

Visualisierung von bibliografischen Metadaten als zusätzlichen Einstieg in die Recherche nach Medien in einem Bibliothekskatalog

Exposé zur Masterarbeit

Peter Breternitz
TH Wildau, Wildau Institute of Technology
Bibliotheksinformatik
peter.breternitz@th-wildau.de

I. EINFÜHRUNG

Gedächtnisinstitutionen wie Bibliotheken verfügen über einen Schatz hochqualitativer bibliografischer Metadaten. Diese werden in klassischen Online-Katalogen oder in Discovery-Systemen zur Anzeige gebracht. Neben einer Facettierung, mit der die gefundenen Ergebnisse einer Suchanfrage fein justiert werden können, scheinen Discovery-Systeme hinsichtlich der Visualisierung von Suchergebnissen nicht mehr zu bieten als die klassischen Online-Kataloge. Vor der Folie des Semantic Webs und dessen Technologien wie zum Beispiel RDF wäre die Frage hier, ob Bibliotheken alle Potentiale ihrer reichhaltigen Metadaten für die Darstellung von Suchergebnissen ausschöpfen? Oder anders: gibt es Alternativen in der Human-Computer-Kommunikation, die das explorative Suchen der NutzerInnen fördern und Ergebnisse zu neuen Wissensräumen zusammenschließen und Erkenntnisse zu Tage bringen, die in der klassischen Suchanzeige verloren gehen beziehungsweise nicht angezeigt werden können.

Mit diesen Topoi beschäftigt sich die zu entstehende Arbeit in einem ersten theoretischen Schritt. Nachfolgend soll in einem praktischen Teil ein proof-of-concept anhand bibliografischer Metadaten einer Gedächtnisinstitution entwickelt werden, dass als Ergebnis einen interaktiven, explorativen und virtuellen Zugang zu den Beständen liefert. Dabei geht es um ein zusätzliches Angebot an die WissenschaftlerInnen. Darüberhinaus könnte das Tool als Werkzeug für den Bestandsaufbau hinsichtlich eines ausgeglichenen Wachstums zum Beispiel der Bestandsgruppen fungieren.

II. PROBLEMSTELLUNG

Die Recherche in einem Bibliothekskatalog oder Discovery-System¹ nach Ressourcen kann über mehrere Sucheinstiege wie Titelstichwörter, Sachschlagwörter, AutorInnen oder über den gesamten Index passieren. Die zurückgelieferte Ergebnismenge auf eine Suchanfrage entspricht meist einer Liste, die nach voreingestellten Kriterien sortiert ist. Discovery-Systeme bieten Facettierung an, um die Ergebnismengen im nachhinein einzuschränken. Die Listenform repräsentiert nicht wie die einzelnen Facetten aufeinanderbezogen sind oder wie die Ressourcen selber verbunden sind untereinander. Informationsräume Searching-Browsing Concept
Interlinked-Ressources
Visuelle Pfade
Verknüpfen von Informationen
Leichter fällt die Suchergebnisse zu überblicken
Schlagwortesuche über das Highlighten der Worte hinaus geht Discovery-System. Darstellung der Ergebnisse - Grenzen der Darstellung. Visualisierung kann die Aussagekraft von Bibliotheksdaten erhöhen. Es werden neue Zusammenhänge, Trends in der Visualisierung sichtbar, die in der klassischen Darstellung verloren gehen [Phe20]. Wenn mehrere Informationsstrukturen parallel bearbeitet werden, können visuelle Übergänge zwischen ihnen helfen, den Kontext zu erhalten. Klassifikatorische Sacherschließung, Stöbern nach ähnlichen Medien, das Gruppieren nach inhaltlichen Gesichtspunkten. Visualisierungen und Textvisualisierungstechniken sind Bereiche, die in den letzten drei Jahrzehnten gewachsen sind und eine Vielzahl an Anwendungen hervorgebracht haben [FHKM17]. Im bibliothekarischen bzw. informationswissenschaft-

¹Discovery-Systemen operieren als One-Stop-Shop-Systeme und basieren auf der Suchmaschinentechnologie. Sie zeichnen sich unter anderem durch einen großen Suchraum, eine intuitive Bedienbarkeit und einem Ranking der Treffer nach Relevanz aus.

lichen Bereich sind hier VisualBib, PivotPath und SeeCollections zu nennen. Visualbib wertet Buchkapitel, Papers, Bücher und Conference sowie Workshop Article von verschiedenen Datenbanken aus, stellt diese graphisch in *narrativ views* dar und hilft den WissenschaftlerInnen dabei Bibliographien zu erstellen [DC]. PivotPath ist eine interaktive Visualisierung zur Erkundung von Informationsressourcen wie Artikeln dar [DRRD12]. SeeCollections ist eine Web-Application zur visuellen Darstellung von Bibliotheksteilbeständen nach Klassifikation, Thema und nach Erscheinungsjahr [Eat17].

III. THEORETISCHER RAHMEN

Hier sollen theoretische Annahmen und Konzepte kurz eingeführt und erklärt werden. Aus bibliothekarischer Perspektive werden in einem Teil wichtige Konzepte wie das Semantic Web und dessen Technologien erklärt. [DP13] Stichpunkte und Fragen für diesen *Theorie-I-Teil* sind:

Semantic Web, FRBR, LRM, RDF, LOD, Warum sind Metadaten von Bibliotheksbeständen besonders für Visualisierungen geeignet? Catalogue Enrichment. Für den zweiten Theorieteil werden Schlaglichter auf Data Science, Digital Humanities geworfen. Dabei spielen auch die verwendeten Technologien eine Rolle. Stichpunkte sind für diesen *Theorie-II-Teil*:

Visualisierungen, Grenzen und Möglichkeiten, Interaktive Grafiken, Data Science Life Cycle Definieren was Visualisierung ist Visualisierung bezeichnet Disziplinen wie Informationsvisualisierung, Datenvisualisierung und visuelle Analyse. Dabei wird sich konzentriert auf Ansätze die Visualisierungen algorithmisch aus den Metadaten der Dokumenten erzeugen.

IV. LITERATURÜBERSICHT

Die Literaturdiskussion führt die beiden Theoriestränge aus dem vorhergehenden Kapitel in den *Theorie-III-Teil* zusammen und diskutiert anhand von Literatur aus der bibliothekarischen Praxis bereits vorhandene Beispiele. Dabei wird hierbei auch die Frage der technischen Umsetzung eine Rolle spielen, aber auch auf insbesondere eventuelle erstellen hingewiesen [Eat17], [Mur15]. RVK-Visual <https://github.com/bvb-kobv-allianz/RVK-VISUAL>

V. EIGENER ANSATZ UND BERÜCKSICHTIGTE METHODEN

Entwicklung Proof-of-Conceptes

VI. ABRISS

Ein grober Abriß der Masterarbeit bezüglich des Inhaltes und des Umfangs ist im Folgendem angegeben:

Einleitung (5 S.):

Tool mit dem man Suchergebnisse visualisieren kann
Tool für die kontrollierte Bestandsentwicklung
Übersicht über die Kapitel - 1- 4 Sätze, was in den

nächsten Kapiteln dargelegt wird

Darlegung warum es sinnvoll ist

Theoretische Teil I (10 S.):

Semantic Web Technologien

FRBR, LRM, RDF, Linked Open Data, Catalogue Enrichment

Ausloten der Möglichkeiten, Warum eignen sich Metadaten aus Bibliotheken und anderen Kultureinrichtungen besonders

Theoretische Teil II (15 S.):

Grenzen und Möglichkeiten von Datenvisualisierungen vor dem Hintergrund Data Science und Digital Humanities

Interaktive Visualisierungen

Theoretische Teil III (15 S.):

Literaturdiskussion Bibliotheken - Metadaten - Visualisierungen

Beispiele - Technologien

Praktischer Teil (15 S.):

Entwicklung eines proof-of-Conceptes anhand von Metadaten einer Bibliothek oder einer Kulturinstitution Aufzeigen des Workflows

Theoretisches Modell der Data Science anwenden auf das Projekt

praktische Methoden - Wie und mit was programmieren

Auswertung welcher Metadatenfelder

Catalogue Enrichment

LITERATUR

- [ALK14] Jae-wook Ahn, Xia Lin, and Michael Khoo. *Dewey Decimal Classification Based Concept Visualization for Information Retrieval*, volume 1311. 2014.
- [Cai16] Alberto Cairo. *The truthful art : data, charts, and maps for communication*. New Riders, Place of publication not identified, 2016.
- [CD] Marco Corbato and Antonina Dattolo. A web application for creating and sharing visual bibliographies. In Alejandra González-Beltrán, Francesco Osborne, Silvio Peroni, and Sahar Vahdati, editors, *Semantics, Analytics, Visualization*, pages 78–94. Springer International Publishing.
- [DC] A. Dattolo and M. Corbato. Visualbib: Narrative views for customized bibliographies. In *2018 22nd International Conference Information Visualisation (IV)*, pages 133–138.
- [DP13] Patrick Danowski and Adrian Pohl. *(Open) Linked Data in Bibliotheken*. Bibliotheks- und Informationspraxis., De Gruyter Saur, Berlin ; Boston, 2013.
- [DRRD12] M. Dörk, N. Henry Riche, G. Ramos, and S. Dumais. Pivotpaths: Strolling through faceted information spaces. *IEEE Transactions on Visualization and Computer Graphics*, 18(12):2709–2718, 2012.
- [Eat17] Mark Eaton. Seeing library data: A prototype data visualization application for librarians. *Journal of Web Librarianship*, 11(1):69–78, 2017.
- [EK12] G. Morton-Owens Emily and L. Hanson Karen. Trends at a glance: A management dashboard of library statistics. *Information Technology and Libraries*, 31(3), 2012.
- [FHKM17] P. Federico, F. Heimerl, S. Koch, and S. Miksch. A survey on visual approaches for analyzing scientific literature and patents. *IEEE Transactions on Visualization and Computer Graphics*, 23(9):2179–2198, 2017.

- [KK] K. Kucher and A. Kerren. Text visualization techniques: Taxonomy, visual survey, and community insights. In *2015 IEEE Pacific Visualization Symposium (PacificVis)*, pages 117–121.
- [LGS06] Bénédicte Le Grand and Michel Soto. *Topic Maps, RDF Graphs, and Ontologies Visualization*, pages 59–79. Springer London, London, 2006.
- [Mur15] Sarah Anne Murphy. How data visualization supports academic library assessment: Three examples from the ohio state university libraries using tableau. *2015*, 76(9):5, 2015.
- [Nev] Zuzana Nevěřilová. Metadata visualization in digital libraries. *Research and Advanced Technology for Digital Libraries*, pages 442–445. Springer Berlin Heidelberg.
- [Phe20] Eric Phetteplace. Effectively visualizing library data. *Reference & User Services Quarterly*, 52(2), 2012-12-20.
- [vEW14] Nees Jan van Eck and Ludo Waltman. Citnetexplorer: A new software tool for analyzing and visualizing citation networks. *Journal of Informetrics*, 8(4):802–823, 2014.
- [Yau13] Nathan Yau. *Data points : visualization that means something*. John Wiley & Sons, Inc., Indianapolis, IN, 2013.