

Datengetriebenes Unterstützungssystem für zukünftige Etatplanung und Mittelallokation einer hybriden Spezialbibliothek

Exposé zur Masterarbeit

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

I. EINFÜHRUNG

Ausgehend von ökonomischen, informationstechnologischen und marktpolitischen Einschnitten in den letzten zwei Jahrzehnten¹, sind Bibliotheken dazu veranlasst, ihr Budget hinsichtlich der Informationsbedarfe ihrer Nutzer:innen behutsamer zu planen und sich in zunehmenden Maß gegenüber ihren Unterhaltsträgern zu rechtfertigen.

Die Relevanz von bibliothekarischen Kennzahlen ist in diesem Zusammenhang größer geworden. Deswegen ist es wichtig Daten aus bibliothekarischen Servicedienstleistungen und Geschäftsprozessen zu aggregieren, zu erheben und statistisch auszuwerten, um auf Basis der daraus erzielten Erkenntnisse handeln zu können.

Ziel der zu entstehenden Arbeit ist die Entwicklung einer interaktiven Business-Intelligence-Applikation als proof-of-concept, mit der systematisch die relevanten Daten einer hybriden Spezialbibliothek aggregiert, statistisch analysiert und mit geeigneten und modernen Datenvisualisierungstechniken² ausgewertet und ausgegeben werden. Vor allem soll sich hier auch auf automatisierte Prozesse zur Gewinnung der Ergebnisse konzentriert werden.

Mit diesen automatisch angefertigten statistischen Datenanalysen sollen zukünftige Entscheidungen im Bibliotheksmanagement wie Erwerbspolitik, Budgetplanung und Mittelallokation hinsichtlich der weite-

ren Entwicklung der Servicedienstleistungen evidenzbasiert und datengetrieben unterstützt werden.

Darüberhinaus soll die Applikation eine Funktion beinhalten, ausgewählte Resultate automatisiert als *factsheet* zu exportieren, um diese als Rechenschaftsbericht gegenüber stakeholdern der Bibliothek präsentieren und verhandeln zu können.

Die Notwendigkeit für diese Applikation ist durch das Fehlen eines zentralen Nachweisortes für bibliothekarische Statistiken in der Bibliothek gegeben. Da sie zudem verschiedene Recherche-Systeme den Wissenschaftler:innen anbietet, wäre eine Engführung bezüglich der statistischen Datenerhebung auf eine Plattform begrüßenswert. Des Weiteren ist das Erfordernis bibliothekarische Geschäftsprozesse zu evaluieren und die Servicedienstleistungen bezüglich der Ziele der Institution noch weiter zu optimieren von großer Relevanz. Mit dem Ende der Konsolidierungsphase der Bibliothek, die im Zuge des *Max-Planck-Institutes für empirische Ästhetik* 2014 gegründet wurde, tritt sie ein in eine Phase, in der ab dem Jahr 2021 Budgetplanungen eine größere Rolle spielen werden.

Die zu entstehende Applikation soll hierbei helfen systematisches Controlling einzuführen und das Bibliotheksmanagement weiter zu professionalisieren. Um künftigen Anforderungen gewachsen zu sein, soll sie modulbasiert programmiert werden und dadurch leicht erweiterbar und/oder für andere Bibliotheken nachnutzbar sein.

II. LITERATURDISKUSSION

Ein gutes Beispiel für ein datengetriebenes Unterstützungssystem findet sich in der Literatur bei Spielberg, der sich mit dem Thema der Bestandspflege an der Universitätsbibliothek Essen befasst und eine Applikation (weiter)entwickelt hat, die die Fachreferent:innen bei der Aussonderung und Erwerbung unterstützt.[9] Ebenso finden sich andere

¹Zu nennen wären hier die Explosion der Zeitschriftenpreise im STM-Bereich, das Aufkommen von Epublishing und die Konzentration auf wenige Verlage

²Visualisierung bezeichnet Disziplinen wie Informationsvisualisierung, Datenvisualisierung und visuelle Analyse. Dabei wird sich konzentriert auf Ansätze die Visualisierungen mittels Visualisierungstechniken algorithmisch aus Daten erzeugen, dargestellt werden können. Visualisierungen können komplexe Sachverhalte herunterbrechen und können so große Datenmengen im Gegensatz zu großen Tabellen leicht verständlich darstellen.[5]

Beispiele, die vorrangig anhand einzelner Fragestellungen hinsichtlich der Bestandsentwicklung[2] oder anderer bibliothekarischer Servicedienstleistungen[4, 3, 6] verschiedene statistische Analysen vollzogen und diese visualisiert haben. Eine Ausnahme bildet die Entwicklung eines Dashboards an der *New York University Health Sciences Libraries*, das versucht möglichst viele Metriken aus bibliothekarischen Dienstleistungen aufzunehmen.[7] Alle Projekte sind meist an größeren akademischen Universitätsbibliotheken mit ganz unterschiedlichen softwaretechnischen Herangehensweisen[1, 11] und Zielen[8] entstanden.

Dennoch fehlen in der gesicherten Literatur Teile die sich mit den Budgetierung explizit befassen und Auskunft über Mittelallokation geben.

Zudem fehlt ein Beispiel in der Literatur, das holistisch alle relevanten Daten, die aus den Geschäftsgängen und Servicedienstleistungen insbesondere einer Spezialbibliothek entstehen, in einem Dashboard aggregiert und auf diesen automatisch statistische Analysen ausführt und diese mit modernen Visualisierungstechniken interaktiv darstellt.

III. PROBLEMSTELLUNG

Die Bibliothek des *Max-Planck-Institutes für empirische Ästhetik* ist Teil des *HeBis-Verbundes*. Die Geschäftsprozesse der Erwerbung und der Katalogisierung findet im *CBS* und im *LBS4* von *OCLC* statt. Das lokale Bibliothekssystem wird gehostet und betreut vom Lokalsystem-Team Frankfurt. Über die Bereitstellung besonderer Funktionalitäten seitens des Lokalsystem-Teams, erhält die Bibliothek unter anderem Ausleih-, Budget- und Umsatzübersichten per email als Text von diesem.

Neben der Verankerung in der deutschen Bibliotheksverbundlandschaft wird die Bibliothek in ihren Aufgaben von der *Max-Planck-Digital-Library* (mpdl) unterstützt. Deren Portfolio umfasst vorrangig die zentrale Lizenzierung von relevanten elektronischen Informationsressourcen, die Bereitstellung von Softwarelösungen, das Betreiben eines Publikationsrepositoriums und das Vorantreiben der OpenAccess-Initiative. Zudem stellt sie noch COUNTER-Statistiken, die sie von den Verlagen bekommt zur Verfügung.

Außer den bereitgestellten Daten, erhebt die Bibliothek u.a. ebenfalls Daten über die Frequentierung des Lesesaals, die Nutzung des nehmenden Fernleihservices, des Dokumentenlieferdienstes *subito* und des Bestandswachstums vor Ort. Nach den Verantwortlichkeiten aufgeteilt, passiert dieses an verschiedenen virtuellen Orten. Eine systematische Auswertung der erhobenen Daten findet nur unzureichend statt. Daher regt sich der Wunsch seitens der Mitarbeiter:innen nach einem gemeinsamen Tool, womit übersichtlich und klar alle notwendigen nutzungs- und sammlungsbezogenen

Statistiken für eine Spezialbibliothek erfasst und dargestellt werden können.³

Es gibt eine Vielzahl von auf Business Intelligence Software basierenden kommerziellen Lösungen, die auf den Bibliotheksbereich zu geschnitten sind. Zu nennen wären *AlmaAnalytics* von *ExLibris* für das NextGenerationSystem *Alma*⁴, *BibControl* von *OCLC*⁵, *CollectionHq* von *Baker & Talor*⁶ von oder *Libinsight* von *SpringShare*⁷. Darüber hinaus gibt es Business Intelligence Applikationen, die von Bibliotheken für Reporting, Datenanalyse und Datenvisualisierung adaptiert werden wie zum Beispiel *Tableau* von der Firma *Tableau Software* oder *Crytal Reports* von *SAP*. Diese Applikationen sind entweder an bestimmte Bibliothekssysteme zurückgebunden, limitiert in ihren Funktionen[10] oder zu generisch. Ferner wird sowohl von *HeBis* bzw. von der LokalBibliotheksSystem-Betreuung als auch von der *mpdl* keine Applikation in dieser Richtung angeboten. Ebenso ist ungewiss, wann die Ablösung des schon betagten CBS/LBS hinzu einem neuen Next-Generation-Library-System im *Hebis-Verbund* stattfindet und ob es Modul zur statistischen Datenerhebung liefert.

IV. INHALTLICHER ABRISS DER MASTERARBEIT

Im Folgenden wird ein skizzenhafter Überblick gegeben, welche Punkte die theoretische Konzeption und praktische Umsetzung beinhalten könnten.

Theoretische Vorarbeit

- Identifizierung der sammlungs- und bestandsbezogenen Geschäftsgänge und Servicedienstleistungen und der in dem Zusammenhang stehenden Daten.
- Sind die Daten ausreichend, um auf ihnen sinnvolle Analysen vollführen zu können? In welchem Format sollen die Daten vorliegen für die statistischen Analysen? Wie müssen diese aufbereitet bzw. bereinigt werden?
- Wie sollen die Daten letztendlich vorgehalten werden?
- Wie sieht es mit Anbinde an LBS-Datenbank für die Abfrage von Live-Daten aus?
- Identifizierung der statistischen Methoden. Welche Fragen werden an die Daten gestellt, welche Ergebnisse sollen diese liefern.
- Wahl des Frameworks für die Applikation/Dashboard. Begründung warum?

³Zwar führt HeBis eine Bestandstatistik, diese ist aber insbesondere für die Evaluation und Optimierung von Geschäftsprozessen einer Spezialbibliothek insuffizient. https://www.hebis.de/de/1ueber_uns/statistik/cbs_statistik.php

⁴<https://www.exlibrisgroup.com/products/alma-library-services-platform/alma-analytics> Stand: 27. Mai 2020

⁵<https://www.oclc.org/de/bibcontrol.html> Stand: 27. Mai 2020

⁶<https://www.collectionhq.com/> Stand: 27. Mai 2020

⁷<https://springshare.com/libinsight/> Stand: 27. Mai 2020

- Identifizierung geeigneter grafischer Visualisierungen. Welche Interaktionen soll die Applikation bieten.
 - Identifizierung und Beschreibung der funktionalen und nichtfunktionalen Anforderungen der Applikation - Lastenheft
 - praktische Überlegungen hinsichtlich der modularen Aufteilung in der Programmierung.
- [10] Golas Ulrike. Statistische Abfragen mit Alma für die Fachreferatsarbeit. *o-bib. Das offene Bibliotheksjournal / Herausgeber VDB*, 5(4), 2018.
 - [11] Laura K. Wiegand and Bob Humphrey. Visualizing Library Statistics using Open Flash Chart 2 and Drupal. *Code4Lib*, (19), 2013.

Praktische Umsetzung

- Umsetzung des Lastenhefts
- modulares Programmieren für die einzelnen Geschäftsgangsdaten/Dienstleistungen
- Programmierung für die automatische Weiterverarbeitung der Daten
- Programmierung der Algorithmen für die statistischen Analysen
- Programmierung der Datenvisualisierungstechnologien

LITERATUR

- [1] J. L. Finch and A. R. Flenner. Using data visualization to examine an academic library collection. *College and Research Libraries*, 77(6):765–778, 2016.
- [2] Michael Hughes. A Long-Term Study of Collection Use Based on Detailed Library of Congress Classification, a Statistical Tool for Collection Management Decisions. *Collection Management*, 41(3):152–167, 2016.
- [3] Jennifer E. Kniewel, Heather Wicht, and Lynn Silipigni Connaway. Use of Circulation Statistics and Interlibrary Loan Data in Collection Management. 2006, 67(1):15, 2006.
- [4] Atalay Kutlay and Cal Murgu. Shiny Fabric: A Lightweight, Open-source Tool for Visualizing and Reporting Library Relationships. *Code4Lib*, (47), 2020.
- [5] Yuvaraj Mayank. Infographics: tools for designing, visualizing data and storytelling in libraries. 34(5):6–9, 2020/05/26 2017.
- [6] Andy Meyer. Using R and the Tidyverse to Generate Library Usage Reports. *Code4Lib*, (39), 2018.
- [7] Emily Morton-Owens and Karen Hanson. Trends at a Glance: A Management Dashboard of Library Statistics. *Information Technology and Libraries*, 31, 2012.
- [8] Eric Phetteplace. Effectively Visualizing Library Data. *Reference & User Services Quarterly*, 52(2), 2012-12-20.
- [9] Eike T. Spielberg. *Der FachRef-Assistent : personalisiertes, fachspezifisches und transparentes Bestandsmanagement*. Thesis, 2017.