Terminologies and other Knowledge Organization Systems

> Gerhard Budin University of Vienna

TSS July 2011, Cologne

Motivation and Purpose – Knowledge Organization Systems

- In this unit of TSS '11, we focus on the role of terminologies as tools to <u>organize</u> and <u>retrieve</u> knowledge (-> Knowledge Organization Systems, KOS)
- Major types of KOS are:
 - Thesauri (in information science: controlled vocabularies for indexing and information retrieval)
 - Classification systems (hierarchical concept systems, usually domain-specific, sometimes universal in scope)
 - taxonomies, nomenclatures (mostly in natural sciences, systematic arrangements of terms seen as "scientific names"
 - Ontologies (in IT -> formal conceptual shared specifications, domain ontologies often created by formalizing the previously listed types of KOS

Knowledge (organization) systems

- Cognitive knowledge systems
- collective knowledge systems, cultural systems, social systems, language and communication systems
- Formal knowledge systems, knowledge representation systems, "semantic systems" (Semantic Web)
- Applications:
 - Knowledge organization as part of knowledge management (Nonaka, Takeuchi, et al)
 - Knowledge organization as daily practice in libraries and information systems (for more than 2000 years)
 - Knowledge organization as formal representations in collective knowledge systems -> Semantic Web applications

What is knowledge organization?

- 1. A part of information and library science, a part of philosophy of science and of epistemology, but also of knowledge management and knowledge engineering
 - Investigating and representing structures of knowledge
 - Epistemological aspects, cognitive science aspects
 - Linguistic and socio-cultural aspects (e.g. folk taxonomies)
 - Historical aspects (e.g. Leibniz, encyclopedism, administrative categorizations in ancient societies, history of science, etc.)
- 2. Practical work: creating and using knowledge organization systems (see further down)
- Knowledge organization is also a crucial process in linguistic action (sprachliches Handeln) – Text organization both in reception and production

Functions of knowledge organization systems

- 1. Instruments of structuring and archiving the content of large scale collections
- 2. Structural components of information systems
- 3. Support of targeted retrieval of information based on conceptual search criteria
- 4. Search aids, visual navigation, query languages
- 5. Communication support tools (cross-lingual, crossdisciplinary, cross-cultural)
- 6. Instruments of corporate knowledge management
- 7. Learning support, orientation support, didactic tools

Properties of knowledge organization systems

- 1. Conceptual structures (hierarchical and nonhierarchical structures)
- 2. Explicitation of conceptual links, definitions (mono- or multilingual)
- 3. Terminological and linguistic standardization
- 4. Increasingly formalized and digital (in particular as "ontologies")
- 5. Different scales (from small KOS to large ones (more than 200.000 concepts)
- 6. Increasingly with visualized structures, interactive user interfaces
- 7. Static or *dynamic* (e.g. ontologies for modelling business processes in companies)

"Ontologies" as formal knowledge systems

- Computer science: From Ontology as a traditional field of philosophy (theory of being, existence, theory of objects, etc.) to formal, digitally represented concept systems/ knowledge systems
- Concepts are explicitly defined terms are assigned
- Relations between concepts are explicitated
- Terms are standardized
- Logical application rules and constraints are specified
- Ontologies as knowledge representation systems

Domain-specific knowledge organization systems

- Medicine, health, bio- and life sciences
- Business, trade
- Industry, engineering
- Natural sciences
- Administration, government
- Culture
- Pedagogy
- Linguistics
- Etc.

Thesaurus Example: EUROVOC

- Thesaurus of the European Union
- Using a ppt presentation by Chr. Laaboudi-Spoiden at the 2010 EUROVOC conference:

General and multidisciplinary

- 21 fields (domains)
 - Two-digit notation + heading
- 127 microthesauri
 - Four-digit notation + heading

Covers both European Union and national point of view

 with emphasis on parliamentary activities

O4 POLITICS 0406 political framework 0411 political party 0416 electoral procedure and voting 0421 parliament 0426 parliamentary proceedings 0431 politics and public safety 0436 executive power and public service

0421 parliament

```
composition of parliament
   NT1 bureau of parliament
                     RT Bureau of the EP [1006]
      NT2 Deputy Speaker of Parliament
                     RT Vice-President of the EP [1006]
      NT2 Speaker of Parliament
                     RT President of the EP [1006]
      NT2 treasurer
                     RT Quaestor of the EP [1006]
   NT1 conference of presidents
   NT1 non-attached member
   NT1 parliamentary chamber
      NT2 bicameral system
           NT3 directly-elected chamber
                     RT universal suffrage [0416]
          NT3 federal chamber
          NT3 Upper House
      NT2 unicameral system
   NT1 parliamentary committee
                     RT committee report [0426]
                     RT EP Committee [1006]
      NT2 ad hoc committee
      NT2 committee of inquiry
                     RT parliamentary inquiry [0421]
                     RT parliamentary scrutiny [0421]
                     RT power of assessment [0406]
                     RT supervisory power [0406]
```

EuroVoc - Multilingual content

- 22 official languages
- 2 other languages
 - Croatian and Serbian



EuroVoc – Language equivalence

Set of 6 797 concepts

- Preferred Terms
 - 24 language equivalences
- Non Preferred Terms
 - Language-dependent

Symmetric equivalence

- Preferred Terms
- Relationships (BT/NT, RT)

LANGUAGE EQUIVALENTS

- BG избирателен мандат
- ES mandato electoral
- CS poslanecký mandát
- DA mandat
- DE Wahlmandat
- ET valitav ametikoht
- EL αιρετό αξίωμα
- EN elective office
- FR mandat électif
- IT mandato elettivo
- LV vēlēts amats
- LT renkamosios pareigos
- HU képviselői mandátum
- MT mandat elettorali
- NL mandaat
- PL mandat przedstawicielski
- PT mandato electivo
- RO mandat electoral
- SK volený mandát

mandaat

ffice				
ilsory mandate nandate al mandate ng cabinet ng government ng member nentary mandai entative manda resignation of an elect resigning member term of office	UF avga avga bun fritt loka mar nati riks	UF aftreden van een gekozen vertegenwoordiger demissionair kabinet demissionair lid duur van het mandaat einde van de ambtstermijn lokaal mandaat nationaal mandaat nationaal mandaat representatief mandaat verplicht mandaat ionellt mandat		

© Laaboudi-Spoiden 2010

04 POLITIKA MT 0421 parlamentas BT1 parlamento narys NT1 daugiapareigystė

RT parlamento paleidimas [0431] rinkimai [0416]

EuroVoc – Previous Data Model

- One ID number per descriptors
 - 22 language equivalents
 - Standard thesaurus relationships
 - BT/NT (Broader/Narrower)
 - RT (Associative)
- Attributes
 - Non Descriptors
 - Scope Notes
 - Share the descriptor ID
 - → "False" equivalence (USE/UF)

Descriptor

```
<RECORD>
```

<DESCRIPTOR_ID>5482</DESCRIPTEUR_ID> <LIBELLE>climate change</LIBELLE> </RECORD>

Non Descriptor

```
<RECORD>
<DESCRIPTEUR_ID>5482</DESCRIPTEUR_ID>
<UF>
<UF_EL>global warming</UF_EL>
<UF_EL>climatic change</UF_EL>
</UF>
</RECORD>
```

EuroVoc – Previous Data Model

Déclinaison						español (es)	¥
Descripteur	climate change			Déclinaison		000000000	
UE	alimatic abanas	Déclinaison		Decimalson			
UF	climatic change	Descripteur	klim	Descripteur	cambio climático		
	global warming	UF	opva	UF	modificación del clima		
SN		SN			recalentamiento del clima		
PERM	change climate	PERM	klim	a, ^{SN}			
	entange, entrate		ænd	Iri PERM	clima, modificación del		
	change, climatic	Polations			clima, recalentamiento del		
	warming, global	MT 5211	natu	ırli	climático, cambio		
Relations		BT1	klim	a Relations			
MT 5211	natural environment	RT	drivh	MT 5211	medio natural		
DT4	- E et		klim	atBT1	clima		
DII	climate			RT	climatología		
RT	climatology	у			efecto invernadero		
	greenhouse effect						

Standards

- ISO 2788-1986: monolingual (1986)
- ISO 5964-1985: multilingual (1985)

© Laaboudi-Spoiden 2010

Descriptor - Non Descriptor

Revised by

→ ISO 25964: Thesauri and interoperability with other vocabularies Preferred Term - Non Preferred Term

New Data Model - Concepts

EuroVoc



New Data Model – Thesaurus Terms



EuroVoc

New Data Model – Thesaurus Terms

URI= <u>http://eurovoc.europa.eu/125206</u>

CAMBIO CLIMÁTICO

PREFERRED TERM



URI= <u>http://eurovoc.europa.eu/125207</u>

17

EuroVoc

Data Model – Summary

climate change

BT1 climate RT climatology [3606] greenhouse effect [5216]

global warming
52 ENVTRONMENT

5211 natural environment

- Concept (multilingual)
 - Concept Terms relation
 - Standard relationships (BT/NT, RT)
 - Relations to the Microthesaurus
 Group of concepts
- Thesaurus terms (language-dependent)
 - Lexical representation of a concept in a given language
 - 2 types of term
 - Preferred term (PT)
 - Non preferred term (NPT)
 - Equivalence relation between PT/NPT (USE/UF)
- What's new ?
 - One URI by concept, term

LANG	SUAGE EQUIVALENTS	
BG ES DA DE ET EL ET EL FR TT LV LT U MT NL PT RO SK SL SV	промяна на климата cambio climático změna klimatu klimaændring Klimaveränderung klimamuutus kλuorikň oλλovň climate change changement climatique cambiamento climatique cambiamento climatique cambiamento climatique cambiamento climatique cambiamento climatique cambiamento climatique cambiamento climatique cambiamento climatique shimats verandering zmiany klimatyczne alteração climática schimbare climatiçã zmena podnebia sprememba podnebja ilmastonmuutos klimatforândring	
HR SR	klimatska promjena промена климе	
climatic change		

UF climatic change

EuroVoc Website

Top menu



Left menu



EuroVoc - Architecture



EuroVoc website – Browse the subject-oriented version

Domains

 O4 POLITICS 0406 political framework 0411 political party 0416 electoral procedure and voting 	CULLAPSE ALL
0421 parliament 0426 parliamentary proceedings 0431 politics and public safety 0436 executive power and public service	
 08 INTERNATIONAL RELATIONS 0806 international affairs 0811 cooperation policy 0816 international balance 0821 defence 	
 D EUROPEAN COMMUNITIES 1006 Community institutions and European civil service 1011 European Union law 1016 European construction 	Browse
 1021 Community finance 12 LAW 1206 sources and branches of the law 1211 civil law 1216 criminal law 1221 justice 1226 organisation of the legal system 1231 international law 1236 rights and freedoms 	 Browse the subject- oriented version

© Laaboudi-Spoiden 2010

EuroVoc website – Thesaurus concept



EuroVoc website – Term details

Preferred Term

national law

UF domestic law

Internal law legislation of the Member States national legal system national legislation national regulations

12 LAW

MT 1206 sources and branches of the law

BT1 legislation

BT2 source of law

RT approximation of laws [1011]

LANGUAGE EQUIVALENTS

- BG национално право
- ES Derecho nacional
- CS národní právo
- DA national ret
- DE nationales Recht
- ET riigisisene õigus
- EL εθνικό δίκαιο EN national law
- EN national law FR droit national
- IT diritto nazionale
- LV valsts tiesību akti
- LT šalies teisė
- HU nemzeti jog
- MT national law (under translation)
- NL nationaal recht
- PL prawo krajowe
- PT direito nacional
- RO drept național
- SK národné právo
- SL nacionalno pravo
- FI kansallinen oikeus
- SV nationell rätt
- HR nacionalno pravo
- SR право државе чланице

Non Preferred Term

legislation of the Member States

USE national law

© Laaboudi-Spoiden 2010

CONTRIBUTE

MAP

EuroVoc website – Map



EuroVoc 4.4

- Lisbon Treaty
 - EC \rightarrow EU
 - European Community → European Union
- Source of the proposals
 - The European Parliament (15%)
 - The national parliament libraries (SE, RO and SK)
 - Terminologists from EC DGT
 - EuroVoc users
 - The Publications Office
- Online publication: first semester 2011

powers of the EC Institutions powers of the EU Institutions EC competition EU competition

delegated decision non-legislative act legislative act implementing regulation special legislative procedure

Your proposals Contribute

 New approved concepts

Ongoing and future developments of EUROVOC

- Available in SKOS
- Visualization
- Online version
- Thesaurus mapping and alignment software support
- Update of content

Other major thesauri

- GEMET General European Multilingual Environmental Thesaurus
- UNESCO Thesaurus
- CEDEFOP Thesaurus (vocational training)
- AAT Art and Architecture Thesaurus
- AGROVOC Thesaurus (FAO)
- General trends:
 - Preparing for Semantic web applications
 - RDF, SKOS, Linked Data, ontologies
 - Networking/mapping/interoperability



🖉 Protege Ontology	Library - Protege Wiki - Windows Ir	Internet Explorer
🕗 🗢 < http	://protegewiki. stanford.edu /wiki/Protege_Ont	Intology_Library 💽 🚱 🔀 Google
Datei Bearbeiten Ar	nsicht Favoriten Extras ?	x 🇞 -
🚖 Favoriten 🛛 🚖 🧣	🖗 Observatoire de linguistique 🦙 APA Sty	Style Essentials 🧧 e-learning glossary. Instruc 😵 xml-intl - Home 🧧 Asia Online - The World Spe
┥ Protege Ontology Lib	orary - Protege Wiki	🛐 🔹 🔂 🔹 🖃 🖶 🔹 Seite 🗸 Sicherheit 👻 Extras 🗸 🔞
	page discussion view source h	Log in / create account
navigation Main Page	Welcome to the Protege Ontol This page is organized into the following OWL ontologies Frame-based ontologies Ontologies in other formats (e.g., DAN If your ontology is available in multiple for Please make insertions in alphabetical of	tology Library! ing groupings: IAML+OIL, RDF Schema, etc.) e formats, please feel free to link to it from multiple sections. al order. Thank you !!!
All Properties All Templates All Forms Ontology Library Plugin Library Multi-upload	OWL ontologies Information on how to open OWL files from th Getting Started with Protege-OWL Web pa Semantic Web search engine Swoogle 🗗	n the Protege-OWL editor 값 is available on the main Protege Web site. See the Creating and Loading Projects 얎 section of the page. Other ways to search for OWL ontologies include using Google: http://www.google.com/search?q=filetype:owl+owl 얇, or the
 Help search Go Search toolbox What links here Related changes Special pages Printable version Permanent link Browse properties 	 AIM@SHAPE Ontologies : Ontologies p Semantic-based systems for Handling, Acc amino-acid.owl : A small OWL ontology Basic Formal Ontology (BFO) : bhakti.owl : An OWL ontology for the tra Biochemical Ontologies : Over 30 ontol primitive skeletons and complex class exp BioPAX : An OWL ontology for biologic BIRNLex : An OWL ontology created by and f BIRN data related to multi-resolutional, or based OBO ontologies. Black Pepper Ontology : A black pepper AMAL JYOTHI College of Engineering 	s pertaining to digital shapes. Source: AIM@SHAPE NoE 🖗 - Advanced and Innovative Models And Tools for the development of koquiring, and Processing knowledge Embedded in multidimensional digital objects. gy of amino acids and their properties. Source: Amino Acid Ontology Web site 🖗. transcendental states of consciousness experienced by practitioners of bhakti-yoga, a form of Vedic consciousness engineering. tologies for knowledge representation and reasoning across scientific domains. Ontologies are normalized into non-disjoint xpressions for DL reasoning. Ontologies describe atoms, bonds, molecules, macromolecules and the processes they participate in. gical pathways, which is primarily used to exchange data between pathway resources. d for the Biomedical Informatics Research Network (BIRN) 🖗 in order to provide a shared semantic framework in which to annotate cross-species studies of neurodegenerative disease. BIRNLex follows the OBO Foundry best practices and makes use of other OWL- per ontology created for DataIndia.ORG 🖗 to provide support for farm automation and agro-processing. Joint project between and Dr. M.G.R. University 🚱.
	BreastCancerOntology @r: An OWL ontolo	stogy for describing some reatures of Breast Cander; Basic model is stable, but still being fieshed out.



For more information on ontologies, knowledge organization systems, on our projects mentioned above, on further reading, related tools, etc.

Please contact

Gerhard Budin University of Vienna Centre for Translation Studies gerhard.budin@univie.ac.at