

Measuring the Quality of Descriptive Languages for Products and Services

Multikonferenz Wirtschaftsinformatik
March 9 – 11, 2004, Essen (Germany)

Dr. Martin Hepp
Florida Gulf Coast University
mhepp@computer.org

Dr. Martin Hepp

Full Paper – Preferred Citation

Hepp, Martin: Measuring the Quality of Descriptive Languages for Products and Services. In: Dorloff, F.D.; Leukel, J.; Schmitz, V.: E-Business – Standardisierung und Integration. Tagungsband zur Multikonferenz Wirtschaftsinformatik 2004. Cuvillier, Goettingen (Germany) 2004, pp. 157-168. Available via <http://ssrn.com/abstract=503602> or www.heppnetz.de.

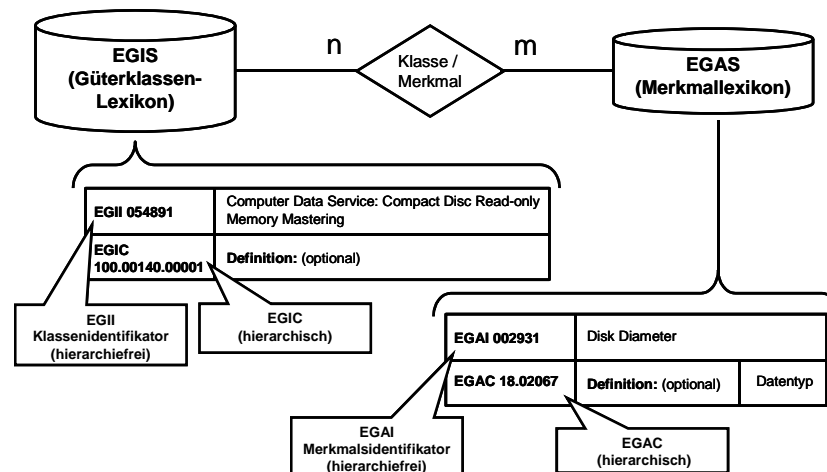
Dr. Martin Hepp

Descriptive Languages for Products and Services

- Key functionality: **Reference** to business meanings (e.g. a product or service) in an unambiguous, machine processable manner.
- Descriptive Languages vs. Ontologies
- Most popular architecture: combination of classes, an attribute library, and class-specific attribute lists

Dr. Martin Hepp

eOTD Architecture



Dr. Martin Hepp

Semantic Standards provide two functions

They can be used for

- the **description** and
- as a **point of reference**.

<CONTENT>
12-45-65
</CONTENT>
...

RECIPE
for 12-50-50:
2 UNITS 12-45-65
1 UNIT 12-50-60
...

Dr. Martin Hepp

Content Quality of Descriptive Languages for Products and Services

- Quality and usefulness is determined by how well it covers the linguistic needs of a business domain.
- Content development and maintenance is a major challenge.

Dr. Martin Hepp

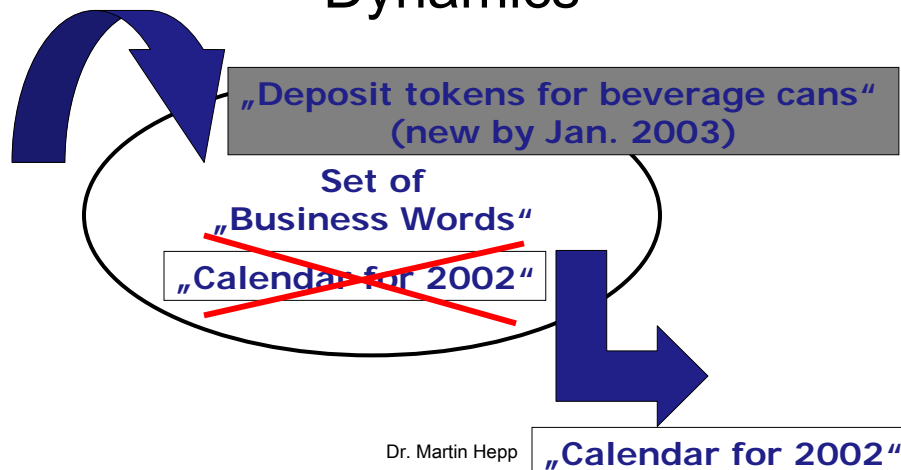
Linguistic Dynamics of Business Meanings

...Language is a living organism that adapts to the development and the trends of society as a whole.

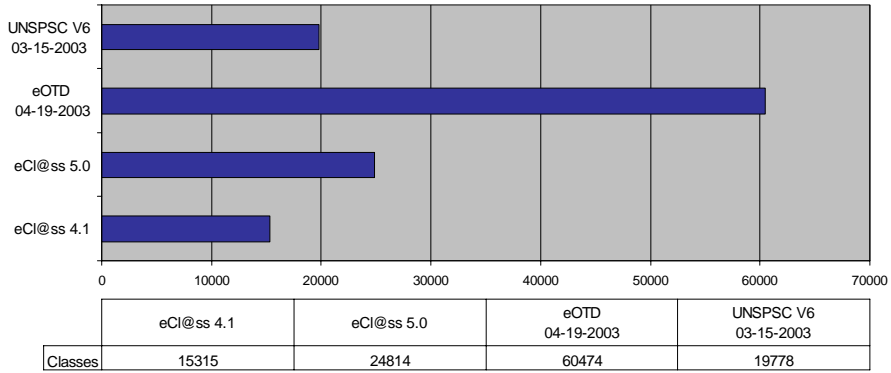
(Umberto Eco
in the preface to Ogden / Richards:
„The Meaning of Meaning“)

Dr. Martin Hepp

Speed of Maintenance vs. Linguistic Dynamics

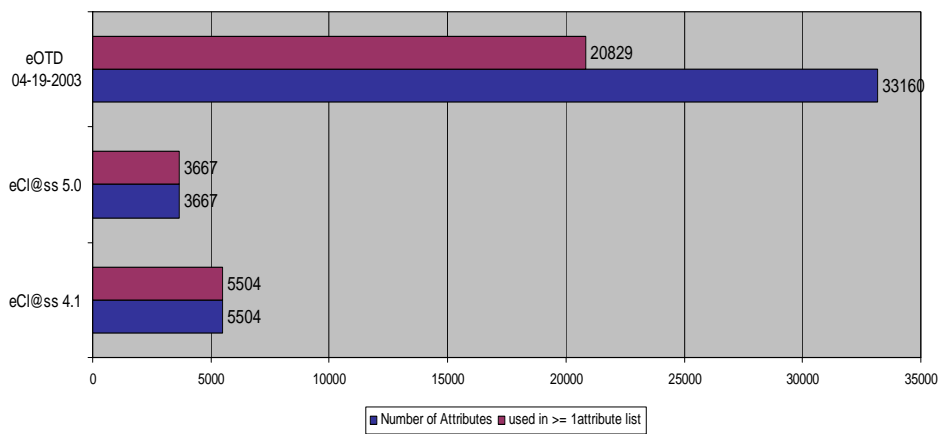


Total Number of Product and Service Classes



Dr. Martin Hepp

Size of the Attribute Library



Dr. Martin Hepp

Metrics for Content Quality

- Distribution of Entries along the Top-Level Hierarchy
- Semantic Weight and Semantic Value

Dr. Martin Hepp

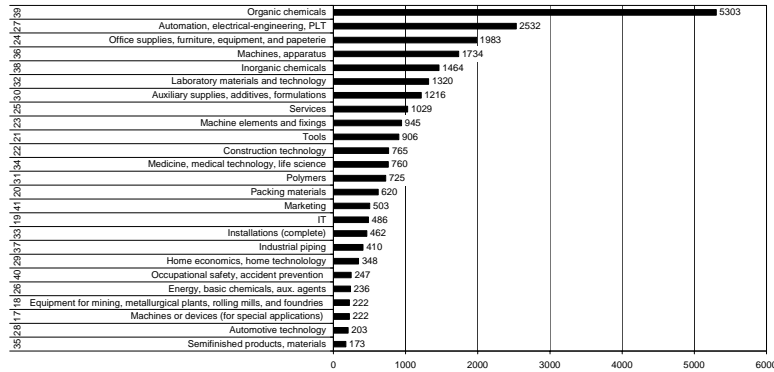
Distribution of Entries along the Top-level Hierarchy

	eCI@ss 4.1	eCI@ss 5.0	eOTD 04-19-2003	UNSPSC V6 03-15-2003
Product classes	15315 (all levels)	24814 (all levels)	60474 (all levels)	19778 (all levels)
Number of top-level categories	22	25	79	55
Mean (percentage)	696 (4.5 %)	992 (4 %)	765 (1.3 %)	360 (1.8 %)
Amount of nodes in the biggest top-level category	3594	5303	14802	2447
Percentage	23.47 %	21.37 %	24.48 %	12.37 %
Amount of nodes in the three biggest top-level categories	6788	9818	24082	5860
Percentage	44.32 %	39.57 %	39.82 %	29.96 %

Dr. Martin Hepp

Distribution of Entries along the Top-level Hierarchy

eCI@ss 5.0: Total Number of Nodes by Top-level Categories



Dr. Martin Hepp

Quality of Attribute Assignment

	eCI@ss 4.1	eCI@ss 5.0	eOTD 04-19-2003
Number of node-specific attribute lists	6507	7913 (7917, see text)	19927 (quality issues, see text)
Percentage (based on the <i>total</i> number of product classes)	42.5 %	31.9 %	33 %

Dr. Martin Hepp

eCl@ss 5.0: Clusters of Attribute Usage Frequency in Attribute Lists

Percentage of Classes with this Attribute	Number of Attributes	Share of the Attribute Library
> 75 %	5	0.14 %
29...74 %	none	
25...28 %	44	1.20 %
3...24 %	none	
< 2 % (but contained in more than one attribute list)	1942	52.96 %
contained in exactly one attribute list	1676	45.70 %

Dr. Martin Hepp

Semantic Value and Weight

Semantic Weight for each attribute:

$$SW_i = \frac{1}{\text{Number of Attribute Lists Containing } A_i}$$

Semantic Value for each class:

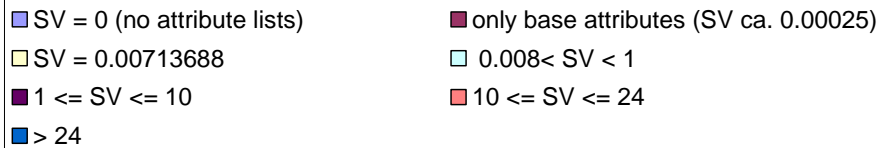
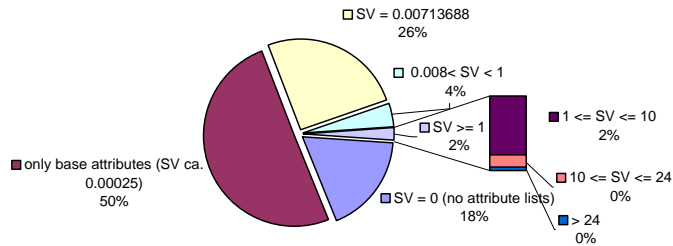
$$SV_j = \sum SWA_i \mid A_i \in S_j$$

Distribution properties!

Mean	0.17994013
Min	0.00019964
Max	87.88928283
Median	0.00025091
Mode	0.00025091

Dr. Martin Hepp

ecl@ss 5.0: Distribution of Semantic Values



Dr. Martin Hepp

Application

1. Corporations can use the metrics to evaluate the content quality of a specific descriptive language with regard to their product range or sourcing needs.

Dr. Martin Hepp

Application

2. Standards bodies could use those metrics to
- monitor the development of content quality,
 - assess the amount of resources necessary to eliminate the shortcomings,
 - rank content maintenance alternatives,
 - motivate industry groups to help improve currently weak segments of the standard.

Dr. Martin Hepp

Discussion, Questions, Comments?

Dr. Martin Hepp

mhepp@computer.org

Working papers available at www.heppnetz.de

Dr. Martin Hepp