Brockhaus 3.0

Zum Wandel der klassischen Gatekeeper-Funktion.

Von Christopher Koska

2008 trennte sich das Bibliographische Institut von der „F.A. Brockhaus AG". Die „Bertelsmann SE & Co. KGaA" übernahm die Rechte an der Marke Brockhaus und übergab die verlegerische Verantwortung für die lexikalischen Inhalte dem konzerneigenen Verlag „wissenmedia" in der „inmediaONE] GmbH" (Brockhaus Wissensservice 2014). „Wegen mangelnder wirtschaftlicher Perspektive" verkündete am 11. Juni 2013 auch die Geschäftsleitung der „inmediaONE] GmbH" die schrittweise Schließung des Unternehmens und des zugehörigen Verlags „wissenmedia" (Bertelsmann 2013). Das Medienecho hallte ein weiteres Mal homofon durch die Gazetten: „Der Brockhaus ist tot" (Frankfurter Allgemeine Zeitung 2014), „Brockhaus beerdigt seine Enzyklopädie" (Handelsblatt 2009) oder „Brockhaus knickt vor Wikipedia und Google ein" (Die Welt 2008). Der Journalist und Literaturkritiker Hellmuth Karasek fasste in seiner Kolumne für das „Hamburger Abendblatt" zusammen: „Nun also stirbt auch der Brockhaus, das letzte deutsche Konversationslexikon, sperriges Wissen [...], dessen Lebenszeit verrinnt wie in einer analogen Sanduhr, dem Zeichen des Sterbens vor dem digitalen Zeitalter. Kein Wunder: In das fast federleichte Smartphone passt zentnerschweres Wissen, überall, wo Empfang ist, abrufbar" (Hamburger Abendblatt 2013).

Insofern sich der Begriff Nachschlagen auf das Aufschlagen von Büchern bezieht, ist es korrekt, dass das Zeitalter des Nachschlagens vorüber ist (HfPh 2014). Die Inhalte des Brockhaus wurden allerdings schon vom Bibliographischen Institut in Mannheim digitalisiert und über die Jahre hinweg umfangreich mit zusätzlichen Metadaten angereichert. Diese für den Nutzer nicht sichtbaren Attribute in den Artikeln vereinfachten zunächst den Redakteuren die Pflege und die Aktualisierung der Inhalte. Über das Content-Management-System der

Christopher Koska, M.A., ist IT-Projektleiter für die „wissenmedia" in der „inmediaONE] GmbH" und wissenschaftlicher Mitarbeiter am Lehrstuhl für Medienethik an der Hochschule für Philosophie in München.

https://doi.org/10.5771/0010-3497-2015-1-65 https://www.inlibra.com/de/page/agb - Open Access - (cc) BY

Redakteure konnte der gesamte Brockhaus schnell nach verschiedenen Fachgebieten, Zeitepochen oder Geobereichen gefiltert werden, um diese dann den entsprechenden Experten zugänglich zu machen. Neben der inhaltlichen Auszeichnung durch Metadaten wurde der XML-Quelltext des Brockhaus auch formal äußerst umfangreich vom Bibliographischen Institut ausgezeichnet.[1] Diese gestalterischen Strukturelemente wurden durch die Berücksichtigung digitaler Brockhaus-Versionen (CD und DVD seit 2002, mobile Endgeräte seit 2004, USB-Stick und Online-Portal seit 2005) zunehmend komplexer. Über ein Variantenkonzept für unterschiedliche Medienformate mussten verschiedene Artikel-Elemente in ein und derselben Quellsubstanz parallel gepflegt werden.

Mit der Übernahme des Brockhaus 2009 durch „wissenmedia" wurde die Brockhaus-Substanz zunächst in die bestehende IT-Infrastruktur der Bertelsmann Lexika portiert und mit Hilfe statistischer Verfahren (Text- und Data Mining) an die hauseigene Topic Map[2] angebunden. Das Konzept der Topic Map erweiterte die semantische Auszeichnung der Lexikonartikel, indem Redakteure substanzübergreifend nach Stichworttypen (Personen, Ereignissen u.v.m.), geografischen Metadaten (latitude, longitude), personenbezogenen Metadaten (wie Geschlecht, Geburts- und Sterbedatum) oder verschiedenen Zeitbereichen suchen konnten. Das Variantenkonzept für die verschiedenen Medienformate wurde hingegen aufgegeben, da es sowohl die redaktionelle Arbeit, als auch die Medienproduktion unnötig belastete. Zur weiteren Optimierung des redaktionellen Aufwands wurden die Brockhaus-Inhalte zudem (über die gemeinsame Topic Map) an einem übergeordneten Priorisierungs- und Aktualisierungsindex angebunden.

Hinter der Kulisse von Ledereinband und Goldschnitt haben sich die redaktionellen Anforderungen an den „Brockhaus" in den Digitalbereich verschoben.

Hinter der Kulisse von Ledereinband und Goldschnitt haben sich die redaktionellen Anforderungen in der über 200-jäh-

1 *Die formale Struktur ermöglicht die strukturelle Darstellung der Inhalte (analog zu HTML für digitale Inhalte im WWW). Klassische Strukturelemente lexikographischer Nachschlagewerke sind die Zugriffsstruktur, die Verweisstruktur und die Rahmenstruktur (vgl. Bergenholtz/Tarp/Wiegand 1998, S. 1889-1909).*

2 *Eine Topic Map ist ein semantisches Netz zur Wissensorganisation. Im Wesentlichen werden darin Begriffe und ihre Beziehungen zueinander beschrieben (vgl. Reimer 2013, S. 172-182).*

https://doi.org/10.5771/0010-3497-2015-1-65 https://www.inlibra.com/de/page/agb - Open Access - (cc) BY

rigen Geschichte des Brockhaus also längst in den Digitalbereich verschoben. Die IT-Infrastruktur war jedoch weiterhin auf die Printproduktion spezialisiert. Deshalb gab Bertelsmann 2009 ein Forschungsprojekt frei, um die breit gefächerten Inhalte von „wissenmedia" (Brockhaus, Bertelsmann Lexika und Chroniken, Wahrig, u.v.m.) in eine hochperformante, skalierbare Wissens-Infrastruktur für die digitale Produktentwicklung zu überführen. Der aus diesem Projekt entstandene „wissensserver" ermöglichte erstmals eine atomare Auslieferung der Inhalte in Echtzeit (Webzugriff über REST-Schnittstelle) und ohne zusätzliche Zwischenschicht. Bis dato notwendige Zwischenschritte für die digitale Medienproduktion, wie die Analyse der formalen Struktur, die Konzeption eines geeigneten Datenbankmodells (für das digitale Produkt), die Spezifikation von Austauschformaten (SGML, XML, HTML, Jason…), die Dokumentation und Transformation der Metadaten (Elemente und Attribute) in das Zielformat, der Export aus dem Redaktions-Content-Management-System und der Import in neue Datenbanken und Retrieval-Systeme, bis hin zur Transformation der Suchergebnisse und der Artikeldarstellung waren so nicht mehr notwendig. Die unterschiedlichsten Plattformen (inklusive Smartphones oder auch Smart TVs) konnten nun „on-the-fly" mit insgesamt 1,6 Millionen Artikeln versorgt werden (One Source für Multi-Channel).

Gedächtnisarchive, wie Museen und Bibliotheken, digitalisieren ihre Objekte und Inhalte. Sie datafizieren sie.

Auf der Grundlage dieser Technologie beteiligte sich „wissenmedia" ab Februar 2013 an dem mehrjährigen EU-Forschungsprojekt „EEXCESS".[3] Entstanden ist dieses Projekt vor dem Hintergrund der anwachsenden Datenberge, die durch europäische Initiativen in den letzten zehn Jahren gefördert wurden, um kulturelle und wissenschaftliche Informationen öffentlich zugänglich zu machen: Gedächtnisarchive, wie Museen und Bibliotheken, digitalisieren ihre Objekte und Inhalte. Sie datafizieren sie.[4] Die klassische Aufgabentrennung zwischen Inhalt und Technik verändert sich dadurch zunehmend. Die tra-

3 *„Enhancing Europe's eXchange in Cultural Educational and Scientific Resources": Mehrjähriges EU-Forschungsprojekt gefördert durch das 7. Forschungsrahmenprogramm (FP7/2007-2013), Grant Agreement n° 600601. www.eexcess.eu (zuletzt aufgerufen am 20.2.2015).*

4 *Erklärungsvideo zu „Linked Open Data" von Europeana https://www.youtube.com/watch?v=uju4wT9uBIA (zuletzt aufgerufen am 24.2.2015).*

https://doi.org/10.5771/0010-3497-2015-1-65 https://www.inlibra.com/de/page/agb - Open Access - (cc) BY

ditionelle Gatekeeper-Funktion (GF), die zunächst in den Händen von Verlagen, Journalisten und Redakteuren lag (GF 1.0), verschiebt sich – über die kollaborativen Autorengemeinschaften des Mitmachwebs (GF 2.0) – auf die informationstechnische Ebene der Datenaggregation (GF 3.0). Aufgrund des homogenen und gut strukturierten Aufbaus verfügt der Brockhaus über großes Potential, um einen qualitativen Relevanzrahmen für die maschinelle Informationsverarbeitung bereitzustellen.[5] Er entzieht sich sukzessive unserem Blickfeld, wenn die technischen Komponenten hinter die Benutzeroberfläche zurücktreten. Tot ist er aber noch immer nicht.

Von November 2014 bis Januar 2015 wurden die ersten EEXCESS-Prototypen evaluiert, unter anderem von Studenten, die die Empfehlungstechnologie bei der Erstellung einer wissenschaftlichen Arbeit nutzten. In einem Szenario wurde über ein „Google-Docs-Plug-In" der von den Studenten verfasste Text automatisch analysiert und im Hintergrund eine Suchanfrage an die Empfehlungstechnologie geschickt. Anschließend wurden über eine Sidebar in der rechten Randspalte des Dokuments zusätzliche Inhalte von bereits verfügbaren Informationsquellen (u.a. Zentralbibliothek für Wirtschaftswissenschaften, Mendeley und Europeana) proaktiv eingeblendet. Die Personalisierung und Kontextualisierung der Suchergebnisse erfordert 1) einen Eingriff in die Privatsphäre der Nutzer, da Nutzerprofile erstellt werden, um möglichst relevante Inhalte anbieten zu können und 2) einen Eingriff in die informationelle Selbstbestimmung der Anwender, da sich die Selektionsmechanismen unter der Motorhaube der grafischen Benutzeroberfläche befinden. Ein Forschungsschwerpunkt des Projekts besteht deshalb auch darin, a) diese Profile zu verschleiern, d.h. sie zu anonymisieren, bevor sie an die Empfehlungstechnologie oder die Content-Anbieter geschickt werden (Hasan et al. 2013) und, um das Eintreten in eine selbstreferentielle Filterblase zu reduzieren (Filipović 2013), ebenso darin, b) die gesammelten Daten für den Anwender möglichst transparent und editierbar aufzubereiten.

Die Gatekeeper-Funktion, die in den Händen von Journalisten lag, verschiebt sich auf die informationstechnische Ebene der Datenaggregation.

5 *Eine anschauliche Algorithmenskizze für die automatische Sprachverarbeitung und für Verfahren des maschinellen Lernens (auf der Basis von Lexika und Wörterbücher) findet sich u.a. bei Hahn (2013, S.252-263).*

https://doi.org/10.5771/0010-3497-2015-1-65 https://www.inlibra.com/de/page/agb - Open Access - (cc) BY

Die ersten Testergebnisse haben gezeigt, dass die Qualität der Empfehlungen noch nicht den hohen Standards des Brockhaus entsprechen und zunächst auf der Ebene der Datenfusion nachgebessert werden muss. Die maschinelle Zusammenstellung der Inhalte wird zum einen von technischen Faktoren, wie der algorithmischen Analyse oder auch der Server-Antwortzeit der Content-Lieferanten, beeinflusst, zum anderen aber ebenso von der semantischen Zuordnung der Metadatenfelder, die in aller Regel von den Institutionen selbst organisiert wird. Beispielsweise verschlagwortete das Science Museum „Webstühle" (engl. loom) als „physikalische Objekte" (RDF-Typisierung „PhysicalObject")[6], das UCL Museum aber als „Bild" (RDF-Typisierung „Image")[7]. Menschen haben keine Schwierigkeiten mit der Schlussfolgerung, dass das UCL Museum das „(Ab)bild eines physikalischen Objekts" beschreibt. Für Maschinen wäre das bei einer facettierten Suchanfrage aber zunächst ein Ausschlusskriterium – das entsprechende Objekt würde also nicht berücksichtigt werden. Im Kontext der Gatekeeper-Funktion 3.0 stellt sich die Frage der Verantwortung, beispielsweise Informationen wahrheitsgetreu wiederzugeben (Pressekodex), in einem technisch und inhaltlich eng verwobenen Medienraum (vgl. Altmeppen 2015). Dies erfordert neue Standards und Regeln, aber auch zusätzliche Kompetenzen von allen Akteuren.

Literatur

Altmeppen, Klaus-Dieter (2015): Zur Handlungsnotwendigkeit im digitalen ubiquitären Medienraum. Aus dem Vortrag bei der Jahrestagung des Netzwerk Medienethik 2015: Vom instrumentellen Artefakt zum ubiquitären digitalen Mediensystem. Zur Veränderung des Verhältnisses von Technik und Journalismus und dessen Folgen für die journalistische Verantwortung. https://twitter.com/cxkoska/status/566201533327740928 (zuletzt aufgerufen am 20.2.2015).

6 *Science Museum, Loom, Type: PhysicalObject (Europeana Data Model), http://www.europeana.eu/portal/record/2022343/8A1A98A25C53A5BD5F285C5B026ACF83BB3D2E38.html?start=2&query=loom&startPage=1&qt=false&rows=24 (zuletzt aufgerufen am 20.2.2015).*

7 *UCL Museum, Loom, Type: Image (Europeana Data Model), http://www.europeana.eu/portal/record/2022350/90CFC376B420D8E9C93FB036C6F9019E1CFA637E.html?start=8&query=loom&startPage=1&qt=false&rows=24 (zuletzt aufgerufen am 20.2.2015).*

https://doi.org/10.5771/0010-3497-2015-1-65 https://www.inlibra.com/de/page/agb - Open Access - (cc) BY

Bergenholtz, Henning/Tarp, Sven/Wiegand, Herbert, Ernst (1998): Datendistributionsstrukturen, Makro- und Mikrostrukturen in neueren Fachwörterbüchern. In: Steger, Hugo/Weigand, Herbert (Hg.): Handbücher zur Sprach- und Kommunikationswissenschaft. Berlin/New York, S.1889-1909.

Bertelsmann (2013): Presseinfo zur schrittweisen Einstellung der inmediaONE] GmbH von Matthias Wulff, Leiter Unternehmenskommunikation Club- und Direktmarketinggeschäfte, am 11.6.

Brockhaus Wissensservice (2014): Brockhaus. https://wm.brockhaus-wissensservice.com/sites/brockhaus-wissensservice.com/files/pdfpermlink/brockhaus-617774a5.pdf (zuletzt aufgerufen am 20.2.2015).

EEXCESS | Enhancing Europe's eXchange in Cultural Educational and Scientific reSources (2013): Take the content tot he user. http://eexcess.eu/ (zuletzt aufgerufen am 24.2.2015).

Europeana (2012): Linked Open Data – What is it? https://www.youtube.com/watch?v=uju4wT9uBIA (zuletzt aufgerufen am 24.2.2015).

Europeana (o.J.): Search ‚loom' (Science Museum) (Search-Funktion) http://www.europeana.eu/portal/record/2022343/8A1A98A25C53A5BD5F285C5B026ACF83BB3D2E38.html?start=2&query=loom&startPage=1&qt=false&rows=24 (zuletzt aufgerufen am 24.2.2015).

Europeana (o.J.): Search ‚loom' (UCL Museum) (Search-Funktion) http://www.europeana.eu/portal/record/2022350/90CFC376B420D8E9C93FB036C6F9019E1CFA637E.html?start=8&query=loom&startPage=1&qt=false&rows=24 (zuletzt aufgerufen am 24.2.2015).

Filipović, Alexander (2013): Die Enge der weiten Medienwelt. Bedrohen Algorithmen die Freiheit öffentlicher Kommunikation? In: Communicatio Socialis, 46. Jg., H. 2, S. 192-208.

Hahn, Udo (2013): Methodische Grundlagen der Informationslinguistik. In: Kuhlen, Rainer/Semar, Wolfgang/Strauch, Dietmar (Hg.): Grundlagen der praktischen Information und Dokumentation. Berlin/Boston, S. 252-263.

Hasan, Omar et al. (2013): A Discussion of Privacy Challenges in User Profiling with Big Data Techniques: The EEXCESS Use Case. http://eexcess.eu/wp-content/uploads/2013/07/2013.07_BigData-2013-04.pdf (zuletzt aufgerufen am 20.2.2015).

Haustein-Tessmer, Oliver (2008): Brockhaus knickt vor Wikipedia und Google ein. In: WeltN24 vom 17.12. http://www.welt.de/wirtschaft/article2893879/Brockhaus-knickt-vor-Wikipedia-und-Google-ein.html (zuletzt abgerufen am 20.2.2015).

HfPh (2014): Zwischen Alphabet und Algorithmus - Das Ende des Nachschlagens und seine Wirkung auf unsere Sprache. Podiumsdiskussion im Rahmen der Veranstaltung „Deutsch 3.0" am 15.5.2014 an der Hochschule für Philosophie in München (HfPh). http://www.goethe.de/lhr/prj/d30/ver/de12147014v.htm (zuletzt aufgerufen am 20.2.2015).

https://doi.org/10.5771/0010-3497-2015-1-65 https://www.inlibra.com/de/page/agb - Open Access - (cc) BY

Karasek, Hellmuth (2013): Der letzte Brockhaus. In: Hamburger Abendblatt vom 15.6. http://www.abendblatt.de/vermischtes/kolumne/article117146790/Der-letzte-Brockhaus.html (zuletzt aufgerufen am 20.2.2015).

Lenz, Daniel (2011): Wir hauchen dem Brockhaus neues Leben ein: Interview mit dem ehemaligen Geschäftsführer von wissenmedia. In: Buchreport vom 29.1. http://www.buchreport.de/nachrichten/verlage/verlage_nachricht/datum/2011/01/29/wir-hauchen-dem-brockhaus-neues-leben-ein.htm (zuletzt aufgerufen am 20.2.2015).

Neller, Marc (2009): Brockhaus beerdigt seine Enzyklopädie. In: Handelsblatt vom 22.7. http://www.handelsblatt.com/unternehmen/it-medien/lexikon-brockhaus-beerdigt-seine-enzyklopaedie/3224592.html (zuletzt aufgerufen am 20.2.2015).

Platthaus, Andreas (2014): Der Untergang des Lexikons. In: Frankfurter Allgemeine Zeitung vom 6.7. http://www.faz.net/aktuell/feuilleton/buecher/themen/wikipedia-hat-brockhaus-beerbt-13028236.html (zuletzt aufgerufen am 20. 2.2015).

Reimer, Ulrich (2013): Wissensorganisation. In: Kuhlen, Rainer/Semar, Wolfgang/Strauch, Dietmar (Hg.): Grundlagen der praktischen Information und Dokumentation. Berlin/Boston, S. 172-182.

https://doi.org/10.5771/0010-3497-2015-1-65 https://www.inlibra.com/de/page/agb - Open Access - (cc) BY