



Terminologies for ENVRI: Why, What & How an introduction

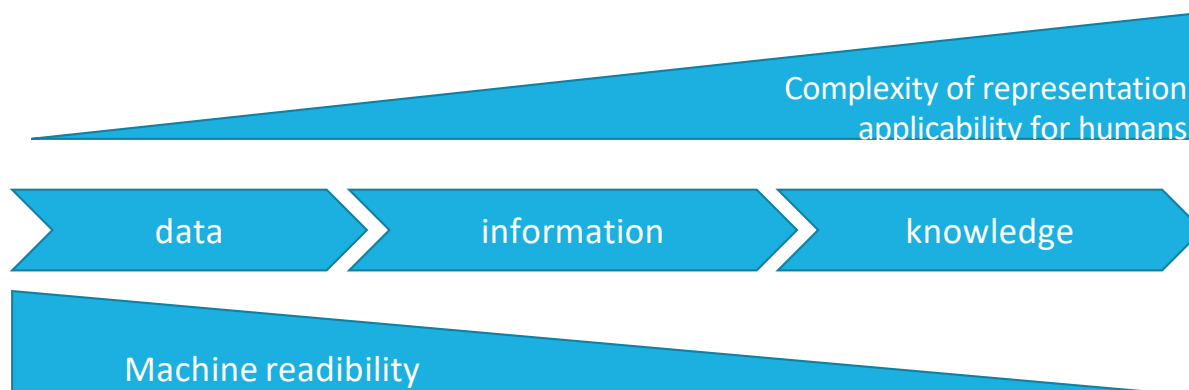
Markus Stocker, Clement Jonquet, Barbara Magagna



ENVRI-FAIR receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824068



From data to knowledge – boosting the I in FAIR



Semantic technologies help to **increase the machine processability of human knowledge** and ultimately increase the I in the FAIRification of data



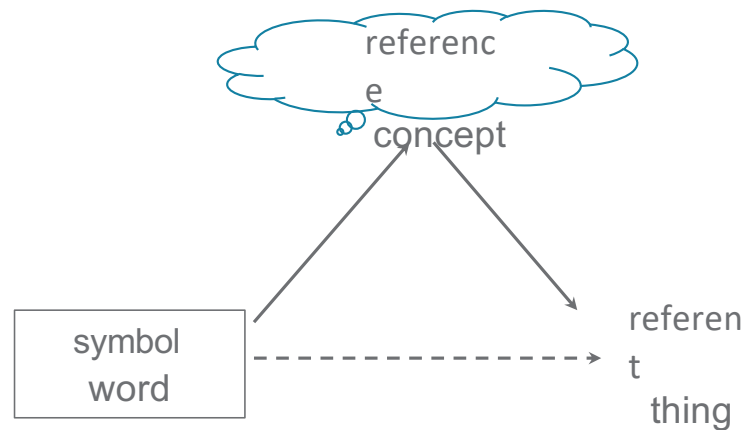
The I principles in FAIR – semantic interoperability

- I1. (Meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation
- I2. (Meta)data use vocabularies that follow FAIR principles
- I3. (Meta)data include qualified references to other (meta)data



Semantics – the Meaning Triangle

- Humans need to communicate to share knowledge. Communication is based on words. The mapping of words to things is indirect. We do it by creating concepts that refer to things.

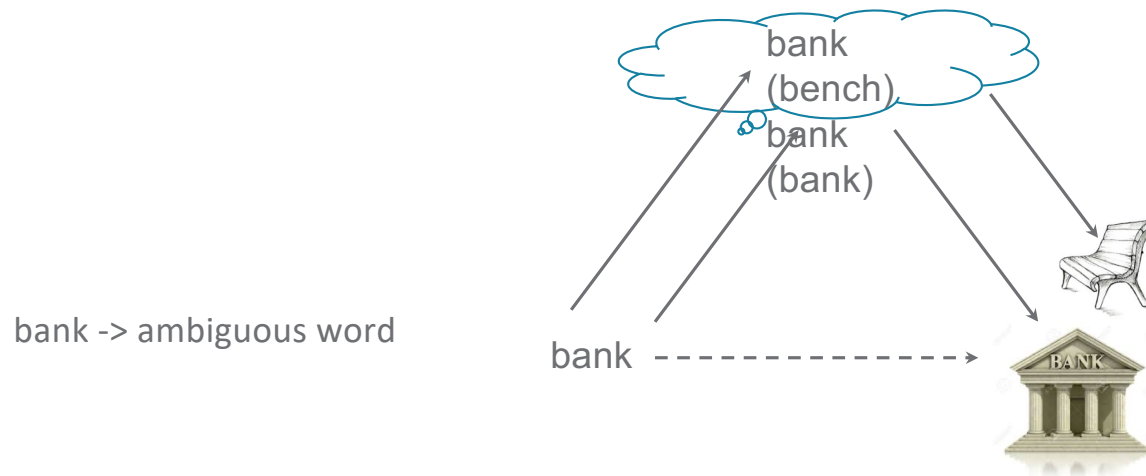


Ogden & Richards (1923)



Problems in human communication

- But human communication is not always straightforward. One word might evoke different interpretations.





Semantic technologies

- Help to disambiguate
 - encoding (thing->concept->word)
 - decoding (word->concept->thing)
- Bank problem solvable by creating two concepts
- Help to extract the implicit human knowledge into an organized, sharable explicit knowledge base



Semantic Web

- World Wide Web -> accumulation of heterogeneous information
- The Semantic Web = extension of the World Wide Web by W3C standards
- Vision of the Web of linked data
- Groundings of Semantic Web technologies
 - semantic standard formats
 - formal logic



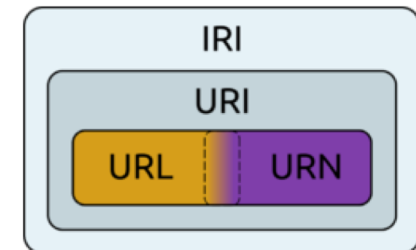
Some definitions for (web) resource identifiers

Uniform Resource Name (URN): a persistent, location-independent identifier based on URN scheme, e.g. urn:isbn:3772305989

Uniform Resource Locator (URL): a reference to a resource that specifies its location on the web.

Uniform Resource Identifier (URI): is the superset and is a reference used to identify a name or location of a resource. Only a string of ASCII characters are allowed.

Internationalized Resource Identifier (IRI): may additionally contain most characters from the Universal Character Set





An environmental domain example: The lightning strike



The lightning strike as an
environmental system
process

It is a lightning process during which electrostatic discharge occurs between a cloud and an object on a planetary surface, or a planetary surface itself, e.g. a tree on a forest floor. (from ENVO: http://purl.obolibrary.org/obo/ENVO_01000901)



RDF

Means

- **Resource:** a thing (*a tree*), an idea or concept (*a lightning process*)
... everything that can have an URI
- **Description:** attributes, features, and relations (predicates) of the resources (*has participant*)
- **Framework:** model, languages and syntaxes for these descriptions



RDF

- Is a triple model decomposing each piece of knowledge (knowlet) into:
subject, **predicate**, **object**

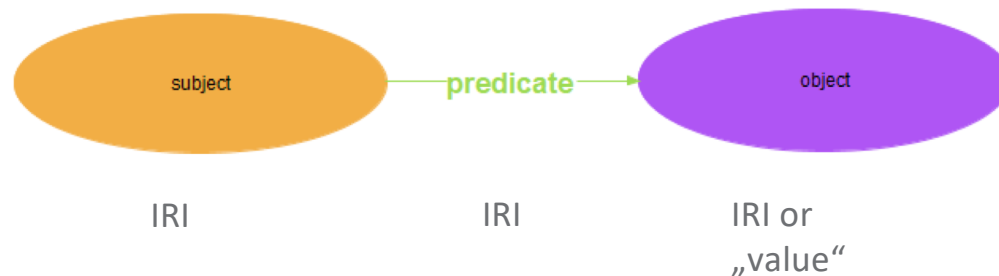


- It is thus also a graph model to link the descriptions of resources



RDF triple

- **subject** is always a resource with an IRI
- **predicates** (properties) are binary relations and their types are identified by IRIs
- **object** (value of property) is a resource with an URI or a literal (string of characters)

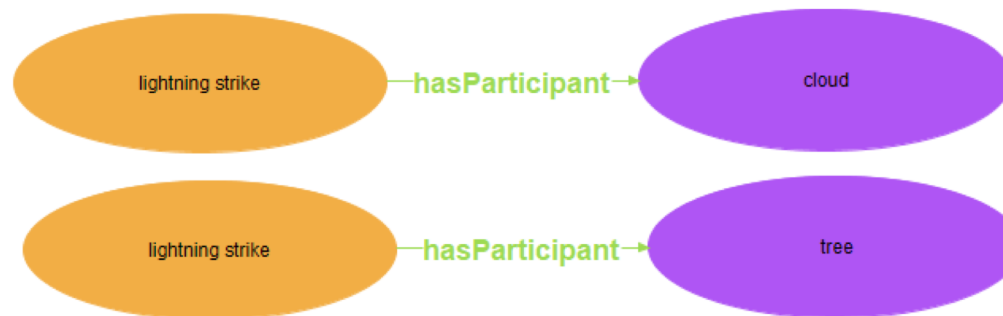




RDF example

- A lightning strike is a process which occurs between a cloud and an object like a tree. We can also say a lightning strike has as participants (in the process) a cloud and a tree.

-> triples: lightning strike, hasParticipant, cloud
lightning strike, hasParticipant, tree





RDF example

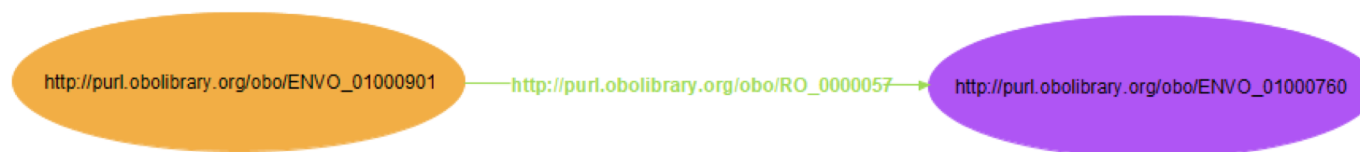
triple:

lightning strike, hasParticipant, cloud ☐

http://purl.obolibrary.org/obo/ENVO_01000901,

http://purl.obolibrary.org/obo/RO_0000057,

http://purl.obolibrary.org/obo/ENVO_01000760





RDF in XML syntax

```
<rdf:RDF
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:obo="http://purl.obolibrary.org/obo#">

  <rdf:Description
    rdf:about="http://purl.obolibrary.org/obo/ENVO_01000901">
    <obo:RO_0000057 rdf:resource=
      "http://purl.obolibrary.org/obo/ENVO_01000760"/>
  </rdf:Description
  </rdf:RDF>
```



Turtle

@prefix rdf:

< <http://www.w3.org/1999/02/22-rdf-syntax-ns#> > .

@prefix envo: <http://purl.obolibrary.org/obo/envo#> > .

<http://purl.obolibrary.org/obo/ENVO_01000901>

obo:RO_0000057 rdf:resource=

< http://purl.obolibrary.org/obo/ENVO_01000760 > .



N-Triples

http://purl.obolibrary.org/obo/ENVO_01000901

http://purl.obolibrary.org/obo/RO_0000057 rdf:resource=

http://purl.obolibrary.org/obo/ENVO_01000760 .



RDFS

RDF Schema provides basic elements for the description of ontologies

- defines classes of resources
- organizes the hierarchy
- defines relations between resources



RDFS example

```
<rdf:RDF xml:base = "http://example.at/2020/process.rdfs"
  xmlns:rdf = "http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
  xmlns    = "http://www.w3.org/2000/01/rdf-schema#">
  <Class rdf:ID="lightning strike">
    <subClassOf rdf:resource="#atmospheric lightning"/>
    <label xml:lang="en">lighting strike</label>
    <comment xml:lang="en">A lightning process during which electrostatic ...
  </Class>
```



SKOS

- **SKOS** is based on RDF
- SKOS concepts are identified with unique URIs and described by natural language labels
- Relations between concepts -> hierarchy
- Mapping concepts also across vocabularies
- SKOS is used for thesauri

Semantic relations	skos:broader skos:narrower skos:related	BT NT RT
Mapping relations	These relations are used to map across vocabularies to enrich semantically each of them. The most popular relationship is skos:exactMatch . skos:broadMatch skos:closeMatch skose:related Match skos:exact Macht	



SKOS example

<http://vocabs.lter-europe.net/EnvThes/20386>

skos:prefLabel "lightning"

skos:broader <http://vocabs.lter-europe.net/EnvThes/10065> (natural induced event)

skos:definition "Lightning is a massive electrostatic discharge between the electrically charged regions within clouds or between a cloud and the Earth's surface."

skos:scopeNote "US LTER controlled vocabulary"

skos:hasExactMatch <https://vocab.lternet.edu/vocab/vocab/?tema=300>



OWL

- RDF is limited in the extraction of implicit knowledge.
- OWL is built upon RDF, but extends its vocabulary with
 - axioms and
 - inferences
- Used for ontologies
- Three variants of OWL:
 - OWL Lite
 - OWL DL
 - OWL Full

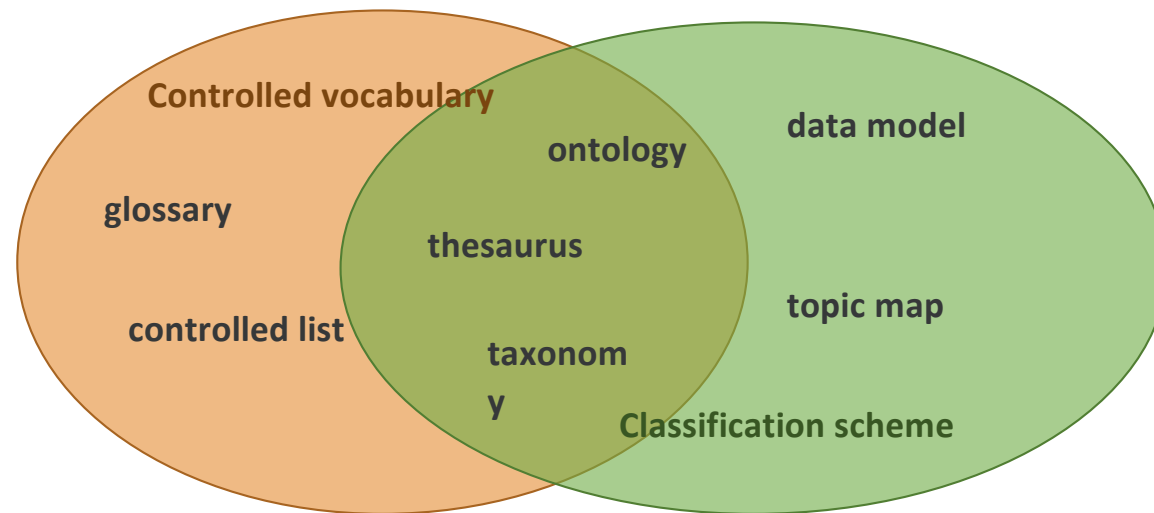


OWL example

```
<owl:Class rdf:about="http://purl.obolibrary.org/obo/ENVO_01000901">
  <rdfs:subClassOf rdf:resource="http://purl.obolibrary.org/obo/ENVO_01000898"/>
  <rdfs:subClassOf>
    <owl:Restriction>
      <owl:onProperty rdf:resource="http://purl.obolibrary.org/obo/RO_0000057"/>
      <owl:someValuesFrom>
        <owl:Class>
          <owl:unionOf rdf:parseType="Collection">
            <rdf:Description rdf:about="http://purl.obolibrary.org/obo/ENVO_01000324"/>
            <owl:Class>
              <owl:intersectionOf rdf:parseType="Collection">
                <rdf:Description rdf:about="http://purl.obolibrary.org/obo/BFO_0000040"/>
                <owl:Restriction>
                  <owl:onProperty rdf:resource="http://purl.obolibrary.org/obo/RO_0002220"/>
                  <owl:someValuesFrom rdf:resource="http://purl.obolibrary.org/obo/ENVO_01000324"/>
                </owl:Restriction>
              </owl:intersectionOf>
            </owl:Class>
          </owl:unionOf>
        </owl:Class>
      </owl:someValuesFrom>
    </owl:Restriction>
  </rdfs:subClassOf>
  <obo:IAO_0000115>A lightning process during which electrostatic discharge occurs between a cloud and an object on a planetary surface, or a planetary surface itself.</obo:IAO_0000115>
  <oboInOwl:hasExactSynonym>CG lightning</oboInOwl:hasExactSynonym>
  <oboInOwl:hasExactSynonym>cloud-ground lightning</oboInOwl:hasExactSynonym>
  <oboInOwl:inSubset>environmental hazards</oboInOwl:inSubset>
  <rdfs:label xml:lang="en">lightning strike</rdfs:label>
</owl:Class>
```



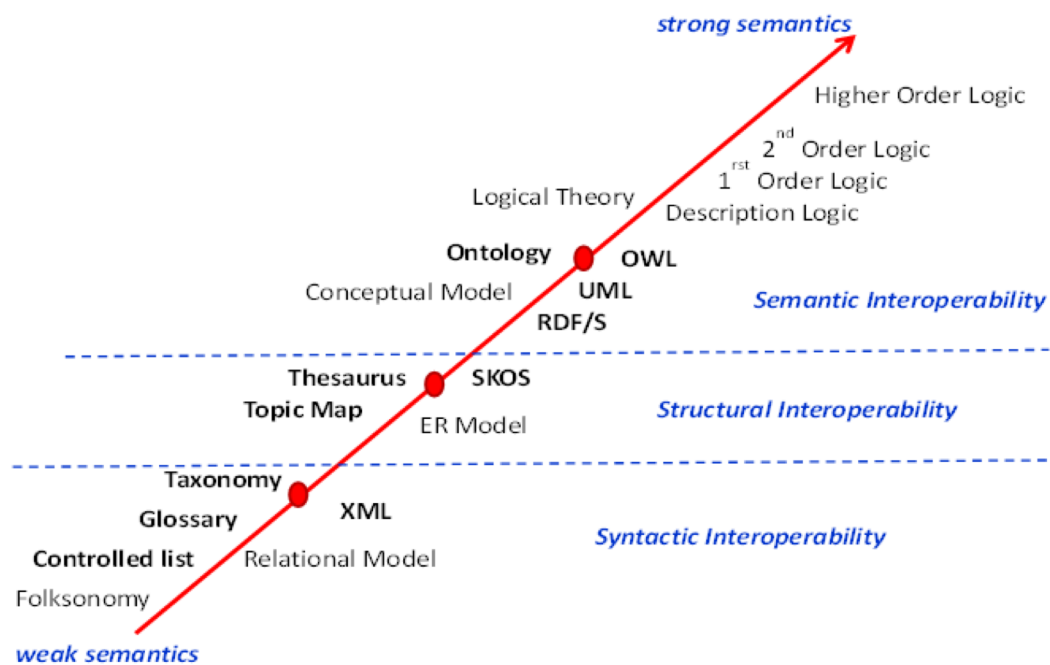
Controlled vocabularies vs ontologies vs terminologies



Terminologies or semantic resources



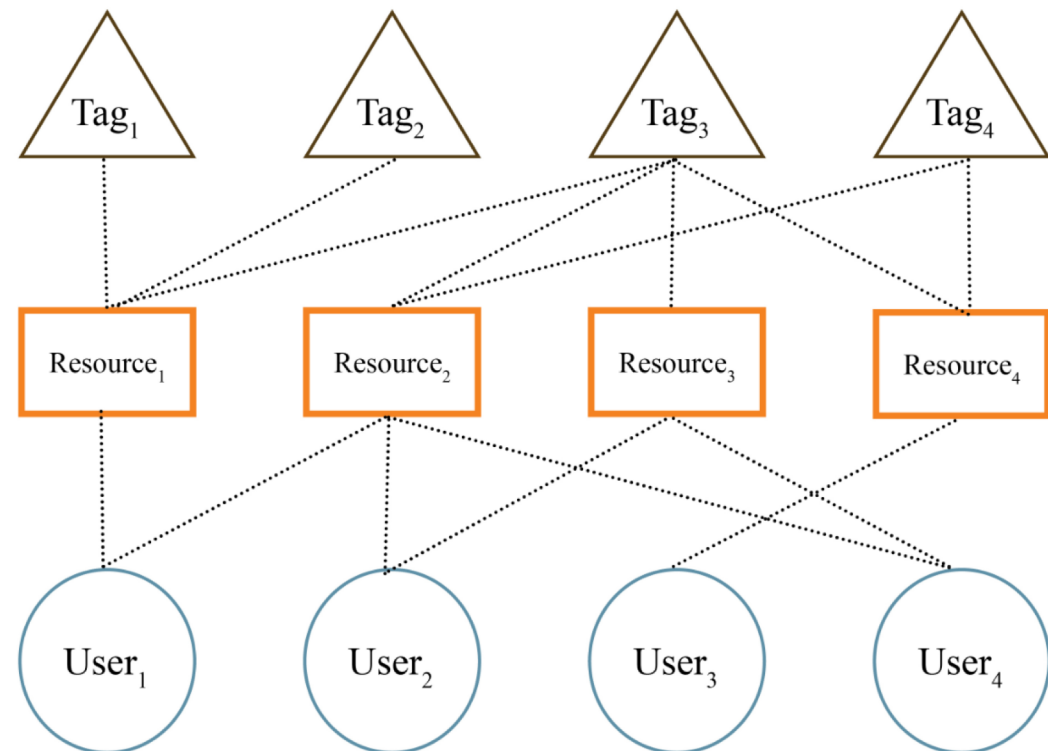
Semantic ladder/spectrum





Folksonomies

- system in which users apply public tags to online items
→ classification system based on those tags and their frequencies
- also known as collaborative tagging, social classification, social indexing, and social tagging





Controlled lists

- Also known as reference list, reference data set or code list
- A simple list of terms used to control terminology
- Each term is unique, all members of the same class
- normally used in databases
- recommend to use SKOS

thing
continuant
independent continuant
material entity
object
meteor
cloud
occurent
process
environmental system process
atmospheric process
atmospheric lightning
lightning strike



Glossaries

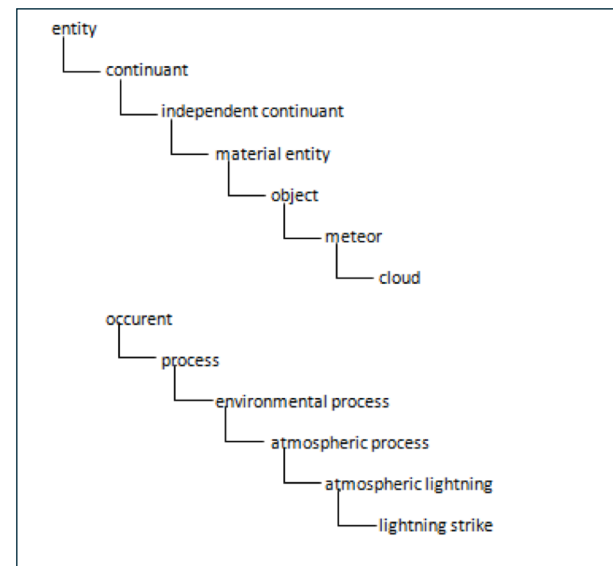
- an alphabetical list of a in a particular domain of knowledge with the definitions for those terms
- also known as a vocabulary or clavis

atmospheric lightning	A sudden electrostatic discharge which occurs during an electrical storm as differentially charged atmospheric entities equalise their charge.
atmospheric process	A process which occurs within an atmosphere.
atmospheric lightning	A sudden electrostatic discharge which occurs during an electrical storm as differentially charged atmospheric entities equalise their charge.
continuant	An entity that exists in full at any time in which it exists at all, persists through time while maintaining its identity and has no temporal parts.
environmental system process	A process in which includes the components of an environmental system as participants.
lightning strike	A lightning process during which electrostatic discharge occurs between a cloud and an object on a planetary surface, or a planetary surface itself.
occurent	An entity that has temporal parts and that happens, unfolds or develops through time.



Taxonomies

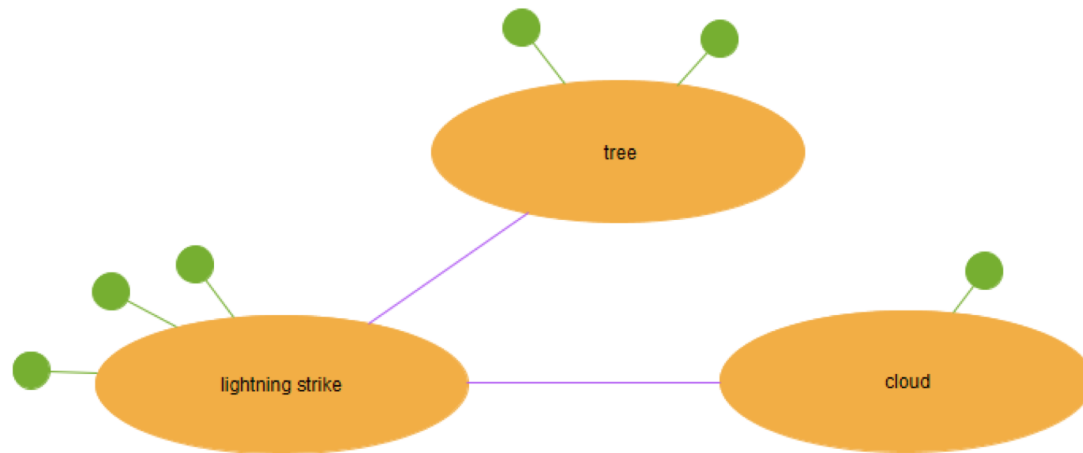
- practice and science of classification of things or concepts
- it contains only terms that are organized into a hierarchical structure





Topic Maps

- A standard for the representation and interchange of knowledge, with an emphasis on the findability of information.
- Uses:
 - (a) topics,
 - (b) associations
 - (c) occurrences





Thesauri

- concept-oriented
- relations between concepts (hierarchy)
- relations between concepts and labels (also in other languages)
- a form of controlled vocabulary with definitions for each concepts
- possibility to link to other concepts (exact match, close match)
- normally expressed in SKOS



Thesaurus example:

lightning

- method
- object of interest
 - entity
 - event
 - extreme event
 - human induced event
 - natural induced event
 - El Nino
 - algal bloom
 - blizzard
 - breeding season
 - burning
 - disturbance (event)
 - drought
 - earthquake
 - fire
 - flood
 - frost
 - growing season
 - insect outburst
 - lightning
 - natural hazard
 - north atlantic oscillation
 - north pacific oscillation
 - ocean circulation
 - pacific decadal oscillation
 - precipitation
 - rain
 - rainfall
 - seed rain
 - storm
 - tide
 - tornado

<http://vocabs.lter-europe.net/EnvThes/20386>

Labels and Description

preferred label: lightning (en)

type: [Concept](#) ▶

definition: Lightning is a massive electrostatic discharge between the electrically charged regions within clouds or between a cloud and the Earth's surface. (en)

Metadata

creator: herbert.schentz@umweltbundesamt.at

created: 2016-03-31T00:00:00Z

modified: 2018-12-18T08:53:13.492+00:00

Standard Relationships

broader concept: [natural induced event](#) ▶

Notes

scope note: US LTER controlled vocabulary (en)

Matching Relationships

has exact match: [c_4327](#) ▶

<http://linkeddata.ge.imati.cnr.it:2020/resource/EARTh/42840> ▶

<https://vocab.lternet.edu/vocab/vocab/?tema=300> ▶

[lightning](#) ▶



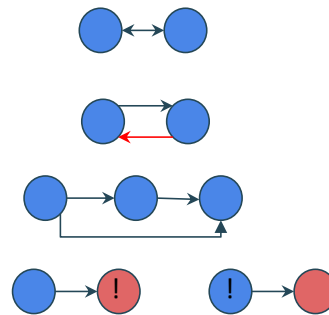
Ontologies

- Sowa: an ontology is a **formal, explicit specification of a shared conceptualization**
- URI concepts related by various well defined kinds of relations
- can be visualized in a graph.
- Most of the ontologies are expressing in OWL



Components of ontologies

- Classes & class definitions
- Properties - object/data/annotation properties
- Algebraic properties
 - symmetric
 - inverse
 - transitive
 - functional/inverse functional
- Axioms
- Instances



Description: 'lightning strike'

Equivalent To +

SubClass Of +

- 'atmospheric lightning'
- 'has participant' some ('planetary surface' or ('material entity' and ('adjacent to' some 'planetary surface')))

General class axioms +

SubClass Of (Anonymous Ancestor)

- 'has participant' some ('environmental feature' or 'environmental material' or 'environmental system')
- process and ('has participant' some ('environmental feature' or 'environmental material' or 'environmental system'))
- 'has participant' min 1 cloud
- 'has output' some plasma
- 'occurs in' some atmosphere
- 'environmental system process' and ('occurs in' some atmosphere)
- 'has participant' some ('environmental feature' or 'environmental material' or 'environmental system')

Instances +

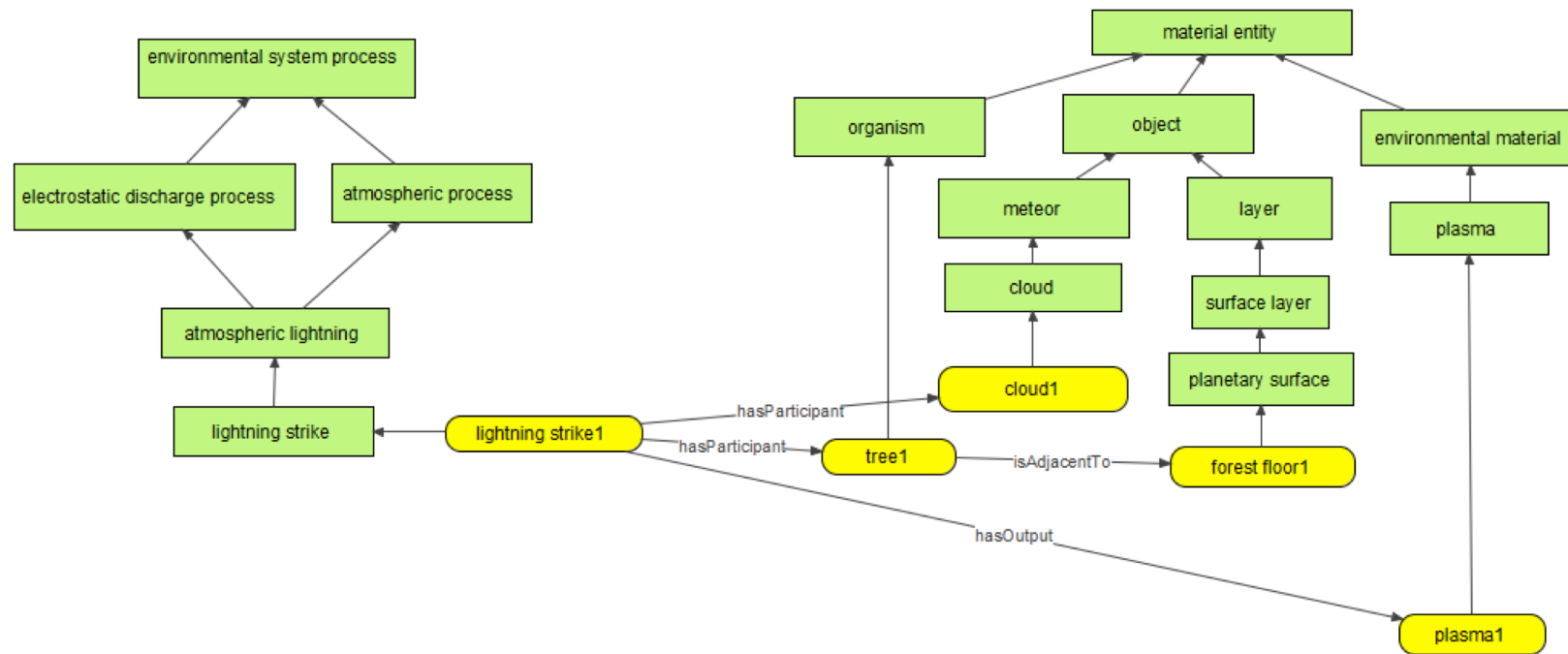
Target for Key +

Disjoint With +

Disjoint Union Of +



A conceptualization of the lightning strike based on





ENVRI
FAIR

THANKS!



envri.eu/envri-fair



[@ENVRIcomm](https://twitter.com/ENVRIcomm)



[ENVRI community](https://www.linkedin.com/groups/ENVRI-community)



facebook.com/ENVRIcomm