# Terminology in ontologies and other lexicographic resources

#### Dra. Guadalupe Aguado de Cea

lupe@fi.upm.es

http://www.oeg-upm.net
Ontological Engineering Group
Facultad de Informática
Universidad Politécnica de Madrid
Campus de Montegancedo sn,
28660 Boadilla del Monte, Madrid, Spain



#### **Outline**

- What is terminology?
- Terminology in scientific domains
- Object of study of terminology
- Concepts and their relations
- Terms and their formation
- Linguistic resources and ontologies
- Final remarks



## What is terminology?

#### The term terminology is polysemic:

- As a product: set of terms from a given subject field.
- As a discipline: set of fundamental principles and conceptual bases that govern the study of specialized terms, their description, analysis and relations.
- As a practice: set of principles oriented towards term compilation



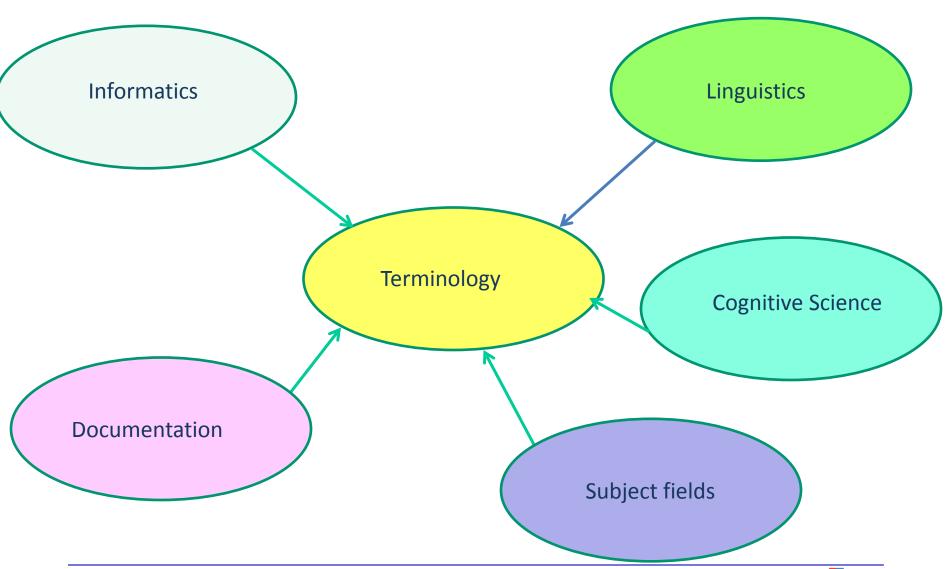
## Terminology in technical & scientific domains

- No professional communication can exist without terminology
- No knowledge transfer can exist without terminology
- Without terminology there is not
  - -Intelectual and material development
  - -Professional research and training
- As a consequence,
  - -no further development would take place
  - A country would isolate from the rest of developed countries

(Picht, 1979)



## Terminology: interdisciplinary subject field



## Object of study in terminology

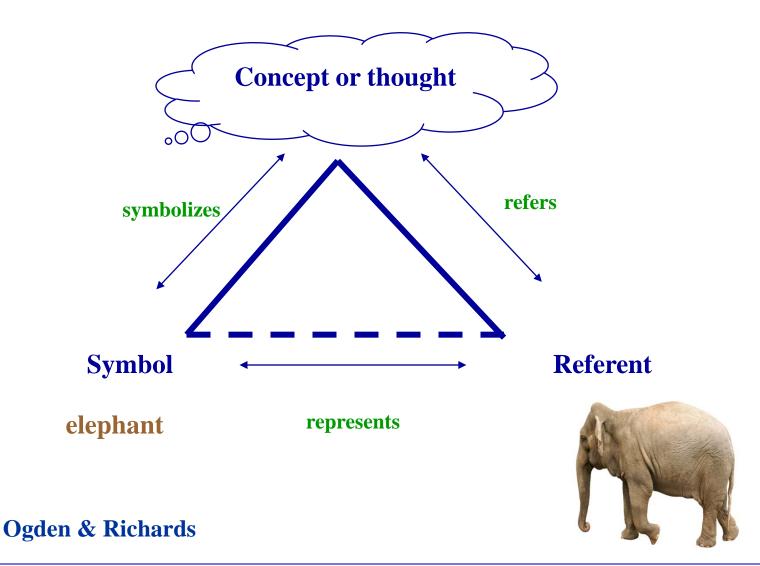
- Concepts
- Terms
- Relation between terms and concepts
- Definitions (not included in this lesson)

#### **NOWADAYS**

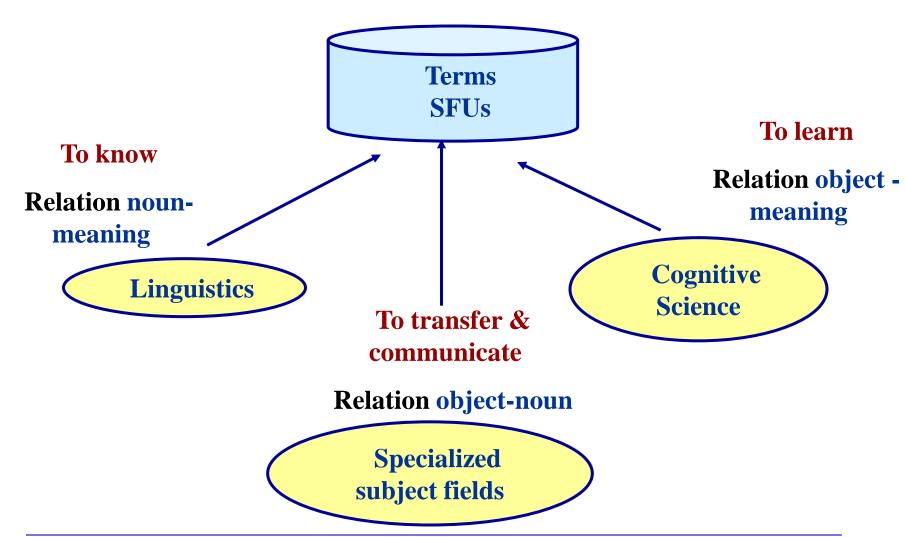
- Emphasis on terminology management:
  - products, tools and applications



#### How do we relate terms and concepts?



## The object of study of terminology



## **Definition of a concept**

- *In general*: Unit of knowledge created by a unique combination of characteristics. ISO 1087- 1 (2000)
- In terminology: Concepts shall be considered mental representation of objects within a specialized context or field ISO/DIS 704 (2009)
- BUT concepts are influenced by social and cultural circumstances given at a certain moment
  - SO this can lead to different classifications in the conceptual system
- Concepts can be seen as:
  - Units of **thought**: represent and recognize the object mentally
  - Units of **knowledge**: represent knowledge in each subject field
  - Units of communication: transmit knowledge by means of linguistic symbols



## **Description of a concept**

• Concepts are described according to their common features, properties or characteristics, either by **intension** or **extension** 

#### Intension

- Set of characteristics which makes up the concept (ISO 1087-1: 2000)
- The bigger the number of common characteristics, the more restricted is the intension.
- The intension of the concept winter in polar countries includes: low temperatures, ice, wind, snow, etc.

#### Extension

- Totality of objects to which a concept corresponds (ISO 1087-1: 2000)
- A general concept has a wide extension as it includes two or more objects by reason of common properties.
- The extension of the concept planet includes: *Mercury, Venus, Earth, Mars, Jupiter, Saturn, Uranus, Neptune and Pluto.*



## Characteristics of a concept

- "Abstraction of a property of an object or of a set of objects" (ISO 1087-1:2000)
- According to the importance in forming a concept
  - essential: indispensable to understand and distinguish a concept
    - The back of a seat distinguishes a stool and a chair.
  - complementary: colour, material, shape, ...



- **intrinsic,** which are observable properties:
  - Shape: oval, round, narrow, wide, ...
  - Material: wooden, stone, metalic, ...
  - Colour: red, blue, green, orange...
  - **Position:** vertical, hanging, slanting
- extrinsic, relation of the object with others
  - **Parts** that make up the object: *ball*, *wire*...
  - Mode of employement or application: analogic, digital, hybrid.
  - **Origin** or how an object comes into existence: producer, inventor, provider, the place of its production, (town, country), ...





## **Concept:** abstraction based on the set of all the characteristics of mechanical mice

#### **Term**: mechanical mouse



- a device;
- ivory-coloured;
- hand-manoeuvred along a firm, flat surface;
- has a ball on its underside;
- has three buttons:
- has a wire for connecting to a computer;
- rollers detect the movement of the ball;
- the ball controls the movement of a cursor on a computer display screen.



- a device;
- blue and grey;
- hand-manoeuvred along a firm, flat surface;
- has a ball on its underside;
- has two buttons:
- has a wire for connecting to a computer;
- without rollers;
- the ball controls the movement of a cursor on a computer display screen.



- a device;
- black-grey;
- hand-manoeuvred along a firm, flat surface;
- has a ball on its underside:
- has two buttons:
- has a wire for connecting to a computer;
- rollers detect the movement of the ball;
- the ball controls the movement of a cursor on a computer display screen.

ISO FDIS 704:2009



# Relations between concepts: hierarchical relations

Close relation between a concept and its characteristics

#### A. GENERIC RELATIONS (genus-species relation) IS\_A

- One of the concepts includes another concept
  - **vertical:** hypernym- hyponym; superordinate –subordinate
  - **horizontal:** two specific ideas of the same generic concept with some distinguishing characteristics



- Broader than (BT)
- Narrower than (NT)
- Associated to (AT)

#### **B. PARTITIVE RELATIONS (part-whole relation) PART\_OF**

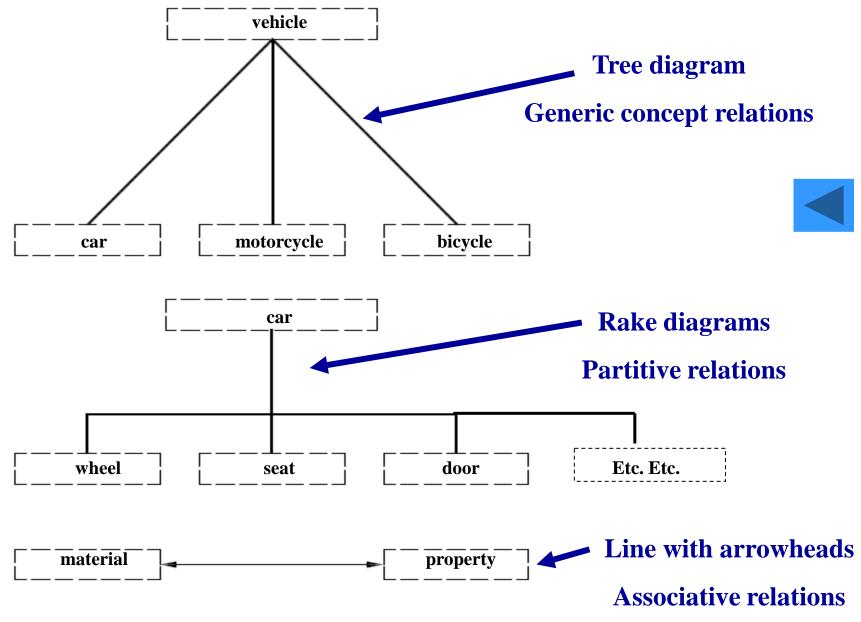
- These relations are also called meronimic (HAS\_PART)
  - Car: wheels, seats, doors, boot, stearing wheel, gearbox...
- Different types of meronimic relations







#### **Graphic representations used in ISO Standards**





#### **Other Meronimic Relations**

Relation	Example
component-object	pedal-bicycle
member-collection	ship-fleet
count-mass	slice (of bread)-bread
material-object	steel-car
phase-activity	pay-buy
place-area	oasis-desert

Winston et al. 1987



Climent, S. 1999 Individuación e información parte-todo. Representación para el procesamiento computacional del lenguaje



# Non-hierarchical relations (associative relations)

- Caused by : (acid rain- nuclear explosion)
- **Product of**: (paper- wood pulp)
- Property of (compressibility -gas)
- Quantitative measure (temperature-heat)
- Instrument for (computer- data processing)
- Counter-agent for (insecticide- insects)
- Container of (toolbox- tools)
- Method of (diamond drilling- drilling)
- Material for (iron-bridge building)
- *Place for* (coal mine- coal exploitation)
- -Associated with (production-consumption)

An associative relation exists when a thematic connection can be established between concepts by virtue of experience.



## How do we express concepts?

- In natural language:
  - Terms (one-word or multi-word terms) that denote or refer to a concept in a subject field
  - Definitions
  - Glosses, etc.
- In artificial language
  - Codes
  - Formulas
- In a multimedia resource:
  - Icons
  - Photos
  - Diagrams
  - Graphs
  - Video-clips
  - Audio-clips
  - Other multimedia representations



## Water in different concept systems

#### **Chemistry**

• molecule-composed compound of two atoms of hydrogen and one atom of oxygen (H2O), considered the universal solvent

#### **Physics**

• fluid which is colourless, odourless and tasteless used as the standard of specific gravity and of specific heat which freezes at 0 °C and boils at 100 °C

#### **Physics**

• chemical compound which is colourless, odourless, and tasteless and whose formula is H2O and which is naturally found in solid state at temperatures at and below 0 °C, in liquid state at temperatures between 0 °C and 100 °C, and as vapour at temperatures above 100 °C

### **Biology**

• chemical substance that is essential to all known forms of life

#### **Metrology**

 chemical compound whose freezing and boiling points are the basis for the Celsius temperature scale, where freezing point equals 0 °C and boiling point equals 100 °C at standard atmospheric pressure

#### **Astrology**

• one of the four elements of life that is associated with the emotional and intuitive processes



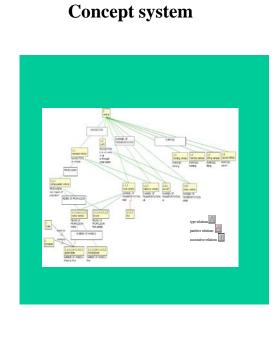
### **Concept systems**

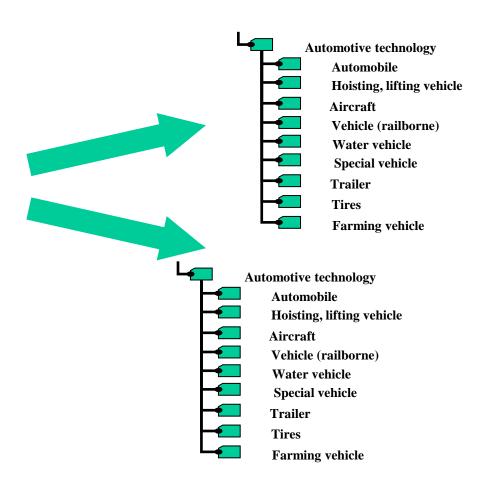
- Model concepts and relations between them based on specialized knowledge of a subject field;
- clarify the relations between *concepts*;
- form the basis for a uniform and standardized terminology;
- facilitate the comparative analysis of concepts and designations across languages and across subject fields;
- facilitate the writing of *definitions*;
- facilitate the inclusion of all relevant *concepts while developing* a terminological resource.



#### **Classification & concepts**

#### **Classification system(s)**

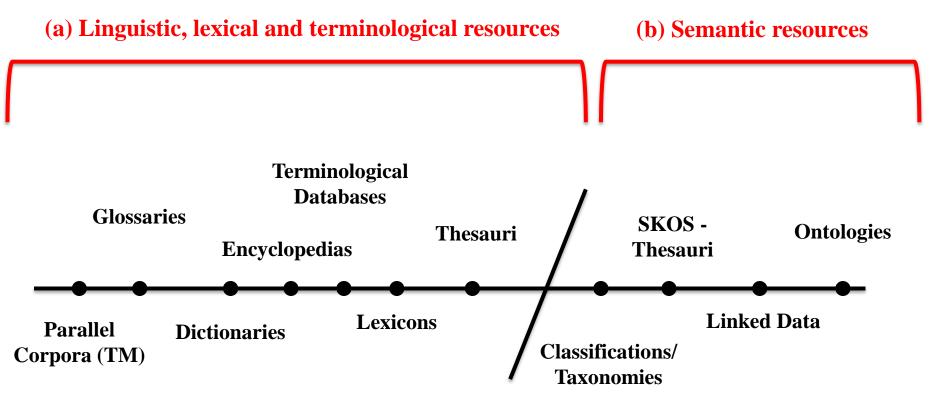




Taken from Reinhard Nerke ISO TC SC3, 22274



### Linguistic and conceptual resources

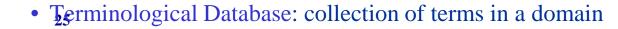


(inspired by Lassila & McGuinness' Ontology Spectrum)



# (a) Linguistic, lexical and terminological resources

- Parallel Corpora (Translation Memories TM): texts that are translations of each other (as opposed to *comparable corpora*, texts in two languages that are similar in content). Parallel corpora can be better exploited by TM, which align translation equivalents.
- **Glossary**: alphabetical list of terms in a particular domain of knowledge with the definitions for those terms.
- (Machine-Readable) Dictionary: alphabetically arranged words along with information about their forms, pronunciations, functions, etymologies, meanings, and syntactical and idiomatic uses.
- (On-line) Encyclopedia: like dictionaries, but with the focus on factual information
- (Computational) Lexicon or lexical database: a resource that organizes lexical units in lexical domains making use of lexico-semantic relations (hyponymy-hypernomymy, meronymy, synonymy, antonymy...)





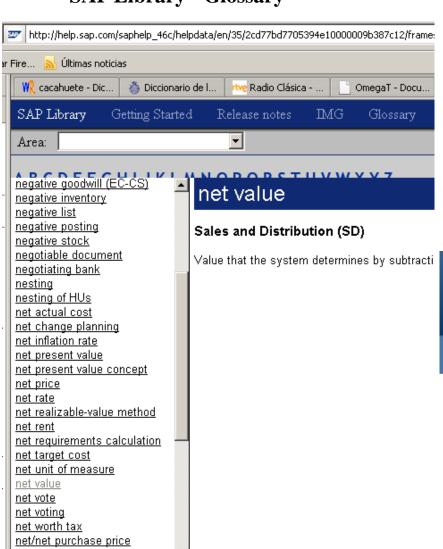
### (b) Semantic resources

- (SKOS) Thesaurus: a controlled list of descriptors (preferred terms) and nondescriptors (non-preferred terms) related by hierarchical, associative or equivalence links
- Classification/taxonomy: systematic arrangement of concepts in groups or categories according to established criteria
- Linked Data (RDFS): method of exposing, sharing, and connecting data via dereferenceable URIs on the Web
- Ontology (OWL): formal specification of a conceptualization



## **Examples of Glossaries**

#### **SAP Library - Glossary**



http://www.duke.edu/~charvey/Classes/wpg/bfglosa.htm

#### **Hypertextual Finance Glossary**

#### **ABCDEFGHIJKLMNOPQRSTUV**

A Fifth letter of a Nasdaq stock symbol specifying Class A shares.

AAII
See: American Association of Individual Investors

ABO
See: Accumulated Benefit Obligation

ABS
See: Automated Bond System

http://www.finance-glossary.com/

## financeglossary

The online dictionary of financial terms

Search

#### Lease

A contract in which the legal owner of property or other asset agrees to another person using that property or asset in return for a regular specified payment (known as rent) over a set term. In addition to buildings, other items such as cars and computers are often leased in order to avoid capital costs in the running of a business.

Related Terms...

Rent



netting (TR-TM)

netting (TR-TM)

#### Wiktionary ['wik∫ənri] n., a wiki-based Open Content dictionary

Contents [hide]
1 English
1.1 Noun
1.1.1 Translations
1.1.2 See also

#### English

Noun

#### income statement (plural income statements)

1. (accounting) A calculation which shows the profit or loss of an accounting unit (company, municipality, foundation etc.) (specific period of time, providing a summary of how the profit or loss is calculated from gross revenue and expenses.

#### income statement

[edit]

Wikip

Wikipedia has an article or 
Income statement

Definitions

 Disambiguation according to PoS

- PoS
- Translations (+ domain info)
- Etimological information
- Lexical relations (synonymy, antonymy...)
- Derived terms
- Related terms
- Illustrations

 Translations

 financial statement of net income
 [hide ▲]

 • Danish: resultatopgørelse (da) c
 • Italian: conto economico (it) m

 • Finnish: tuloslaskelma (fi)
 • Portuguese: demonstração do resultado do exercício (pt) m

 • French: compte de résultat (ff) m
 • Spanish: estado de resultados (es) m

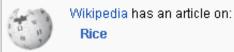
 • German: Gewinn- und Verlustrechnung (de) f
 • Swedish: resultaträkning (sv) c



#### **Entry for Rice**

#### Etymology

[edit]



Old French  $ris < \text{Italian } riso < \text{Latin } oriza < \text{Ancient Greek ŏρυζα } (óruza) < Old Persian; compare Sanskrit वीहि <math>(vr\bar{t}hi)$ , Tamil  $\mathfrak{A}ph$  (aricy).

#### Pronunciation

IPA: /Jats/





Audio (UK) (file)

Rhymes: -ais

[edit]

[edit]



Noun

#### rice (uncountable)

- 1. Cereal plants (Oryza sativa) of the grass family whose seeds are used as food.
- 2. The seeds of this plant used as food.

#### **Translations**

plants

[show ▼]

[show ▼]

[edit]

#### seeds used as food

The translations below need to be checked and inserted above into the appropriate translation tables, removing any numbers. Numbers do not necessarily match those in definitions. See instructions at Help:How to check translations.

- Hausa: àlbárkà kumbe
   Sinhalese: ವಿಜ್ (bat)
- Volapük: risat



#### **Entry for Rice**

Translations [edit]

```
plants
                                                                                                                                                              [hide ▲]

    Marathi: भात (bhāt)

    Afrikaans: rys (af)

    Mongolian: тутрага (tutraga)

    Albanian: oriz m

    Navajo: alóós

    Arabic: أرز (ar) ('arz, 'áruzz) شرز (ar) (ruzz) m

    Nepali: please add this translation if you can

    Armenian: pphົບດັ (hy) (brindz).

 Norwegian: ris (no) m

    Basque: arroz, irris

    Oriya: please add this translation if you can

    Bengali: ভাভ (bn)

    Persian: برنیج (fa)

 Bosnian: riža (bs) f, pirinač (bs) m

 Polish: ryż<sup>(pl)</sup> m

    Burmese: || || || || || || || (my) (htămin:)

    Portuguese: arroz (pt) m

    Catalan: arròs (ca) m

    Punjabi: ਚਾਵਲ (ćāval), ਚੌਲ (ćaul)

    Chinese:

    Rohingya: soil.

        Mandarin: 稻 (cmn) (dào)

    Romanian: orez (ro)

    Croatian: ríža (hr) f.

    Russian: рис (ru) (ris) m

 Danish: ris (da)

    Sanskrit: ताणड्ल

 Dutch: rijst (nl) m

    Serbian:

    Estonian: riis (et)

                                                                                             Cyrillic: рижа f, пиринач m
                                                                                             Roman: riža f, pirinač m

    Finnish: riisi (fi), riisikasvi (fi)

                                                                                     • Shan: 10 10 10 10 (khaw3)

 French: riz (ff) m.
```



#### **Entry for Rice**

Derived terms [edit] · arborio rice rice bowl ricebird brown rice. rice grass · rice-paper plant golden rice rice leafhopper ricer Indian rice · rice paper Spanish rice jollof rice · rice pudding · sticky rice · mealie rice rice rat · white rice Patna rice rice weevil · wild rice · brewer's rice broken rice Related terms [edit] Menominee rijsttafel risotto See also [edit] jambalaya basmati. paella khir bhelpuri · pilaf pilau California roll. mirin rangoli dosa mochi risotto gumbo nasi goreng sake idli onigiri samshu

pad thai



sushi

idli





¹lease ◀)) noun \'lēs\	
Definition of LEASE	
<ul> <li>1 : a contract by which one conveys real estate, equipment, or facilities for a specified term and for a specified rent; also : the act of such conveyance or the term for which it is made</li> </ul>	
2 : a piece of land or property that is leased	
3 : a continuance or opportunity for continuance <a lease="" life="" new="" on=""></a>	
Examples of LEASE	
<ul> <li>They took out a five-year lease on the house.</li> </ul>	
<ul> <li>We hold leases on both of our cars.</li> </ul>	
Origin of LEASE	
Middle English les, from Anglo-French, from lesser	

- Disambiguation according to PoS
- PoS
- Definitions
- Etimological information
- Illustrations
- Examples of use



## **Example of DBPedia (Encyclopaedia)**



- Definition
- •Comparable texts (not parallel!!)

#### About: Finance lease

An Entity of Type: unknown, from Named Graph: http://dbpedia.org, within Data Space: dbpedia.org

A finance lease or capital lease is a type of lease.



## **Examples of DBPedia (Encyclopaedia)**



- •Semantic relations
- •Labels in different languages
- •Links to other resources

dbpprop:hasPhotoCollection	or http://www4.wiwiss.fu-berlin.de/flickrwrappr/photos/Finance_lease
dbpprop:reference	<ul> <li>http://www.aasb.com.au/public_docs/aasb_standards_2005/compilations/AASB117_07-04</li> <li>http://www.fasb.org/pdf/fas13.pdf</li> <li>http://www.investorwords.com/722/capital_lease.html</li> </ul>
rdfs:comment	<ul> <li>Finansielle leasing er en leasingform der det er leasingtakeren som har den finansielle risik praksis et selskap muligheten til å finansiere en eiendel selv om man strengt tatt aldri kjøp leasingforhold gi leasingtageren kontroll over eiendelen i store deler av dennes levetid. Leas samme fordeler (og ulemper) som en eier. Se også operasjonell leasing.</li> <li>Operasjonell leasing er en leasingform der det er utleieren som har den finansielle risikoen. kontraktens løpetid, men påtar seg ingen forpliktelser utover å betale den avtalte leasingleis føres som en kostnad og objektet vil ikke aktiveres i leasingtakerens regnskaper. Se også</li> <li>Beim Finanzierungsleasing als typischem Leasing überwälzt der Leasinggeber das Investit Der Geber trägt somit nur das Kreditrisiko und eventuell vereinbarte Dienstleistungen. Der I Vertragslaufzeit nicht dinglicher Eigentümer des Vermögensgegenstandes, wenn ihm die S zugerechnet werden kann, da der Leasing-Geber kein Interesse an einem Rückerhalt des E</li> <li>A finance lease or capital lease is a type of lease.</li> </ul>
rdfs:label	<ul> <li>Finanzierungsleasing</li> <li>Finansiell leasing</li> <li>Operasjonell leasing</li> <li>Finance lease</li> </ul>
owl:sameAs	■ fbase:Finance lease
skos: <b>subject</b>	<ul> <li>category:Financial_terminology</li> <li>category:Business_law</li> </ul>
foaf:page	http://en.wikipedia.org/wiki/Finance_lease
is dbpprop:redirect of	<ul> <li>dbpedia:Capital_lease</li> <li>dbpedia:Finance_leasing</li> </ul>
is foaf:primaryTopic of	http://en.wikipedia.org/wiki/Finance_lease



## **Example of Computational Lexicon**

**PRINCETON** UNIVERSITY

### WordNet

## •Diambiguation according to PoS

#### Noun

- •Definition
- Synsets
- •Examples of use
- •Lexical relations (hyperonymy-hyponymy, meronymy, synonymy, antonymy...)

- S: (n) lease, rental, letting (property that is leased or rented out or let)
  - o direct hyponym I full hyponym
    - S: (n) car rental, hire car, rent-a-car, self-drive, u-drive, you-drive (a rented car) "she picked up a hire ca hotel"
    - S: (n) sublease, sublet (a lease from one lessee to another)
  - direct hypernym I inherited hypernym I sister term
    - S: (n) property, belongings, holding (something owned; any tangible or intangible possession that is owned)
       property"; "he is a man of property";
  - o derivationally related form
    - W: (v) lease [Related to: lease] (grant use or occupation of under a term of contract) "I am leasing my co:
    - W: (v) lease [Related to: lease] (let for money) "We rented our apartment to friends while we were abr
    - W: (adj) rental [Related to: rental] (available to rent or lease) "a rental car"
- S: (n) lease (a contract granting use or occupation of property during a specified time for a specified payment)
- S: (n) lease, term of a contract (the period of time during which a contract conveying property to a person is in effect)



### WordNet 3.0 Vocabulary Helper

pollution	Search
-----------	--------

Help for Eva Word Lookup Interfaces

#### Synonyms/Hypernyms (Ordered by Estimated Frequency) of noun pollution

3 senses of pollution

#### Sense 1

pollution -- (undesirable state of the natural environment being contaminated with harmful substances as a consequence of human activities)

- environmental condition -- (the state of the environment)
- impurity, impureness -- (the condition of being impure)

#### Sense 2

befoulment, defilement, pollution -- (the state of being polluted)

• dirtiness, uncleanness -- (the state of being unsanitary)

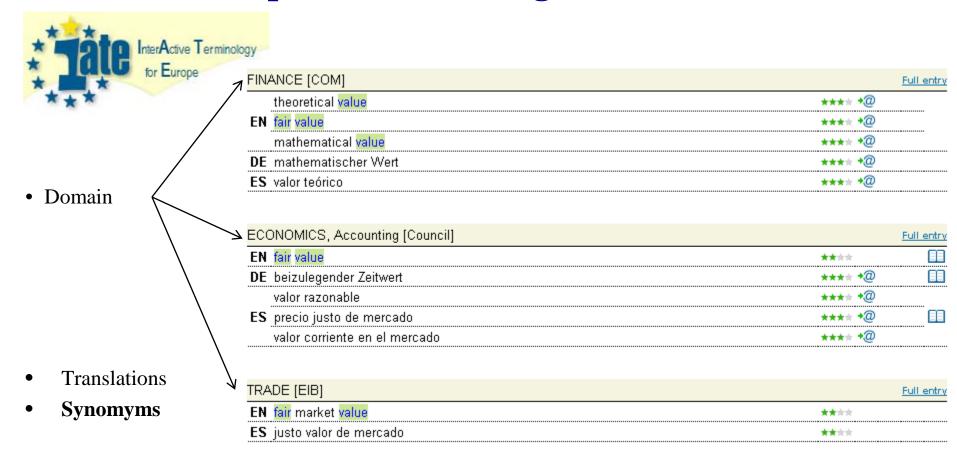
#### Sense 3

contamination, pollution -- (the act of contaminating or polluting, including (either intentionally or accidentally) unwanted substances or factors)

- soiling, soilure, dirtying -- (the act of soiling something)
- 1. (1) pollution -- (undesirable state of the natural environment being contaminated with harmful substances as a consequence of human activities)
- 2. befoulment, defilement, pollution -- (the state of being polluted)
- 3. contamination, pollution -- (the act of contaminating or polluting; including (either intentionally or accidentally) unwanted substances or factors)



## **Example of Terminological Database**





## **Example of Terminological Database**



en

ECONOMICS, Accounting

Definition "In International Accounting Standards, the amount for which an asset could be exchanged

between a knowledgeable, willing buyer and a knowledgeable, willing seller in an arm's-length

transaction."

Dictionary of Accounting Terms, Derek French Definition Ref.

fair value Тегт

> Reliability 2 (Minimum reliability)

Date: 03/11/1997

Definitions

•Reliablity ratio

 Date of inclusion in termbase

•Sources of reference for terms. definitions, etc.

de

Definition Betrag, zu dem ein Vermögenswert zwischen sachverständigen, vertragswilligen und voneinander

unabhängigen Gzschäftspartnern getauscht werden könnte.

Note CONTEXT: Rechnungswesen; MISC: Zentraler Begriff des International Accounting Standard (IAS)

> 39 und des RL-Vorschlags zur Änderung der RL 78/660/EWG und 83/349/EWG über die im Jahresabschluß bzw. im konsolidierten Abschluß von Gesellschaften bestimmter Rechtsformen zulässigen Wertansätze (KOM (2000) 80 endg.); der b.Z. soll insb. die korrekte Bewertung und Bilanzierung derivativer Finanzinstrumente sicherstellen; für diese Instrumente entspricht der b.Z.

häufig dem Marktwert, nicht jedoch z.B. dem Anschaffungswert (= Kaufpreis);DIV: RSZ

16/01/2001

beizulegender Zeitwert Тегт

Reliability 3 (Reliable)

Term Ref. International Auccounting Standards 1999 (Deutsche Ausgabe) über: Deutsches

Rechnungslegungs Standards Committee e.V. (DRSC)

in anderen Kontexten als Rechnungswesen ist u.U. "angemessener Wert" auch korrekt; ggf. kann bei der ersten Nennung der EN Ausdruck "fair value" in Klammern hinzugesetzt werden

Date 03/11/1997

Language Usage



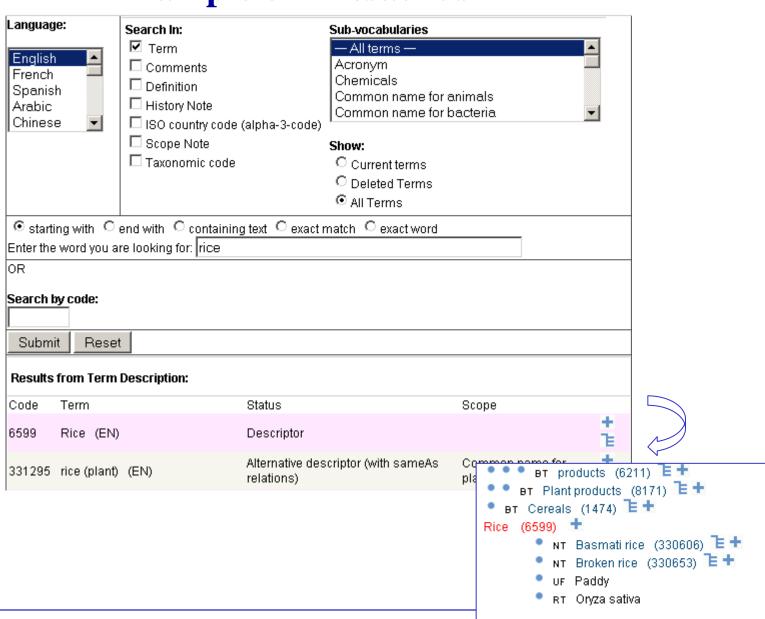
## **Example of Thesaurus**

#### **AGROVOC**

•Fuzzy relations:

Broader than
Narrower than
Used for
Related to

•Info about preferred terms



## An excerpt from INSPEC Thesaurus

Cellular radio	used for (UF): cellular communication cellular telephones Groupe Speciale Mobile (GSM) microcellular radio pan-european radio vodafone
Land mobile radio Radiotelephony	These twp terms are broader terms (BT) to "cellular radio". If you searched under these terms, you will retrieve a larger set of documents
Radio applications Telecommunication	These terms are top terms (TT) in the hierarchy
Channel allocation Land mobile radio Personal communication networks Radio access networking Space division multiple access	All these terms are related terms (RT) to "cellular radio"
DI January 1985	Date when "cellular radio" was added
mobile radio systems	previous term (PT) used before 1985
B6250F; D4045	class codes

Table 1: Inspec Thesaurus (1999) -- an excerpt on cellular phones



# Reusing and sharing KOS in the Semantic Web: SKOS

- Simple Knowledge Organization System
  - Simple, flexible, extensible, machine-understandable representation for sharing KOS
  - Goal: to enable easy publication of controlled structured vocabularies for the semantic web
    - Thesauri
    - Classification schemes
    - Subject heading systems
    - Taxonomies
    - Other 'controlled language'
  - How: by using a common data model for sharing and linking knowledge organization systems
  - BUT SKOS is **not** a formal knowledge representation language.

Many exist and are in use in cultural heritage, medicine, libraries, ...



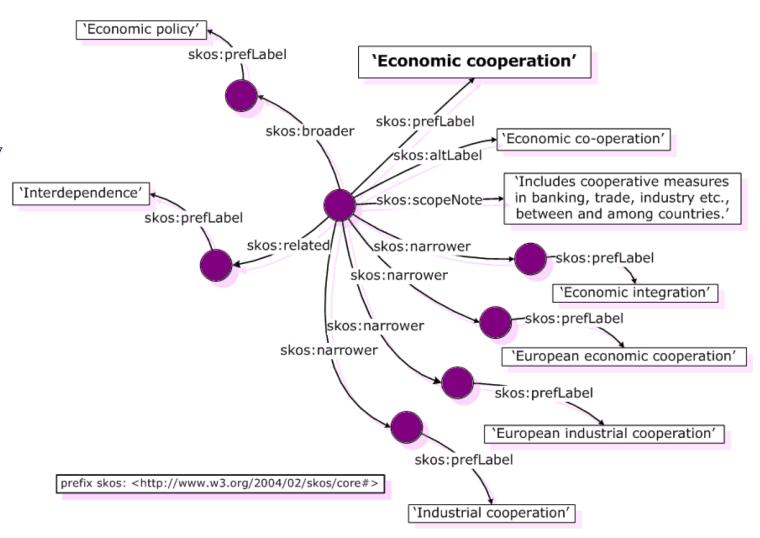
#### **Elements in SKOS**

Semantic
 Relationships
 Broader/Narrower
 Terms

**Related Terms** 

• Lexical Labels
Preferred, alternativ
and hidden labels

 Additional documentation Notes, comments, descriptions

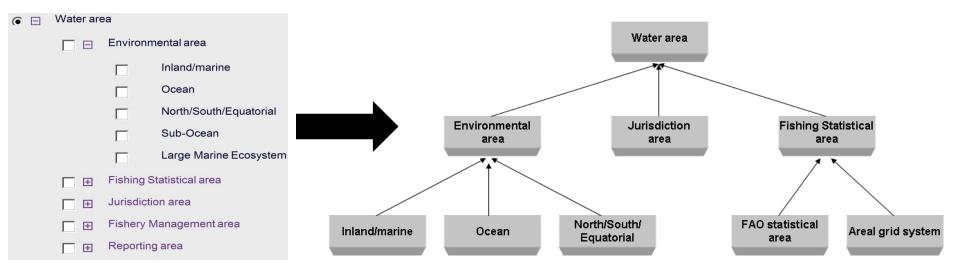


SKOS concept schemes are not formal ontologies!!!!



#### **Classification Scheme**

• A classification scheme<sup>1</sup> is the descriptive information for an arrangement or division of objects into groups based on characteristics, which the objects have in common. E.g. water area classification scheme<sup>2</sup>.



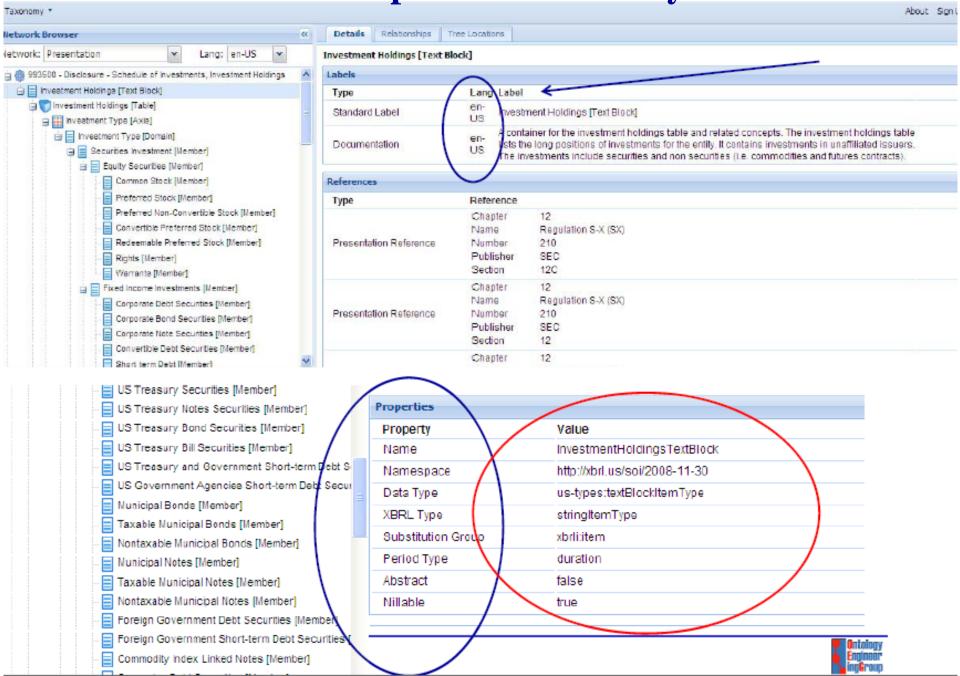


<sup>1.</sup> International Standard Organization (ISO). Information technology - Metadata registries - Part 1: Framework, 2004. Report ISO/IEC FDIS 11179-1.



<sup>2.</sup> http://www.fao.org/figis/servlet/RefServlet

**Example of a Taxonomy** 



### **Example of Ontology**

