

## **Neue Formen der Relevanz-Sortierung in bibliothekarischen Informationssystemen: Das DFG-Projekt LibRank**

### **Christiane Behnert**

Hochschule für Angewandte Wissenschaften Hamburg  
Fakultät Design, Medien, Information  
Department Information  
Finkenau 35  
22081 Hamburg  
christiane.behnert@haw-hamburg.de

### **Dr. Timo Borst**

Deutsche Zentralbibliothek für Wirtschaftswissenschaften  
Abteilung Innovative Informationssysteme und Publikationstechnologien  
Düsternbrooker Weg 120  
24105 Kiel  
t.borst@zbw.eu

*Zusammenfassung:* Das von der DFG geförderte Projekt LibRank erforscht neue Rankingverfahren für bibliothekarische Informationssysteme, die aufbauend auf Erkenntnissen aus dem Bereich Web-Suche qualitätsinduzierende Faktoren wie z.B. Aktualität, Popularität und Verfügbarkeit von einzelnen Medien berücksichtigen. Die konzipierten Verfahren werden im Kontext eines in den Wirtschaftswissenschaften häufig genutzten Rechercheportals (EconBiz) entwickelt und in einem Testsystem systematisch evaluiert. Es werden Rankingfaktoren, die für den Bibliotheksbereich von besonderem Interesse sind, vorgestellt und exemplarisch Probleme und Herausforderungen aufgezeigt.

*Schlüsselwörter:* Relevanz-Ranking, Information Retrieval, Discovery-System, Bibliothekskatalog

### **LibRank: New approaches to relevance ranking in Library Information Systems – a research project funded by the German Research Foundation (DFG)**

*Abstract:* The project explores new approaches to relevance ranking in library information systems. Based on insights into the ranking of web search engines, quality-indicating factors, e.g. freshness, popularity and availability, are taken into consideration. The conceptualized methods are being implemented into a test system based on a library's search portal (EconBiz) and systematically evaluated. Ranking factors suitable to library information systems are discussed with problems and challenges pointed out.

*Keywords:* Relevance Ranking, Information Retrieval, Discovery System, Library Catalogs

## **Inhalt**

1. Ausgangslage .....	2
2. Ziele und Methodik von LibRank .....	3
3. Anwendungskontext EconBiz .....	5
4. Relevanzfaktoren für bibliothekarische Informationssysteme .....	6
5. Herausforderung Datenheterogenität .....	12
6. Fazit und Ausblick.....	14

## 1. Ausgangslage

In der heutigen Zeit ist das Informationssuchverhalten stark geprägt von Websuchmaschinen wie Google, Bing oder Yahoo. Studien zeigten auf, dass Nutzer in der Regel ihre Suchanfragen mit einer geringen Anzahl von Suchbegriffen formulieren und dabei die Treffer auf der ersten Seite einer Trefferliste favorisieren,<sup>1</sup> zumeist sogar lediglich die ersten drei Treffer.<sup>2</sup> Der Grund hierfür liegt in dem Vertrauen in die Suchmaschinen, dass die Suchergebnisse, die am besten zur Suchanfrage passen, also die vermeintlich größte Relevanz besitzen, zuerst angezeigt werden und die weniger relevanten Treffer auf den hinteren Rängen zu finden sind.<sup>3</sup>

Das Suchverhalten und die Erwartungen der Nutzer an die Trefferdarstellung sind allgemein auf den Bibliothekskontext übertragbar.<sup>4</sup> Nutzer sind an die einfache Suchoberfläche (den sogenannten „Google-Suchschlitz“) gewöhnt<sup>5</sup>, die erweiterte Suchmaske wird vergleichsweise selten bzw. seitens der Informationsexperten verwendet. Auch hier umfassen die Suchanfragen durchschnittlich nicht mehr als 2 bis 3 Suchbegriffe.<sup>6</sup> Dabei werden Voreinstellungen, z.B. die automatische UND-Verknüpfung der Suchbegriffe, in der Regel nicht verändert.<sup>7</sup> Ebenso werden die Treffer auf den oberen Rängen stark präferiert.<sup>8</sup>

In traditionellen Bibliothekskatalogen (OPACs) werden die Suchergebnisse üblicherweise innerhalb eines Bestandes nach Erscheinungsjahr sortiert - das Relevanz-Ranking folgt hier lediglich einem relativ starren Muster. Die Erwartungen an die Trefferdarstellung sind jedoch nicht nur in Bezug auf das Ranking geprägt, auch die Qualität der Suchergebnisse lässt Defizite erkennen. Dies ist damit zu begründen, dass das Retrieval in traditionellen OPACs auf Basis des Booleschen Modells und des *exact match*-Ansatzes funktioniert, wodurch nur die Dokumente (bzw. Dokumentrepräsentationen) gefunden werden, die mit den Suchbegriffen (syntaktisch) exakt übereinstimmen. Es ist jedoch lange bekannt, dass Nutzer ohne bibliothekarisches Wissen mit der richtigen Verwendung von Booleschen Operatoren nicht vertraut sind und daher die Suchanfrage oft nicht zufriedenstellende Ergebnisse liefern kann.<sup>9</sup> Im Gegensatz dazu verfolgen Suchmaschinen den Ansatz des *best match*<sup>10</sup>: Mithilfe

<sup>1</sup> Vgl. Jansen, Bernard J.; Spink, Amanda.: How are we searching the World Wide Web? A comparison of nine search engine transaction logs. In: Information Processing & Management 42(1) 2006, S. 248–263. Und: Vgl. Höchstötter, Nadine; Koch, Martina: Standard parameters for searching behaviour in search engines and their empirical evaluation. In: Journal of Information Science 35(1) 2008, S. 45–65.

<sup>2</sup> Vgl. Höchstötter, Nadine; Lewandowski, Dirk: What users see - Structures in search engine results pages. Information Sciences 179(12) 2009, S. 1796–1812.

<sup>3</sup> Vgl. Pan, Bing; Hembrooke, Helene; Joachims, Thorsten; Lorigo, Lori; Gay, Geri; Granka, Laura: In Google We Trust: Users' Decisions on Rank, Position, and Relevance. In: Journal of Computer-Mediated Communication 12(3) 2007, S. 801–823.

<sup>4</sup> Vgl. Hennies, Markus; Dressler, Juliane: Clients information seeking behaviour: An OPAC transaction log analysis. In: click 06, ALIA Biennial Conference. Perth, AU 2006.

<sup>5</sup> Vgl. Yu, Holly; Young, Margo: The impact of web search engines on subject searching in OPAC. In: Information Technology & Libraries 23(4) 2004, S. 168–180.

<sup>6</sup> Vgl. Niu, Xi; Hemminger, Bradley M.: Beyond text querying and ranking list: How people are searching through faceted catalogs in two library environments. In: Proceedings of the American Society for Information Science and Technology 47(1) 2010, S. 1–9.

<sup>7</sup> Vgl. Asher, Andrew D.; Duke, Lynda M.; Wilson, Suzanne: Paths of discovery: Comparing the search effectiveness of EBSCO Discovery Service, Summon, Google Scholar, and conventional library resources. In: College & Research Libraries 74(5) 2013, S. 464–488.

<sup>8</sup> Vgl. Antelman, Kristin; Lynema, Emily; Pace, Andrew K.: Toward a twenty-first century library catalogue. In: Information Technology & Libraries 25(3) 2006, S. 128–139.

<sup>9</sup> Vgl. Khoo, Christopher S. G. G.; Wan, Kwok-Wai: A simple relevancy-ranking strategy for an interface to Boolean OPACs. In: The Electronic Library 22(2) 2004, S. 112–120.

linguistischer Verfahren können Suchbegriffe (syntaktisch) modifiziert werden, sodass auch Trefferdokumente erscheinen, die ähnliche Begriffe enthalten, z.B. erzielt die Anfrage „statistische methoden“ auch Treffer mit dem Begriff „Statistik“. Durch semantische Erweiterungen können beispielsweise auch Synonyme mitberücksichtigt werden, ohne dass die Suchanfrage durch ODER-Verknüpfungen künstlich erweitert werden muss.

Damit OPACs langfristig den Nutzererwartungen gerecht werden können, gilt es, die technologischen Möglichkeiten hinsichtlich Suche und Relevanz-Ranking gezielt auszuschöpfen.<sup>11</sup> Mit dem Einsatz von Suchmaschinentechnologie in Bibliothekssystemen wurde begonnen, diesem Defizit entgegen zu wirken.<sup>12</sup> Die sogenannten Kataloge der nächsten Generation (auch *Next Generation Catalogs* oder *Discovery Systeme*) ermöglichen bereits die Suche in elektronischen Dokumenten durch Volltextindexierung sowie den direkten Zugriff auf lizenzierte oder frei verfügbare Inhalte, die mit dem traditionellen Bibliothekskatalog als Nachweisinstrument der analogen Bestände nicht sichtbar sind. Mithilfe dieser integrativen Systeme wird Nutzern die Suche über das gesamte Bibliotheksangebot ermöglicht, ohne dass diese die unterschiedlichen Datenquellen manuell ansteuern müssen.

Such- und Zugriffsmöglichkeiten weisen dadurch bereits eine höhere Qualität auf, jedoch werden die Erwartungen an das Ranking nur teilweise erfüllt. Die einzelnen Algorithmen, die dem Ranking in Software-Tools, wie beispielsweise in den kommerziellen Lösungen *Summon* von ProQuest (ehemals als Serviceangebot von Serials Solutions) oder *Primo* von ExLibris, zugrunde liegen, sind weder vollständig transparent noch von den jeweiligen bibliothekseigenen IT-Abteilungen ohne weiteres anpassbar. Zwar kommen durch den zunehmenden Einsatz von Suchmaschinen wie Lucene/SOLR mittlerweile auch deren standardmäßig verwendete Rankingparameter und -Scores zum Einsatz (vorzugsweise der textstatistische tf/idf-Algorithmus), jedoch wurde bisher keine systematische Evaluierung der überhaupt infrage kommenden oder verwendeten Rankingfaktoren durchgeführt. An diesem Punkt setzt das Projekt LibRank an.

## 2. Ziele und Methodik von LibRank

Das Ziel des Projektes ist die Entwicklung von neuartigen Rankingverfahren für Bibliotheksinformationssysteme auf Basis von Suchmaschinentechnologie und entsprechendem Nutzerverhalten. Hierbei sind zum einen qualitätsinduzierende Faktoren, wie z.B. Aktualität, Popularität und Verfügbarkeit von einzelnen Medien zu berücksichtigen, zum anderen müssen die Suchtreffer den Nutzererwartungen entsprechend präsentiert werden. Die Bereitstellung eines Demonstrators, mit dem Bibliotheken und Wissenschaftseinrichtungen die für ihre eigenen Daten optimalen Rankingfaktoren ermitteln können, ist ebenfalls Ziel des Projekts. Neben dem Demonstrator werden auch die während des Projekts aufgebauten Testkollektionen zum Zweck der Transparenz und Nachnutzbarkeit bereitgestellt.

---

<sup>10</sup> Vgl. Baeza-Yates, Ricardo; Ribeiro-Neto, Berthier: *Modeling*. In: *Modern Information Retrieval: the concepts and technology behind search*. 2. edition, Harlow [u.a.] 2011, S. 57–130.

<sup>11</sup> Vgl. Lewandowski, Dirk: *Ranking library materials*. In: *Library Hi Tech* 27(4) 2009, S. 584–593.

<sup>12</sup> Vgl. Lewandowski, Dirk: *Suchmaschinen als Konkurrenten der Bibliothekskataloge: Wie Bibliotheken ihre Angebote durch Suchmaschinentechnologie attraktiver und durch Öffnung für die allgemeinen Suchmaschinen populärer machen können*. In: *Zeitschrift für Bibliothekswesen und Bibliographie*, 53(2) 2006, S. 71–78.

Die Deutsche Forschungsgemeinschaft (DFG) fördert das Gemeinschaftsprojekt LibRank<sup>13</sup> seit März 2014 für die Projektlaufzeit von zwei Jahren. Das Projekt wird gemeinsam mit der Hochschule für Angewandte Wissenschaften Hamburg (HAW) und der Deutschen Zentralbibliothek für Wirtschaftswissenschaften (ZBW) durchgeführt, womit Kompetenzen aus der Informationswissenschaft und -praxis gebündelt werden. Auf der Grundlage der genannten wissenschaftlichen Ziele wird grundsätzlich die Verbesserung des Rankings in dem ZBW-eigenen Fachportal EconBiz<sup>14</sup> angestrebt. Mit dieser Anwendung können die konzeptuellen Überlegungen unter realen Praxisbedingungen getestet und das Relevanzmodell als Projektergebnis schließlich in EconBiz – bzw. angesichts seiner prinzipiellen Übertragbarkeit auch in anderen bibliothekarischen Informationssystemen – implementiert werden.

Zur Erreichung dieser Ziele wurden zunächst für den Bibliothekskontext mögliche Relevanz-Rankingfaktoren identifiziert (siehe Abschnitt 4), für deren Aufstellung die grundlegenden Konzepte und Ideen von Ranking in Websuchmaschinen untersucht wurden. Diese Rankingfaktoren werden im weiteren Projektverlauf mithilfe eines gewichteten Relevanzmodells in einem auf dem EconBiz-Datenbestand basierenden Testsystem systematisch evaluiert. Da EconBiz und allgemein moderne bibliothekarische Informationssysteme Suchmaschinentechnologie verwenden, konnte ein bereits getestetes Framework zur Evaluierung der Retrieval-Effektivität von Suchmaschinen<sup>15</sup> teilweise für die Evaluierung des EconBiz-Testsystems adaptiert und angepasst werden.

In den Testläufen werden mithilfe eines an der HAW entwickelten Tools zur Relevanzbewertung (*Relevance Assessment Tool - RAT*)<sup>16</sup> Suchanfragen an das Testsystem geschickt und die jeweiligen Suchergebnisse von Juroren nach ihrer Relevanz bewertet. Die Relevanzbewertungen werden analysiert und das Ranking bzw. die Gewichtungen einzelner Faktoren basierend auf den Ergebnissen des entsprechenden Testdurchlaufs angepasst und in dem darauffolgenden Durchgang erneut bewertet. Dabei werden die Rankingfaktoren sowohl individuell als auch in Kombination mit anderen Faktoren betrachtet, um einerseits die Faktoren mit dem stärksten Einfluss auf die Retrieval-Effektivität festzustellen und andererseits Korrelationen zu ermitteln.

Für die Bewertung der Suchergebnisse werden reale Suchanfragen aus EconBiz den Juroren mit einer kurzen Beschreibung des jeweiligen Informationsbedürfnisses präsentiert, wobei die Beschreibungstexte von Studierenden der Wirtschaftswissenschaften erstellt wurden, um das Informationsverhalten der Fachcommunity realitätsnah abzubilden. Ein Beispiel:

- Suchanfrage: *Prinzipal-Agenten-Theorie*
- Beschreibung: *Was besagt diese Theorie und wo findet sie Anwendung?*

Die Suchtreffer (Titel datensätze, ggf. mit Abstract bzw. Beschreibung) werden in einer zufälligen, nicht dem Rankingalgorithmus entsprechenden Reihenfolge angezeigt, um eine mögliche Beeinflussung der Juroren auszuschließen. Dies geschieht vor dem Hintergrund, dass die Ergebnisse

---

<sup>13</sup> Siehe auch den Webaufttritt des Projekts unter: <http://www.librank.info/de/>

<sup>14</sup> Der Sucheinstieg und weitere Informationen sind zu finden unter: <http://www.econbiz.de/>

<sup>15</sup> Vgl. Lewandowski, Dirk: A Framework for Evaluating the Retrieval Effectiveness of Search Engines. In: Next Generation Search Engines: Advanced Models for Information Retrieval. Hershey, PA. 2012, S. 456–479.

<sup>16</sup> Vgl. Lewandowski, Dirk; Sünkler, Sebastian: Relevance Assessment Tool: Ein Werkzeug zum Design von Retrievaltests sowie zur weitgehend automatisierten Erfassung, Aufbereitung und Auswertung von Daten. In: Social Media und Web Science : das Web als Lebensraum ; 2. DGI-Konferenz ; 64. Jahrestagung der DGI ; Düsseldorf, 22. bis 23. März 2012. Frankfurt am Main 2012, S. 237–249.

auf den ersten Positionen in einer Trefferliste in der Regel bevorzugt geklickt werden, wie bereits in Abschnitt 1 erwähnt. Die Juroren bewerten die Ergebnisse mithilfe eines im RAT implementierten Schiebereglers auf einer Skala von 0 (nicht relevant) bis zu 100 möglichen Punkten (höchst relevant). Diese Relevanzbewertungen werden neben den Suchanfragen und Trefferdokumenten als Testkollektionen nach Projektabschluss über ein geeignetes Repository zur Nachnutzung zur Verfügung gestellt.

### 3. Anwendungskontext EconBiz

Grundthema von LibRank ist die Übertragbarkeit von Beobachtungen und Methoden zum Informationsverhalten bei der allgemeinen Websuche auf ein bibliothekarisches Informationssystem. Hierfür wurde das wirtschaftswissenschaftliche Fachportal EconBiz aus verschiedenen Gründen ausgewählt:

- Als Suchanwendung beinhaltet EconBiz den unseres Wissens nach größten offenen Suchindex (9,5 Millionen Titeldaten, Stand März 2015) für wirtschaftswissenschaftliche Fachinformationen, zu denen neben herkömmlicher Verlagsliteratur die in den Wirtschaftswissenschaften verbreiteten Arbeitspapiere sowie Veranstaltungshinweise gehören.<sup>17</sup> M.a.W., EconBiz liefert praktisch für jede fachlich einschlägige Suchanfrage eine gewisse Menge an Ergebnissen und somit überhaupt den Bedarf und die Notwendigkeit einer relevanzbasierten Sortierung zur Bewältigung des ‚information overload‘.
- Insbesondere wenn man berücksichtigt, dass sich EconBiz als wirtschaftliches Fachportal an eine relativ spezielle Zielgruppe richtet, wird es mit monatlich zwischen 150.000 und 300.000 Besuchen gut genutzt.<sup>18</sup>
- Gegenüber allgemeinen Websuchmaschinen ist die Nutzergruppe bei EconBiz nachweislich relativ homogen, vornehmlich bestehend aus wirtschaftsstudierenden und -forschenden Personen<sup>19</sup>. Damit erhöht sich im Prinzip die Wahrscheinlichkeit, dass eine für das Informationsverhalten dieser Zielgruppe optimierte Relevanzsortierung tatsächlich auch flächendeckend angenommen, zumindest aber analysiert werden kann. Die in dem Projekt durchgeführten Evaluierungsläufe lassen sich auf ein für die Nutzergruppe typisches - durch eine Reihe von internen Studien<sup>20</sup> bereits belegtes - Informationsverhalten abstimmen, so dass im Rahmen von LibRank keine zusätzlichen Studien zum Informationsverhalten durchgeführt werden müssen.
- EconBiz basiert auf VuFind<sup>21</sup>, einer Bibliothekssoftware zur facettierten Recherche in bibliothekarischen Metadaten, die wiederum auf der Suchmaschinentechnologie Lucene/SOLR aufsetzt. Aufgrund der Quelloffenheit des Systems ergibt sich die Möglichkeit bzw. Notwendigkeit,

<sup>17</sup> Vgl. Pianos, Tamara: Hin zum Fachportal? Hin zur Kundschaft? Oder beides? In: Bibliothek : Forschung und Praxis 38(1) 2014, S. 47–55.

<sup>18</sup> Vgl. Pianos (Anm. 17).

<sup>19</sup> Vgl. Pianos, Tamara: EconBiz - Meeting user needs with new technology. In: Liber Quarterly 20(1) 2010, S. 4–24.

<sup>20</sup> Siehe z.B. die Ergebnisse einer von der ZBW in Auftrag gegebenen Studie, zusammengefasst in: Siegfried, Doreen; Flieger, Elisabeth: World Wide Wissenschaft – Wie professionell Forschende im Internet arbeiten. ZBW - Leibniz-Informationszentrum Wirtschaft 2011. Zugriff:

[http://www.zbw.eu/presse/pressemitteilungen/docs/world\\_wide\\_wissenschaft\\_zbw\\_studie.pdf](http://www.zbw.eu/presse/pressemitteilungen/docs/world_wide_wissenschaft_zbw_studie.pdf)

<sup>21</sup> <http://vufind-org.github.io/vufind/>

die Berechnung des Relevanz-Scores einzusehen und ggf. zu optimieren. Im Gegensatz hierzu sind bei den *Discovery-Systemen* kommerzieller Anbieter die Grundlagen und Verfahren zur Relevanzberechnung nur bedingt intransparent und veränderbar.<sup>22</sup>

- Mit VuFind bzw. Lucene/SOLR als Open Source-Basis ist eine Verbreitung, Nachnutzung und ggf. Weiterentwicklung der LibRank-Ergebnisse durch andere Informationseinrichtungen am ehesten gegeben: Die Software wurde dezidiert für die Verarbeitung bibliothekarischer Metadaten entwickelt und wird mittlerweile von einer Reihe entsprechender Einrichtungen eingesetzt.<sup>23</sup> Die auf GitHub gestellten (Weiter-)Entwicklungen dokumentieren darüber hinaus eine aktive Entwickler-Gemeinde.<sup>24</sup>

#### 4. Relevanzfaktoren für bibliothekarische Informationssysteme

Bevor Faktoren für Relevanz-Ranking entwickelt werden können, ist die wissenschaftliche Durchdringung des Themas Relevanz unabdingbar. Da Relevanz von höchst subjektiver und dynamischer Natur ist, wird sie von jedem Menschen anders wahrgenommen: Sind Informationen für den einen relevant, sind sie es für eine andere Person nicht zwingend in gleichem Maße; selbst für denselben Informationssuchenden kann aktuell relevante Literatur zu einem späteren Zeitpunkt gänzlich irrelevant sein.<sup>25</sup> Nutzungskontexte und Informationsbedürfnisse ändern sich, wobei dies nicht notwendig durch eine Suchanfrage erkenntlich ist. Hier zeigt sich ein Merkmal von Information-Retrieval-Systemen (IRS): Informationsbedürfnisse können (noch) nicht in natürlicher Sprache ausgedrückt und an das IRS übermittelt werden<sup>26</sup>, sondern sind zu einer Suchanfrage zu formulieren. Für viele Nutzer stellt dies eine große Herausforderung dar; beispielsweise knüpfen an dieser Stelle im Kontext Katalogsuche Bibliothekare als vermittelnde Instanzen an und geben Hilfestellungen bei der Recherche.

Aufgrund der subjektiven Aspekte bei der Informationsbeschaffung ist es ausgeschlossen, Relevanz aus rein objektiver, logischer Sichtweise heraus zu betrachten. Selbst das Ranking in Websuchmaschinen basiert lediglich auf wahrscheinlicher Relevanz<sup>27</sup>, d.h. es existiert keine richtige oder falsche Relevanzsortierung von Dokumenten, weil Dokumente für jeden Informationssuchenden eine andere Relevanz besitzen können. Obwohl viele unterschiedliche Definitionen von Relevanz in der Literatur zu finden sind<sup>28</sup>, fasst die folgende Aussage eine

<sup>22</sup> Allein EBSCO hat unlängst in einem kurzen Beitrag Einblicke in die Relevanzberechnung beim „EBSCO Discovery Service“ geboten: [http://support.epnet.com/knowledge\\_base/detail.php?id=3971](http://support.epnet.com/knowledge_base/detail.php?id=3971) (letzter Zugriff: 30.04.2015)

<sup>23</sup> [https://vufind.org/wiki/installation\\_status](https://vufind.org/wiki/installation_status)

<sup>24</sup> <https://github.com/vufind-org/vufind/graphs/contributors>

<sup>25</sup> Vgl. Saracevic, Tefko: Relevance: A review of the literature and a framework for thinking on the notion in information science. Part II. In: *Advances in Librarianship* (30) 2006, S. 3–72.

<sup>26</sup> Hier spielt das Thema Anfrageinterpretation eine signifikante Rolle, denn das Ranking einer Dokumentenmenge kann nur zufriedenstellend sein, wenn die Treffermenge an sich mit hoher Wahrscheinlichkeit relevante Dokumente enthält. Ohne das eigentliche Informationsbedürfnis zu kennen, ist es für ein IRS schwierig, auf Basis von zwei bis drei Suchbegriffen dieses zu interpretieren. Für einen Überblick zum Thema siehe z.B. Lewandowski, Dirk: *Query Understanding*. In: *Handbuch Internet-Suchmaschinen 2 : Neue Entwicklungen in der Web-Suche*. Heidelberg, 2011, S. 55–75.

<sup>27</sup> Vgl. Lewandowski, Dirk: *Suchmaschinen verstehen*. Berlin ; Heidelberg 2015, S. 30.

<sup>28</sup> Mizzaro hat mehr als 150 Studien zum Thema Relevanz analysiert. Vgl. Mizzaro, Stefano: Relevance: The whole history. In: *Journal of the American Society for Information Science and Technology* 48(9) 1997, S. 810–832.

realistische Erkenntnis prägnant zusammen: „*Nobody has to explain to users of IR systems what relevance is, even as they struggle (sometimes in vain) to find relevant stuff. People understand relevance intuitively.*“<sup>29</sup>

Auf welche Weise Nutzer die Relevanz von Informationen beurteilen, wurde von Barry und Schamber bereits in den 1990er Jahren analysiert. Sie identifizierten die folgenden zehn Kriterien:<sup>30</sup>

1. *Depth/Scope/Specificity*: Die Informationen sind ausführlich, beinhalten eine Zusammenfassung oder Interpretation.
2. *Accuracy/Validity*: Die Informationen sind genau, korrekt bzw. valide.
3. *Clarity*: Die Informationen werden klar und leicht verständlich präsentiert.
4. *Currency*: Die Informationen sind aktuell, auf dem neuesten Stand.
5. *Tangibility*: Die Informationen sind bewiesen, es werden harte Fakten genannt.
6. *Quality of Sources*: Die Quelle ist reputabel, vertrauenswürdig.
7. *Accessibility*: Der Zugang zu den Informationen ist ohne Anstrengungen oder Kosten möglich.
8. *Availability of Information/Sources of Information*: Die Informationen bzw. Quellen sind verfügbar.
9. *Verification*: Die Informationen sind konsistent oder werden durch andere Informationen innerhalb des Forschungsfelds oder der Disziplin gestützt.
10. *Affectiveness*: Die Informationen rufen eine emotionale Reaktion hervor, wie z.B. Vergnügen oder Unterhaltung.

Auf Basis dieser Relevanzkriterien lassen sich die Relevanzfaktoren bezüglich Ranking ableiten, jedoch sind sie nicht direkt auf die einzelnen Rankingalgorithmen übertragbar. Dennoch sind Faktoren wie Aktualität (*currency*) und Verfügbarkeit (*availability*) hier ebenso enthalten und werden im Folgenden neben vier weiteren Faktorengruppen vorgestellt.

Abbildung 1 gibt zunächst einen Überblick über die sechs identifizierten Relevanzfaktorengruppen. Ausgehend von *textstatistischen Verfahren*, auf deren Basis überhaupt erst eine Treffermenge als Antwort jedes textbasierten IRS auf die vom Nutzer gestellte Suchanfrage generiert werden kann, schließen sich qualitätsinduzierende Faktoren an. Diese können hauptsächlich der Gruppe *Popularität* zugeordnet werden, wobei die Gruppen *Aktualität* sowie *Standort & Verfügbarkeit* ebenfalls einen Mehrwert für das Ranking darstellen. Die letzten beiden Faktorengruppen vereinen *Dokumenteigenschaften* und Informationen über den *Hintergrund der Nutzer*.

---

<sup>29</sup> Saracevic, Tefko: Relevance reconsidered. In: Information science: Integration in perspectives. Proceedings of the Second Conference on Conceptions of Library and Information Science (CoLIS 2). Copenhagen: Royal School of Librarianship 1996, S. 215.

<sup>30</sup> Vgl. Barry, Carol; Schamber, Linda: Users' criteria for relevance evaluation: A cross-situational comparison. In: Information Processing & Management 34(2-3) 1998, S. 219–236.



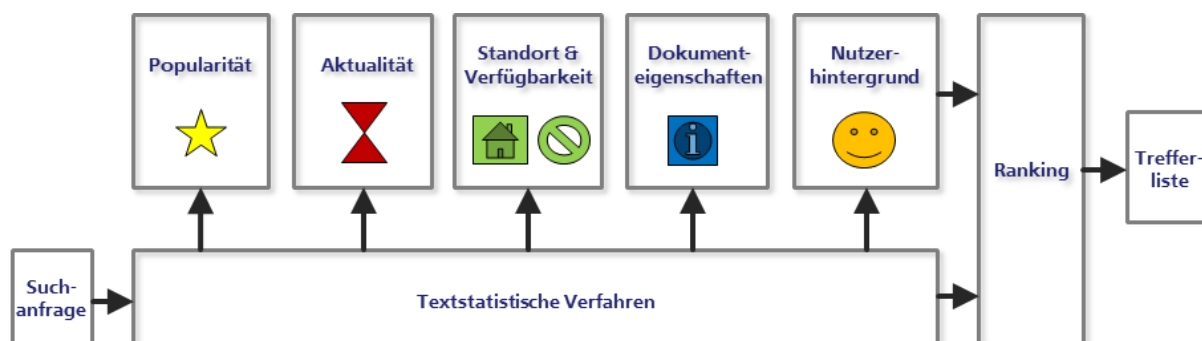


Abbildung 1: Gruppen von Rankingfaktoren

### Textstatistische Verfahren

In textbasierten Retrieval-Systemen werden ausgehend von einer Suchanfrage die Suchbegriffe mit einer Dokumentenmenge abgeglichen und in letzterer passende Treffer vom IRS identifiziert (*text matching*). Die zwei gängigsten statistischen Rankingverfahren hierbei sind das Messen der relativen Häufigkeit eines Suchbegriffs in einem Dokument (*term frequency - TF*) und die inverse Dokumentenhäufigkeit (*inverse document frequency - IDF*), welche die relative Worthäufigkeit innerhalb einer Menge von Dokumenten berücksichtigt. Wichtig hierbei ist jedoch, dass die alleinige Häufigkeit eines Wortes nicht sehr aussagekräftig ist, denn Stoppwörter, z.B. Artikel, Konjunktionen usw., treten im Allgemeinen am häufigsten auf. Aus diesem Grund werden die Begriffe gewichtet – das meist genutzte Gewichtungsverfahren ist eine Kombination der beiden Methoden (*TF-IDF*), mit der seltene Wörter höher gewichtet werden, wie beispielsweise Eigennamen von Personen oder Ländern.<sup>31</sup>

Aufgrund dieser Gewichtung ist das exakte Auftreten der Suchbegriffe in der Reihenfolge, in der sie in der Suchanfrage gestellt wurden, nicht mehr ausschließlich erforderlich, da nun auch die teilweise Übereinstimmung der Begriffe zählt – also *best match* anstelle von *exact match*. Zwar wird die Reihenfolge von Begriffen auch weiterhin für die Gewichtung berücksichtigt, aber auch die Position der Suchbegriffe innerhalb des Dokuments und das Auftreten in einem Metadatenfeld spielen unter anderem eine große Rolle.<sup>32</sup> So wird beispielsweise das Titelfeld höher gewichtet als das Abstract und dieses wiederum höher als der eigentliche Text.

Websuchmaschinen und Discovery-Systeme im Bibliotheksbereich verwenden zwar Verfahren nach dem *best match*-Ansatz, jedoch reichen textstatistische Methoden alleine nicht aus, um ein gutes Ranking zu erzielen. Insbesondere bei Bibliothekskatalogen stoßen textstatistische Verfahren schnell an Grenzen, da die Metadatenfelder im Vergleich zum Volltext zu wenig Text enthalten, als dass auf deren Basis die genannten Verfahren effektiv für das Relevanzranking angewandt werden können.<sup>33</sup> Zudem sind die Objekte in den heutigen Bibliotheksinformationssystemen sehr heterogen, da neben Metadaten zu gedruckten Büchern auch Artikel im Volltext oder Audiomaterialien enthalten sind. Aus diesem Grund ist ein Ranking, welches allein auf textstatistischen Verfahren beruht, nicht mehr zeitgemäß.

<sup>31</sup> Vgl. Larson, Ray R.: Information retrieval systems. In: Understanding Information Retrieval Systems. Boca Raton 2012, S. 15–30.

<sup>32</sup> Vgl. Dopichaj, Philipp: Ranking-Verfahren für Web-Suchmaschinen. In: Handbuch Internet-Suchmaschinen. Heidelberg 2009, S. 101–115.

<sup>33</sup> Vgl. Lewandowski (Anm. 11) S. 588ff.



## Popularität

Die begrenzte Reichweite von rein statistischen Verfahren zur Berechnung des Rankings haben die Gründer von Google bereits im Jahr 1998 mit der Entwicklung des PageRank-Verfahrens verdeutlicht.<sup>34</sup> Obwohl Googles Ranking-Algorithmus ständig weiterentwickelt wird, liegt das linkbasierte Ranking auch heute noch diesem komplexen System zugrunde. Linkbasiertes Ranking beruht auf der Annahme, dass das Ranking einer Webseite rekursiv aus dem Ranking der verlinkenden Seiten berechnet werden kann.

Die Qualität einer Website oder eines Dokuments kann als Popularitätsindikator gesehen werden, wobei Popularität dem Prinzip der „Weisheit der Vielen“<sup>35</sup> folgt: Je mehr Nutzer ein Dokument positiv bzw. als relevant bewerten, umso wahrscheinlicher ist, dass es auch für den einzelnen Nutzer relevant ist. Dieser Ansatz findet sich beispielsweise in der Klickhäufigkeit, d.h. wie häufig eine Webseite oder ein Dokument angeklickt, also geöffnet wurde, wieder. Allerdings sollte die Zahl der Klicks im Zusammenhang mit der Verweildauer betrachtet werden. Diese beruht auf der Annahme, dass Dokumente, die nach nur wenigen, d.h. 2 oder 3, Sekunden geschlossen werden, als irrelevant bewertet werden. Eine längere Verweildauer, z.B. von einigen Minuten, spricht jedoch für eine gewisse Relevanz und ist demzufolge als ein Indikator für Qualität anzusehen.<sup>36</sup> Ein wesentlich stärkerer Indikator als die Zahl der Klicks ist die Zahl der Downloads, da diese ein höheres Nutzungsinteresse implizieren als das reine Ansehen bzw. Öffnen von Dokumenten.

Die Nutzungshäufigkeit bzw. das Nutzungsinteresse zu analysieren, ist im Bibliotheksbereich bei weitem keine Neuerscheinung. Eine Form ist die Auswertung von Downloadstatistiken: Im Zusammenhang mit Bestandsaufbau und -pflege können Downloadzahlen elektronischer Dokumente wie Ausleihzahlen konventioneller Bibliotheksmaterialien ebenso als Maße für Rankingfaktoren dienen. So werden beispielsweise diejenigen Bücher, die unter Berücksichtigung des Themengebiets (siehe Schlagworte, Klassifizierung) häufiger ausgeliehen bzw. online aufgerufen werden als andere, inhaltlich vergleichbare Bücher, entsprechend höher gewichtet. Auf diese Weise rücken besonders beliebte Bücher zu dem gesuchten Thema auf den oberen Rängen der Trefferliste leichter in das Blickfeld der informationssuchenden Person. Ausleihzahlen als Popularitätsfaktor wurden z.B. an der Bibliothek der North Carolina State University, USA, im Zusammenhang mit Empfehlungen eingesetzt und getestet, mit dem Ergebnis, dass diese Form der Treffersortierung als zweithäufigste nach dem Faktor „Erscheinungsjahr“ großen Zuspruch fand.<sup>37</sup> Ausgewählte Popularitätsindikatoren wurden z.B.

---

<sup>34</sup> Vgl. Brin, Sergey; Page, Lawrence: The anatomy of a large-scale hypertextual web search engine. In: Computer Networks and ISDN Systems 30(1-7) 1998, S. 107–117.

<sup>35</sup> Vgl. Surowiecki, James: Die Weisheit der Vielen: Warum Gruppen klüger sind als Einzelne und wie wir das kollektive Wissen für unser wirtschaftliches, soziales und politisches Handeln nutzen können. 1. Aufl. München: 2005.

<sup>36</sup> Die reale Nutzungsintention bzw. inwieweit das Dokument das Informationsbedürfnis des Nutzers befriedigen konnte, kann alleine durch Zählen von Klicks und Messen der Verweildauer nicht ermittelt werden: Liegen zwischen dem Aufrufen von zwei Dokumenten mehrere Minuten, kann dies unterschiedliche Ursachen haben (z.B. das Öffnen eines anderen Dokuments im Hintergrund oder Ablenkung durch ein Telefongespräch) und bedeutet nicht zwangsläufig, dass das Dokument tatsächlich gelesen wurde. Diese Problematik besteht allerdings allgemein bei Nutzungsanalysen, jedoch liegen hier relativ hohe Datenbasen zugrunde, wodurch allgemeine Schlussfolgerungen getroffen werden können.

<sup>37</sup> Vgl. Antelman; Lynema; Pace (Anm. 8).

auch von den Bibliotheken der Universitäten Heidelberg (HEIDI)<sup>38</sup> und Bremen (E-Lib)<sup>39</sup> in das jeweilige Relevanz-Ranking integriert.

Neben der Nutzungshäufigkeit als Indikator für Qualität spielen im akademischen Kontext allgemein Zitationszahlen eine große Rolle. Für Zeitschriften sei hier exemplarisch der Journal Impact Factor<sup>40</sup> zu nennen, der Einfluss von Autoren kann z.B. mit dem h-Index<sup>41</sup> bemessen werden. Bibliometrische Methoden kommen insofern zur Anwendung, als dass - unabhängig von Rankingalgorithmen - auf vorhandene Maße zur Qualität, Autorität, Popularität oder Relevanz zurückgegriffen wird.

Ein Problem bezüglich der praktischen Umsetzung der oben genannten Faktoren besteht in der direkten Vergleichbarkeit und Übertragbarkeit: Es stellt sich beispielsweise die Frage, ob ein Download mit einer Ausleihe gleichzusetzen ist oder 20 Zitationen eines Aufsatzes doppelt so viel wiegen sollten wie 10 Zitationen eines anderen. Aus diesem Grund kann nicht nur ein einzelner Rankingfaktor das ausschlaggebende Maß für Popularität sein, sondern die Faktoren müssen in Kombination und mit Blick auf ihre Gewichtung betrachtet werden, auch um einer durch Mängel in der Datenbasis hervorgerufenen Benachteiligung bestimmter Dokumente (z.B. Bücher, die aus unterschiedlichen Gründen keine Ausleihzahlen haben, oder Artikel, für die keine Zitationsdaten vorliegen) entgegen zu wirken.

### *Aktualität*

Die Sortierung von Suchergebnissen in traditionellen Bibliothekskatalogen erfolgt üblicherweise nach dem Erscheinungsjahr. Aktualität als Rankingfaktor ist somit kein neuer Ansatz, sondern neben der alphabetischen Sortierung die nächstliegende Methode. Dass diese sich bezüglich gedruckter Monografien und den Möglichkeiten der formalen Erschließung bewährt hat, lässt sich nicht bestreiten. Allerdings ist die Menge an Suchergebnissen, die heutzutage von Discovery-Systemen geliefert wird, wesentlich umfangreicher und Nutzer erwarten zudem eine durchmischte Trefferliste, also beispielsweise sowohl Monografien als auch Artikel aus (elektronischen) Zeitschriften, sowohl aus dem gedruckten Bestand der Bibliothek als auch aus den lizenzierten elektronischen Ressourcen.

Während die Fähigkeit, aktuelle Ergebnisse zu liefern, ein sehr wichtiges Qualitätsmerkmal von Suchmaschinen darstellt<sup>42</sup>, sind im wissenschaftlichen Kontext die Informationsbedürfnisse oft nicht nur mit den aktuellsten Dokumenten zu befriedigen. Je nach Fachdisziplin, wie z.B. in der Philosophie oder historischen Wissenschaften, und in Abhängigkeit des individuellen Forschungs- bzw. Studieninteresses, nehmen ältere Werke einen besonderen Stellenwert ein (*Klassiker*). Zudem können Artikel auch nach etlichen Jahren (erneut) an Popularität gewinnen, wenn ein Themengebiet

---

<sup>38</sup> Vgl. Maylein, Leonhard; Langenstein, Anette: Neues vom Relevanz-Ranking im HEIDI-Katalog der Universitätsbibliothek Heidelberg. In: B.I.T. Online 16(3) 2013, S. 190–200.

<sup>39</sup> Vgl. Haake, Elmar: "Nur die ersten drei zählen!" Erfolgsfaktoren beim Betrieb von Discovery Systemen / Neues Relevanz Ranking. Vortrag auf dem 103. Deutschen Bibliothekartag. Bremen 2014. Zugriff: <http://www.opus-bayern.de/bib-info/volltexte//2014/1598/>

<sup>40</sup> Vgl. Garfield, Eugene: The history and meaning of the journal impact factor. In: JAMA : The Journal of the American Medical Association 295(1) 2006, S. 90–93.

<sup>41</sup> Vgl. Hirsch, J. E.: An index to quantify an individual's scientific research output that takes into account the effect of multiple coauthorship. In: Scientometrics 85(3) 2010, S. 741–754.

<sup>42</sup> Vgl. Lewandowski, Dirk; Höchstötter, Nadine: Web searching: A quality measurement perspective. In: Web Search. Berlin ; Heidelberg 2008, S. 309–340.

mit einem gewissen zeitlichen Abstand zum Erscheinungsdatum ein hohes Interesse innerhalb einer wissenschaftlichen Community weckt (*Hot Topic*).<sup>43</sup>

An dieser Stelle wird ebenfalls deutlich, dass nicht nur Faktoren *innerhalb* der Gruppe Popularität in Kombination berücksichtigt werden müssen - insbesondere *in Verbindung* mit Aktualitätsfaktoren (z.B. Publikationsjahr) sind diese von besonderer Bedeutung für das Relevanz-Ranking.

### **Standort und Verfügbarkeit**

Die Nutzung von Standortdaten ist spätestens seit dem Aufkommen mobiler Endgeräte eine unverzichtbare Datenquelle für das Ranking in Suchmaschinen. Ein Beispiel: Jemand befindet sich in London und sucht mithilfe eines Smartphones den Begriff „Pizzeria“. Die Suchmaschine interpretiert diese Anfrage so, dass sie erwartungsgemäß die nächstgelegene Pizzeria in London als bestes Suchergebnis ausgibt, ungeachtet der eventuell sehr hohen Popularität einer Pizzeria in Rom. Der physische Standort des Nutzers und die Entfernung zum gesuchten Objekt beeinflussen in diesem Fall in hohem Maße das Ranking.<sup>44</sup>

Das Konzept von Standortberücksichtigung bei der Websuche lässt sich auch mit Blick auf Verfügbarkeitsinformationen auf den Bibliotheksbereich übertragen. Zum einen kann unterschieden werden zwischen dem Standort des Nutzers (vor Ort, also in den Räumen der Bibliothek bzw. im Netz der Bibliothek) und des Mediums generell (Erwerbungsdaten, Lizenzinformationen); zum anderen über die aktuelle Verfügbarkeit des Mediums (Ausleihdaten, Lizenzinformationen).

Im Rahmen von LibRank wird diesen Faktoren anhand der Berücksichtigung zweier Nutzermodelle Rechnung getragen (Tabelle 1). Dabei legen wir die Annahme zugrunde, dass Nutzer, die sich nicht vor Ort befinden, elektronische Dokumente, die sofort verfügbar sind (durch lizenzierte E-Books bzw. E-Journals), bevorzugen, weil diese gegebenenfalls sofort das Informationsbedürfnis befriedigen.

<b>Nutzermodell</b>	<b>Annahme über präferierte Dokumenttypen</b>
Nutzer ist vor Ort	Gedruckt und elektronisch
Nutzer ist nicht vor Ort	Elektronisch

**Tabelle 1: Nutzermodelle und deren präferierte Dokumenttypen**

Anhand dieser Rankingfaktoren zeigt sich, dass sich bestimmte Faktoren in Kombination und deren jeweilige Gewichtungen widersprechen können, da beispielsweise beliebte Bücher, die ohnehin hohe Ausleihzahlen verzeichnen, demzufolge höher gerankt werden, wodurch unter Umständen keine Exemplare verfügbar sind, was wiederum zu Frustration bei den Nutzern führen kann.<sup>45</sup> Generell muss bei der Kombination von unterschiedlichen Rankingfaktoren nicht nur deren bestmögliche (individuelle) Gewichtung, sondern das Ranking in seiner Gesamtheit betrachtet werden. Die richtige Balance zu finden, kann unter Umständen auch nur eine Kompromisslösung sein. „Learning to rank“-

<sup>43</sup> Vgl. Van Raan, Anthony F. J.: Sleeping Beauties in science. In: *Scientometrics* 59(3) 2004, S. 467–472.

<sup>44</sup> Vgl. Jones, Rosie; Zhang, Wei Vivian; Rey, Benjamin; Jhala, Pradhuman; Stipp, Eugene: Geographic intention and modification in web search. In: *International Journal of Geographical Information Science* 22(3) 2008, S. 229–246.

<sup>45</sup> Vgl. Langenstein, Anette; Maylein, Leonhard: Relevanz-Ranking im OPAC der Universitätsbibliothek Heidelberg. In: *B.I.T. Online* 12(4) 2009, S. 408–413.

Ansätze, die demgegenüber auf ein prinzipielles Optimum bei der Parametrisierung und Gewichtung von Rankingfaktoren zielen<sup>46</sup>, werden im Rahmen des Projekts nicht behandelt.

### *Dokumenteigenschaften*

Faktoren dieser Gruppe bedienen sich der formalen Eigenschaften eines Informationsobjekts. Das Vorhandensein zusätzlicher Informationen generiert einen Mehrwert in Bezug auf die Relevanzbeurteilung, z.B. mithilfe von Abstracts<sup>47</sup> und Inhaltsverzeichnissen<sup>48</sup>. Auch das Verlinken von den der Publikation zugrundeliegenden Forschungsdaten kann als ein Merkmal für Qualität gesehen und als Rankingfaktor berücksichtigt werden. Weitere Faktoren dieser Gruppe sind das Dateiformat bzw. der Dokumenttyp und die Sprache des Dokuments, wobei die Sprache durchaus ein wichtiges Kriterium für Relevanz ist, da selbst das populärste, inhaltlich bestpassende, sofort und kostenfrei verfügbare Dokument nicht zwingend auch die höchste Relevanzbewertung erzielt, wenn der Inhalt vom Nutzer aus sprachlicher Sicht nicht verständlich ist.<sup>49</sup>

### *Nutzerhintergrund*

Bereits bei der Anwendung von Popularitätsfaktoren werden teilweise *Nutzungsdaten* ausgewertet, die nach dem Prinzip der Weisheit der Vielen Rückschlüsse auf die wahrscheinlich vorhandenen Präferenzen eines einzelnen ziehen lassen. Im Gegensatz dazu dienen hier *Nutzerdaten* als Grundlage für das Ranking. Hier geht es z.B. darum, den (akademischen) Hintergrund der Nutzer zu berücksichtigen: beim „fachspezifischen Boosting“ werden Dokumente aus dem jeweiligen Forschungsumfeld höher gewichtet.<sup>50</sup>

Hinter dem Stichwort *personalisiertes Ranking*<sup>51</sup> verbirgt sich der Ansatz, individuelles Such-, Klick-, Lese-, Ausleih-, Nutzungsverhalten auszuwerten und Erkenntnisse daraus in das Ranking, angepasst auf den jeweiligen Nutzer, einfließen zu lassen. Die Voraussetzung hierfür ist einerseits die Zuordnung der Aktivitäten eines Nutzers zu demselben (Nutzungsprofil) und zum anderen sein Einverständnis, dass solche Daten gesammelt und verwendet werden dürfen. Dies kann beispielweise mit der Anmeldung im Bibliothekskonto erfolgen.

## 5. Herausforderung Datenheterogenität

Die fast zwangsläufig entstehende Heterogenität bei der Sammlung und Integration von (Meta-) Daten aus verschiedenen Quellen ist ein klassischer Topos verteilter bzw. aggregierender

---

<sup>46</sup> Siehe z.B. Liu, Tie-Yan: Learning to Rank for Information Retrieval. In: Foundations and Trends® in Information Retrieval, 3(3) 2009, S. 225–331.

<sup>47</sup> Vgl. Nicholas, David; Huntington, Paul; Jamali, Hamid R.: The use, users, and role of abstracts in the digital scholarly environment. In: The Journal of Academic Librarianship 33(4) 2007, S. 446–453.

<sup>48</sup> Diese sind heutzutage in Bibliothekskatalogen häufig verlinkt, was unter den Begriff *Catalogue Enrichment* fällt.

<sup>49</sup> Im Bereich Websuche greift dieser Faktor mit höherer Gewichtung (Filter durch Spracheinstellung des Browsers) als im Kontext Bibliothek, da sich in der Praxis derartige Anforderung beim Bestandsaufbau und in den Erwerbungsprofilen widerspiegeln.

<sup>50</sup> Vgl. Maylein; Langenstein (Anm. 38) S. 190–200.

<sup>51</sup> Vgl. Riemer, Kai; Brüggemann, Fabian: Personalisierung der Internetsuche - Lösungstechniken und Marktüberblick. In: Handbuch Internet-Suchmaschinen. Heidelberg 2009, S. 148–171.

Informationssysteme<sup>52</sup> und auch bei EconBiz zu beobachten: Über die Hälfte der indexierten Daten stammen aus fremden Quellen (siehe Abbildung 2, Stand: April 2015) und haben damit in der Regel - und nach allen praktischen Erfahrungen - verschiedene Metadatenformate bzw. unterschiedliche Ausprägungen desselben Metadatenelements. Durch das Aggregieren verschiedener, aber fachlich verwandter Datenquellen entstehen Duplikate, also Mehrfachnachweise zu einem Werk – auch dies erschwert eine auf textuelle Inhalte bezogene Relevanzbewertung.

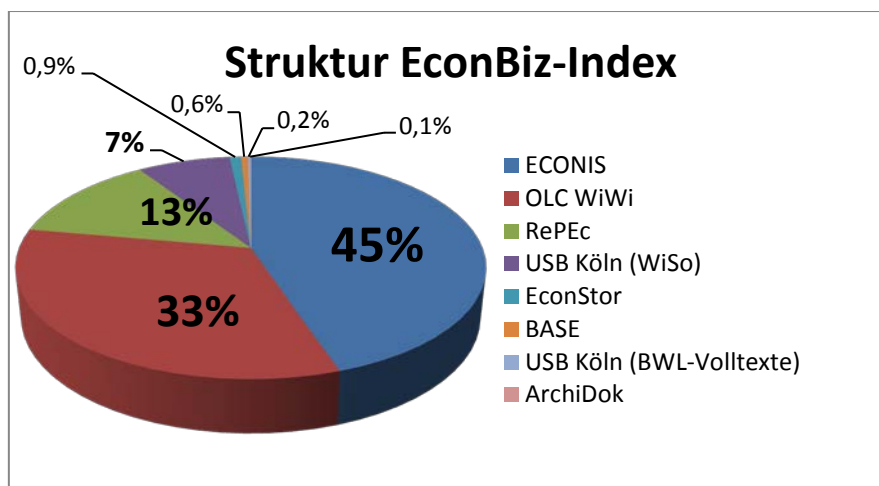


Abbildung 2: Struktur des EconBiz-Index

Selbst die Daten innerhalb einer Quelle oder eines Schemas sind nur bedingt als homogen zu bezeichnen: Gerade für die Beurteilung der inhaltlichen Relevanz wichtige Metadaten, wie z.B. „Abstract“ oder „Beschreibung“, sind nicht immer vorhanden, Deskriptoren bzw. Schlagwörter werden nicht einheitlich vergeben, oder es fehlen gerade im wissenschaftlichen Kontext wichtige Informationen wie die Klassifikation eines Artikels bzw. Journals als begutachtet. Schließlich liegen die für die textstatistische Relevanzbewertung wichtigen Volltexte nicht immer vor bzw. können als lizenzierte Verlagspublikationen zwar für die interne Relevanzberechnung einbezogen, den Nutzern nach außen jedoch nicht transparent zur Verfügung gestellt werden.

Dies berücksichtigend stellt sich die Grundfrage, wie die o.g. Heterogenitätsfaktoren möglichst ausgeblendet, minimiert oder kompensiert werden können. Der zunächst naheliegende Ansatz, die vorhandenen Metadaten zu vereinheitlichen, liegt angesichts des enormen redaktionellen Aufwands und der Masse an Metadaten außerhalb des Projekts und dürfte auch kein praktikables Vorgehen für andere Einrichtungen sein. Stattdessen verfolgen wir den Ansatz, bei den vorhandenen deskriptiven Metadaten vor allem auf die relativ "verlässlichen", also statischen Daten, wie z.B. das Publikationsjahr, zurückzugreifen, sowie den vorhandenen Gesamtindex um weitere Metadaten möglichst flächendeckend anzureichern.

Zu letzterem zählt, dass wir zum einen auf *systemexterne*, vor allem bibliometrische Relevanzfaktoren wie die Anzahl an Zitationen (Datenquelle: CitEc<sup>53</sup>) sowie – darauf aufbauend – den

<sup>52</sup> Vgl. Mayr, Philipp; Stempfhuber, Maximilian; Walter, Anne-Kathrin: Auf dem Weg zum wissenschaftlichen Fachportal – Modellbildung und Integration heterogener Informationssammlungen. In: 27. DGI-Online-Tagung, Frankfurt am Main 2005. S. 29-43.

<sup>53</sup> <http://citec.repec.org/>

h-Index von Autoren zurückgreifen. Darüber hinaus haben wir auf *systembedingte*, vornehmlich webstatistische Relevanzfaktoren abgestellt, wie z.B. die Nutzung elektronischer Ressourcen in Form von Downloads, Ausleihzahlen oder die Verweildauer auf einer Einzeltrefferanzeige.

## 6. Fazit und Ausblick

Die Entwicklung und systematische Evaluierung von neuen Rankingverfahren ist durch eine hohe Komplexität gekennzeichnet: Die Identifikation unterschiedlicher Relevanzfaktoren wirft einerseits Fragen zur Datensammlung und -aufbereitung auf, zum anderen Fragen zur Operationalisierung von Relevanz in Bezug auf Qualität, Autorität und Popularität. Die Darstellung möglicher Relevanzfaktoren für bibliothekarische Informationssysteme zeigte auch, dass die Faktoren nur in Kombination betrachtet werden sollten und es auf ihre Gewichtung ankommt, damit das Relevanz-Ranking eine sinnvolle Sortierung liefert.

Erfolgreiches Relevanz-Ranking kann nur als ganzheitliches Konzept begriffen werden, in dem die Funktion von Relevanz, Kenntnisse über Nutzerverhalten, die Anwendung statistischer bzw. bibliometrischer Methoden und weitere Faktoren gleichermaßen erforscht werden. Eine Vielzahl an Entscheidungen muss bereits vor der eigentlichen Bestimmung des Relevanzmodells und der tatsächlichen Implementierung der Rankingalgorithmen getroffen werden, z.B. in Bezug auf das Sammeln und Aufbereiten von Download- und Klickdaten (Zeitraum) oder Möglichkeiten zur Integration vorhandener administrativer Kennziffern (z.B. Erwerbungs- und Zugangsdaten).

Nicht alle Problemstellungen und Ansätze können während der Projektlaufzeit differenziert untersucht werden. So schließen sich weitere Forschungsfragen an, z.B. das Thema alternative Metriken (altmetrics) in Bezug auf Popularität im Zusammenhang mit bibliometrischen Methoden.

Nach Ende der Projektlaufzeit wird ein Demonstrator inklusive der getesteten Rankingverfahren für eigene Analysen und Tests zur Verfügung gestellt.