
Maschinelle Indexierung in den Wirtschaftswissenschaften

Zum praktischen Einsatz automatisierter Erschließungsverfahren und ihre Auswirkungen auf das Thesaurusmanagement

*Dr. Andreas Oskar Kempf,
ZBW – Leibniz-Informationszentrum Wirtschaft
DGI Praxistage, Frankfurt am Main, 09. November 2018*

Gliederung

- Kontext des ZBW-internen Projekts Automatische Indexierung
- Entstehungs-, Veränderungs- & Verwendungszusammenhänge des STW
Exkurs: Einblick in Entwicklungsperspektiven der Fachreferatsarbeit
- Ansätze der Weiterentwicklung des Thesaurus vor dem Hintergrund des Einsatzes automatisierter Sacherschließungsverfahren
- (Zwischen-)Fazit

Kontext des ZBW-internen Projekts Automatische Indexierung

ZBW-externe & -interne Rahmenbedingungen

- Erhöhtes Publikationsaufkommen
- Fachspezifisches Publikationsverhalten

Kontext des ZBW-internen Projekts Automatische Indexierung

ZBW-externe & -interne Rahmenbedingungen

- Erhöhtes Publikationsaufkommen
 - Fachspezifisches Publikationsverhalten
 - Rückgang der Personalressourcen im Fachreferat
 - Ausbau der ZBW zu Forschungsbibliothek
 - Neukonzeption: Aufbau einer eigenen maschinellen Erschließungsinfrastruktur
 - *Projektziel:*
Aufbau eigener Kompetenzen im Bereich Text- und Data-Mining, die als Wissenstransfer in forschungsbasierte anwendungsorientierte Bibliotheksservices einfließen
-

Kontext des ZBW-internen Projekts Automatische Indexierung

Sacherschließungskonzept

Verschiedene Sacherschließungsebenen:

- Intellektuelle Sacherschließung
- Fremddatenübernahme
- Automatische Sacherschließung

Aktuell: Forschungsbasierte Weiterentwicklung und Evaluation von Kurztext-Verfahren, die eine inhaltliche Erschließung von Publikationen auf Basis von Titeln und Autoren-Keywods ermöglichen (vgl. Toepfer/Kempf 2016; Toepfer 2017)

- *Ausblick: Test der semiautomatischen Sacherschließung mit dem DA 3 & Integration forschungsbasierter automatischer Sacherschließungsverfahren*
-

Standard Thesaurus-Wirtschaft (STW)

Erfahrungen im Umgang mit unterschiedlichen Veränderungsprozessen

- Technologischer Wandel & Wandel der Informationsumgebung
 - Geänderte Nutzerbedürfnisse
- Öffnung des Informationssystems & Systemöffnung durch Web-Publikation
 - Diversifikation der Anwendungskontexte
- Wandel in der Fachterminologie
 - Kontinuierlicher Bedarf an Pflege & Weiterentwicklung des Vokabulars

Standard Thesaurus-Wirtschaft (STW)

Entstehungskontext – Wandel der Informationsumgebung

Produktion einer gemeinsamen CD-ROM zur
simultanen Recherche OHNE einheitliche Inhaltserschließung



Kooperationsprojekt (1995-1997):

„Vereinheitlichung der Wirtschaftsnomenklatur zu einer gemeinsamen Indexierungssprache (Thesaurus)...“ (vgl. Gastmeyer 2000: 108)

Projektziele:

- Institutionenübergreifende internationale Vereinheitlichung der Inhaltserschließung in den Wirtschaftswissenschaften (terminologische Normierung)

Standard Thesaurus-Wirtschaft (STW)

Informationssystem

- **Fachliche Abdeckung:** Bilingualer Fachthesaurus für die Wirtschaftswissenschaften und angrenzende Fachgebiete (u.a. Recht, Technik)
- **Allgemeinheit/Spezifität:** Kernbereich relativ spezifisch; Randbereiche relativ allg. (vgl. GND)
- **Umfang:** Knapp 6.000 Deskriptoren & 20.000 Synonymverweise
- **Strukturierung:** Subthesauri, umfangreiches Zugangsvokabular (Äquivalenzklassen)

Standard Thesaurus-Wirtschaft (STW)

Informationssystem – Strukturkonzeption

Zum Teil umfangreiches Zugangsvokabular in der sog.

Äquivalenzklasse:

Zusammenfassung bedeutungsgleicher (Synonyme) bzw. - ähnlicher Bezeichnungen (Quasi-Synonyme) zu einem Begriff. Von diesen Bezeichnungen wird eine Variante zur Vorzugsbezeichnung bestimmt, die den Begriff, bzw. die gedankliche Einheit, repräsentiert. (Beziehung zw. Bezeichnung(en) und Begriff)

Im Einzelnen können dies sein:

- Lexikalische Varianten (Schreibweisen, z.B. alte/neue Rechtschreibung, Akronyme)
- Quasisynonyme (weibliche & männliche Form)
- „Synonyme“, die strenggenommen UB, OB, VB sind.

The screenshot shows the ZBW website interface for the Standard Thesaurus Wirtschaft (STW). The page is titled 'Elektronisches Geld' and includes the following information:

- Navigation:** Home, STW Relaunch, Alphabetische Liste der Deskriptoren, Mappings, Versionen, Web Services, Downloads, Mehr zum STW.
- Category Navigation:** V Volkswirtschaft, B Betriebswirtschaft, W Wirtschaftssektoren, P Produkte, N Nachbarwissenschaften, G Geographische Begriffe, A Allgemeinwörter.
- Elektronisches Geld EB**
- Electronic money (engl.)**
- benutzt für:** Mobiles Zahlungssystem, Internet-Zahlungsverkehr, Elektronisches Zahlungssystem, Elektronisches Zahlungsmittel, E-Geld, Mobile payment, Mobile money, M-payment, Electronic payment, Electronic cash, E-payment.
- Unterbegriffe:** Virtuelle Währung EB.
- Oberbegriffe:** Bargeldloser Zahlungsverkehr EB, Geld EB.
- Verwandte Begriffe:** Chipkarte EB, Electronic Banking EB, Online-Handel EB.
- Thesaurus Systematik:** V.05.01 Geld und Währungssystem.
- Links zu anderen Thesauri und Vokabularen:** Electronic Cash (aus GND), Point of Sale Banking (aus GND), Elektronischer Zahlungsverkehr (aus GND), Electronic Cash (aus TheSoz).
- Persistenter Identifier (für Bookmarks und zum Verlinken):** <http://zbw.eu/stw/descriptor/18022-0>

Standard-Thesaurus Wirtschaft (v 9.06, 2018-08-15) • Anregungen und Kommentare an das [Thesaurus Team](#) • Mailinglisten: [stw-announce](#), [stw-user](#)
ZBW - Leibniz-Informationszentrum Wirtschaft - Impressum

Der Standard-Thesaurus Wirtschaft steht unter der Open Database License (ODbL) 1.0. Darüber hinausgehende Rechte können Sie durch die ZBW erhalten.

Standard Thesaurus-Wirtschaft (STW)

Öffnung des Informationssystems – Interoperabilität

Abbildung des STW auf andere Normdaten auf Basis von Äquivalenz-, Ober-/Unterbegriff- und Assoziationsrelation (Crosskonkordanz).

- Erweiterung des Zugangsvokabulars zur **integrierten inhaltlichen Suche**

id	concept	synonyms
4042669-5	Notenbank	Central bank Zentralbank National bank Nationalbank Notenbank Reserve bank central bank

id	concept	synonyms
10774-0	Zentralbank	Central bank Zentralbank National bank Nationalbank Notenbank Reserve bank central bank

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

Zentralbank EB

Central bank (engl.)

benutzt für: Notenbank, Nationalbank, Reserve bank, National bank

Oberbegriffe

- Offentliche Verwaltung EB

Verwandte Begriffe

- Lender of Last Resort EB
- Zentralbankunabhängigkeit EB

Thesaurus Systematik

- V.05.01.02 Geldpolitik
- W.14.01 Banken

<http://zbw.eu/stw/descriptor/10774-0>

Link zu diesem Datensatz	http://d-nb.info/gnd/4042669-5
Sachbegriff	Notenbank
Quelle	M
synonyme	Währungsbank Zentralbank Zentralnotenbank Zettelbank Nationalbank
Oberbegriffe	Bank
DDC-Notation	332.11

<http://d-nb.info/gnd/4042669-5>

Standard Thesaurus-Wirtschaft (STW)

Öffnung des Informationssystems – Interoperabilität

STW-Konkordanzen (Mappings) zu:

- **Gemeinsame Normdatei (DNB)**
- **DBpedia & Wikidata** (in Arbeit)
- **Thesaurus Sozialwissenschaften (GESIS)**,
- **AGROVOC (FAO)**
- **EuroVoc (EU) & JEL-Klassifikation (Journal of Economic Literature)**



The screenshot shows the ZBW website interface. At the top left is the ZBW logo and the text 'Leibniz-Informationszentrum Wirtschaft' and 'Leibniz Information Centre for Economics'. At the top right are links for 'RDF/XML', 'RDF/Turtle', and 'English'. Below the logo is a search bar. On the left side, there is a navigation menu with links: Home, STW Relaunch, Alphabetsche Liste der Deskriptoren, Mappings, Versionen, Web Services, Downloads, and Mehr zum STW. Below this menu is a list of categories: V Volkswirtschaft, B Betriebswirtschaft, W Wirtschaftssektoren, P Produkte, N Nachbarwissenschaften, G Geographische Begriffe, and A Allgemeinwörter. On the right side, the main heading is 'Standard-Thesaurus Wirtschaft' and 'STW Mappings'. Below this, there is a paragraph: 'Hier finden Sie Crosskonkordanzen zu anderen Thesauri und sonstigen Vokabularen, die auch heruntergeladen werden können.' Below this paragraph is a list of mapping options: Gemeinsame Normdatei (GND), Wikidata, DBpedia, Thesaurus Sozialwissenschaften (TheSoz), AGROVOC, WKD - Arbeitsrechtthesaurus, JEL Klassifikation, and SDMX Subject-Matter Domains Klassifikation. At the bottom of the page, there is a footer with the text: 'Standard-Thesaurus Wirtschaft (v 9.06, 2018-08-15) • Anregungen und Kommentare an das Thesaurus Team • Mailinglisten: stw-announce, stw-user ZBW - Leibniz-Informationszentrum Wirtschaft - Impressum'. There is also a logo for 'W3C XHTML + RDFa' and a small text: 'Der Standard-Thesaurus Wirtschaft steht unter der Open Database License (ODbL) 1.0. Darüber hinausgehende Rechte können Sie durch die ZBW erhalten.'

<http://zbw.eu/stw/versions/latest/mapping/about.de.html>

Standard Thesaurus-Wirtschaft (STW)

Systemöffnung

Web-Publikation des STW im SKOS-Format

➤ STW-Zugang und Suche im Web

<http://zbw.eu/stw/version/latest/about.de.html>

The screenshot shows the website for the Standard Thesaurus Wirtschaft. At the top left is the ZBW logo and the text 'Leibniz-Informationszentrum Wirtschaft Leibniz Information Centre for Economics'. At the top right are links for 'RDF/XML', 'RDF/Turtle', and 'English'. Below the logo is a search bar. On the left side, there is a navigation menu with links: Home, STW Relaunch, Alphabetische Liste der Deskriptoren, Mappings, Versionen, Web Services, Downloads, Mehr zum STW, and a list of categories: V Volkswirtschaft, B Betriebswirtschaft, W Wirtschaftssectoren, P Produkte, N Nachbarwissenschaften, G Geographische Begriffe, and A Allgemeinwörter. The main content area features the title 'Standard-Thesaurus Wirtschaft' and 'Version 9.06 (Änderungsübersicht)'. Below this is a 'Subthesauri' section with a list of categories: V Volkswirtschaft, B Betriebswirtschaft, W Wirtschaftssectoren, P Produkte, N Nachbarwissenschaften, G Geographische Begriffe, and A Allgemeinwörter. A highlighted section titled 'Der Standard-Thesaurus Wirtschaft erleichtert mit seinen genannten Schlagwörtern die thematische Suche im Fachportal EconBiz (per EB Icon)' contains text about the thesaurus's vocabulary and search capabilities. Below this is another section titled 'Die Thesaurus-Informationen sind zusätzlich als Daten im RDFa-Format in die Webseiten eingebettet, um eine vielfältige Nachnutzung durch die wirtschaftswissenschaftliche Fachgemeinschaft und im Semantic Web zu unterstützen.' At the bottom, there is a footer with contact information, a mailing list link, and a Creative Commons BY-NC-ND license notice.

Leibniz-Informationszentrum
Wirtschaft
Leibniz Information Centre
for Economics

RDF/XML RDF/Turtle English

Standard-Thesaurus Wirtschaft
Version 9.06 (Änderungsübersicht)

Subthesauri

- V Volkswirtschaft
- B Betriebswirtschaft
- W Wirtschaftssectoren
- P Produkte
- N Nachbarwissenschaften
- G Geographische Begriffe
- A Allgemeinwörter

Der Standard-Thesaurus Wirtschaft erleichtert mit seinen genannten Schlagwörtern die thematische Suche im Fachportal EconBiz (per EB Icon).

Der Thesaurus enthält Vokabular zu allen ökonomischen Themenstellungen: knapp 6.000 Schlagwörter und über 20.000 zusätzliche Synonyme als Zugangshilfe, um individuelle Formulierungen bei der Suche aufzufangen. Auch Fachbegriffe aus benachbarten Bereichen wie z.B. Recht, Soziologie oder Politik sowie Geobegriffe sind dort zu finden. Wenn Sie Schlagwörter aus diesem Vokabular auswählen, können Sie sicher sein Treffer zu erhalten, die zu dem gewünschten Sachgebiet passen.

Die Thesaurus-Informationen sind zusätzlich als Daten im RDFa-Format in die Webseiten eingebettet, um eine vielfältige Nachnutzung durch die wirtschaftswissenschaftliche Fachgemeinschaft und im Semantic Web zu unterstützen.

Standard-Thesaurus Wirtschaft (v 9.06, 2018-08-15) • Anregungen und Kommentare an das Thesaurus Team • Mailinglisten: [stw-announce](#), [stw-user](#)
ZBW - Leibniz-Informationszentrum Wirtschaft - [Impressum](#)

W3C Der Standard-Thesaurus Wirtschaft steht unter der Open Database License (ODbL) 1.0. Darüber hinausgehende Rechte können Sie durch die ZBW erhalten.

Standard Thesaurus-Wirtschaft (STW)

Systemöffnung

Web-Publikation des STW im SKOS-Format

- STW-Zugang und Suche im Web
- Download-Angebot und maschinelle
Nachnutzung (ODbL 1.0)

<http://zbw.eu/stw/version/latest/download/about.de.html>

The screenshot shows the ZBW website interface. At the top left is the ZBW logo and the text 'Leibniz-Informationszentrum Wirtschaft' and 'Leibniz Information Centre for Economics'. At the top right are links for 'RDF/XML', 'RDF/Turtle', and 'English'. Below the logo is a navigation menu with items: Home, STW Relaunch, Alphabetische Liste der Deskriptoren, Mappings, Versionen, Web Services, Downloads, and Mehr zum STW. A secondary menu lists categories: V Volkswirtschaft, B Betriebswirtschaft, W Wirtschaftssektoren, P Produkte, N Nachbarwissenschaften, G Geographische Begriffe, and A Allgemeinwörter. The main content area is titled 'Standard-Thesaurus Wirtschaft' and includes a 'Downloads' section with a link for 'Anmeldung zur Mailingliste stw-announce'. Below this is a registration form with fields for 'E-Mail:', 'Name:', and 'Institution:', and a larger text area for 'vorgesehene Verwendung:'. A disclaimer text is provided below the form, followed by an 'absenden' button. Further down, there is a 'Download Dateien' section for 'STW v 9.06 (SKOS)' with links for 'RDF/XML', 'Ntriples', and 'Turtle (besser lesbar)'. At the bottom, there is a 'Mapping GND - STW (SKOS)' section with similar links and a note about its creation by DNB and ZBW.

Standard Thesaurus-Wirtschaft (STW)

Systemöffnung

Web-Publikation des STW im SKOS-Format

- STW-Zugang und Suche im Web
- Download-Angebot und maschinelle Nachnutzung (ODbL 1.0)
- Zuordnung einer eindeutigen HTTP-Adresse als persistentem Identifikator zu jedem Deskriptor...

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

Umweltbericht EB   

Environmental reporting (engl.)

benutzt für: Ökobilanzierung, Ökobilanz, Umweltbilanz, Umweltberichterstattung, Life cycle assessment, Environmental management accounting, Environmental disclosure, Corporate environmental accounting

Oberbegriffe

- Nachhaltigkeitsbericht EB 

Verwandte Begriffe

- Öko-Audit EB 
- Umweltbewertung EB 
- Umwelt-Informationssystem EB 
- Umweltökonomische Gesamtrechnung EB 

Thesaurus Systematik

- B.01.03 Betriebliche Information und Kommunikation
- ▼
- B.01.06 Umweltmanagement ▼

Links zu anderen Thesauri und Vokabularen

- = Umweltbericht (aus GND)
- ~ Umweltbilanz (aus GND)
- ≡ Umweltökonomie (aus DBpedia) W
- = Ökobilanz (aus Thesoz)
- = Umweltgesamtrechnung (aus Agrovoc)

Persistent Identifier (für Bookmarks und zum Verlinken)

- <http://zbw.eu/stw/descriptor/19090-6>

Standard Thesaurus-Wirtschaft (STW)

Systemöffnung

Web-Publikation des STW im SKOS-Format

- STW-Zugang und Suche im Web
- Download-Angebot und maschinelle Nachnutzung (ODbL 1.0)
- Zuordnung einer eindeutigen HTTP-Adresse als persistentem Identifikator zu jedem Deskriptor und Verknüpfung mit semantischen Inhalten

The screenshot shows the web interface for the 'Umweltbericht' descriptor. At the top, there are navigation links for 'RDF/XML', 'RDF/Turtle', 'Concept history (RDF/Turtle)', and 'English'. The main content area displays the descriptor 'Umweltbericht' with its English translation 'Environmental reporting' and a list of related terms. Below this, there is a section for 'Oberbegriffe' (superordinate terms) with a link to 'Nachhaltigkeitsbericht'. A code block shows the RDF/XML representation of the descriptor, including its URI and SKOS exact match relationships with other thesauri. At the bottom, there are links to other thesauri and a persistent identifier for bookmarks.

```
<rdf:RDF xml:base="http://zbw.eu/stw/">
  <rdf:Description rdf:about="descriptor/19090-6">
    <skos:exactMatch rdf:resource="http://d-nb.info/gnd/4202417-1"/>
  </rdf:Description>
  <rdf:Description rdf:about="descriptor/19090-6">
    <skos:exactMatch rdf:resource="http://lod.gesis.org/thesoz/concept/10039271"/>
  </rdf:Description>
  <rdf:Description rdf:about="descriptor/19090-6">
    <skos:exactMatch rdf:resource="http://aims.fao.org/aos/agrovoc/c_9000050"/>
  </rdf:Description>
</rdf:RDF>
```

Links zu anderen Thesauri und Vokabularen

- = Umweltbericht (aus GND)
- ~ Umweltbilanz (aus GND)
- ≡ Umweltökonomie (aus DBpedia) W
- ≡ Ökobilanz (aus Thesoz)
- = Umweltgesamtrechnung (aus Agrovoc)

Persistenter Identifier (für Bookmarks und zum Verlinken)

- <http://zbw.eu/stw/descriptor/19090-6>

Standard Thesaurus-Wirtschaft (STW)

Systemöffnung

Web-Publikation des STW im SKOS-Format

- STW-Zugang und Suche im Web
- Download-Angebot und maschinelle Nachnutzung (ODbL 1.0)
- Zuordnung einer eindeutigen HTTP-Adresse als persistentem Identifikator zu jedem Deskriptor und Verknüpfung mit semantischen Inhalten
- Einbettung der Inhalte in Webseiten mgl.

<https://tinyurl.com/y7shrmly>

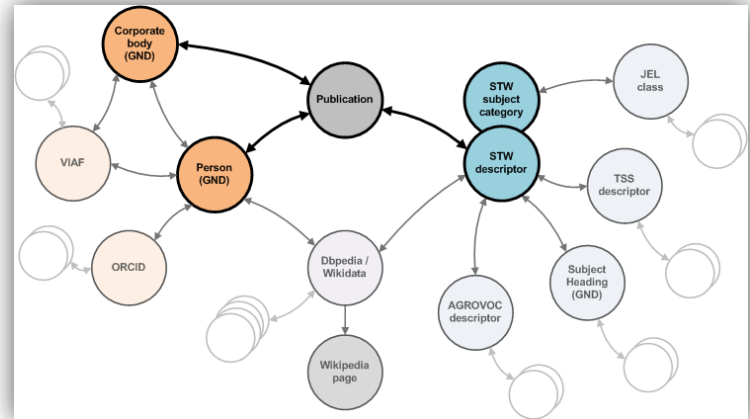
The screenshot displays the EconStor search results page. At the top, the EconStor logo and tagline 'Make Your Publications Visible.' are visible. Below the logo, there is a navigation bar with links for 'STARTSEITE', 'ÜBER ECONSTOR', 'VERÖFFENTLICHEN', 'SUCHEN', 'BROWSEN', 'REFERENZEN', 'FAQ', and 'OPEN ACCESS'. The main heading is 'Suchen in EconStor'. A search bar contains the text 'Informationsverhalten' and a green 'Suchen' button. Below the search bar, there is a section for 'Erweiterte Suchvorschläge' with a dropdown arrow. Underneath, there are several search suggestions, including 'Informationsverhalten', 'Informationsverbreitung', 'Informationsversorgung', 'Informationseffizienz', 'Bildungsverhalten', and 'Mediennutzung'. A dropdown menu is open, showing a list of related terms such as 'Informationsverhalten, de', 'Information behaviour, en', 'information behaviour, en', 'informationssuche, de', 'information seeking behaviour, en', 'informationskompetenz, de', 'informationsverhalten, de', 'information-seeking behavior, en', and 'information literacy, en'. To the right of the suggestions, there is a 'Sortieren nach' section with a dropdown menu set to 'absteigend'. Below this, there is a pagination bar with the text '|< < 1 2 3 4 5 6 7 8 9 10 > >|'. The search results are displayed in a table format with columns for 'Research Report', 'Article', and 'Working Paper'. Each result includes a title, author information, year, and a link to the document. For example, the first result is a 'Research Report' titled 'Neuaufstellung des Zentraler-Orte-Konzepts in Nordrhein-Westfalen' by Greiving, Stefan, Flex, Florian, published in 2016. The second result is an 'Article' titled 'Hat die Einführung der Abgeltungsteuer Aktienkurse beeinflusst?' by Eichfelder, Sebastian, Lau, Mona, published in 2016. The third result is a 'Working Paper' titled 'Elections and market provision of information' by Bruns, Christian, published in 2013. Each result also indicates the number of hits found and a 'Schlagewörter' button.

Standard Thesaurus-Wirtschaft (STW)

Systemöffnung

Web-Publikation des STW im SKOS-Format

- STW-Zugang und Suche im Web
- Download-Angebot und maschinelle Nachnutzung (ODbL 1.0)
- Zuordnung einer eindeutigen HTTP-Adresse als persistentem Identifikator zu jedem Deskriptor und Verknüpfung mit semantischen Inhalte
- Einbettung der Inhalte in Webseiten mgl.
- Anschlussfähigkeit an einen Entitäten-basierten Verknüpfungsansatz gegeben



Aus: Kempf/Neubert. 2016.

Standard Thesaurus-Wirtschaft (STW)

Umgang mit sprachlichem Wandel in der Fachterminologie

Änderungsumfang u.a.:

- Einführung neuer Thesaurusbegriffe
- Anpassung der Vorzugsbezeichnungen an geänderten Sprachgebrauch
- Korrekturen am Zugangsvokabular
- Änderungen der Fachgebietsgliederung

Organisationsform:

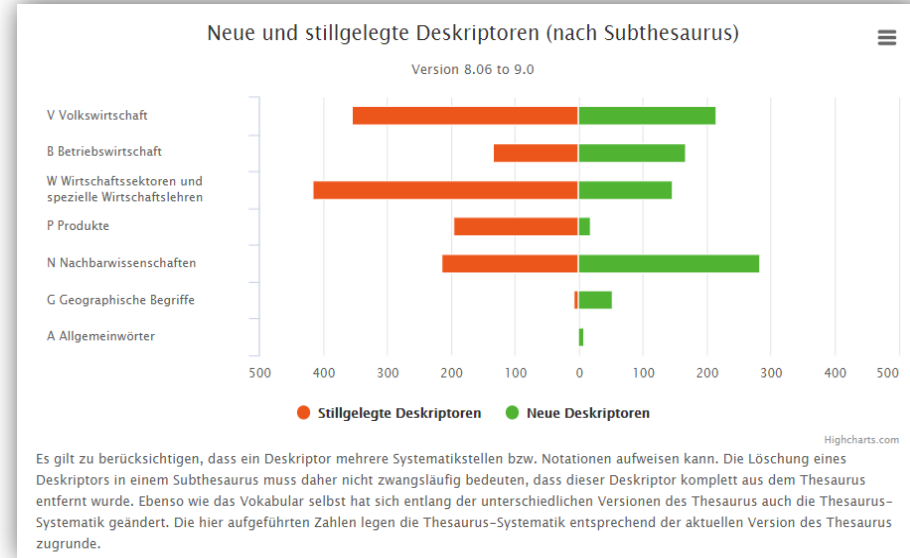
- Vorschlagsliste zu Kandidatenvokabular (DE, ND, FSW) & Vokabularänderungen
- Abzüge der Freitext-Kategorie (#5206)
- Sitzungen der STW-Redaktion

	Begriff	Status-Vorschlag Einordnung in die Systematik	kurze Beschreibung oder ggf. Quelle/Links	Vorschlag von	Rückmeldung/ Kommentar	Beschlüsse DE (Deskr.) ND (Nicht-Deskr.) FSW (Freies SW) WV (Wiedervorl.) NEIN (keine Berücksichtigung)
1	Trans-Pacific Partnership (TPP)		Freihandelsabkommen im Asiatisch-Pazifischen Raum http://en.wikipedia.org/wiki/Trans-Pacific_Strategic_Economic_Partnership	V	Lässt sich mit Freihandelsabkommen + Land indizieren.	NEIN
2	Firmenalter (firm age)	B.01.05.01 Unternehmensentwicklung		W	Beobachten. Lässt sich nicht durch andere Begriffe ausdrücken.	FSW
3	Knowledge-based View (KBV) - Wissensbasierte Unternehmenssicht		http://de.wikipedia.org/wiki/Wissensbasierte_Unternehmenssicht	X	Lässt sich z.B. mit DE Resource-based view + DE Wissen ODER DE Wissensintensives Unternehmen indizieren	NEIN
4	Product Piracy	BF Produktpiraterie		Y	BF Produktpiraterie	ND
5	Cooperation Council for the Arab States of the Gulf (CCASG) / Gulf Cooperation Council (GCC)	BF 1091347319! = Arabische Golfstaaten	http://en.wikipedia.org/wiki/Cooperation_Council_for_the_Arab_States_of_the_Gulf	Z	Hier ist die Körperschaft in 5204 zu verwenden + GEO (Arabische Golfstaaten) Körperschaft kann nicht BF bei Geo-Deskriptor werden.	NEIN

Standard Thesaurus-Wirtschaft (STW)

Umgang mit sprachlichem Wandel in der Fachterminologie

- Begriffe & Begriffsbez. sind in sozialwiss. Thesauri häufig in dynamische Diskussionsprozesse eingefasst.
- Entscheidung der STW-Redaktion zu vollständigem Relaunch (Übertragung Zuständigkeit SSG BWL)
- Wandel des STW zu flexibel nachnutzbarem Wissensrepräsentationssystem verlangt nach erhöhter Transparenz (skos-history Projekt) (vgl. Neubert 2016)



http://zbw.eu/stw/versions/latest/relaunch/charts/changed_descriptors.de.html

STW-Anwendungskontext

ZBW-Projekt Automatische Indexierung

Exkurs: Entwicklungsperspektiven für die Fachreferatsarbeit (vgl. Kempf 2018)

- Mitarbeit bei der Erstellung von Dokumentkorpora zu Trainings- & Testzwecken
- Revision der hauseigenen Sacherschließungsregeln

STW-Anwendungskontext

ZBW-Projekt Automatische Indexierung

Exkurs: Entwicklungsperspektiven für die Fachreferatsarbeit (vgl. Kempf 2018)

- Mitarbeit bei der Erstellung von Dokumentkorpora zu Trainings- & Testzwecken
- Revision der hauseigenen Sacherschließungsregeln
- Intellektuelle Mehrfachbewertung unterschiedlicher automatischer Indexierungsverfahren auf Deskriptor- und Gesamtindexatebene

STW-Anwendungskontext

ZBW-Projekt Automatische Indexierung

Exkurs: Entwicklungsperspektiven für die Fachreferatsarbeit (vgl. Kempf 2018)

Intellektuelle Mehrfachbewertung unterschiedlicher Indexierungsverfahren

- Webanwendung
- 4 Vorschlagsquellen:
 - 3x automatische Verfahren & 1x intellektuell erstellte Indexate zum Vergleich
- 8 Teilnehmer

EconBiz ID: 10011371997
Titel: A reconsideration of multiple equilibria in the analysis of one-period government bonds with default risk
Autor: Guo, Yanling
Jahr: 2015
Veröffentlicht in: Economics: The Open-Access, Open-Assessment E-Journal

Automatisch generierte Deskriptoren:

Open access
Öffentliche Anleihe
Budgetrestriktion
Staatsbankrott
Öffentliche Schulden
Theorie

Gesamtbewertung	++	+	0	-	---
Wie gut beschreibt das automatisch generierte Indexat den relevanten Inhalt der Publikation?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Weiter

(a) Gesamtbewertung

EconBiz ID: 10011371997
Titel: A reconsideration of multiple equilibria in the analysis of one-period government bonds with default risk
Autor: Guo, Yanling
Jahr: 2015
Veröffentlicht in: Economics: The Open-Access, Open-Assessment E-Journal

Descriptor	++	+	0	-	---	×
Open access	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>
Öffentliche Anleihe	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Budgetrestriktion	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Staatsbankrott	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Öffentliche Schulden	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Theorie	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Finanzierungstheorie	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Schuldenkrise	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
+						

Freie Kommentare und weitere Schlagwörter

Gesamtbewertung abschicken

(b) Einzelbewertung

STW-Anwendungskontext

ZBW-Projekt Automatische Indexierung

Exkurs: Entwicklungsperspektiven für die Fachreferatsarbeit (vgl. Kempf 2018)

- Mitarbeit bei der Erstellung von Dokumentkorpora zu Trainings- & Testzwecken
- Revision der hauseigenen Sacherschließungsregeln
- Intellektuelle Mehrfachbewertung unterschiedlicher automatischer Indexierungsverfahren auf Deskriptor- und Gesamtindexatebene
- Mitarbeit bei der Erstellung der Bewertungskriterien und -heuristik
- Intellektuelle Bewertung von Stichproben im Vorfeld der regelmäßigen Releases bzw. Einspielungen

STW-Anwendungskontext

ZBW-Projekt Automatische Indexierung

Exkurs: Entwicklungsperspektiven für die Fachreferatsarbeit (vgl. Kempf 2018)

1: Intellektuelle Erfassung des Dokumentinhalts

2: Bewertung einzelner Schlagwörter

3: Bewertung des Gesamtindexats

The screenshot shows the ZBW-A1-Release-Tool interface. At the top, it displays 'Econbiz-ID: 10002772736' and the title 'Differences in desired and actual fertility : an economic analysis of the Spanish case'. Below the title, there are keywords: 'fertility', 'religion', 'family planning', 'desired number of children', and 'unemployment'. An abstract follows, discussing family size and fertility in Spain. Below the abstract is a table of descriptors with columns for 'Descriptor', 'main_ai', 'lazy', 'maui', and 'binary_relevance'. The table contains several rows with descriptors like 'Family planning', 'Fertility', 'Children', 'Religion', 'Index construction', 'Fertilizer', 'China', and 'Spain', each with a rating icon. At the bottom, there is an 'Add' button and a 'Total rating' section with 'Reject', 'Accept, Fair', and 'Accept, Good' options.

Descriptor	main_ai	lazy	maui	binary_relevance
15939-4 Family planning	✓	✓	✓	
15941-3 Fertility	✓	✓	✓	
15925-1 Children	✓	✓	✓	
15671-6 Religion	✓	✓	✓	
15097-5 Index construction	✓	✓	✓	
13427-6 Fertilizer	✓	✓	✓	
17565-2 China	✓	✓	✓	
17161-5 Spain	✓	✓	✓	

Unterschiedliche Verfahren:
wörterbuchbasiert
statistisch, assoziativ
nächste Nachbarn

Automatisch erstellter **Hauptvorschlag**
unter Einbezug der Subthesauri

STW-Anwendungskontext

ZBW-Projekt Automatische Indexierung

Exkurs: Entwicklungsperspektiven für die Fachreferatsarbeit (vgl. Kempf 2018)

Ausblick:

- *Qualitätsmanagement: Themenspezifische Qualitätsbestimmung auf Zeitschriften- und Serienebene (Indexierungsprofil)*
- *Test-Vorbereitung zur Verwendung des Assistenzsystems DA 3*

STW-Anwendungskontext

ZBW-Projekt Automatische Indexierung

Exkurs: Entwicklungsperspektiven für die Fachreferatsarbeit (vgl. Kempf 2018)

Erstes Fazit:

- Rekurs auf die Sacherschließungsexpertise der Fachreferentinnen und Fachreferenten
- Rekurs auf die Bestandsmanagementexpertise der Fachreferentinnen und Fachreferenten
- Einbindung & Verwendung von Kompetenzen aus der klassischen Fachreferatsarbeit

Ansätze zur terminologischen Weiterentwicklung

Fokus: Autoren-Keywords

↓ Untersuchung der Überschneidungsmenge zw. Autoren-Keywords und STW zur Auslotung des Potentials für 1:1-Nachnormierung

An exploration study to detect important factors influencing on compulsive buy : a case study of Shahrvand chain stores				
Ali Alikhani, Jamshid Salehi Sadaghiani, Maryam Jabbarzadeh and Mehdi Asgarielo				
Erscheinungsjahr:	2013			
Weitere Verfasser/innen:	Alikhani, Ali; Sadaghiani, Jamshid Salehi; Jabbarzadeh, Maryam; Asgarielo, Mehdi			
Veröffentlicht in:	Management science letters - Growing Science, ISSN 1923-9343, ZDB-ID 25920091. - Vol. 3.2013, 7, p. 2185-2188			
Beschreibung:	graph. Darst.			
Sprache:	Englisch			
Schlagwörter:	Consumer behavior Food industry Emotional Instability Attractive Appearance			
Publikationsform (Subkategorien):	Aufsatz in Zeitschriften	Article in Journal	Fallstudie	Case study
Publikationsform:	Aufsatz	Aufsatz	Aufsatz	Aufsatz
	STW ID 10300-3	STW ID 13160-6	STW-ID	STW-ID

Aus: Wortmann et al. 2014

Ansätze zur terminologischen Weiterentwicklung

Fokus: Autoren-Keywords

- ↓ Untersuchung der Überschneidungsmenge zw. Autoren-Keywords und STW zur Auslotung des Potentials für 1:1-Nachnormierung
- ↓ Verwendung von Autoren-Keywords neben Dokumenttiteln als Textgrundlage für automatische Kurztext-Verfahren

The screenshot shows the ZBW-A1-Release-Tool prototype interface. The title "Differences in desired and actual fertility : an economic analysis of the Spanish case" and the keywords "fertility", "religion", "family planning", "desired number of children", and "unemployment" are circled in red. Below the abstract is a table of descriptors.

Descriptor	main_ai	lazy	maui	binary_relevance
15939-4 Family planning	✓			
15941-3 Fertility	✓			
15925-1 Children	✓			
15671-6 Religion	✓			
15097-5 Index construction	✓			
13427-6 Fertilizer	✓			
17565-2 China			✓	
17161-5 Spain		✓		

<https://github.com/zbw/releasetool>

Ansätze zur terminologischen Weiterentwicklung

Fokus: Autoren-Keywords

- ↓ Untersuchung der Überschneidungsmenge zw. Autoren-Keywords und STW zur Auslotung des Potentials für 1:1-Nachnormierung
- ↓ Verwendung von Autoren-Keywords neben Dokumenttiteln als Textgrundlage für automatische Kurztext-Verfahren
- Keine Generierung neuen Vorschlagsvokabulars

The screenshot shows the ZBW-A1-Release-Tool interface. The document title is "Differences in desired and actual fertility : an economic analysis of the Spanish case". The keywords are "fertility", "religion", "family planning", "desired number of children", and "unemployment". The abstract discusses the gap between desired and actual fertility in Spain, influenced by preferences and economic conditions. Below the abstract is a table of descriptors.

Descriptor	main_ai	lazy	maui	binary_relevance
15939-4 Family planning	✓			
15941-3 Fertility	✓			
15925-1 Children	✓			
15671-6 Religion	✓			
15097-5 Index construction	✓			
13427-6 Fertilizer	✓			
17565-2 China				
17161-5 Spain				

<https://github.com/zbw/releasetool>

Ansätze zur terminologischen Weiterentwicklung

Prozessunterstützte Generierung von Vorschlagsvokabular

Prozessunterstützte Selektion von Termvorschlägen aus unterschiedlichen Textquellen und Evaluation durch die STW-Redaktion

Quellbereiche:

- Autoren-Keywords
- Titel von Publikationen
- Abstracts von Publikationen
- EconBiz Such-Logfiles

Prozessschritte:

- Extrahieren > Gewichten > Filtern > Anreichern
- Evaluieren

Evaluierung verschiedener Quellen
zur Anreicherung kontrollierter
Vokabulare durch Generierung von
Vorschlägen

Thesis zur Erlangung eines Abschlusses als
Master of Science

vorgelegt von
Alexander Prange

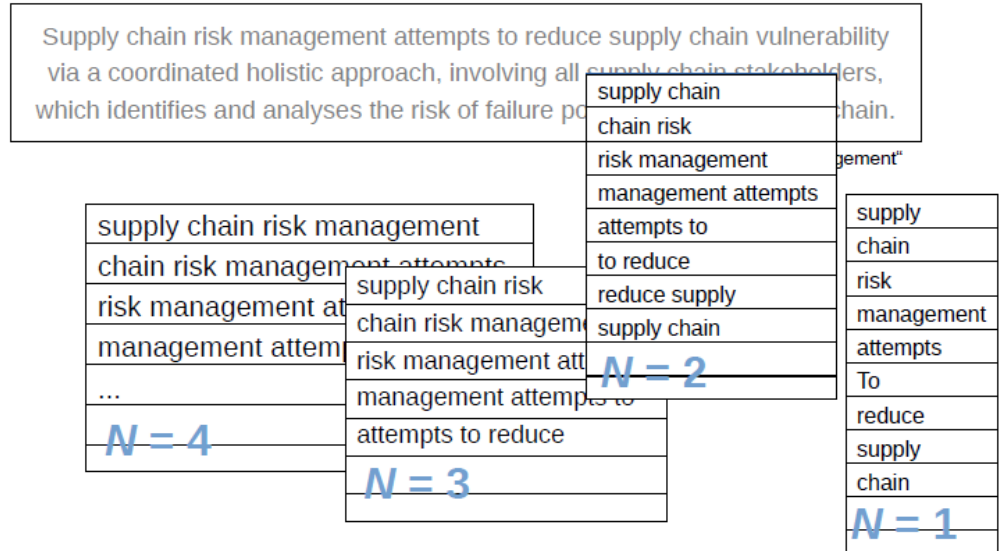
Christian-Albrechts-Universität zu Kiel
Leibniz-Informationszentrum Wirtschaft
Arbeitsgruppe Knowledge Discovery

Erstgutachter: Prof. Dr. Ansgar Scherp
Betreuer: Martin Toepfer

Ansätze zur terminologischen Weiterentwicklung

Prozessunterstützte Generierung von Vorschlagsvokabular: Extrahieren

- Generierung von n-Grammen




Vgl. Prange 2016

Ansätze zur terminologischen Weiterentwicklung

Prozessunterstützte Generierung von Vorschlagsvokabular: Gewichten

- Generierung von n-Grammen
- Subtraktion der durchschnittlichen Häufigkeit der Oberterme von der Häufigkeit des Terms



Term	Anzahl	X
supply chain risk management	15	15
supply chain risk	15	0
supply chain	25	10

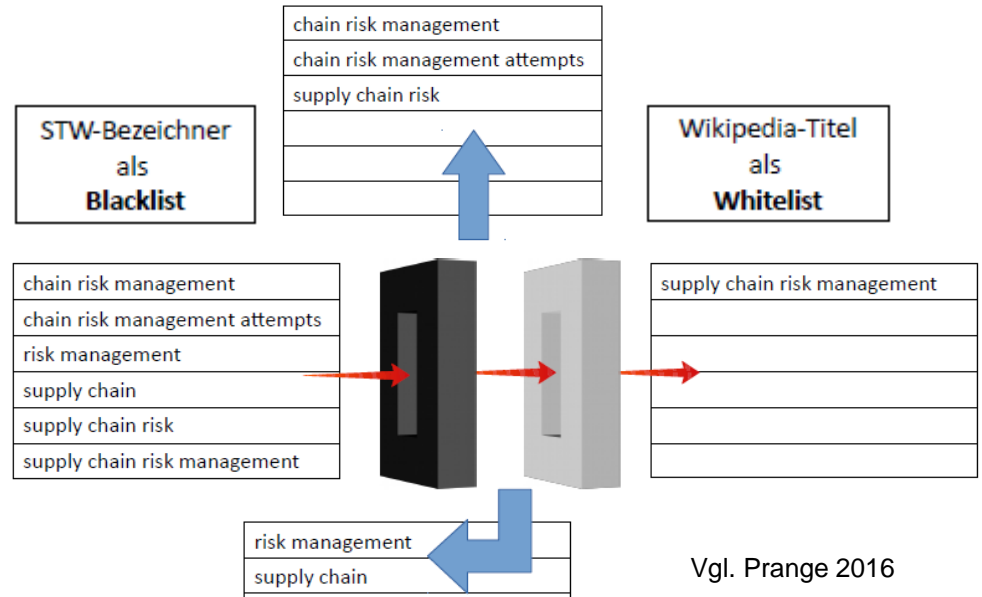
$X = 25 - \text{Mittelwert}(15, 15)$

Vgl. Prange 2016

Ansätze zur terminologischen Weiterentwicklung

Prozessunterstützte Generierung von Vorschlagsvokabular: Filtern

- Generierung von n-Grammen
- Subtraktion der durchschnittlichen Häufigkeit der Oberterme von der Häufigkeit des Terms
- Filterung (STW- & Wikipedia-Filter)



Ansätze zur terminologischen Weiterentwicklung

Prozessunterstützte Generierung von Vorschlagsvokabular: Anreichern

- Generierung von n-Grammen
- Subtraktion der durchschnittlichen Häufigkeit der Oberterme von der Häufigkeit des Terms
- Filterung (STW- & Wikipedia-Filter)
- Anreicherung durch *Matches*, *Directs*, *Relation* & *One-Hops*

Agrovoc

Matches

[information systems](#)

Direct Links

Origin	Link	STW
information systems	skos:exactMatch	Information system

One-Hop Links

Origin	Link 1	Resource	Link 2	STW
information systems	agrontology:hasObjectOfActivity	computers	skos:exactMatch	Computer
information systems	agrontology:influences	information services	skos:exactMatch	Information service provider
information systems	skos:narrower	databases	skos:exactMatch	Database
information systems	skos:narrower	expert systems	skos:exactMatch	Expert system
information systems	skos:narrower	management information systems	skos:exactMatch	Management information system

Vgl. Prange 2016

Ansätze zur terminologischen Weiterentwicklung

Prozessunterstützte Generierung von Vorschlagsvokabular: Evaluieren

Evaluation der Vorschläge durch die STW-Redaktion

Korpus	Terme	Deskriptor	Synonym	Anteil
Abstracts	61	1	3	6,56%
Titles	61	2	7	14,75%
Keywords	61	1	15	26,23%
Queries	61	1	7	13,11%
Gesamt	201	4	28	15,92%

Aus: Prange 2016: 58

Ergebnis:

Autoren-Keywords, Suchanfrage und Titeldaten eignen sich relativ gut für eine Terminologie-Extraktion mit dem Ziel der Vokabular-Anreicherung.

(Zwischen-)Fazit

- Auswirkungen automatischer Sacherschließungsverfahren auf das Thesaurusmanagement sollten in einem größeren Kontext unterschiedlicher Veränderungszusammenhänge (z.B. Systemöffnung) betrachtet werden.
 - Der verstärkte Einsatz automatischer Sacherschließung lässt es erforderlich erscheinen, geeignete neue Quellbereiche (prozessunterstützt) einzubeziehen.
 - Die Diversifikation der Sacherschließungsebenen sollte mit einer Diversifikation des Vorschlagswesens für neue Term- und Begriffsvorschläge korrespondieren.
 - Evaluation, Diskussion & Entscheidung über neues Thesaurusvokabular sollten weiterhin redaktionell organisiert sein.
-

Quellen

- Gastmeyer, Manuela. Der Einsatz des Standard-Thesaurus Wirtschaft im HWWA. Ein Instrument zur Qualitätssicherung von wirtschaftswissenschaftlichen Fachinformationen. In: Auskunft. Mitteilungsblatt Hamburger Bibliotheken, Vol. 20, No. 2, pp. 108-130.
- Dies.; Wannags, Max-Michael; Neubert, Joachim. Relaunch des Standard-Thesaurus Wirtschaft – Dynamik in der Wissensrepräsentation. In: Information – Wissenschaft & Praxis. Berlin: De Gruyter, Vol. 67 (2016), No. 4, pp. 217-240, doi:10.1515/iwp-2016-0039
- Kempf, Andreas Oskar. 2018. Qualität erfordert Fachexpertise. Entwicklungsperspektiven für die Fachreferatsarbeit vor dem Hintergrund des Qualitätsmanagements automatisierter Sacherschließungsprozesse. Hands-On Lab Bewertung maschineller Indexierung – Qualität ist kein Zufall, 107. Deutscher Bibliothekartag, 12.-15. Juni 2018, Berlin https://opus4.kobv.de/opus4-bib-info/files/15801/20180620_bibtag_kem_final_upload.pdf
- Neubert, Joachim. Skos-history: Exploring Web Standards for Change Tracking in Knowledge Organization Systems. In: Proceedings of the 25th International Conference Companion on World Wide Web, New York City: ACM, 2016, pp. 275-276, <http://hdl.handle.net/111108/248>
- Prange, Alexander. Evaluierung verschiedener Quellen zur Anreicherung kontrollierter Vokabulare durch Generierung von Vorschlägen. Thesis zur Erlangung eines Abschlusses als Master of Science. 2016, Christian-Albrechts-Universität zu Kiel, ZBW – Leibniz-Informationszentrum Wirtschaft.
- Toepfer, Martin. Machine Learning Architectures for Scalable and Reliable Subject Indexing. Fusion, Knowledge Transfer, and Confidence. In: Research and Advanced Technology for Digital Libraries, 21st International Conference on Theory and Practice of Digital Libraries, TPDL 2017, Thessaloniki, Greece, September 18-21, 2017, Proceedings, Lecture Notes in Computer Science, Cham: Springer, 2017, pp. 644-647, doi:10.1007/978-3-319-67008-9_61
-

Quellen

Toepfer, Martin; Kempf, Andreas Oskar. Automatische Indexierung auf Basis von Titeln und Autoren-Keywords – ein Werkstattbericht. In: 027.7 Zeitschrift für Bibliothekskultur. Basel: Universität Basel, Vol. 4 (2016), No. 2, pp. 84-97, doi:10.12685/027.7-4-2-156

Wortmann, Karin; Gross, Thomas; Bahls, Daniel. Sacherschließung in der ZBW. Anwendung automatischer Verfahren – Werkstattbericht, 103. Deutscher Bibliothekartag in Bremen, 2014, https://opus4.kobv.de/opus4-bibinfo/files/1571/2014_06_03_Beitrag2_automSacherschliessung_ZBW.pdf

ZBW – Deutsche Zentralbibliothek für Wirtschaftswissenschaften Leibniz-Informationszentrum Wirtschaft. Strategie der ZBW 2015-2020, 2015, Kiel, Hamburg. <https://www.zbw.eu/fileadmin/pdf/ueber-uns/2015-strategie.pdf>

ZBW – Leibniz-Informationszentrum Wirtschaft. Einfach. Machen. Jahresbericht 2017, 2018, Kiel, Hamburg. <https://www.zbw.eu/fileadmin/pdf/ueber-uns/jb-2017.pdf>