

Content Interoperability

re-purposing structured content for AAC and eLearning
taking structured content as light-weight ontologies

Helmut Beckmann, Hochschule Heilbronn

Christian Galinski, Infoterm



Experiences from projects

- **IN-SAFETY** **Infrastructure and SAFETY" (2006–2008)**
- **ISO/CDB** **ISO Concept DataBase (as from 2009)**
- **ASK-IT** **Ambient Intelligence System of Agents for Knowledge-based and Integrated Services for Mobility Impaired users (2004–2009)**
- **OASIS** **Open architecture for Accessible Services Integration and Standardization (2008–2011)**
- **CEN/WS/eCAT** **Workshop "Multilingual eCataloguing and eClassification in eBusiness (as from 2002)**

→ global content interoperability

eProcurement 1

Product classification systems

eCl@ss

Class code:19010290

Preferred name: Notebook (unclassified)

Definition:

Keywords:

Laptop, Mini notebook, Sub-/Ultra-Portable-
Notebooks, Sub-Notebook, Subnotebook

Common Procurement
Vocabulary (CPV)

22815000-6 Notebooks



[AAE670001](#) - additional link address
[BAE864003](#) - Audio input present
[BAE862003](#) - Audio output present
[BAF128002](#) - Average elapsed time
[BAF478002](#) - Capacity
[BAF046003](#) - CD and DVD present
[AAD931001](#) - customs tariff number
[BAF068002](#) - Data transfer rate
[BAE156002](#) - Dimensions
[BAD335003](#) - Diskette drive present
[BAE830002](#) - display format
[BAF172002](#) - Extension rule
[BAF707003](#) - Flash storage installed
[BAA271005](#) - GTIN
[BAF239003](#) - Hard disk storage present
[BAA001003](#) - Manufacturer name
[BAD847003](#) - Manufacturer product number
[BAE926003](#) - Max. battery weight
[AAC860002](#) - Min. capacity of hard disc storage
[BAF734003](#) - Modem for communication present
[BAF752003](#) - Network present
[BAE604003](#) - Operating system present
[BAE717003](#) - optional connection peripheral device
[BAA316003](#) - Product name
[BAA002002](#) - Product type description
[BAG017003](#) - Service
[BAG053003](#) - Software included
[BAB542002](#) - Supplier name
[BAA059004](#) - Supplier product number
[BAE712003](#) - Total number of available connections
[BAF502003](#) - Type of communication
[BAE711002](#) - Type of connection
[BAF458002](#) - Type of integrated peripheral devices
[BAF580002](#) - Type of performance characteristics
[BAF141002](#) - Type of peripheral input device
[BAE636002](#) - Type of processor
[BAA261002](#) - Type of support
[BAG249003](#) - Video output present
[BAF340003](#) - With guarantee

eProcurement 2

Product classification systems

GPC

10001145 - Personal Computers – Portable

Attribute: 20001087 - If Wireless
Value: 30002960 - NO
Value: 30002518 - UNIDENTIFIED
Value: 30002654 - YES

Attribute: 20001148 - Screen Size (Inches)
Value: 30013229 - <14
Value: 30013231 - >16.9
Value: 30013230 - 14-16.9
Value: 30002515 - UNCLASSIFIED
Value: 30002518 - UNIDENTIFIED

Attribute: 20001053 - Type of Personal Computer – Portable
Value: 30014546 - NETBOOK
Value: 30007661 - NOTEBOOK
Value: 30007662 - TABLET PC
Value: 30002515 - UNCLASSIFIED
Value: 30002518 - UNIDENTIFIED

UNSPSC

43211503 Notebook computers



Requirements

Standardization need

Complexity ↑

Process standards
ECR, ebXML, SCOR, RosettaNet, ...

Transaction standards
UBL, RosettaNet, EANCOM®, GS1-XML, OAGIS, ...

Product catalogue standards
BMEcat, cXML, xCBL, ebXML, RosettaNet, ...

Product classification standards
CPV, eCI@ss, GPC, UNSPSC, ...

Identification standards
DUNS®, EAN/GTIN, EPC, ILN/GLN, ...

Requirements

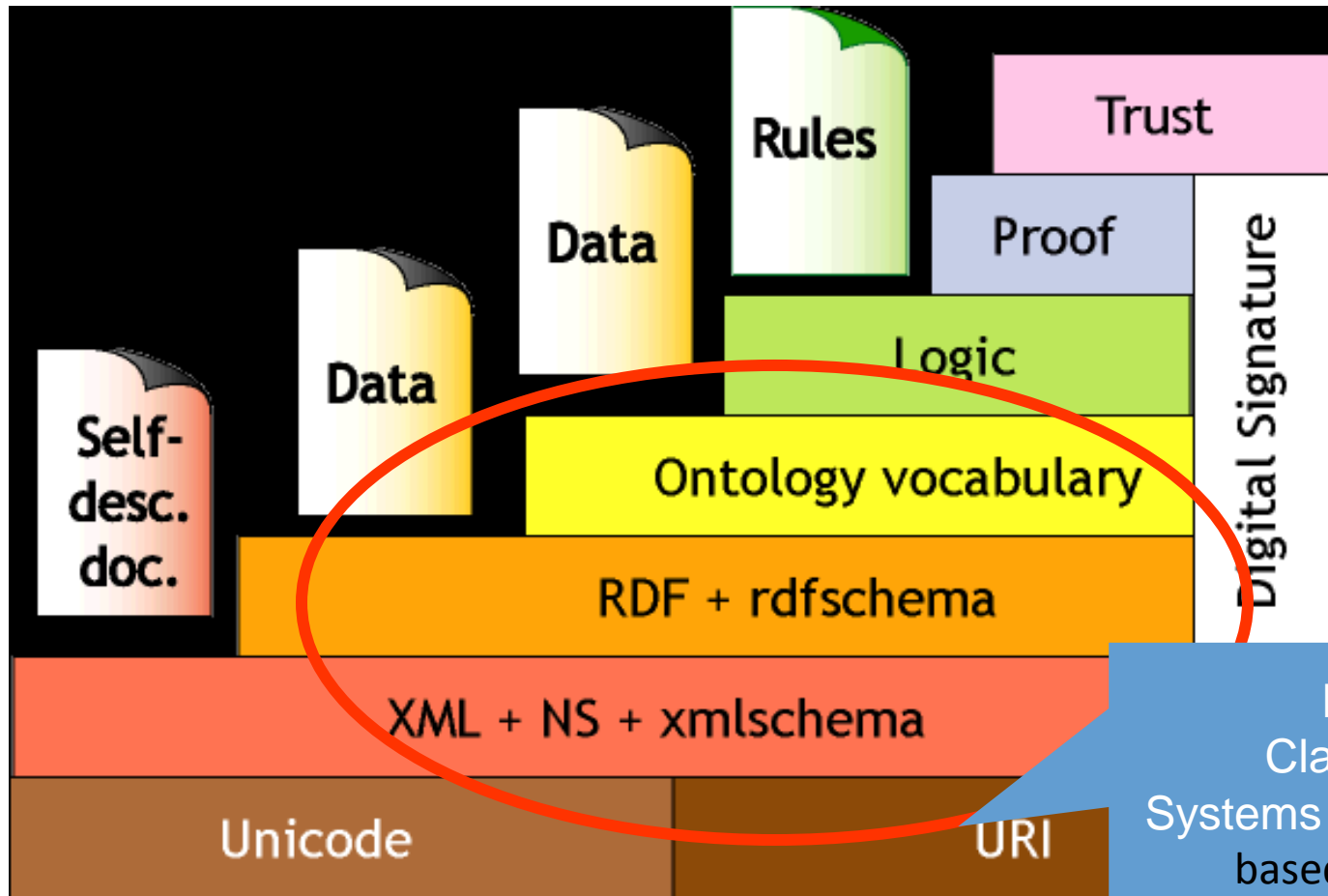
eAccessibility&eInclusion

Interoperability → incl. content re-purposability

Solution approach: Enhancement of existing data models for eProcurement and/or ontologies for all kinds of structured content (at the level of lexical semantics), such as

- Product Classification Systems → mapping
- Terminologies and other language resources
- AAC (augmentative and alternative communication)

The Semantic Web Architecture



Product Classification Systems (PCS) as OWL-based lightweight Ontologies and according mapping approaches

Product Classification Systems

seen as lightweight ontologies

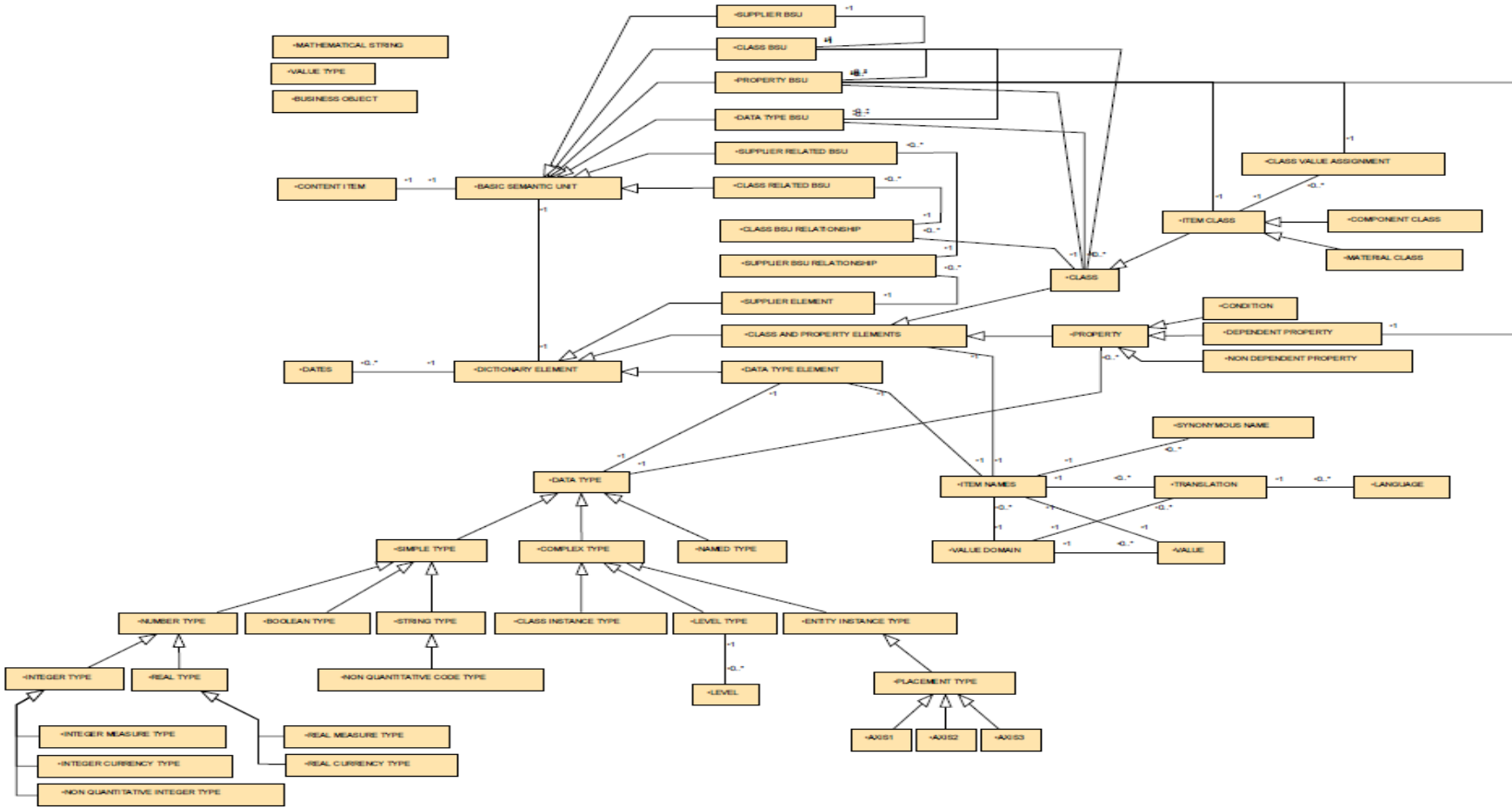
Approach: Mapping of lightweight ontologies

- Ontologies are serving for structuring and exchanging of data or information. They typically consist of
 - Concepts or classes,
 - Types,
 - Instances,
 - Relations,
 - Inheritance and
 - Axioms.
- Taking into account these formal parts of ontologies, a product classification system can be seen as a **“lightweight ontology”**, containing terms, a taxonomy (a classification framework for all products and services) and relations between terms and properties, which describe these terms, like classes.

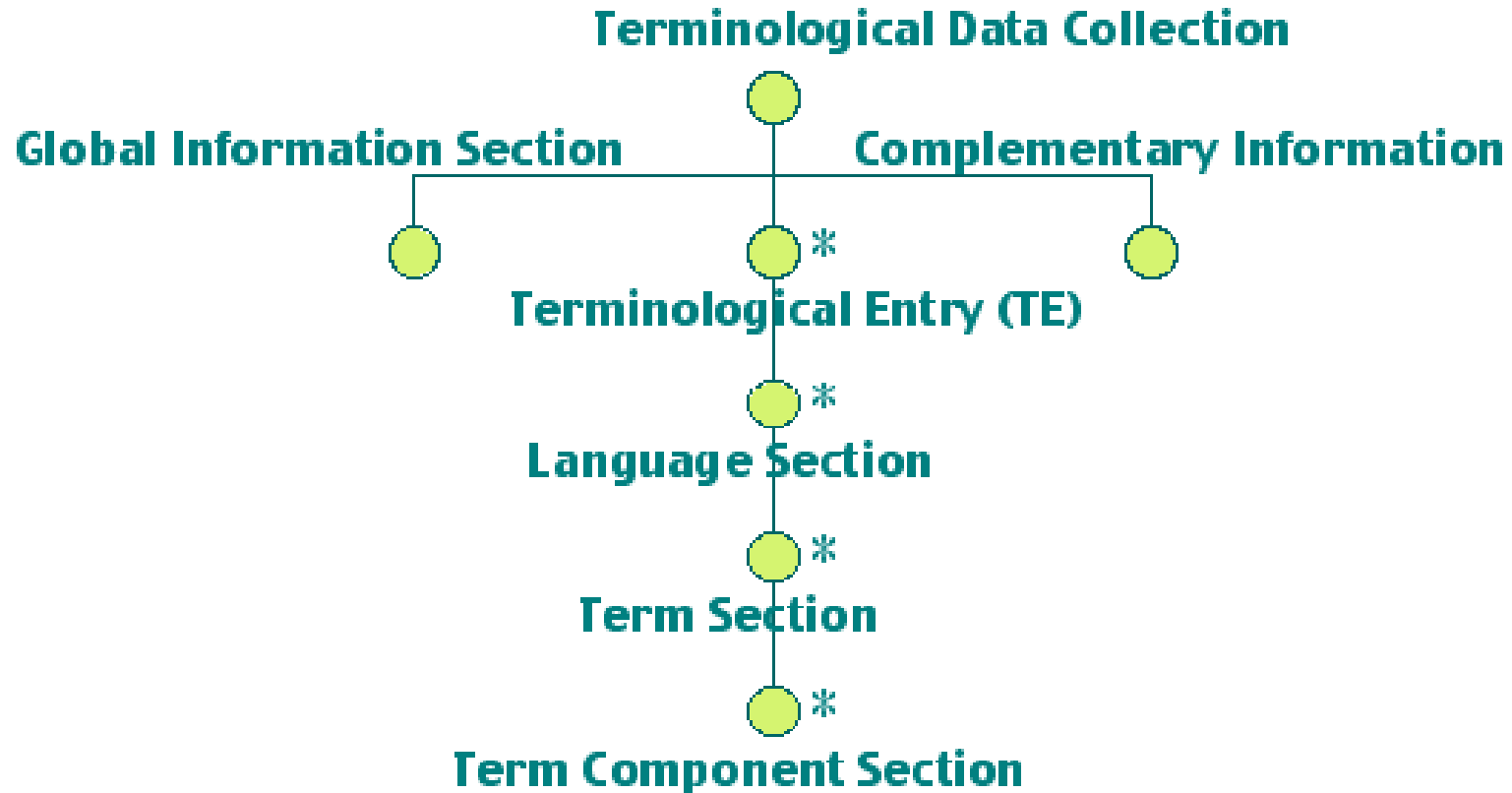
→ The ontology mapping problem is very closely related to the product classification mapping problem, since a product classification can be seen as a specialized ontology using a (mono-) hierarchical structure for the concepts, e.g. product classes, within the ontology.

eAccessibility & eInclusion

Approach: data model extensions



Terminological metamodel



Source: Klaus-Dirk Schmitz 2005

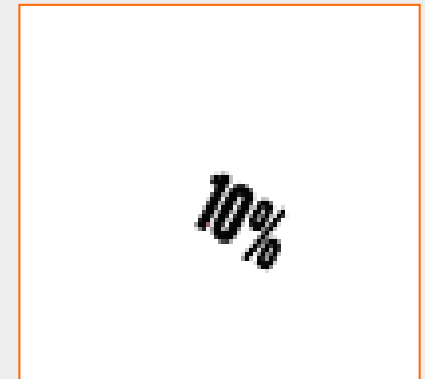
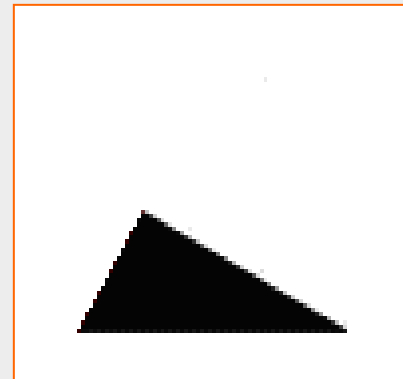
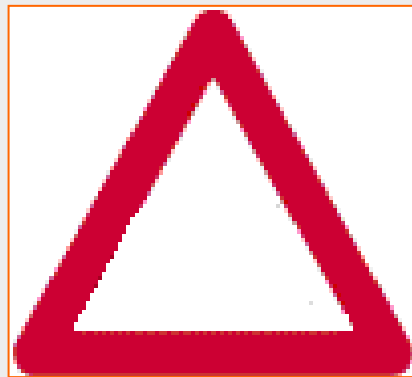
“Morphology” of pictogrammes

■ Representation elements

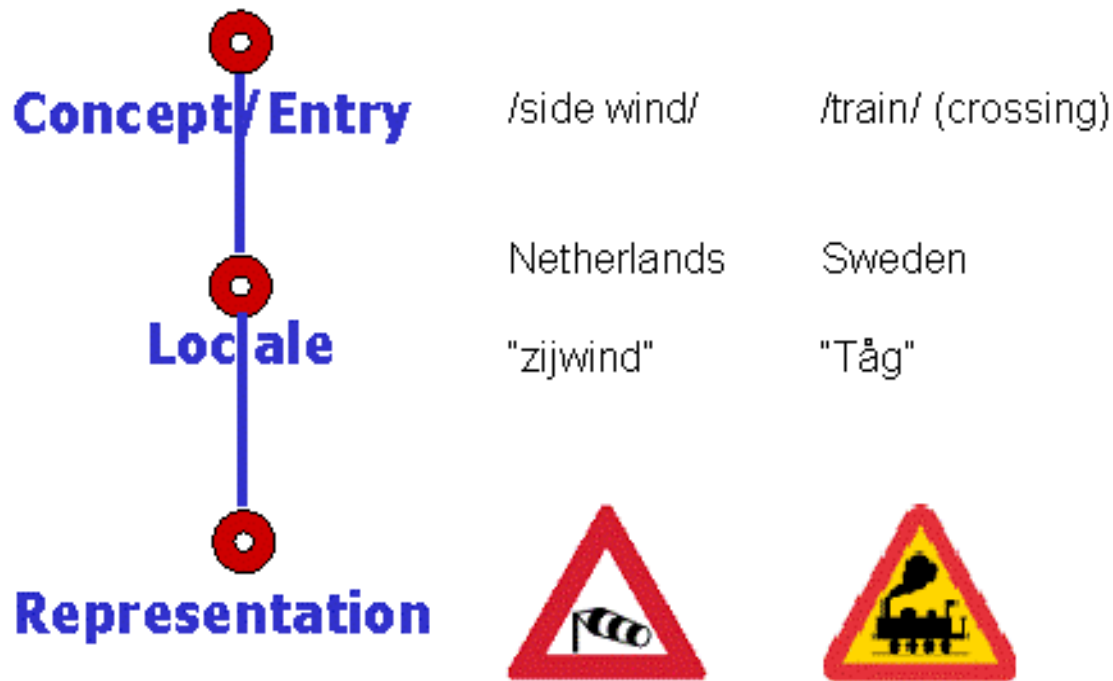
- morphology
- non-verbal search?










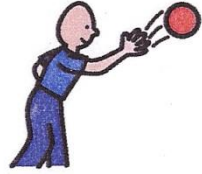



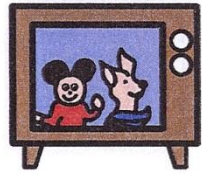


=



„Locales“ → representation



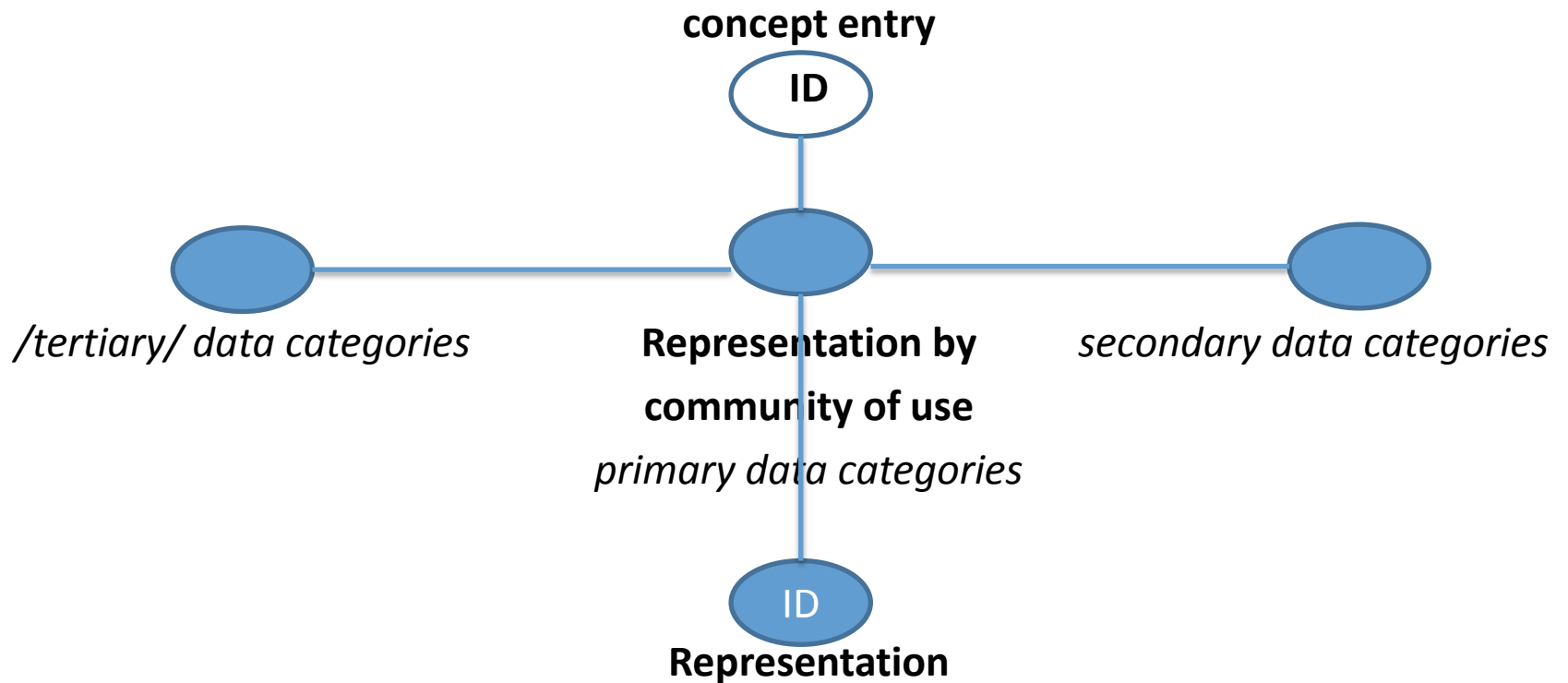
AAC

Bonjour! 	Je m'appelle Calum		
j'aime 	les voitures et 	les trains 	
j'aime 	faire du cheval 	lancer des choses 	et les balles 
j'aime 	écouter la musique 	et lesdessins animés 	
J'aime pas 	rester calme 		



Adapted metamodel

for all kinds of structured content
(at the level of lexical semantics)



Representation

- Representation autonomy
- Repeatability by kind of representation
- Repeatability within kind of representation
- Linguistic / non-linguistic representations
- /symbolic/ vs. descriptive representations
- Low/high complexity (primitive/composite)

Representations

	PD	SD	TD
R1	term, synonym, abbreviation... definition...	use, gender, declination, concept system...	ID, source... ID, source...
R2	letter symbol	use... concept system...	ID, source...
R3	graphical symbol	use...	ID, source...
R4	formula complex formula	use concept system...	ID, source...
R5	class (entry – e.g.) (product) characteristics	use... class structure...	ID, source...
R6	AAC entry	use...	ID, source...
R7	learning object	use...	ID, source...

Structured content

Learning object: Kanji Flashcard 「休」

Front side: kanji, examples and additional information →

10 → キュウ
11 → やす・む/やす・まる/やす・める
12 → rest, take a day off, relax
13 → ■ person ■ tree
14 → 1: ていきゅうび a regular holiday {for a store}
2: ひとやすみする to take [have] a (short) rest
3: やすまる to be [feel] rested; to be relieved
4: やすむ to rest; to take a day off; to sleep
5: やすみ (a) rest; a holiday; (a) vacation; a day off
6: きゅうじつ a day off; a holiday

1: 定休日
2: 一休みする
3: 休まる
4: 休む
5: 休み
6: 休日

13 → 5
115 → 6
7 → 7
body → 8
1039 → 9
crouch → 9

6-2-4 1

← Back side: lexems, pronunciation, meanings and additional information

Source: Whiterabbit kanji flashcard

→ **CLIL – Content and Language Integrated Learning**

AAC for cognitively disabled

Swedish	födelsedag	pappa	ge	jag	gåva	hund	lille
Dutch	verjaardag	vader	geven	mij	kado	hond	klein
English	birthday	father	give	me	present	dog	small
PCS							
Picto							
Bliss							

This image shows six different ways to express the same sentence: "For my birthday, father gives me a small dog". Vertically, each single concept is represented in different languages (keywords) and in different symbol systems (PCS, Picto, Bliss).

Structured content:

Traffic informatics



Way to the airport – turn right in 5 km



Way to the train station – down to the right



ZONE = verbal

red ring = (morphology) prohibition sign

30 = micro-proposition: max speed 30km/h

→ variable message sign boards

communicating with car-driver system

Content interoperability

Three levels of standardization:

- **Ontology interoperability**
incl. meta ontology language ISO/NWI 17347
- **Data modeling methods incl. metamodels**
- **Data categories, data models**

➔ Repurposing of structured content for eLearning

➔ **ONE generic approach? YES**

Conclusion

- Terminology science beyond linguistics and LSP
- Concept representations not only terms, words...
- Terminological principles and methods contribute to modern content management
- Mobile technology enables mobile content
- Applications in all domains and subjects
- Dynamics in non-linguistic applications
- Mobile technology is changing user's attitudes
- **Increasing need for global content integration**

Open issues

- **New dimensions federating content repositories**
- **Increasingly different kinds of content are combined with or embedded in each other**
- **New dimensions of content management**
- **Mobile technology = mobile web & mContent**
- **Unsolved issues and lack of business models**
- **Data quality (incl. total reliability of certain data)**
- **Standards → standards-based certification**
- **eProcurement as a lever for implementation**

Questions / Discussion?



Questions?



Infoterm

**International Information Centre for
Terminology**

**Gymnasiumstrasse 50
1190 Vienna – Austria
Tel: +43-1-4277 58026
Fax: +43-1-4277 58027
infopoint@infoterm.org
<http://www.infoterm.info>**



International
Organization for
Standardization

Symbolic representation

- **GPL lexemes:**
from letter via morpheme and word to collocation and micro-sentences
what about non-verbal communication?
- **Domain/subject communication designations:**
SPL: from letter symbol via morpheme and term/synonym/ abbreviation/... to terminological phrasemes and micro-statements
what about non-linguistic representations?
- **Spoken language and other modalities?**

Descriptive representation

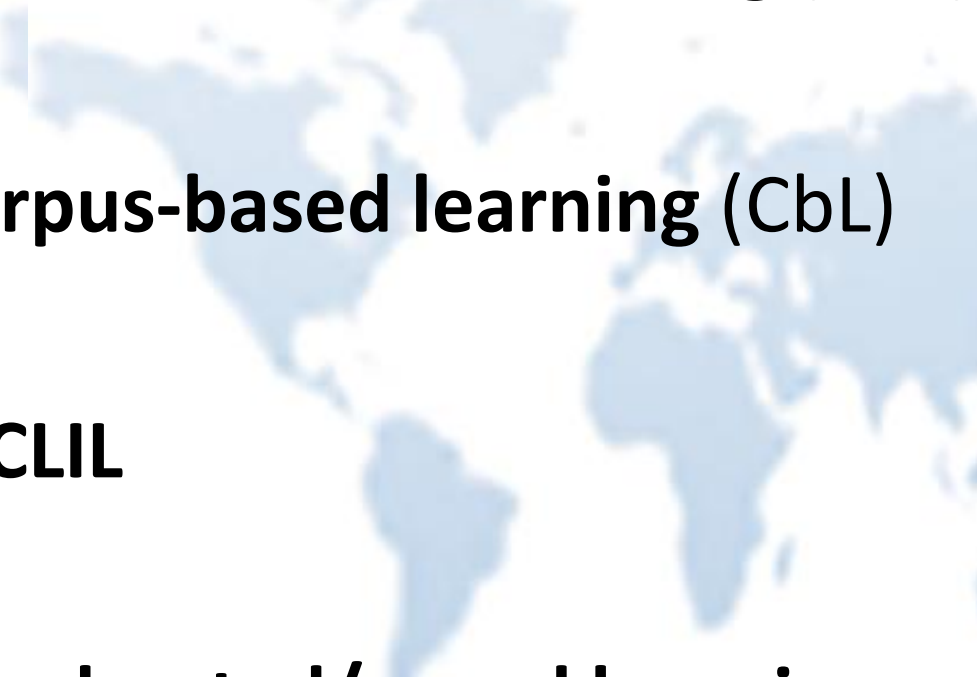
- **GPL explanation, contexts etc.**
what about non-verbal communication?
- **/descriptions/ in domain/subject communication :**
SPL: definitions, explanations, contexts, notes, expamples, ...
... non-linguistic descriptive representations?
- **Spoken language and other modalities?**

Items of structured content

in eLearning as learning objects

- **Low-complexity items („primitives“):**
 - Referring only to one concept
 - Referring only to one object or /class/ of objects (e.g. one model or group of models of a product)
- **High-complexity items („composites“):**
 - Referring to more than one concept, e.g. by
 - comprising examples of usage
 - including information on a higher grammatical level
 - Referring to complex objects, such as components
 - Mixing information on concepts and objects

Combined methodology

- **Resource-based learning (RbL)**
- + 
- **Corpus-based learning (CbL)**
- +
- **MCLIL**
- +
- **Accelerated/speed learning**

Resource-based learning (RBL):

*Info
term*

based on structured content and MCLIL approaches

Any entity of structured content
with or without context / co-text
is **potentially re-usable for eLearning**

Any teaching / learning
starts with a phase for beginners
and may continue
with a phase for advanced students

There can be language teaching / learning
for advanced domain experts and
domain knowledge acquisition for advanced language learners
and **anything between...**