Leibniz-Informationszentrum
Wirtschaft
Leibniz Information Centre
for Economics

Neue Ansätze zur Relevanzsortierung in bibliothekarischen Informationssystemen

Dr. Timo Borst
Informationssysteme und Publikationstechnologien
ZBW Leibniz-Informationszentrum Wirtschaft

Leipzig, Bibliothekskongress, 14. März 2016

Leibniz-Informationszentrum
Wirtschaft
Leibniz Information Centre
for Economics

Überblick

I. Relevance Ranking in bibliothekarischen Informationssystemen

II. Der Ansatz des DFG-Projekts LibRank

III. Rankingfaktoren und -modell

IV. Evaluierungen

V. Ergebnisse und „lessons learned“

Relevance Ranking in bibliothekarischen Informationssystemen

- Bibliothekarische Informationssysteme müssen sich an allgemeine Nutzererwartungen im Web anpassen (hinsichtlich Suchen, Trefferlisten, -sortierung, -präsentation)…
- …dabei aber gleichzeitig den spezifischen Inhalten und Nutzererwartungen der Communities Rechnung tragen (es handelt sich ganz überwiegend um wissenschaftliche Inhalte und/oder Nutzungsszenarien!)
- Erhöhter Bedarf an
  - methodisch fundierter Relevanzbestimmung
  - Transparenz
  - Nachnutzbarkeit der Ergebnisse

Relevance Ranking in bibliothekarischen Informationssystemen

„kurz und bündig“:

Connect from Off-Campus | Contact Us | Site Search

& Help | Services | Locations & Hours | Giving | About Us | My Account

arch »

How does Summon's relevance ranking work?

Summon's sophisticated relevance ranking algorithm is a 'trade secret' of the software developers, so we are unable to explain exactly how it works. However, it gives priority to "exact title" matches on your search keywords.

For instance: a search for "practical grammar" will return items with the exact title "practical grammar" at the top. A search for "the practical grammar" will return almost identical results, but with priority given to titles with "the" in them.

See more FAQs in:

Summon Search

Last updated: 01 September 2010

Quelle: http://library.queensu.ca/help/faq/8043

Relevance Ranking in bibliothekarischen Informationssystemen

„etwas ausführlicher“:

- Maximizing Accuracy with Field Ranking
- Minimizing Full Text Influence
- Enhanced Subject Precision
- Value Ranking
- Local Library Collections
- Delivering Relevant Results

Quelle: http://support.epnet.com/knowledge_base/detail.php?id=3971

Ansatz von LibRank (allgemeine IR)

1. Identifizierung geeigneter **Rankingfaktoren** für bibliothekarische Informationssysteme

2. Entwicklung eines **Rankingmodells**

3. Entwicklung eines **Evaluierungsframeworks** für das Relevance Ranking speziell in Bibliotheksanwendungen

Ansatz von LibRank (bezogen auf EconBiz)

1. Aufbau einer **Testkollektion** mit prototypischen Informationsbedürfnissen („Search Tasks“) und entsprechenden Relevanzbewertungen basierend auf einem repräsentativen Ausschnitt der in EconBiz durchsuchbaren Bestände als Grundlage zur Entwicklung von Relevanz-Modellen

2. **Implementierung des Relevanz-Modells** in EconBiz und Entwicklung entsprechender Komponenten für Lucene/SOLR

3. **Implementierung von Relevanzprofilen** in Abhängigkeit von den jeweiligen Informationsbedürfnissen der verschiedenen Nutzergruppen

Vorgehen bei LibRank

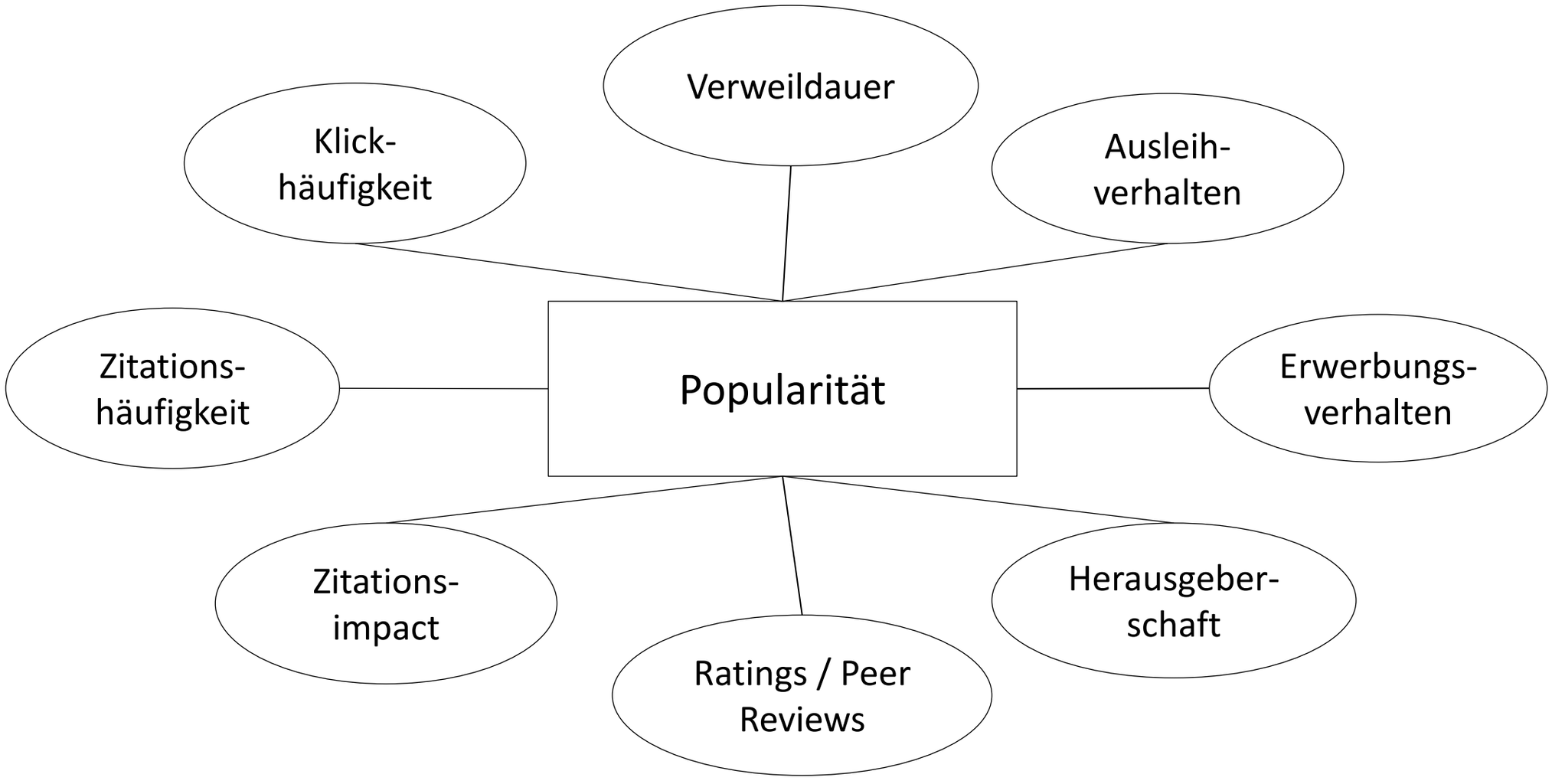
- Analytisch-konzeptionell: (Mögliche) Rankingfaktoren identifizieren
- Praktisch-empirisch: Datenquellen sondieren und „anzapfen“
- Vorgehen (iterativ): Baseline festlegen, Testset bilden, Rankings nach verschiedenen Faktoren und durch verschiedene Nutzergruppen bewerten lassen, Rankings auswerten
- Dokumentation und Verbreitung

(Mögliche) Rankingfaktoren

Abb.: Behnert, C., & Borst, T. (2015). Neue Formen der Relevanz-Sortierung in bibliothekarischen Informationssystemen: Das DFG-Projekt LibRank. BIBLIOTHEK Forschung Und Praxis, 39(3), S. 389. doi:10.1515/bfp-2015-0052

(Mögliche) Rankingfaktoren

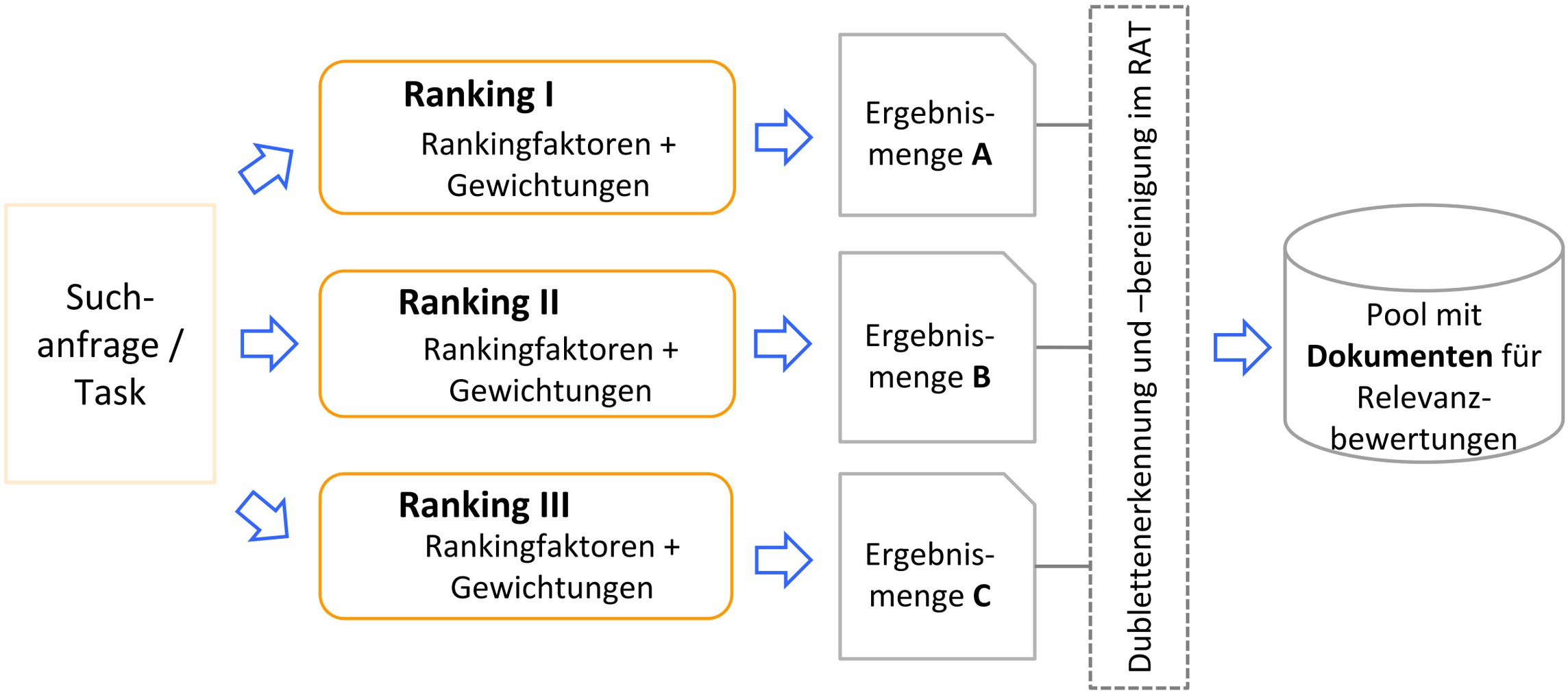
Popularität *- Aktualität - Standort & Verfügbarkeit - Dokumenteigenschaften - Nutzerhintergrund*



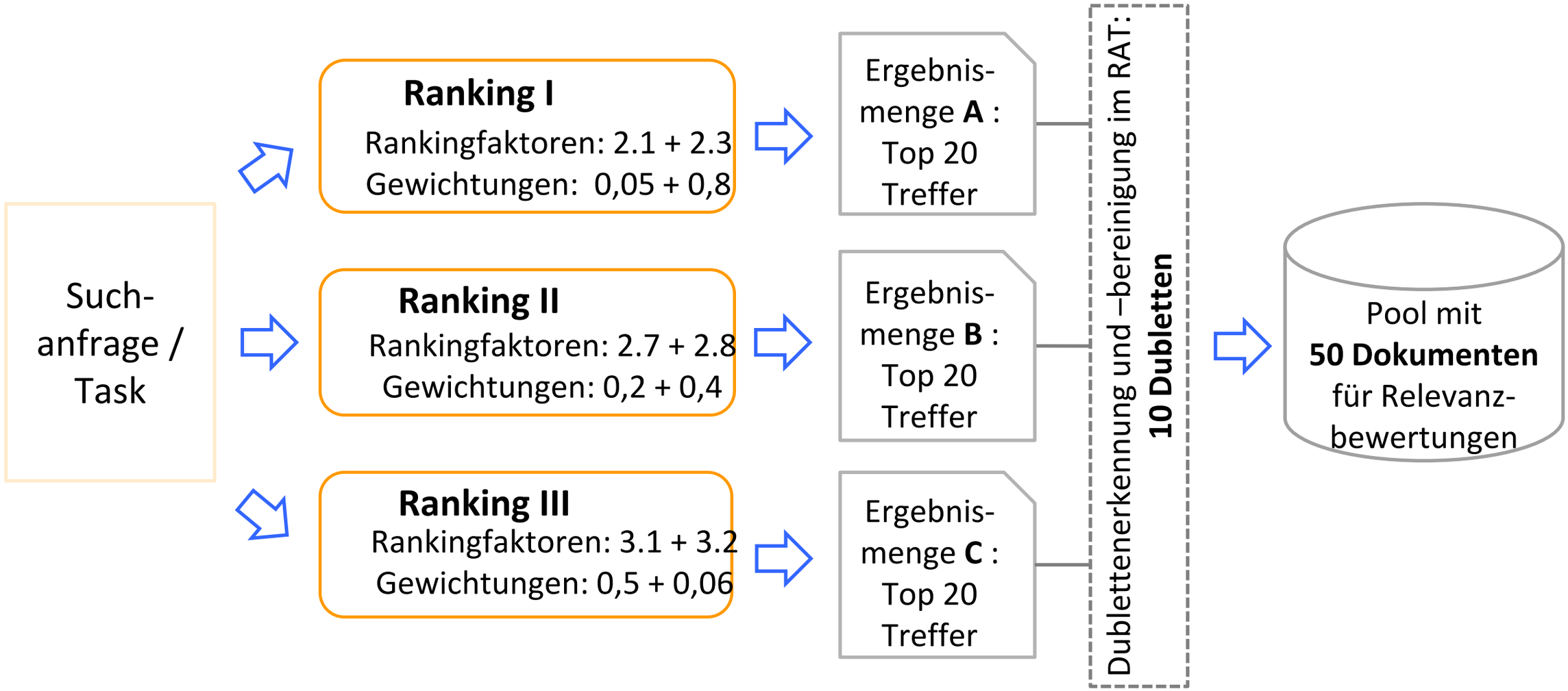
Rankingmodell (?)

$$\mathscr{R} = f_{\text{Popularität}} * 0.75 + f_{\text{Aktualität}} * 0.6 + [\ldots]\ f_{\text{Nutzerprofil}} * 0.45$$

Evaluierungsschema



Evaluierung (Beispiel)



Evaluierungsläufe (Übersicht)

TEST-LAUF	ANZAHL JURORENEN	JURORENGRUPPE	ANZAHL DER TASKS			JURORENVERGÜTUNG	ANZAHL RANKINGS (CUT-OFF)	ANZAHL BEWERT. DOKUMENTE
			GESAMT	PRO JUROR	BEARBEITET			
#1	4	Fachreferenten der ZBW	120	30	83	keine	10 (20)	16.680
#2	8	Fachreferenten der ZBW	120	15	109	keine	10 (20)	22.590
#3	45	Studierende (Wirtschafts-fächer)	450	min.1 (5 Tasks je Paket)	363	20€-Amazon-Gutschein pro bearbeitetes Aufgabenpaket	10 (20)	72.490
Σ	57	-	**690**	-	**555**	-	-	**111.760**

(ca. 1 % des gesamten EconBiz-Bestands)

Evaluierungsläufe (Übersicht zu Rankingmodell)

	Evaluierungslauf #1	Evaluierungslauf #2	Evaluierungslauf #3
Ranking 1	*Nur textstatistische Verfahren „Text statistics“*		
Ranking 2	*EconBiz-Ranking = Baseline*		
Ranking 3	7 Faktoren inkl. Subfaktoren	7 Faktoren inkl. Subfaktoren	*Gelernte Variante*
Ranking 4	Nur Popularität	Nur Popularität	Nur Popularität
Ranking 5	Nur Aktualität	Nur Aktualität	Nur Aktualität
Ranking 6	Nur Verfügbarkeit	Nur Verfügbarkeit	Nur Verfügbarkeit
Ranking 7	Nur Dokumenteigenschaften	Nur Dokumenteigenschaften	Nur Dokumenteigenschaften
Ranking 8	Nutzungshäufigkeit + Ersch.-Datum	Nutzungshäufigkeit + Ersch.-Datum	Verfügbarkeit + Ersch.-Datum + Nutzungshäufigkeit + Zitationen
Ranking 9	Nutzungshäufigkeit + Zitationen	Nutzungshäufigkeit + Zitationen	Verfügbarkeit + Ersch.-Datum + Nutzungshäufigkeit + Zitationen
Ranking 10	Zitationen + Erscheinungsdatum	Zitationen + Erscheinungsdatum	Verfügbarkeit + Ersch.-Datum + Nutzungshäufigkeit + Zitationen

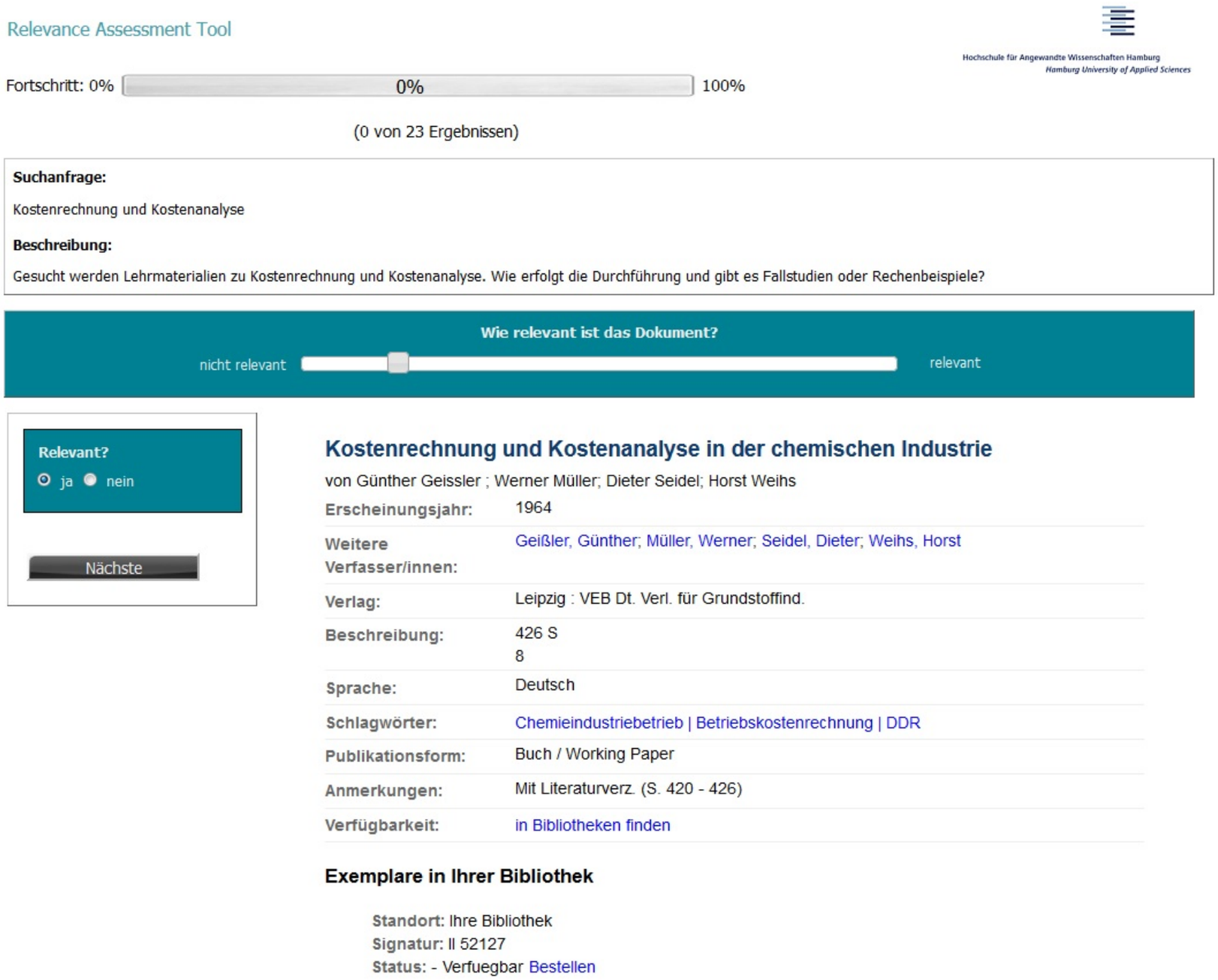
Datenbasis

- Essentiell zur Implementierung eines Rankingmodells in einer konkreten Informationsumgebung

- Man wird i.d.R. nicht zu allen Rankingfaktoren jeweils konkrete Daten erhalten (zumal nicht regelmäßig-aktualisiert)

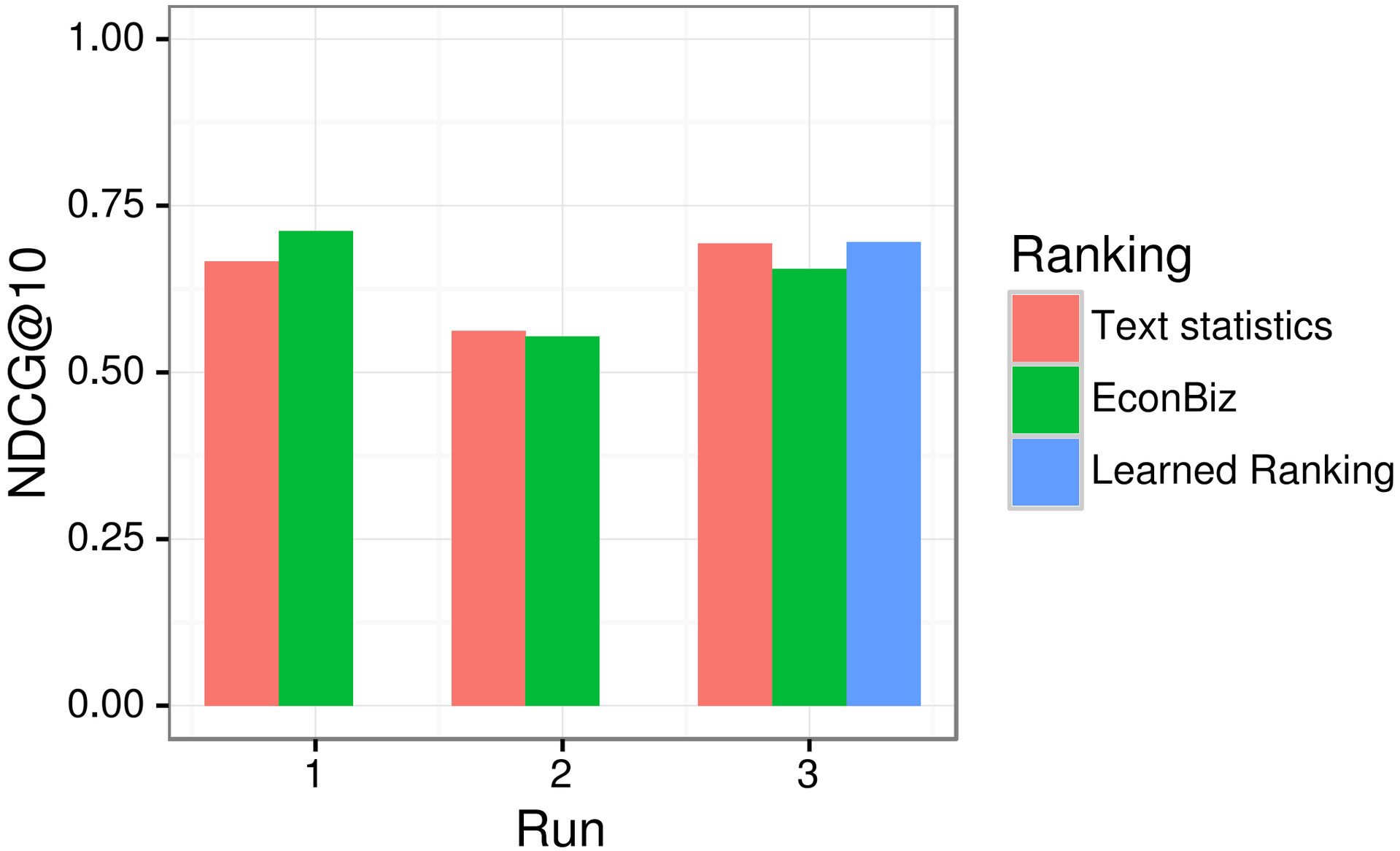
- Im Kontext von EconBiz wurden Daten speziell zur Popularität erhoben bzw. aufbereitet (Logfiles, Zitationsdaten von CitEc,…)

Relevanzbewertungen mit dem RAT



- Suchanfrage mit Beschreibung
- Binäre und Skalenbewertung
- Fortschrittsbalken und Anzahl der zu bewertenden Dokumente (Surrogate)

Performanceauswertung



Ergebnisse und „lessons learned“

- Obwohl in der Informationspraxis wichtig, gelten das Zustandekommen und die Anzeige von Suchergebnissen in IR- und Discovery-Systemen eher als ein „Betriebsgeheimnis“

- Rein analytisch-konzeptionell lassen sich verschiedene Rankingkategorien und -faktoren identifizieren

- Welche davon bei der Praxis der (intuitiven) Relevanzbewertung wirksam sind, lässt sich nach unseren Ergebnissen und den uns zur Verfügung stehenden Methoden nicht zweifelsfrei feststellen

- Rein textstatistische Verfahren schneiden nicht signifikant schlechter ab, bilden aber gleichzeitig die Basis für alle anderen Rankings (Primat der „topicality“)

Ergebnisse und „lessons learned“

- In der Praxis wird man i.a. mit einer ausgewogenen Textstatistik + einer Mischung aus Aktualität und Popularität von Werken ganz gut fahren

- Die einzig-allgemeingültige „Relevanzformel“ haben wir zumindest nicht gefunden

- Alternative/Weiterführende Ansätze:
  - Im Rahmen eines konkreten IR-Systems „Relevanzsignale“ identifizieren (z.B. erhöhter Download), die z.B. mit Query Recommendation in Verbindung stehen
  - Verstärkter Einsatz von Learning-to-rank-Ansätzen
  - Spezifische Nutzermodelle (z.B. prinzipielle Bevorzugung von Verfügbarkeit bei Terminarbeiten, oder peer reviewed Publikationen)
  - Verfeinerung der textstatistischen Verfahren (Ranking der Volltexte)
  - Query understanding and reformulation
  - Relevanzsignale aus den Sozialen Medien

Timo Borst

t.borst@zbw.eu

ZBW Leibniz-Informationszentrum Wirtschaft
Leibniz Information Centre for Economics