

Sacherschließung in der ZBW Anwendung automatischer Verfahren

Werkstattbericht

Karin Wortmann / Thomas Gross / Daniel Bahls

ZBW – Leibniz-Informationszentrum Wirtschaft

03.Juni 2014

Überblick

1. Strategische Ausrichtung der Sacherschließung in der ZBW
2. Standard-Thesaurus Wirtschaft (STW)
 - Anwendungsbereiche des STW im Rahmen von Crosskonkordanzen
3. Automatische Sacherschließung
 - Rückblick
 - Ausblick

Sacherschließungskonzept

Aufgaben und Ziele der ZBW

- § 2 (1) Stiftungszweck: „Die Stiftung sammelt und erschließt weltweit erscheinende wirtschaftswissenschaftliche Literatur.[...]“ (Satzung der Stiftung „Deutsche Zentralbibliothek für Wirtschaftswissenschaften - Leibniz-Informationszentrum Wirtschaft“)
 - Als überregionale Informationsinfrastruktureinrichtung verpflichtet sich die ZBW der zuverlässigen, umfassenden und nachhaltigen Versorgung der wirtschaftswissenschaftlichen Forschungsgemeinschaft mit wirtschaftswissenschaftlichen Informationen. Die Grundlage hierfür bilden der einmalige Bestand, die Bereitstellung von einfachen und neuartigen Zugangswegen zu diesem Bestand sowie die exzellente Kunden- und Serviceorientierung der ZBW. (Strategie ZBW)
 - Die ZBW bietet einfache, vielfältige und auf exzellenten Forschungsergebnissen basierende Zugänge zu ihrem Bestand. Diese Zugangsmöglichkeiten, die auf den hochwertig aufbereiteten Metadaten basieren, sind derart gestaltet, dass der Bestand lokal, national und international genutzt werden kann. (Strategie ZBW)
-

Sacherschließungskonzept

Änderung der Rahmenbedingungen für die Sacherschließung

- Zunehmende Digitalisierung und Vernetzung von Informationssystemen
- Nachnutzbare Metadaten in Dokumenten und Katalogen
- Existierende Recherchertools und Suchmaschinentechnologien, die zu qualitativ befriedigenden Rechercheergebnissen führen
- Starker Anstieg der zu erschließenden wirtschaftswissenschaftlichen Publikationen
- Begrenzte personelle Ressourcen
- Kompetenzen im Haus durch Aufbau eines neuen Programmbereichs Medieninformatik

Sacherschließungskonzept

Unbefriedigender Zustand durch veränderte Rahmenbedingungen

- Immer geringerer Sacherschließungsgrad der neuesten wirtschaftswissenschaftlichen Literatur
- Negative Auswirkungen auf die Dienste der ZBW (EconBiz, EconDesk)
- STW zunehmend in der eigenen Datenbasis nicht präsent

Neues Sacherschließungskonzept

Intellektuelle Erschließung unter stärkerer Einbeziehung von automatischen Verfahren und Berücksichtigung von Crosskonkordanzen und vorhandener Fremderschließung

- Intellektuelle Sacherschließung mit dem STW
- Crosskonkordanzen
- Fremddatenübernahme, z.B. Autoren-Keywords
- (Teil-)Automatisierte Verfahren der Sacherschließung

Standard-Thesaurus Wirtschaft

- wurde in den 1990er Jahren unter öffentlicher Förderung von vier wirtschaftswissenschaftlichen Institutionen entwickelt
- wird heute von der ZBW gepflegt und weiterentwickelt
- umfasst über 6.000 Deskriptoren in Deutsch und Englisch
- bildet ein feinmaschiges Begriffsnetz
 - mehr als 16.000 Ober-/Unterbegriffsbeziehungen
 - fast 11.000 verwandte Begriffe
 - zusätzlicher Zugang über Thesaurussystematik mit über 500 Stellen

Suche mit Thesaurusunterstützung

Ökobilanz → Umweltbilanz

[Liste der Deskriptoren](#)
[Download](#)
[Andere Versionen](#)
[Wiki](#)


- ▶ [A Allgemeinwörter](#)
- ▶ [B Betriebswirtschaft](#)
- ▶ [G Geographische Begriffe](#)
- ▶ [N Nachbarwissenschaften](#)
- ▶ [P Produkte](#)
- ▶ [V Volkswirtschaft](#)
- ▶ [W Wirtschaftssectoren und spezielle Wirtschaftslehren](#)

Standard-Thesaurus Wirtschaft

Neue Version 8.06 (Änderungen)

Subthesauri

- [A Allgemeinwörter](#)
- [B Betriebswirtschaft](#)
- [G Geographische Begriffe](#)
- [N Nachbarwissenschaften](#)
- [P Produkte](#)
- [V Volkswirtschaft](#)
- [W Wirtschaftssectoren und spezielle Wirtschaftslehren](#)

Der *Standard-Thesaurus Wirtschaft* soll mit seinen genormten Schlagwörtern die thematische Suche im ZBW-Katalog [ECONIS](#) erleichtern (einfach per  Icon).

STW Begriffsnetz

[Home](#)

[Alphabetische Liste der Deskriptoren](#)
[Wiki](#)

- ▶ [A Allgemeinwörter](#)
- ▶ [B Betriebswirtschaft](#)
- ▶ [G Geographische Begriffe](#)
- ▶ [N Nachbarwissenschaften](#)
- ▶ [P Produkte](#)
- ▶ [V Volkswirtschaft](#)
- ▶ [W Wirtschaftssectoren und spezielle Wirtschaftslehren](#)

Umweltbilanz

Environmental accounting (engl.)

benutzt für: Ökobilanz, Umweltbezogene Rechnungslegung

Oberbegriffe

- [Sozialbilanz](#) 
- [Umweltmanagement](#) 

Verwandte Begriffe

- [Corporate social responsibility](#) 
- [Öko-Audit](#) 
- [Ökosozialprodukt](#) 

Thesaurus Systematik

- [B.01.06 Umweltmanagement](#) ▼
- [B.03 Betriebswirtschaftliches Rechnungs- und Prüfungswesen](#) ▼

Persistenter Identifier (für Bookmarks und zum Verlinken)

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

STW-Web Services und Crosskonkordanzen

Umweltbericht

Environmental reporting (engl.)

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

Oberbegriffe

- [Nachhaltigkeitsbericht](#)  


Verwandte Begriffe

- [Öko-Audit](#)  
- [Produktlebenszyklus](#)  
- [Stoffstrommanagement](#)  
- [Umwelt-Informationssystem](#)  
- [Umweltkosten](#)  
- [Umweltökonomische Gesamtrechnung](#)  

Thesaurus Systematik


- [B.01.03 Betriebliche Information und Kommunikation](#) ▼
- [B.01.06 Umweltmanagement](#) ▼
- [B.03 Betriebswirtschaftliches Rechnungs- und Prüfungswesen](#) ▼

Links zu anderen Thesauri und Vokabularen

- = [Umweltbilanz](#) (aus SWD)
- ≅ [Umweltökonomie](#) (aus DBpedia) 
- = [Ökobilanz](#) (aus Thesoz)
- = [Environmental accounting](#) (aus Agrovoc)

Integrierte Suche in EconBiz unter Verwendung der Crosskordanz STW-SWD

Search (all) **Internet Sources** **Events**

umweltbilanz  **SEARCH**
[Advanced Search](#)

search pdf full-texts (not available for the entire collection)
 search also for related terms

[Your search term has been extended to synonyms and translations \(display/hide\)](#)
Extended query string with operators and brackets:

umweltbilanz OR ("life cycle assessment" OR "ökobilanz" OR "environmental reporting" OR "umweltbericht" OR "environmental disclosure" OR "ökobilanzierung" OR "corporate environmental accounting" OR "umweltberichterstattung" OR "environmental management accounting" OR "umweltbilanz" OR "life-cycle-assessment" OR "environmental accounting")

Search terms of extended query for direct search:
[umweltbilanz](#) ["life cycle assessment"](#) ["ökobilanz"](#) ["environmental reporting"](#) ["umweltbericht"](#) ["environmental disclosure"](#) ["ökobilanzierung"](#) ["corporate environmental accounting"](#) ["umweltberichterstattung"](#) ["environmental management accounting"](#) ["umweltbilanz"](#) ["life-cycle-assessment"](#) ["environmental accounting"](#)

Intellektuelle Sacherschließung

Klassische intellektuelle Indexierung mit dem STW

Voraussetzung:

- Keine Fremderschließung/Autoren-Keywords vorhanden
- Keine Crosskonkordanzen vorhanden
- Automatische Indexierung führt zu wenig brauchbaren Ergebnissen

Automatische Sacherschließung

Ziel:

- Inhalterschließung elektronischer Dokumente
- Indexierung mit genormtes Vokabular des Standard-Thesaurus Wirtschaft (STW)
- Ermöglichung von homogener Recherche im Portal EconBiz

Rückblick – Reccomind-Projekt

- Laufzeit: 2009 bis 2011
- Datensatz: 40.000 Titel mit Volltexten (Working Paper, Dissertationen...)
 - Trainingsset = 26.000 Titel vs. Testset = 14.000 Titel
- begriffsorientiertes Verfahren
 - statistischer Ansatz („Decisiv Categorization“) + STW
- Auswertung: Indexierungskonsistenz (ReferentIn vs. Maschine) → 36%
- STW-Belegung Automatik: nur 17% des STW werden benutzt; oft nur Oberbegriffe / Allgemeinwörter → Mensch: 71%
- deutsch (-) vs. englisch (+), BWL (-) vs. VWL (+), formellastig (-) vs. textlastig (+), Themencluster

Rückblick – Recommind-Projekt

→ „lessons learned“

- Akzeptanz des Systems durch Kollegium als Grundlage
- umfassende Projektverortung und –beteiligung im Haus (Katalogisierung...)
- Aufbau von internem IT-Know-how (Abhängigkeiten verringern)
- Polyhierarchie des STW (Standard-Thesaurus-Wirtschaft)
- wichtig: ausreichende Trainingsmenge an Dokumenten (Fokus auf Dokumentenqualität, Urheberrecht)

Neues Sacherschließungskonzept

Einsatz automatischer Verfahren

- Erstellung eines Konzepts zur automatischen Indexierung in der ZBW
 - Formulierung eines klaren Projektauftrags
 - Erstellung eines umfassenden „Anforderungsprofils“
 - Lastenheft
 - Evaluierung bestehender Systeme
- Ziel: im ersten Schritt bis Ende 2015 (halb-)automatische Nachnormierung von Autoren-Keywords mit STW-Schlagworten im Produktivbetrieb

Wo stehen wir?

- Frühjahr 2013: Identifizierung der in Frage kommenden Zeitschriften mit Autoren-Keywords
- Seit September 2013 Aufnahme der Autoren-Keywords aus wissenschaftlichen Fachzeitschriften durch die Abteilung Integrierte Erwerbung & Katalogisierung
- Keine intellektuelle Erschließung der einzelnen Aufsätze mehr mit dem STW durch die wissenschaftlichen Referentinnen und Referenten

Sacherschließungskonzept

Autoren-Keywords

Economic schools of thought on the environment: investigating unity and division

*Clive L. Spash and Anthony Ryan**

How do ecological and heterodox economists differ, if at all, from each other and from neoclassical economists addressing environmental problems? In 2009 we probed this question by conducting an international survey across these communities, namely at conferences of the European Society for Ecological Economics, the Association for Heterodox Economics, and the European Association of Environmental and Resource Economists. The research was designed to gain insight into the extent to which ecological economics can be described as heterodox and a distinct field from orthodox environmental and resource economics. Conflicting visions of ecological economics have led to a prevalence of neoclassical articles and thought mixed in amongst more heterodox work. We introduce a novel classification of work in the field of environmental policy in order to test for the existence of differences in terms of methodological and ideological approaches. How heterodox economists understand environmental issues is also an important question to answer if there is to be more collaboration between them and ecological economists. The findings have implications for cooperation and the future direction of both ecological and heterodox economics.

Keywords: Ecological economics, Heterodox, Neoclassical, Methodology, Ideology
JEL classifications: R40, R59, Q00

Screenshot aus ECONIS



[Elektronische Ressource]

Aufsatz: [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](#)

Sonst. Personen: [Alikhani, Ali](#) ; [Sadaghiani, Jamshid Salehi](#) ; [Jabbarzadeh, Maryam](#) ; [Asgarielo, Mehdi](#)

Umfang: graph. Darst.

Anmerkung: Systemvoraussetzung: **Acrobat** Reader

In: [Management science letters, ISSN 1923-9343, ZDB-ID 25920091. - Waterloo, Ont : Growing Science. - Bd. 3.2013, 7, S. 2185-2188](#)

Art/Inhalt: Fallstudie / Case study / Online-Publikation / Online publication

Autoren-Keywords: Consumer behavior / Food industry / Emotional Instability / **Attractive Appearance**

Link: <http://dx.doi.org/10.5267/j.msl.2013.05.046>
http://www.growing-science.com/msl/Vol3/msl_2013_166.pdf

DOI: 10.5267/j.msl.2013.05.046

Nachnormierung von Keywords

Erste Auswertungen

- **2.835** Dokumente unter der Lupe
davon (leider nur)
451 auch manuell verschlagwortet
 - **35.146** Autoren-Keywords insgesamt
davon **20.134** *unique* – nach Entfernen der Dubletten
 - Versuch einer automatischen Indexierung anhand *einfacher*
Nachnormierung der Keywords
 - Betrachtung der Zurodnung von Keywords zu STW-Begriffen
-

Nachnormierung von Keywords

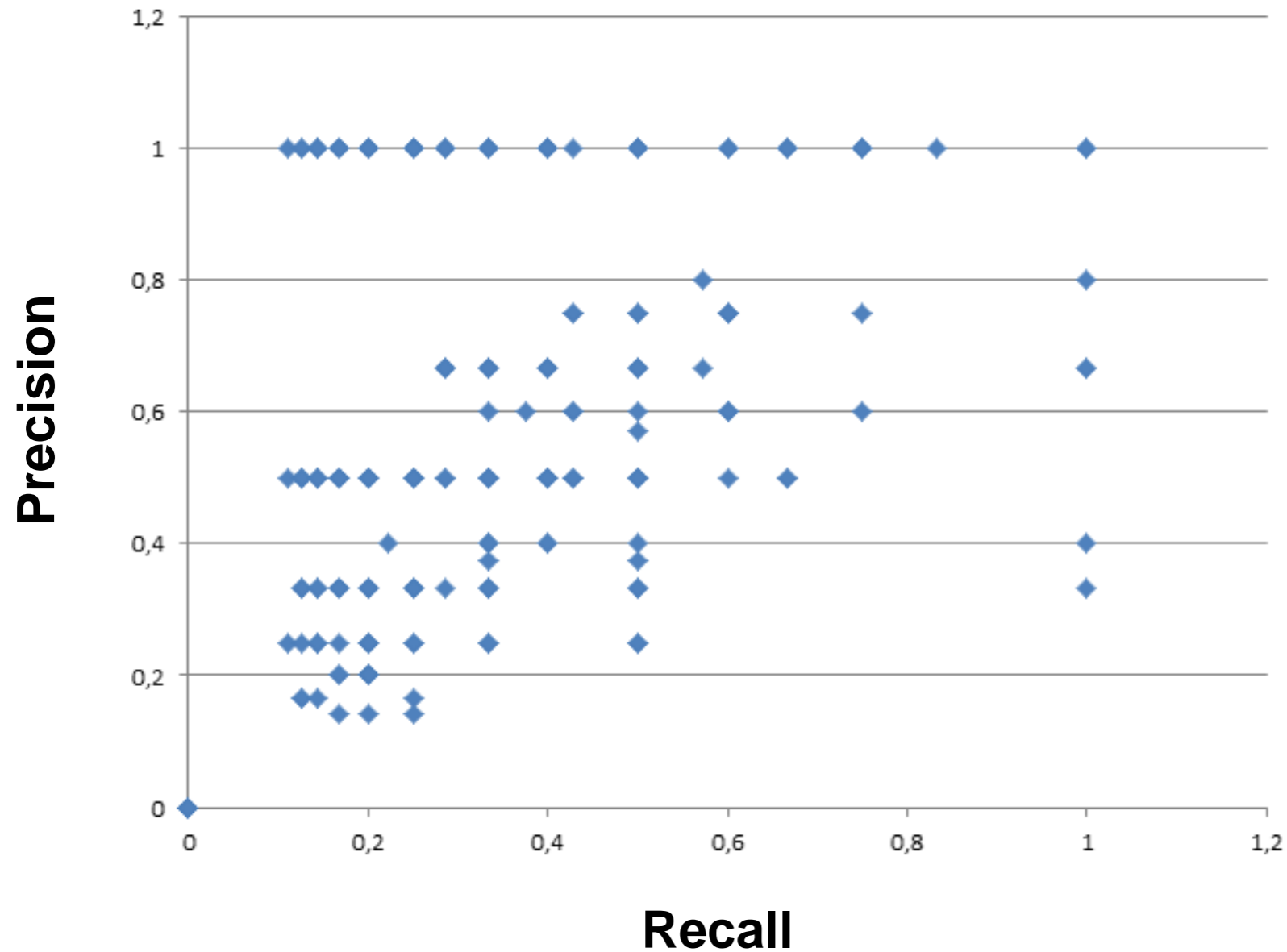
Erste Auswertungen – Top 20

Keyword	Freq
innovation	111
china	96
entrepreneurship	89
corporate governance	81
leadership	74
economic growth	70
trust	58
india	57
philippines	54
supply chain management	54
performance	54

Keyword	Freq
education	53
unemployment	52
gender	52
germany	51
monetary policy	46
productivity	44
financial crisis	44
corporate social responsibility	43
service quality	43
growth	41

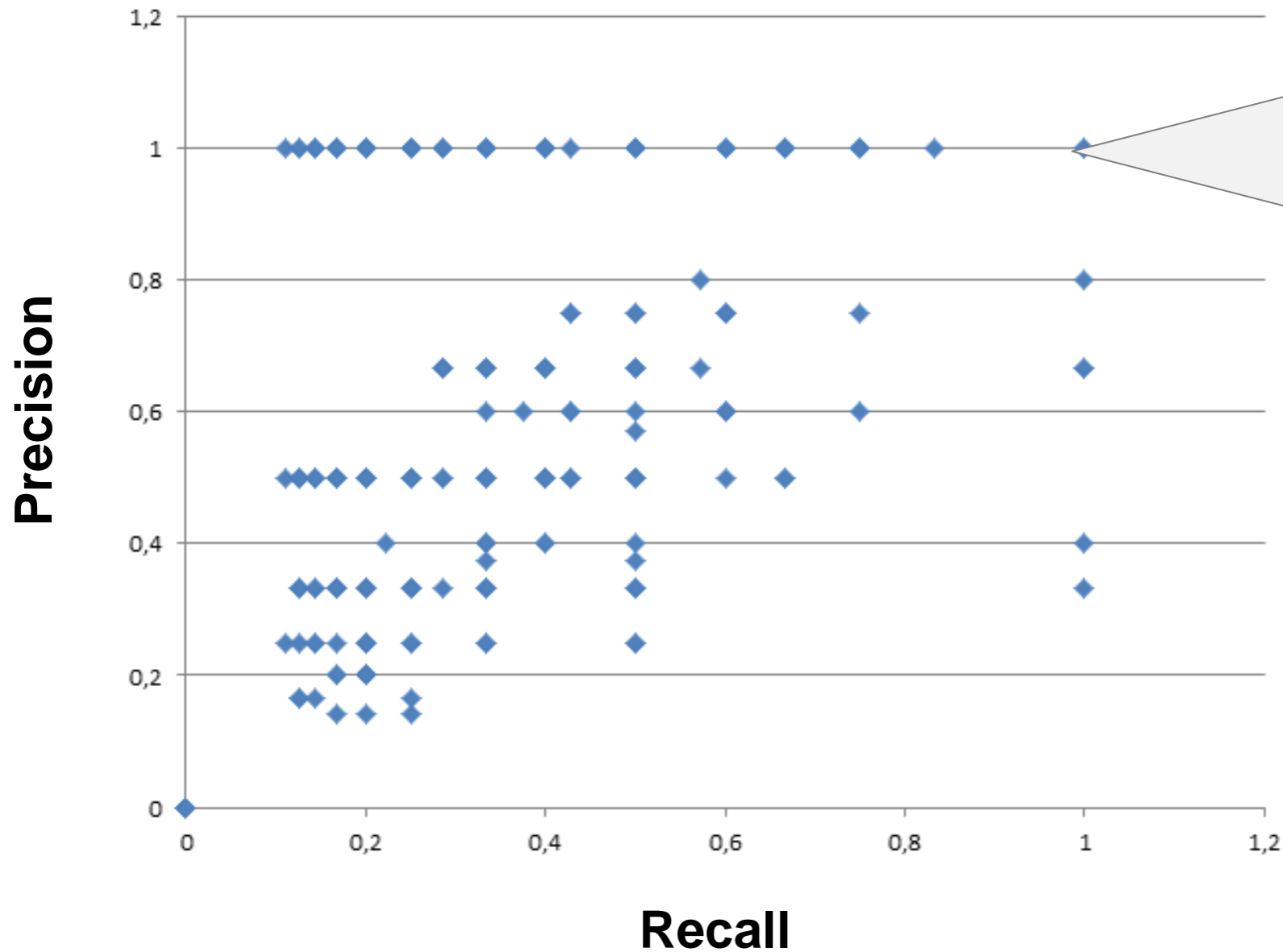
Nachnormierung von Keywords

Erste Auswertungen – Precision / Recall



Nachnormierung von Keywords

Erste Auswertungen – Precision / Recall



Anchoring of consumers' inflation expectations

Manuell sowie automatisch

microdata

Inflationserwartung

Keywords:

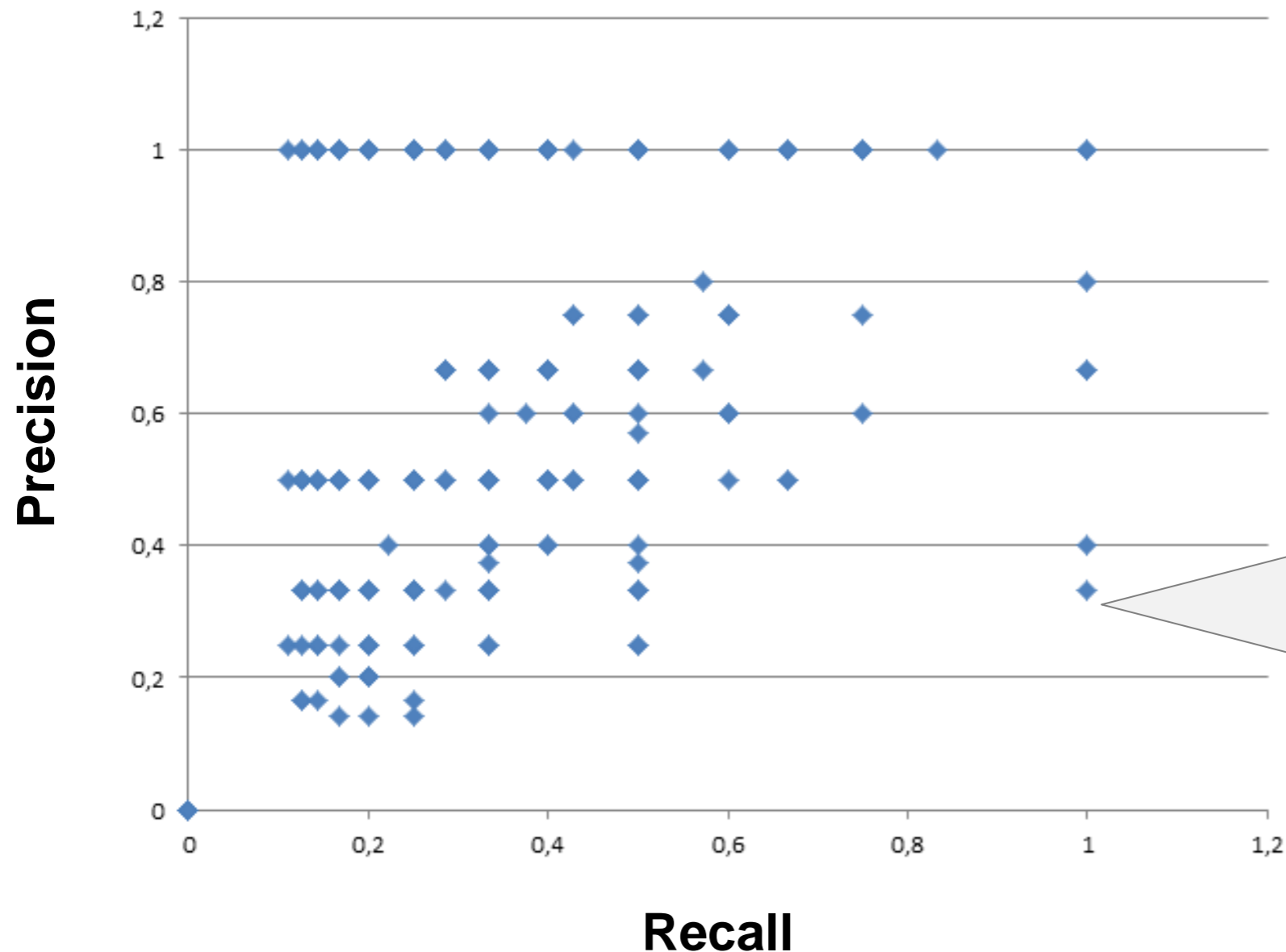
Anchoring;

inflation expectations;

microdata; news

Nachnormierung von Keywords

Erste Auswertungen – Precision / Recall



Beyond the management view

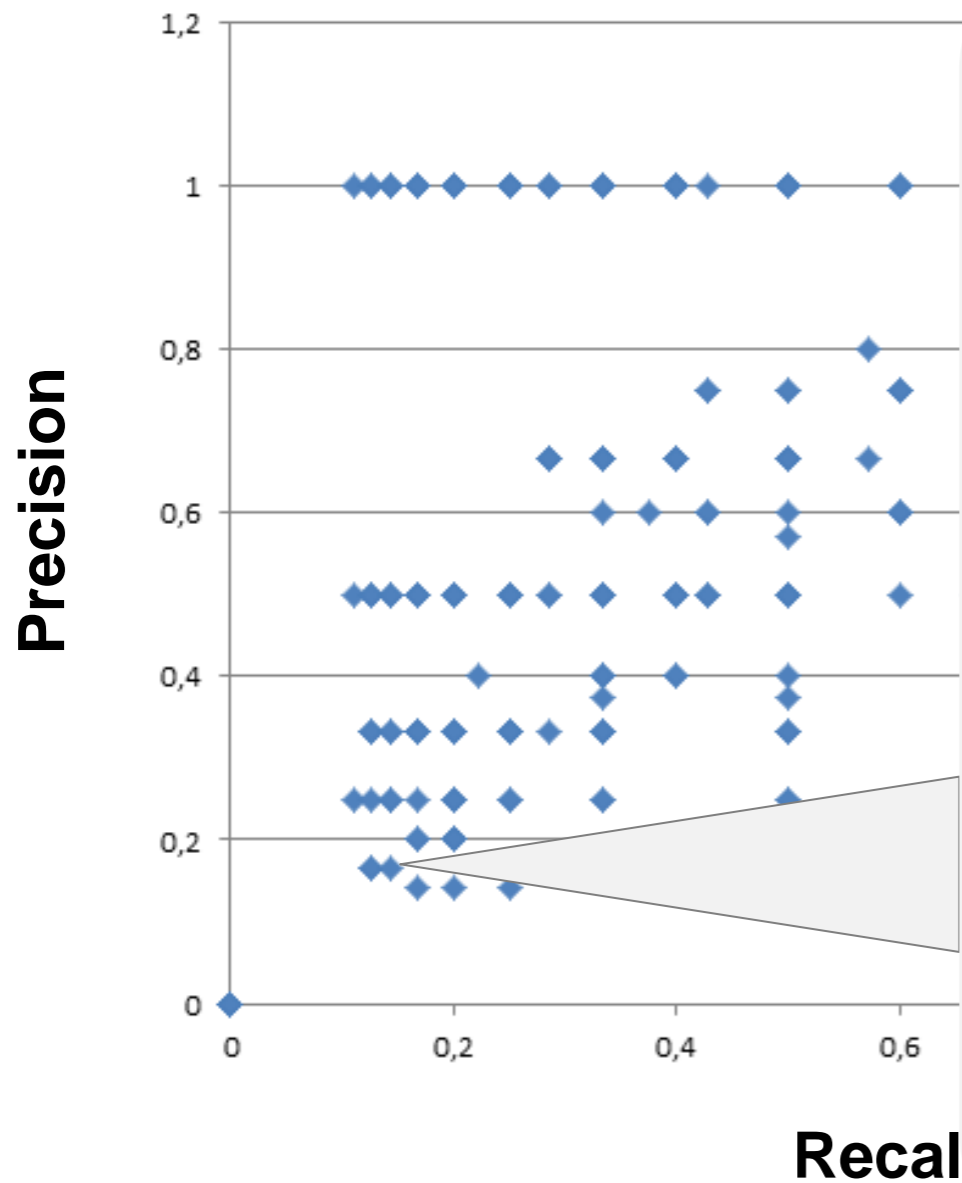
Manuell sowie automatisch:
**Ressourcenorientierter
Ansatz**

Automatisch zudem:
**Wettbewerbsvorteil
Transaktionskosten**

Keywords:
**the resource-based view;
transaction cost economics;
sustained competitive
advantage**

Nachnormierung von Keywords

Erste Auswertungen – Precision / Recall



The economic impact of non-communicable disease in China and India

Manuell:

China, Indien, Prognose, Alternde Bevölkerung, Wirtschaftswachstum, Markoökon. Einfluss, Gesundheitskosten, Sterblichkeit, Krankheit

Automatisch:

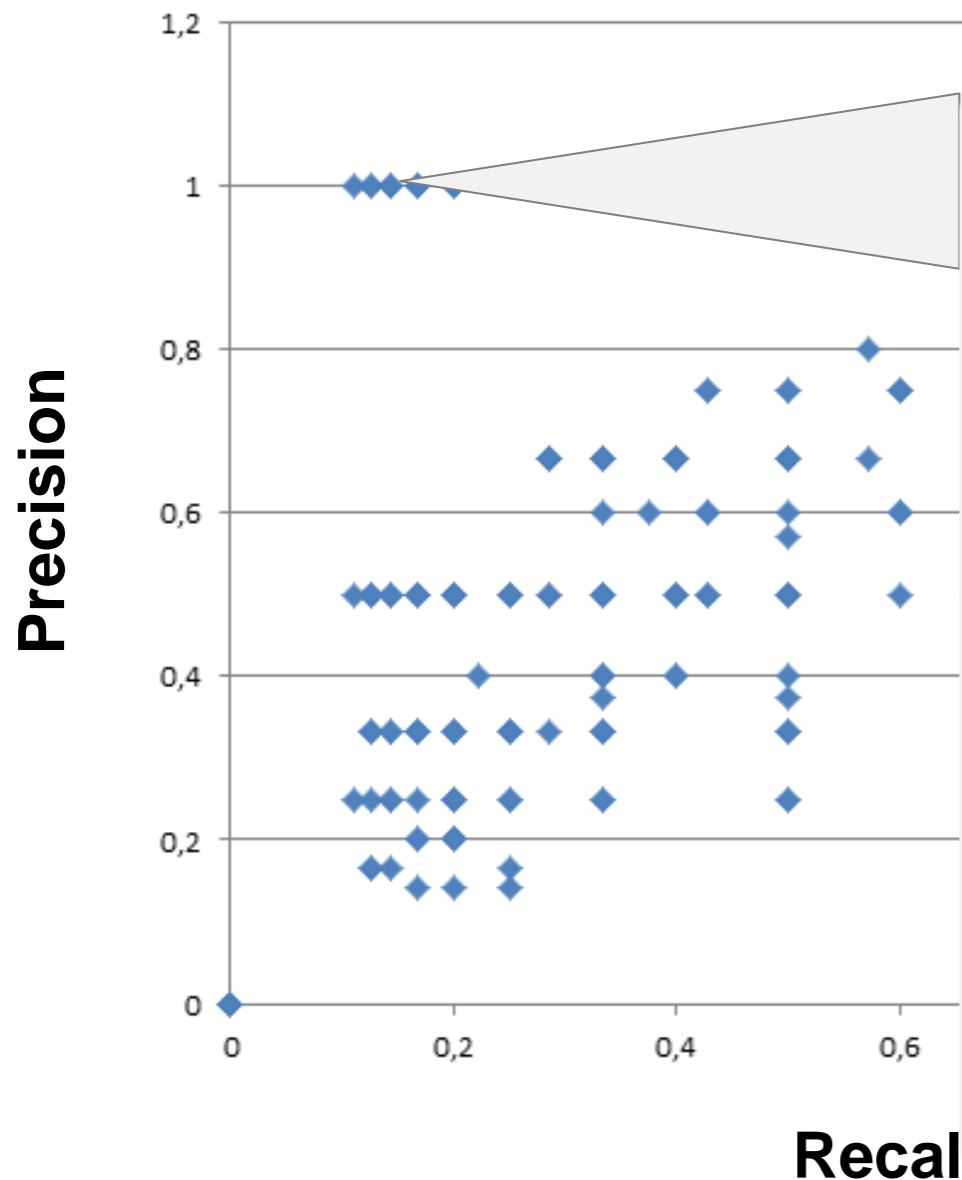
Kosten, Krankheit, Gesundheit, Entwicklung (ökon.)

Keywords:

health and economic development, non-communicable disease, growth models, cost-effectiveness

Nachnormierung von Keywords

Erste Auswertungen – Precision / Recall



Sustainability of small and medium scale enterprises in rural Ghana

Manuell:

Ghana, ländlicher Raum, KMU, Unternehmensfinanzierung, Mikrofinanzierung

Automatisch:

Mikrofinanzierung

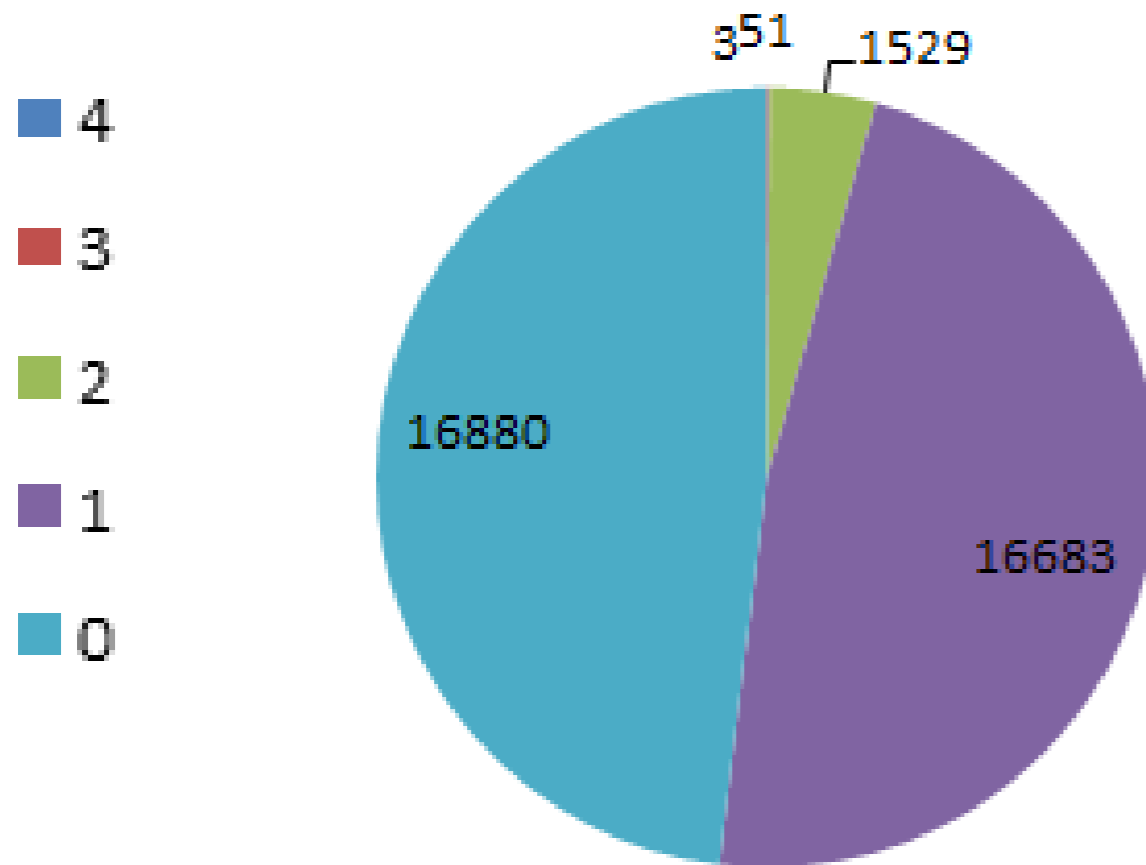
Keywords:

microfinance, small and medium enterprises, microfinance institutions

Nachnormierung von Keywords

Erste Auswertungen

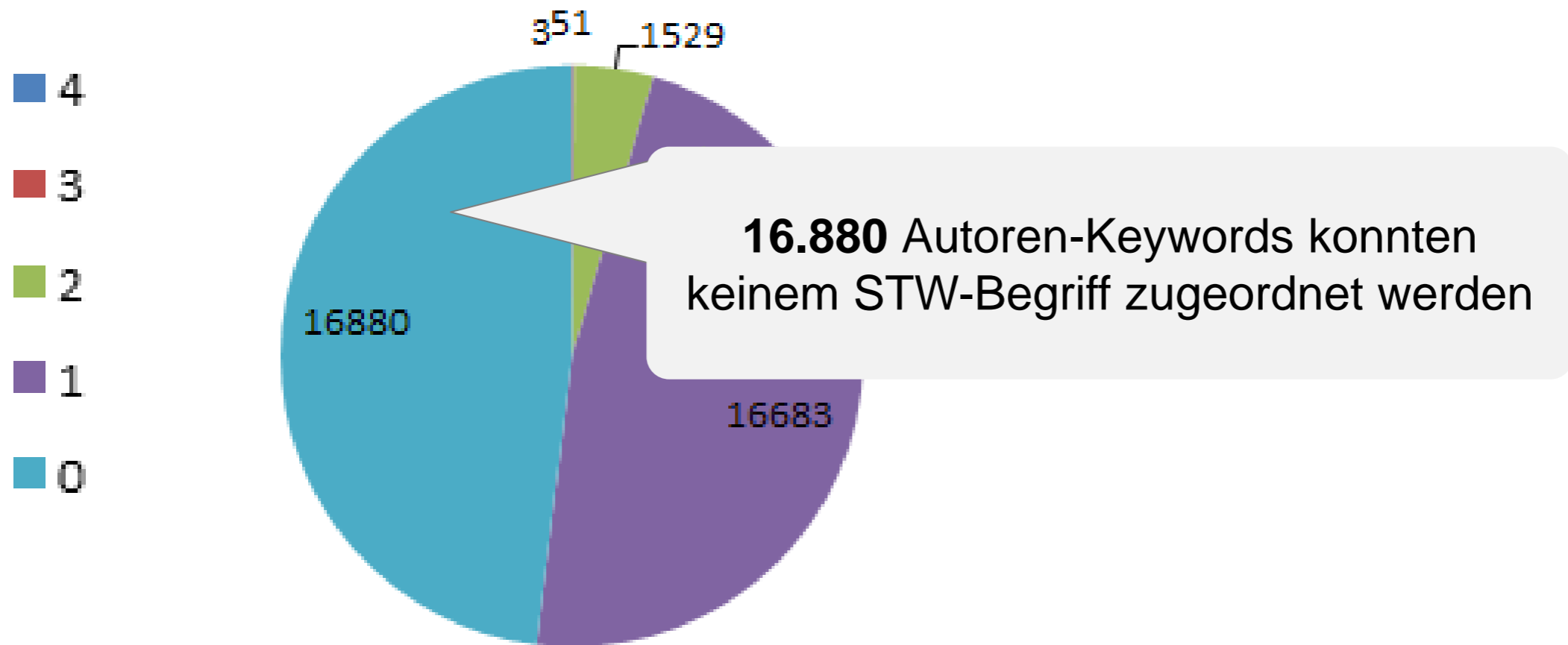
STW Treffer in Keywords



Nachnormierung von Keywords

Erste Auswertungen

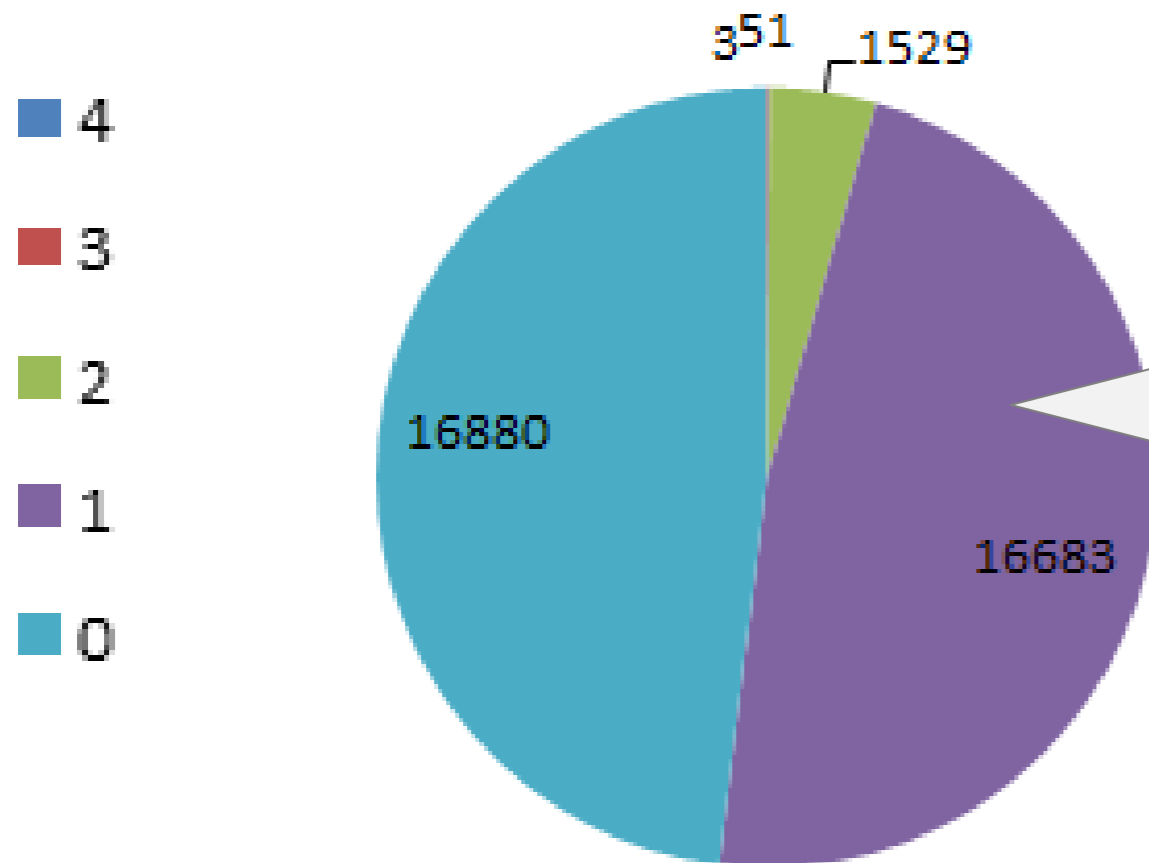
STW Treffer in Keywords



Nachnormierung von Keywords

Erste Auswertungen

STW Treffer in Keywords



16.683 Autoren-Keywords haben genau einen Treffer im STW

Davon **10.548** als **ganze Treffer**
z.B.:

[government failure]

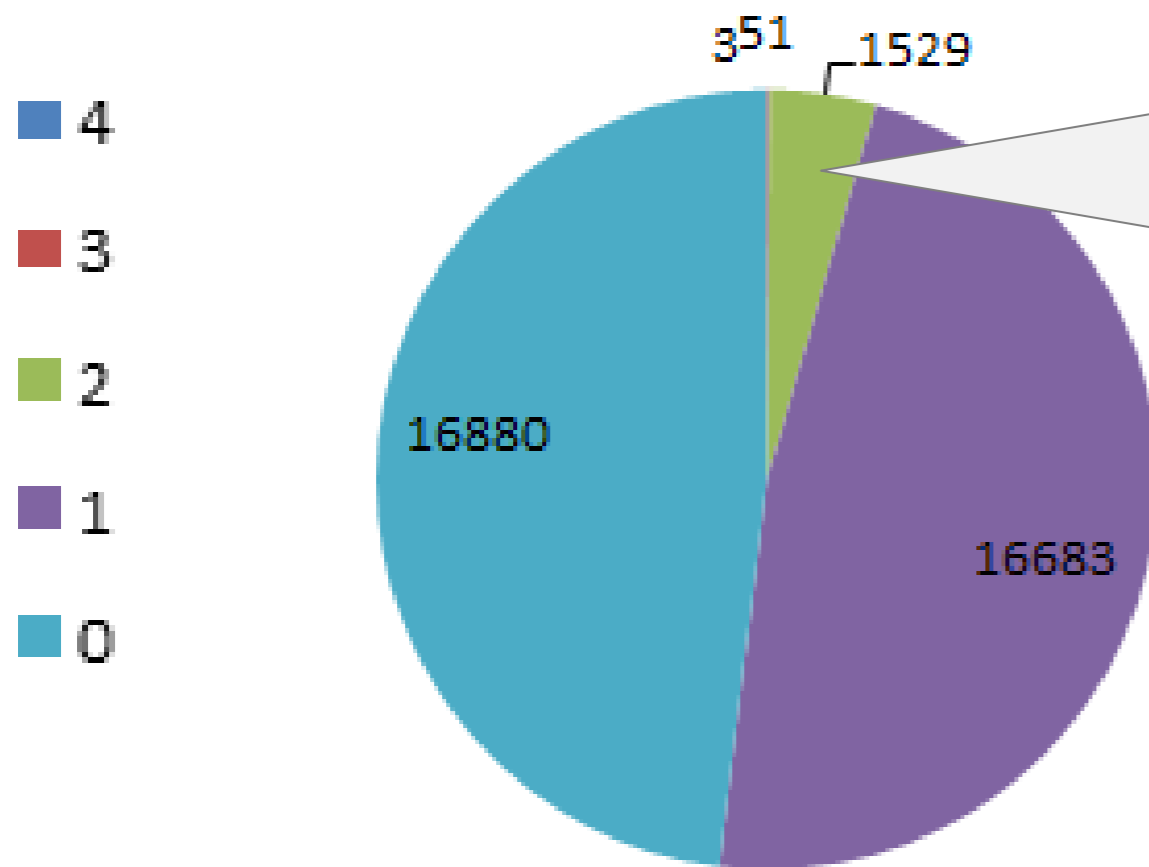
und **6.135** als **partielle Treffer**
z.B.:

[regression] discontinuity design

Nachnormierung von Keywords

Erste Auswertungen

STW Treffer in Keywords



1529 Autoren-Keywords haben genau **zwei** Treffer im STW

z.B.:

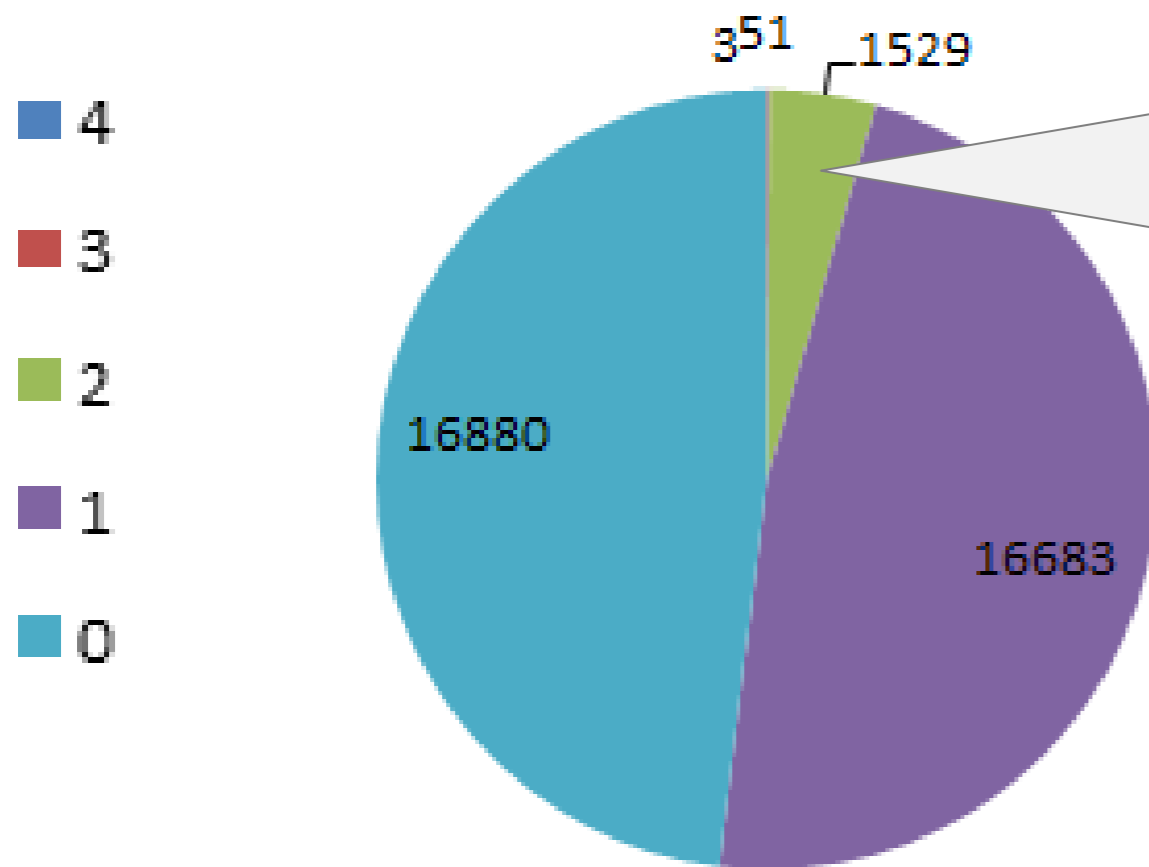
[coordination] [games]

[credit risk] [management]

Nachnormierung von Keywords

Erste Auswertungen

STW Treffer in Keywords



51 Autoren-Keywords haben genau **drei** Treffer im STW

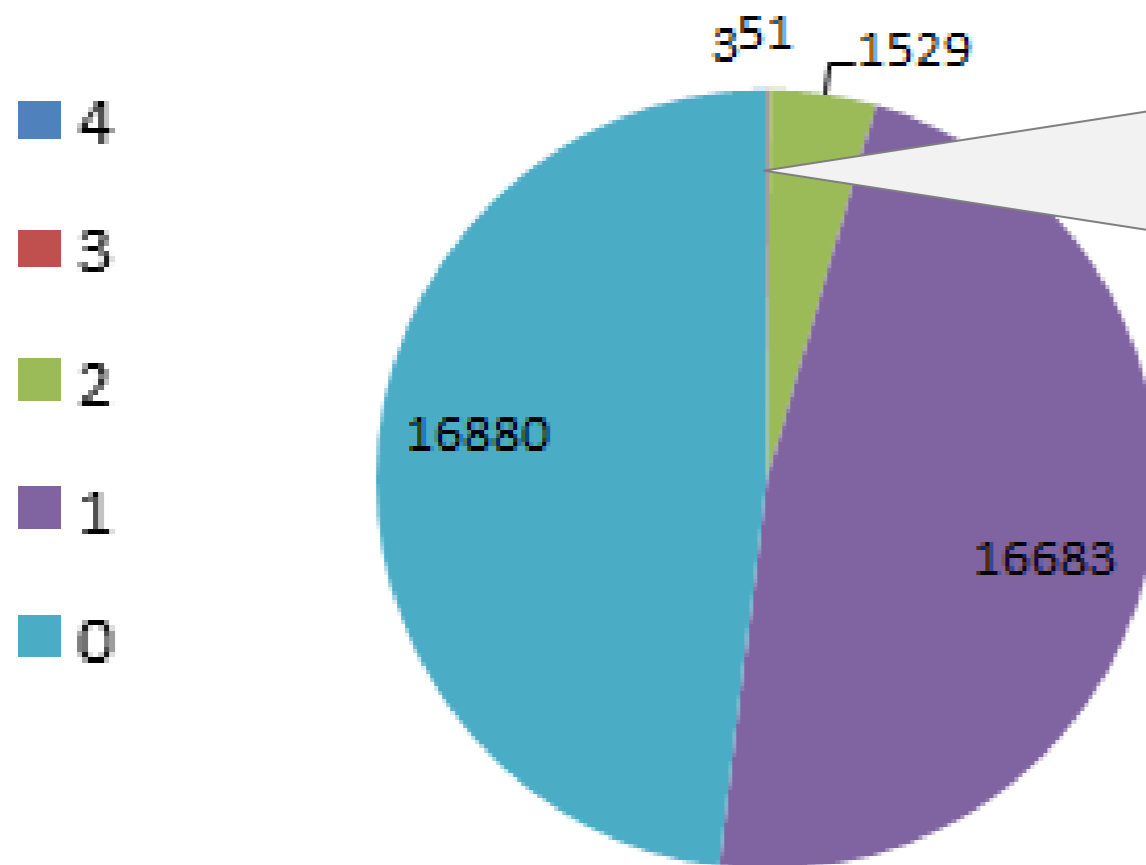
z.B.: interdependent multidimensional
[time] and **[income]** **[poverty]** and
affluence

[World] **[labor]** mobility **[costs]**

Nachnormierung von Keywords

Erste Auswertungen

STW Treffer in Keywords



3 Autoren-Keywords haben genau **vier** Treffer im STW:

öffentliche **[Ausgaben]** für Alterssicherung, **[Gesundheit]**, Pflege, **[Arbeitslose]**, **[Bildung]** und Familien

[ministry] of **[science]** **[research]** and **[technology]**

[energy] / **[electricity]** **[consumption]** and **[economic growth]**

Nachnormierung von Keywords

Erste Auswertungen – **nächste Schritte**

- Erweiterung des Dokumenten-Korpus'
- Einbeziehung von Titel und/oder Abstracts
- Fokus auf Einrichtung einer skalierbaren Umgebung
 - *ein* erweiterbarer Korpus (Versionierung und Updates)
 - Flexibilität bei der Verwendung automatischer Systeme
 - Entkopplung und Adapter
 - Autotagger, MAUI, Recommind,...
 - Evaluierungsroutinen
 - Semi-Automatik und Review-Werkzeuge

Vielen Dank – noch Fragen?

Karin Wortmann

ZBW – Leibniz Informationszentrum Wirtschaft

k.wortmann@zbw.eu

Thomas Gross

ZBW – Leibniz Informationszentrum Wirtschaft

t.gross@zbw.eu

Daniel Bahls

ZBW – Leibniz Informationszentrum Wirtschaft

d.bahls@zbw.eu
