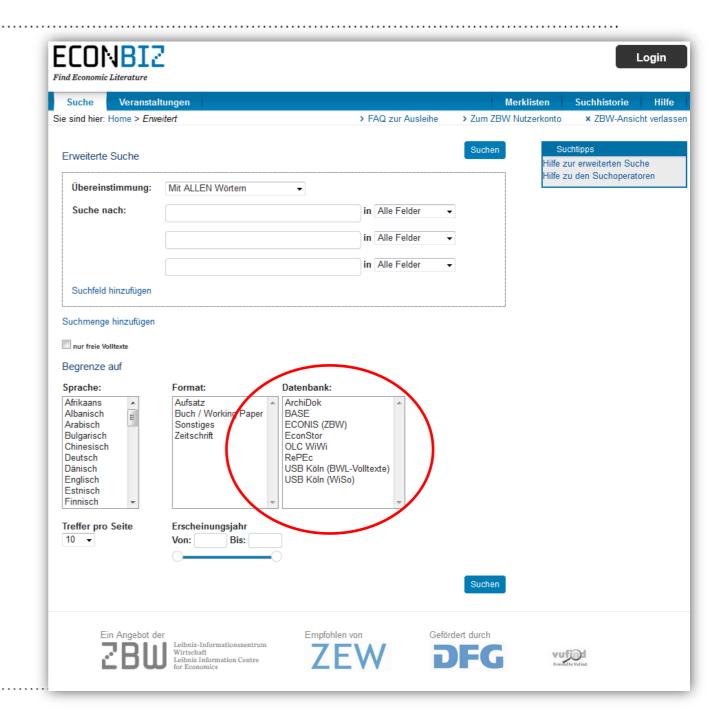
# The Missing Link – A Vocabulary Mapping Effort in Economics

<u>Andreas Oskar Kempf</u>, Joachim Neubert, Manfred Faden ZBW – Leibniz Information Centre for Economics – German National Library of Economics 14th European Networked Knowledge Organization Systems (NKOS) Workshop September 18th 2015



## Why do we do vocabulary mappings in general?

- Mappings enable an integrated search in a distributed search environment.
- Mappings translate search terms into the vocabulary of the target KOS.

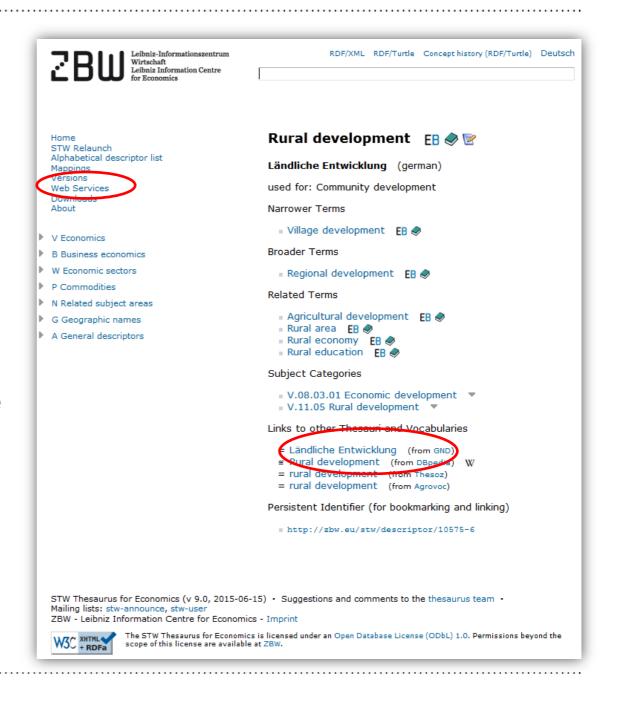






## For what reason did we at ZBW do mappings in the past?

- ... to offer an integrated search space for our search portal for economics EconBiz,
  - e. g. Integrated Authority File
- ... to link the STW with other vocabularies for the development of semantic web applications.



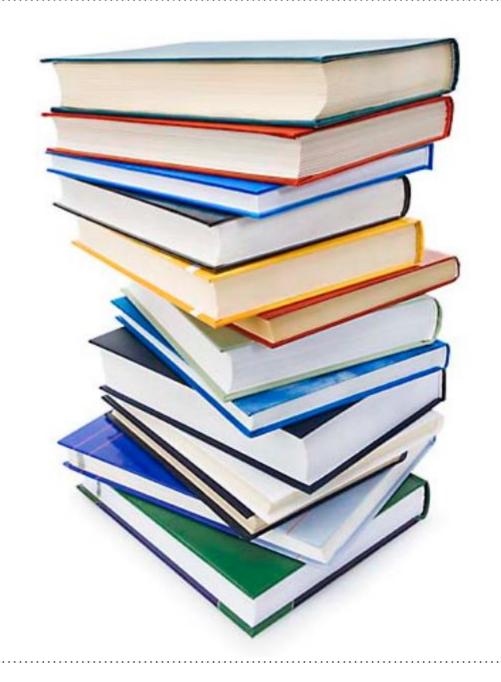




## What is new about the current mapping effort?

#### **Context:**

- increasing numbers of publications and decreasing personnel resources.
  - ... complementary approaches to conventional subject indexing are needed,
    - i. a. reuse of user-generated content.

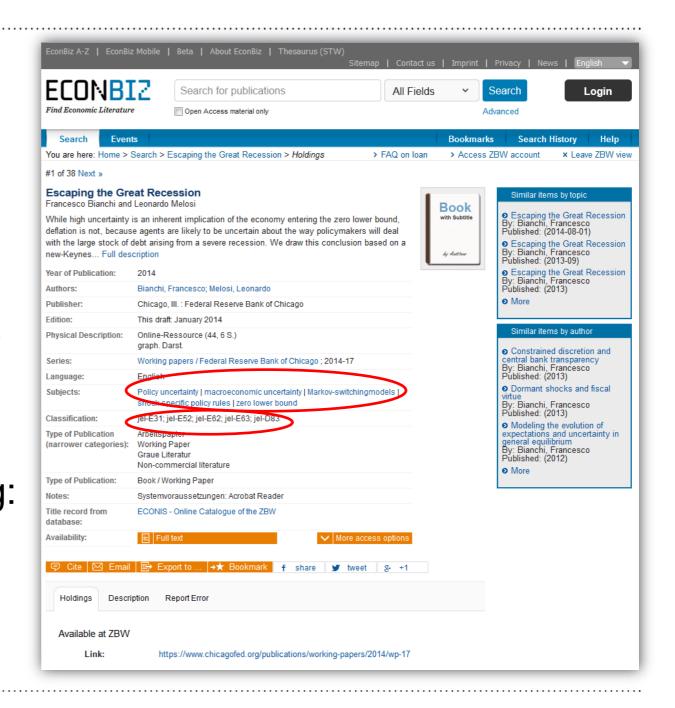




## Current reuse scenario of user-generated content at ZBW:

Regarding working paper series:

- Verbal subject indexing:
   inclusion of author keywords
   into bibliographic records if
   available.
- Classificatory subject indexing: inclusion of **JEL classes** into bibliographic records if available.







### Future reuse scenario for a JEL – STW (systematic display) mapping effort:

- ... building on the fact that
   economists are usually quite
   familiar with the JEL classification
   codes.
- ... animate economists to use STW subject headings in order to provide a more fine-grained content description with a standardized vocabulary.

STW Thesaurus for Economics
(subject categories)



Journal of Economic
Literature
Classification

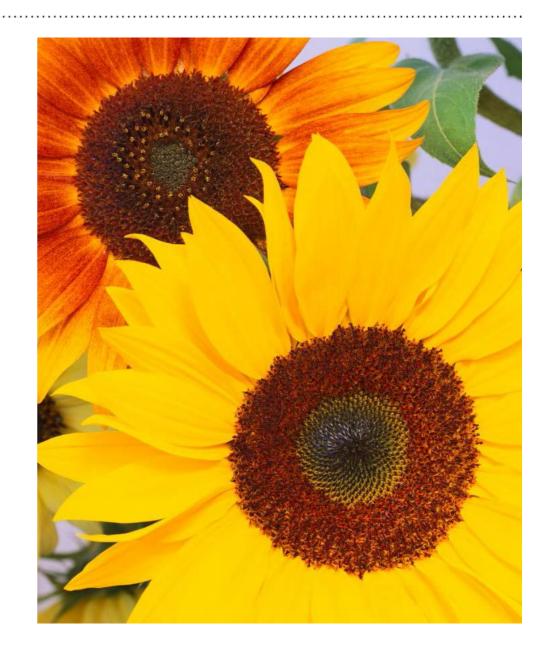




## Research question

Regarding the use case we have in mind to what extent is a useful mapping between both KOS possible?

Dealing with this question on the one hand includes a **theoretical reflection** on the structure of both KOS. On the other hand it includes the presentation of a specific **iterative semi-automatic mapping** 





approach.



#### Outline

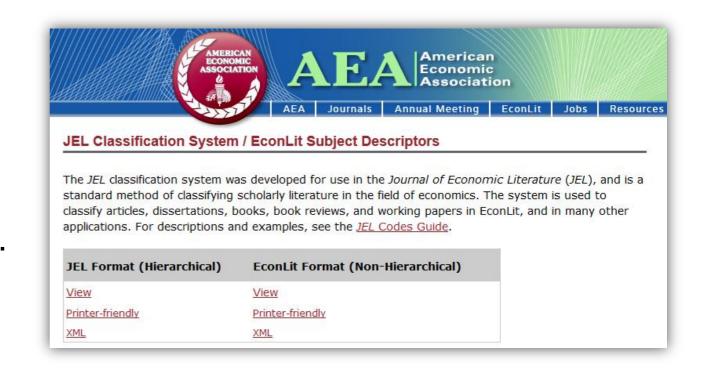
- Introduction
- Knowledge organization systems in economics
- Definition of interoperability and structural models for mapping
- Mapping process
- Empirical examples
- Results
- Conclusion and future outlook



## JEL Classification System

#### Institutional background:

- It is published by the American
   Economic Association (AEA),
   which publishes the American
   Economic Review and maintains
   the searchable database EconLit.
- The AEA Executive Committee regularly reports on changes of JEL classes in the American Economic Review.







#### JEL Classification

#### Scope:

 It represents an Anglo-American understanding of economics mainly focusing on (national) economics [ger.: VWL].

#### Structural characteristics:

 It is a precombined classification system with a monohierarchical structure and polydimensional ordering principles.





#### **JEL Classification Codes Guide**

The information on these pages is to provide helpful reference information and guidelines about the <u>JEL Classification System and EconLit Subject Descriptors</u>.

You can drill down to the more detailed secondary level of classifications by clicking on the JEL code links (for example, A3, H7); these contain more descriptive guidelines, keywords, and also examples of items with that specific classification.

#### A General Economics and Teaching

B History of Economic Thought, Methodology, and Heterodox Approaches

**Guideline:** Covers studies about schools of economic thought from pre-classical to contemporary developments (B1-B3) and current heterodox approaches to various economic issues (B5). Also covers studies about methodology (B4). The categories for B5 were created to accommodate the recent substantial developments in non-mainstream (non-neoclassical) economic approaches to contemporary economic problems. Studies about an economic subject (or subjects) adopting a given heterodox approach should be cross-classified under the appropriate B5 category and also under the appropriate subject category (or categories).

Keywords: History of Economic Thought

#### **B00 General**

**B1 History of Economic Thought through 1925** 

**B2 History of Economic Thought since 1925** 

**Guideline:** Studies that are also relevant to contemporary economics or economies should be cross-classified here and under the other appropriate categories.

Keywords: Economic Thought, Neo Classical.

#### **B20** General

**B21 Microeconomics** 

**B22 Macroeconomics** 

**B23 Econometrics; Quantitative and Mathematical Studies** 

**B24 Socialist; Marxist; Sraffian** 

B25 Historical; Institutional; Evolutionary; Austrian

**B26 Financial Economics** 

**B29 Other** 

**B3 History of Economic Thought: Individuals** 

#### STW Thesaurus for Economics

#### Institutional background:

 Developed in cooperation thanks to a project funded by the Ministry for Economy in the 1990s.

#### Scope:

It covers all economics-related subject areas and, on a broader level, the most important related subjects (e.g. social sciences).













Leibniz-Informationszentrum Wirtschaft Leibniz Information Centre for Economics

#### **STW Thesaurus for Economics**

Version 9.0 (Changes)

#### Subthesauri

- V Economics
- B Business economics
- W Economic sectors
- P Commodities
- N Related subject areas
- G Geographic names
- A General descriptors





#### STW Thesaurus for Economics

#### Structural characteristics:

STW is a **polyhierarchical** bilingual thesaurus.

#### Types of relations:

- equivalent relations, including synonyms and quasi-synonyms (UF),
- hierarchical relations, including broader (BT) and narrower terms (NT)
- associate relations, including related terms (RT)

#### Links to other vocabularies:

Mappings to GND, TheSoz, AGROVOC, (DBpedia)



Home STW Relaunch Alphabetical descriptor list Mappings Versions Web Services Downloads

- V Economics
- B Business economics
- W Economic sectors
- P Commodities
- N Related subject areas
- G Geographic names
- A General descriptors

#### Financial statement 🔣 🥏 📝



RDF/XML RDF/Turtle Concept history (RDF/Turtle) Deutsch

Jahresabschluss (german)

used for: Annual accounts

Narrower Terms

- Balance sheet EB 🕏
- Consolidated financial statements EB 🔗
- Income statement FR 🔊

#### Broader Terms

- Accounting EB Reporting EB
- Related Terms
  - Cash flow statement EB 🥏
  - Corporate annual report FR
  - Financial statement analysis FR 🔊
  - Financial statement audit EB
  - Management commentary EB

#### Subject Categories

■ B.03.01.02 Financial statement

Links to other Thesauri and Vocabularies

- = Jahresabschluss (from GND)
- > Versicherungsbilanz (from GND)
- > Kurzfristige Rechnungslegung (from GND)
- < Rechnungslegung (from GND)
- ≅ Jahresabschluss (from DBpedia) W
- = year-end financial statement (from Thesoz)
- = rendering of accounts (from Thesoz)

Persistent Identifier (for bookmarking and linking)

http://zbw.eu/stw/descriptor/12381-2

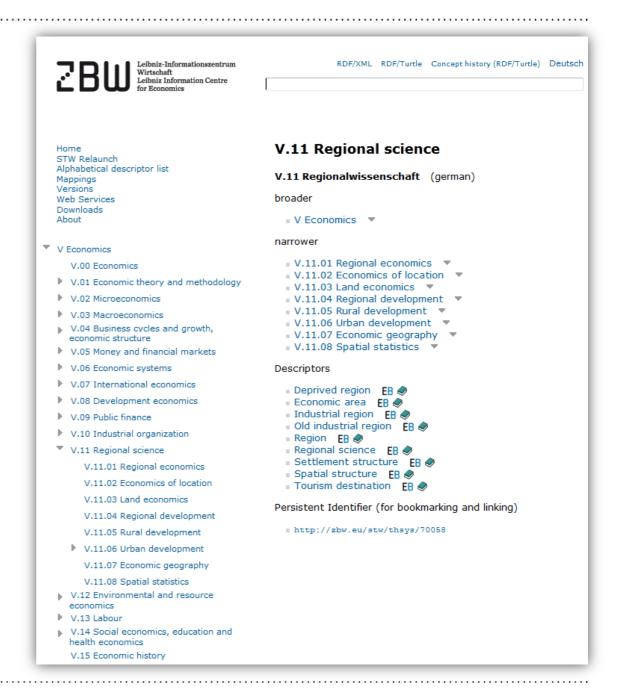


## STW subject categories

#### **Structural characteristics:**

The STW subject categories (in total 497) constitute a monohierarchical structure with polydimensional
 for subthesaurus V + B – consistently subject-specific ordering principles for vertical and horizontal subdivision.

	Subthesaurus V	Subthesaurus B
1st level	1	1
2nd level	15	10
3rd level	62	38
4th level	43	21
Total	121	70







## JEL Classification vs. STW Subject Categories

	JEL Classification	STW Subject Categories			
Definition	Concept group (ISO 25964-2: 3.18, "group of concepts selected by some specified criterion…")	Concept group (ISO 25964-2: 3.18)			
Scope	Domain-specific (USA, UK)	Domain-specific (GER > international) Here: Restriction to the subthesauri: V: Economics and B: Business economics.			
Type of concept group	Classes of a classification Purpose: All-embracing systematization of a discipline	Subject categories of a thesaurus Purpose: Systematization of the thesaurus vocabulary			
Structural characteri stics	<ul><li>Precombined classification</li><li>Monohierarchical</li><li>Polydimensional ordering principles</li></ul>	<ul><li>Monohierarchical</li><li>Polydimensional ordering principles</li></ul>			

 Because of the structural heterogeneity between the two vocabularies mapping relations for the most part are not expected to be relations of full equivalence.
 Rather they are presumed to oftentimes consist of inexact equivalent relations





#### Outline

- Introduction
- Knowledge organization systems in economics
- Definition of interoperability and structural models for mapping
- Mapping process
- Empirical examples
- Results
- Conclusion and future outlook





## Definition of interoperability

ISO 25964: Thesauri and interoperability with other vocabularies

Developed by an international working group (2008-2013)

- Part 1: Thesauri for information retrieval (published 2011)
   Contains guidelines for establishing monolingual and multilingual thesauri.
- Part 2: Interoperability with other vocabularies (published 2013)
   Deals with mappings between thesauri and other types of vocabularies for information retrieval.



## Definition of interoperability

ISO 25964-2:2013(E)

#### 3.38

#### interoperability

ability of two or more systems or components to exchange information and to use the information that has been exchanged.

NOTE Vocabularies can support interoperability by including mappings to other vocabularies, by presenting data in standard formats and by using systems that support common computer protocols.

#### 3.40

mapping, gerund (verbal noun) process of establishing relationships between the concepts (3.17) in one vocabulary and those of another

#### 3.41

#### mapping, noun

(product of mapping process) relationships between a concept (3.17) in one vocabulary and one or more concepts (3.17) in another





## Two different types of vocabularies

#### **Structural unity:**

The mapped vocabularies have the same structure.

The equivalence of the concepts of such vocabularies is expressed by their identical structural position in the vocabulary. All the relationships of the concepts correspond to each other (e.g. multilingual thesauri of public institutions)

#### **Structural disunity:**

The mapped vocabularies do not have the same structure.

Equivalence of concepts has nothing to do with their position in the vocabularies. The mapping process produces either exact equivalence pairs or inexact equivalent pairs.

#### Different types of equivalences:

(Real) exact equivalence: =EQ

Inexact equivalence: ~EQ (e.g. voc.have emerged from different cultural communities)

Partial equivalence: The concept is broader: BM ("Broader Mapping")

The concept is narrower: NM ("Narrower Mapping")

The concepts are somehow related: RM ("Related Mapping").





## Structural models for mapping

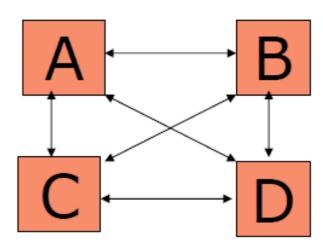
ISO 25964-2:2013(E)

## Three different structural models for mapping across vocabularies Model 1: Structural unity (6.2)

"All the participating vocabularies share exactly the same structure of hierarchical and associate relationships between concepts…" (e.g. multilingual thesauri)

#### Model 2: Direct-linked (6.3)

The direct-linked model addresses linkages betweent two or more vocabularies that do not share the same structure. As well as differing in scope, language and structure, the vocabularies may include other types of vocabulary (classification scheme, name authority list, etc.).





## Structural models for mapping

#### Model 3: Hub structure (6.4)

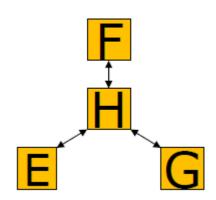
One vocabulary is designated as "hub", or conprehensive structure to which each of the other vocabularies is mapped as "satellite". The concepts of the different vocabularies are only mapped to the concepts of the one vocabulary which has the role of a hub. This model is appropriate if there is one vocabulary with a dominating position.

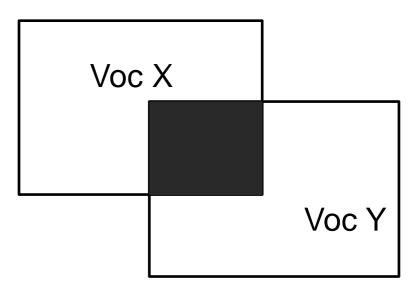
#### **Model 4: Selective Mapping (6.5)**

In cases where there is only small overlap expected, it could be unnecessary to map the vocabularies comprehensively.

➤ In real applications combinations of these types often occur and the boundaries might be blurred (see ibd. p.20):

#### ISO 25964-2:2013(E)





Selected mapping in area of overlap.



#### **Outline**

- Introduction
- Knowledge organization systems in economics
- Definition of interoperability and structural models for mapping
- Mapping process
- Empirical examples
- Results
- Conclusion and future outlook



## Mapping process

#### **Previous work:**

- Outdated mapping JEL > STW (descriptor level)
   KoMoHe project context (2004-2007)
   Mapping relations:
  - equivalent relations (=)
  - broader/narrower relations (>/<)
  - associate relations (^)
  - compound mappings (+)
  - including a relevance rating ( high, medium, low)

22	8	A10	>0	g	Zeitkonsistenz
22	8	A10	>0	g	Ökonomisches Prinzip
22	8	A10	>0	m	Wirtschaftsmodell
22	8	A10	>0	m	Wirtschaftswissenschaft
22	8	A11	<0	g	Wirtschaftsforschung
22	8	A11	>0	g	Wirtschaftswissenschaftler
22	8	A12	<	g	Wirtschaftswissenschaft
22	8	A13	٨	m	Wirtschaftsethik
22	8	A14	=	h	Wirtschaftssoziologie
22	8	A19	<	g	Wirtschaftswissenschaft
22	8	A1	=	m	Wirtschaftstheorie
22	8	A20	<0	m	Pädagogik
22	8	A20	0	h	Wirtschaftspädagogik
22	8	A20	^0	m	Arbeitswissenschaft
22	8	A20	^0	m	Berufsbildung
22	8	A20	^0	m	Wirtschaftsstudium
22	8	A20	^0	m	Wirtschaftswissenschaft
22	8	A21	<	g	Wirtschaftsstudium
22	8	A22	<	g	Wirtschaftsstudium
22	8	A23	<	g	Wirtschaftsstudium
22	8	A29	<	a	Wirtschaftsstudium

- Outdated concordanceSTW (classification system) > JEL
  - On the third level of JEL classes,
  - No specified mapping relations.

STW classification system	STW subject categories	JEL classes			
V02-000 Microeconomics	V.02 V.02.05	B21 D00 			
V02-010 Houlsehold economics	V.02.01	D10 D11			



## Mapping process

#### What is new?

- ...mapping on the level of the STW subject category system (instead of the STW classification level),
- ...referring to a web-based interactive mapping platform,
- ...using the SKOS vocabulary to build and to manage the mapping Note: this goes along with the assumption that both vocabularies could be mapped bilaterally,
- ...referring to an iterative mapping process of a first and a second iteration and an approach of vocabulary enrichment of the mapping with additional keywords (JEL) and subject headings from STW together with equivalent concept relations from past vocabulary mappings.



#### **Outline**

- Introduction
- Knowledge organization systems in economics
- Definition of interoperability and structural models for mapping
- Mapping process
- Empirical examples
- Results
- Conclusion and future outlook



#### Selection of STW subject categories:

- STW subthesaurus V Economics:
  - V.02 Microeconomics (1 subject category) V.02.01 - V.02.05 (5 s.c.)
  - V.15 Economic history (1 s.c.) V.15 - (-)
- STW subthesaurus B Business economics
  - B.07 Marketing (1 s.c.) B.07.01 - B.07.06 (6 s.c.)
  - B.09 Business information systems (1 s.c.) B.09.01 - B.09.03 (3 s.c.)



STW Relaunch Alphabetical descriptor list Mappings Web Services Downloads About

V Economics

V.00 Economics

V.01 Economic theory and methodology

V.02 Microeconomics

V.02.01 Household economics

V.02.02 Theory of the firm

V.02.03 Welfare economics

V.02.04 Economics of information

V.02.05 Economy of time

- V.03 Macroeconomics
- V.04 Business cycles and growth;
- economic structure
- V.05 Money and financial markets
- V.06 Economic systems
- V.07 International economics
- V.08 Development economics
- V.09 Public finance
- V.10 Industrial organization
- V.11 Regional science
- ▶ V.12 Environmental and resource
- economics
- V.13 Labour
- ▶ V.14 Social economics, education and health economics
  - V.15 Economic history

#### V.02 Microeconomics

V.02 Mikroökonomik (german)

#### broader

■ V Fconomics ▼

#### narrower

- V.02.01 Household economics ▼ ■ V.02.02 Theory of the firm
- V.02.03 Welfare economics ▼
- V.02.04 Economics of information
- V.02.05 Economy of time

#### Descriptors

- Allocative efficiency EB 🕏 ■ Behavioral economics EB 🥏
- Bounded rationality EB 🥏
- Comparative statics EB
- Elasticity EB 🥏
- Marginal analysis
   EB 🔊
- Microeconometrics 👍 🥏
- Microeconomics EB Opportunity cost EB
- Price elasticity EB Prospect theory EB
- Rationality EB
- Self-interest FR ■ Theory of value EB 🥏

Persistent Identifier (for bookmarking and linking)

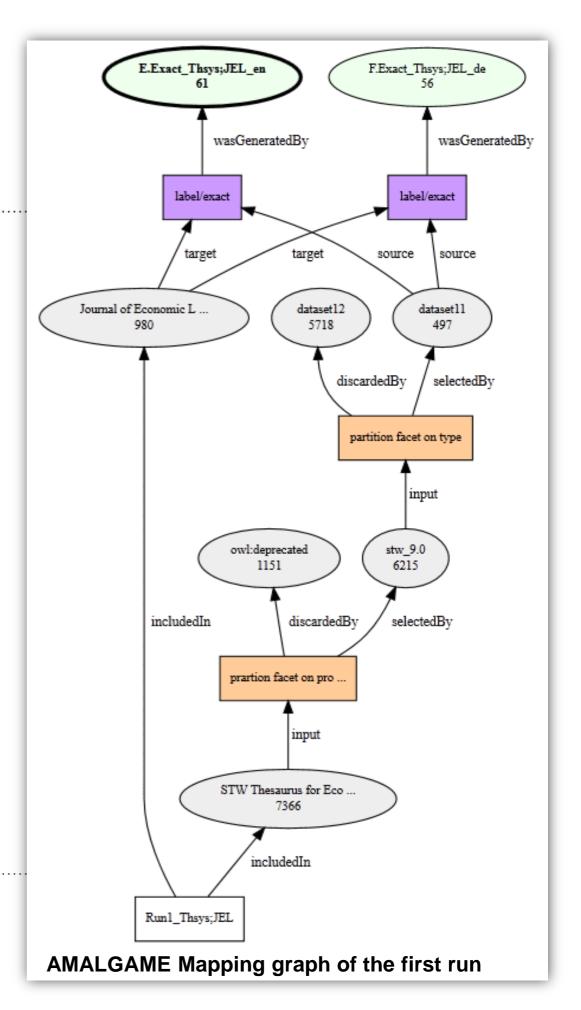
http://zbw.eu/stw/thsys/70180



#### Mapping procedure:

- Use of the interactive alignment server AMALGAME
   (AMsterdam ALignment GenerAtion MEtatool)
  - Upload of the STW (v 9.0) in SKOS
     <a href="http://zbw.eu/stw/versions/latest/downloa/d/about.de.html">http://zbw.eu/stw/versions/latest/downloa/d/about.de.html</a>
  - Upload of the JEL classification in SKOS <u>http://zbw.eu/beta/external\_identifiers/jel/</u>
     about.en.html
  - Exact language dependent string match of STW subject categories and JEL classes.





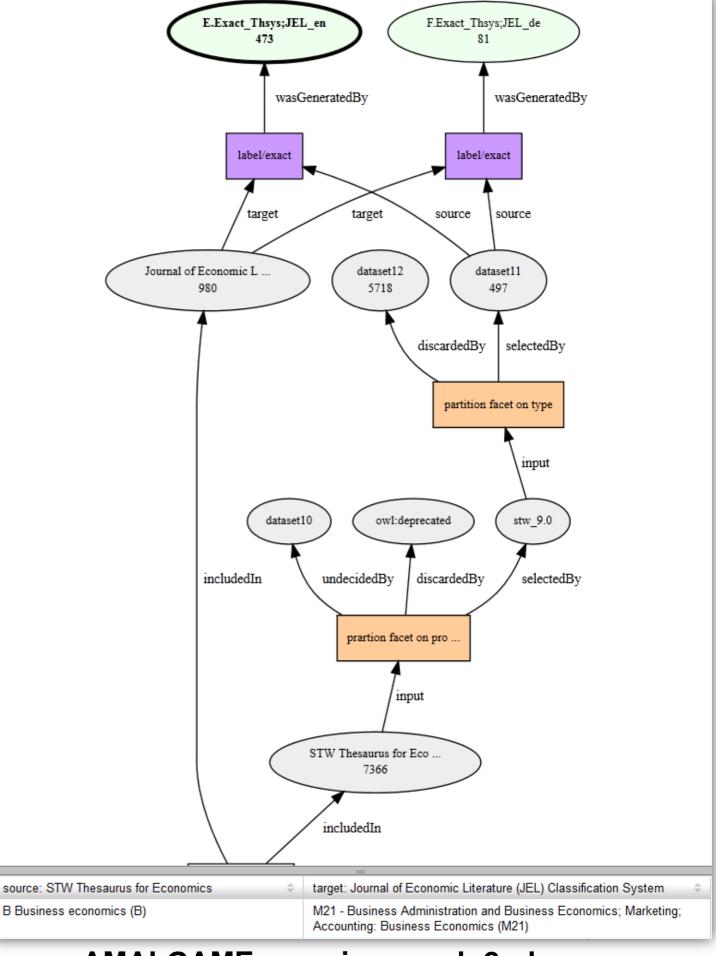
#### Second run:

- STW subthesaurus V:
  - V.02 Microeconomics
  - V.15 Economic history
- STW subthesaurus B:
  - B.07 Marketing
  - B.09 Business inform. Systems

## **Enrichment of STW subject categories and JEL classes:**

- STW subject categories by:
  - STW descriptors + synonyms
  - Vocabulary mapping relations (exactMatch) + synonyms (GND, TheSoz, DBpedia, AGROVOC)
- JEL classes by:
  - JEL keywords

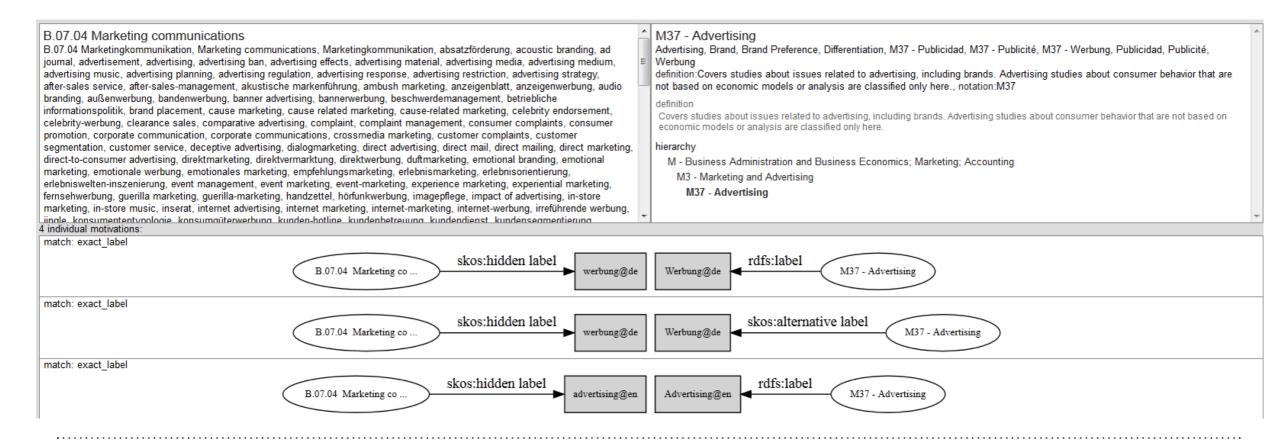




AMALGAME mapping graph 2nd run.

#### **Enrichment of STW subject categories and JEL classes:**

- STW subject categories by:
  - STW descriptors + synonyms
  - Vocabulary mapping relations (exactMatch) + synonyms (GND, TheSoz, DBpedia, AGROVOC)
- JEL classes by:
  - JEL keywords
     (The definitions do not get considered.)







#### **Outline**

- Introduction
- Knowledge organization systems in economics
- Definition of interoperability and structural models for mapping
- Mapping process
- Empirical examples
- Results
- Conclusion and future outlook





		Intell.	1 <sup>st</sup> run	(+)	Certain	Wrong (-)	2 <sup>nd</sup> run	(+)	Certain	Wrong (-)
V.02	Microeconomics	eval. 5	(=)		overlap ~		1 (13)	1	overlap ~	7
								•		'
V.02.01	Household economics	5					3 (21)		6	11
V.02.02	Theory of the firm	4					2 (10)	1		7
V.02.03	Welfare economics	6					3 (10)		2	5
V.02.04	Economics of information	4					2 (4)			2
V.02.05	Economy of time	7					1 (9)			8
V.15	Economic history	74					9 (15)		4	2
B.07	Marketing	3	1 (5)		4		2 (20)	1	1	
B.07.01	Marketing management	6					(-)			
B.07.02	Product Management	2					1 (15)		5	9
B.07.03	Pricing strategy	1					- (11)		2	9
B.07.04	Marketing communications	2					1 (2)			1
B.07.05	Distribution	1					- (12)		2	9
B.07.06	Market research	3					- (2)		1	1
B.09	Business information systems	1					- (5)			5
B.09.01	Information system components	1					- (4)		1	3
B.09.02	IS development and management	1					1 (6)			5
B.09.03	Corporate information systems	1					- (4)			4

.....





#### **Outline**

- Introduction
- Knowledge organization systems in economics
- Definition of interoperability and structural models for mapping
- Mapping process
- Empirical examples
- Results
- Conclusion and future outlook





#### Conclusion and future outlook

- String match can only generate mapping candidates it is blind to structural differences.
- The approach of vocabulary enrichment including JEL keywords, STW descriptors, synonyms, translations and equivalent terms and their synonyms from past vocabulary mappings led to a substantial increase of mapping candidates also included in the intellectual mapping.

  Note: A new use case for already established vocabulary alginments.
- Vocabulary enrichment has also led to new mapping candidates worth revisioning the current intellecutal mapping.
- Optional mapping procedure in the future:
   The STW as access vocabulary to JEL classes and "semantic bridge to a core system" (see Gödert 2010 and the CrissCross project).





## Thank you for your attention!

#### **Contact:**

Dr. Andreas Oskar Kempf <a href="mailto:a.kempf@zbw.eu">a.kempf@zbw.eu</a>

Joachim Neubert j.neubert@zbw.eu

Manfred Faden m.faden@zbw.eu

http:zbw.eu/stw







#### References

- BERTRAM, J. (2005) Einführung in die inhaltliche Erschließung. Ergon Verlag.
- GODERT, W. (2010) Semantische Wissensrepräsentation und Interoperabilität. In: Information Wissenschaft & Praxis, p. 5-28.
- HUBRICH, J. (2008) CRISSCROSS: SWD-DDC-Mapping. In: Mitteilungen der VÖB 61, Nr. 3, p.50-58.
- ISO 25964 Thesauri and interoperability with othr vocabularies. Part 1 (2011) +2 (2013).
- MAYR, P./PETRAS, V. (2007) Building a terminology network for search: the KoMoHe project. Proceedings of the International Conference on Dublin Core and Metadata Applications.
- SCHEVEN, E. (2013) ISO 25964 and GND, LIS'workshop 2013, Luxemburg



