

# Fernsehen für alle! oder: Was kann ein Katalog?

Ulrich Johannes Schneider, Sebastian Stoppe

Vielen Dank, dass wir unser Projekt kurz vorstellen können.<sup>1</sup> Wir bereiten den Projektantrag gerade in der Universitätsbibliothek Leipzig vor, und zugleich ist er in Vorbereitung mit Ihnen. Bibliotheken machen Kataloge für alle möglichen Medien. Wir denken und werden das jetzt gleich vorstellen, dass wir das auch für Produktionsarchive machen könnten. Wir wissen sicher, dass mit unserem Projekt im Bewilligungsfall der Wissenschaft geholfen wird, für die wir eigentlich hauptsächlich arbeiten und der gegenüber sich Bibliotheken als Dienstleister verstehen.

So begreift sich jedenfalls die UB Leipzig, die eine der ältesten wissenschaftlich gegründeten Bibliotheken der Welt ist. Sie hat ein schönes altes Gebäude, die Bibliotheca Albertina, unter ihren sieben neu gebauten bzw. neu aufgebauten Gebäuden ist seit zehn Jahren auch eine 24-Stunden-Bibliothek direkt am Augustusplatz. Wir erschließen sehr viele, sehr verschiedene Medien, darunter zum Beispiel 86.000 Münzen, 15.000 Portraitstiche, 5.000 orientalische Handschriften, Ostraka und Papyri und vieles mehr. Und für all diese Medien haben wir Kataloge

entwickelt.<sup>2</sup> Wir entwickeln auch Kataloge für andere Bibliotheken<sup>3</sup>, zum Beispiel Musikbibliotheken und andere, derzeit insgesamt 20 zum Teil sehr spezialisierte Bibliotheken in ganz Deutschland. Wir machen deren Kataloge, weil Metadaten sozusagen unser tägliches Geschäft sind. Was Sie vielleicht interessieren mag: Wir haben auch eine Sammlung japanischer Videospiele, die bei uns in Zusammenarbeit mit dem Institut für Japanologie erschlossen wird.<sup>4</sup> Das ist eine heiße Sache und sehr kompliziert. Wo man die Metadaten zu einem japanischen Videospiele herkriegt, das möchten Sie gar nicht wissen. Da muss man Fansseiten katalogisieren und indexieren, aber auch dieser Herausforderung stellen wir uns.

Wir sind 2017 „Bibliothek des Jahres“ geworden. Das ist eine Auszeichnung, die wir vor allen Dingen für unsere Digital-Kompetenz bekommen haben. Vor 13 Jahren hatten wir 230 Mitarbeiter, darunter fünf in der EDV. Jetzt haben wir 210 Mitarbeiter. Wir mussten 20 Stellen abbauen und haben dennoch jetzt knapp 40 Beschäftigte im Bereich Digitale Dienste. Da sehen Sie die Schwerpunktverlagerung, ohne die



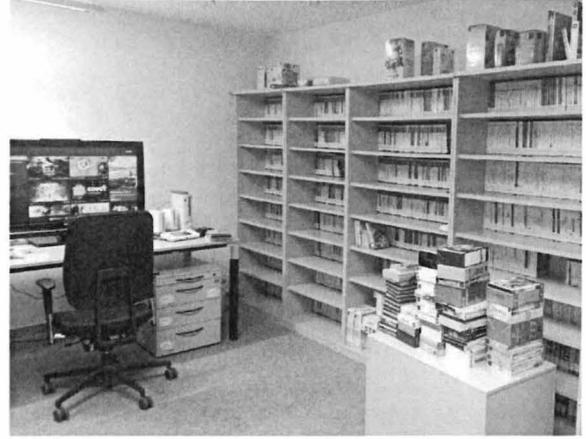
Prof. Dr. Ulrich Johannes Schneider und Sebastian Stoppe  
Universitätsbibliothek Leipzig  
Beethovenstraße 6  
04107 Leipzig  
+49 341-97 30577  
schneider@ub.uni-leipzig.de

<sup>1</sup> Der Text beruht auf dem gleichnamigen Vortrag bei der Frühjahrstagung des vfm in Leipzig am 10.4. 2019. Wir haben den Text im ursprüngliche Redestil belassen, ihn aber grundlegend sprachlich redigiert. Aus Gründen der besseren Lesbarkeit verwenden wir nur einen Genus, natürlich sind in diesen Fällen alle Geschlechter ausdrücklich mitgemeint.

Universitätsbibliothek Leipzig



Universitätsbibliothek Leipzig erschließt 5.000 japanische Videospiele



wir uns nicht trauen würden, hier vor Ihnen zu sprechen. Wir haben auch einen „Hackathon Coding Da Vinci“<sup>5</sup> durchgeführt, eben hier in Leipzig.

Wir möchten kurz resümieren, wie wir als Bibliothek sehen, was an Sie als Archive an Forderungen herangetragen wird. Also da gibt es einmal die Mediathekenfans, die nicht verstehen, dass da nicht alles drin ist. Das ist zum Beispiel eine Forderung, ein stark artikuliertes Konsumbedürfnis: Die Leute wollen bedient werden. Warum kann ich Heinz Erhardt jetzt nicht sehen? Gesellschaftliche Interessen gibt es auch. Das ist dann schon eine etwas gewichtigere Argumentation aus einer anderen Ecke, die sagt, wir wollen Transparenz, wir wollen Zugänglichkeit. Natürlich gibt es institutionelle Rahmenbedingungen und es steht fest, dass die allgemeine Freigabe aller öffentlich-rechtlichen Inhalte nicht ursprünglich zum Aufgabenbereich der Archive gehört. Es gibt natürlich auch finanzielle Hinderungsgründe. Wenn wir Ihnen gerade schildern, wie der Umbau der IT-Technik innerhalb einer Bibliothek stattfindet, können Sie sich ja vorstellen, dass man sich strategische und finanzielle Gedanken unterschiedlichster Art machen kann.

Dann kennen wir auch diese öG-öG-Leute, die rufen: Öffentliches Geld macht etwas zum öffentlichen Gut. Mit anderen Worten: Was öffentlich bezahlt worden ist, sollte auch öffentlich zur Verfügung stehen. Das halten wir für überzogen. Wir als Bibliotheken müssen ja auch Rechte von Verlegern schützen und damit auch die von Urhebern. Also wir stehen wir als Institution, die einerseits die Öffentlichkeit und insbesondere die wissenschaftliche allgemeine Öffentlichkeit informiert, auch immer in diesem Spagat, einerseits alles öffentlich zugänglich zu machen und andererseits die Schutzrechte nicht zu verletzen.

Aus unserer Sicht sind Wissenschaftler wirklich nicht berechenbar. Die wollen eigentlich immer alles haben. Sie wollen die Rundfunkkultur in Deutschland im Ganzen und im Detail erforschen. Sie sagen uns, wir behandeln das ganze Archiv als eine Forschungsdatenbank und das möchten wir eben durchsuchen.

Es sollen Inhalte auch innerhalb der Sendung gefunden werden. Also muss die Granularität sehr genau sein. Das ist ein berechtigtes Interesse der Wissenschaft, das aber schon auch im bibliothekarischen Bereich an allerlei Grenzen stößt, denn nicht alle Dinge sind so genau erschlossen, wie man das haben möchte. Und wenn jetzt die Wissenschaft sagt, alle Archivdaten sind potenzielle Forschungsdaten, würden wir als Bibliothek sagen, dass Kataloge machen nicht heißt, dass wir das in der Fläche freigeben. Die Recherche, die Einzelrecherche ist etwas anderes, als eine Recherche, die man über einen Katalog als komplette Datenbank vermittelt.

Die Tatsache bleibt und sollte uns auch immer zu denken geben, dass Forschungsinteressen sich sehr wandeln. So wie man früher Bücher ohne die Umschläge retrodigitalisiert hat, und Zeitschriften ohne die Werbeseiten. Das bedauern die Kulturwissenschaftler heute sehr, weil die genau darüber gerne arbeiten würden. Also muss man aufpassen. Dieser Wandel der wissenschaftlichen Interessen kann uns als Gedächtnisinstitution an Stellen erwischen, an die wir gar nicht gedacht haben. Im geänderten Urheberrecht gibt es jetzt mit dem Einzelplatz in Bibliotheken eine bessere Lösung. Wir digitalisieren zum Beispiel im Landesdigitalisierungsprogramm des Freistaats Sachsen auch die Filme des Tanzarchivs der DDR. Diese können wir aber zum großen Teil nicht öffentlich ins Netz stellen, aber wohl über einen Einzelplatz in der Bibliothek zugänglich machen. Das ist vielleicht nicht ideal für die Wissenschaft. Wir sehen darin als Bibliothek jedoch einen enormen Fortschritt. Immer wenn das Urheberrecht in Frage steht, kommen die Verwertungsrechteinhaber stark nach vorne und sagen: Wir vertreten hier die Individuen und die gesellschaftliche Kreativität. Dagegen wird es unseres Erachtens auch künftig immer Anwälte des wissenschaftlichen Interesses geben. Und als Bibliothek verstehen wir uns auch auf dieser Seite stehend.

Wenn wir hier von Wissenschaftlern gesprochen haben, dann meinen wir nicht nur Wissenschaftler

<sup>2</sup> <https://www.ke-nom.de>, <http://www.portraitindex.de>, <https://www.is-lamic-manuscripts.net>, <https://www.refaiya.uni-leipzig.de> sowie <https://papyri.uni-leipzig.de>. Siehe auch <https://www.ub.uni-leipzig.de/forschungsbibliothek/digitale-sammlungen>.

<sup>3</sup> <https://finc.info>

<sup>4</sup> <https://diggr.link>

<sup>5</sup> <https://codingdavinci.de/events/ost/>

<sup>6</sup> <https://adlr.link>



Universitätsbibliothek Leipzig organisiert Hackathon Coding da Vinci Ost

an der Universität Leipzig, sondern durchaus Wissenschaftler bundesweit. Wir betreiben seit drei Jahren den Fachinformationsdienst *adlr.link*.<sup>6</sup> Das ist ein Suchportal für die Medien-, Kommunikations- und Filmwissenschaften, eine Art Spezialkatalog. Wir haben da zurzeit ungefähr anderthalb Millionen Datensätze drin verzeichnet. Und es ist nicht nur ein Rechercheinstrument, sondern auch ein Lieferdienst für die Wissenschaftler. Diese können sich bei *adlr.link* anmelden und hinterlegen bei Ihrer Anmeldung die Information, welche Bibliothek in Deutschland sie normalerweise nutzen. Und wir zeigen ihnen dann bei der Recherche an, ob ein Buch bei ihrer Bibliothek schon vorhanden ist oder ob eine Zeitschrift lizenziert ist oder nicht. Wenn das nicht der Fall ist, dann ist es unser Ziel, unseren Nutzern die benötigte Literatur schnell und unkompliziert zur Verfügung zu stellen. Und dabei wollen wir neue, innovative Wege gehen, also nicht über die alten Fernleihsysteme, sondern wann immer es möglich ist, versuchen wir den Wissenschaftlern Inhalte gleich digital weiterzuleiten. Wenn es lizenzrechtlich erlaubt ist, dann schicken wir etwa einen Zeitschriftenaufsatz, den der Wissenschaftler an seiner Bibliothek nicht bekommt, ihm direkt per E-Mail. Und der freut sich, dass er in relativ kurzer Zeit seinen ganzen benötigten Literaturkorpus zusammen hat. Und natürlich ist *adlr.link* auch eine Suchmaschine, mit der man Suchergebnisse facettieren, also nach Erscheinungsjahr, Medientyp oder ähnlichem einschränken kann, mit der man Suchen abspeichern kann, mit der man auch relativ viel auswerten kann. Zurzeit sind es 36 unterschiedliche Datenquellen, die bei uns im Suchportal recherchierbar sind, welche auch ganz unterschiedliche Datenschemata beinhalten. Da haben wir zum einen sehr gute Verbunddaten aus unseren Bibliotheksverbänden: Das sind natürlich wunderbare bibliothekarische Metadaten. Aber wir sind auch schon den Weg gegangen, dass wir reine Excel-Listen von Zeitschriftenmetadaten in unserem Index verarbeitet und so zum Beispiel die Media Perspektiven für *adlr.link* erschlossen haben.

Man kann die bei uns recherchieren, die Artikel erscheinen dann in den Suchergebnissen mit einem direkten Link auf den Volltext.

Aber wir beschränken uns nicht nur auf Texte, sondern wir wollen auch audiovisuelle Inhalte erschließen. Das sind zum einen natürlich sehr viele DVDs und Videos, die in den Bibliotheken im Bestand sind, aber auch Spezialsammlungen. Wir haben zum Beispiel eine Bildersammlung der Margaret Herrick Library in Los Angeles über die frühe Zeit der Filmindustrie in Hollywood. Wir haben ausgewählte Filme aus dem Internet Archive, bei denen die Metadatenqualität leider sehr schlecht ist, aber die Filme sind da und man kann sie auch bei uns finden und sofort aufrufen.

Und was wir gerade planen, ist, dass wir Ergebnislisten auch als Forschungsdaten verfügbar machen wollen. Bis jetzt ist das so: Man sucht etwas und erhält eine Ergebnisliste. Die möchten wir zukünftig für die Nutzer grafisch aufbereiten: Wenn jemand nach einem bestimmten Begriff sucht, dann wollen wir nicht nur zeigen, wie viel Literatur es zu diesem Suchbegriff gibt, sondern auch in welchen Zeiträumen die Literatur erschienen ist. Gibt es da bestimmte Schwerpunkte? Gibt es irgendwelche Peaks von Erscheinungsjahren, wo zu einem bestimmten Thema besonders viel publiziert wurde? Das sind bei unserem Katalog bisher alles hauptsächlich Sekundärquellen, aber die Wissenschaft macht ja vor Sekundärquellen nicht halt, und gerade die Medien- und Kommunikationswissenschaft nicht. Wesentliches Instrument sind tatsächlich die Primärquellen. Und da sind wir bei den Archiven. Wir wissen aus vielen Gesprächen mit Wissenschaftlern aus dem Fach, dass diese sich bei YouTube ihr Untersuchungsmaterial besorgen, weil zum Beispiel ganze Folgen von „Die Pyramide“ mit Dieter Thomas Heck aus den 80er-Jahren dort vorhanden sind. Und dann behelfen sich die Wissenschaftler teilweise damit, ihre Primärquellen und damit ihre Untersuchungsgegenstände in solchen Videoportalen zu finden. Das kann kein dauerhafter Zustand sein.

Nun gibt es ja die Erklärung der Intendanten, welche unterstreicht, wie wichtig die audiovisuellen Medien, die in den öffentlich-rechtlichen Archiven liegen, als wissenschaftliches Quellenmaterial sind und dass sie ein besonderes Kulturgut darstellen. Dass es hier eine grundsätzliche und einheitliche Regelung gibt, dass die Wissenschaft Zugang zu den Archiven erhalten kann, ist erstmal ganz wunderbar. Ein ganz praktisches Problem hat ein Wissenschaftler jedoch, wenn er jetzt eine bestimmte Sendung sucht, weil er darüber forschen möchte, aber nicht genau weiß, in welchem Archiv jetzt das Sendungsmaterial eigentlich liegt. Das ist dann bei neun ARD-Anstalten mitunter ein bisschen schwierig herauszufinden und bindet natürlich auch Ressourcen bei den Archiven.

Gleichwohl ist es so, wie wir das ja bei dieser Tagung nun auch gehört haben, dass die Archive ihre Metadaten ständig verbessern. Wenn die Archive etwa durch Audiominig ganze Transkripte in die Metadaten der Sendungen aufnehmen können, dann ist das natürlich interessant für die Forschung, etwa weil man herausbekommen kann, welche Themen in Nachrichtensendungen vorkamen oder was in Interviews zu einem Thema in einem O-Ton geäußert wurde. Sowas kann man über ein Suchportal zum Beispiel sehr gut vorakzessieren, sodass man als Wissenschaftler gleich bei der Recherche angezeigt bekommt, dass es da bestimmte Zeiträume gibt, wo in den Nachrichtensendungen besonders viel über ein Thema berichtet wurde. Und nicht zuletzt kann man natürlich solche Metadaten auch verlinken mit Norm- oder mit Geodaten, sodass man dann auch noch ganz andere Recherchen damit anstellen könnte.

International ist schon vieles möglich. In Frankreich gibt es eine Pflichtabgabe für audiovisuelle Medien, die es in Deutschland nicht gibt. In den Niederlanden ist es mittlerweile auch so, dass 70 Prozent des Bild- und Tonerbes öffentlich zugänglich sind. In den USA gibt es die Fair Use-Regel, die viele dazu nutzen, um bei YouTube ganze Sendungen für wissenschaftliche Zwecke hochzuladen. Und auch das Internet Archive beruft sich ja vielfach auf diese Fair Use Regel, um ganze Sendungen mitzuschneiden. Wenn man im Internet Archive ein bisschen recherchiert, dann findet man zum Beispiel auch die Nachrichtensendungen zu den Terroranschlägen vom 11. September 2009 mit Transkripten, die man sich sofort dort angucken kann und mit denen man dort forschen kann.

Unser Projekt ist ambitioniert und kann nur in Zusammenarbeit mit Ihnen wirklich fruchten. Was wir vorhaben, ist einen Archivkatalog für ARD und ZDF zu erstellen, in dem man jetzt nicht das gesamte Material, was die Archive haben, durchsuchbar macht, sondern das, was in den Programmen von ARD und ZDF einmal gesendet und damit veröffentlicht wurde. Uns geht es da hauptsächlich um die

Metadaten, nicht um den Content selber. Wir wollen jetzt nicht gleich noch einen Link einbauen, sodass sich die Wissenschaftler von zuhause sämtliche Sendungen anschauen können, das ist gar nicht unser Ziel. Aber wenn ein Wissenschaftler bei uns recherchieren kann und dann allein durch die Metadaten viele Informationen über die Sendung erhält und auch die Information, in welchem Archiv das liegt, dann ist demjenigen schon mal sehr geholfen.

Warum wollen ausgerechnet wir das machen? Das hat damit zu tun, dass wir viel Erfahrung im Aufbau eigener Indizes gesammelt haben. Unser Katalogindex „finc“ enthält 150 Millionen Datensätze. Es ist ein Open Source-Index und wir haben eine Open Source-Suchmaschine namens VuFind, mit der man diesen Index durchsuchen kann. Wir haben den Fachinformationsdienst und damit seit drei Jahren sehr gute Erfahrungen gemacht, wie man Wissenschaftler mit dem umfassenden Nachweis und der Bereitstellung von Sekundärliteratur wunderbar zufriedenstellen kann. Und nicht zuletzt haben wir durch den Fachinformationsdienst auch den direkten Draht zu den Wissenschaftlern. Schließlich entwickeln wir unsere Software auch stets weiter. Insbesondere für den Fachinformationsdienst mussten wir die VuFind-Suchmaschine eigens anpassen, um sie für unsere Bedürfnisse herzurichten. Wir brauchten eine Nutzer-Authentifikation, damit wir auch wissen, dass es auch wirklich Wissenschaftler sind, die sich bei uns anmelden. Also auch dort haben wir Erfahrungen.

Und was ist der Vorteil für die Rundfunkanstalten? Sie haben wunderbare, Sie haben gut strukturierte Metadaten in Ihren Archiven. Das ist eigentlich, wenn man so will, der Schatz, an den wir ein wenig ran möchten. Das sind Metadaten, die das audiovisuelle Kulturerbe Deutschlands repräsentieren. Durch unsere Suchmaschine kann man die Archive tatsächlich von diffusen Rechercheanfragen entlasten, weil die Wissenschaftler dann selbst recherchieren können. Für die Archive ändert sich im Prinzip wenig, weil wir nur die Daten nutzen möchten, die ohnehin für die Produktionsarchive erstellt werden und die die Rundfunkanstalten ja schon intern nutzen. Es sollen also keine zusätzlichen Daten erfasst werden.

Und am Ende könnte dann tatsächlich noch so was stehen, was Mitte der 80er-Jahre schon einmal angedacht wurde, nämlich eine deutsche Mediathek oder – was es eigentlich noch eher trifft, weil es ja um Metadaten geht – eine Nationalmediographie des öffentlich-rechtlichen Rundfunks. •