

# Informationstechnologische Unterstützung der Archivierung biographischer Interviews und Erinnerungszeugnisse

Munir Salman, Felix Engel, Almut Leh und Matthias Hemmje

## 1. Ausgangspunkt

Wie die Wissenschaft insgesamt, so stehen auch die traditionellen Gedächtnisinstitutionen – Archive, Bibliotheken und Museen – vor der Herausforderung der digitalen Transformation. Als Zugang zum kulturellen und wissenschaftlichen Erbe bilden sie die Grundlage für die geistes- und kulturwissenschaftliche Forschung. Auch im Prozess der Digitalisierung kommt ihnen mit der Forderung nach einer digitalen Forschungs- und Informationsinfrastruktur eine Schlüsselfunktion zu. Konkrete Herausforderungen sind dabei unter anderem auch die digitale Langzeitarchivierung, das Management digitaler Forschungsdaten und die Verfügbarmachung der digitalen Daten. Dies umfasst unter anderem inhaltlich deren Anreicherung mit Meta-Informationen, welche zum einen den suchenden Zugriff und zum anderen deren Nachnutzung auch langfristig sichern sollen.

Ein Beispiel einer solchen Institution ist das im Folgenden beschriebene Archiv „Deutsches Gedächtnis“ für qualitative Forschungsdaten und subjektive Erinnerungszeugnisse, das audiovisuelle, textliche und bildliche Dokumente teils in digitaler, teils in analoger Form handhabt. Das Archiv „Deutsches Gedächtnis“ ist das Forschungsdatenrepositorium des Instituts für Geschichte und Biographie der FernUniversität in Hagen (vgl. zum Folgenden Leh 2015; 2018). Hier werden biographische Interviews archiviert und zur Verfügung gestellt, die seit den späten 1970er Jahren im Rahmen von Oral-History-Projekten am Institut und seinen Vorläuferprojekten durchgeführt wurden. Hinzu kommen Interviews aus Forschungen Dritter, die dem Format des lebensgeschichtlichen narrativen Interviews entsprechen. Zurzeit sind es rund 3.100 Interviews aus mehr als einhundert Projekten verschiedener Disziplinen, davon zwei Drittel aus eigenen und ein Drittel aus externen Projekten. Die Audiointerviews sind vollständig digitalisiert, während bei den rund 650 Videointerviews die Digitalisierung der älteren analogen Aufzeichnungen noch aussteht. Rund zwei Drittel aller Interviews sind transkribiert, einige allerdings nur teilweise wörtlich, in anderen Teilen zusammenfassend. Außerdem werden Interviewprotokolle zum Kontext archiviert. Im Online-Archiv „Deutsches Gedächtnis“, einer Kooperation mit dem Center für Digitale Systeme (CeDiS) der Freien Universität Berlin, ist ein aktuell kleiner, aber wachsender Teil der Interviews nach vorheriger Anmeldung online zugänglich.

Neben den biographischen Interviews bietet das Archiv „Deutsches Gedächtnis“ Zugang zu Autobiographien, Familienchroniken, Tagebüchern, Fotos und Briefsammlungen. Rund 1.000 solcher schriftlichen Bestände sind aktuell im „Deutschen Gedächtnis“ archiviert. Zum Teil handelt es sich um Originale, zum Teil um Kopien der

Dokumente. Zunehmend werden auch die Textquellen in elektronischer Form (Scans oder Textdateien) archiviert.

Das „Kempowski-Archiv“ verfügt über 8.000 biographische Dokumente unveröffentlichter Autobiographien in Deutschland. Am Institut für Geschichte und Biographie der FernUniversität in Hagen wurde Anfang der 1990er Jahre eine Nebenstelle des „Archives für unpublizierte Autobiographien“ aufgebaut, in der gut dreihundert Dokumente dieser Sammlung unter dem Namen „Kempowski-Archiv“ einer weiteren wissenschaftlichen Nutzung zur Verfügung gestellt werden.

Das „Roeßler-Archiv“ ist eine Sammlung von rund 76.000 Schulaufsätzen, die das Pädagogenehepaar Wilhelm und Elfriede Roeßler in den 1950er Jahren mit Unterstützung der Schulämter und Kultusministerien im gesamten Bundesgebiet schreiben ließ. Schüler und Schülerinnen aller Altersstufen und Schultypen wurden aufgefordert, innerhalb des Deutschunterrichtes unbenotete Erlebnis- und Besinnungsaufsätze zu bestimmten Themen zu schreiben, wie Familie und deren Umwelt, Verhältnis zum eigenen Körper, Verhalten gegenüber Erwachsenen und untereinander, Einstellung zu Schule, Beruf, Freizeit und verschiedenen Aspekten des öffentlichen Lebens sowie Erinnerungen an die Kriegs- und Nachkriegszeit (vgl. Abels/Krüger/Rohrman 1989). Eine erste Auswertung auf Basis von 20.000 Aufsätzen veröffentlichte Wilhelm Roeßler 1957 unter dem Titel „Jugend im Erziehungsfeld“. Das Ehepaar Roeßler überließ seine umfangreiche Sammlung Ende der 1980er Jahre der FernUniversität. Die Aufsätze sind heute als „Roeßler-Archiv“ im Archiv „Deutsches Gedächtnis“ zugänglich.

## **2. Problembeschreibung und Herausforderungen**

Von den biographischen Interviews im Archiv „Deutsches Gedächtnis“ sind 80 Prozent Audiointerviews – überwiegend analog aufgezeichnet, inzwischen vollständig digitalisiert im mp3- und wav-Format – und 20 Prozent Videointerviews, von denen etwa die Hälfte digitalisiert vorliegt. Darüber hinaus sind um die 70 Prozent der Interviews als Skript verfügbar. Die Erschließung der biographischen Quellen erfolgt zum größten Teil manuell. In der elektronischen Datenbank sind alle Dokumente – Interviews, Text- und Bilddokumente – verzeichnet und mit bis zu 124 Kriterien beschrieben. Nicht in der Datenbank erfasst sind die Schulaufsätze. Die Kriterien geben Auskunft über die Herkunft des Materials, die biographischen Daten des Interviewpartners bzw. des Autobiographen vom Geburtsjahr über die familiäre und regionale Herkunft, den schulischen und beruflichen Werdegang sowie die familiäre Entwicklung bis hin zu politischen und religiösen Orientierungen des Interviewpartners oder Autobiographen, seiner Eltern und seines Lebenspartners. Bei den Interviews ist außerdem der Stand der Archivierung festgehalten, wie etwa Art und Menge der Datenträger, Zeitpunkt und Dauer des Interviews, Bearbeitungsstand usw. Für die schriftlichen Dokumente gibt es ein Textfeld „Kurzbeschreibung“, das Raum für eine freie Beschreibung bietet. Die erschlossenen Inhalte sind/werden in einer Access-Datenbank aufgenommen; diese werden wiederum mit dem System SECONDO des Lehrgebiets Datenbanksysteme für neue Anwendungen der FernUniversität in Hagen

manuell synchronisiert.<sup>1</sup> Das System SECONDO stellt die Metadaten und die Transkriptionen über eine webbasierte Benutzungsschnittstelle zur Verfügung. Die Suche nach Interviews und Inhalten kann auf Basis der Metadatenbeschreibung und als Volltextsuche über die Transkripte erfolgen.

Aktuell existiert lediglich für die Interviews eine Online-Archivumgebung, wobei angesichts des hohen Aufwandes der Erschließung der Interviews für die Online-Bereitstellung – namentlich die Zeitkodierung für eine synchrone Darstellung von Audio/Video und Transkript – die Zahl der online bereitgestellten Interviews nur langsam wächst (Leh/Tausendfreund 2017). Für die Textquellen des Kempowski-Archivs und des Roeßler-Archivs existiert keine Online-Umgebung. Da die externe Nachfrage nach der Nutzung der vorhandenen Archivinhalte ständig steigt, bestehen neben der Verbesserung der Suchfunktionen weitere Herausforderungen im Bereich der Skalierbarkeit (technische Einschränkung, Access Datenbank) und der Absicherung der Zugänglichkeit (Zugriffsrechte, Benutzerfreundlichkeit und Usability) der bestehenden Systeme und der darunterliegenden IT-Infrastrukturen.

### 3. Ansatzpunkte und Ziele für die weitere Arbeit

Im Folgenden werden Ansatzpunkte und Ziele eines künftigen Arbeitsprogramms umrissen, das die genannten Herausforderungen adressiert. In Aussicht genommen wird zunächst ein phasenweises Vorgehen mit definierten Zielen. Die Herausforderung liegt dabei in einer ordentlichen Konzeption und Umsetzung der einzelnen Zwischenziele. Zuerst sollen daher die Schnittstellen des Archives „Deutsches Gedächtnis“ mit den oben beschriebenen Archivierungssystemen analysiert und auf Erweiterbarkeit untersucht werden. Anschließend soll die Erschließung elektronischer und traditioneller autobiographischer Quellen mit Berücksichtigung der verschiedenen Hintergrundinformationen erfolgen. Dabei sollen möglichst automatische, semi-automatische und manuelle Erschließungsmethoden im Sinne der Annotationen und Verschlagwortung berücksichtigt werden. Darüber hinaus sollen die Inhalte (Digitalisate) mit relevanten anderen Quellen, Informationen und Medieninhalten verknüpft werden. Die verbesserte Unterstützung der Suchfunktionen nach Digitalisaten stellt dabei eine zentrale Herausforderung dar, die sowohl den Einsatz von Merkmalsextraktionsverfahren als auch von semantischen Webtechnologien zum Anreichern von Metadaten mit maschinenlesbaren semantischen Erschließungen, Beschreibungen und Klassifikationen der Inhalte erforderlich macht. Durch den Einsatz von semantischen Webtechnologien sollen dabei die semantische Integration wie auch die automatische Suche und der Zugriff auf multimediale Inhalte (wie biographische Interviews, Autobiographien, Fotos, Tagebücher, Briefsammlungen usw.) sowie die Präsentation, Darstellung, Bündelung, Publikation, Archivierung und Erhaltung dieser multimedialen Inhalte als Informations- und Wissensressourcen mit nachhaltigem Mehrwert (engl. *Assets*) unterstützt werden. Darüber hinaus soll das System zukünftig die Erschließung von multimedialen Digitalisaten durch den Einsatz von Methoden der Merkmalsextraktion (engl. *Feature Extraction*) und des maschinellen Lernens (engl.

---

<sup>1</sup> Ralf Hartmut Güting stellt dem Archiv „Deutsches Gedächtnis“ dankenswerterweise das von ihm entwickelte Datenbanksystem samt Support zur Verfügung. Näheres zu Secondo: <http://dna.fernuni-hagen.de/Secondo.html/index.html>.

*Machine Learning*, z. B. Klassifikation, engl. *Classification*, Erkennung von Namen, engl. *Named Entity Recognition* usw.) unterstützen.

Die Langzeitarchivierung und Erhaltung der digitalen Archivbestände auf der Basis von Emulation und Migration sollen möglichst so unterstützt und erweitert werden, dass die Inhalte (Digitalisate) über viele Jahre erreichbar, ausspielbar und verwendbar bleiben (trotz möglicherweise veralteter Formate). Der Zugriff auf die Archive soll darüber hinaus online über das Internet möglich sein. Dabei sollen ethische Kriterien wie die Anonymisierung berücksichtigt und sichergestellt werden. Durch die Nutzung cloudbasierter Infrastrukturen soll das System skalierbar und erweiterbar sein. Das angestrebte Integrationskonzept soll darüber hinaus ein globales Schema für die Abbildung und die Integration von heterogenen Metadaten (wie z. B. aus sozialen Netzwerken) innerhalb der integrierten Archivsysteme liefern. Auch soll überprüft werden, inwieweit die Open-Source-Technologien DSpace (DSpace - A Turnkey Institutional Repository Application 2018) und andere vergleichbare für Archivsysteme vorgesehene Standardtechnologien bereits genutzt werden bzw. eine Nutzung von DSpace oder ähnlicher Open-Source-Systeme für die Archivierung als Basissystem für das Archiv „Deutsches Gedächtnis“ und vergleichbare Archive möglich ist.

Ein Teil des vorgestellten Archivs besteht aus digitalisierten Audiointerviews. Ob überhaupt ein Open Archival Information System (das heißt ein bezüglich des ISO-Standards 14721 OAIS-konformes) Archivsystem für diese Digitalisate vorhanden ist und wie dieses aufgebaut ist, ist zu klären. Auch hier soll ein zukünftig zu entwickelndes System genutzt werden, um die Eingliederung der Digitalisate in die jeweiligen funktionalen Bereiche und Einheiten der bereits vorhandenen Archive zu ermöglichen. Benötigt wird der Zugriff auf die Audiointerviews im Wesentlichen für zwei Benutzergruppen: für Informatiker und für Historiker. Auf den Audiointerviews soll eine maschinelle, webbasierte, semantische Suche auf der Basis eines festgelegten, fachspezifischen Vokabulars möglich sein, die durch das System unterstützt wird. In diesem Zusammenhang sind für die Repräsentation der lokalen Schemata der verteilten Archive und deren semantische Integration in ein globales Schema auch die Themen Taxonomien/Ontologien und das Abbilden von Taxonomien/Ontologien aufeinander relevant. Gegenstand der künftigen Arbeit ist für den Anwendungsfall der automatischen Suche auf Audioaufnahmen ebenfalls die Untersuchung der Potentiale und Herausforderungen der Merkmalsextraktion und des Maschinellen Lernens auf der Basis einer Transformation von Sprachaufnahmen in Text (engl. *speech to text*). Die Ergebnisse der Suche sollen jedoch nicht direkt zur Verfügung gestellt werden. Geplant ist vielmehr, Links zu angefragten Ressourcen in den Archiven bereitzustellen. Die Möglichkeit einer direkten Ausgabe der Suchergebnisse ist im Rahmen der Arbeit im Zusammenhang mit der Empfehlung geeigneter Referenzarchitekturen, Standards und Technologien dann nachfolgend ebenfalls anzusprechen. Die in zukünftigen Arbeiten zu adressierenden Aufgaben bewegen sich somit im Spannungsfeld zwischen Metadatenmanagement, Semantik-Web-Technologien, Archivsystemen, Langzeitarchivierung und verteilten Informations- und Wissensmanagementsystemen im Anwendungsbereich der Digitalisierung, Verwaltung, Speicherung und Zugriffsunterstützung sowie langfristigen Sicherung und Erhaltung von multimediale Forschungsdaten für die biographische Forschung in verschiedenen Disziplinen.

Um die oben ausgeführten Herausforderungen bearbeiten zu können, wurden aktuell folgende Teilaufgaben für die weitere Arbeit definiert: Recherche verwendeter

Technologien in den genannten Archiven/Sammlungen, Recherche und Entwurf von geeigneten Metadaten-Standards und -Schemata für die Domäne der biographischen Forschung, Identifizierung geeigneter Infrastruktur-Provider. Das zu erstellende Lösungskonzept soll die Langzeitarchivierung unter Einhaltung nationaler und internationaler Standards (DIN, ISO usw.) sowie unter Einhaltung ethischer Kriterien im Hinblick auf die Integration im gesamten System unterstützen. Weiterhin soll die Entwicklung einer *Proof-of-Concept*-Implementierung angestrebt werden.

Ziel der zukünftigen Arbeit ist es somit, zunächst eine ausführliche Analyse durchzuführen, um festzustellen, welche relevanten Archivinhalte und verwandte Repositorien von zugehörigen Ressourcen bereits existieren, aus welchen funktionalen Bereichen und Einheiten im Sinne des ISO-Standards 14721 OAIS diese bestehen und wie diese funktionalen Bereiche und Einheiten sowohl konzeptuell als auch technisch in ein zentrales System integriert werden können. Sofern es gelingt, ein OAIS-konformes Archivsystem zu erstellen, wäre nachfolgend auch eine Zertifizierung desselben auf der Basis verwandter Standards (z. B. ISO 16363 Audit and Certification) ohne weiteres möglich. Ein weiteres Ziel ist es, die Potentiale und Probleme für eine technische und semantische Integration von verteilten Archivinhalten (wie z. B. das Archiv „Deutsches Gedächtnis“, zu dem wiederum das Kempowski- und das Roeßler-Archiv gehören) mit zum Teil unterschiedlichen Erschließungsgraden in eine zentrale Zugriffs- und Verwaltungs-, Produktions- und Archivierungsplattform zu integrieren, um die oben genannten Aufgaben (Erschließung, Suche, Finden, Präsentation, Bündelung, Publikation, Paketierung, Archivierung, Erhaltung) über einen zentralen Zugangspunkt gebündelt durchführen zu können.

Die Integration der Systeme soll entsprechend über Open-Source-Technologien sowie über standardisierte Schnittstellen wie zum Beispiel die RESTful API (DSpace - A Turnkey Institutional Repository Application 2018), die OAI-PMH (REST API - DSpace 6.x Documentation - DuraSpace Wiki 2018), die SWORD (SWORD - The SWORD website 2018), OpenSearch (Home - OpenSearch 2018), LinkedData-Schnittstelle (Becker 2014) usw. realisiert werden. Da der ISO-Standard 14721 OAIS in der aktuellen Version zu verteilten Archiven noch keine ausreichend flexiblen Topologien für die Verteilung von Inhalten und Archivfunktionen vorsieht, sondern aktuell nur mit kooperierenden Archivtopologien eher Spezialfälle von Verteilungstopologien beschreibt, soll in der weiteren Arbeit die Frage, welche der beispielhaft betrachteten Archivinfrastrukturen als funktionale Bereiche und funktionale Einheiten im Sinne von OAIS integriert werden können insbesondere mit Blick auf mögliche OAIS-konforme Verteilungstopologien betrachtet werden.

#### **4. Vorarbeiten im Bereich Digitale Bibliotheken, Medienarchive und Langzeitarchivierung**

Das Lehrgebiet für Multimedia und Internetanwendungen der FernUniversität in Hagen beschäftigt sich seit vielen Jahren mit Themen rund um die Langzeitarchivierung von digitalen Daten. Dies umfasst inhaltlich deren Anreicherung mit Informationen, welche deren Nachnutzung auch in der Langzeit sichern sollen. Im Zuge dieser Tätigkeiten beteiligt sich das Lehrgebiet auch an der Weiterentwicklung von Archivierungsstandards in einer internationalen Kooperation.

Forschungsbeiträge leistete das Lehrgebiet bereits in unterschiedlichen, von der EU geförderten Projekten. Beiträge des Lehrgebietes befassten sich in diesen Projekten mit Themen zur OAIS-konformen Darstellung von Forschungsdaten sowie deren Überführung in Archive. Aktuell liegt der Forschungsschwerpunkt des Lehrgebietes auf der Verarbeitung von Informationen aus multimedialen Daten, unter anderem auch mittels maschineller Lernverfahren, um deren umfassende Erschließung, Erhaltung und Nachnutzbarkeit langfristig sicherzustellen, was nicht zuletzt die Reproduzierbarkeit von Analysen und Experimenten sichert. Dies umfasst die Extraktion von Informationen, deren semantische Repräsentation und deren Integration in verfügbare Wissensressourcen (z. B. Ontologien) sowie Visualisierung und Zugriff. Aktuell arbeitet das Lehrgebiet dazu an einer webbasierten Software, die das Management und den einfachen Zugriff auf diese Daten und zugehörigen Wissensressourcen aus verschiedenen Quellen wie z. B. aus sozialen Netzwerken und unterschiedlichen Archivsystemen realisiert. Im Bereich der digitalen Bibliotheken, Medienarchive und der Langzeitarchivierung sind in diesem Zusammenhang die nachfolgend beschriebenen und bereits erfolgreich abgeschlossenen Projekte zu erwähnen.

Das Projekt *Early Modern Thought Online (EMTO)* war ein DFG-gefördertes Projekt, das die Digitalisierung philosophischer Handschriften behandelte und diese der Erforschung der Philosophie der frühen Neuzeit und angrenzender Fachgebiete auf eine einfache (unkomplizierte) Weise zur Verfügung stellte. Indem der Zugriff auf Fachwissen erleichtert und optimiert wurde, wurde auch ein inhaltlicher Mehrwert erzielt. Entwickelt wurde dafür ein System, das mehr als 14.000 Digitalisate aus europäischen Datenquellen mittels einer webbasierten Benutzungsschnittstelle zur Verfügung stellt. Die Quellen, bestehend aus nationalen und internationalen Bibliotheken, wurden mit Hilfe einer Mediatorarchitektur in das System integriert. Die Digitalisate wurden von Experten mit Metadaten und Annotationen angereichert und beschrieben.

Das Projekt *Collate* war ein EU-gefördertes Projekt zur Unterstützung der Filmarchivierung. Dabei wurde eine webbasierte Plattform entwickelt, die mehr als 6.000 historische Filme inklusive Metadaten, Zensurdokumenten und mehr als 18.000 Dokumentbilder als zusätzliche Ressourcen verwaltet und bereitstellt. Die Langzeitarchivierung und Verfügbarkeit dieser Inhalte wird durch OAI-PMH und europäische Standards auf Basis des OAI-Referenz-Modells sichergestellt. Dabei stellt der OAI-PMH-Standard eine zukunftsweisende Möglichkeit dar, ein OAIS-konformes Archivsystem auch mit weiteren Archivsystemen jederzeit technisch integrieren zu können. Auf diese Art und Weise wird auch eine langfristige Erweiterbarkeit der entstehenden Lösung sichergestellt. Das System ist modular mit Hilfe von Open-Source-Software aufgebaut und kann jederzeit skaliert und durch cloudbasierte Speicher- und Verarbeitungsressourcen erweitert werden.

Das EU-geförderte Projekt *Sustaining Heritage Access through Multivalent Archiving (SHAMAN)* war ein Infrastrukturprojekt zur digitalen Langzeitarchivierung von digitalen Kulturgütern, mit dem Ziel, einen übergreifenden Ansatz für die digitale Archivierung (engl. *Digital Preservation, DP*) der nächsten Generation, basierend auf Cloudtechnologien zu schaffen. Darüber hinaus stellte das Projekt eine Reihe von Werkzeugen zur Analyse, Aufnahme, Verwaltung und Wiederverwendung von Informationsobjekten und Daten in verteilten (heterogenen) Bibliotheken und Archiven bereit. Um diese Ziele zu erreichen, stellte die SHAMAN-Kerninfrastruktur cloudbasierte, linguistische, semantische und Analyse-, Erschließungs- und Zugriffsmethoden

zur Unterstützung von DP-Prozessen bereit. Das SHAMAN-Rahmenwerk besteht aus integrierten Daten, digitalen Bibliotheken und persistenten Archiven sowie kontextuellen Repräsentations- und Annotationsmechanismen für semantisch integrierte Datentypen (Dokument-, Media-, CAD- und Forschungsdaten) sowie generell wissensbasierten Informationssammlungen.

Das EU-geförderte Projekt *Sensor Enabled Affective Computing for Enhancing Medical Care Sensecare* hat zum Ziel, zukünftige Prozesse und Systeme im Bereich des Gesundheitswesens mit Hilfe von sensorischen und maschinellen Lerntechnologien zu verbessern und voranzubringen, um z. B. emotionale (affektive) und kognitive Erkenntnisse über das Wohlbefinden der Patienten zu gewinnen und ihnen eine effektivere Behandlung in verschiedenen medizinischen Bereichen zu ermöglichen. Hierfür wurden Technologien und Methoden entwickelt, die die wachsenden Gesundheitskosten im Bereich der Demenz und der damit verbundenen kognitiven Beeinträchtigungen, die die europäischen Bürger belasten und die bis 2030 schätzungsweise über 250 Milliarden Euro kosten werden (Alzheimer Europe - Research - European Collaboration on Dementia - Cost of dementia - Prognosis to 2030 2018), zu verringern (Engel et al. 2016). Um diese Ziele zu erreichen, stellt *Sensecare* eine cloudbasierte Online-Plattform zur Archivierung und Bereitstellung solcher Mess- und Analysedaten aus dem Gesundheitswesen zur Verfügung.

Außerdem ist das Lehrgebiet für Multimedia und Internetanwendungen Mitglied des *Network of Expertise in long-term Storage and availability of digital Resources in Germany*, kurz NESTOR, ein deutsches Netzwerk für die Langzeitarchivierung und Langzeitverfügbarkeit digitaler Ressourcen. Es vernetzt unterschiedliche Institutionen, die sich mit der Langzeitarchivierung digitaler Daten befassen. Weiterhin dient NESTOR als Anlaufstelle und als Informationsplattform für Fragen rund um die digitale Langzeitarchivierung (nestor - Home 2018).

Aus allen oben beschriebenen Vorarbeiten entstand in den vergangenen Jahren eine Technologie zur Unterstützung eines Wissensmanagement-Ökosystem-Portals (engl. *Knowledge Management Ecosystem Portal*, KM-EP). Die KM-EP-Technologie ist in den Projekten *Realising an Applied Gaming Ecosystem (RAGE)* (RAGE 2018; Salman et al. 2017) und *EDISON* (Demchenko et al. 2016) bereits produktiv eingesetzt worden und soll nun auch für die Erarbeitung von Lösungen in den eingangs am Beispiel des Archivs „Deutsches Gedächtnis“ beschriebenen Handlungsfeldern herangezogen werden.

## 5. Zusammenfassung und Ausblick

Traditionelle Gedächtnisinstitutionen – Archive, Bibliotheken und Museen – stellen die Grundlage für die geistes- und kulturwissenschaftliche Forschung dar. Sie stehen derzeit vor der Herausforderung der digitalen Transformation. Aktuell besteht das Problem dieser Einrichtungen darin, dass Inhalte aus unterschiedlichen Gründen nur zum Teil online verfügbar sind, während die externe Nachfrage nach der Nutzung der vorhandenen Archivinhalte ständig steigt. Herausforderungen sind dabei insbesondere im Umfeld der Verbesserung der Suchfunktionen, aber auch der Skalierbarkeit (technische Einschränkung, Access Datenbank) und der Absicherung der Zugänglichkeit (Zugriffsrechte, Benutzerfreundlichkeit und Usability) zu sehen.

Ziel der zukünftigen Forschungsarbeit des Lehrgebiet für Multimedia und Internetanwendungen ist es, zunächst eine ausführliche Analyse durchzuführen, um festzustellen, welche relevanten Archivinhalte und verwandte Repositorien von zugehörigen Ressourcen bereits existieren, aus welchen funktionalen Bereichen und Einheiten diese bestehen und wie diese funktionalen Bereiche und Einheiten sowohl konzeptuell als auch technisch in ein zentrales System integriert werden können.

#### LITERATUR

- Abels, Heinz, Heinz-Hermann Krüger und Hartmut Rohrmann (1989): „Jugend im Erziehungsfeld“. Schüleraufsätze aus den fünfziger Jahren im Roeßler-Archiv, in: BIOS - Zeitschrift für Biographieforschung und Oral History, Jg. 2, 139-150.
- Alzheimer Europe - Research - European Collaboration on Dementia - Cost of dementia - Prognosis to 2030 [WWW Document] (2018).  
[https://www.alzheimer-europe.org/Research/European-Collaboration-on-Dementia/Cost-of-dementia/Prognosis-to-2030/\(language\)/eng-GB](https://www.alzheimer-europe.org/Research/European-Collaboration-on-Dementia/Cost-of-dementia/Prognosis-to-2030/(language)/eng-GB) (12.4.2018).
- Becker, Pascal-Nicolas (2014): Repositorien und das Semantic Web - Repositorieninhalte als Linked Data bereitstellen 94.
- Demchenko, Yuri, A.S.Z. Belloum, Wouter Los, Tomasz Wiktorski, Andrea Manieri, Holger Brocks, Jana Becker, Dominic Heutelbeck, Matthias Hemmje und Steve Brewer (2016): EDISON Data Science Framework: A Foundation for Building Data Science Profession for Research and Industry, in: 2016 IEEE International Conference on Cloud Computing Technology and Science (CloudCom). Presented at the 2016 IEEE International Conference on Cloud Computing Technology and Science (CloudCom), IEEE, Luxembourg, Luxembourg, pp. 620-626. <https://doi.org/10.1109/CloudCom.2016.0107>.
- DSpace - A Turnkey Institutional Repository Application (2018). Duraspace.org.  
<https://duraspace.org/dspace/> (12.4.2018).
- EMTO: Lehrgebiet Multimedia und Internetanwendungen (2018). <http://www.lgmmia.fernuni-hagen.de/forschung/projekte/abgeschlossene-projekte/emto.html> (11.6.2018).
- Engel, Felix, Raymond Bond, Alfie Keary, Maurice Mulvenna, Paul Walsh, Huiru Zheng, Haiying Wang, Ulrich Kowohl und Matthias Hemmje (2016): SenseCare: Towards an Experimental Platform for Home-Based, Visualisation of Emotional States of People with Dementia, in: Marco X. Bornschlegl, Felix C. Engel, Raymond Bond, Matthias Hemmje (Hg.): Advanced Visual Interfaces. Supporting Big Data Applications, Lecture Notes in Computer Science. Springer International Publishing, 63-74.  
[https://doi.org/10.1007/978-3-319-50070-6\\_5](https://doi.org/10.1007/978-3-319-50070-6_5)
- Home - OpenSearch [WWW Document] (2018). <http://www.opensearch.org/Home> (12.4.2018).
- Jannidis, Fotis, Hubertus Kohle und Malte Rehbein (Hg.) (2017): Digital Humanities: Eine Einführung. J.B. Metzler. <https://doi.org/10.1007/978-3-476-05446-3>
- Leh, Almut (2015): Vierzig Jahre Oral History in Deutschland. Beitrag zu einer Gegenwartsdiagnose von Zeitzeugenarchiven am Beispiel des Archivs „Deutsches Gedächtnis“, in: Westfälische Forschungen. Zeitschrift des LWL-Instituts für westfälische Regionalgeschichte, hrsg. von Bernd Walter und Thomas Küster, 65, 255-268.
- Leh, Almut und Doris Tausendfreund (2017): Curation and Dissemination of Lifestory Interviews for the Humanities in: Proceedings of the Second Conference on Biographical Data in a Digital World 2017 (Linz, Austria, November 6-7, 2017), edited by Antske Fokkens, Serge ter Braake, Ronald Sluijter, Paul Arthur and Eveline Wandl-Vogt, 9-15.  
<http://ceur-ws.org/Vol-2119/paper2.pdf>.
- Leh, Almut (2018): Zeitzeugenkonserven. Interviews für nachfolgende Forschergenerationen im Archiv „Deutsches Gedächtnis“, in: Archivar, 71. Jg., Heft 2, 155-157.

nestor - Home (2018).

[https://www.langzeitarchivierung.de/Subsites/nestor/EN/Home/home\\_node.html](https://www.langzeitarchivierung.de/Subsites/nestor/EN/Home/home_node.html)  
(12.6.2018).

RAGE (2018). Home RAGE. <http://rageproject.eu/> (7.6.2018).

REST API - DSpace 6.x Documentation - DuraSpace Wiki (2018).

<https://wiki.duraspace.org/display/DSDOC6x/REST+API> (12.4.2018).

Salman, Munir, Jana Mertens, Michael Fuchs, Dominic Heutelbeck, Matthias Hemmje und Binh Vu (2017): Social Network-Based Knowledge, Content, and Software Asset Management Supporting Collaborative and Co-Creative Innovation, in: Proceedings of the Collaborative European Research Conference CERC 2017. Presented at the CERC 2017, Conference Proceedings, University of Applied Sciences – Karlsruhe, Germany.

SWORD – The SWORD website (2018). <http://swordapp.org/> (12.4.2018).

### **Zusammenfassung**

Traditionelle Gedächtnisinstitutionen – Archive, Bibliotheken und Museen – stehen vor der Herausforderung der digitalen Transformation. Insbesondere, da Inhalte aus unterschiedlichen Gründen nur zum Teil online verfügbar sind, während die externe Nachfrage nach der Nutzung der vorhandenen Archivinhalte ständig steigt. Der Beitrag befasst sich daher mit Ansatzpunkten und Zielen eines künftigen Arbeitsprogramms, das die genannten Herausforderungen adressiert.