

# Organizing Economic Information

---

An Overview of Application and Reuse Scenarios of an Economics Knowledge Organization System

---

***Tobias Rebholz**, Andreas Oskar Kempf, Joachim Neubert*  
*ZBW – Leibniz Information Centre for Economics –*  
*German National Library of Economics*  
*Beyond the Numbers, Federal Reserve Bank of St. Louis*  
*October 6-8, 2016*

# Contents

---

- 0. ZBW**
- 1. STW**
- 2. Subject indexing**
- 3. Inter-vocabulary mapping**
- 4. STW for information retrieval support**
- 5. STW web services**



# ZBW - Leibniz Information Centre for Economics

- World's largest information and research infrastructure for online/offline economic literature
- More than 4 million volumes and more than 26,000 periodicals and journals
- **ECONBIZ** - the search engine for economics  
*Find Economic Literature*
- **ECONSTOR** - publication server for scholarly economic literature  
*Make Your Publications Visible*
- Application-oriented research in the field of Science 2.0, Web Science, and Knowledge Discovery
- Financing: The ZBW is jointly financed by the federal and state governments



---

# 1. STW - Thesaurus for Economics

# STW - Thesaurus for Economics

---

## Institutional background

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

## Scope

- Covers all economics-related subject areas and the most important related subjects
- Comprises a systematic structure of domain-specific subject categories

## Maintenance & Development

- Regularly updated by an editorial team of domain experts from the ZBW



# STW - Thesaurus for Economics

## Structure



- Polyhierarchical

## Languages

- Bilingual: German & English




## Types of relations

- Equivalent relations, including synonyms and quasi-synonyms (UF)
- Hierarchical relations, including broader (BT) and narrower terms (NT)
- Associative relations, including related terms (RT)



**Financial statement** EB  

**Jahresabschluss** (german)  
used for: Annual accounts






Narrower Terms

- Balance sheet EB 
- Consolidated financial statements EB 
- Income statement EB 

Broader Terms

- Accounting EB 
- Reporting EB 

Related Terms

- Cash flow statement EB 
- Corporate annual report EB 
- Financial statement analysis EB 
- 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)

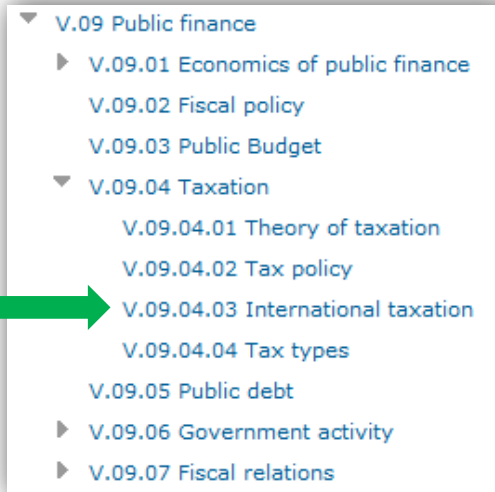
Persistent Identifier (for bookmarking and linking)

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

# STW subject categories

## Structural characteristics

- The STW subject categories (in total 497) constitute a monohierarchical structure
- Navigation tree
  - Seven main groups ▶ sub-thesauri
  - Allows thematically browsing in a certain subject field



*Navigation tree*

## V.09.04.03 International taxation

### V.09.04.03 Internationale Besteuerung (german)

broader

- V.09.04 Taxation ▼

Descriptors

- Cross-border taxation EB
- Cross-border tax evasion EB
- Destination principle EB
- Double taxation EB
- European corporate taxation EB
- EU tax law EB
- Export tax EB
- Group taxation EB
- International corporate taxation EB
- International tax law EB
- Rules of origin EB
- Tax competition EB
- Tax harmonization EB
- Tax haven EB

Persistent Identifier (for bookmarking and linking)

- <http://zbw.eu/stw/thsys/71077>





---

## 2. Subject indexing

# Overall indexing situation

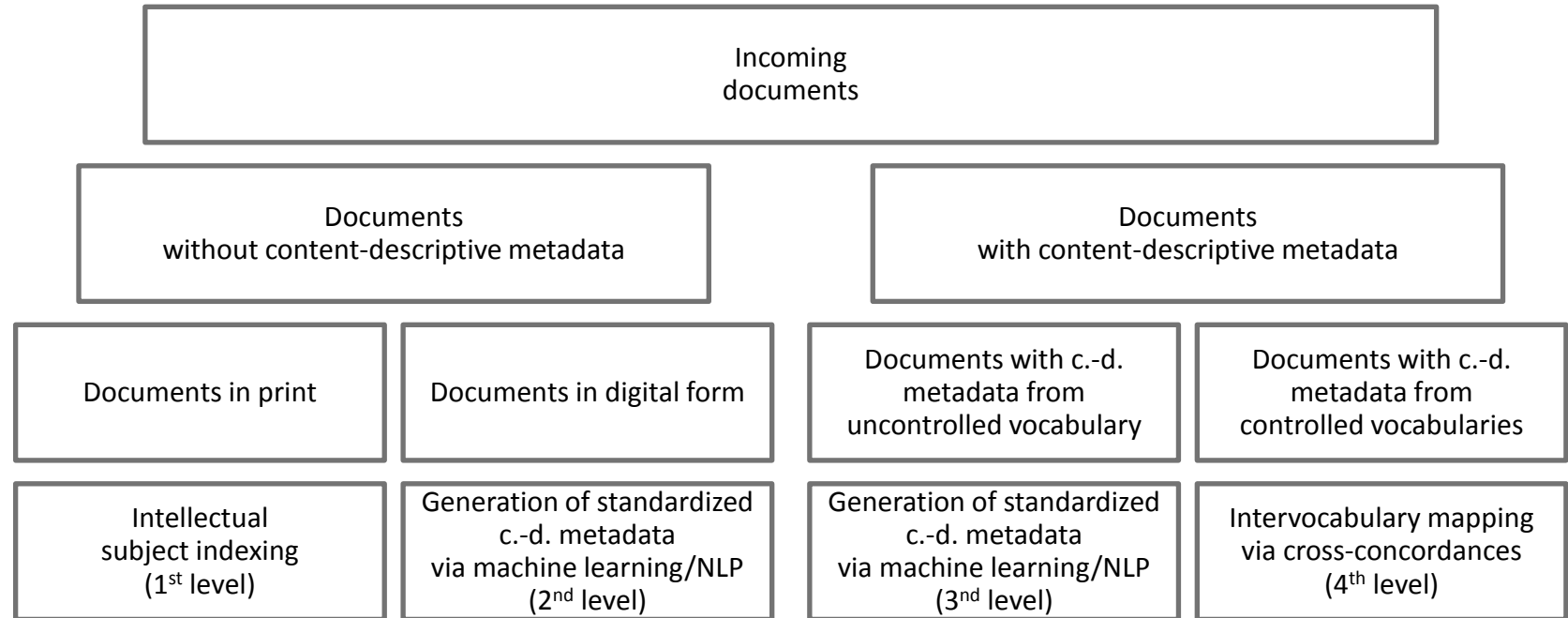
---

- **The process of subject indexing is changing more and more into an interplay between various partly interwoven indexing components, in which different indexing methods are applied.**

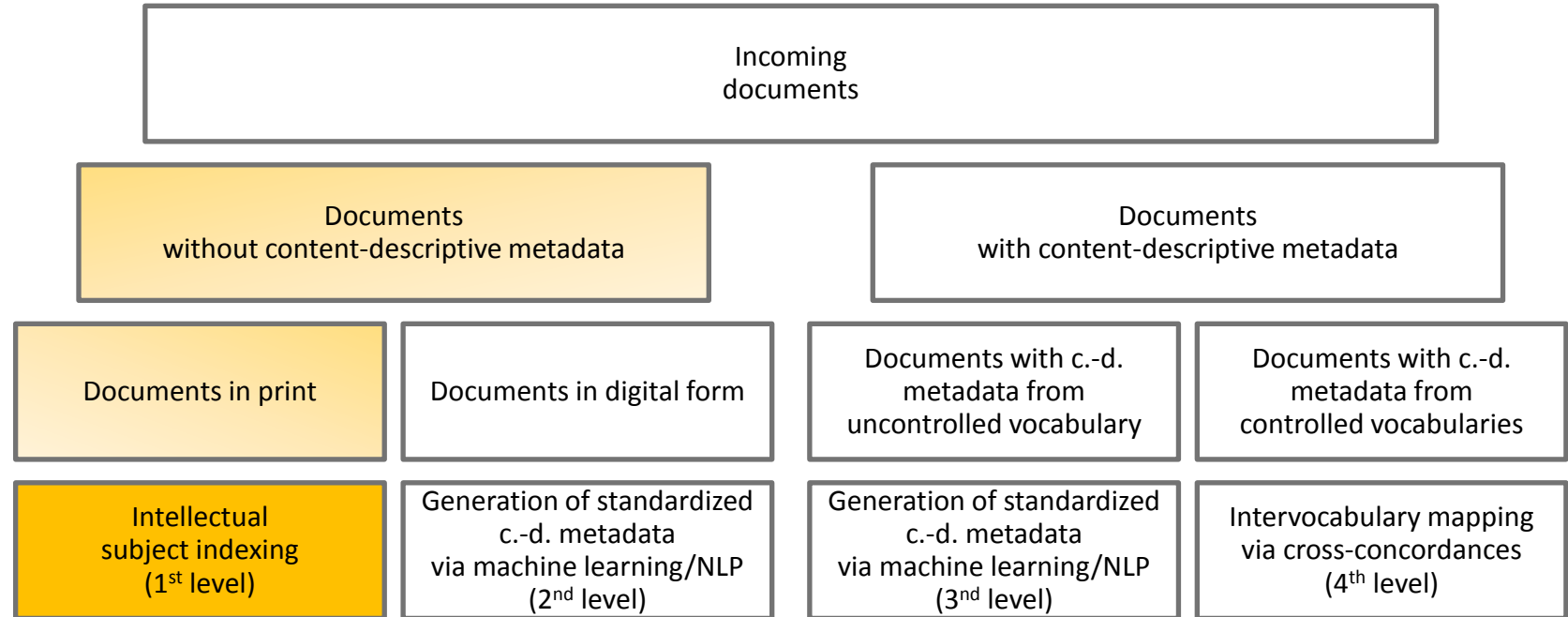


<http://www.lkrs.de/images/Startseite/puzzleschieber.jpg>

# Different indexing scenarios



# Different indexing scenarios (1<sup>st</sup> level)



# Intellectual subject indexing

## Intellectual subject indexing by information professionals (1<sup>st</sup> level):

- Limited to a subset of incoming documents
  - Subject metadata is not yet available
  - TDM on electronic fulltext is prohibited due to legal restrictions
- Used as the essential basis for further development of the thesaurus

The Finnish Great Depression: From Russia with Love<sup>®</sup>

By YURIY GORODNICHENKO, ENRIQUE G. MENDOZA, AND LINDA L. TESAR<sup>®</sup>

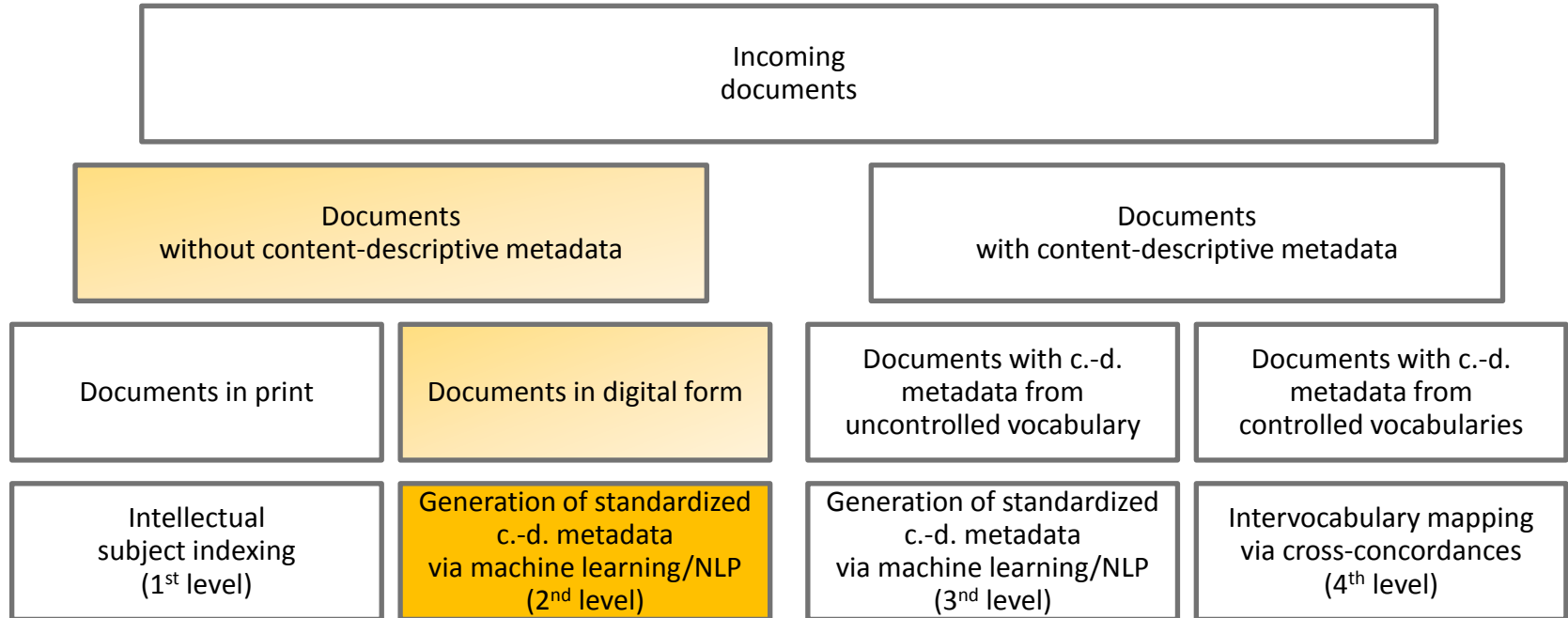
Understanding economic depressions has long been one of the central challenges in macroeconomics. Their massive costs as well as disagreement over their causes and propagation are subject to continuous debate. We examine the Finnish Great Depression of the early 1990s to shed new light on important transmission mechanisms that can drive depressions through disruption of international trade relationships. Our analysis of the Finnish Great Depression can also be useful for understanding the macroeconomic implications of large structural shocks affecting trade arrangements and the terms of trade in other countries, particularly in the case of the transition economies of Eastern Europe in the aftermath of the collapse of the Soviet Union.

During the 1991-1993 period, Finland experienced the deepest economic contraction in industrialized countries since the 1930s and the deepest recorded peacetime recession in Finland. As illustrated in panel A of Figure 1, real GDP fell between 1990 and 1993, real GDP declined by 11 percent, and consumption declined by 10 percent, and investment fell to 50 percent of its 1990 level. The declines were even more dramatic when we consider deviations from trend. In fact, this metric value added in the private sector fell about 20 percent below trend. Over the same period, Finland experienced a quadrupling of unemployment, from 6.5 percent to a peak of 16.5 percent, and the stock market lost 60 percent of its value.

We argue that a major cause of the Finnish Great Depression was the costly restructuring of the manufacturing sector and a sudden, sharp increase in energy costs caused by the demise of the Soviet Union. The barter-type trade arrangement between the Union of the Soviet Socialist Republics (USSR) and Finland skewed Finnish manufacturing production and investment toward particular industries and effectively allowed Finland to export noncompetitive products in exchange for energy imports at an overvalued exchange rate. The collapse of the USSR provides a unique natural experiment for which we know with precision the timing, nature, and size of the exogenous shocks that hit the Finnish economy. Furthermore, unlike previous

**florid title**  
**no keywords**  
**no abstract**

# Different indexing scenarios (2<sup>nd</sup> level)



# Subject indexing by computers (1)

---

## Subject indexing by computers based on fulltext/abstracts (2<sup>nd</sup> level):

- Based on text- and data-mining approaches, machine learning algorithms which are able to model human indexing behavior are tested and under active development
- Indexing assistant automatically suggests descriptors
- Quality management, which takes into account the quality of intellectual indexing

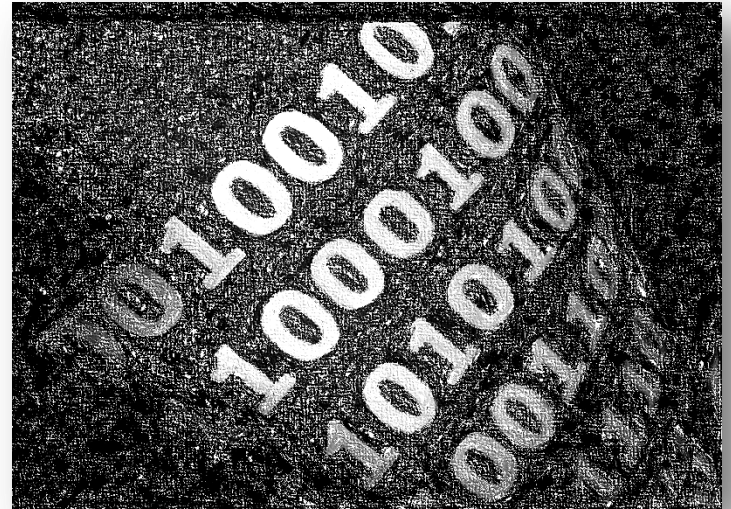
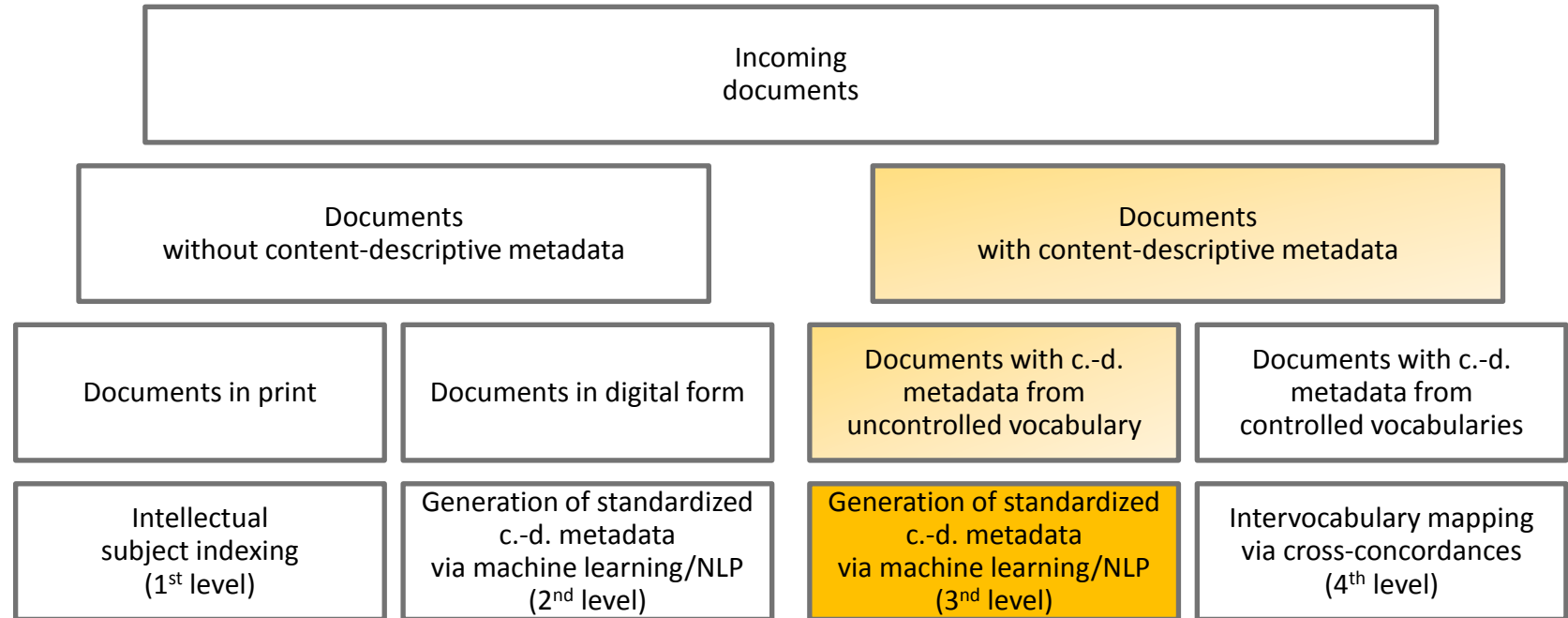


Photo: flickr: Hillary - CC BY-SA 2.0

# Different indexing scenarios (3<sup>th</sup> level)

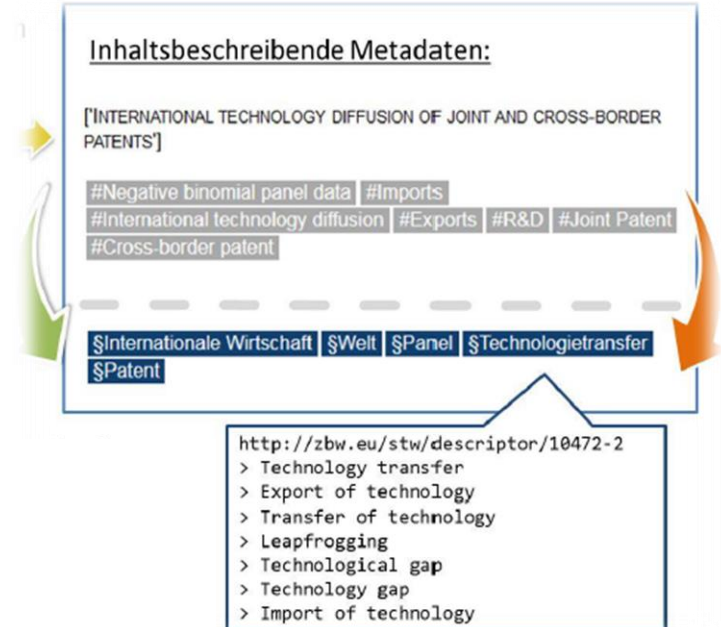




# Subject indexing by computers (2)

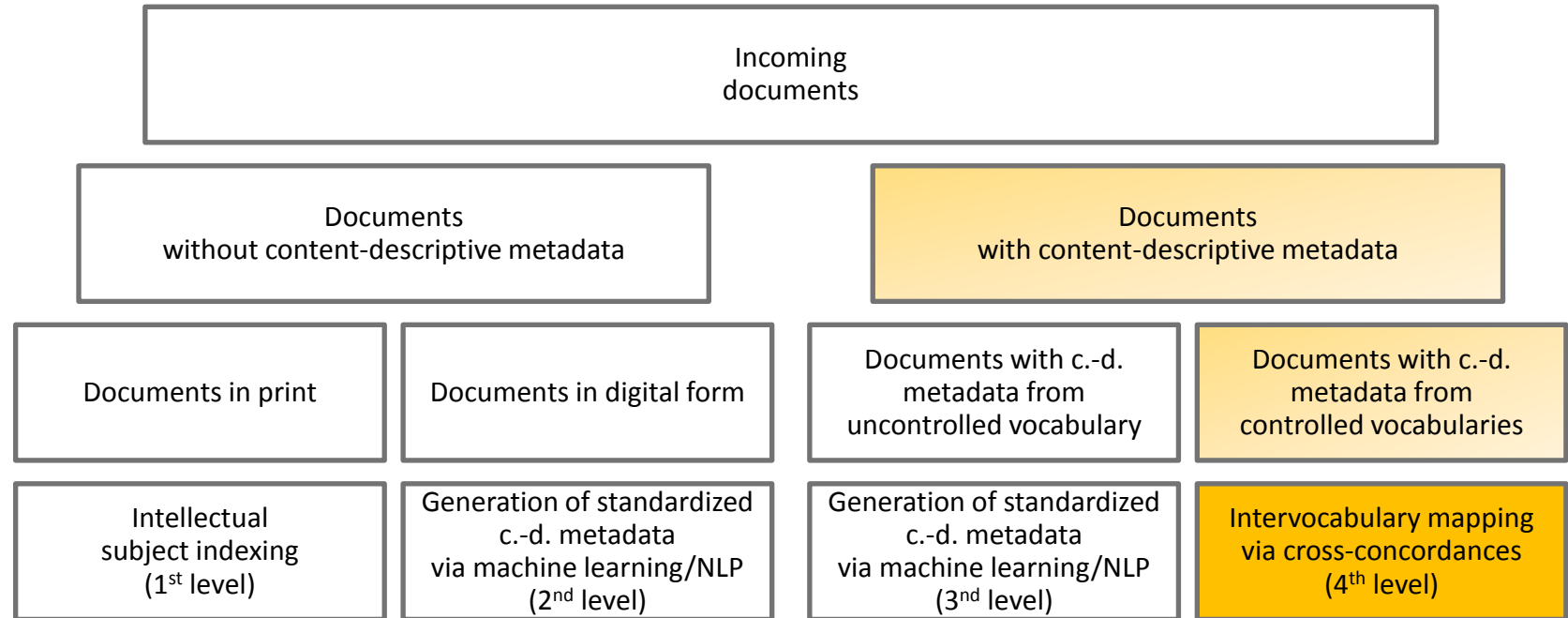
## Subject indexing by computers based on shorttext (3<sup>th</sup> level):

- Basic idea: to convert author keywords into STW subject headings based on string match and similarity metrics in combination with/without document title
- Now: Title + keyword indexing with STW
- Allows us to index print publications or publications without the rights for TDM
- 1,7 Million records (title + STW) as training corpus for this method



ZBW - Toepfer, Kempf 2016

# Different indexing scenarios (4<sup>th</sup> level)



# Cross-concordances

---

## Mappings to other vocabularies (4<sup>th</sup> level):

- Convert descriptors from other vocabularies into STW descriptors
- Supports truly collaboratively organized subject indexing beyond library boundaries
- Allows an integrated search across various databases indexed with different controlled vocabularies in search portals like EconBiz
- At the beginning mappings were built up exclusively intellectually
- Now: semi-automatic matching procedure

STW



INTEGRATED AUTHORITY FILE (GND)



AGROVOC

Thesaurus Sozialwissenschaften

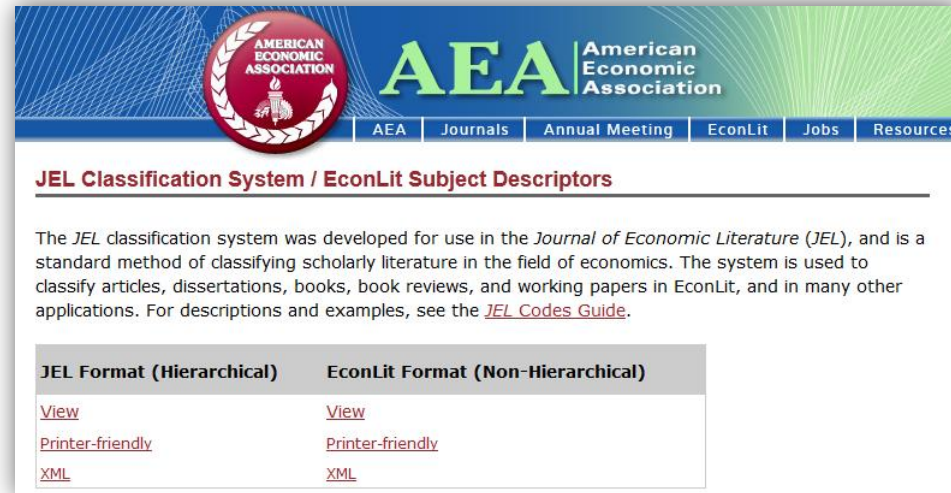
---

# 3. Inter-vocabulary mappings

# Inter-vocabulary mapping ► STW - JEL

## Reuse scenario for a JEL – STW mapping effort:

- Economists are usually quite familiar with the JEL classification codes
- Long-term objective: to animate economists to use STW subject headings in order to provide a more fine-grained content description with a standardized vocabulary



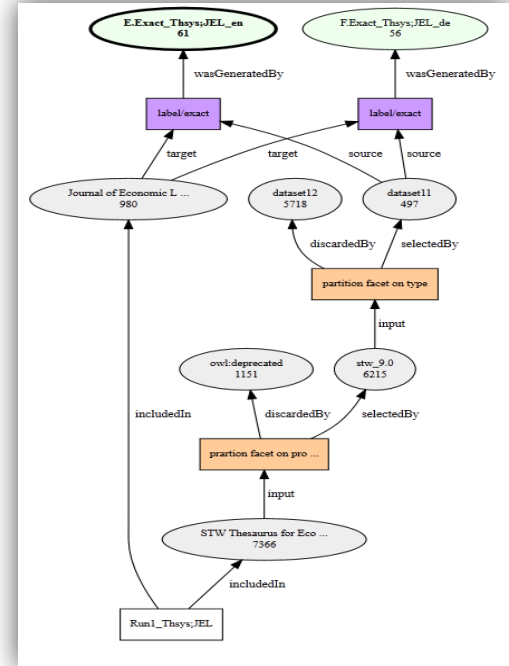
The screenshot shows the AEA website header with the logo and navigation links: AEA, Journals, Annual Meeting, EconLit, Jobs, Resources. The main heading is "JEL Classification System / EconLit Subject Descriptors". Below this, a paragraph explains that the JEL system was developed for use in the *Journal of Economic Literature* (JEL) and is a standard method of classifying scholarly literature in economics. It is used to classify articles, dissertations, books, book reviews, and working papers in EconLit, and in many other applications. For descriptions and examples, see the [JEL Codes Guide](#).

JEL Format (Hierarchical)	EconLit Format (Non-Hierarchical)
<a href="#">View</a>	<a href="#">View</a>
<a href="#">Printer-friendly</a>	<a href="#">Printer-friendly</a>
<a href="#">XML</a>	<a href="#">XML</a>

# STW – JEL Mapping (1)

## Mapping procedure

- Use of the interactive alignment server **AMALGAME** (AMsterdam ALignment GenerAtion METatool)
- Iterative semi-automatic mapping process
- SKOS vocabulary needed
- Enrichment of STW subject categories and JEL classes
- Exact language dependent string match of STW subject categories and JEL classes.
- Evaluation tool - subsets of alignments can be evaluated manually



# STW – JEL Mapping (2)

## STW subject categories enriched by

- STW descriptors + synonyms
- Mapped (exactMatch) concepts from other vocabularies – descriptors + synonyms

## JEL classes enriched by

- JEL keywords scraped from JEL guide <https://www.aeaweb.org/jel/guide/jel.php>

B.07.04 Marketing communications  
B.07.04 Marketingkommunikation, Marketing communications, Marketingkommunikation, absatzförderung, acoustic branding, ad journal, advertisement, advertising, advertising ban, advertising effects, advertising material, advertising media, advertising medium, advertising music, advertising planning, advertising regulation, advertising response, advertising restriction, advertising strategy, after-sales service, after-sales-management, akustische markenführung, ambush marketing, anzeigenblatt, anzeigenwerbung, audio branding, außenwerbung, bandenwerbung, banner advertising, bannerwerbung, beschwerdemanagement, betriebliche informationspolitik, brand placement, cause marketing, cause related marketing, cause-related marketing, celebrity endorsement, celebrity-werbung, clearance sales, comparative advertising, complaint, complaint management, consumer complaints, consumer promotion, corporate communication, corporate communications, crossmedia marketing, customer complaints, customer segmentation, customer service, deceptive advertising, dialogmarketing, direct advertising, direct mail, direct mailing, direct marketing direct-to-consumer advertising, direktmarketing, direktvermarktung, direktwerbung, duftmarketing, emotional branding, emotional marketing, emotionale werbung, emotionales marketing, empfehlungsmarketing, erlebnismarketing, erlebnisorientierung, erlebniswelten-inszenierung, event management, event marketing, event-marketing, experience marketing, experiential marketing, fernsehwerbung, guerilla marketing, guerilla-marketing, imagezettel, hörfunkwerbung, imagepflege, impact of advertising, in-store marketing, in-store music, inserat, internet advertising, internet marketing, internet-marketing, internet-werbung, irreführende werbung, inplane, konsummentenpsychologie, konsumüterwerbung, kunden-hotline, kundenbetreuung, kundendienst, kundensegmentierung, individual motivations, match: exact\_label

M37 - Advertising  
Advertising, Brand, Brand Preference, Differentiation, M37 - Publicidad, M37 - Publicité, M37 - Werbung, Publicidad, Publicité, Werbung  
definition:Covers studies about issues related to advertising, including brands. Advertising studies about consumer behavior that are not based on economic models or analysis are classified only here., notation:M37  
definition  
Covers studies about issues related to advertising, including brands. Advertising studies about consumer behavior that are not based on economic models or analysis are classified only here.  
hierarchy  
M - Business Administration and Business Economics; Marketing: Accounting  
M3 - Marketing and Advertising  
M37 - Advertising



---

# 4. STW for information retrieval support



# Index enhancement by search engine

---

## Optimizing the search engine behind EconBiz

- Fast access to the title records is provided by a customized open-source search engine (Solr)
- For the field with STW descriptors, index entries are produced with all synonyms of the descriptors
- Synonym enhancement is active for all “simple” searches
- Main advantage: index-enhancement delivers search results very quickly
- Boosting rules can be applied, which for example rank results higher when a search term appears in the title field of the record or in a thesaurus-enhanced keyword field

Solr 

vufind 

**ECONBIZ**  
*Find Economic Literature*

---

# 5. STW web services

# Linked data – STW web services

The ZBW has early on published the STW data additionally in the form of web services

- STW web services execute predefined and **pre-optimized SPARQL queries** - each service for a particular use case
- **Synonym service** returns the alternative search terms, optionally including synonyms derived from the mappings to other vocabularies
- **Autosuggest service** supports data input

<http://zbw.eu/beta/econ-ws>

## Table of Contents

- General Concepts
- Terminology Services
  - /suggest - Suggestions for Resources (starting with a given string)
  - /suggest2 - Suggestions for Resources (optimized for large datasets)
  - /synonyms - Alternative Terms for a Search Term (from matching labels)
  - /combined1 - Combined Info for Retrieval Applications
  - /mappings - Mappings for a Concept
  - /concept-history - History of a SKOS Concept
- Auxiliary Services
  - /assigned\_descriptors - Retrieves the descriptors assigned to a work
- *Deprecated Services*
  - /concepts - Search for Thesaurus Concepts
  - /narrower - Narrower Concepts for a Concept
  - /labels - All Labels for a Concept
- Common Parameters
- Error Handling
- Implementation
- Changes

# Thank you for your attention!

---

## Contact:

**Tobias Rebholz**

[t.rebholz@zbw.eu](mailto:t.rebholz@zbw.eu)

**Andreas Oskar Kempf**

[a.kempf@zbw.eu](mailto:a.kempf@zbw.eu)

**Joachim Neubert**

[j.neubert@zbw.eu](mailto:j.neubert@zbw.eu)

<http://zbw.eu/stw/>

<http://www.zbw.eu/en/stw-info/>

[Join the EconBiz Partner Network](#)

