

Geschichte aus erster Hand – Der Aufbau eines nationalen Zeitungsportals unter Berücksichtigung der Bedürfnisse verschiedener Nutzergruppen

Landes, Lisa

l.landes@dnb.de

Deutsche Nationalbibliothek, Deutschland

Dinger, Patrick

p.dinger@dnb.de

Deutsche Nationalbibliothek, Deutschland

Ziele und technische Grundlagen des Infrastrukturprojektes „Deutsches Zeitungsportal“

Historische Zeitungen wurden in den letzten Jahren von deutschen Kulturerbe-Einrichtungen verstärkt digitalisiert und zugänglich gemacht. Dadurch stehen der Forschung hunderte Millionen digitalisierter Zeitungsseiten zur Verfügung – ein Reichtum an Primärquellen, dem man mit den herkömmlichen geisteswissenschaftlichen Forschungsmethoden („Close Reading“) kaum gerecht werden kann. Zunehmend gewinnen daher Analysemethoden der Digital Humanities, wie z.B. „Distant Reading“ (Burckhardt u.A. 2018), an Bedeutung, um die bei der Massendigitalisierung entstandenen Daten auswerten zu können.

Die Aufgabe eines nationalen Zeitungsportals ist es jedoch, nicht nur für Forschende, sondern auch für eine interessierte Öffentlichkeit einen niedrighschwelligen Zugang zu entwickeln. Während mit ANNO, Delpher oder dem British Newspaper Archive im europäischen Raum mehrere Projekte – teilweise in Kooperation mit kommerziellen Partnern – entstanden sind, die große digitale Zeitungsbestände in einem nationalen Portal verbinden, existieren in Deutschland bislang nur lokale und regionale Portale einzelner Bibliotheken oder Regionen.¹ Ein übergreifendes Portal, das einen zentralen Zugriff bietet, ist bislang noch ein Forschungsdesiderat (Blome 2018: B.6-33). Dank verschiedener DFG-Förderinitiativen zur Digitalisierung historischer Zeitungen sowie der Förderung der Errichtung eines nationalen Zeitungsportals wird sich das nun ändern. Anfang 2019 wurde das Projekt „Deutsches Zeitungsportal“ gestartet. Es wird

im Rahmen der Deutschen Digitalen Bibliothek (DDB) umgesetzt. Ende 2020 soll das Portal, an dessen Aufbau vier Projektpartner² beteiligt sind, online gehen.

Das entstehende Portal wird eine auf die Besonderheiten von Zeitungen zugeschnittene zentrale Präsentations- und Rechercheoberfläche bieten und die folgenden Kernfunktionalitäten umsetzen:

- übergreifende Volltextsuchen in den digitalisierten Zeitungsbeständen
- unterschiedliche Einstiegspunkte, z.B. über Kalender und Zeitungstitel
- einen unmittelbar in die Portalumgebung integrierten Viewer für Volltexte und Images
- persistente Referenzierbarkeit und damit Zitationsfähigkeit der digitalen Zeitungsobjekte

Das Deutsche Zeitungsportal ist ein Infrastrukturprojekt, das auf den bestehenden Netzwerken, Techniken und Workflows der DDB aufbaut. Das heißt, die Zeitungen, die im Zeitungsportal zur Verfügung gestellt werden, stammen aus verschiedenen Einrichtungen, zumeist Bibliotheken, die Erschließungsinformationen, Bilddateien und Volltexte der Zeitungen aus ihren Beständen an das Zeitungsportal liefern. In der ersten Ausbaustufe des Zeitungsportals stellen Bibliotheken ihre Metadaten im METS/MODS-Format bereit, sodass über Verlinkungen in der Datenstruktur sowohl die Bilddateien (i.d.R. im jpg-Format) als auch die Volltexte (i.d.R. im ALTO-xml-Format) in das Portal übernommen werden. Dort werden diese Bestände zusammengeführt und können über einen Suchschlitz durchsucht werden. Basis der Suche bilden die Volltexte, welche vollständig in einen zentral vorgehaltenen SOLR-Index aufgenommen werden. Durch die standardisierten, maschinennutzbaren Inhalte des Zeitungsportals, die über eine Schnittstelle (API) heruntergeladen werden können, sollen neue Nutzungsszenarien, wie z.B. Big-Data-Analysen, ermöglicht werden (Altenhöner 2018: 146). Da es sich bei den Zeitungstexten und -bildern ausschließlich um rechtfreies Material handelt, können die Dokumente in der eigenen Forschung nachgenutzt und weiterverarbeitet werden.

Zum Start des Portals werden Bestände aus mindestens sechs deutschen Bibliotheken³ verfügbar sein; vorsichtigen ersten Schätzungen zufolge wird es sich um 250 verschiedene Zeitungstitel im Umfang von ca. 15 Mio. Zeitungsseiten handeln. Nach der Inbetriebnahme des Portals sollen die Inhalte beständig wachsen, um möglichst viele historische Zeitungen in diesem Portal zu vereinen. Das Fernziel des Unterfangens ist es also, den Forschenden (und der allgemeinen Öffentlichkeit) einen einheitlichen und stabilen Zugang zu vielen (wenn möglich: allen) historischen Zeitungen aus deutschen Kulturerbe-Einrichtungen zu geben. Zudem wird angestrebt, in einer zweiten Ausbaustufe ab 2021 das Zeitungsportal für andere Datenformate wie TEI-XML zu öffnen, um z.B. digitale

Editionen und Annotationen in den Korpus aufnehmen zu können.

Um Zeitungen aus vielen unterschiedlichen Quellen zusammenzuführen, sind bei der Aufbereitung der (Meta-)Daten vielfältige Standardisierungsprozesse notwendig. So wird im Rahmen des Projekts ein zeitungsspezifisches Anforderungsprofil für das METS/MODS-Metadatenformat entwickelt und eine Verknüpfung der Metadaten mit der Zeitschriftendatenbank (ZDB) umgesetzt, deren Identifier für die eindeutige Identifizierung der Zeitungstitel benutzt werden.

Durch den Aufbau der Datenverarbeitungsstrukturen für Zeitungen und die dafür nötigen Standardisierungsprozesse sollen auch Impulse für die Digitalisierung, Erschließung und Referenzierung der Zeitungsbestände der kooperierenden Datenpartner ausgehen, ein Effekt, der sich bereits beim Aufbau der Deutschen Digitalen Bibliothek gezeigt hat. Darüber hinaus soll der Aufbau eines Zeitungsportals Kulturerbe-Einrichtungen dazu anregen, weitere Digitalisierungsvorhaben von historischen Zeitungen anzugehen. Um ein Portal mit für die Zukunft gerüsteten offenen Schnittstellen zu schaffen, ist die perspektivische Integration der IIIF-Technologie sowie eine Anbindung an die Europeana Newspapers Collection geplant.

Ein Zeitungsportal für die Wissenschaft und die interessierte Öffentlichkeit

Das Medium Zeitung zeichnet sich dadurch aus, dass es alle Bereiche des Lebens abdeckt. Digitalisierte historische Zeitungen bieten den (Geistes-)Wissenschaften die Möglichkeit, eine Vielzahl von Forschungsfragen zu adressieren (Blome 2018). Das Interesse aus der Wissenschaft fließt beim Aufbau des Zeitungsportals auf mehreren Wegen in die Konzeption ein: So waren die oben genannten vier Kernfunktionalitäten Ergebnis eines Workshops mit WissenschaftlerInnen, der im Herbst 2014 im Rahmen des DFG-Pilotprojektes „Digitalisierung historischer Zeitungen“ in Bremen stattfand. Auch der laufende Entwicklungsprozess des Zeitungsportals wird von einem internationalen wissenschaftlichen Beirat begleitet.⁴ Im Gegensatz zu führenden Zeitungsprojekten der Digital Humanities wie *Impresso*, *NewsEyes* oder *Oceanic Exchanges*, welche ausschließlich ein wissenschaftliches Publikum im Blick haben, richtet sich das Angebot des Zeitungsportals auch an allgemeininteressierte NutzerInnen. Wie bereits bestehende, internationale Zeitungsportale zeigen, werden ihre Angebote sehr gut angenommen und von unterschiedlichsten Nutzergruppen besucht (für Nutzergruppen des österreichischen Zeitungsportals ANNO vgl. z.B. Müller 2016: 86). Vor allem die Volltextsuche, also die Möglichkeit, nicht nur in

den Zeitungstiteln und anderen Metadaten, sondern in den eigentlichen Zeitungstexten zu suchen, macht das Zeitungsportal zu einem sehr niedrigschwelligen Angebot: Auch ohne große Recherchekenntnisse können NutzerInnen Artikel oder Informationen zu allen vorstellbaren Themen finden – sei es zum eigenen Sportverein, zu einer berühmten Persönlichkeit oder zur Geschichte der eigenen Familie.

Anforderungen der unterschiedlichen Nutzergruppen

Das Projekt „Deutsches Zeitungsportal“ steht somit vor der Herausforderung, Anforderungen aus der Wissenschaft und aus der allgemeinen Öffentlichkeit zu analysieren, sie soweit wie möglich zu vereinen und ihnen allen bei der Umsetzung des Zeitungsportals möglichst gerecht zu werden.

Einer der ersten Schritte bei der Entwicklung des Zeitungsportals war es darum, die Anforderungen von Seiten der Wissenschaft sowie von allgemeininteressierten NutzerInnen zu erheben. Dabei kamen zu Beginn des Projekts drei unterschiedliche Methoden zum Einsatz: eine webbasierte Nutzerumfrage, fünfzehn, jeweils einstündige Interviews mit ausgewählten NutzerInnen und ein zweitägiger Workshop mit dem wissenschaftlichen Beirat.

Die Umfrage hatte als Ziel, potenzielle Zielgruppen des Zeitungsportals und ihre Anforderungen und Erwartungen kennenzulernen. Sie umfasste 23 Fragen zu Themengebieten wie Rechercheanlässe, Funktionalitäten und thematische Interessen und wurde für fünfeinhalb Wochen über die Homepage der DDB und über 27 weitere digitale Kanäle verbreitet. Das große Interesse, das dem Zeitungsportal entgegengebracht wird, zeigte sich hierbei schon rein zahlenmäßig – mit 2.422 ausgefüllten Fragebögen.

Die Usability-Tests mit fünfzehn ausgesuchten potenziellen NutzerInnen wurden als halbstrukturierte Interviews durchgeführt, bei denen Methoden wie Protokolle lauten Denkens, Beobachtung und Aufzeichnung des Klickverhaltens, Aufzeichnung der Gestik und Mimik zum Einsatz kamen. Der Schwerpunkt der Interviews lag auf der Interaktion mit einem Klick-Dummy des Zeitungsportals, der aufgrund gezielter Fragen und kleiner Aufgaben auf seine Verständlichkeit und Usability einerseits und die Erwartungen der Interviewten andererseits überprüft wurde.

Die Nutzerumfrage und die Nutzerinterviews richteten sich an die allgemeine Öffentlichkeit und wurden in Zusammenarbeit mit einer Marktforschungsagentur durchgeführt, die die Ergebnisse analysiert und aufbereitet hat.

Bei dem Workshop mit dem wissenschaftlichen Beirat handelte es sich um ein zweitägiges Treffen, bei dem im Juni 2019 die vorliegenden Konzeptpapiere und der Klick-Dummy vorgestellt und mit den WissenschaftlerInnen

unter Einbeziehung anderer Zeitungsportale diskutiert wurden. Die Empfehlungen, die im Lauf des Workshops entwickelt wurden, sind ebenfalls in die Konzeption des Portals eingeflossen.

Der Vortrag widmet sich der Auswertung dieser Erhebungen: Wo finden sich Gemeinsamkeiten? Wo gibt es Unterschiede? Was ist zu tun, wenn sich die Anforderungen aus Wissenschaft und allgemeiner Öffentlichkeit widersprechen? Welche Wünsche lassen sich überhaupt realisieren und wo liegen Grenzen, seien diese technisch oder urheberrechtlich bedingt?

Ein Beispiel für eine übereinstimmende Anforderung ist der Wunsch nach möglichst umfassenden Inhalten. Sowohl die allgemeine Öffentlichkeit als auch die Wissenschaftscommunity wünscht sich ein Zeitungsportal, das weitreichende Bestände anbietet, sodass sich die Recherche im besten Fall über ein einziges Portal erledigen lässt. Zwar ist genau dies der Anspruch und das angestrebte Alleinstellungsmerkmal des Deutschen Zeitungsportals, aber die Umsetzung dieses Zieles kann nicht allein vom Zeitungsportal erreicht werden. Viele Faktoren und Stakeholder – wie die schiere Menge des Materials, die Zerstreuung der Bestände in viele unterschiedliche Einrichtungstypen, die von der DFG überhaupt nicht erreicht werden, und nicht zuletzt das Urheberrecht, das die Digitalisierung und Verbreitung der besonders interessanten Bestände aus dem 20. Jahrhundert einschränkt – spielen hier eine Rolle. Ein kompletter Nachweis aller deutschen digitalisierten Zeitungsbestände kann darum eher als Vision der Community der Kulturerbe-Einrichtungen und ihrer Träger beschrieben werden, denn als ein kurzfristig erreichbares Ziel (Bürger 2018: 131f.).

Eine wichtige Erkenntnis der Diskussion war es, dass das Zeitungsportal möglichst transparent mit seinen Inhalten umgehen muss: Wenn nicht alle historischen Zeitungen zu finden sind, muss für die NutzerInnen deutlich werden, welche Inhalte verfügbar sind und nach welchen Kriterien das vorhandene Zeitungskorpus zusammengestellt wurde.

Unterschiedliche Erwartungen der Nutzergruppen wurden besonders im Bereich Nutzerinterface formuliert. Während die allgemeinen NutzerInnen sich hier eher eine einfach gestaltete Oberfläche wünschen und erwarten, alle Suchanfragen über einen Suchschlitz eingeben zu können, wurde in der Diskussion mit der wissenschaftlichen Begleitgruppe der Wunsch nach anspruchsvolleren Funktionen laut, z. B. nach der Möglichkeit, sich ein individuelles Zeitungskorpus zusammenzustellen und dieses gezielt durchsuchen zu können. Diese Anforderungen werden teilweise von der aus dem DDB-Hauptportal übernommenen Funktion „Meine DDB“ erfüllt, mit der sich Favoriten und Suchanfragen speichern lassen. Zudem könnten zukünftig erweiterte Funktionen implementiert werden, die, ohne die Startseite zu überladen, für WissenschaftlerInnen und erfahrene NutzerInnen einen komfortablen Zugang bieten, der komplexere Suchanfragen erlaubt. Der Suchschlitz für die Volltextsuche soll jedoch über alle Iterationen das bestimmende Element bleiben. Im Suchschlitz selber

können auch komplexe Abfragen gemäß der von der Suchmaschine vorgegebenen Syntax eingegeben werden.

Ausblick

Zum Abschluss des Vortrags wird der aktuelle Projektstand vorgestellt, inklusive eines ersten Blicks auf den Prototypen des Zeitungsportals, das Ende 2020 für die Öffentlichkeit freigeschaltet werden soll. Der Prototyp ist nicht nur aus technischen Gesichtspunkten (Tests, Qualitätssicherung) wichtig, sondern soll als Grundlage für eine zweite Runde der Nutzerforschung dienen: Die Erkenntnisse, die zu Beginn des Projektes gewonnen wurden, sollen bis Ende 2020 anhand einer weiteren Nutzerbefragung überprüft werden. So ist geplant, den Prototyp sowohl von den allgemeininteressierten NutzerInnen als auch vom wissenschaftlichen Beirat evaluieren zu lassen und daraus Erkenntnisse für die Schwerpunkte der nächsten, für 2021–2022 geplanten Projektphase zu gewinnen.

Fußnoten

1. Hervorzuheben sind insbesondere das Portal der Staatsbibliothek zu Berlin ZEFYS <http://zefys.staatsbibliothek-berlin.de> [letzter Zugriff 05.09.19], digiPress der Bayerischen Staatsbibliothek <https://digiPress.digital-sammlungen.de/> [letzter Zugriff 05.09.19] und das Zeitungsportal NRW <https://zeitpunkt.nrw/> [letzter Zugriff 05.09.19].
2. Deutsche Nationalbibliothek (Projektleitung), Sächsische Landesbibliothek – Staats- und Universitätsbibliothek Dresden (SLUB), Staatsbibliothek zu Berlin – Preußischer Kulturbesitz, FIZ Karlsruhe – Leibniz-Institut für Informationsinfrastruktur.
3. Dabei handelt es sich um die folgenden sechs Bibliotheken, von denen die ersten fünf auch am vorgeschalteten DFG-Pilotprojekt beteiligt waren: Sächsische Landesbibliothek – Staats- und Universitätsbibliothek Dresden (SLUB), Staatsbibliothek zu Berlin – Preußischer Kulturbesitz (SBB), Bayerische Staatsbibliothek (BSB), Universitäts- und Landesbibliothek Sachsen-Anhalt (ULB), Staats- und Universitätsbibliothek Bremen (SuUB), Staats- und Universitätsbibliothek Hamburg Carl von Ossietzky (SUB).
4. Mitglieder des wissenschaftlichen Beirats sind: Astrid Blome (Institut für Zeitungsforschung Dortmund), Christa Müller (Österreichische Nationalbibliothek), Claudia Resch (Österreichische Akademie der Wissenschaften), Estelle Bunout (Universität Luxemburg), Fotis Jannidis (Universität Würzburg), Günter Mühlberger (Universität Innsbruck), Jana Keck (Universität Stuttgart), Jörg Lehmann (Universität Tübingen), Marc Priewe (Universität Stuttgart), Maria Elisabeth Müller (Staats- und Universitätsbibliothek Bremen), Marian Dörk

(Fachhochschule Potsdam/Urban Complexity Lab), Marten Düring (Universität Luxemburg), Pim Huijnen (Universität Utrecht), Thomas Werneke (Zentrum für Zeithistorische Forschung Potsdam).

Bibliographie

ANNO (AustriaN Newspapers Online): <http://anno.onb.ac.at> [letzter Zugriff 29. August 2019].

Altenhöner, Reinhard (2018): „Auf dem Weg zu einem nationalen Zeitungsportal. Eine materialspezifische Kooperation als Treiber eines neuen Dienstes für Wissenschaft und Forschung“ in: Bonte, Achim / Rehnolt, Juliane (eds.): *Kooperative Informationsinfrastrukturen als Chance und Herausforderung: Festschrift für Thomas Bürger zum 65. Geburtstag*. Berlin, Boston: De Gruyter 144-160, DOI: 10.1515/9783110587524-019.

Bürger, Thomas (2016): „Zeitungsdigitalisierung als Herausforderung und Chance für Wissenschaft und Kultur“ in: *Zeitschrift für Bibliothekswesen und Bibliographie* 63, H 3, 123-132, DOI: 10.3196/186429501663332.

Blome, Astrid (2018): „Zeitungen“ in: Busse, Laura / Enderle, Wilfried / Hohls Rüdiger u.A. (eds.): *Clio-Guide. Ein Handbuch zu digitalen Ressourcen für die Geschichtswissenschaften*. Berlin: Clio-online und Humboldt-Universität zu Berlin (Veröffentlichungen von Clio-online, Nr. 2), B.6-1-D.6-36, DOI: 10.18452/19244.

The British Newspaper Archive: <https://www.britishnewspaperarchive.co.uk> [letzter Zugriff 29. August 2019].

Burckhardt, Daniel / Geyken, Alexander / Saupe, Achim / Werneke, Thomas (2019): „Distant Reading in der Zeitgeschichte. Möglichkeiten und Grenzen einer computergestützten Historischen Semantik am Beispiel der DDR-Presse“ in: *Zeithistorische Forschungen/Studies in Contemporary History* 16: 177-196, DOI: 10.14765/zzf.dok-1345.

Delpher: <https://www.delpher.nl/> [letzter Zugriff 29. August 2019].

DFG-Antrag „Errichtung eines nationalen Zeitungsportals auf der Basis der organisatorischen und technischen Infrastruktur der Deutschen Digitalen Bibliothek (DDB) – DDB-Zeitungsportal. Online verfügbar unter: <https://pro.deutsche-digitale-bibliothek.de/node/985> [letzter Zugriff 29. August 2019].

DFG-Ausschreibung „Digitalisierung historischer Zeitungen des deutschen Sprachgebiets“ im Rahmen des LIS-Förderprogramms von 2018 (https://www.dfg.de/download/pdf/foerderung/programme/lis/ausschreibung_zeitungsdigitalisierung.pdf) [letzter Zugriff 10. September 2019].

DFG-Ausschreibung „Digitalisierung historischer Zeitungen des deutschen Sprachgebiets“ im Rahmen des LIS-Förderprogramms von 2019(<https://www.dfg.de/download/pdf/foerderung/programme/lis/>

[ausschreibung_zeitungsdigitalisierung_2019.pdf](#)) [letzter Zugriff 10. September 2019].

digipress: <https://digipress.digitale-sammlungen.de/> [letzter Zugriff 05. September 2019].

Empfehlungen zur Digitalisierung historischer Zeitungen in Deutschland (Masterplan Zeitungsdigitalisierung): http://www.zeitschriftendatenbank.de/fileadmin/user_upload/ZDB/z/Masterplan.pdf [letzter Zugriff 05. September 2019].

Impresso: <https://impresso-project.ch/> [letzter Zugriff 29. August 2019].

Müller, Christa (2016): „ANNO – Der Digitale Zeitungslesesaal der Österreichischen Nationalbibliothek: Aktuelle und zukünftige Entwicklungen im Überblick“ in: *BIBLIOTHEK – Forschung und Praxis* 40(1). Berlin: de Gruyter 83-89, DOI: 10.1515/bfp-2016-0012.

NewsEye: <https://www.newseye.eu/> [letzter Zugriff 29. August 2019].

Oceanic Exchanges: <http://oceanicexchanges.org/> [letzter Zugriff 29. August 2019].

ZDB (Zeitschriftendatenbank): <https://zdb-katalog.de/> [letzter Zugriff 29. August 2019].

ZEfYS: <http://zefys.staatsbibliothek-berlin.de> [letzter Zugriff 05. September 2019].

Zeitpunkt NRW: <https://zeitpunkt.nrw/> [letzter Zugriff 05. September 2019].