

DISCOVERY-SYSTEME, DIE IHREM NAMEN EHRE MACHEN

VERBESSERUNGSPOTENZIALE FÜR BIBLIOTHEKARISCHE SUCHMASCHINEN

Hands-on Lab, Bibliothekskongress 2019, 18. März

Kristof Keßler, Technische Universität Braunschweig, Universitätsbibliothek, Braunschweig

Anne Christensen, Leuphana Universität Lüneburg, Medien- und Informationszentrum , Lüneburg

Dr. Jarmo Schrader, Stiftung Universität Hildesheim, Universitätsbibliothek, Hildesheim

Dr. Jan F. Maas, Staats- und Universitätsbibliothek Hamburg Carl von Ossietzky, Hamburg

WILLKOMMEN ZUM HANDS-ON LAB
DISCOVERY-SYSTEME, DIE IHREM NAMEN EHRE MACHEN

VORSTELLUNG DES VORGEHENS

1. Ergänzung Themen	5 bis 10 Minuten
2. Bildung von Themenclustern durch Moderationsteam	5 Minuten
3. Aufteilung auf Gruppen Diskussionen	5 bis 10 Minuten
4. Diskussionszeit pro Themenbereich	15 + 5 Minuten
4.1 Während der Diskussion: Wichtige Punkte auf Kärtchen notieren	
4.2 Nach 15 Minuten Diskussionszeit: Zusammenfassung der Diskussionspunkte	
5. Wechsel des Themenbereiches → zweimalige Wiederholung der Punkte 3. und 4.	
6. Zusammenfassung der Diskussionsergebnisse	5 Minuten

Eine Zusammenfassung der Ergebnisse wird erstellt und auf OPUS veröffentlicht.

Themenbereich: Metadaten

Facetten

EDV-gerechte
DatenDublette
DatenQualität
des IndexNormdaten-
integration

Datenmigration

Heterogene
DatenErweiterung
MetadatenschemaOpen-Access-
Publikationen

Themenbereich: Technik

Ranking nach
AusleihzahlenIntransparentes
RankingSuchtech-
nologienLinked
Open DataTestsets für
RankingSharding/
Förderierte Suche

Themenbereich: Nutzbarkeit

Schlechte
Verfügbarkeits-
informationen„Das Ding
macht Spaß!“

Nutzbar

Leicht

Vollständigkeit

Daten-/Quellen
Transparenz

DISKUSSIONSERGEBNISSE - METADATEN

Metadatenaufbereitung und -anreicherung

- Bibliothekarische Metadaten sollten besser genutzt werden. Norm- und andere Erschließungsdaten von Print-Titeln automatisch auf elektronische Ausgaben vererben.
- Deduplizierung: Sicherstellen, dass im Index vorhandene, aber mager erschlossene Datensätze in denjenigen Fällen verworfen werden, wo besser erschlossene Datensätze existieren.
- Moving Walls sollten korrekt abgebildet werden.
- Es werden Metadaten benötigt, die von technischen Systemen interpretierbar, also EDV-gerecht, sind: zum Beispiel sollten statt textuellen Auflagenkennzeichnungen normierten Zahlenwerte genutzt werden.
- Aktueller Stand der Sacherschließung: Ein "Babylon" verschiedener, teils historischer Erschließungen (sowohl formal als auch inhaltlich). Muss schrittweise aufgeräumt werden. Konkordanzen ein üblicher Weg, der aber viele Probleme mit sich bringt.
- Facetten: Daten hinter Facetten müssen
 - im Idealfall 100% Abdeckung besitzen
 - sinnvolle Facetten bilden
 - sinnvolle Namen besitzen
 - zutreffen
- Es ist ein stärkerer Einsatz von automatischen Verfahren (Formalerschließung, Sacherschließung, Dublettenreduzierung) gewünscht.
- Es ist einerseits der Inhalt von Such-Indizes zu überdenken:
 - Wie viel ist zu viel?
 - Müssen wir wirklich auch solche Inhalte einbinden, die sich sehr leicht auch anderweitig finden lassen (z.B. Patente, OA-Titel)?
- Aber andererseits ist der Datenumfang des Discovery-Systems ein wichtiges Nutzungskriterium: Es werden möglichst viele Daten auch aus verschiedenen Quellen benötigt.

Metadatenqualität

- Grundlegende Fragen:
 - Welche Ziele sollen die Metadaten unterstützen?
 - Wie kann Qualität gemessen werden?
 - Kann/Sollte diese nach der Erfüllung von Katalogisierungsrichtlinien gemessen werden?
 - Wer entscheidet über das Regelwerk der Erfassung?
 - Kann eine Validierung auch die inhaltliche Bedeutung der Daten prüfen?
- Es sollte bei der Qualitätsmessung nach ISO-Normen vorgegangen werden.
- Wie ermittelt man die Kennzahlen, gegen die man misst? Die Kennzahlen für Qualität sollten auf Basis der Ziele erstellt werden und das Regelwerk der Datenerfassung sollte auf die Erfüllung der Ziele in EDV-gerechter Form ausgelegt sein.
- Ziele müssen definiert sein, also z.B. Unterstützung von Discovery-Systemen.
- Ein Beispiel für ein System, das entsprechend der ursprünglichen Ziele sehr gut funktioniert, ist die EZB für die Administration der Besitznachweise. Diese ist hingegen nicht für die zuverlässige Anreicherung von Artikeln mit Bestandsdaten geeignet.
- Es entstehen Unschärfen durch Regelwerke, die für einen anderen Zweck als Discovery-Systeme erstellt worden sind.
- Wenn es das Ziel ist, Discovery-Systeme zu unterstützen, dann müssten die Regeln der Erfassung geändert werden (oder sie müssten umgangen werden).
- Mapping zwischen Regelwerken ist möglich, bringt aber immer Probleme mit sich.
- Mehr Quellen an Daten führen zu noch mehr Heterogenität und sind somit noch schwieriger zusammenzuführen. Es gilt nicht nur die manuell erfassten Daten zu integrieren, sondern auch die durch Systeme erfassten Daten, z.B. E-Book-Pakete.
- Es müssen sowohl manuelle als auch die maschinell erfassten Daten validiert werden. Die Frage, ob manuell erstellte oder maschinell erfasste Daten "besser" oder "schlechter" sind, ist irrelevant. Es sollten die gleichen Qualitätsmaßstäbe angelegt werden.
- Wenn man sich auf "bessere", also zielführendere, Regeln einigt, so müssten diese auch retrospektiv angewendet werden. Wer hat die Zeit diese Umwandlungen vorzunehmen? Auch wenn dies größtenteils automatisch geschehen würde, müssten Prüfungen vorgenommen werden. Oder es wird die Entscheidung getroffen, dass die Altdaten doch nicht angepasst werden.

Weitere Themen

- Auto Suggest muss von den Metadaten im Suchindex während der Eingabe von Suchbegriffen unterstützt werden.
- Fehlertoleranz: Welche Fehler tolerieren wir als Bibliothekar*innen, welche tolerieren die Nutzer*innen?
- Fachdatenbanken mit speziellen Suchmethoden (individuelle Klassifikationen o.ä.) zumindest mit Database Recommendern im Discovery-System suchspezifisch verlinken.
- Doppelarbeit ist ein Problem: Viele Objekte sind mehrfach katalogisiert. Aber wohl doch nicht so viele Dubletten wie gedacht, zumindest bei der Zusammenführung von GBV und SWB. Ein Nationalkatalog erscheint daher in Deutschland, im Gegensatz zu z.B. Dänemark, schwierig aufzubauen zu sein. Die EKI war einmal eine Hoffnung als Grundlage für die Zusammenfassung in dieser Hinsicht.
- Föderierte Suche: Besser alles in einem Index unterbringen.

DISKUSSIONSERGEBNISSE - TECHNIK

Grundsätzliche Aspekte

- Wissen wir überhaupt was Nutzer suchen? (Nicht allein auf einen elektronischen Katalog oder eine Discovery-System bezogen, sondern auch auf die allgemeine Interaktion mit Nutzern.)
- Wie erkennt man “lose Enden” bei Suchvorgängen? Sind viele Suchen gut oder schlecht? Ist ein Klick zu einem externen Link als gutes Ergebnis zu bewerten oder ist dies bei schneller Rückkehr in die Suche als nicht so gutes Ergebnis zu bewerten?
- Sollen Nutzer richtig suchen lernen oder sollen die Systeme sich anpassen und ihnen dies abnehmen?
- Es ist bei Ranking-Diskussionen auch immer zu bedenken, dass bei kommerziellen Anbietern in der Regel überhaupt keine Transparenz besteht und deren Ranking-Algorithmus als “Geschäftsgeheimnis” gilt.

Zufriedenheit der Nutzer*innen

- Am Ende muss der Anspruch sein, zufriedene Nutzer*innen zu haben.
- Was wäre eine Messgröße für die Zufriedenheit der Nutzer*innen hinsichtlich des Rankings? Eine erfolgreiche Suche, aber was genau ist ein Erfolg? Ist ein Linkout ein Erfolg? Ist eine Vormerkung, eine Bestellung ein Erfolg? Können auch schon Vorschaubilder ein Erfolg sein? Näherungsweise könnten diese Aktionen gezählt werden, um zumindest eine Tendenz zu erhalten. Print-Nutzung bliebe weitgehend unberücksichtigt.
- Zufriedenheit könnte direkt abgefragt werden, z.B. durch Auswahl eines passenden Smileys (Fröhlich, neutral, wütend) nach Beenden einer Suchsession.
- Probleme bei der Interpretation der Bewertung durch Nutzer werden anhand einer Studie sichtbar gemacht, die Bewertungen durch Stiftung Warentest und Amazon-Bewertungen vergleicht. Diese fallen sehr stark auseinander. Eine Teilbegründung dafür kann sein, dass bei Amazon die zufriedene Mitte fehlt. Diese hat in der Regel keine Motivation zu bewerten. In der Regel sind Personen, die sehr positive oder sehr negative Erfahrungen gemacht haben motivierter Bewertungen abzugeben. Man muss sich also die Fragen stellen, was Bewertungen aussagen und was die Motivation dahinter sein könnte. Sollte man also insgesamt in dieser Hinsicht nicht mehr von Datenanalyse, sondern Dateninterpretation auf Basis begründeter Annahmen sprechen?

Kontextabhängigkeit des Rankings

- Relevanz ist abhängig vom Kontext der Nutzer*innen. Dies soll aber nicht das Plädoyer dafür sein, massenhaft Daten von diesen zu sammeln. Sollte man am Paradigma der Datensparsamkeit rütteln?
- Aus unterschiedlichen Kontexten ergeben sich teilweise gegensätzliche Ansprüche an das Ranking (z.B. Suche nach Standardwerken vs. Suche nach neuester Literatur).
- Eine Berücksichtigung des Kontext hätte aber auch zur Folge, dass die Treffermengen und das Ranking bei gleichen Suchbegriffen sehr unterschiedlich sein können. Ist es wirklich gewollt, dass im Extremfall jede Person bei gleichen Suchbegriffen andere Ergebnisse erhält?
- Nutzer*innen sollen sich auch nicht beobachtet fühlen.
- Allerdings ist die Annahme bei jüngeren Nutzer*innen, dass diese erwarten, dass das “Ding läuft”, ob es etwas über sie weiß oder nicht ist für sie meist irrelevant.

Testen von Rankings

- Das Ranking sollte getestet werden. Die Definition von Testsets könnte dabei helfen. Diese sind auf bibliotheksübergreifender Ebene vielleicht unzweckmäßig, da immer bestandsabhängig. Es sollten auch Automatismen entwickelt und genutzt werden, die unerwünschte Ranking-Effekte offenlegen.
- Die Erstellung von Testsets ist allerdings äußerst aufwändig.
- Erst einmal müsste man eine Prozess etablieren, der die Nutzererwartung greifbar machen kann. Dann wären diese Erwartungen auch nur statisch und auf den Moment bezogen. Zum einen entwickeln sich Erwartungen weiter, zum anderen entwickeln sich die Inhalte der Suchindizes weiter.
- Um die Nutzerzufriedenheit zu gewährleisten ist es daher wichtiger ständig kleine Optimierungen vorzunehmen.
- Der Wunsch nach umfassenden Testsets ist nachvollziehbar aber er erscheint nicht zielführend und die Umsetzung in umfassender Größe erscheint unrealistisch.
- Bei der Optimierung des Rankings sollte eine Verschlechterung vermieden werden. Dies könnte durch kleine Testsets für spezifisch wichtige Fälle von Datensätzen geschehen. Diese sollten im Prozess der Änderung des Rankings überprüft werden.

Weitere Themen

- Ist TF/IDF geeignet für bibliografische Metadaten?
- Bericht von Versuchen mit einem Regler für die nutzerseitige Beeinflussung des Rankings auf Basis eines Schiebereglers von "1. Semester" bis "Spitzenbedarf". → Keine Nutzung
- Wichtig ist die Unterscheidung zwischen verschiedenen Sucharten, z.B. Known-Item-Searches sollten anders behandelt werden. Beispiel: Nach der Einführung eines Discovery Systems in einer öffentlichen Bibliothek erhält man nun bei der Suche nach Bibi Blocksberg Band 3 statt einem Treffer Dutzende Treffer.
- Die grundlegenden Informationen sollten in der Hilfe erklärt werden.
- Tools für Bibliothekar*innen und Nutzer*innen um das Ranking zu verstehen und einzustellen könnten hilfreich sein. (Aber siehe Hinweis zu keiner Nutzung eines entsprechenden Schiebereglers.)
- Wichtige mögliche Rankingfaktoren:
 - Nutzerverhalten
 - Rankingunterschiede nach Nutzergruppe / Fach
 - Aktualität der Literatur
 - Bibliometrische Kennzahlen
 - Popularität, Ausleihstatistiken
- Aber: Ranking nach Ausleihzahlen wird eher kritisch gesehen.
- Best Practices für Rankings werden benötigt
- Wie können Nutzer*innen auf technischem Weg unterstützt werden, gute Suchen einzugeben? Könnten automatisch generierte Suchvorschläge hier hilfreich sein?
- Es ist die Annahme, dass Nutzer*innen oft nur auf die ersten drei Treffer schauen. Sollte man also die Darstellung der Ergebnisse im Vergleich zu Listen anpassen?
- Ist das Highlighting von eingegebenen Suchbegriffen in der Trefferliste hilfreich?
- Verschiedene Auflagen und Ausgaben sollten möglichst verknüpft werden.

DISKUSSIONSERGEBNISSE - SPAß & GEBRAUCHSTAUGLICHKEIT

Abwesenheit von Stress - Unsicherheit ist eine Spaßbremse

- Nur tatsächlich vorhandene Literatur sollte nachgewiesen werden? (Aber steht dies nicht in Widerspruch dazu, dass Discovery-Systeme sehr umfangreich sein sollten?)
- Systeme sollten so gestaltet werden, dass diese schnell erlernbar sind.
- Es sollte eine Metaebene zur Trefferliste eingeführt werden, die erklärt wie Treffer zustande gekommen sind sowie ob und wie eingegebene Suchbegriffe verändert oder erweitert wurden.
- Es könnten verschiedene Suchmodi angeboten werden, z.B. Erstsemestern könnten vorrangig Lehrbücher angezeigt werden. Ist dies Art der Personalisierung gewünscht? (Siehe auch Hinweis zu keiner Nutzung eines entsprechenden Schiebereglers.)
- Eine Differenzierung in Known-Item- und thematische Suchen sollte vom System vorgenommen werden können.
- Feedback-Möglichkeiten zum Ranking sollten angeboten werden.

Bibliothekarische Sicht

- Haben Bibliothekar*innen andere Wünsche an Transparenz (z.B. Ranking) als Nutzer*innen?
- Es sollte dafür gesorgt werden, dass Signaturen, Zeitschriftentitel bzw. Zeitschriftenabkürzungen gut gesucht und gefunden werden können, damit der alte Katalog tatsächlich obsolet wird.
- Die Möglichkeit einer Feldsuche sollte beibehalten werden.
- "All in One"-Lösung sind trotzdem gewünscht. Es sollte also keine Trennung von Discovery-System und Dienstkatalog vorgenommen werden.
- "Perfektes" Discovery kann alles, was der OPAC auch unterstützt (und mehr).

Grundannahmen der Gebrauchstauglichkeit

- Die Geschwindigkeit ist ein wichtiges Kriterium. Nichts ist schlimmer als ein langsamer Katalog.
- Discovery-Systeme sollten keine Schulungen benötigen, um sie verwenden zu können.
- Wichtig sind klare Informationsangaben, insbesondere zur Verfügbarkeit. Wobei es sein kann, dass Nutzer*innen auch nur zur Bibliographie suchen und in dem Moment die Verfügbarkeit irrelevant ist.
- Im Moment ist es, insbesondere bei Artikeln, so, dass die Verfügbarkeit erst nach der Suche über Linkresolver abgefragt werden kann. Idealer wäre es, wenn diese bereits im Index vorhanden wäre und darauf gefiltert werden könnte. Allerdings könnte es dann auch dazu kommen, dass Nutzer den Weg des wenigstens Widerstandes gehen und nur Publikationen nutzen auf die direkter Zugriff besteht. Auch Linkresolver benötigen eine gut gepflegte Knowledgebase.
- Die Nutzbarkeit von Facetten basiert auch auf der Vollständigkeit der Datengrundlage.
- Visualisierungen, z.B. die Wege zum Regal aufzuzeigen, sollten angeboten werden.

Nutzerverhalten

- Nutzer*innen sind subjektiv mit Discovery-Systemen deutlich zufriedener als mit klassischen OPACs, allerdings ist der Grund unklar. Möglichkeiten:
 - höhere Usability
 - bessere (modernere) Optik
 - viel mehr Daten
- Veränderung des Nutzerverhaltens: die "Generation Google" will keine klassischen OPACs mehr, sondern intuitiv bedienbare Discovery-Systeme.

VIELEN DANK FÜR IHR INTERESSE UND IHRE BEITRÄGE