

Martin Faßnacht

Liberté, Egalité, Solidarité

Die retrospektive und prospektive Digitalisierung der ThPQ

◆ War es bis vor nicht allzu langer Zeit unumgänglich, besonders Fachliteratur in Bibliotheken zu entleihen, so bieten heute via Internet abrufbare Digitalisate öffentlichen Zugang zu schier unerschöpflichen Quellen schriftlichen Schaffens – dies unabhängig von Ort, Zeit und Stand. Welchen konzeptionellen, logistischen und technischen Aufwand sowie Mehrwert ein solches Projekt bedeutet, zeigt unser Autor, Leiter der Abteilung Fachinformationsdienste und Koordinator des Index Theologicus an der Universitätsbibliothek Tübingen, exemplarisch anhand der aktuell in Umsetzung begriffenen Digitalisierung der Theologisch-praktischen Quartalschrift. (Redaktion)

Vor nicht allzu langer Zeit war der Besuch einer Bibliothek der einschlägige Weg, Zugang zu wissenschaftlicher Literatur zu erhalten. Hatte die Leserin / der Leser die Bibliothekspforten durchschritten, stand ihr / ihm der Weg zum Buch frei. Oder doch nicht ganz? Während meiner Mitarbeit am Institut für neutestamentliche Textforschung in Münster bekannte mir eine italienische Kollegin, eine ihrer befreienden Erfahrungen im wissenschaftlichen Informationssystem sei der freie Zugang zum Buch in einer deutschen Bibliothek gewesen. Keine Bestellscheine ausfüllen, kein Schlange stehen vor Bibliotheksschaltern, keine Wartezeit bei der Bereitstellung der Bücher, kein Vertrösten auf den nächsten Tag! Nun ist es auch in einer deutschen Bibliothek noch keine Ewigkeit her, seit die Nutzer mit Mantel und Tasche in die Bibliothek marschieren, am Regal das Buch entnehmen und sogleich entleihen können – und in vielen italienischen Bibliotheken dürfte sich diese Praxis mittlerweile ebenso etabliert haben. Auch wenn es nach wie vor Bücher gibt, die bestellt wer-

den müssen – vor allem die Schützenswerten –, steht fest: Der freie Zugang über ein Freihandmagazin hat viele Erwartungen der Nutzer an die Zugänglichkeit benötigter Literatur mit einem Schlag erfüllt.

Erwartungen und Bedürfnisse der Nutzer in Bezug auf den Zugang haben sich in den letzten Jahren rasant verändert. Im Zeitalter der Digitalität gehen sie weit über die analogen Bedürfnisse hinaus und sind nur noch durch Breitbandgeschwindigkeit und Hardware limitiert. Heutzutage muss man immer seltener Bibliotheksportale durchschreiten, vielmehr genügt es, Onlineportale aufzurufen, um die benötigte Literatur direkt auf seinem mobile Device lesen und nutzen zu können, weil neue Inhalte durchgängig elektronisch produziert und publiziert werden. Das gilt für die alten Druckausgaben naturgegeben nicht. Da sich aber auch die vorgängig analog publizierten Inhalte dem digitalen Paradigma nicht verweigern können, ohne immer seltener wahrgenommen zu werden, haben sich viele Kulturinstanzen der Aufgabe angenommen, diese Inhalte in das

digitale Format zu transformieren. Auch die Universitätsbibliothek Tübingen – u. a. mit ihren Fachinformationsdiensten – ist daran aktiv beteiligt.

1 Liberté, Egalité, Solidarité

Elektronisch verfügbare Texte rufen nach wie vor Kritik auf den Plan. Und es ist angesichts der digitalen Möglichkeiten sicherlich klug, nicht über-optimistisch zu werden. Lesegeohnheiten, Vorteile des haptischen Mediums, Energieverbrauch und Langzeitspeicherung elektronischer Information sind dabei einige Stichwörter. Das digitale Paradigma ist aber dem unbedarften Leser nicht von außen aufgezungen, sondern bietet sich dem Leser selbstverständlich an. Es ist vergleichbar dem Wikipedia-Kritiker, der seine Studierenden übertrieben vor dem Gebrauch der Online-Enzyklopädie auf seinem Fachgebiet warnt und in allen anderen Fällen selbst darin nachschlägt.

Was ist also der Vorteil elektronisch verfügbarer Texte? Oder präziser formuliert: elektronischer Texte, die im Open Access frei zugänglich sind?

Die drei gering modifizierten Schlagwörter der französischen Revolution fassen prägnant die enormen Vorteile zusammen: *Freiheit*, unabhängig örtlicher Gegebenheiten zu jeder Tages- und Nachtzeit schnell auf die Texte zugreifen zu können; *Chancengleichheit*, ohne Zugehörig-

keit zu einer bestimmten Statusgruppe, für die heutzutage Lizenzen erworben werden (z. B. der Gruppe von Lehrenden und Studierenden einer Universität), oder ohne die Privilegien wohnortnaher Informationsstrukturen lesen, sich bilden und forschen zu können; *Solidarität* und wirtschaftliche Sparsamkeit, den elektronischen Zugang nur einmal bezahlen zu müssen und nicht von jedem Nutzer (oder seiner Bibliothek) einzeln und durch jährliche Lizenzierung¹ auf alle Ewigkeit immer wieder.

Es kommt also weder auf Wohnort noch Status und auch nicht auf das Budget des Einzelnen oder der einzelnen Bibliothek an, ob ein Buch oder eine Zeitschrift vor Ort zugänglich ist oder nicht. Und das ist nicht nur im Hinblick auf die Solidarität mit Nutzern aus Entwicklungsländern, sondern auch schon im Hinblick auf die sehr unterschiedliche Ausstattung der Bibliotheken in Deutschland ein wichtiger Gesichtspunkt.²

2 Digitalisierungsprojekte an der Universitätsbibliothek Tübingen im Bereich der Theologie

Die Universitätsbibliothek Tübingen hat im Bereich der Theologie ein erstes Digitalisierungsprojekt 2010 gestartet. Zentrale Quellen zur „Deutschen Theologie des 19. Jahrhunderts“ wurden online gestellt.³ Die Vorgabe lautete, dass nur Quellen in das Projekt aufgenommen werden, die nicht schon anderweitig digital zur Verfügung

¹ Solche Lizenzierungsmodelle bieten JSTOR (Journal Storage, eine in New York ansässige Non-Profit-Organisation), DigiZeitschriften (ein von deutschen Kultureinrichtungen getragener Verein) und ATLA (American Theological Library Association) an.

² Die Chancen elektronischer Open Access Publikationen im Hinblick auf neue Publikationsformate, Forschungsmethoden, interaktive Inhalte, neue Fragestellungen sowie den Möglichkeiten linguistischer Verfahren im Bereich der Digital Humanities werden hier nicht dargestellt. Siehe dazu den Beitrag von Franz Böhmisch in dieser Ausgabe.

³ <http://idb.ub.uni-tuebingen.de/digitue/theo/> [Abruf: 10.01.2020].

standen. Zwanzig Zeitschriften und 58 Bücher wurden ausgewählt, insgesamt wurden mehr als 500.000 Einzelseiten und ca. 46.000 Metadaten erzeugt.

Ein zweites Projekt schloss sich mit der Digitalisierung von zwei renommierten Rezensenzeitschriften an. 66.000 Rezensionsartikel und 109.000 Metadaten aus dem Theologischem Literaturblatt (1880–1943) und der Theologischen Literaturzeitung (1876–1995) wurden im Open Access publiziert. Wurden die beiden ersten Projekte anteilig von der Deutschen Forschungsgemeinschaft (DFG) und der Universitätsbibliothek Tübingen finanziert, so konnte für das dritte Projekt die Diözese Rottenburg-Stuttgart als Ko-Finanzier gewonnen werden: während die Jahrgänge 1819–1900 der Theologischen Quartalschrift (ThQ) schon im ersten Digitalisierungsprojekt digitalisiert wurden, stehen nun auch alle Artikel von 1901–2005 digital zur Verfügung.

Die Auswahl der Quellen in den bisherigen Projekten basiert auf verschiedenen Kriterien: Im ersten Projekt wurden die Quellen nach inhaltlichen Kriterien ausgewählt, im zweiten Projekt lagen formale Kriterien zugrunde (Gattung Rezension).

Das jetzt anstehende vierte und bei der DFG zur Förderung eingereichte Digitalisierungsprojekt, in dessen Rahmen auch die Jahrgänge 1848–2015 der Theologisch-Praktischen Quartalschrift (ThPQ) digitalisiert werden sollen, hat ein völlig anderes Auswahlkriterium. Da die theologische Fachcommunity in Deutschland eine umfangreiche Retrodigitalisierung theologischer Fachzeitschriften als ein wichtiges Desiderat ansieht, hat die UB Tübingen ein Massendigitalisierungsprojekt initiiert,

an dem außer der ThPQ noch weitere 66 Zeitschriften teilnehmen. Die Auswahlkriterien lauten: Die Zeitschriften müssen im Index Theologicus (IxTheo) ausgewertet werden (Qualitätskriterium) und sind noch nicht oder nur teilweise schon digital verfügbar. Um das Projekt logistisch handhaben zu können, wurden nur deutschsprachige Zeitschriften in das Projekt aufgenommen. Nach Beendigung des Projekts stehen mehr als 150.000 Aufsätze und Rezensionen neu im Open Access weltweit zur Verfügung. Damit wird systematisch die Lücke fehlender retrospektiver Open Access Stellung von theologisch wichtigen Zeitschriften verkleinert.

Ein wichtiger Unterschied zu den früheren Digitalisierungsprojekten besteht darin, dass sich die Digitalisierung nicht nur retrospektiv auf das Material bis zu einem bestimmten Zeitpunkt bezieht, sondern nach dem Prinzip der Moving Wall jährlich ein weiterer Jahrgang, der aus der Embargofrist fällt, online gestellt werden kann.⁴ Dadurch wird das Geschäftsmodell der Printzeitschriften nicht gefährdet, gleichzeitig werden die Inhalte aber zeitverzögert elektronisch verfügbar. Für ein solches „mitlaufendes“ Verfahren braucht es andere Tools und Workflows als bei abgeschlossenen Digitalisierungsprojekten.

3 Arbeitsschritte in einem Digitalisierungsprojekt

Im Rahmen eines Zeitschriften-Digitalisierungsprojektes fallen prinzipielle Arbeitsschritte an:

1. Jede einzelne Seite einer Printzeitschrift wird gescannt und im TIFF-Format

⁴ Beispiel: Eine Zeitschrift mit einem Embargo von drei Jahren überführt im Jahre 2020 den Jahrgang 2016 in den Open Access, 2021 den Jahrgang 2017 etc.

unter einem eindeutigen, sprechenden Dateneinamen abgespeichert.

2. Jeder einzelne Aufsatz und jede einzelne Rezension wird formal katalogisiert: Erfassung der Metadaten wie Autor, Autoren-Identifikationsnummer (GND)⁵, Titel, Untertitel, Seitenzahl sowie die Zuordnung zu Heft, Band und Jahrgang.

Weiterführende Literatur:

Zu Open Access:

Bundesministerium für Bildung und Forschung, Open Access, <https://www.bildung-forschung.digital/de/open-access-2471.html> [Abruf: 10.01.2020]. Hier wird das Thema aus verschiedenen Blickwinkeln und mit verschiedenen Medien beleuchtet: Forschungspolitische Sicht, rechtliche Fragen, Qualität, Open Access Initiativen, innovative Projekte etc. Spannend darin der Podcast von *Holger Klein* zum Thema Open Access, <https://www.bildung-forschung.digital/de/open-access---der-podcast-zum-thema-2761.html> [Abruf: 10.01.2020].

Zur Digitalisierung:

Fridtjof Küchemann, Leseforscher zur Digitalisierung. Der Kontakt zu unserer Kultur steht auf dem Spiel, Artikel vom 15.10.2018, in: Frankfurter Allgemeine Zeitung, Rubrik Debatten, <https://www.faz.net/aktuell/feuilleton/debatten/interview-acht-leseforscher-zur-digitalisierung-15833105.html> [Abruf: 10.01.2020]. Kritischer Blick auf mögliche Auswirkungen digitalen Lesens, der eine heftige und kontroverse Debatte ausgelöst hat.

Wenn die Ressourcen dazu vorhanden sind, außerdem:

3. Inhaltliche Erschließung: Jeder Aufsatz wird verbal mit Schlagwörtern und/oder klassifikatorisch mit Notationen erschlossen.

Die beiden ersten Arbeitsschritte sind obligatorisch, da sonst die Scans einem Aufsatz nicht zugeordnet werden können. Es muss ja bestimmt werden können, dass z. B. die ersten zehn Scans zu Aufsatz A und die nächsten fünf Scans zu Aufsatz B gehören. Die Katalogisierung der Aufsätze wie die Zuordnung der einzelnen Scans zu den Aufsätzen kann in einem großen Digitalisierungsprojekt nicht händisch erfolgen. Hier werden halb-automatische Verfahren eingesetzt, um große Stückraten pro Stunde zu erreichen.

4 Art des Digitalisierungsprojektes: retrospektiv und/oder prospektiv?

Die Wahl der Tools hängt von der Art des Digitalisierungsprojektes ab. Ist es abgeschlossen oder sollen nach dem Moving Wall Prinzip auch zukünftig regelmäßig ganze Jahrgänge ins Open Access überführt werden?

In retrospektiven, abgeschlossenen Projekten genügt es, die Präsentation der Digitalisate verschiedener Zeitschriften in einer großen Digitalisierungsumgebung unter einheitlicher URL zu organisieren. Die UB Tübingen verwendet dazu die an der Universitätsbibliothek Heidelberg entwickelte Software DWork.⁶ DWork ist darauf ausgelegt, als Frontend für die Nutzer eine gut funktionierende Präsentationsschicht bereit zu stellen und zugleich

⁵ Im deutschsprachigen Bereich (D-A-CH) wird zur Autorenidentifikation die Gemeinsame Normdatei (GND) verwendet.

⁶ <http://dwork.uni-hd.de> [Abruf: 10.01.2020].

als Backend für die Organisation der Metadaten und Scans zu dienen. Für prospektive Projekte ist DWork jedoch nicht geeignet. Hier brauchen die Herausgeber und Schriftleiter selbst die Möglichkeit, Metadaten und PDFs der einzelnen Aufsätze eingeben, hochladen und gemäß den Absprachen mit dem Verlag mit einem Zeitstempel versehen zu können, so dass nach Ablauf der Embargofrist der Aufsatz automatisch ins Open Access überführt wird, ohne dass nach Jahren noch etwas an den Datensätzen verändert werden muss. Für eine solche Anforderung ist die Software Open Journal Systems (OJS) geeignet.⁷

OJS bietet die Möglichkeit, jeder am Digitalisierungsprojekt beteiligten Zeitschrift einen Internetauftritt mit eigenständiger URL und hoher Sichtbarkeit sowie der Indizierung durch die wichtigsten Suchmaschinen wie Google zu garantieren. Als Backend für die prospektive Erfassung von Aufsatzmetadaten und dem Hochladen von PDFs neuer Hefte ist es eine weltweit tausendfach genutzte und bewährte Software. Als Backend für retrospektive Digitalisierungsworkflows wie z. B. der automatischen Zuordnung von Scans und Katalogisat ist OJS allerdings ungeeignet. Die UB Tübingen wird deswegen bei dem kommenden Digitalisierungsprojekt sowohl DWork als auch OJS einsetzen und somit das Beste beider Tools nutzen. In DWork als Backend werden die Workflows abgearbeitet, danach werden die Me-

tadaten samt PDFs nach OJS als Frontend migriert.

5 Workflow Digitalisierung ThPQ

Die folgend dargestellten Arbeitsschritte sind für jede am Projekt beteiligte Zeitschrift durchzuführen und werden modellhaft an der ThPQ skizziert:

1. Scannen der ThPQ Jahressbände von 1848 bis einschließlich 2015 durch einen externen Dienstleister: hochgerechnet ca. 42.500 Seiten.

2. Katalogisierung von hochgerechnet ca. 4.400 Aufsätzen und Rezensionen der ThPQ nach bibliothekarischen Qualitätsstandards im gemeinsamen Katalog K10plus des Südwestdeutschen Bibliotheksverbundes (SWB) und des Gemeinsamen Bibliotheksverbundes (GBV). 63 % können dabei mit einem halb-automatischen Verfahren bearbeitet werden. Das Software Programm der Firma ImageWare (Bonn) übernimmt die Metadaten strukturiert aus dem Inhaltsverzeichnis.⁸ In 37 % der Fälle sind die Inhaltsverzeichnisse der ThPQ in Fraktur gedruckt oder anderweitig nicht für das bildgestützte ImageWare Verfahren geeignet und müssen händisch katalogisiert werden.

3. Anlegen eines Projektes für jeden Jahrgang der ThPQ in DWork.

4. Jahrgangsweises Einlesen der 42.500 ThPQ-Scans nach DWork.

⁷ OJS ist ein Redaktions- und Präsentationssystem für digitale Zeitschriften, mit dem der gesamte Redaktionsprozess von der Einreichung eines Artikels über das Review-Verfahren bis zur digitalen Publikation abgebildet werden kann. Die Vergabe von Zeitstempeln steuert auf Artikel-ebene die Open Access Stellung.

Ein Beispiel für eine Zeitschrift mit einjährigem Embargo ist das von Marianne Heimbach-Steins im Ashendorf-Verlag herausgegebene Jahrbuch für Christliche Sozialwissenschaften: <https://www.uni-muenster.de/Ejournals/index.php/jcsw/index> [Abruf: 10.01.2020].

⁸ Siehe hierzu *Martin Faßnacht / Winfried Gebhard*, Index Theologicus – neue Produktionsverfahren bei der Bibliographieerstellung, in: b. i. t. online 19 (2016), 511–514. <https://www.b-i-t-online.de/heft/2016-06-nachrichtenbeitrag-fassnacht.pdf> [Abruf: 10.01.2020].

5. Semi-automatisches Editieren der Seitennummern in DWork.

6. Export der Metadaten der einzelnen ThPQ Aufsätze und Rezensionen aus dem SWB und Import nach DWork.

7. Strukturdatenerfassung: Zusammenfassung von Titelseite, Vorwort etc. als Strukturdatum Frontmatter; Strukturdatum Inhaltsverzeichnis; Zusammenfassung von nicht wissenschaftlichen Textbeiträgen, Verzeichnissen etc. als Strukturdatum Backmatter.

8. Zuordnung der Scans zu den Metadaten der Aufsätze/Rezensionen. Aufgrund der sowohl in DWork als auch in den SWB-Metadaten vorhandenen Seitennummern werden die Zuordnungen der Scans zu den Metadaten automatisch erstellt.

9. Qualitätskontrolle: Prüfung der Zuordnung der Metadaten und Digitalisate.

10. Konvertierung der Scans aus dem TIFF-Format ins JPG-Format; Erkennung und Überführung der Bilddateien in maschinenlesbaren Text (OCR); Generierung eines PDFs für jeden Aufsatz/Rezension; Automatische Erzeugung eines persistenten Identifiers (DOI)⁹ pro Aufsatz/Rezension, der unabhängig von einer bestimmten URL zum Volltext führt.

11. Anlegen eines OJS Mandanten für die ThPQ.

12. Export der Metadaten und der zugehörigen PDFs aus DWork und Import nach OJS.

13. Einspielung in den K10plus als Online-Katalogisat. Dabei wird aus dem Metadatensatz des Print-Katalogisates ein Online-Katalogisat mit URL und DOI als Link auf das PDF erzeugt. Mittels eines Abrufzeichens werden die Aufsätze und Rezensionen automatisch in den IxTheo übernommen.

6 Präsentation und Distribution

Die retrospektiv erzeugten Digitalisate und Metadaten aller Jahrgänge der ThPQ von 1848 bis 2015 werden in einem eigenständigen OJS Mandanten präsentiert. Dieser Mandant kann zukünftig auch für die aktuellen Hefte, die noch einem Embargo unterliegen, genutzt werden. Einzelne Aufsätze können trotz Embargo als Volltext freigegeben werden.

Für die Auffindbarkeit ist neben einem sichtbaren, persistenten und durch Suchmaschinen ausgewerteten Online-Auftritt auch die Indizierung in international einschlägigen Datenbanken unverzichtbar. Diese sind als weltweiter Hub erster Anlaufpunkt für die wissenschaftliche Literaturrecherche. Alle Aufsätze und Rezensionen werden deswegen in der internationalen Open Access Bibliografie Index Theologicus nachgewiesen. Das ist nur durch die oben dargestellten Arbeitsschritte 2 und 13 möglich. Durch die Katalogisierung im gemeinsamen Verbundkatalog K10plus können die Aufsätze und Rezensionen sowohl im IxTheo als auch in den Bibliothekskatalogen der einzelnen Verbundbibliotheken recherchiert und gefunden werden.

Die Projektbeteiligung nicht nur der ThPQ, sondern auch der anderen ausgewählten Zeitschriften ist ein großer Dienst an der weltweiten Fachcommunity und trägt dazu bei, dass die deutschsprachige, theologische Forschung der vergangenen Jahrzehnte und Jahrhunderte auch zukünftig einfach, schnell, kostenfrei, umfassend und auf moderne Weise neu genutzt werden kann. In geisteswissenschaftlichen Fächern wie der Theologie ist das ein sehr großer Mehrwert und ein bleibendes Ver-

⁹ Ein Digital Object Identifier (DOI) ist ein eindeutiger, dauerhafter, digitaler Identifikator.

dienst aller am Projekt beteiligten Herausgeber und Verlage.

Der Autor: *Dr. Martin Faßnacht*, geb. 1964, Ausbildungen zum Tischler, Theologen und Bibliothekar; berufliche Tätigkeit: 2000–2011: Wissenschaftlicher Mitarbeiter am Seminar für Zeit- und Religionsge-

schichte des Neuen Testaments und am Institut für neutestamentliche Textforschung, Münster; seit 2011: Leiter des Index Theologicus/FID Theologie und Abteilungsleitung Fachinformationsdienste (FID) an der Universitätsbibliothek Tübingen; GND: 122054482; ORCID: <http://orcid.org/0000-0002-2672-4387>.