

Relevance Ranking in elektronischen Bibliothekskatalogen (OPACs)

Fabian Fricke

fabian.fricke@bluewin.ch

Von verschiedener Seite wird immer wieder gefordert, dass die Ergebnisse in den Trefferlisten der OPACs ebenso wie in denjenigen von Internetsuchmaschinen nach ihrer Relevanzwahrscheinlichkeit geordnet sein sollten (Relevance Ranking): Diejenigen Treffer, die dem in der Suchanfrage ausgedrückten Informationsbedürfnis des Nutzers am besten entsprechen, sollten in der Trefferliste auch zuoberst stehen. Der überwiegende Teil der Nutzer sieht eine solche Funktion aufgrund ihrer Vertrautheit mit Internetsuchmaschinen heutzutage als selbstverständlich an. In OPACs hingegen stellt Relevance Ranking noch nicht den Standard dar, und wo es angeboten wird, ist die Qualität der Sortierung nicht immer überzeugend. Dies liegt auch daran, dass sich die Suchbedingungen in OPACs von denjenigen im World Wide Web unterscheiden und dass die Algorithmen von Internetsuchmaschinen nicht einfach auf OPACs übertragen werden können. Wie aber muss ein Ranking-Algorithmus aussehen, der im spezifischen Kontext von Bibliothekskatalogen gute Resultate liefert? Zur Beantwortung dieser Frage soll die vorliegende Arbeit einen Beitrag leisten.

Dabei beleuchtet sie in einem ersten Teil anhand zahlreicher Nutzerstudien das Verhalten der OPAC-Nutzer, insoweit es für das Thema Relevance Ranking relevant ist. Es zeigt sich, dass viele Nutzer bei der Suche in OPACs auf Schwierigkeiten stoßen, bei deren Bewältigung ein gutes Relevance Ranking helfen kann. Insbesondere haben Nutzer oft Probleme, präzise Suchanfragen zu formulieren und mit großen Treffermengen zurechtzukommen. Die Ergebnisse legen zudem nahe, dass sich das Sucherlebnis in OPACs so weit wie möglich an demjenigen in Internetsuchmaschinen orientieren sollte und dass Transparenz und Nachvollziehbarkeit der Sortierung den Nutzern wichtige Bedürfnisse sind.

Im zweiten Teil der Arbeit werden zahlreiche mögliche Kriterien für das Relevance Ranking in OPACs unter die Lupe genommen, eingeteilt in die Gruppen Textabgleich, Popularität, Neuheit, Verfügbarkeit, Personalisierung und weitere Eigenschaften der Ressource. Es zeigt sich, dass die Methode des Textabgleichs – die Basis des klassischen Relevance Rankings –, den Bedingungen eines Bibliothekskatalogs angepasst, zwar ein Gerüst für das Relevance Ranking liefern kann, jedoch durch weitere Kriterien insbesondere aus den Gruppen Popularität (z. B. Ausleihhäufigkeit, Zahl der besitzenden Bibliotheken) und Neuheit ergänzt werden sollte. Klar wird zudem, dass es nicht einen einzigen Relevance-Ranking-Algorithmus geben kann, der für alle Nutzer und Bestände die besten Resultate liefert: Jede Bibliothek wird selbst entscheiden müssen, welche Faktoren in ihren spezifischen Umständen welches Gewicht haben sollen. Neben einer kompletten Personalisierung ist auch die Entwicklung mehrerer Algorithmen für verschiedene Nutzergruppen eine Option für die Zukunft; solange man mit einem einzigen Algorithmus arbeitet, sollte sich dieser vornehmlich an den Bedürfnissen weniger erfahrener Nutzer orientieren. Für erfahrene Rechercheure sollten auf jeden Fall neben Relevance Ranking weiterhin andere Sortierungsarten angeboten werden.