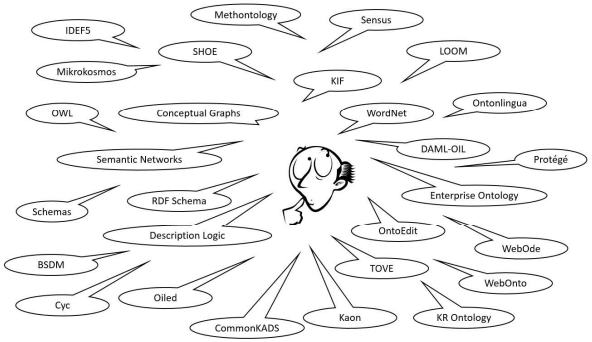
Ontology

How to represent the world in a computer-readable format?



Ontology 本体论





Ontology 本体论

Ontos + **Logos**

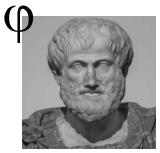
Ontology 本体论

1 “the branch of metaphysics dealing with the nature of being”

形而上学的分支处理存在的本质

OXFORD UNIVERSITY PRESS

➥ ***Epistemological Principles***



哲学

Understanding the “world”

了解“世界”

2 “the branch of knowledge engineering dealing with conceptualization in a computer readable form”

概念化的规范

➥ ***Formalization & Representation***

正式化与表征

知识工程

Knowledge Engineering

Ontology 本体论

- **There is today (more or less) an agreement on the 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). »

    "Ontologies: Principles, Methods and Applications" M.Ushold & M.Gruninger. Knowledge Engineering Review, Vol.11, n°2, June1996

- **There is today an agreement on the objective:**
  - **Communication and Knowledge Sharing between Human and/or Software Agents**

    « The main purpose of an ontology is to enable communication between computer systems in a way that is independent of the individual system technologies, information architectures and application domain. »
    www.ontology.org

Ontology 本体论

"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

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

"On the Semantic Web, vocabularies define the concepts and relationships (also referred to as "terms") used to describe and represent an area of concern. Vocabularies are used to classify the terms that can be used in a particular application, characterize possible relationships, and define possible constraints on using those terms."

OWL 2 Web Ontology Language
Document Overview (Second Edition)

W3C Recommendation 11 December 2012

https://www.w3.org/TR/owl2-overview/#

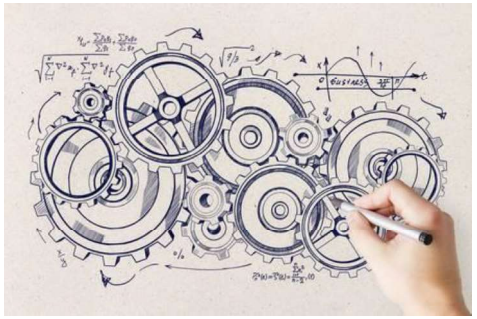
Ontologies are formalized vocabularies of terms, often covering a specific domain and shared by a community of users. They specify the definitions of terms by describing their relationships with other terms in the ontology.

Ontology 本体论

"An ontology is a shared description of **concepts** and **relationships** of a domain expressed in a formal and computer readable language" C. Roche

本体论是以正式（计算机可读）语言表达的领域的概念和关系的共享描述

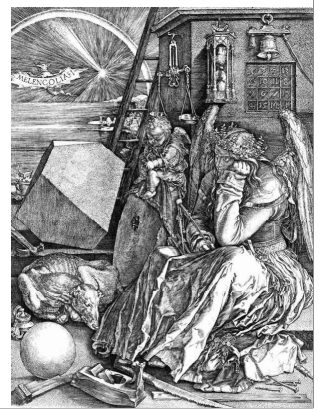
3. CONCEPTUAL SYSTEM



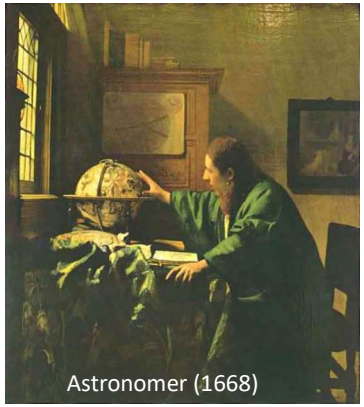
What do we need?

✓ Organize the world

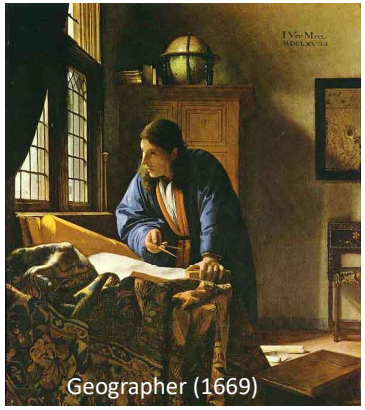
✓ Represent the world



Understanding & Representing



Astronomer (1668)

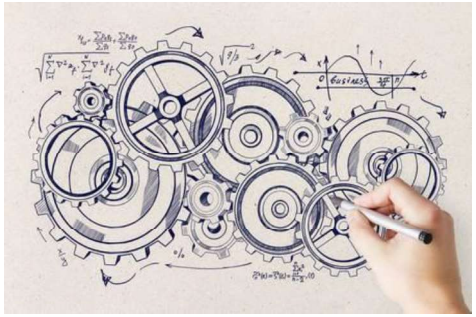


Geographer (1669)

Epistemological Principles ⟷ **Language of Representation**

Johannes Vermeer (1632-1675)

3. CONCEPTUAL SYSTEM



2.3.1 Concept

2.3.2 Representation Languages

Concept

- What is a concept?
- What is it for?
- Which language to define a concept?

Concept

- An abstract idea; a general notion. (Oxford Dictionary)
- An idea or abstract principle (Collins)
- Concepts are mental representations, abstract objects or abilities that make up the fundamental building blocks of thoughts and beliefs. They play an important role in all aspects of cognition (Wikipedia)
- Intellectual representation targeting reality into abstract and general determinations and not in its concrete singularity. (Dictionnaire de philosophie. Armand Colin)

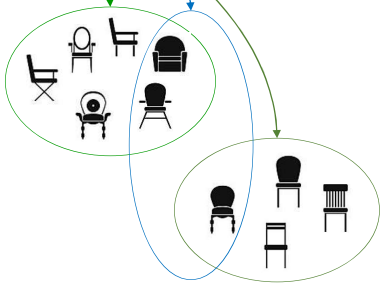
Unit of knowledge (识单位) created by a unique combination of characteristics [ISO 1087-1]

概念 concept
通过对特征(3.2.4)的独特组合而形成的知识单元。 注：概念不受语种限制，但受社会或文化背景的影响。

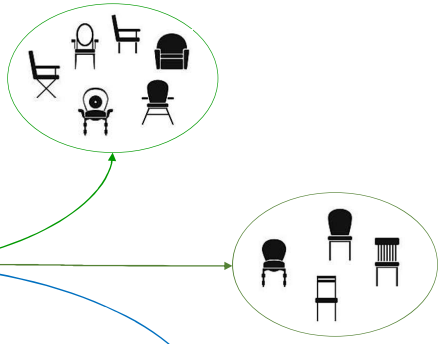
Concept 概念

Unit of knowledge
知识单位
about a plurality of things

✓ Understanding the World
✓ Organizing the objects



Concept 概念



extensional **Sets** of objects

intensional verifying a common **property** ?

Theories of Concept

Set of objects verifying a common **property**

As you perceive them

- a long thin bar or stick
- made of steel
- a 3-centimeter-long bar
- with thread

Description

As you conceive them

- connecting rod
- used to join together female-threaded components
- double-end threaded

"double-end threaded studs"

Definition

Theories of Concept

Set of objects verifying a common **property**

Through their relationships

Through their parts

Description

Epistemological Principles

As they are **perceived**

describe objects

As they are **thought**

define concepts

Theories of Concept

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

- **"essential" characteristic**

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

- **"descriptive" characteristic**

✓ **Relations between objects**: Things are defined not according to their "nature" but through relationships between them

- **relation**

Class

✓ **Set of objects verifying a same property**

There is no Terminology without a Concept Theory

Categories of Thought

- ✓ **Essential Characteristic**
- ✓ **Descriptive Characteristic**
- ✓ **Concept: combination of essential characteristics**
- ✓ **Class: set of objects verifying a same property**
- ✓ **Relation**

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

ISO Epistemological Principles

Concept (概念): **Unit of knowledge** (识单位) created by a unique combination of **characteristics** [ISO 1087-1]

> **概念 concept**
> 通过对特征(3.2.4)的独特组合而形成的知识单元。
> 注：概念不受语种限制，但受社会或文化背景的影响。

Characteristic : abstraction of a property of an object (3.1.1) or of a set of objects

> **特征 characteristic**
> 一个客体(3.1.1)或一组客体特性的抽象结果。
> 注：特征是用来描述概念(3.2.1)的。

Delimiting characteristic : essential characteristic used for distinguishing a concept from related concepts

> **区别特征 delimiting characteristic**
> 一个概念(3.2.1)同其他相关概念相区别的本质特征(3.2.6)。
> 注：可以用"靠背"这个特征(3.2.4)来区别"椅子"和"凳子"两个概念(3.2.1)。

Theories of Concept

ISO Unit of knowledge created by a unique combination of essential characteristics

- What is an essential characteristic?
- How to find essential characteristics?

If an essential characteristic is removed from an object, the object is no more what it is

☞ Looking for essential differences between objects

/without arms/

/with arms/

Theories of Concept

ISO Unit of knowledge created by a unique combination of essential characteristics

- What is an essential characteristic?
- How to find essential characteristics?
- How to combine them?

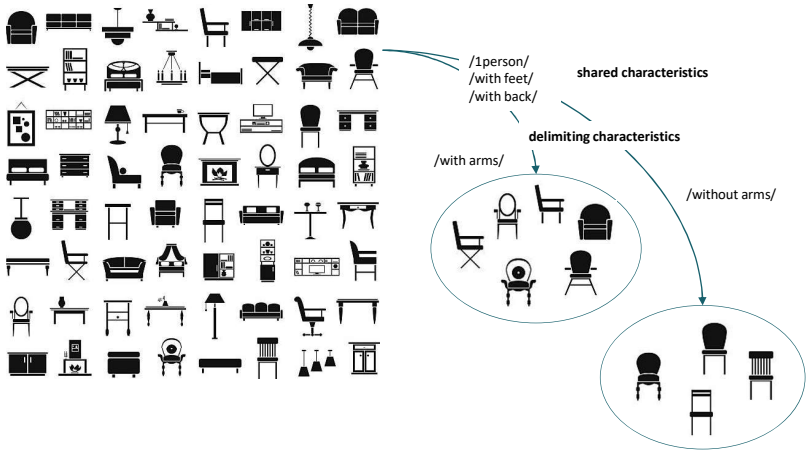
If an essential characteristic is removed from an object, the object is no more what it is

Looking for essential differences between objects

A concept is a set of essential characteristics stable enough to be named in a given natural language

Theories of Concept

Unit of knowledge created by
a unique combination of essential characteristics



Concept

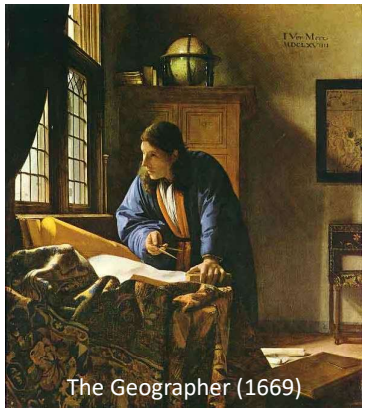
concept
unit of knowledge created by a unique combination of *characteristics*
[ISO 1087: 2019]

"A concept can be viewed as an idea or notion; a unit of thought" [SKOS]

"Resources may be divided into groups called classes" [RDFS]

Representing

There is no knowledge without language


The Geographer (1669)

Categories of Thought

versus

Categories of Language

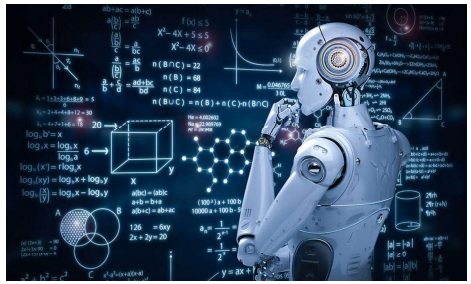
Languages

1 Theories of Concept

?

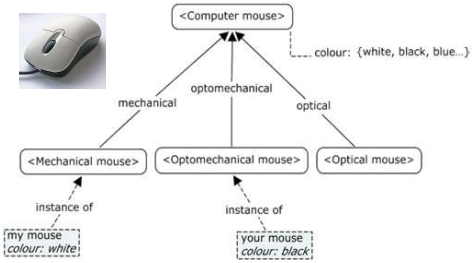
Can all the categories of thought be represented in *any* representation language equally well?

Computational → 2 Representation languages?



Representation Languages

There is no knowledge without language



- In which language?
- Are all languages equivalent?
- Why a formal language?
- Why a computational language?

Which Language?

✗ **The expression of knowledge is limited to the well formed formulas of the theory**

✗ **The formal languages are not equivalent**

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

- Power of expression : categories of thought supported by the categories of the language
- Logical Properties
- Operationalization

Which Formal Language?

✓ **Graphical Notation**

✓ **Artificial Intelligence**

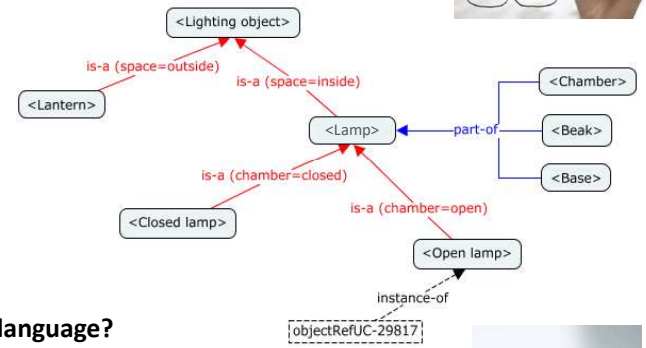
✓ **Logic**

Graphical Notation

✓ **Easy to use**

✓ **Human Readable**

✓ **Semi-Formal**



- **Categories of language?**
- **Methodology?**
- **Coherency?**
- **Operationalization?**

Artificial Intelligence

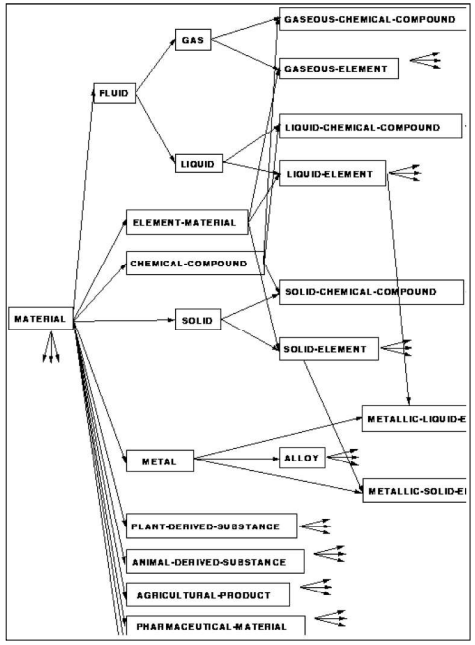
Schema (Frame) - Minsky

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))

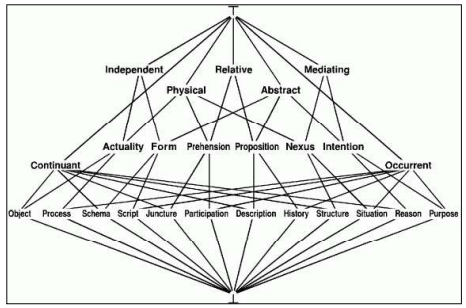
Clear, powerful, readable both by human and computer



Formal Language

Syntax and Semantics:

- ✓ Clear
- ✓ Precise
- ✓ Formally specified



A concept (category) is an unary predicate.

Form(x) ::= Independant(x) ∧ Abstract(x)

➡ **Properties of Axiomatic System**

Definitions are:

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

Logic

A concept is a well formed formula

Definition:

Pocket-Portable-Time-Piece(x) ::= Time-Piece(x) ∧ Portable(x) ∧ Pocket(x)

Properties:

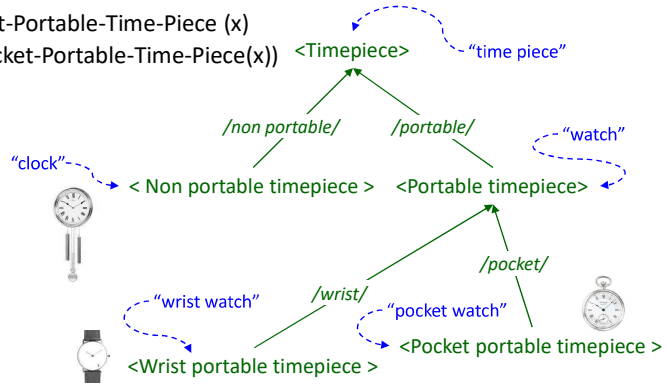
⊨ ¬ (Pocket(x) ∧ Wrist(x))

Reasoning:

Portable(x) ⟶ (Wrist-Portable-Time-Piece (x)
∨ Pocket-Portable-Time-Piece(x))

Computer readable:

Description Logic



Logic is necessary

Which Language?

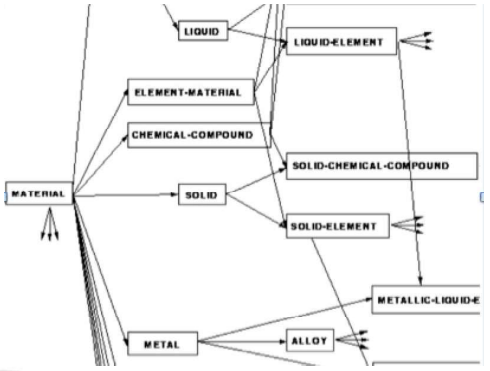
Epistemological Problems

✗ **Logic**

Man(x), Reasonable (x), Sick(x)

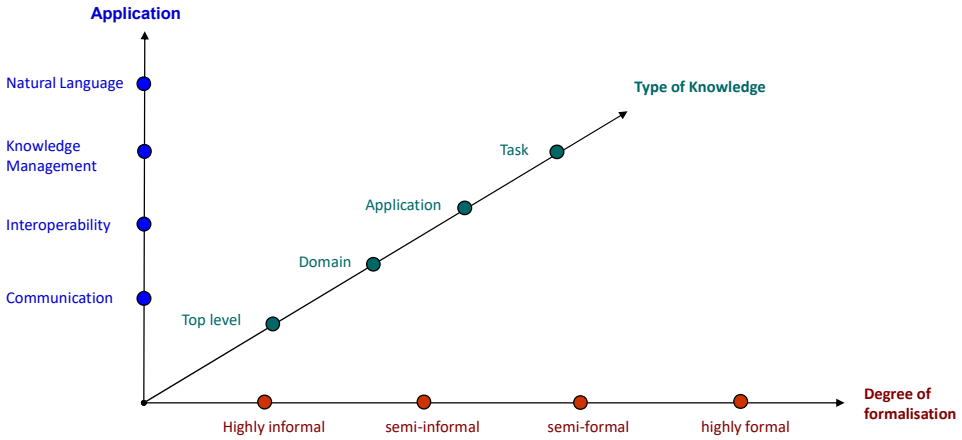
✗ **Frame Languages**

Is mercury a liquid metal? no



Mercury is a heavy, silvery-white liquid metal. Compared to other metals, it is a poor conductor of heat, but a fair conductor of electricity.

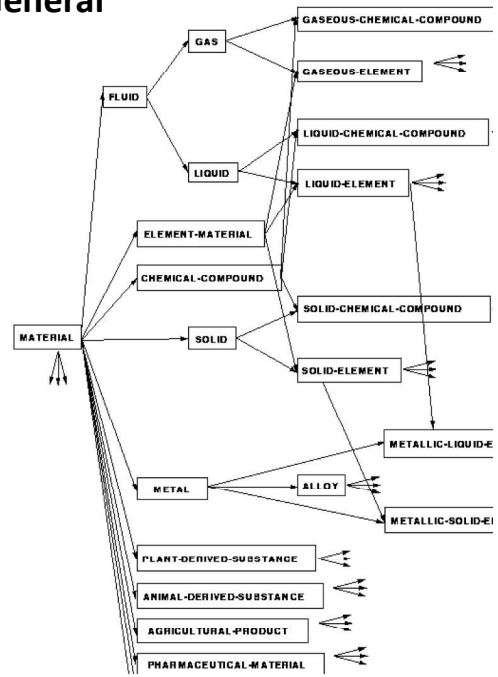
Typology



Examples

General

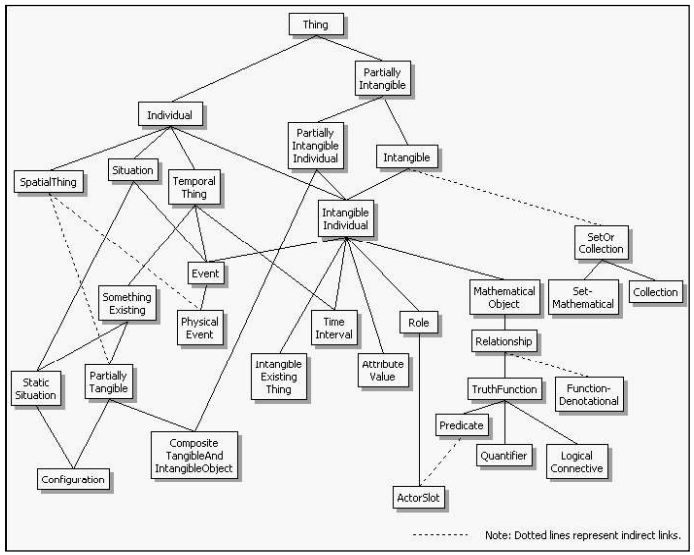
- **Mikrokosmos**



Examples

General

- **Upper Cyc**



Cycorp

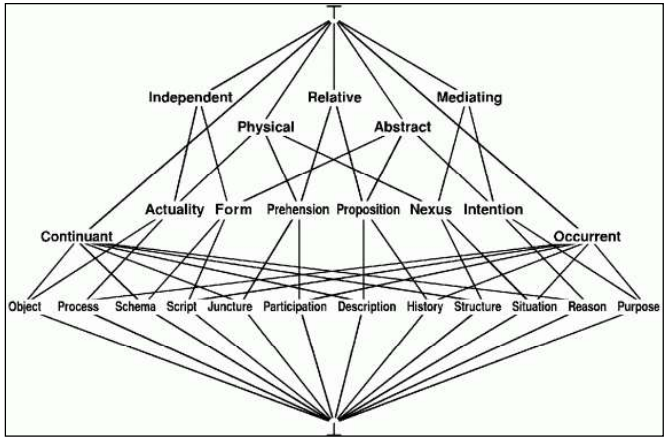
https://www.cyc.com/

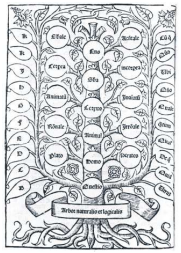
Examples

General

- **« KR Ontology » Sowa**

http://www.jfsowa.com/ontology/kronto.htm





A concept is a unary predicate
A concept is defined as a conjunction of unary predicates

'form(x)' = independant(x) ∧ abstract(x)

http://www.jfsowa.com/krbook/index.htm

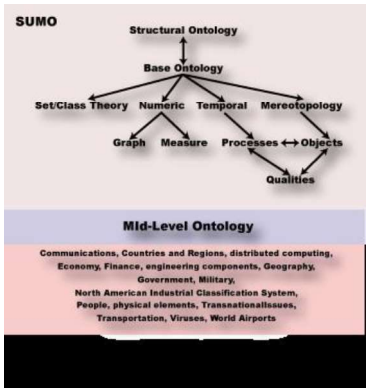
Examples — General

Suggested Upper Merged Ontology (SUMO)

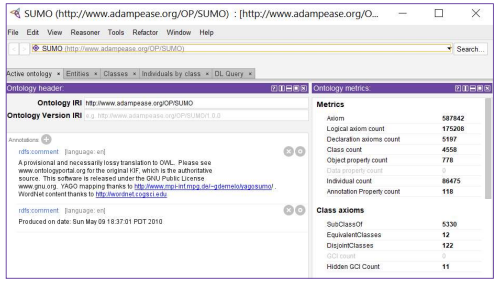
http://www.adampease.org/OP/SUMO.owl

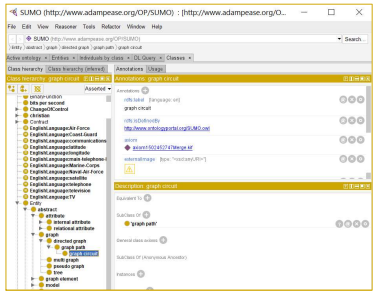
The **Suggested Upper Merged Ontology** (**SUMO**) is an upper ontology intended as a foundation ontology for a variety of computer information processing systems. SUMO defines a hierarchy of *classes* and related rules and relationships. These are expressed in a version of the language SUO-KIF which has a LISP-like syntax.

The Suggested Upper Merged Ontology (SUMO) and its domain ontologies form the largest formal public ontology in existence today. They are being used for research and applications in search, linguistics and reasoning. SUMO is the only formal ontology that has been mapped to all of the WordNet lexicon. SUMO is written in the SUO-KIF language. SUMO is free and owned by the IEEE. The ontologies that extend SUMO are available under GNU General Public License. Adam Pease is the Technical Editor of SUMO.



http://www.adampease.org/OP/SUMO.owl

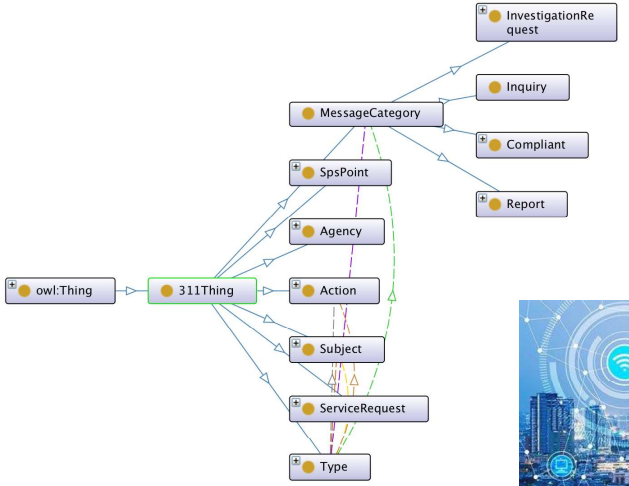


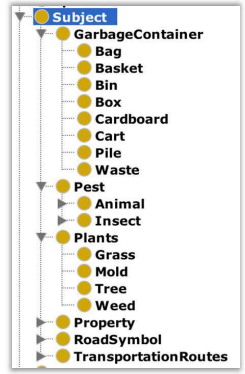


Examples — Smart City

Open 311 Ontology

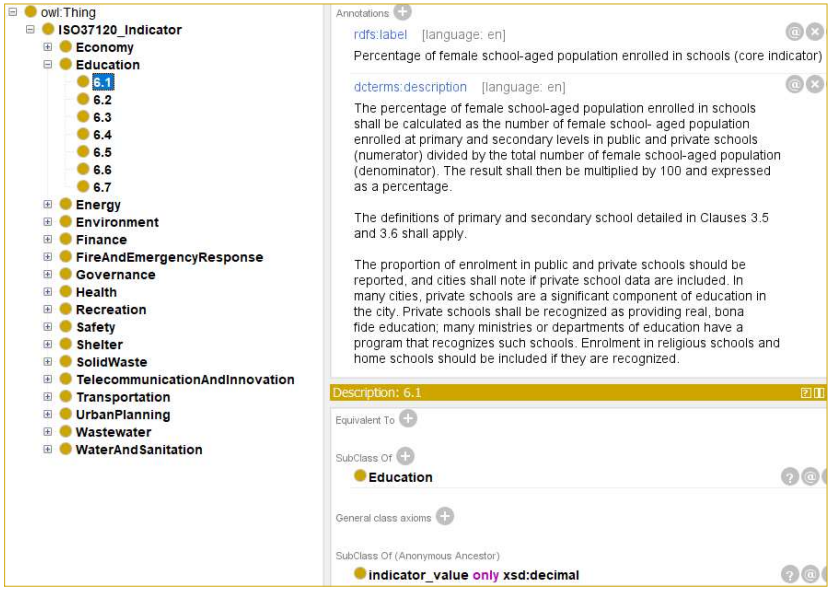
http://ontology.eil.utoronto.ca/o311o.owl





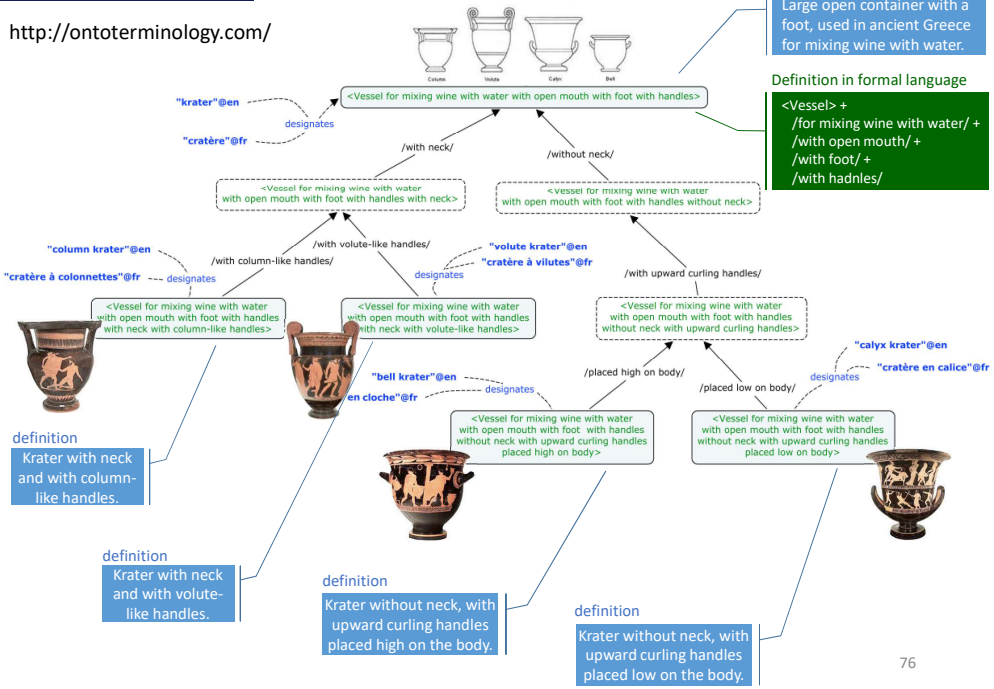
Examples — Smart City

http://ontology.eil.utoronto.ca/ISO37120.owl



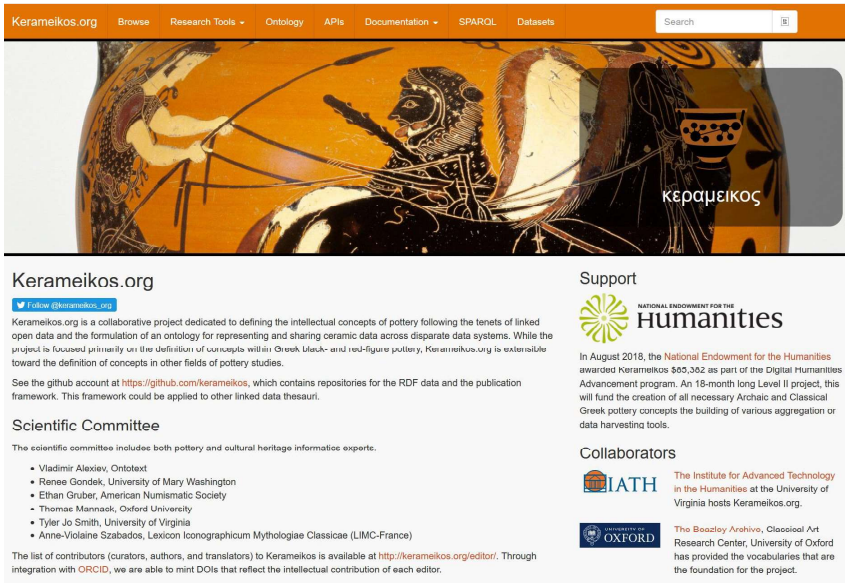
Examples — Digital Humanities

http://ontoterminology.com/

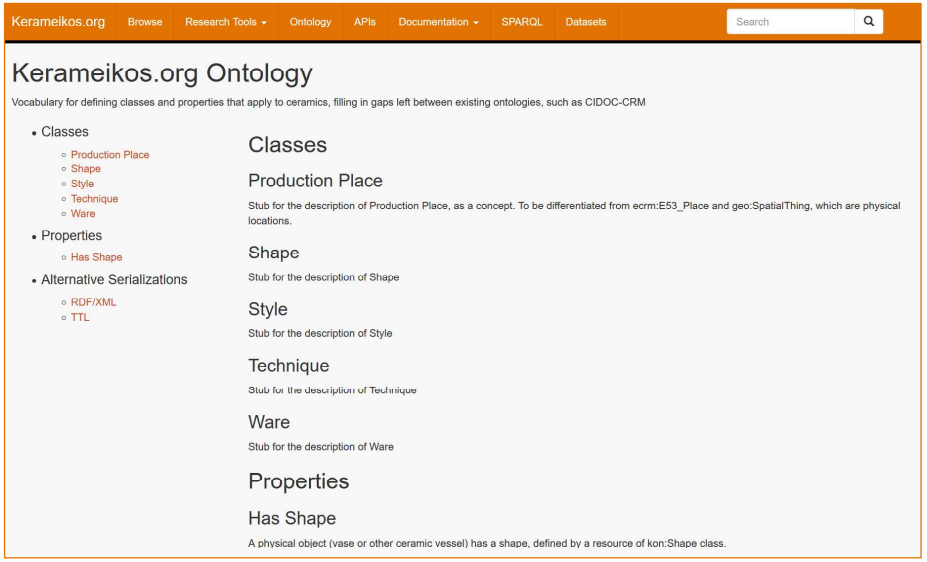


Examples — Digital Humanities

http://kerameikos.org/



Examples — Digital Humanities

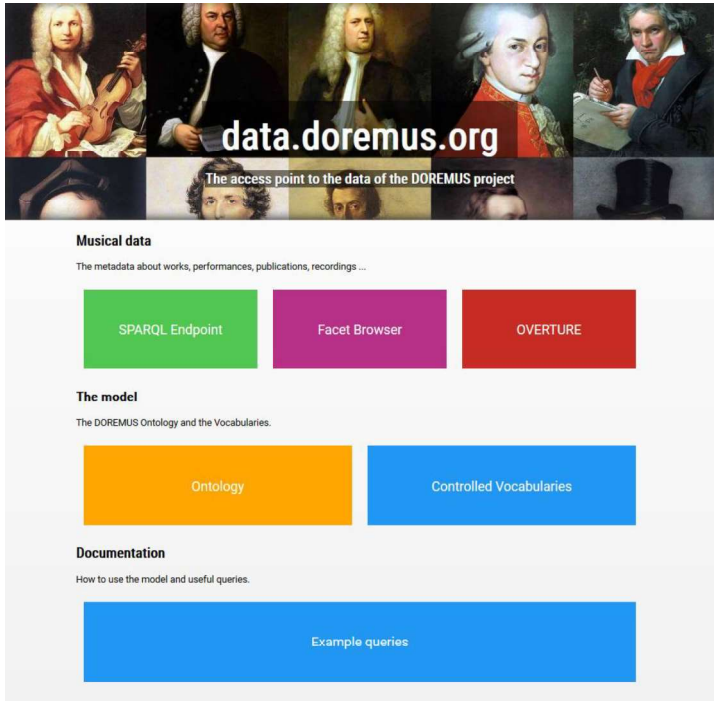


Examples — Digital Humanities

bell_krater (kon:Shape)

skos:prefLabel	Glockenkrater *(de)*, Bell Krater *(en)*, crátera de campana *(es)*, cratère en cloche *(fr)*, cratere a campana *(it)*, звонест кратер *(mk)*, klokkrater *(nl)*
skos:definition	Der Glockenkrater ist eine rot-figurige Innovation. Er weist einen niedrigen oder manchmal modifizierten Standfuß mit dem leichten Ansatz eines Stiels auf. Die Gefäßform erinnert an eine umgedrehte Glocke mit einer leicht nach außen gebogenen Mündung with gerundeten Lippe. Er hat massive, horizontale, zylindrische Henkel, die parallel unter dem Rand angebracht sind. *(de)*
skos:definition	The bell-krater is an innovation belonging to the red-figure technique. The body rises from a low disk-foot or sometimes a modified disk-foot into the hint of a stem before expanding into the shape of an inverted bell with a mildly flaring mouth with a torus lip. It has sturdy, horizontal, cylindrical handles that are located high up on the body opposite one another and are slightly upturned. *(en)*
dcterms:source	https://zenon.dainst.org/Record/000188572
rdf:type	skos:Concept
rdfs:seeAlso	https://www.beazley.ox.ac.uk/tools/pottery/shapes/bell.htm
skos:broader	http://kerameikos.org/id/krater
skos:changeNote	http://kerameikos.org/id/bell_krater#provenance
skos:exactMatch	http://vocab.getty.edu/aat/300198857
skos:exactMatch	http://www.wikidata.org/entity/Q1531905
skos:exactMatch	https://www.britishmuseum.org/collection/term/x5427
skos:inScheme	http://kerameikos.org/id/

http://data.doremus.org/



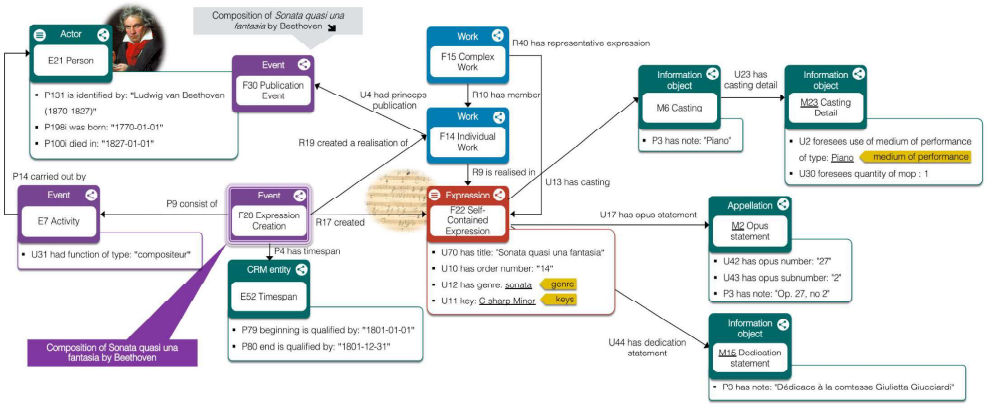
DOREMUS Ontology

The DOREMUS Ontology is an extension the **FRBRoo** model for describing music catalogs.

This ontology defines 83 classes and 279 properties.

Data model overview

Musical composition



FRBRoo

Functional Requirements for Bibliographic Records

Functional Requirements for Bibliographic Records (FRBR) is a conceptual entity–relationship model developed by the International Federation of Library Associations and Institutions (IFLA) that relates user tasks of retrieval and access in online library catalogues and bibliographic databases from a user's perspective.

The FRBRoo ("FRBR-object oriented") initiative is a joint effort of the CIDOC Conceptual Reference Model and Functional Requirements for Bibliographic Records international working groups;

The FRBRoo is a formal ontology intended to capture and represent the underlying semantics of bibliographic information and to facilitate the integration, mediation, and interchange of bibliographic and museum information.

CIDOC CRM

CONCEPTUAL REFERENCE MODEL

http://www.cidoc-crm.org/

The CIDOC Conceptual Reference Model (CRM) is a theoretical and practical tool for information integration in the field of cultural heritage. It can help researchers, administrators and the public explore complex questions with regards to our past across diverse and dispersed datasets.

The CIDOC CRM achieves this by providing definitions and a formal structure for describing the implicit and explicit concepts and relationships used in cultural heritage documentation and of general interest for the querying and exploration of such data.

Such models are also known as formal ontologies. These formal descriptions allow the integration of data from multiple sources in a software and schema agnostic fashion.

The **CIDOC CRM** represents an 'ontology' for cultural heritage information i.e. it describes in a formal language the explicit and implicit concepts and relations relevant to the documentation of cultural heritage. The primary role of the CIDOC CRM is to serve as a basis for mediation of cultural heritage information and thereby provide the semantic 'glue' needed to transform today's disparate, localised information sources into a coherent and valuable global resource.

http://www.cidoc-crm.org/sites/default/files/CIDOC%20CRM_v.7.0.1_%2018-10-2020.pdf

The CIDOC CRM is an ontology in the sense used in computer science.

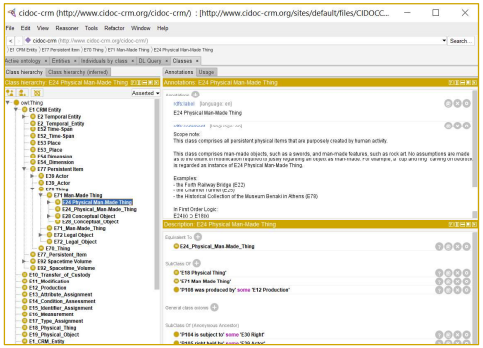
CIDOC CRM Class Hierarchy

E1	CRM Entity
E2	- Temporal Entity
E3	- - Condition State
E4	- - Period
E5	- - - Event
E7	- - - - Activity
E8	- - - - - Acquisition
E96	- - - - - - Purchase
E9	- - - - - Move
E10	- - - - - Transfer of Custody
E11	- - - - - Modification
E12	- - - - - - Production
E79	- - - - - - Part Addition
E80	- - - - - - Part Removal
E13	- - - - - Attribute Assignment
E14	- - - - - - Condition Assessment
E15	- - - - - - Identifier Assignment
E16	- - - - - - Measurement
E17	- - - - - - Type Assignment
E65	- - - - - Creation
E83	- - - - - - Type Creation
E66	- - - - - Formation
E85	- - - - - Joining
E86	- - - - - Leaving
E87	- - - - - Curation Activity

Property id	Property Name	Entity – Domain	Entity - Range
P1	is identified by (identifies)	E1 CRM Entity	E41 Appellation
P48	- has preferred identifier (is preferred identifier of)	E1 CRM Entity	E42 Identifier
P102	- has title (is title of)	E71 Human-Made Thing	E35 Title
P2	has type (is type of)	E1 CRM Entity	E55 Type
P137	- exemplifies (is exemplified by)	E1 CRM Entity	E55 Type
P177	- assigned property type	E13 Attribute Assignment	E55 Type
P3	has note	E1 CRM Entity	E62 String
P79	- beginning is qualified by	E52 Time-Span	E62 String
P80	- end is qualified by	E52 Time-Span	E62 String
P190	- has symbolic content	E90 Symbolic Object	E62 String
P4	has time-span (is time-span of)	E2 Temporal Entity	E52 Time-Span
P5	consists of (forms part of)	E3 Condition State	E3 Condition State
P7	took place at (witnessed)	E4 Period	E53 Place
P8	took place on or within (witnessed)	E4 Period	E18 Physical Thing
P12	occurred in the presence of (was present at)	E5 Event	E77 Persistent Item
P111	- added (was added by)	E79 Part Addition	E18 Physical Thing
P113	- removed (was removed by)	E80 Part Removal	E18 Physical Thing
P11	- had participant (participated in)	E5 Event	E39 Actor
P14	- - carried out by (performed)	E7 Activity	E39 Actor
P22	- - - transferred title to (acquired title through)	E8 Acquisition	E39 Actor
P23	- - - transferred title from (surrendered title through)	E8 Acquisition	E39 Actor
P28	- - - custody surrendered by (surrendered custody through)	E10 Transfer of Custody	E39 Actor
P29	- - - custody received by (received custody through)	E10 Transfer of Custody	E39 Actor
P96	- - by mother (gave birth)	E67 Birth	E21 Person
P99	- - dissolved (was dissolved by)	E68 Dissolution	E74 Group
P143	- joined (was joined by)	E85 Joining	E39 Actor
P144	- - joined with (gained member by)	E85 Joining	E74 Group
P145	- - separated (left by)	E86 Leaving	E39 Actor
P146	- - separated from (lost member by)	E86 Leaving	E74 Group

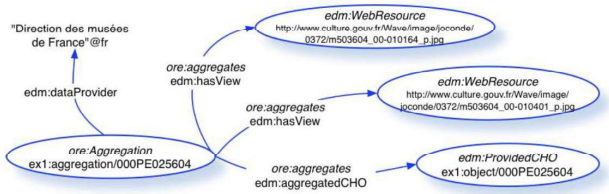
http://www.cidoc-crm.org/sites/default/files/CIDOCCRM_ecrm.owl

```
−<owl:Class rdf:about="http://www.cidoc-crm.org/cidoc-crm/E57_Material">
   <owl:equivalentClass rdf:resource="http://erlangen-crm.org/current/E57_Material"/>
  −<rdfs:subClassOf>
      <owl:Class rdf:about="http://www.cidoc-crm.org/cidoc-crm/E55_Type"/>
   </rdfs:subClassOf>
  −<rdfs:comment xml:lang="en">
      Scope note: This class is a specialization of E55 Type and comprises the concepts of materials. Instances of E57 Material may denote
      properties of matter before its use, during its use, and as incorporated in an object, such as ultramarine powder, tempera paste,
      reinforced concrete. Discrete pieces of raw-materials kept in museums, such as bricks, sheets of fabric, pieces of metal, should be
      modelled individually in the same way as other objects. Discrete used or processed pieces, such as the stones from Nefer Titi's temple,
      should be modelled as parts (cf. P46 is composed of). This type is used categorically in the model without reference to instances of it,
      i.e. the Model does not foresee the description of instances of instances of E57 Material, e.g.: "instances of gold". It is recommended
      that internationally or nationally agreed codes and terminology are used. Examples: - brick - gold - aluminium - polycarbonate - resin
   </rdfs:comment>
   <skos:notation rdf:datatype="http://www.w3.org/2001/XMLSchema#string">E57</skos:notation>
   <rdfs:label xml:lang="en">E57 Material</rdfs:label>
 </owl:Class>
```



Europeana Data Model

EDM adheres to the modelling principles that underpin the approach of the Web of Data ("Semantic Web").



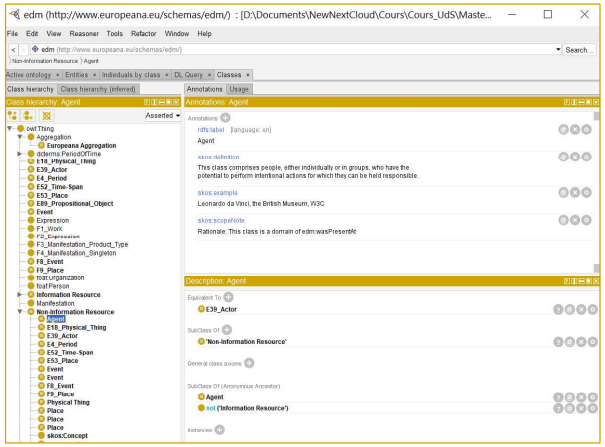
Class name: edm:Agent	
URI	http://www.europeana.eu/schemas/edm/Agent
Label	Agent
Definition	This class comprises people, either individually or in groups, who have the potential to perform intentional actions for which they can be held responsible.
Subclass of	edm:NonInformationResource
Equivalent class	E39_Actor (CIDOC CRM)
Example	Leonardo da Vinci, the British Museum, W3C
Comment	This class is a domain of edm:wasPresentAt

https://pro.europeana.eu/files/Europeana_Professional/Share_your_data/Technical_requirements/EDM_Documentation/EDM_Definition_v5.2.7_042016.pdf

Europeana Data Model

http://www.europeana.eu/schemas/edm/rdf/edm.owl

```
−<owl:Class rdf:about="http://www.europeana.eu/schemas/edm/Agent">
   <rdfs:label xml:lang="en">Agent</rdfs:label>
  −<skos:definition>
      This class comprises people, either individually or in groups, who have the potential to perform intentional
      actions for which they can be held responsible.
   </skos:dcfinition>
   <skos:example>Leonardo da Vinci, the British Museum, W3C</skos:example>
   <rdfs:subClassOf rdf:resource="http://www.europeana.eu/schemas/edm/NonInformationResource"/>
   <owl:equivalentClass rdf:resource="http://www.cidoc-crm.org/rdfs/cidoc-crm#E39_Actor"/>
  −<skos:scopeNote>
      Rationale: This class is a domain of edm:wasPresentAt
   </skos:scopeNote>
 </owl:Class>
```



LIBRARY OF CONGRESS

https://id.loc.gov/ontologies/premis-3-0-0.html

PREMIS 3 Ontology

Ontology for PREMIS 3, the international standard for metadata to support the preservation of digital objects and ensure their long-term usability.

URI: http://www.loc.gov/premis/rdf/v3/
Instance Of: OWL Ontology
Version: 3.0.0
Version IRI: http://www.loc.gov/premis/rdf/v3/2018/10/12/

Class List

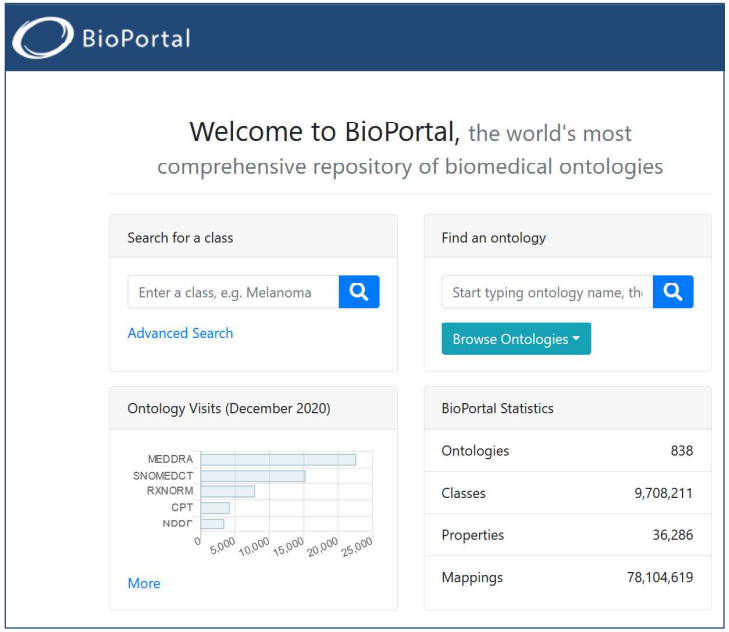
Action	Identifier	Representation
+ Agent	Inhibitor	+ RightsBasis
Bitstream	InstitutionalPolicy	RightsStatus
Copyright	IntellectualEntity	Rule
Dependency	License	Signature
EnvironmentCharacteristic	+ Object	SignatureEncoding
Event	Organization	SignificantProperties
File	OutcomeStatus	SoftwareAgent
Fixity	Person	Statute
HardwareAgent	+ PreservationPolicy	StorageLocation
		StorageMedium

Property List

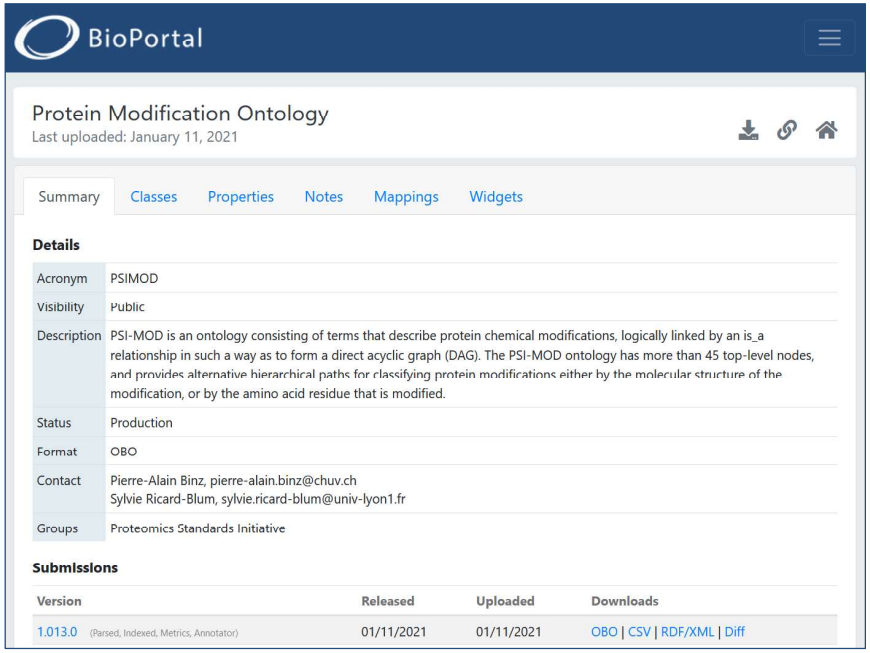
act	governs	prohibits
allows	identifier	purpose
basis	inhibitedBy	rationale
characteristic	inhibits	relationship
citation	jurisdiction	restriction
compositionLevel	key	rightsStatus
dependency	medium	signature
determinationDate	+ note	size
documentation	originalName	startDate
encoding	outcome	storedAt
endDate	outcomeNote	terms
fixity	policy	validationRules
		version

Class:	premis:Action
URI:	http://www.loc.gov/premis/rdf/v3/Action
Label:	Action
Comment:	Operation type to perform on an Object. Effectively performing this action may produce an Event.

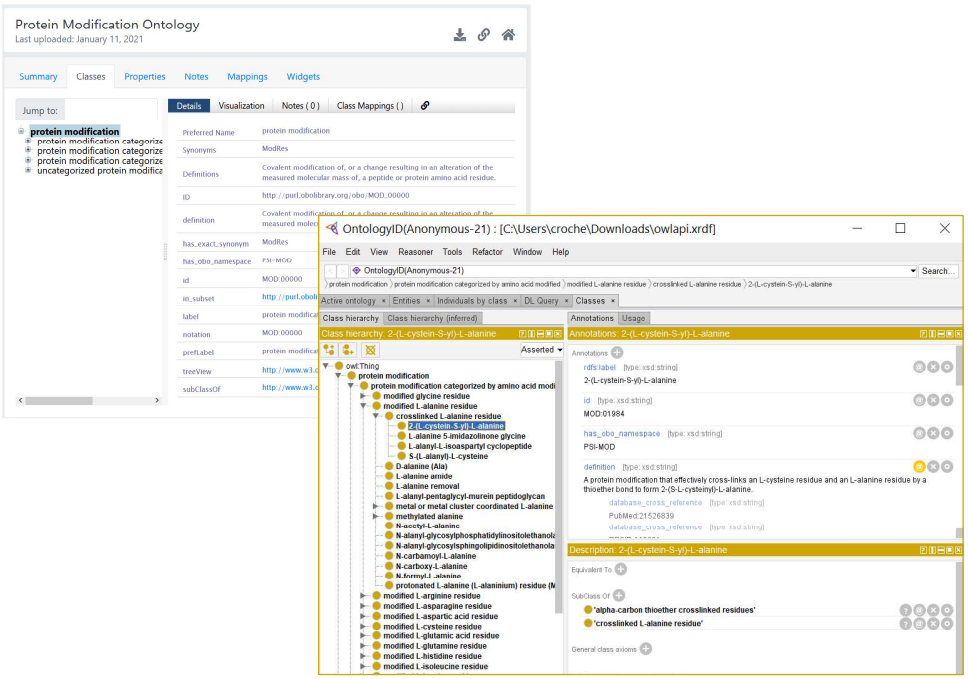
Examples — Medicine



Examples — Medicine



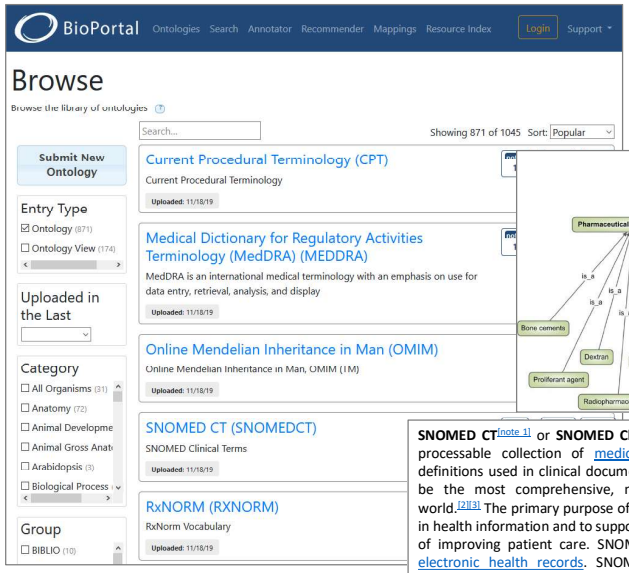
Examples — Medicine

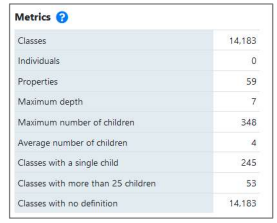


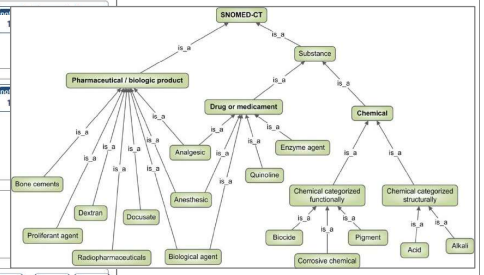
Examples — Medicine

SNOMED

http://www.snomed.org/





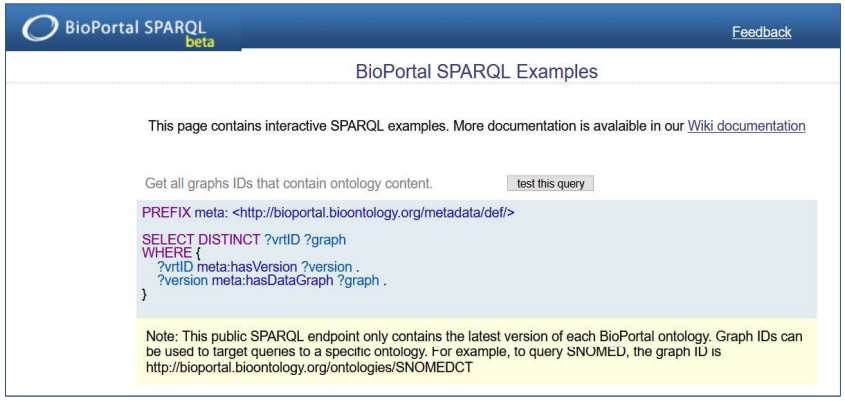


SNOMED CT[note 1] or **SNOMED Clinical Terms** is a systematically organized computer processable collection of medical terms providing codes, terms, synonyms and definitions used in clinical documentation and reporting. SNOMED CT is considered to be the most comprehensive, multilingual clinical healthcare terminology in the world.[2][3] The primary purpose of SNOMED CT is to encode the meanings that are used in health information and to support the effective clinical recording of data with the aim of improving patient care. SNOMED CT provides the core general terminology for electronic health records. SNOMED CT comprehensive coverage includes: clinical findings, symptoms, diagnoses, procedures, body structures, organisms and other etiologies, substances, pharmaceuticals, devices and specimens.

Examples

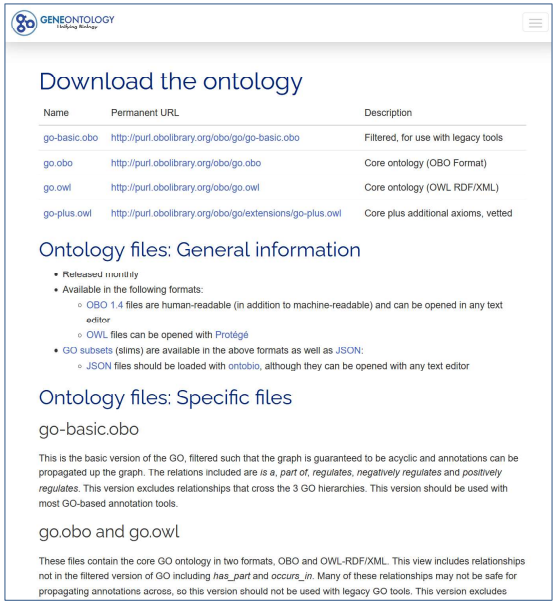
Medicine

http://sparql.bioontology.org/examples



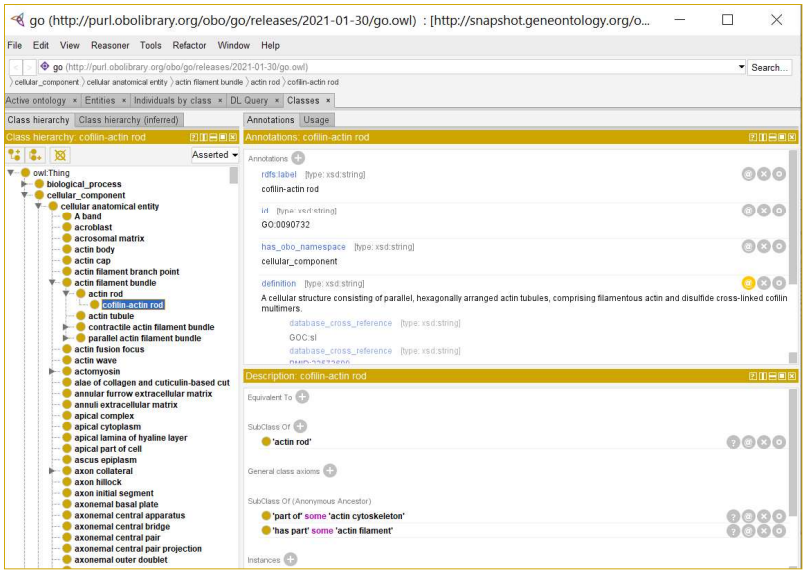
GENEONTOLOGY
Unifying Biology

http://geneontology.org/docs/download-ontology/

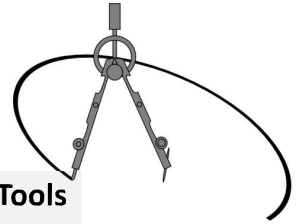


GENEONTOLOGY
Unifying Biology

http://snapshot.geneontology.org/ontology/go.obo



Tools & Environments



- **Graphical Tools:** - CmapTools
- **Formal Tool:** - Protégé
- **Dedicated Tools:** - Tedi

Downloads

CmapTools

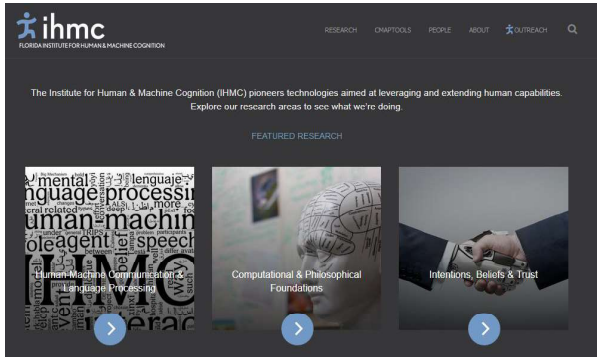
http://cmap.ihmc.us/

Protégé

https://protege.stanford.edu/

CmapTools

http://cmap.ihmc.us/



Concept maps are graphical tools for organizing and representing knowledge.

They include **concepts**, usually enclosed in circles or boxes of some type, and **relationships** between concepts indicated by a connecting line linking two concepts.

CmapTools

☑ **Epistemological Principles**

Concept:	*Concept* as *a perceived regularity in events or objects, or records of events or objects, designated by a label.*
Proposition:	*Propositions* are *statements about some object or event in the universe, either naturally occurring or constructed. Propositions contain two or more concepts connected using linking words or phrases to form a meaningful statement.*

CmapTools

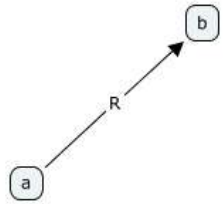
☑ **Epistemological Principles**

Concept:	*Concept* as *a perceived regularity in events or objects, or records of events or objects, designated by a label.*
Proposition:	*Propositions* are *statements about some object or event in the universe, either naturally occurring or constructed. Propositions contain two or more concepts connected using linking words or phrases to form a meaningful statement.*

☑ **Representation**

Concept: *Node*

Proposition: *labeled link*



☑ Epistemological Principles

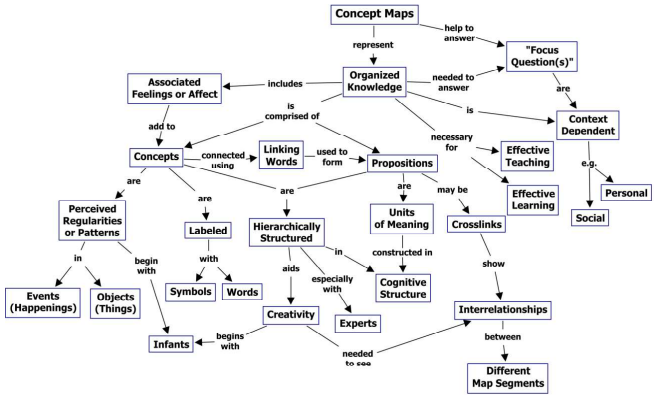
Concept: *Concept as a perceived regularity in events or objects, or records of events or objects, designated by a label.*

Proposition: *Propositions are statements about some object or event in the universe, either naturally occurring or constructed. Propositions contain two or more concepts connected using linking words or phrases to form a meaningful statement.*

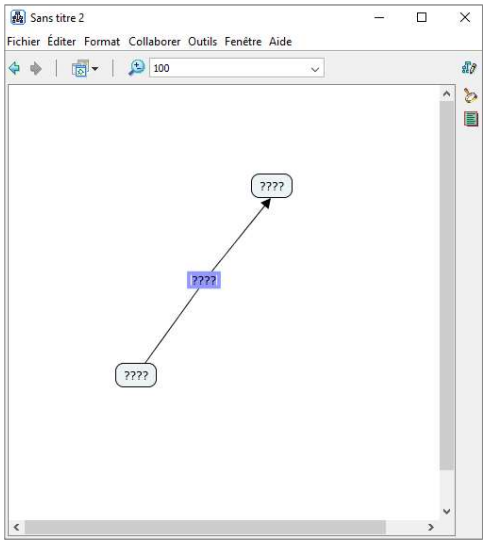
☑ Representation

Concept: *Node*

Proposition: *labeled link*



Do it



How to Represent the Categories of Thought?

Langue of Intellection

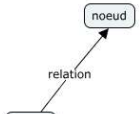
Categories of Thought :

- **Object**
- **Characteristic** : - essential
  - descriptive
- **Concept**
- **Class**
- **Relation** : - generic
  - partitive
  - functonial
  - associative
  - .....

Representation Language

Language Categories of CmapTools

- **Node**
- **Labelled link** (binary relationhips)

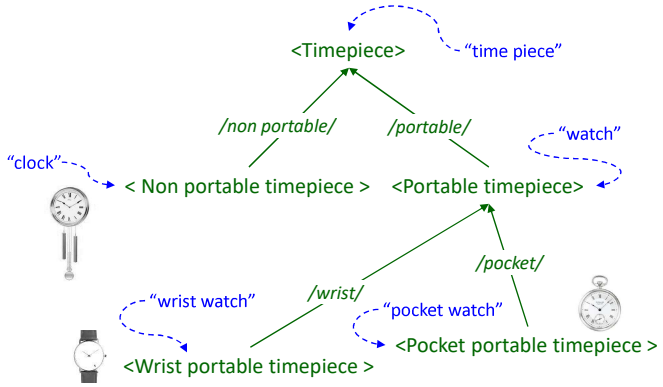


Semiotic code

each category of thought will be represented differently

concept	<Concept>	<Concept>
individual	*individual*	*individual*
Essential characteristic	/characteristic/	isa (characteristic)
relations	R(a,b)	
Descriptive characteristic	at(e,v)	

CmapTools



No verification

Exercise

Ontology of Seats

Christophe Roche
Condillac Research Group
Savoie Mont-Blanc University
Liaocheng University

Exercise

armchair

椅子

chaise

fauteuil

seat

couch

pouf

stool

banc

长凳

tabouret

bench

chair

canapé

couch

长椅

Exercise

Definition of:
- "chaise"? French
- "chair"? English
- "椅子"? Chinese

- A **term** is a designation of a **concept**
- A **concept** is defined as a unique combination of **essential characteristics**

the definition of a term is the definition of the concept denoted by the term expressed in a natural language

Give the definition of the terms:

"seat", "chair", "armchair", "bench", "couch", "stool"

Exercise

Where to start?

- With terms?

- With concepts?

- With characteristics?

- By drawing?

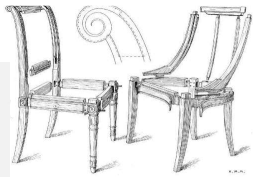
Do it!

Exercise

Essential characteristics:
- *for one person* or *for several persons*
- *with feet* or *without feet*
- *with back* or *without back*
- *with arms* or *without arms*
- *for interior* or *for exterior*
- *confortable* or *not confortable*

Descriptive characteristics:
- material
- colour
- weight

Parts of :
- arms
- feet
- back



Exercise

Essential characteristics:
- *for one person* or *for several persons*
- *with feet* or *without feet*
- *with back* or *without back*
- *with arms* or *without arms*
- *for interior* or *for exterior*
- *confortable* or *not confortable*

"A chair is a seat for one person, with back, with feet and without arms"

"chair" → { /for one person/ , /with feet/ , /with back/ , /without arms/ }

Array of differences

Objects	Concepts	Axis of analysis		Axis of analysis		Axis of analysis		Axis of analysis		Terms	
Objects	Concepts	for one person	several persons	with feet	without feet	with back	without back	with arms	without arms	Designations (English)	Designations (French)
	<Seat 1 person with feet with back without arms>	X		X		X			X	"chair"	"chaise"
	<Seat 1 person with feet with back with arms>	X		X		X		X		"armchair"	"fauteuil"
	<Seat 1 person with feet without back without arms>	X		X			X		X	"stool"	"tabouret"
	<Seat several persons with feet with back with arms>		X	X		X		X		"couch"	"canapé"
	<Seat several persons with feet without back without arms>		X	X			X		X	"bench"	"banc"

<Seat for one person with feet with back without arms>

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

"chair" : Seat for one person with feet and back without arms.

Le Trésor de la Langue Française informatisé

"chaise" : siège à dossier et généralement sans bras.
"chair": seat with backrest and generally without arms

Array of differences

Objects	Concepts	Axis of analysis		Axis of analysis		Axis of analysis		Axis of analysis		Terms	
Objects	Concepts	for one person	several persons	with feet	without feet	with back	without back	with arms	without arms	Designations (English)	Designations (French)
	<Seat 1 person with feet with back without arms>	X		X		X			X	"chair"	"chaise"
	<Seat 1 person with feet with back with arms>	X		X		X		X		"armchair"	"fauteuil"
	<Seat 1 person with feet without back without arms>	X		X			X		X	"stool"	"tabouret"
	<Seat several persons with feet with back with arms>		X	X		X		X		"couch"	"canapé"
	<Seat several persons with feet without back without arms>		X	X			X		X	"bench"	"banc"

"chair" : Seat for one person with feet and back without arms.

<Seat for one person with feet with back without arms>

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

"armchair" : Seat for one person with feet and back with arms.

<Seat for one person with feet with back without arms>

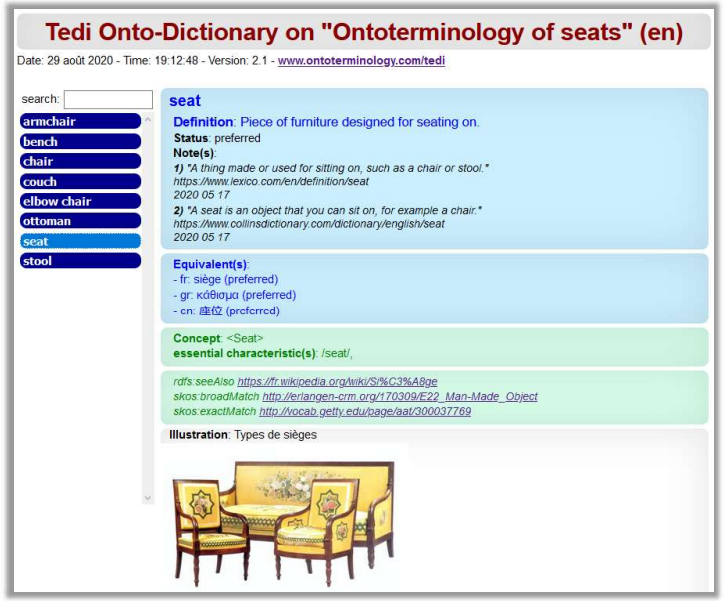
::= <Seat> + /for one person/ + /with feet/ + /with back/ + /with arms/

"bench" : Seat for several persons with feet, without back, and without arms.

<Seat for one person with feet with back without arms>

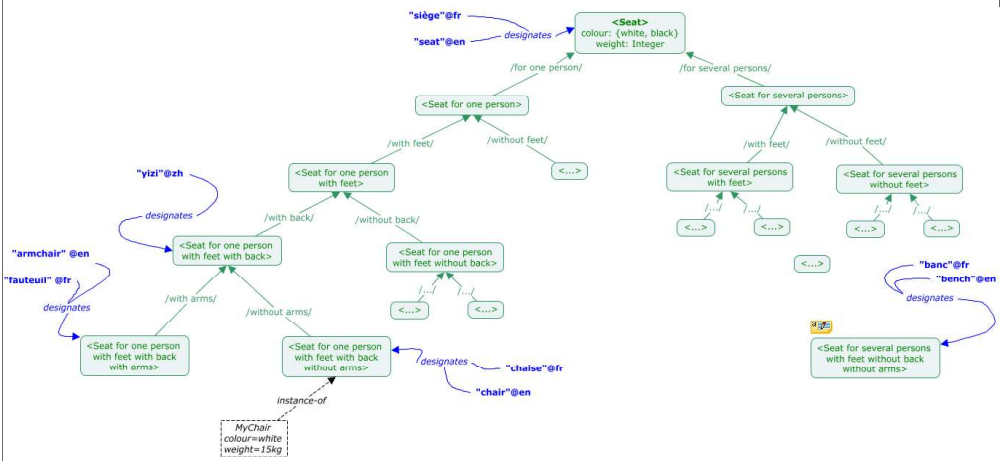
::= <Seat> + /for one person/ + /with feet/ + /with back/ + /with arms/

http://ontoterminology.com/e-dictionaries

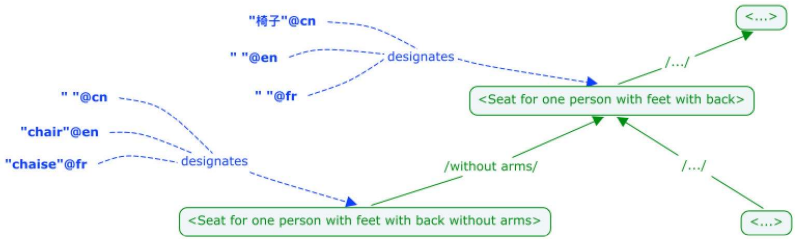


TEDI

Graphical Notation



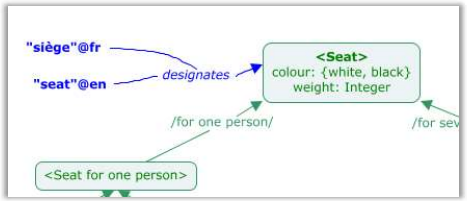
Graphical Notation

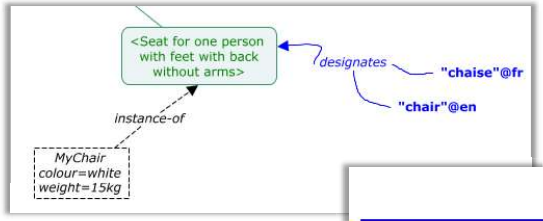


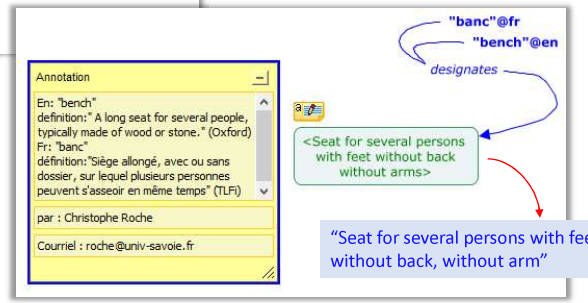
Some concepts can have no designation:

- **in a given language** (e.g. the term "椅子" in Chinese designates the concept of <Seat for one person with back and feet>, which has no designation in English ("chair" or "armchair")
- **in any languages** (e.g. concepts whose purpose is organizing the ontology)

Graphical Notation



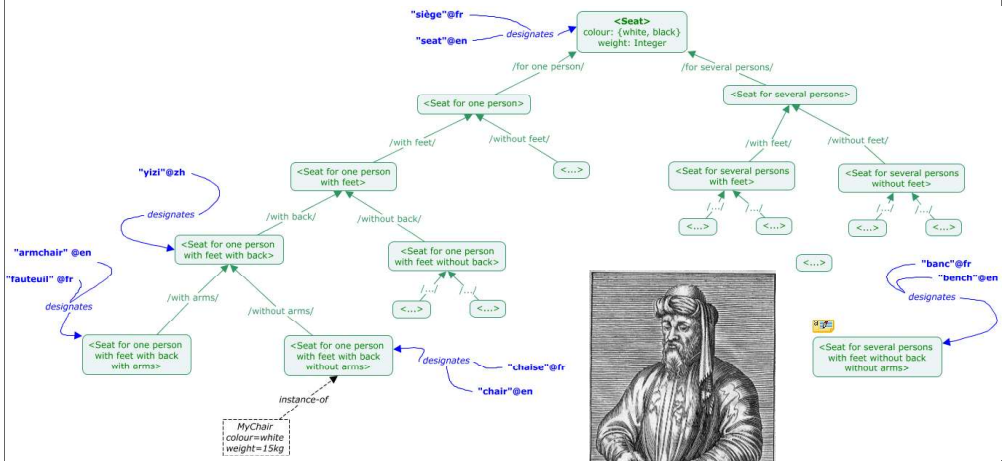




"Seat for several persons with feet, without back, without arm"

Graphical Notation

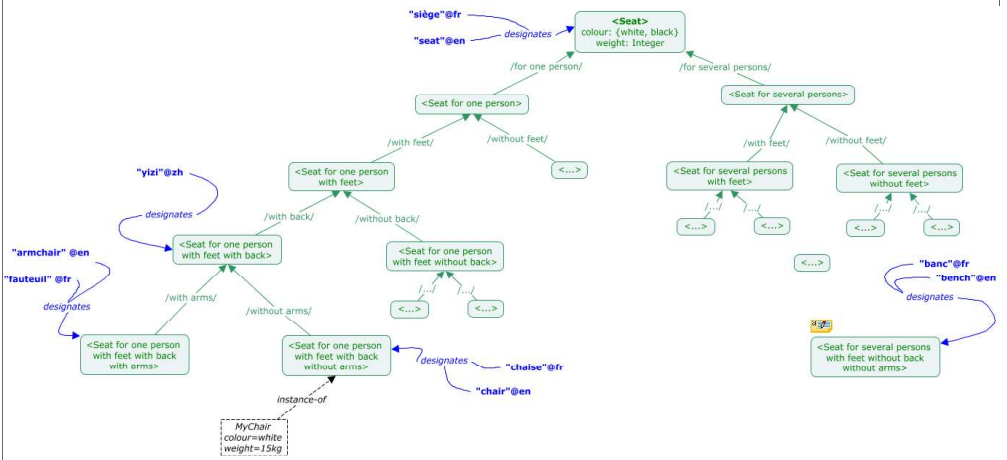
Porphyry Tree



234-305 AD

The ***Isagoge*** (Greek: Εἰσαγωγή, *Eisagōgḗ*) or "Introduction" to Aristotle's "Categories", written by Porphyry in Greek and translated into Latin by Boethius, was the standard textbook on logic for at least a millennium after his death.

Graphical Notation: a Trap?

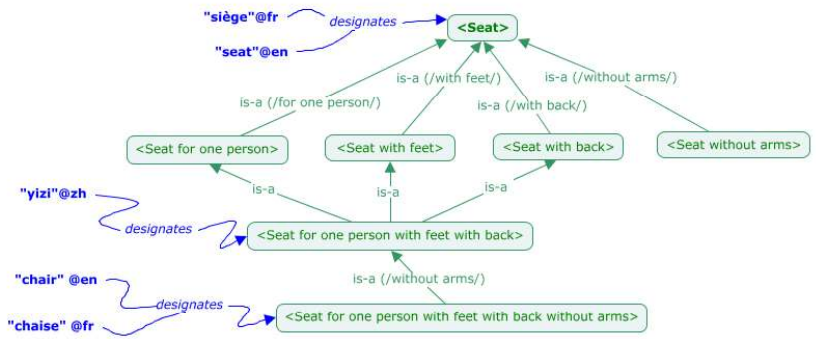


<Seat with back>

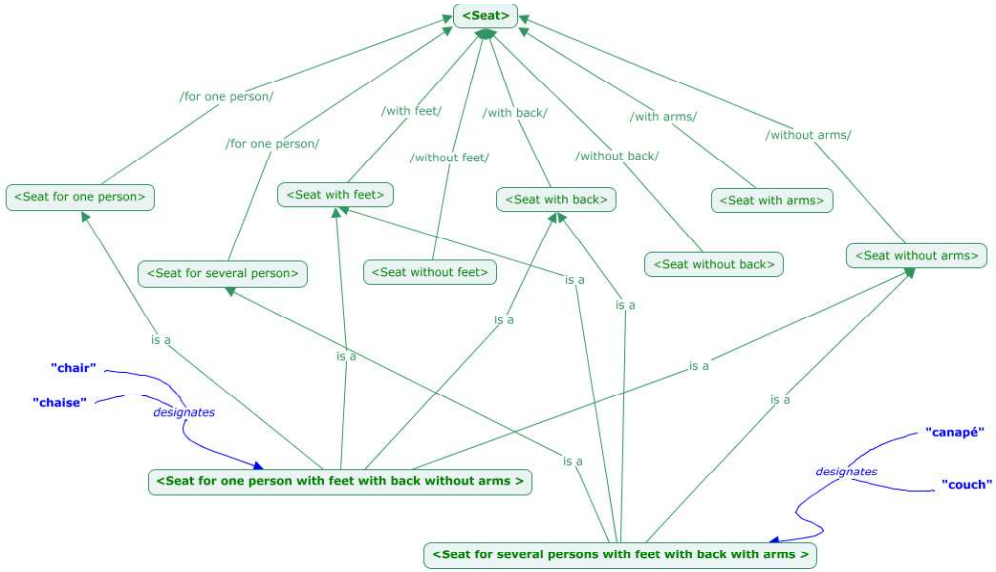
Can I insert it into the tree?

yes or no

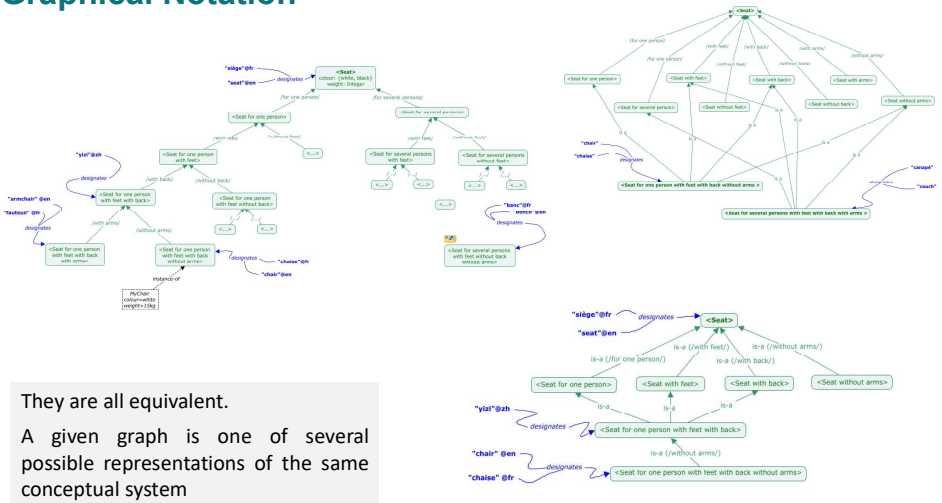
Graphical Notation



Graphical Notation



Graphical Notation



They are all equivalent.

A given graph is one of several possible representations of the same conceptual system

<Concept> = { essential characteristics}

Finding out Differences

Non Physical Objects: Emotions

"A natural instinctive state of mind deriving from one's circumstances, mood, or relationships with others." Oxford Dictionary

thankfulness

modesty

revulsion

curiosity

rage

arrogance

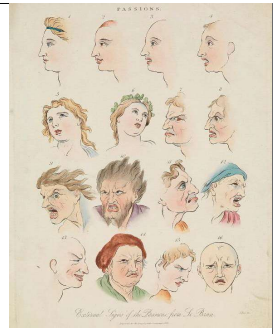
sympathy

Is it possible to find essential characteristics?

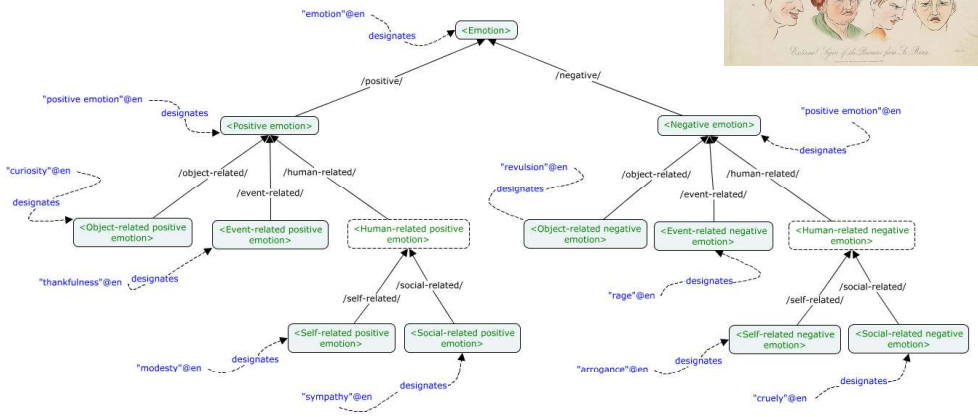
yes or no

Naming is Classifying

Ontology of Emotions



"A natural instinctive state of mind deriving from one's circumstances, mood, or relationships with others." Oxford Dictionary



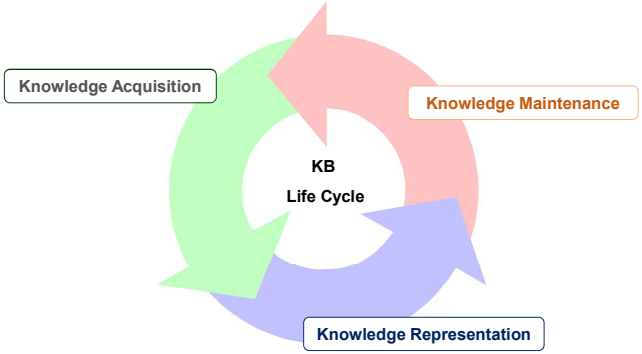
https://en.wikipedia.org/wiki/Emotion_classification

Methodology

Ontology Life Cycle

IEEE Standard 1074-1995

This standard provides the set of Activities that constitute the Processes that are mandatory for the development and maintenance of software, whether stand-alone or part of a system.



Methodology

Uschold & King

« Towards a Methodology for Building Ontologies » AIAI-TR-183

Ontology Engineering

Methodology

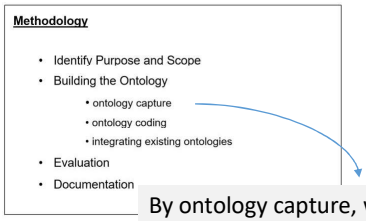
- Identify Purpose and Scope
- Building the Ontology
  - ontology capture
  - ontology coding
  - integrating existing ontologies
- Evaluation
- Documentation

Methodology

Uschold & King

« Towards a Methodology for Building Ontologies » AIAI-TR-183

Ontology Engineering



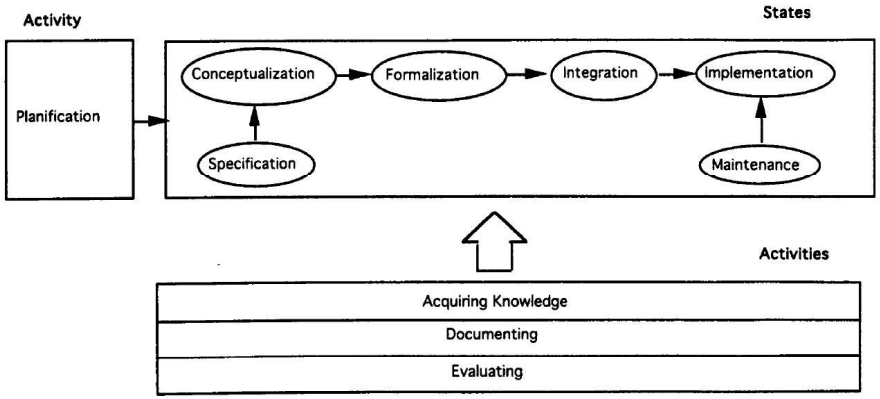
By ontology capture, we mean:

1. identification of the key concepts and relationships in the domain of interest, i.e. scoping
2. production of precise unambiguous text definitions for such concepts and relationships;
3. identification of terms to refer to such concepts and relationships;
4. agreeing on all of the above.

Methodology

"*Overview Of Methodologies For Building Ontologies*" Fernández-López, M. (1999). Proceedings of the IJCAI-99 workshop on Ontologies and Problem-Solving Methods (KRR5) Stockholm, Sweden, August 2, 1999

"METHONTOLOGY: from ontological art towards ontological engineering" Mariano Fernández-LópezAsuncion Gomez-PerezAsuncion Gomez-PerezNatalia JuristoNatalia Juristo. From: AAAI Technical Report SS-97-06. Compilation copyright © 1997, AAAI (www.aaai.org). All rights reserved.



Competency Questions

Reference: "Towards Competency Question-Driven Ontology Authoring" V. Presutti et al. (Eds.): ESWC 2014, LNCS 8465, pp. 752–767, 2014. c Springer International Publishing Switzerland 2014

"A competency question (CQ) is a natural language sentence that expresses a pattern for a type of questions people expect an ontology to answer."

1) **For specification**: to determine the scope, the granularity, the main classes, properties, relations

2) **For Validation**: from competency questions to authoring tests using SPARQL

Evaluation

Reference: *A survey on ontology evaluation methods* Raad, J., & Cruz, C. (2015).

"Ontologies nowadays have become widely used for knowledge representation, and are considered as foundation for Semantic Web. However with their wide spread usage, a question of their evaluation increased even more."

Criteria:

Accuracy: is a criterion that states if the definitions, descriptions of classes, properties, and individuals in an ontology are correct.

Consistency: refers to whether it is possible to obtain contradictory conclusions from valid input definitions.

Completeness: measures if the domain of interest is appropriately covered in this ontology.

Conciseness: is the criteria that states if the ontology includes irrelevant elements with regards to the domain to be covered.

Adaptability: measures how far the ontology anticipates its uses. An ontology should offer the conceptual foundation for a range of anticipated tasks.

Clarity: measures how effectively the ontology communicates the intended meaning of the defined terms. Definitions should be objective and independent of the context.

Computational efficiency: measures the ability of the used tools to work with the ontology, in particular the speed that need to fulfil the required tasks.

One good ontology does not perform equally well with regards to all these criteria. The first task of the evaluator is therefore to choose the criteria relevant for the given evaluation and then to choose the proper evaluation methods to assess how well the ontology meets these criteria.

Evaluation

Tools: OntOlogy Pitfall Scanner!

http://oops.linkeddata.es/

OOPS! (OntOlogy Pitfall Scanner!) is an on-line service intended to help ontology developers, mainly newcomers, during the ontology validation activity, who is unfamiliar with description logics and ontology implementation language.

OOPS! helps to detect some of the most common pitfalls appearing within ontology developments. For example:

- The domain or range of a relationship is defined as the intersection of two or more classes. This warning could avoid reasoning problems in case those classes could not share instances.
- No naming convention is used in the identifiers of the ontology elements. In this case the maintainability, the accessibility and the clarity of the ontology could be improve.
- A cycle between two classes in the hierarchy is included in the ontology. Detecting this situation could avoid modelling and reasoning problems.

Evaluation

OOPS! **The system provides three indicators: critical, important, minor.**

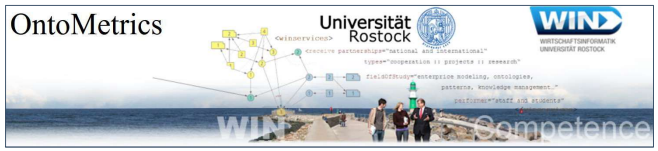
Critical: It is crucial to correct the pitfall. Otherwise, it could affect the ontology consistency, reasoning and applicability, among others. For example, defining wrong inverse relationships.

Important: Though not critical for ontology function, it is important to correct this type of pitfall. For example, equivalent classes or properties not explicitly declared,

Minor: It does not represent a problem. However, correcting it makes the ontology better organized and user friendly. For example, missing annotations.

Evaluation

Tools: **OntoMetrics**



https://ontometrics.informatik.uni-rostock.de/ontologymetrics/

a web-based tool that validates and displays statistics about a given ontology.

- **Base Metrics**
  - Class Axioms
  - Object Property Axioms
  - Data Property Axioms
  - Individual Axioms
  - Annotation Axioms
- **Schema Metrics**
- **Knowledgebase Metrics**
- **Class Metrics**
- **Graph Metrics**

Base Metrics comprise of simple metrics, like the counting of classes, axioms, objects etc. These metrics show the quantity of ontology elements.

Schema Metrics address the design of the ontology.

Attribute Richness

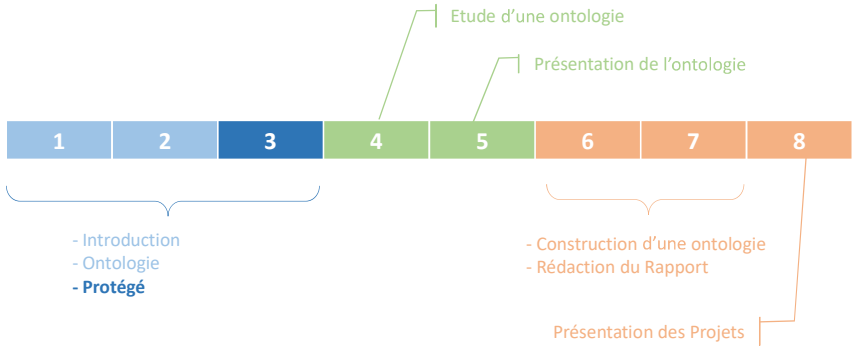
The number of attributes (slots) that are defined for each class can indicate both the quality of ontology design and the amount of information pertaining to instance data.

The attribute richness (AR) is defined as the average number of attributes (slots) per class. It is computed as the number attributes for all classes (att) divided by the number of classes (C).

$$AR = \frac{|ATT|}{|C|}$$

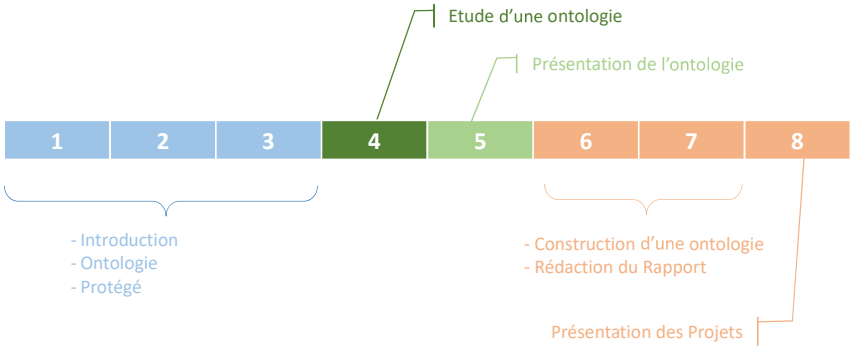
Info 911

Ingénierie des Connaissances

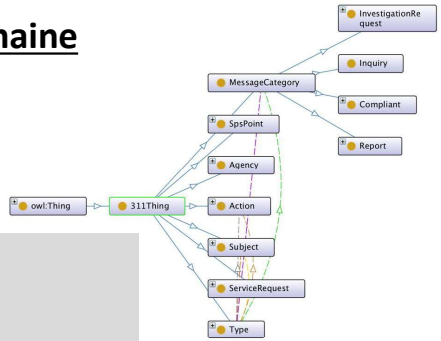


Info 911

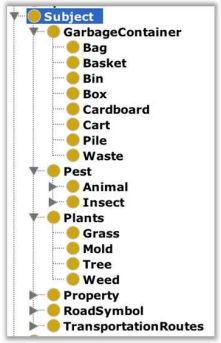
Ingénierie des Connaissances



Etude d'une Ontologie de Domaine

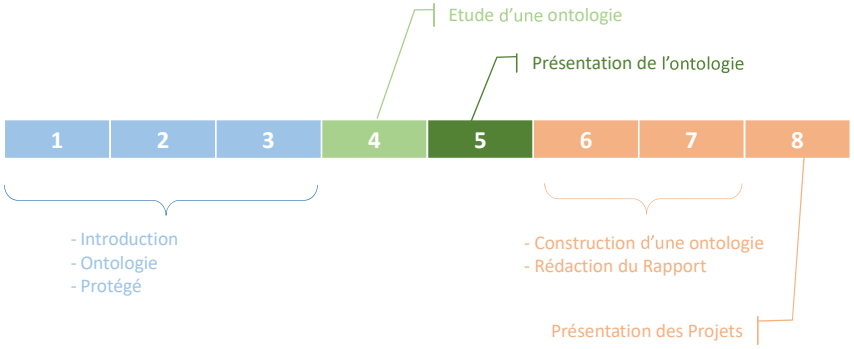


1) **Présentation du domaine**
2) **Vocabulaires utilisés**
3) **Description de l'ontologie de domaine :**
   1) Hiérarchie des concepts
   2) Définition des concepts
   3) Peuplement de l'ontologie (le cas échéant)
4) **Requêtes SPARQL**



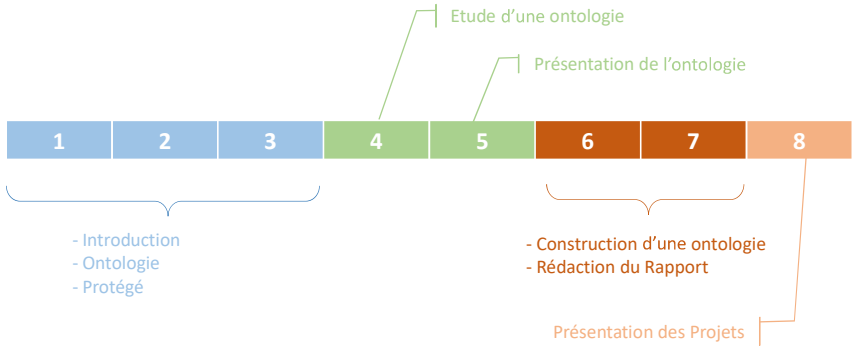
Info 911

Ingénierie des Connaissances



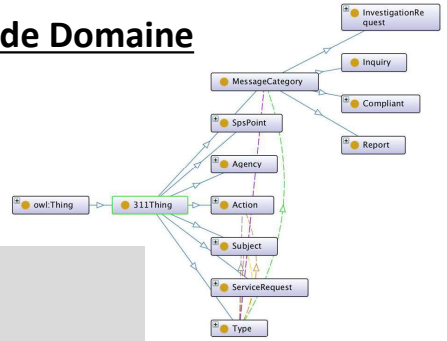
Info 911

Ingénierie des Connaissances



Construction d'une Ontologie de Domaine à l'aide de Protégé

Protégé



1) **Présentation du domaine choisi**
2) **Questions de Compétences (>= 5)**
3) **Vocabulaires utilisés**
4) **Description de l'ontologie de domaine :**
   1) Hiérarchie des concepts
   2) Définition des concepts (restrictions de propriétés)
   3) Peuplement de l'ontologie (s'il y a lieu)
   4) Vérification : Raisonneur(s)
   5) Définition de règles (s'il y a lieu)
5) **Evaluation**
6) **Requêtes SPARQL (Questions de Compétences)**
7) **Rapport**