Hochschule für Angewandte Wissenschaften Hamburg
Hamburg University of Applied Sciences

Relevance Ranking

State of the Art in Web Search and Library Catalogs

Technical Report
2015

www.librank.info

Technical Report

Relevance Ranking

State of the Art in Web Search and Library Catalogs

Christiane Behnert

2015

This report covers first results of the research project LibRank funded by the German Research Foundation (DFG - Deutsche Forschungsgemeinschaft). A revised and abridged version of this report is published as:

Behnert, C., & Lewandowski, D. (2015): Ranking Search Results in Library Information Systems — Considering Ranking Approaches Adapted From Web Search Engines. In: The Journal of Academic Librarianship. http://dx.doi.org/10.1016/j.acalib.2015.07.010

Author: Christiane Behnert

Institution: Hamburg University of Applied Sciences (HAW)
Faculty of Design, Media, Information
Department Information
Finkenau 35
22081 Hamburg

E-Mail: christiane.behnert@haw-hamburg.de

HAW Hamburg
Fakultät DMI

www.librank.info

Table of Contents

Introduction

In order to understand the factors behind relevance ranking, this report surveys the state of the art in web search engine ranking factors and how they can be adopted to library information systems. The exemplary search results ranking performed by web search engines can be a useful model for other information systems providers, especially libraries, to emulate. Since people are now used to web search interfaces and relevancy-ranked results lists, they expect searching in library catalogs to be as easy, and the presentation of results to be as good, as when they search the web. This section begins with definitions of the terms *relevance* and *ranking* and describes the differing search behavior and expectations typical of library catalog use.

The reason search results are ranked in an Information Retrieval (IR) system derives from the assumption that information-seeking users should get *all* the information relevant to their search query and *only* that information. In order to help the user judge the relevance of a single search result, the results are presented in a certain way – the most relevant documents are presented first and less relevant documents are presented later. This raises the question: How does the IR system "know" which documents are (most) relevant to satisfy an individual's information needs? A clear definition of the term *relevance* is problematic, and differing views on the meaning of *relevance* can lead to misunderstandings (Bade, 2007; Mizzaro, 1997; Saracevic, 2006).

Although mathematical and statistical methods of varying complexity do exist to determine the relevance of a search result (one rather simple measurement is term frequency, or *TF*), such methods use algorithms to integrate *assumptions of relevance*. But it is the subjective relevance of a result to the user that matters in the end (Bade, 2007), "because an information-retrieval system exists only to serve its users" (Swanson, 1986, p. 390). This concept of subjective relevance can be referred to as pertinence, defined as the user's cognitive ability to understand the knowledge obtainable from a search result (Stock & Stock, 2013, p. 118).

Within these two main classes of relevance – the *objective* or system-based relevance and the *subjective* or user-based relevance (pertinence), five types or manifestations have been identified by Saracevic (1996) and summarized by Borlund (2003) as follows, whereas only the first type refers to objective, logical relevance:

"(1) System or algorithmic relevance, which describes the relation between the query (terms) and the collection of information objects expressed by the retrieved information object(s); (2) a topical-like type, associated with aboutness; (3) pertinence or cognitive relevance, related to the information need as perceived by the user; (4) situational relevance, depending on the task interpretation; and (5) motivational and affective, which is goal-oriented." (Borlund, 2003, p. 914)

As pointed out by Saracevic (2006), "affective relevance is time dependent over all manifestations except algorithmic relevance" (Saracevic, 2006, p. 34). User needs change, their contexts change. Even if the same user has the same query for the same task, a new relevant document may occur within the results set because it had just been published or otherwise become available (Baeza-Yates & Ribeiro-Neto, 2011b, p. 4), or the same user may even find a document that had previously been relevant to not be anymore because the task or information need has changed in the meantime.

Barry & Schamber (1998, p. 227) identified general criteria by which users assess the relevance of information. These criteria are as follows:

1. *Depth/Scope/Specificity*: information is in-depth, provides a summery or interpretation
2. *Accuracy/Validity*: information is accurate, correct or valid
3. *Clarity*: information is presented clearly and easily understood
4. *Currency*: information is recent, up-to-date
5. *Tangibility*: information is proven, hard data are provided
6. *Quality of Sources*: source is reputable, trusted
7. *Accessibility*: information are obtained without effort or costs
8. *Availability of Information/Sources of Information*: sources are available
9. *Verification*: information are consistent or supported by other information within the field
10. *Affectiveness*: information cause emotional response, such as pleasure or entertainment

These relevance criteria are user-based and can be seen as a basis for extracting theoretical relevance ranking factors, but they do not necessarily correspond to the applied technical factors, although there are certain overlaps, for example the criteria 'currency' and 'availability' that are described as ranking factors in section 2.

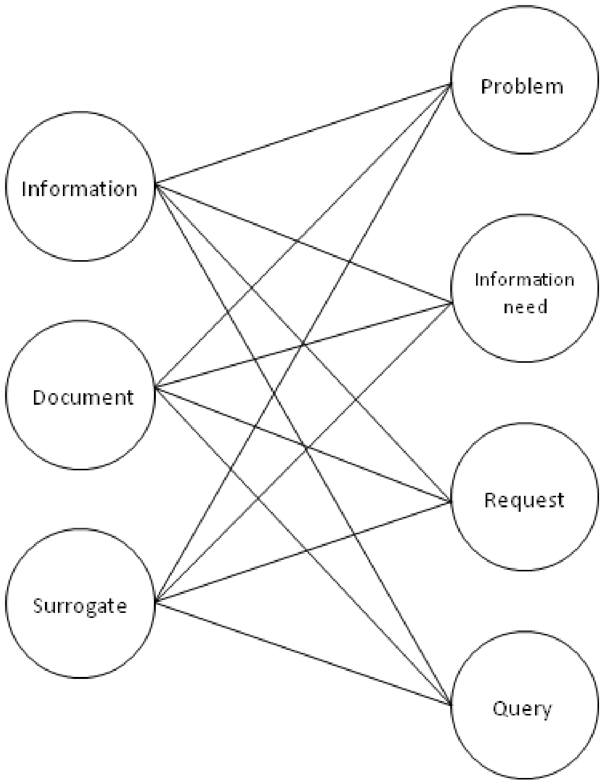

Fig. 1: Relations of relevance.
Modified after Mizzaro (1997, p. 812)

The research on relevance has been discussed in numerous papers reviewed by Mizzaro (1997). He basically defines relevance as "a relation between two entities of two groups" (Mizzaro, 1997, p. 811), as shown in Fig. 1. On the left side, we have the group of results as the output or outcome of the information seeking process. On the right side, we see the group of user-specific entities, as the *trigger for* or the actual *input of* the information seeking process. Relevance as a relational concept may exist between every single entity between these groups, i.e. the relevance between information and information need, between a surrogate and a query and so forth. Thus, relevance plays a central role in the overall information-seeking process and therefore in the process of online searching, as well.

According to Järvelin & Ingwersen (2012), one representative model for user behavior in the context of Web Information Retrieval in particular is Bates' *berrypicking*. Bates sees the information seeking process as an *evolving* search where the user modifies the query at each stage of the search because with every new search results set, new information is obtained and new knowledge is gained. The idea refers to seeking information as picking berries – "[t]he berries are scattered on the bushes; they do not come in bunches. One must pick them one at a time". (Bates, 1989, p. 410)

Several studies have been conducted to analyze search behavior in the context of Web IR (i.e., methods of Information Retrieval in the context of the World Wide Web) and the findings are that queries usually consist only of one or two words, whereas Boolean operators are rarely or only implicitly used (Höchstötter & Koch, 2008). Furthermore, users only look at the first result page and consider mainly the top-ranked hits (Jansen & Spink, 2006; Lewandowski, 2008b; Pan et al., 2007; Schmidt-Mänz & Koch, 2005; Spink, Wolfram, Jansen, & Saracevic, 2001).

Studies also showed that web search often acts as a starting point in the information seeking process (Rowlands et al., 2008, p. 296): Before users start searching in library catalogs, they tend to obtain information of the desired materials via web search, and then carry on searching in the Online Public Access Catalog (OPAC) (Pera, Lund, & Ng, 2009) or the library's website (De Rosa et al., 2005, 2010). Thus, one major implication for library systems is that they "need to look and function more like search engines" (Connaway & Dickey, 2010, p. 5).

When searching the library OPAC, generally the same search and browsing behavior as in search engines could be observed (Hennies & Dressler, 2006): Users consider the top results on the first result page to be the most relevant ones (Antelman, Lynema, & Pace, 2006, p. 135). Queries also usually consist of only few words, i.e. one, two, or three words (Niu & Hemminger, 2010; R. Schneider, 2009). Studies also show that users rely on default settings (Asher, Duke, & Wilson, 2013; S. Jones, Cunningham, McNab, & Boddie, 2000) and, more importantly, that they expect a library catalog to have the same search capabilities and options for displaying results that they are familiar with from web search engines (Yu & Young, 2004).

Academic researchers often use specialized scientific web search engines such as Google Scholar to find journal articles and other sources of information. In the library context, scholarly articles have not been as easily searchable nor have they been directly available (Lewandowski, 2010a). Traditional OPACs with "second-generation" features (e.g. cross-references, exact match Boolean search) (Antelman et al., 2006) still lack a single search interface that allows searching across multiple databases (Luther, 2003), which users expect, having grown accustomed to it when searching on the web. Instead, articles are searchable in separate databases or portals. End users are frequently uncertain which database to choose. Thus, searching on the web and in traditional library OPACs cannot equally satisfy users' information seeking needs. Besides, search results are presented differently (Lewandowski, 2008b). Results presented by library catalogs are bibliographic records, i.e., metadata. We have (1) the metadata of printed and other physically tangible materials, for example books, periodicals, CDs, DVDs, maps, and (2) the metadata of digital contents, for example licensed e-journals and even links to other external content such as audio and video files.

Library materials increasingly comprise more than just printed monographs and journal articles. Now, "web content" such as links to licensed e-journals, e-books, research data and infographics are also included. The characteristics of modern library materials with regard to web content properties are listed in Table 1. Traditional IR techniques alone are insufficient for these types of library content. Because of the change in user behavior when submitting search queries and the expectation that result quality will be indicated by means of a ranking, it is important to implement ranking factors in library information systems inspired by Web IR. Traditional OPACs lack relevance ranking, despite the fact that "[a]lphabetizing makes for easy lookups, but ranking is better for human interest" (White, 2007. p. 600). As a consequence, the integration of search engine technology into library catalogs via discovery software is an essential component of solving OPAC ranking problems (Lewandowski, 2006, 2009, 2010b; K. G. Schneider, 2006).

	Properties of Library Materials
Documents	
Languages	Documents in different languages
File types	Several file types
Document length	Document length varies; one documentary unit exists for each
Document structure	Structured documents
Spam	Suitable document types are defined in the process
Hyperlinks	Documents / entities are connected
Web characteristics	
Amount of data, size of databases	Exact amount of data can be determined when using formal criteria
Coverage	Complete coverage according to the defined sources
Duplicates	Duplicates are singled out
User behavior	
User interests	Both defined but heterogeneous user groups / interests (depending on user context, academic background)
Type of queries	Short queries, users expect Google-like search interface
IR system	
User interface	Easy-to-use interface suitable for end users required
Ranking	Relevance ranking is required because of user expectations
Search functions	Possibility for narrowing search results required (advanced search, facets)

Table 1: Properties of modern library materials
(adapted from Lewandowski 2005, p. 140)

Ranking features have already been implemented in *next-generation catalogs* and *discovery tools*, which enable users to not only find but also access licensed materials. Along with enriched content, faceted navigation and spell-checking, one of the defining features of discovery systems is relevance ranking (Yang & Hofmann, 2011; Yang & Wagner, 2010). Discovery tools such as Serial Solutions' *Summon* or ExLibris' *Primo* provide ranked search result lists using web technology that corresponds more closely to user expectations than traditional catalogs (Breeding, 2006, 2007). Results are ranked based on what the software provider refers to as relevance ranking, which is not necessarily the same term understood by users or librarians. The reason is that ranking factors and algorithms used in commercial discovery software are either undisclosed or not completely transparent because they represent a valuable competitive advantage to their owners (Bade, 2007).

With open source software such as *VuFind* and *Blacklight,* libraries can take things one step further. These applications give libraries control over the technology and the ability to set up their own relevance rankings (Oberhauser, 2010; Parry, 2010). Furthermore, libraries are not seeking profits (unlike commercial library software providers), which may be one reason that the ranking factors and algorithms used are not concealed. This makes it easier for library staff to react to their users' requirements, which may change over time, by altering individual ranking factors themselves. Furthermore, transparency of the applied ranking factors can be seen as one way to maintain informational autonomy (Niedermair, 2014).

Below, we discuss ranking factors used by web search engines and their potential adaptation for use in library information systems. The ranking factors are categorized in six groups, each illustrated with an overview of the group and the individual factors. The first group, *text statistics*, comprises factors which are primarily derived from traditional IR methods. Text statistics include the fundamental ranking factors for all text-based retrieval systems, because there always has to be a query text that can be matched with the documents' representation if any search results at all are to be obtained. Since such ranking factors alone cannot lead to a *quality-induced* ranking, there are other factors building on this first group, as shown in Fig. 2. These factors consider the "wisdom of crowds" and rank results based on a document's *popularity*. It is shown that popularity is hard to measure as it takes into account a lot of information that indicates popularity but cannot provide explicit reasons on a sound theoretical basis. Popularity is rather seen as a way to measure quality or credibility (Lewandowski, 2012). Another group is *freshness.* The up-to-dateness of a document is not only important in Web IR, it is also the standard ranking concept used in traditional library catalogs since their inception. Within the group *locality & availability,* ranking factors consider the physical location of both the user and the document since mobile data connectivity now enables access independent of physical location. Apart from these four major ranking groups, we introduce two others which provide additional valuable information for relevance ranking. The group *content properties* includes characteristics of the document content, while the factors contained within the last group, *user background*, derive from characteristics of the user. In the last section of this report, we summarize the discussed ranking factors and offer a tabular overview of the presented ranking factors in an appendix.

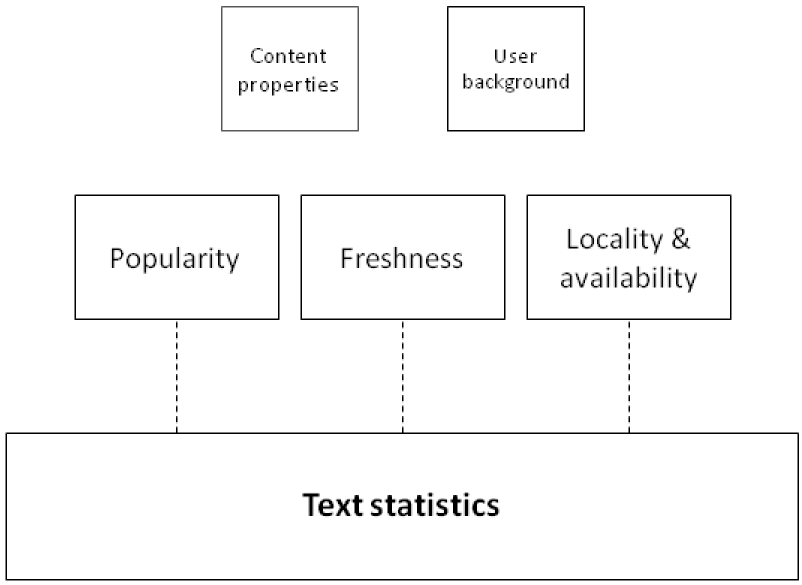

Fig. 2: Overview of ranking factor groups

1 Text statistics

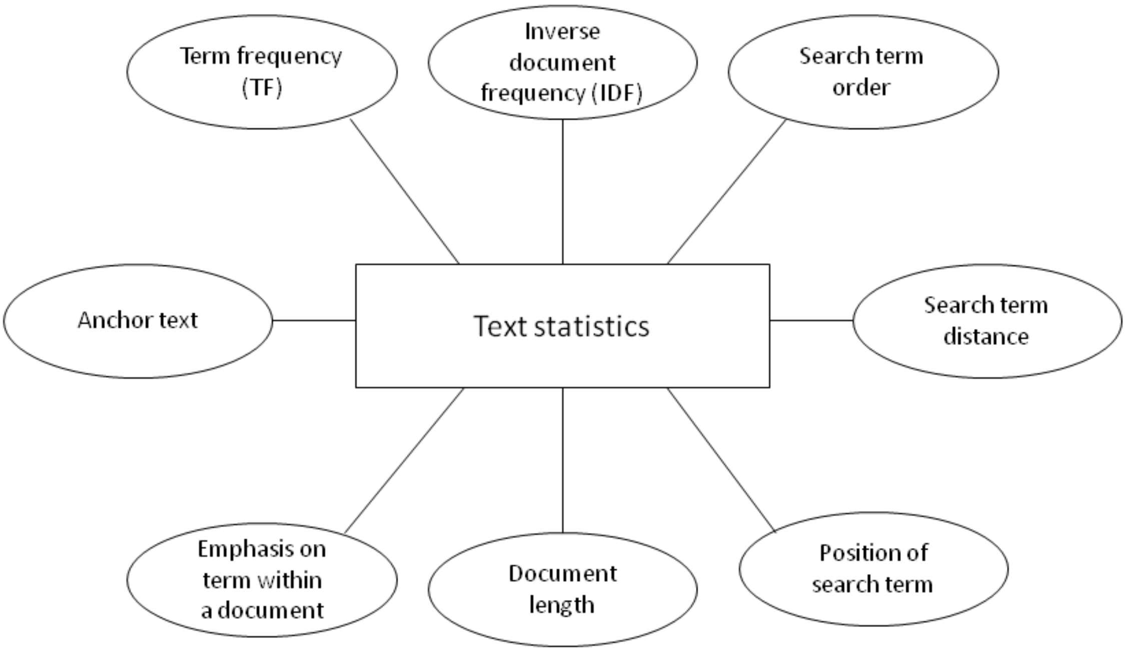

Fig.3: Ranking factors of the group *text statistics*

Here, the process of retrieving documents is called text matching. It involves formulating a query and retrieving the documents whose keywords match the query. Two simple statistical text matching methods for generating a ranking are *term frequency* (TF), which is the relative frequency of a search term within a document, and *inverse document frequency* (IDF), which takes into account the relative frequency of a term in a document collection. The more often a term occurs, the higher its frequency. However, the importance of a term within a document is not indicated exclusively by the frequency with which it occurs. If it were, frequently occurring stop words (*the, a, and*) would be ranked most important. To counteract this effect, terms are weighted. The most popular weighting scheme combines TF and IDF to give less frequent terms a higher weighting. With TF-IDF, “the most common form of vector space weighting” (R. R. Larson, 2012, p. 21), partial text matching is possible within the vector model instead of the exact matching proposed by the Boolean model, which sees a document either as relevant or not relevant to a query. (Baeza-Yates & Ribeiro-Neto, 2011a) Web search engines and discovery systems provide search results based on partial matching or *best match*, but traditional OPACs processed queries solely based on the Boolean model, although it had already been realized in the early 1990s that “users have difficulty searching Boolean OPACs effectively” (Khoo & Wan, 2004, p. 112).

The application of different weighting schemes for different types of queries and tasks can improve precision (Zhitomirsky-Geffet, Feitelson, Frachtenberg, & Wiseman, 2009, p. 533). Nevertheless, TF-IDF weighting is a rather insufficient relevance ranking method. “[Lists] ranked by tf *idf weighting are designed to appeal to people without special claims, people who can make only the easier relevance judgments - students, librarians, readers unfamiliar with a literature, hired judges in information retrieval experiments”. (Bade, 2007, p. 839) In addition, metadata do not provide enough text for applying term frequency in a suitable manner. Traditional ranking of bibliographical records is also based on *position of search terms* (Yang & Hofmann, 2010, p. 143). Consequently, documents with search terms appearing in prominent fields such as the title are weighted higher (Lewandowski, 2005b, p. 143).

Other statistical measures include *search term distance* and *search term order* (Dopichaj, 2009). If a query consists of more than one search term, the documents with the terms closest to each other are more likely to be relevant. A search term at the beginning of the query (order) is also weighted higher. For example, the query “information retrieval” should not prefer “information […] retrieval” or produce documents with the phrase “retrieval information”. The factors search term order and distance should not be neglected, since phrases provide more information content than single words and – as shown above – are often used in a query consisting of only one or a few words (Vechtomova & Karamuftuoglu, 2008, p. 1487).

Furthermore, documents within a certain length span may be preferred, which means their contents should be neither too long nor too short, but rather meaningful (Lewandowski, 2005a, p. 94), for example the size of a website or the number of (printed) pages can act as a ranking factor. Nonetheless, the significance of the *document length* may vary from one research discipline to another, as for example short papers can be found more often in the natural sciences than in the humanities.

In addition, *emphasized text* within a document may be preferred, for example bold or italic terms in title, heading or body text are weighted higher. Relevant terms can also be emphasized using *anchor text*, which is the visible text of a hyperlink (Dopichaj, 2009). Anchor text is considered helpful in web search because it acts as a summary of the target content and usually consists of only a few words, a characteristic it shares with search queries (Eiron & McCurley, 2003). The anchor texts of all documents pointing to a specific document may be seen as an alternative representation of that document.

Library catalogs also provide bibliographic metadata with hyperlinks that refer to other publications, for example to a series title or other works by the same author. Yet text matching, in particular exact matching, is still the foundation of the relevance ranking used in current library catalogs. The problem is that metadata alone do not contain sufficient text, and catalog listings are highly variable with respect to text quantity. A monograph's metadata-only listing and an open-access journal article that includes both metadata and the body copy, for instance, will vary greatly in this regard. This precludes using a single general-purpose ranking algorithm (Lewandowski, 2009; Oberhauser, 2008).

Relevance ranking solely based on statistical measures quickly reaches its limits. As mentioned above, relevance is a subjective matter and differs from one individual to the next. To account for these circumstances, a variety of other factors need to be taken into account. A very popular approach is to determine the demand among users for the respective resource. Such popularity factors are described in the next section.

2 Popularity

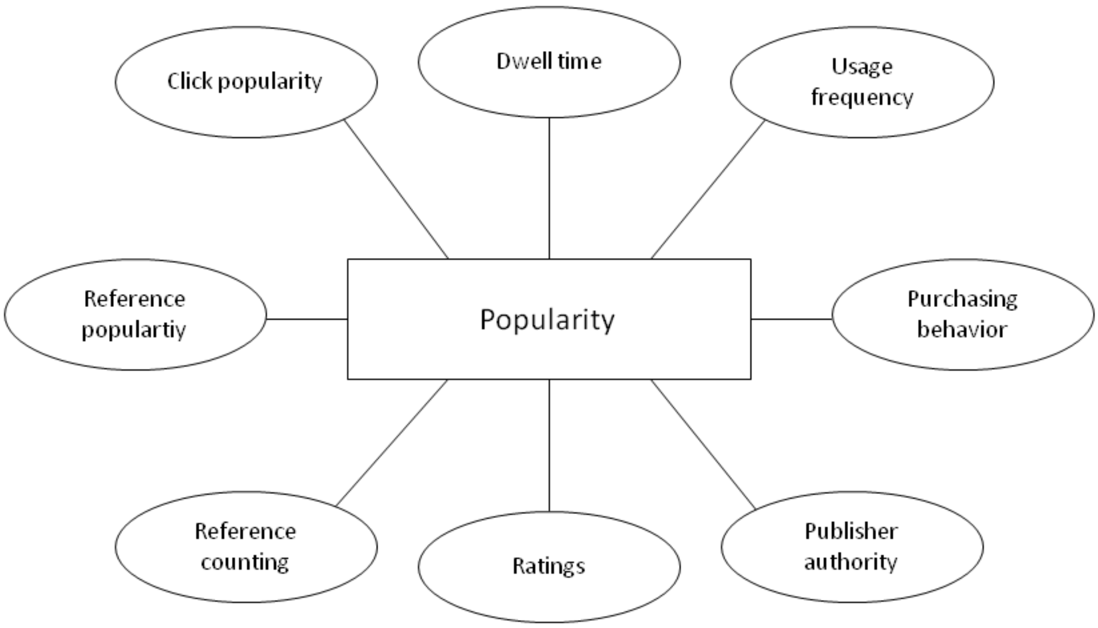

Fig. 4: Ranking factors of the *popularity*

This factor group is based on the "wisdom of crowds" principle, i.e., the knowledge and experiences of many are seen to be more significant than the wisdom of an individual (Surowiecki, 2005). That means, the more people who find a document relevant, the more likely it is to be relevant for an individual user. In this model, popularity indicates quality and therefore relevance. It should be noted that on a theoretical level, this model has many flaws (see Lewandowski, 2012), but on a practical level, it often works quite well.

Popularity for relevance ranking can basically be derived from the number of times a document is used. There is no exact and clear definition of one document use. Use can be defined as the number of clicks on a web page or the actual downloading of a document. But the actual usage of printed library materials is nearly impossible to measure empirically. The factors that imply high usage and therefore popularity are as follows.

Popularity by usage

Click popularity is applicable to digital content and indicates that documents visited by many users must be popular and therefore should be ranked higher. Click data provide implicit relevance feedback, because the individual user signals to the IR system that "more documents like this one" are sought (Jung, Herlocker, & Webster, 2007, p. 791). In web search, click popularity is about the number of clicks a particular web page receives as derived from log data, (Yeadon, 2001). However, clicking decisions cannot provide conclusive feedback because they are influenced by the trust in the way results are displayed (first results receive more clicks) and the quality of the result set (Joachims, Granka, Pan, Hembrooke, & Gay, 2005, p. 161).

In the library context, click frequency is comparable with the number of clicks on electronic books and articles as well as bibliographic records and enriched metadata (e.g. table of contents, abstract, publisher content description). By choosing to look at an abstract, it is certain that the user has an interest in a search result. However, counting the clicks on electronic resources is difficult to transfer to printed copies, as actual reading behavior cannot be measured, at least not in an automatic way. Look-ups could only be measured by observing patrons on-site, i.e. within library facilities, which would be prohibitively expensive. Circulating items are obviously out of reach for such observations. Another point is that the number of clicks only makes sense in conjunction with *dwell time*. It can be assumed that if a user opens a document only to close it again within a second, he or she judges it to be irrelevant. The amount of time a user dwells on a document should be large enough to indicate actual content browsing or intensive reading which would suggest quality (Lewandowski, 2011, p. 63), although the direct dwell time as the only "measure of document preference is likely to fail" (Kelly & Belkin, 2004, p. 383).

Besides click popularity, a strong willingness to actually use, browse, read or further utilize a document can be assumed if it is downloaded. The number of times a full text article or book chapter is downloaded, i.e. the *usage frequency*, is another indicator of its popularity. Libraries in particular collect usage statistics for their electronic resources to calculate metrics such as cost per use.

Initiatives such as COUNTER[1] and SUSHI[2] have helped make usage statistics consistent, credible and comparable, even though they are provided by different vendors (Pesch, 2007).

Usage frequency could also be analyzed for bibliographic records in the library catalog. The willingness to use an electronic document or at least a certain interest in it is indicated by the number of exports from the catalog to reference management software such as Mendeley (Bar-Ilan et al., 2012; Haustein et al., 2014). Such impact data can be derived by applying *altmetrics* (alternative metrics), that are described below in the context of bibliometric methods.

The approach of counting full text downloads is not equally adaptable to printed books and journals. Instead, circulation statistics can be used. Although more frequently circulated books indicate usefulness and therefore popularity, the number of copies available of a book is another way to assess the importance of a work (Yang & Hofmann, 2011, p. 270). Circulation data including circulation rates, periods or even the number of extended or renewed loans cannot indicate usage per se, as they do not cover non-circulating items, such as current journal volumes or rare materials. In addition, the actual usage of loaned items in the private sphere of an individual user is practically unmeasurable. Nevertheless, circulation statistics can instead be used as a source of data for ranking the most popular items, even if these data do exclude non-circulating items, as a log analysis at the North Carolina State University Libraries showed (Antelman et al., 2006, p. 134).

That the generation of circulation statistics does not necessarily depend on the library system's provider was demonstrated by the University Library of Regensburg, Germany. A software tool was implemented to harvest detailed usage statistics for their text book collections. These data were considered to be much more valuable to the subject specialists for their purchasing decisions. (Knüttel & Deinzer, 2013)

Basing acquisition decisions on user demands is one way to adapt to falling circulation rates of printed books (*patron-driven acquisition* or *demand-driven-acquisition*). In this approach, an e-book is purchased by the library only once it has already been selected by the user. (Fischer, Wright, Clatanoff, Barton, & Shreeves, 2012)

Popularity by authority

Another approach to meet user demands is analyzing circulation statistics. Thus, *purchasing behavior* is oriented not only on the collection mandate or approval profile, but on usage frequency as well, i.e. works with a large number of purchased copies can be ranked higher (Yang & Hofmann, 2011, p.

[1] http://www.projectcounter.org
[2] http://www.niso.org/workrooms/sushi/

270). In the library context, acquisition decisions are made by librarians or, in an academic library, by members of the faculty (e.g. professors). Therefore, the quality of a work or document is indicated by their choice of selection (authority). *Purchasing behavior* as a ranking factor can be derived either from the number of local copies in stock and with the distinction between copies within a textbook collection, in open or closed stacks, or on a rather global level with the number of published editions, even taking sales figures from the publisher into account. In addition, the number of libraries owning the particular item, for instance on an international level or within a certain library network, can be interpreted as an indicator of popularity (Maylein & Langenstein, 2013, p. 200).

Purchasing behavior may also be influenced by *publisher authority*. Some examples include publishing houses with good reputations, theses from renowned academic institutions, and well-known working paper series. Reputation in this sense can, for instance, be measured by the number of items bought from a certain publisher by the library. The expert status of reviewers also indicates quality or authority. Papers appearing in peer-reviewed journals can be ranked higher than non-peer-reviewed articles. The authority of a publisher in the web search context can be explained by the size or size range of the website, i.e. documents hosted by larger websites may be preferred because they are assumed to be more authoritative than smaller websites (Lewandowski, 2005b, p. 143).

Apart from the above-mentioned popularity ranking at the North Carolina State University Libraries (Antelman et al., 2006, p. 134), the popularity factors included in the “next generation” E-LIB catalog at the State and University Library of Bremen in Germany are another example of best practice. These factors comprise the number of purchased copies (more than 3 copies indicate increased demand) as well as the number of published editions (indicator of global or international demand) and the click frequency on titles in the search result list. The ranking modifications have had positive effects: Textbooks and articles in strong demand due to searches involving specific topics or popular items can be found more easily as they receive top spots in results lists. (Haake, 2012, 2014)

According to Chickering & Yang (2014), who evaluated and compared 14 major open source and proprietary discovery tools implemented mainly by academic libraries, only Primo by ExLibris considers the number of clicks on bibliographic records for popularity ranking. This leads to the conclusion that in the library context, factors indicating popularity by means of click data have not yet reached their full potential.

Another way of taking user preferences into account is through explicit user-submitted *ratings* or *recommendations*. In contrast to implicit recommendations derived from analyzing user behavior (e.g. clicking, tagging), explicit ratings are directly communicated by the user (Stock & Stock, 2013, p. 416), for example by “liking” as popularized by Facebook, Google’s “+”, or ranking on a scale by awarding stars. Documents with high ratings that have been assessed by many users imply a certain degree of quality and are therefore ranked higher than non-rated or poorly-rated documents. Thus, some of the factors already presented above, such as click rates and the number of tagging or rating users can be applied to relevance ranking (Stock & Stock, 2013, p. 630).

Recommendations provide a readers' perspective on content that may also serve as an indicator for "hot topics" in the context of freshness (Haustein, Golov, Luckanus, Reher, & Terliesner, 2010). Nonetheless, analyzing user preferences via recommendations of course requires the active participation of the user. If the reviewer's identity is visible next to the rating, not every user is willing to provide such information (Stock & Stock, 2013, p. 421).

Ranking factors based on data provided through ratings or recommendations can be used for ranking in library information systems as well. One benefit of "next generation" catalog search interfaces is the integration of user ratings (Vaughan, 2012, p. 38) to "provide the means to help evaluate a given piece" (Breeding, 2007, 35). For example, the recommender tool *bibtip* is used by a number of academic and public libraries in both Germany and other countries. [3] It records user behavior patterns by capturing anonymous session data (Mönnich & Spiering, 2008, para. 3).

The ability to leave comments about a particular work could serve as a means of recommendation as well (Yu & Young, 2004, p. 176). Following this thought, acquisition requests or bibliographies submitted by faculty members for librarians to purchase could also be seen as a list of recommended works.

The idea that highly recommended books awaken the interest of other users has been confirmed in practice by commercial online booksellers such as Amazon. Recommendations need not be restricted to books; see for example star ratings within the music download platform iTunes or the Internet Movie Database (IMDb). Adding external data of this kind to local usage data, for example circulation and acquisition data (Dellit & Boston, 2007, p. 10), can offer a more global perspective. But one limitation may be that libraries prefer not to rely upon proprietary software and may be unwilling to finance the interfaces required to integrate such data into their local catalogs.

Another practice of measuring the popularity of documents is by applying bibliometric methods, which are important for ranking with regard to *reference popularity*.

[3] http://www.bibtip.com/en/references.html [Accessed: 17.07.2014]

Bibliometric methods

Measuring impact based on *reference counting*, also known as straight citation, or link counting *without* considering other aspects influencing impact has been criticized by scientific communities. The factor *reference popularity* ranks documents by the number of incoming links or citations in relation to other documents or entities. But to look at the actual impact, the content quality of the website or document a link points to should be taken into account in addition to the document's connections to other works. Citation counting and citation impact can be considered on three different levels: for the specific journal the article is published in, for the item itself (e.g. article or book section) and for the author. An overview of impact measures with regard to their historical development, points of criticism and suggested alternatives is provided by Smith (2012). Here is a brief selection of general approaches:

The idea of a Journal Impact Factor (JIF) initially proposed by Garfield (1955) was applied to ranking journals by the frequency and impact of citations using the Science Citation Index (SCI) over forty years ago (Garfield, 1972). The JIF aims to determine the reputation of a journal by measuring the average number of citations per article published by the journal over the previous two years. Although it considers journals with a rather small number of published articles that are nonetheless very influential in their fields (Garfield, 2006), comparability between different research disciplines is not guaranteed, as different citation or publishing conventions by researchers across different disciplines are not considered. Thus, articles and journals should be ranked based on their JIF in relation to the respective field. For example, the query "statistical methods" can relate to information retrieval models as well as the social sciences and the information sought may be obtained from articles in journals of different research areas. Keeping in mind differences of citing or publishing habits depending on the research discipline may allow for better relevance judgments without focusing solely on the JIF.

Another approach for measuring the impact of a journal is stated by the *Eigenfactor*. It determines how central a journal is within a journal citation network based on the number of citations a journal receives. The more connected a journal, the more central it is. In contrast to the JIF, the Eigenfactor score does not include self-citations. Citations are counted across a five year window instead of a two year window. (West, Bergstrom, & Bergstrom, 2010)

Ranking journal articles based on the journal's impact means ranking by *journal reputation*. To determine the impact of the *individual researcher*, the *h*-index (Hirsch 2005) can be applied, among others. It is a simple means of measuring citations based on the number of papers and citations: For example, one author has the index $h = 15$, if he or she has a minimum of 15 published papers that have received at least 15 citations in other publications. One point of criticism is that authors of papers published without co-authors do not get extra credit, making it difficult to compare individual research outcomes. Considering these circumstances, Hirsch (2010) presented an advanced h-index, *hbar*, which takes co-authorship into account. With the exception of a few modifications, the h-index can be seen as an improvement over other straight citation counts. (Cronin & Meho, 2006) Nevertheless, differences in fields of research are not completely reflected either. For example "[e]conomists write fewer papers than do physicists, and individual papers get many citations. As a result, the [h-index] starts to look uncomfortably like a publication count." (Ellison, 2010, p. 2)

One bibliometric method used for (re-)ranking documents of a search result set is *Bradfordizing,* which is an application of Bradford's law of scattering (Bradford, 1934). Results are divided into three zones based on the source journals, each zone consisting of one third of all articles on a topic. The idea of Bradfordizing is that documents are ranked based on the core journals, i.e. documents that are part of the core zone are ranked higher than documents of the second or third zone. (Mayr, 2011, 2013) "This re-ranking method is interesting because it is a robust and quick way of sorting the central publication sources for any query to the top positions of a result set." (Mayr, 2013, p.3) Another idea to re-rank documents by using the Bradfordizing technique would be to invert this method. Journals that contain only a few articles on a specific topic could be interesting for the expert user because it can be assumed that experts know the papers from the core journals relevant to their field of research anyway. On the other hand, this might make it easy to miss relevant articles not published in the usual sources. This approach again illustrates the importance of the user's context within the information seeking process.

Measuring the web impact of organizations is not only possible by counting the links to their websites, but also by counting the number of their titles mentioned on a web page and the number of URL citations as well. Since the "numbers of links pointing to a research website are not a good indicator of researcher productivity" (Thelwall & Sud, 2011, p. 1490), measuring impact based on citation counting or link counts *without* considering other aspects influencing impact has been criticized by scientific communities. An overview of limitations of impact factors regarding long-term scientific influence is given by Wang, Song, & Barabási (2013) and includes the phenomenon of established journals being favored and the low degree of comparability of publications across disciplines.

An alternative approach to focus on individual researchers is provided by *altmetrics*, which aims to measure academic influence by tracking the use of social media tools (Priem, Piwowar, & Hemminger, 2012; Priem, Taraborelli, Groth, & Neylon, 2011). This can be achieved, for example by deriving data from web-based social reference management software like Mendeley or performing social media analysis on weblogs (e.g. on the micro blogging platform Twitter) or from other new scholarly communication tools (Bar-Ilan et al., 2013). Prerequisites for such measures are integrated bookmarking and tagging tools as well as high-quality metadata (Haustein & Siebenlist, 2011). One advantage of altmetrics is that the *reader-specific* view on documents is reflected by the number of bookmarks, in contrast to the *author-specific* view offered by citation counts (Haustein et al., 2010). Search results would then be ranked based on the author's social profile.

Link-based ranking

Ranking documents based on citation analysis can be seen as the major underlying concept for ranking web pages. Although a majority of web links are actually links for navigational and other purposes instead of research links (A. G. Smith, 2004), link-based ranking aims to solve "one of the key problems of Web IR […]" which "requires some kind of relevance estimation. In this context, the number of hyperlinks that point to a page provides a measure of its popularity and quality." (Baeza-Yates & Maarek, 2011, p. 470) Nevertheless, the sole number of links is an insufficient measure of quality, as noted above in the context of citation counts without relation to other entities, e.g. other works by the particular author.

The structure of the web can be described as a graph, in which individual web pages are represented by nodes, and links between web pages are the edges (Baeza-Yates & Maarek, 2011, p. 452). The quality of web pages is assessed by the links pointing to an individual page within the web graph. PageRank, proposed by Page, Brin, Motwani, & Winograd (1998) as part of the ranking algorithm used by Google, is one well-known model that works this way.

PageRank forecasts the probability that a user who is randomly following one link after another will visit a website (*random surfer model*). The more backlinks (hyperlinks from other web pages) a document has and the higher the backlinks' PageRank, the higher the PageRank of the document itself. This simplified calculation is illustrated in Fig. 5: For example, a page is assigned a PageRank of 53 based on two other sites, ranked 100 and 9 respectively, which link to it. The PageRank-100 site has two outbound links (100 / 2 = 50), and the PageRank-9 site has three (9 / 3 = 3). Following the idea of citation analysis in an academic context, the importance or popularity of a web page is indicated by the number of backlinks from pages with a high number of backlinks and so forth, but in correlation with the content quality the links point to, i.e. different weights are allocated to different links.

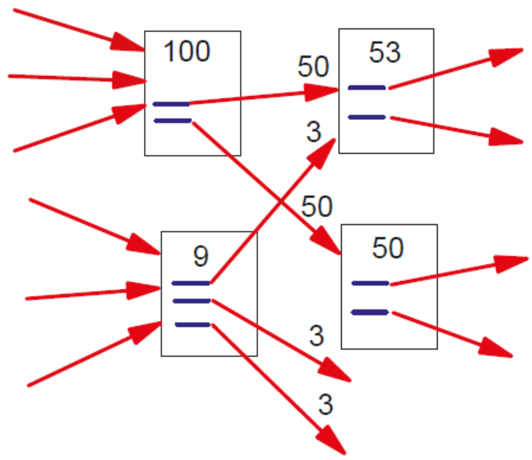

Fig. 5: Simplified PageRank Calculation (Page et al., 1998, p. 4)

This section has shown several methods for measuring popularity to generate relevance rankings. It is important to recognize that these measures cannot indicate absolute popularity. However, they can provide useful information that allows us to determine popularity within a complex system. Furthermore, measurements for electronic resources cannot equally be applied to printed materials in the library context, see for instance the incomparability of download statistics and circulation data. This presents a special challenge, and apart from possible algorithms normalizing for these effects, the distinction must be clearly communicated to the user.

3 Freshness

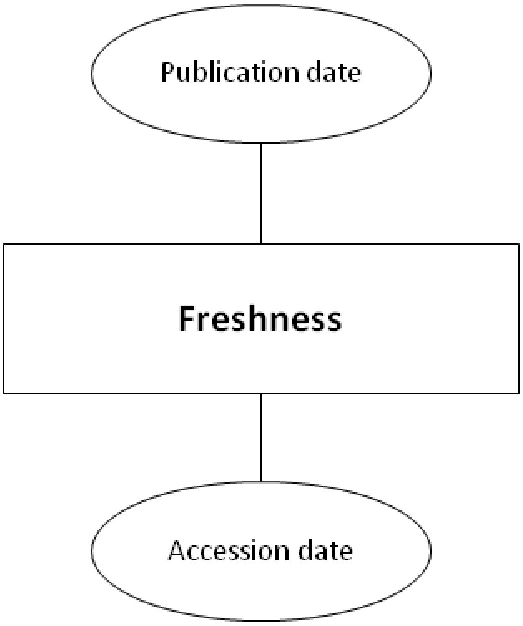

Fig. 6: Ranking factors of the group *freshness*

Freshness (sometimes also called up-to-dateness) is a very important factor in the context of relevance ranking. It can be assumed, that users in general seek *current* information, especially for academic research. Freshness is one major indicator of the overall quality of a web search engine (Lewandowski & Höchstötter, 2008). The ability to produce current search results depends on the update frequency of the web index used. For economic and technical reasons, it seems practically impossible to update the (main) index for every single website on a daily basis. (Lewandowski, 2008a) Instead, crawling frequency depends on factors such as the size and popularity of a website or its past update frequency (Lewandowski, Wahlig, & Meyer-Bautor, 2006).

Thus, an important part of the ranking methods is not only popularity (see *reference popularity*), but immediacy as well: Documents with a recent *publication or accession date* may be preferred and ranked higher. An overview of measurements for scoring documents as part of relevance ranking is presented by Acharya et al. (2005). They describe types of history data that can be used for ranking, which includes the inception date as well as content updates or changes and other metrics that correlate with document usage. At this point, it is already foreseeable, that combinations of different factors (of different groups) play an essential role in relevance ranking.

Although the same freshness factors used by web search engines can and should be applied to library materials, there is an issue regarding the need for freshness. Traditional catalogs commonly sort their results by publication year, or in alphabetic order by author or title (Oberhauser, 2010, p. 30). But in an academic context, current resources are not always the only relevant ones. The importance of freshness is determined in part by the nature of the respective academic discipline. In the sciences, for instance, usually the most recent paper and freshest results are sought. In the humanities, a seminal article published decades ago (e.g. historical sources) can remain or re-emerge as a “hot topic” (Chen, Luesukprasert, & Chou, 2007). On the other hand, the subjective view of relevance to the user counts most of all in the end. As mentioned above, relevance may change after a period of time. “Hot topics” can be indicated, for instance, by programs or proceedings of recent or future (international) conferences, where the results of recently finished or even ongoing studies are presented. Tags attached to journal articles can also be a real-time indicator of “hot topics”, as Haustein et al. (2010) concluded in their proposal for applying social bookmarking data to journal evaluation. When searching for current events, new products or neologisms, tags can also be a useful tool, as they reflect user behavior in a flexible way (Peters, 2011, p. 43).

Library catalogs enable the distinction between scientific disciplines and research topics by assigning subject headings or using another form of classification, which is usually part of a bibliographic record. This type of data should thus be available for further measurements. Additionally, the accession data including date of licensing for electronic books or journals are part of the technical metadata stored in the system. Another useful indicator of freshness could be the publication or accession date of enriched metadata, for example of a book review.

Nonetheless, library systems that utilize discovery software should continue integrating the year of publication as a ranking factor based on the document level, but the need for freshness (based on the query level) should be considered as well. More importantly, this factor should not be the major one (Maylein & Langenstein, 2013, p. 196), as the combination of both freshness and popularity influences ranking significantly. These factors have already been effectively used by web search engines and, even if only partially, incorporated into library information systems. Especially in the library context, the enormous potential of implementing popularity factors inspired by web search has been recognized.

4 Locality and availability

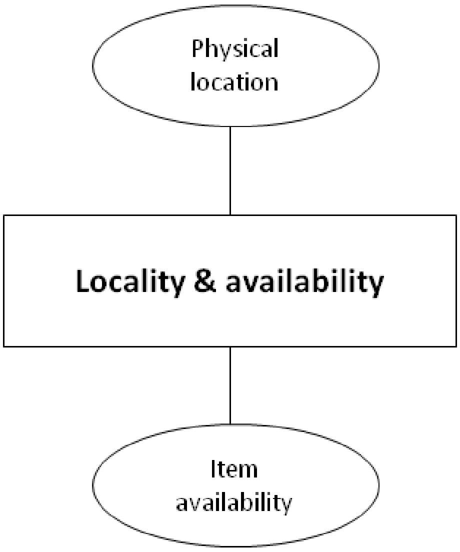

Fig. 7: Ranking factors of the group *locality and availability*

Next to popularity and freshness, another factor for ranking search results is the locality or availability on both the item and the user level. These factors can be applied to electronic resources and printed copies in an analogous fashion. Taking locality into account can be a major advantage for ranking algorithms, as geographic data provide contextual information that is useful in determining the actual information need (Baeza-Yates, Broder, & Maarek, 2011, p. 5). The *physical location* of the user or the item influences the search engine's query interpretation. Thus, web pages that are "closer" to the location of the user would be preferred (for a detailed discussion on measuring distance for ranking purposes, see R. Jones, Zhang, Rey, Jhala, & Stipp, 2008). For example, the query "nearest pizza" assumes that the user is seeking information on the nearest pizza restaurant. A user in London would not expect a pizza shop in Rome to be among the search results (regardless of its popularity score). Location information is also used for personalized ranking (Baeza-Yates & Maarek, 2012), which will be covered in section 6.

For documents to provide geographic information, they can include metadata such as longitude, latitude, region, type (e.g. city, lake), spatial relationship (e.g. a region name *x* in a query is contained in the country *y* of or equal to the document) (Larson & Frontiera, 2004). Nonetheless, the geographic intent of the user determined by the query is associated with identifying synonyms and disambiguating location names (e.g. "San Jose" that can be related to more than 900 places world-wide) due to reformulating the query (R. Jones, Zhang, Rey, Jhala, & Stipp, 2008, p. 3). Another way of determining location is by interpreting the user's selection behavior as an implicit indication.

In an academic search context, the query 'employment statistics' would return statistics of the user's region or country, and the query term 'conference on topic *t*' would prefer websites with upcoming events at the nearest locations.

Geographical relevance is a factor that should not be ignored, nor should it be weighted too highly, as the "best performance is achieved when the importance of non-geographical relevance outweighs the importance of geographical relevance by a large factor" (Andogah, Bouma, & Nerbonne, 2012, p. 18). It is important, though not trivial, to balance between popularity and locality factors (Baeza-Yates & Maarek, 2012, p. 502). One exemplary method for location-enhanced ranking for web pages is presented by Zhao, Jin, Zhang, & Wen (2014). They developed an algorithm called *MapRank*, which considers both textual (keywords) and location relevance of a web page, but excludes popularity factors.

Geographic search should be integrated into library information systems as well. One of the prerequisites for storing such data successfully in the system is fulfilled by using standardized formats such as the MARC21 catalog format (Buckland et al., 2007). Libraries must however face such challenges to provide geographic data with regard to their users' expectations in searching (Abresch, Hanson, Heron, & Reehling, 2008).

Both acquisition and circulation data indicate not only "hot topics," but also provide useful information on current location or availability. Even if data on printed materials are not taken into account, the availability of electronic resources would still cover a large part of the library holdings (Maylein & Langenstein, 2013, p. 200). A document that is physically unavailable may be less relevant to a user located within the library building, for instance when every copy of the desired book is circulating and the information is needed right away. An available copy or different licensed journal article is likely to be more relevant in this case, since it would satisfy the user's information need immediately. Conversely, situations in which only electronically available resources are relevant may arise, for instance when users are located outside of the library facility. For example, a student who is home for semester break but needs to complete a paper for a seminar may not be able to visit the library in person. Works that the user can actually access, in this case electronic resources, are the only relevant documents and would therefore be preferred by the ranking mechanism.

On the item level, deriving the necessary data to determine availability would be accompanied by the integration of circulation and acquisition data for popularity ranking factors. On the user level, the "location" information can be obtained via authentication or IP address.

5 Content properties

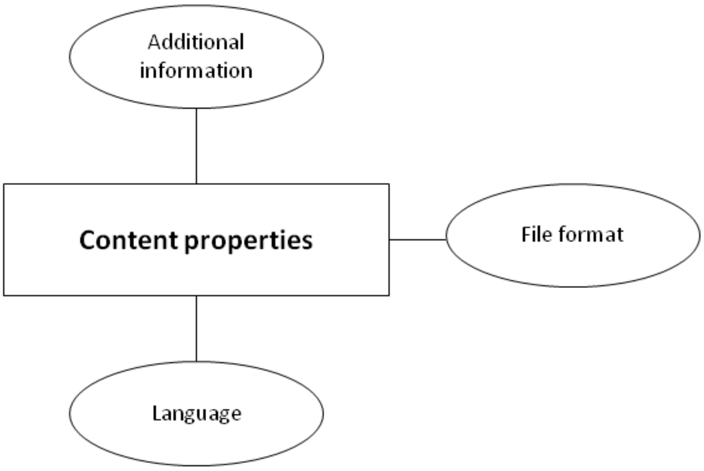

Fig. 8: Ranking factors of the group *content properties*

This group contains factors that refer to the formal properties of a document's content. One property is the availability of *additional information*, i.e. documents with additional content are weighted higher. Website information may include metatags, such as keywords, whereas table of contents, annual indexes, or reviews would be part of enriched metadata for bibliographic records. Such data indicate reliability due to a high degree of indexing, whereas abstracts provide a summarized content overview of an academic article allowing an immediate relevance judgment. They are popular among scholars, according to a study conducted by Nicholas, Huntington & Jamali (2007).

Another indicator for relevance can be the availability of underlying research data. Research data offer evidence of the transparency of the applied scientific methods and allow reproducibility of the study results. Some publishers have integrated the availability of underlying research data in their publishing policies, as a prerequisite for publishing the paper at all.

Ranking documents based on their *file format* illustrates the distinction between web pages and library materials. In the context of web search engines, HTML is preferred over PDF or DOC formats, "because the user can see these files in his browser without opening another program or plug-in" (Lewandowski, 2005b, p. 144). Although HTML is the standard format for webpages, PDF documents are increasing in number, especially in academics. One of PDF's advantages is that the document can be cited, which is necessary when publishing research. In addition, PDF would be the preferred file type for electronic full text, made available through the library's licensed journals or full text databases. With regard to printed copies, the format can relate to the circulation status, as large format items (e.g. maps) are may not be suitable for circulation.

Tools for managing references, such as Mendeley, Citavi or Endnote, support different formats (e.g. RIS, BibTex, RefWorks), but this factor can be seen as negligible because the possibility of not being able to download a citation seems rather small.

The *language* of a document represents a factor that, in combination with the language of both the query and the search interface, can improve the precision of the IR system. As Leveling, Ghorab, Magdy, Jones, & Wade (2010) showed in an experiment with log data of *The European Library*, a large percentage of queries are submitted in English although the amount of users coming from English-speaking countries is comparatively small. We can therefore assume that a user located in Germany submitting queries in German via a German search interface most likely expects search results in German, which is of particular relevance for terms occurring in multiple languages such as "computer". Thus, these search results should be ranked higher than results in other languages because those would probably not be relevant at all. A preferred language can in some cases be set in personal profiles, for example in the user account settings for the library catalog.

Although these content properties are rather formal ranking factors, the combination of different factors in differently weighted scores should be taken into account. Since the language of search results should be the choice of the user (Baeza-Yates & Maarek, 2012), such content information might influence relevance judgment as well.

6 User background

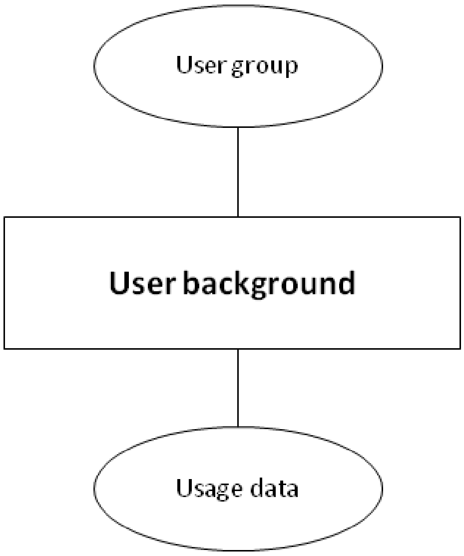

Fig. 9: Ranking factors of the group *user background*

Besides the information need, knowledge of the particular user background should be considered in making the relevance judgment. As noted above in the context of popularity factors, analyzing usage data is based on the “wisdom of crowds” principle. Documents relevant to a particular *user group* are also relevant to an individual user with the same user background. Ranking based on this idea is described as personalized ranking. (Lewandowski, 2005a, p. 138; Riemer & Brüggemann, 2009)

Libraries provide access to information with the objective of satisfying their users’ information needs. In order to achieve this goal, it is beneficial to know who the library’s users are and what (academic) backgrounds they possess. One of the many user studies describes the user range at an academic library to be “from digital native students […] to middle-aged researchers” (Pianos, 2010, p. 5) who show different levels of search skills. Data on the different user groups and their academic status can be derived from their library ID. This requires some kind of authentication or login into one’s user account before starting to search the catalog. An interesting approach using this type of data has been taken by the Heidelberg University Library in Germany, which weights documents within the user’s specific field of research or study higher (Maylein & Langenstein, 2013, p. 196).

In contrast with digital libraries, web search engines basically deal with a rather heterogeneous user base that consists of untrained end users (Lewandowski, 2005a, p. 34) as well as experienced users (e.g. with an academic background) and trained information professionals. The search behavior of end users differs from that of the latter group, as noted above (see section *Introduction*). This knowledge has been gained over the course of multiple user studies which, for example, analyzed click frequency or clickstream data from the user's browsing activity.

Click-through data provide information to construct a "user profile, which stores the user's interests and preferences" (de Campos, Fernández-Luna, Huete, & Vicente-López, 2013, p. 176) for personalized ranking based on individual user behavior.

The idea of click-through data for personalized ranking as mentioned by Joachims (2002) later lead to the conclusion that such data convey reasonably accurate information about user preferences (Radlinski & Joachims, 2005, p. 243). Furthermore, data on post-search clicks can be used to determine the intentions of users who submit search queries (Chapelle et al., 2011, p. 587).

With regard to click frequency, the user's dwell time on a web page or document can be seen as an indicator of his or her interest, although this cannot be put on the same level as actual preference. Following this approach for personalized re-ranking, Xu, Jiang, & Lau (2011) developed an algorithm that can predict a user's potential dwell time based on his or her interest presumed by the dwell time measured on previously read documents at the concept word level.

A user profile can be constructed by counting the number of web pages the user browsed in conjunction with term frequency. In an experimental study conducted by Sugiyama, Hatano, & Yoshikawa (2004) based on term weights within a website a user visited, the constructed user profile achieved the best retrieval accuracy (Sugiyama et al., 2004, p. 683). This approach was also taken by de Campos et al. (2013), who developed, combined and evaluated four new search personalization strategies. The results showed an improvement in performance with personalizing strategies, as they consider the user information needs in a more suitable way (de Campos et al., 2013, p. 876).

Browsing behavior was assessed to be more revealing than query behavior, because a higher volume of browsing data is generated, representing a more robust data source (Bilenko, White, Richardson, & Murray, 2008). Nonetheless, Kashyap, Amini, & Hristidis (2012) adopted a contrasting approach that involves analyzing social links and the query history of a user for personalized web search results.

Personalized ranking on the basis of analyzing usage data must take privacy and data protection into consideration, i.e. gaining informed consent from users is obligatory (Baeza-Yates & Maarek, 2012). For personalized web search, this can be achieved, for example, through implicitly or explicitly accepting browser cookies or the terms of use that a tool or application demands. In the library context, the authentication process can be used to obtain informed consent from users. This kind of click-through data include clicks on bibliographic records, abstracts, table of contents or full text, to name a few examples. The central tenets of data protection require that users be informed of the type of data being collected and how personalized ranking is implemented.

7 Conclusion

In this article, we have provided an overview of ranking factors for use in web search engines and the ways in which they can be adopted to library information systems. The first factor group, *text statistics*, describes basic statistical measures that form the foundation of relevance ranking, but are not sufficient to successfully implement relevance ranking. For relevance ranking to be efficient, it must take other factors into account as well. *Popularity* factors represent a rather complex approach. We illustrated how complex the respective measures for popularity can be, starting with the definition of usage and going on to illustrate the comparability problem of usage statistics and academic impact based on citation analysis. The combination of several different factors has a major role in relevance ranking. In conjunction with the factors *freshness* and *locality and availability*, relevance can be indicated on different levels, taking into account not only the *content properties* (item level), but the *user background* (user level) as well.

It can be assumed that basic ranking algorithms and methods will not change drastically in the near future of web search. The ability of search engines to better understand search queries and user intent via semantic components will also influence relevance ranking. Several approaches towards semantic search and ranking issues have been illustrated (see for example Agrawal, Sharma, Kumar, Parshav, & Goudar, 2013; Jindal, Bawa, & Batra, 2014; Shepherd, 2007). In addition, the combination of natural language processing and artificial intelligence may replace conventional keyword searching over the long term. We have demonstrated that libraries are increasingly integrating search engine technology in their catalogs — how more advanced technologies can be successfully adapted as well remains to be seen.

References

Abresch, J., Hanson, A., Heron, S. J., & Reehling, P. J. (2008). *Integrating geographic information systems into library services : A guide for academic libraries* (p. 318). Hershey, PA [u.a.]: Information Science Publ.

Acharya, A., Cutts, M., Dean, J., Haahr, P., Henzinger, M., Hoelzle, U., … Tong, S. (2005). *Information Retrieval based on historical data*. USA: United States Patent ; US 2005/0071741 A1.

Agrawal, J., Sharma, N., Kumar, P., Parshav, V., & Goudar, R. H. (2013). Ranking of searched documents using semantic technology. *Procedia Engineering*, *64*, 1–7. doi:10.1016/j.proeng.2013.09.070

Andogah, G., Bouma, G., & Nerbonne, J. (2012). Every document has a geographical scope. *Data & Knowledge Engineering*, *81-82*, 1–20. doi:10.1016/j.datak.2012.07.002

Antelman, K., Lynema, E., & Pace, A. K. (2006). Toward a twenty-first century library catalogue. *Information Technology & Libraries*, *25*(3), 128–139.

Asher, A. D., Duke, L. M., & Wilson, S. (2013). Paths of discovery: Comparing the search effectiveness of EBSCO Discovery Service, Summon, Google Scholar, and conventional library resources. *College & Research Libraries*, *74*(5), 464–488.

Bade, D. (2007). Relevance ranking is not relevance ranking or, when the user is not the user, the search results are not search results. *Online Information Review*, *31*(6), 831–844. doi:10.1108/14684520710841793

Baeza-Yates, R., Broder, A. Z., & Maarek, Y. (2011). The new frontier of web search technology : Seven challenges. In S. Ceri & M. Brambilla (Eds.), *Search Computing II* (pp. 3–9). Berlin ; Heidelberg: Springer.

Baeza-Yates, R., & Maarek, Y. (2011). Web retrieval. In *Modern Information Retrieval: the concepts and technology behind search* (2. ed., pp. 445–514). Harlow [u.a.]: Addison-Wesley/Pearson.

Baeza-Yates, R., & Maarek, Y. (2012). Usage data in web search: Benefits and limitations. In A. Ailamaki & S. Bowers (Eds.), *Scientific and Statistical Database Management SE - 33* (Vol. 7338, pp. 495–506). Berlin ; Heidelberg: Springer. doi:10.1007/978-3-642-31235-9_33

Baeza-Yates, R., & Ribeiro-Neto, B. (2011a). Modeling. In *Modern Information Retrieval: the concepts and technology behind search* (2. ed., pp. 57–130). Pearson Education.

Baeza-Yates, R., & Ribeiro-Neto, B. (2011b). *Modern information retrieval : the concepts and technology behind search* (2. ed.). Harlow [u.a.]: Addison-Wesley/Pearson.

Bar-Ilan, J., Haustein, S., Peters, I., Priem, J., Shema, H., & Terliesner, J. (2012). Beyond citations: Scholars' visibility on the social web. In É. Archambault, Y. Gingras, & V. Larivière (Eds.), *Proceedings of 17th International Conference on Science and Technology Indicators* (pp. 98–109). Montréal, CAN: Science-Metrix and OST. Retrieved from http://2012.sticonference.org/Proceedings/vol1/Bar-Ilan_Beyond_98.pdf

Bar-Ilan, J., Sugimoto, C., Gunn, W., Haustein, S., Konkiel, S., Larivière, V., & Lin, J. (2013). Altmetrics : Present and future – Panel. In *ASIST*. Montréal, CAN.

Barry, C. L., & Schamber, L. (1998). Users' criteria for relevance evaluation: A cross-situational comparison. *Information Processing & Management*, *34*(2-3), 219–236. doi:10.1016/S0306-4573(97)00078-2

Bates, M. J. (1989). The design of browsing and berrypicking techniques for the online search interface. *Online Information Review*, *13*(5), 407–424. doi:10.1108/eb024320

Bilenko, M., White, R. W., Richardson, M., & Murray, G. C. (2008). Talking the talk vs. walking the walk : Salience of information needs in querying vs. browsing. In *SIGIR´08*. Singapore.

Borlund, P. (2003). The concept of relevance in IR. *Journal of the American Society for Information Science and Technology*, *54*, 913–925. doi:10.1002/asi.10286

Bradford, S. C. (1934). Sources of information on specific subjects. *Engineering*, *137*(3550), 85–86.

Breeding, M. (2006). Technology for the Next Generation. *Computers in Libraries*, *26*(10), 28–30.

Breeding, M. (2007). The birth of a new generation of library interfaces. *Computers in Libraries*, *27*(9), 34–37.

Buckland, M., Chen, A., Gey, F. C., Larson, R. R., Mostern, R., & Petras, V. (2007). Geographic search: Catalogs, gazetteers, and maps. *College & Research Libraries*, *68*(5), 376–387.

Chapelle, O., Ji, S., Liao, C., Velipasaoglu, E., Lai, L., & Wu, S.-L. (2011). Intent-based diversification of web search results: metrics and algorithms. *Information Retrieval*, *14*(6), 572–592. doi:10.1007/s10791-011-9167-7

Chen, K.-Y., Luesukprasert, L., & Chou, S. T. (2007). Hot topic extraction based on timeline analysis and multidimensional sentence modeling. *IEEE Transactions on Knowledge and Data Engineering*, *19*(8), 1016–1025. doi:10.1109/TKDE.2007.1040

Chickering, F. W., & Yang, S. Q. (2014). Evaluation and comparison of discovery tools: An update. *Information Technology & Libraries*, *33*(2), 5–30. doi:10.6017/ital.v33i2.3471

Connaway, L. S., & Dickey, T. J. (2010). The digital information seeker: Report of findings from selected OCLC, RIN and JISC user behaviour projects. OCLC Research. Retrieved from http://www.jisc.ac.uk/media/documents/publications/reports/2010/digitalinformationseekerreport.pdf

Cronin, B., & Meho, L. (2006). Using the h-index to rank influential information scientistss. *Journal of the American Society for Information Science and Technology*, *57*(9), 1275–1278. doi:10.1002/asi.20354

De Campos, L. M., Fernández-Luna, J. M., Huete, J. F., & Vicente-López, E. (2013). XML search personalization strategies using query expansion, reranking and a search engine modification. In *Proceedings of the 28th Annual ACM Symposium on Applied Computing - SAC '13* (pp. 872–877). New York: ACM Press. doi:10.1145/2480362.2480531

De Rosa, C., Cantrell, J., Cellentani, D., Hawk, J., Jenkins, L., & Wilson, A. (2005). Perceptions of libraries and information resources: A report to the OCLC membership. Dublin, OH: OCLC Online Computer Library Center.

De Rosa, C., Centrell, J., Carlson, M., Gallagher, P., Hawk, J., & Sturtz, C. (2010). Perceptions of libraries: Context and community ; a report to the OCLC membership. Dublin, OH: OCLC Online Computer Library Center.

Dellit, A., & Boston, T. (2007). Relevance ranking of results from MARC-based catalogues: from guidelines to implementation exploiting structured metadata. *National Library of Australia Staff Papers*, 1–14. Retrieved from http://www.nla.gov.au/openpublish/index.php/nlasp/article/view/1052/1321

Dopichaj, P. (2009). Ranking-Verfahren für Web-Suchmaschinen. In D. Lewandowski (Ed.), *Handbuch Internet-Suchmaschinen* (pp. 101–115). Heidelberg: AKA Verlag.

Eiron, N., & McCurley, K. S. (2003). Analysis of anchor text for web search. In *Proceedings of the 26th annual international ACM SIGIR conference on Research and development in informaion retrieval - SIGIR '03*. New York: ACM Press. doi:10.1145/860500.860550

Ellison, G. (2010). *How does the market use citation data? The Hirsch index in economics*. *NBER WORKING PAPER SERIES*. Cambridge, MA: NBER Working Paper Series ; 16419.

Fischer, K. S., Wright, M., Clatanoff, K., Barton, H., & Shreeves, E. (2012). Give 'em what they want: A one-year study of unmediated patron-driven acquisition of e-books. *College & Research Libraries*, *73*(5), 469–492.

Garfield, E. (1955). Citation indexes for science: A new dimension in documentation through association of ideas. *Science*, *122*(3159), 108–111. doi:10.1126/science.122.3159.108

Garfield, E. (1972). Citation analysis as a tool in journal evaluation: Journals can be ranked by frequency and impact of citations for science policy studies. *Science*, *178*(4060), 471–479. doi:10.1126/science.178.4060.471

Garfield, E. (2006). The history and meaning of the journal impact factor. *JAMA : The Journal of the American Medical Association*, *295*(1), 90–93. doi:10.1001/jama.295.1.90

Haake, E. (2012). Relevanzranking als Erfolgsfaktor für Discoverysysteme. In *16. GBV-Verbundkonferenz*. Hannover.

Haake, E. (2014). "Nur die ersten drei zählen!" Erfolgsfaktoren beim Betrieb von Discovery Systemen / Neues Relevanz Ranking. In *103. Deutscher Bibliothekartag*. Bremen.

Haustein, S., Golov, E., Luckanus, K., Reher, S., & Terliesner, J. (2010). Journal evaluation and Science 2.0: Using social bookmarks to analyze reader perception. In *Eleventh International Conference on Science and Technology Indicators; Book of Abstracts* (pp. 117–119). Leiden.

Haustein, S., Peters, I., Bar-Ilan, J., Priem, J., Shema, H., & Terliesner, J. (2014). Coverage and adoption of altmetrics sources in the bibliometric community. *Scientometrics*, 1–19. doi:10.1007/s11192-013-1221-3

Haustein, S., & Siebenlist, T. (2011). Applying social bookmarking data to evaluate journal usage. *Journal of Informetrics*, *5*(3), 446–457. doi:http://dx.doi.org/10.1016/j.joi.2011.04.002

Hennies, M., & Dressler, J. (2006). Clients information seeking behaviour: An OPAC transaction log analysis. In *click 06, ALIA Biennal Conference*. Perth, AU.

Hirsch, J. E. (2005). An index to quantify an individual's scientific research output. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(46), 16569–72. doi:10.1073/pnas.0507655102

Hirsch, J. E. (2010). An index to quantify an individual's scientific research output that takes into account the effect of multiple coauthorship. *Scientometrics*, *85*(3), 741–754. doi:10.1007/s11192-010-0193-9

Höchstötter, N., & Koch, M. (2008). Standard parameters for searching behaviour in search engines and their empirical evaluation. *Journal of Information Science*, *35*(1), 45–65. doi:10.1177/0165551508091311

Jansen, B. J., & Spink, A. (2006). How are we searching the World Wide Web? A comparison of nine search engine transaction logs. *Information Processing & Management*, *42*(1), 248–263. doi:10.1016/j.ipm.2004.10.007

Järvelin, K., & Ingwersen, P. (2012). User-oriented and cognitive models of information retrieval. In M. J. Bates (Ed.), *Understanding Information Retrieval Systems* (pp. 47–64). Boca Raton, Fl.: CRC Press.

Jindal, V., Bawa, S., & Batra, S. (2014). A review of ranking approaches for semantic search on Web. *Information Processing & Management*, *50*(2), 416–425. doi:10.1016/j.ipm.2013.10.004

Joachims, T. (2002). Optimizing search engines using clickthrough data. *Proceedings of the Eighth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining - KDD '02*, 133–142. doi:10.1145/775066.775067

Joachims, T., Granka, L., Pan, B., Hembrooke, H., & Gay, G. (2005). Accurately interpreting clickthrough data as implicit feedback. In *Proceedings of the 28th annual international ACM SIGIR conference on Research and development in information retrieval - SIGIR '05* (pp. 154 – 161). New York: ACM Press. doi:10.1145/1076034.1076063

Jones, R., Zhang, W. V., Rey, B., Jhala, P., & Stipp, E. (2008). Geographic intention and modification in web search. *International Journal of Geographical Information Science*, *22*(3), 229–246. doi:10.1080/13658810701626186

Jones, S., Cunningham, S. J., McNab, R., & Boddie, S. (2000). A transaction log analysis of a digital library. *International Journal on Digital Libraries*, *3*(2), 152–169. doi:10.1007/s007999900022

Jung, S., Herlocker, J. L., & Webster, J. (2007). Click data as implicit relevance feedback in web search. *Information Processing & Management*, *43*(3), 791–807. doi:10.1016/j.ipm.2006.07.021

Kashyap, A., Amini, R., & Hristidis, V. (2012). SonetRank: leveraging social networks to personalize search. In *Proceedings of the 21st ACM international conference on Information and knowledge management - CIKM '12* (pp. 2045–2049). New York: ACM Press. doi:10.1145/2396761.2398569

Kelly, D., & Belkin, N. J. (2004). Display time as implicit feedback. In *Proceedings of the 27th annual international conference on Research and development in information retrieval - SIGIR '04* (pp. 377–384). New York: ACM Press. doi:10.1145/1008992.1009057

Khoo, C. S. G. G., & Wan, K.-W. (2004). A simple relevancy-ranking strategy for an interface to Boolean OPACs. *The Electronic Library*, *22*(2), 112–120. doi:10.1108/02640470410533380

Knüttel, H., & Deinzer, G. (2013). Optimierung des Lehrbuchbestandes durch detaillierte Nutzungsstatistiken. *GMS Medizin-Bibliothek-Information*, *13*(3), 1–7. doi:10.3205/mbi000285

Larson, R. R. (2012). Information retrieval systems. In M. J. Bates (Ed.), *Understanding Information Retrieval Systems* (pp. 15–30). Boca Raton, FL.: Taylor & Francis.

Larson, R. R., & Frontiera, P. (2004). Spatial ranking methods for geographic information retrieval (GIR) in digital libraries. In R. Heery & L. Lyon (Eds.), *Research and Advanced Technology for Digital Libraries SE - 5* (pp. 45–56). Berlin ; Heidelberg: Springer. doi:10.1007/978-3-540-30230-8_5

Leveling, J., Ghorab, M. R. E., Magdy, W., Jones, G. J. F., & Wade, V. P. (2010). *DCU-TCD@LogCLEF 2010: Re-ranking document collections and query performance estimation*. Retrieved from http://hdl.handle.net/2262/41262

Lewandowski, D. (2005a). *Web Information Retrieval: Technologien zur Informationssuche im Internet*. Frankfurt am Main: Dt. Ges. für Informationswiss. und Informationspraxis.

Lewandowski, D. (2005b). Web searching, search engines and Information Retrieval. *Information Services and Use*, *25*(3/4), 137–147.

Lewandowski, D. (2006). Suchmaschinen als Konkurrenten der Bibliothekskataloge: Wie Bibliotheken ihre Angebote durch Suchmaschinentechnologie attraktiver und durch Öffnung für die allgemeinen Suchmaschinen populärer machen können. *Zeitschrift Für Bibliothekswesen Und Bibliographie*, *53*(2), 71–78.

Lewandowski, D. (2008a). A three-year study on the freshness of web search engine databases. *Journal of Information Science*, *34*(6), 817–831. doi:10.1177/0165551508089396

Lewandowski, D. (2008b). Search engine user behaviour: How can users be guided to quality content? *Information Services & Use*, *28*, 261–268. doi:10.3233/ISU-2008-0583

Lewandowski, D. (2009). Ranking library materials. *Library Hi Tech*, *27*(4), 584–593. doi:10.1108/07378830911007682

Lewandowski, D. (2010a). Google Scholar as a tool for discovering journal articles in library and information science. *Online Information Review*, *34*(2), 250–262. doi:10.1108/14684521011036972

Lewandowski, D. (2010b). Using search engine technology to improve library catalogs. *Advances in Librarianship*, *32*, 35–54. doi:10.1108/S0065-2830(2010)0000032005

Lewandowski, D. (2011). Query understanding. In D. Lewandowski (Ed.), *Handbuch Internet-Suchmaschinen 2 : Neue Entwicklungen in der Web-Suche* (pp. 55–75). Heidelberg: AKA Verlag.

Lewandowski, D. (2012). Credibility in web search engines. In S. Apostel & M. Folk (Eds.), *Online Credibility and Digital Ethos: Evaluating Computer - Mediated Communication* (pp. 131–145). Hershey, PA: IGI Global.

Lewandowski, D., & Höchstötter, N. (2008). Web searching: A quality measurement perspective. In A. Spink & M. Zimmer (Eds.), *Web Search* (pp. 309–340). Berlin ; Heidelberg: Springer.

Lewandowski, D., Wahlig, H., & Meyer-Bautor, G. (2006). The freshness of web search engine databases. *Journal of Information Science*, *32*(2), 131–148. doi:10.1177/0165551506062326

Luther, J. (2003). Trumping Google? Metasearching's promise (cover story). *Library Journal*, *128*(16), 36–39.

Maylein, L., & Langenstein, M. (2013). Neues vom Relevanz-Ranking im HEIDI-Katalog der Universitätsbibliothek Heidelberg. *B.I.T. Online*, *16*(3), 190–200.

Mayr, P. (2011). Bradfordizing als Re-Ranking-Ansatz in Literaturinformationssystemen. *Information : Wissenschaft Und Praxis*, *62*(1), 3–10.

Mayr, P. (2013). Relevance distributions across Bradford zones: Can Bradfordizing improve search? In *14th International Society of Scientometrics and Informetrics Conference* (pp. 1493–1505).

Mizzaro, S. (1997). Relevance: The whole history. *Journal of the American Society for Information Science and Technology*, *48*(9), 810–832. doi:10.1002/(SICI)1097-4571(199709)48:9<810::AID-ASI6>3.0.CO;2-U

Mönnich, M., & Spiering, M. (2008). Adding value to the library catalog by implementing a recommendation system. *D-Lib Magazine*, *14*(5/6). Retrieved from http://www.dlib.org/dlib/may08/monnich/05monnich.html

Nicholas, D., Huntington, P., & Jamali, H. R. (2007). The use, users, and role of abstracts in the digital scholarly environment. *The Journal of Academic Librarianship*, *33*(4), 446–453. doi:10.1016/j.acalib.2007.03.004

Niedermair, K. (2014). Gefährden Suchmaschinen und Discovery-Systeme die informationelle Autonomie? *Mitteilungen Der VÖB*, *67*(1), 109–125.

Niu, X., & Hemminger, B. M. (2010). Beyond text querying and ranking list: How people are searching through faceted catalogs in two library environments. *Proceedings of the American Society for Information Science and Technology*, *47*(1), 1–9. doi:10.1002/meet.14504701294

Oberhauser, O. (2008). Sachliche Erschliessung im Österreichischen Verbundkatalog: Status und Perspektiven. *Mitteilungen Der VÖB*, *61*(3), 59–77.

Oberhauser, O. (2010). Relevance Ranking in den Online-Katalogen der "nächsten Generation." *Mitteilungen Der VÖB*, *63*(1/2), 25–37.

Page, L., Brin, S., Motwani, R., & Winograd, T. (1998). *The PageRank citation ranking: bringing order to the web*. *Technical report, Stanford Digital Library Technologies Project*. Stanford, CA: Stanford University. Retrieved from http://ilpubs.stanford.edu:8090/422/

Pan, B., Hembrooke, H., Joachims, T., Lorigo, L., Gay, G., & Granka, L. (2007). In Google We Trust: Users' Decisions on Rank, Position, and Relevance. *Journal of Computer-Mediated Communication*, *12*(3), 801–823. doi:10.1111/j.1083-6101.2007.00351.x

Parry, M. (2010). After losing users in catalogs, libraries find better search software. *Education Digest*, *75*(February), 54–57.

Pera, M. S., Lund, W., & Ng, Y.-K. (2009). A sophisticated library search strategy using folksonomies and similarity matching. *Journal of the American Society for Information Science and Technology*, *60*(7), 1392–1406. doi:10.1002/asi.21072

Pesch, O. (2007). Usage statistics: About COUNTER and SUSHI. *Information Services & Use*, *27*(4), 207–213.

Peters, I. (2011). Folksonomies und Kollaborative Informationsdienste : Eine Alternative zur Websuche? In D. Lewandowski (Ed.), *Handbuch Internet-Suchmaschinen* (pp. 29–53). Heidelberg: AKA Verlag.

Pianos, T. (2010). EconBiz - Meeting user needs with new technology. *Liber Quarterly*, *20*(1), 4–24.

Priem, J., Piwowar, H. A., & Hemminger, B. M. (2012). Altmetrics in the wild: Using social media to explore scholarly impact. In *altmetrics12: ACM Web Science Conference 2012 Workshop*. Evanston, IL.

Priem, J., Taraborelli, D., Groth, P., & Neylon, C. (2011). Altmetrics: a manifesto. Retrieved July 15, 2014, from http://altmetrics.org/manifesto/

Radlinski, F., & Joachims, T. (2005). Query chains: Learning to rank from implicit feedback. In *Proceeding of the eleventh ACM SIGKDD international conference on Knowledge discovery in data mining - KDD '05* (pp. 239–248). New York: ACM Press. doi:10.1145/1081870.1081899

Riemer, K., & Brüggemann, F. (2009). Personalisierung der Internetsuche - Lösungstechniken und Marktüberblick. In D. Lewandowski (Ed.), *Handbuch Internet-Suchmaschinen* (pp. 148–171). Heidelberg: AKA Verlag.

Rowlands, I., Nicholas, D., Williams, P., Huntington, P., Fieldhouse, M., Gunter, B., … Tenopir, C. (2008). The Google generation: the information behaviour of the researcher of the future. *Aslib Proceedings*, *60*(4), 290–310. doi:10.1108/00012530810887953

Saracevic, T. (1996). Relevance reconsidered. In *Information science: Integration in perspectives. Proceedings of the Second Conference on Conceptions of Library and Information Science (CoLIS 2)* (pp. 201–208). Copenhagen.

Saracevic, T. (2006). Relevance: A review of the literature and a framework for thinking on the notion in information science. Part II. *Advances in Librarianship*, *30*, 3–72. doi:10.1016/S0065-2830(06)30001-3

Schmidt-Mänz, N., & Koch, M. (2005). Patterns in Search Queries. In D. Baier, R. Decker, & L. Schmidt-Thieme (Eds.), *Data Analysis and Decision Support* (pp. 122–129). Berlin ; Heidelberg: Springer. doi:10.1007/3-540-28397-8_15

Schneider, K. G. (2006). How OPACs suck, Part 1: Relevance rank (or the lack of it). *ALA TechSource*. Retrieved March 17, 2014, from http://www.alatechsource.org/blog/2006/03/how-opacs-suck-part-1-relevance-rank-or-the-lack-of-it.html

Schneider, R. (2009). OPACs, Benutzer und das Web. *Zeitschrift Für Bibliothekswesen Und Bibliographie*, *56*(1), 9–19.

Shepherd, S. J. (2007). Concepts and architectures for next-generation information search engines. *International Journal of Information Management*, *27*(1), 3–8. doi:10.1016/j.ijinfomgt.2006.06.005

Smith, A. G. (2004). Web links as research indicators: analogues of citations? *Information Research*, *9*(4). Retrieved from http://www.informationr.net/ir/9-4/paper188.html

Smith, D. R. (2012). Impact factors, scientometrics and the history of citation-based research. *Scientometrics*, *92*(2), 419–427. doi:10.1007/s11192-012-0685-x

Spink, A., Wolfram, D., Jansen, M. B. J., & Saracevic, T. (2001). Searching the web: The public and their queries. *Journal of the American Society for Information Science and Technology*, *52*(3), 226–234. doi:10.1002/1097-4571(2000)9999:9999<::AID-ASI1591>3.0.CO;2-R

Stock, W. G., & Stock, M. (2013). *Handbook of Information Science*. Berlin; Boston: De Gruyter Saur. doi:10.1515/9783110235005.416

Sugiyama, K., Hatano, K., & Yoshikawa, M. (2004). Adaptive web search based on user profile constructed without any effort from users. In *Proceedings of the 13th conference on World Wide Web - WWW '04* (pp. 675–684). New York: ACM Press. doi:10.1145/988672.988764

Surowiecki, J. (2005). *The wisdom of crowds*. New York, NY [u.a.]: Anchor Books.

Swanson, D. R. (1986). Subjective versus objective relevance in bibliographic retrieval systems. *The Library Quarterly*, *56*(4), 389–399.

Thelwall, M., & Sud, P. (2011). A comparison of methods for collecting web citation data for academic organizations. *Journal of the American Society for Information Science and Technology*, *62*(8), 1488–1497. doi:10.1002/asi.21571

Vaughan, J. (2012). Investigations into library web-scale discovery services. *Information Technology & Libraries*, *31*(1), 32–82.

Vechtomova, O., & Karamuftuoglu, M. (2008). Lexical cohesion and term proximity in document ranking. *Information Processing & Management*, *44*(4), 1485–1502. doi:10.1016/j.ipm.2008.01.003

Wang, D., Song, C., & Barabási, A.-L. (2013). Quantifying long-term scientific impact. *Science*, *342*(6154), 127–132. doi:10.1126/science.1237825

West, J. D., Bergstrom, T. C., & Bergstrom, C. T. (2010). The Eigenfactor Metrics TM: A network approach to assessing scholarly journals. *College & Research Libraries*, *71*(3), 236–244.

White, H. D. (2007). Combining Bibliometrics, Information Retrieval, and Relevance Theory, Part 2 : Some Implications for Information Science. *Journal of the American Society for Information Science and Technology*, *58*(4), 583–605. doi:10.1002/asi.20542

Xu, S., Jiang, H., & Lau, F. C. M. (2011). Mining user dwell time for personalized web search re-ranking. In *IJCAI'11 Proceedings of the Twenty-Second international joint conference on Artificial Intelligence* (pp. 2367–2372). AAAI Press. doi:10.5591/978-1-57735-516-8/IJCAI11-394

Yang, S. Q., & Hofmann, M. A. (2010). The next generation library catalog: A comparative study of the OPACs of Koha, Evergreen, and Voyager. *Information Technology & Libraries*, *29*(3), 141–150.

Yang, S. Q., & Hofmann, M. A. (2011). Next generation or current generation? A study of the OPACs of 260 academic libraries in the USA and Canada. *Library Hi Tech*, *29*(2), 266–300. doi:10.1108/07378831111138170

Yang, S. Q., & Wagner, K. (2010). Evaluating and comparing discovery tools: how close are we towards next generation catalog? *Library Hi Tech*, *28*(4), 690–709. doi:10.1108/07378831011096312

Yeadon, J. (2001). Web site statistics. *VINE*, *31*(3), 55–60. doi:10.1108/03055720010804168

Yu, H., & Young, M. (2004). The impact of web search engines on subject searching in OPAC. *Information Technology & Libraries*, *23*(4), 168–180.

Zhao, J., Jin, P., Zhang, Q., & Wen, R. (2014). Exploiting location information for web search. *Computers in Human Behavior*, *30*, 378–388. doi:10.1016/j.chb.2013.04.023

Zhitomirsky-Geffet, M., Feitelson, D. G., Frachtenberg, E., & Wiseman, Y. (2009). A unified strategy for search and result representation for an online bibliographical catalogue. *Online Information Review*, *33*(3), 511–536. doi:10.1108/14684520910969934

Appendix – Ranking factors overview

Ranking factor	Description	Web Search Engines	Library Materials			Example / Note
			Metadata	Enriched Metadata	Full Text / Printed Materials	
1. Text statistics						
Term frequency (TF)	Relative frequency of search term in a document	Terms within website content, metatags	Terms within bibliographic data such as title, author, keywords	Terms within enriched data such as abstract, table of contents	Terms within full text	Metadata do not provide enough text; amount of text per catalog listing varies greatly, the same ranking algorithm cannot be applied
Inverse document frequency (IDF)	Relative frequency of term in all documents (rarely occurring terms are preferred)	Terms within website content (all documents)	N/A	N/A	Terms within a whole document collection	Search term “relevance” is weighted higher than terms “and” or “the”, although such terms occur more frequently
Search term order	If a query consists of more than one search term, documents with the term at the beginning of the query will be ranked higher	Search term order within website content, metatags	Search term order within bibliographic data such as title, author, keywords	Search term order within enriched data such as abstract, table of contents	Search term order within full text	Query “information retrieval” should not produce result “retrieval information”
Search term distance	If a query consists of more than one search term, documents with the terms closest to each other will be ranked higher	Search term distance within website	Search term distance within for example title field	Search term distance within for example abstract, table of content	Search term distance within full text	Query: “information retrieval” ranks results with “information retrieval” first

Position of search term	If term appears in beginning of document, it is ranked higher	Position of search term within website	Position of search term in bibliographic metadata	Position of search term in abstract or table of content	Position of search term in the document	Document with search terms appearing in title field is weighted higher than document with terms in abstract
Document length	Documents within a certain range of lengths are preferred (content must be neither too short nor too long)	Size of website content	N/A	N/A	Size of item (number of pages)	Dependent on discipline, for instance short papers are more common in natural sciences than the humanities
Emphasis on terms within a document	Terms that are emphasized are weighted higher	Emphasis on terms within website	Emphasis on terms within bibliographic data	Emphasis on terms within additional information	Emphasis on terms within a document	Bold or italic text
Anchor text	Terms appearing in anchor text are ranked higher	Anchor text of hyperlinks	Anchor text of hyperlinks as references to other entities	Anchor text, e.g. in review, abstract	Anchor text within full text	References to conference title, series title
2. Popularity						
Click popularity	Documents that have been visited by many users are preferred (in conjunction with dwell time)	Number of clicks on web content	Number of clicks on bibliographical record	Number of clicks on table of contents, other additional information (e.g. publisher's content description)	Number of clicks on full text (article, book chapter)	Difficult to compare with printed books (look-ups only measurable by observing patrons on-site)
Dwell time	The amount of time the document has been viewed (amount of time that implies reading)	Dwell time	Dwell time on bibliographical record	Dwell time on additional information (e.g. abstract, review)	Dwell time on whole document	Dwell time on electronic item difficult to compare with printed book (look-ups only measurable by observing patrons on-site)

Usage frequency	Items that have been downloaded or loaned more often than others are ranked higher	*see click popularity in conjunction with dwell time* / number of PDF downloads	Number of citation downloads to reference manager	Number of citation and abstract downloads	Number of full text downloads / number of loans at the library or within a library network	A standard for retrieving download statistics is COUNTER; Circulation data include circulation rate, period, number of extended or renewed loans
Purchasing behavior	Works that have a large number of purchased copies (locally and globally) are weighted higher	N/A	N/A	N/A	**Local:** number of copies in stock (in open or closed stacks, text book collection, on course reserve) **Global:** number of editions or different formats (content also published as audio book or DVD), number of item owning libraries	Quality/authority is indicated by selection for acquisition by librarians / faculty members
Publisher authority	Works by a highly reputed publisher are weighted higher	Documents from larger web sites (or within a certain size range) are preferred	N/A	Reviews (reviewed by experts indicates quality/authority)	Number of items by a certain publisher; peer-reviewed vs. non peer-reviewed journals	Quality/authority is indicated by the reputation of the publisher, e.g. theses by an elite university
Ratings	Documents rated or recommended by others are ranked higher	Explicit ratings through likes, Google+, star ratings, etc.	N/A	N/A	Recommendations for books, articles by explicit user ratings or comments	Authority indicated by recommendations in the form of acquisition requests, bibliographies by faculty members
Reference counting	Document is ranked by the number of incoming links or citations	Link counting measures the number of incoming links (regardless of the content quality the link points to)	N/A	N/A	Citation counting for journal, item, author (without relation to other works)	Bibliometric measures including, e.g. citation counts, number of published papers, for example: Journal Impact Factor, Eigenfactor, Author's H-Index; Alternative Metrics
Reference popularity	Document is ranked by the number of links or citations in relation to other documents or entities	Link popularity measures linking within the Web graph (e.g. Google's PageRank)	N/A	N/A	Citation impact for journal, item, author (citation counting with relation to other works)	

3. Freshness						
Publication date	Current documents are preferred	Documents that have recently been published are preferred	N/A	Publication date of review	Publication date of books, articles (both print and electronic)	"Need for freshness" depends on discipline, research area (data derivable from subject heading, classification)
Accession date	Documents that have recently been accessioned are preferred	Websites that have recently been accessioned are ranked higher	N/A	Accession date of review or other additional information	Date of acquisition or licensing by the library	
4. Locality & availability						
Physical location	Physical location of the item and the user	Pages that are "closer" to the location of the user are preferred (geo targeting)	N/A	N/A	Physical location of user and item (home, central library / library branch or campus	Location information derivable from IP address
Availability of the webpage / item	Current availability of the item	Website available or currently unavailable?	N/A	Table of contents or review available?	Is the (e-)book or (e-)journal volume currently available (licensed, open access, printed item currently on loan or status missed)?	Current availability derivable from circulation data (preferable on a daily basis)
5. Content properties						
Additional information	Documents with other information than basic descriptions available are weighted higher	Metatags such as keywords, description exist	Bibliographic record must be complete	Abstracts, table of contents, annual indexes, reviews	Availability of underlying research data / supplements	High degree of indexing indicates reliability of information sources

File format	Documents written in a particular file format are preferred to other formats	Standard HTML is preferred to PDF or DOC	Citation format for reference manager	Citation format (with abstract) for reference manager	Full text preferably as PDF / Format can relate to circulation status (large format items, e.g. maps, are potentially non-circulating items)	Documents for academic use must be quotable (PDF generally preferred to HTML)
Language	Documents in the preferred language(s) are ranked higher	Same language of query term and used interface	Language of bibliographic description	Language of additional information, e.g. abstract, review	Language of full text / printed item	Preferred language(s) set in personal profile setting or derived from usage data
6. User background						
User group	Ranking based on preferences of a particular user group	Heterogeneous user base, untrained end users	N/A	N/A	Usage data of a particular user group (e.g. circulation data from professors, students, graduate students)	User data derivable from IP (university membership), library card number or student ID
Usage data	Personalized ranking based on personal (social) profile of the individual user	Clickstream data from navigation, browsing history	Clickstream data from navigation in bibliographic records	Clickstream data from navigation abstracts, reviews, table of contents, other additional information	Usage data (e.g. click data on full texts, or circulation data of a particular user)	Informed consent by user required; privacy and data protection must be ensured