

100010001000111110100101  
101101011010000011101101  
0123.0TAGUNG1DES1110010  
10111ARBEITSKREISES1001  
10110011011011001011010  
01010ARCHIVIERUNG01  
1101110VON110000000001  
010111UNTERLAGEN110111  
10111011AUS1100000011100  
110110DIGITALEN1100101  
111001SYSTEMEN0010111  
1101100011000000100  
110012.0UND013.0MÄRZ11111  
001101020191PRAG011  
00101001010010100111  
10011111NATIONALARCHIV101  
100101101101100110011  
100110000 100111100  
101000100 111001101  
101001100110100110110



National Archives  
Czech Republic

110101101000011101101  
010ARCHIVIERUNG01  
01110VON110000000001  
1011UNTERLAGEN110111  
111011AUS1100000011100  
1101DIGITALEN1100101  
100SYSTEMEN0010111  
0010101PRAHA20190  
00101001010010100111

Editor: Karolína Šimůnková, Milan Vojáček

Kolektivní monografie vznikla na základě konference Archivierung von Unterlagen aus digitalen Systemen (Archivace záznamů z digitálních systémů) a souvisejících seminářů, které se konaly ve dnech 12. a 13. března 2019 v Národním archivu v Praze.

© Národní archiv, 2019  
[www.nacr.cz](http://www.nacr.cz)

ISBN 978-80-7469-100-3 (pdf)

23. Tagung des Arbeitskreises

# ARCHIVIERUNG VON UNTERLAGEN AUS DIGITALEN SYSTEMEN

12. und 13. März 2019, Nationalarchiv Prag

## Inhaltverzeichnis

### **LEITIDEEN, PROJEKTE, NETZWERKE**

<b>Ein Überblick zur Archivierung digitaler Unterlagen in Deutschland</b>	<b>9</b>
Kai NAUMANN (Landesarchiv Baden-Württemberg)	
<b>Der Stand und die Perspektiven der digitalen Archivierung in der Tschechischen Republik</b>	<b>16</b>
Milan VOJÁČEK (Nationalarchiv Prag)	
<b>Das Projekt DIMAG. Sachstand 2019</b>	<b>21</b>
Christian KEITEL (Landesarchiv Baden-Württemberg)	

### **BEWERTUNG UND ÜBERNAHME**

<b>Qualitative Bewertung im Hinblick auf Informationsgehalt, Verwendung und Kostenmetrik Juni 2019</b>	<b>35</b>
Martin KAISER (Koordinationsstelle für die dauerhafte Archivierung elektronischer Unterlagen – KOST)	
<b>Übernahme von landwirtschaftlichen Daten aus Fachapplikationen</b>	<b>41</b>
Bernhard STÜSSI, Flurina CAMENISCH (Staatsarchiv Graubünden)	
<b>DCAT-AP.de. Der Metadatenstandard zum Austausch offener Verwaltungsdaten</b>	<b>53</b>
Paul FLAMME (Staatsarchiv Hamburg)	

### **ERHALTUNGSSTRATEGIEN UND ERHALTUNGSMAßNAHMEN**

<b>Alles Blockchain oder was? Nutzen und Grenzen der Distributed-Ledger-Technologie in der digitalen Archivierung</b>	<b>67</b>
Björn STEFFENHAGEN (Landesarchiv Sachsen-Anhalt)	
<b>Vom Umgang mit fehlerhaften Daten im Ingest. Ein Werkstattbericht aus der Entwicklung der DA-NRW Software Suite (DNS).</b>	<b>81</b>
Jens PETERS (LVR-InfoKom)	

## **ERSCHLIEßUNG UND NUTZUNG**

**„Ich packe mein AIP und nehme mit...“ Dateistrukturierung im Digitalen Magazin des Landesarchivs Thüringen (ThELMA)** 91

Svenia POHLKAMP, Konrad MECKEL (Landesarchiv Thüringen)

**Von der Archivvorstufe zum Archiv: (Forschungs-)Datenmanagement im Kontext von Gestaltung und Kunst** 99

Jürgen ENGE, Tabea LURK (Hochschule für Gestaltung und Kunst FHNW Basel)

**Die Indexierung von audiovisuellen Archiven unter Benutzung von Methoden der automatischen Sprach- und Bilderkennung** 120

Pavel IRCING, Jan ŠVEC (Westböhmische Universität in Pilsen)

**Service Design für den digitalen Lesesaal. Ein Werkstattbericht der Staatsarchive Basel-Stadt und St. Gallen** 129

Esther BAUR – Lambert KANSY (Staatsarchiv Basel-Stadt), Martin LÜTHI (Staatsarchiv St. Gallen)

## **AUDS-CAMP**

**WIE GEHEN WIR MIT DIGITALEN MEDIEN ALS BEILAGE EINER PAPIERNEN AKTE UM?**

**Wie gehen wir mit digitalen Medien als Beilage einer papiernen Akte um?** 143

Zusammenfassung: Kai NAUMANN (Landesarchiv Baden-Württemberg)

**ELEKTRONISCH SEI DIE AKTE, HILFREICH UND GUT. BERATUNG VON BEHÖRDEN BEIM UMSTIEG AUF DIE E-AKTE**

**Elektronisch sei die Akte, hilfreich und gut. Beratung von Behörden beim Umstieg auf die E-Akte** 147

Zusammenfassung: Christine FRIEDERICH, Martin SCHLEMMER (Landesarchiv Nordrhein-Westfalen)

**MIT BPMN UND REST ZU EINER KLAREN UND OAIS-REFERENZMODELL-KONFORMEN SOFTWARE-ARCHITEKTUR**

**Mit Bpmn und Rest zu einer klaren und Oais-Referenzmodell-Konformen Software-Architektur** 153

Zusammenfassung: Frank OBERMEIT (Landesarchiv Sachsen-Anhalt)

**Mit BPMN und REST zu einer OAIS-konformen Software-Architektur** 155

Frank OBERMEIT

**VOLKSZÄHLUNG 2021 UND DIE ROLLE DER DIGITALEN ARCHIVE**

**Archivcamp „Volkszählung 2021 und Rolle der digitalen Archive“** 167

Zusammenfassung: Christian KEITEL (Landesarchiv Baden-Württemberg)

<b>Statistik im Archiv – Eine schwierige Beziehung</b>	<b>169</b>
Christian KEITEL (Landesarchiv Baden-Württemberg)	
<b>Zur Problematik der elektronischen Archivierung von Volkszählungen 2011 und 2021 in der Tschechischen Republik</b>	<b>177</b>
Karolína ŠIMŮNKOVÁ, Martin ŠISLER (Nationalarchiv Prag)	
 <b>VALIDIERST DU NOCH ODER ARCHIVIERST DU SCHON?</b>	
<b>Validierst du noch oder archivierst du schon? Der Bedarf eines Workflow-Managements in der Formatverifikation</b>	<b>187</b>
Zusammenfassung: Christine TRÄGER, Daniel WITTMANN (Landesarchiv Thüringen)	
 <b>ARCHIVSTANDARDS UND BEST PRACTICES</b>	
<b>Archivstandards und Best Practices</b>	<b>201</b>
Zusammenfassung: Christine GIGLER (Archiv der Erzdiözese Salzburg)	
<b>Komplexe Realitäten – Komplexe Metadaten. Wie man Standards, Best Practices und Realität zusammenführt</b>	<b>202</b>
Marta RIESS (Internationale Atomenergie-Organisation)	
<b>Aktuelle Standards und Normen in Records Management und beweissicherer Langzeitspeicherung</b>	<b>210</b>
Tomasz KUSBER (Fraunhofer FOKUS), Ulrike KORTE (Bundesamt für Sicherheit in der Informationstechnik), Steffen SCHWALM (msg systems)	
<b>Die nestor-AG „Kooperation der Archive“ und der nestor-Archivstandard</b>	<b>233</b>
Christine GIGLER (Archiv der Erzdiözese Salzburg)	
 <b>DIGITAL OBJECTS AND THEIR PRESERVATION IN ACADEMIC WORLD</b>	
<b>Digital Objects and their Preservation in Academic World</b>	<b>241</b>
Zusammenfassung: Zdeněk VAŠEK (Karls-Universität in Prag)	
 <b>Autorinnen und Autoren</b>	<b>243</b>



# VON DER ARCHIVVORSTUFE ZUM ARCHIV

## (Forschungs-)Datenmanagement im Kontext von Gestaltung und Kunst

*Jürgen Enge, Tabea Lurk*

### Einleitung

Das digitale Archiv der Hochschule für Gestaltung und Kunst (HGK) FHNW Basel wird seit 2017 gemeinsam vom Center for Digital Matter (DIGMA) und der Mediathek aufgebaut. Es beabsichtigt, die Geschichte der Hochschule, ihre digitalen Lehr- und Forschungsmittel sowie die ihrer affilierten Kontexte dokumentarisch zu erschliessen, dauerhaft aufzubewahren und gemeinsam mit anderen Wissensressourcen (z. B. Bibliotheksressourcen) durchsuchbar zu machen. Auch wenn die Hochschule von überschaubarer Grösse ist, zeichnet sie sich durch eine bemerkenswerte Dynamik und Agilität aus, deren Wesenhaftigkeit in den entsprechenden Infrastrukturen reflektiert werden sollte. Am deutlichsten wird dies vielleicht in der Art und Weise, wie mit Datenlieferant\*innen und ihren Beständen umgegangen wird.

Zudem kann das digitale Archiv der HGK als strategisches Instrument begriffen werden, das die Hochschule und ihre Angehörigen auf sanfte Weise an die Digitalisierung heranführt: Die Übernahmeprozesse gestalten sich stets dialogisch; neue Werkzeuge, digital motivierte Gedankengänge und Abläufe werden vermittelt, die künftig frühzeitig(er) berücksichtigt werden oder sogar gemeinsame Entwicklungsprojekte anstossen.

Die Aufnahme der Konvolute und Sammlungen erfolgt stets fallbasiert, auf Initiative von (Einzel-)Personen und nicht aufgrund eines Erlasses. Mit Blick auf die Aktenbildung steht uns dieser interessen geleitete Weg offen, da der Grossteil der aufbewahrten Inhalte keine obligationsrechtlichen Zwecke bedient: Daten, welche aufgrund gesetzlicher Aufbewahrungsvorschriften archiviert werden müssen, werden vom Stab bzw. dem FHNW Präsidium aufbewahrt oder den jeweiligen Stellen zur Übernahme angeboten.<sup>1</sup> Die gesetzliche Revisionsfrist von Abschlussarbeiten beträgt lediglich 10 Jahre und wird an der HGK bisher von den Instituten selbst gemanagt, da im Falle der Revision neben den von uns dokumentierten Thesen und Ressourcen zumeist auch praktische Lehr- und Lernresultate des jeweiligen Abschlusses zur Disposition stehen. Dass die medial überlieferten Inhalte der HGK hochschulintern dennoch

.....  
1 Unsere Bestände enthalten z. B. keine Studierendendaten oder Studieninformationen wie sie die Fachgruppe 8 „Archive der Hochschulen und wissenschaftlichen Institutionen“ im „Verband deutscher Archivarinnen und Archivare“ (VdA) behandelt (vgl. Fussnote 8).

ausgesprochen wichtig erscheinen und kulturelle, bildungspolitische oder strategische Zwecke erfüllen, steht auf einem anderen Blatt.<sup>2</sup> Aber genau hier, an den semantisch wertvollen Inhalten, die fakultativ gesammelt werden können, setzt das digitale Archiv der HGK an<sup>3</sup>. Entsprechend folgt die Aufnahme von Inhalten einfachen Prämissen:

- *Institutionelle Affilierung*: Die angebotenen Inhalte müssen an der HGK entstanden, von ihr/Hochschulangehörigen in Auftrag gegeben oder nachweislich mit Personen, Instituten oder Themen der HGK assoziiert sein, sodass ihre Aufbewahrung der Forschung und Lehre dient und Bezüge benannt werden können.<sup>4</sup>
- *Geklärte Rechtelage*: Die einfachen Nutzungsrechte an den Inhalten müssen der HGK übertragen werden. Inhalte, die dauerhaft für alle Nutzungsgruppen unzugänglich bleiben, finden nur in absoluten Ausnahmefällen Einzug ins digitale Archiv.<sup>5</sup> Kontrollierte Zugangsformen sind hingegen selbstverständlich möglich.
- *Ökonomische Vertretbarkeit*: Es fallen keine Sonder- oder Akquisitionskosten an, die jenseits der Aufwände für die Erschliessung, Vereinnahmung, (dauerhafte) Archivierung in der Mediathek zu Buche schlagen.
- *Kulturelle, historische, regionale, strategische Relevanz*: Die Inhalte sind nachweislich von besonderem kulturellem, historischem oder regionalem Wert – sie tragen also zur Entwicklung und Bedeutung der HGK bei und ermöglichen Formen der Vernetzung, die der strategischen Ausrichtung der HGK dienen.<sup>6</sup>

## Bestandsübersicht

Ein Blick auf die Bestände verdeutlicht die Heterogenität der Inhalte, wobei ein Bestand als «Niederschlag einer schöpferischen, geschäftlichen oder anderweitigen Tätigkeit» verstanden wird, wie ihn Philipp Messner aus der institutionellen Praxis der Archivtheorie ableitet.<sup>7</sup> Neben der Provenienz spielt die Sacherschliessung (Pertinenzprinzip) eine wichtige Rolle.

- .....
- 2 Die an Kunsthochschulen entstehenden Arbeiten gewinnen nicht nur bei später erfolgreichen Künstler\*innen und Gestalter\*innen an Wert, sondern sind insgesamt von historischer Relevanz. (vgl. z. B. die Lehr- und Vermittlungsforschung)
  - 3 Vgl. Baumann, Laurent; Messerli, Franco; Mumenthaler, Samuel: *Memoriav Bulletin*: Zugang. Bern 2015. Online: [https://memoriav.ch/wp-content/uploads/2015/10/Bulletin-22\\_Web.pdf](https://memoriav.ch/wp-content/uploads/2015/10/Bulletin-22_Web.pdf) [Abruf 4. 10. 2019].
  - 4 Dies trägt der Forschungsrelevanz von künstlerischen resp. gestalterischen Tätigkeiten Rechnung, wie sie in den «Strategischen Leitlinien FHNW 2025» (2019, S. 12) niedergeschrieben sind.
  - 5 Ein solcher Bestand sind z. B. Thesen oder Abschlussarbeiten, die auch dann (sofern möglich) vollständig aufbewahrt werden, wenn die Institute sie aufgrund von mangelnder Qualität nicht verbreitet/publizieren möchten.
  - 6 Jahresrückblick FHNW. Lurk, Tabea (2019): Ankommen und Zusammenwachsen. In: Langkilde, Kirsten; Hochschule für Gestaltung und Kunst FHNW (Hrsg.) *Imagine* (Jahresbericht 2018). Basel: Peter Median Verlag, S. 261-270.
  - 7 Messner, Philipp: Eine andere archivische Praxis, *Brand-New-Life – Magazin für Kunstkritik*. Zürich 2017. Online: <http://brand-new-life.org/b-n-l-de/eine-andere-archivische-praxis/> [Abruf 4. 10. 2019]. .



Unter den hochschuleigenen, digitalen Sammlungen sind zu erwähnen:

- die *Videosammlung des Instituts Kunst*, die der Künstler, Kurator, Kunsttheoretiker und ehemalige Leiter des Instituts Kunst, René Pulfer, seit den 1990er Jahren bis 2014 aufgebaut hat, bis die HGK auf das Dreispitz Areal gezogen ist und die Mediathek neu organisiert wurde. Hierbei handelt es sich um ca. 2500 diskbasierte Datenträger mit vornehmlich Film- und Videoarbeiten sowie Referenzbeispielen der Videokunst. Die Sammlung wird bis heute, wenn auch mit unterschiedlichem Engagement und Strategien, weitergeführt.
- das Archiv des *Instituts Industrial Design* (2006-2015), das projektbasierte Arbeiten und Entwürfe dokumentiert und den wiederkehrend stattfindenden Modulen zuordnet. Die sowohl formal (Verknüpfung von Studienverlauf, Lehrangeboten und Studierendendaten)<sup>8</sup> als auch inhaltlich gepflegte Struktur klassifiziert die Inhalte entlang einem hierarchisch gegliederten Thesaurus und wurde aufgrund der hochwertigen Datenstruktur zum Vorbild für diverse Bestands-Modellierungen.
- Teile der *Thesenarchive der Institute Kunst* (1995/2012-2018), *Hyperwerk* (2002-2016, ohne 2007) und *Masterstudio* (2010-2019).
- seit 2017 die *DiplomHGK*- bzw. *NextGeneration*-Website, welche kurze Abstracts, Bild-, Audio-, Video- und Textmaterial aller Absolvent\*innen eines Jahrgangs der HGK präsentiert und die gleichnamige Jahresausstellung begleitet.
- seit Mitte 2018 die Inhalte der *Campus-der-Künste*-Website, die als eine zentrale Plattform der Aussenkommunikation funktioniert und wichtige Veranstaltungen, News etc. ankündigt.<sup>9</sup>
- der Teilbestand *ACT Basel* (2004-2018), des gleichnamigen, schweizweit ausgetragenen Performance-Festivals *ACT*. Aussergewöhnlich ist dieses Konvolut, weil das Festival hochschulübergreifend ausgetragen wird und die Teilnehmenden aller Hochschulen an allen Orten ihre Performances auführen können. Das Konvolute enthält über 900 Ansetzungen von über 340 Künstler\*innen und/oder Künstlergruppen mit etwas mehr als 14.000 Medien (Video-, Bild-, Ton- und Textmaterialien).
- seit 2019 fotografisches Dokumentationsmaterial, das Hans-Peter Huser in seiner Arbeit für die HGK Kommunikation zwischen 2007 und 2018 fotografiert oder in Auftrag gegeben hat und alle wichtigen Anlässe der Hochschule in diesem Zeitraum dokumentiert.

Zu den Inhalten aus Forschungsprojekten und den forschungsrelevanten Sammlungen gehören:

- Resultate des Forschungsprojekts *To Perform* (2012 ff.), das sich mit der Aufführungspraxis der Neuen Musik beschäftigte und federführend von der Hochschule für Musik FHNW realisiert wurde. Im Rahmen von fünf Fallstudien wurden die Komponisten Vinko Globokar, Dieter Schnebel, Heinz Holliger, Christoph Caskel, Peter Maxwell Davies und Péter Eötvös bei der Erarbeitung einer Aufführung (inkl. Instrumentenbau, Spielanleitungen etc.)

8 Zur Relevanz dieser Strukturverknüpfung vgl. Nippert, Klaus: Archivisches Dokumentationsprofil für Studierendendaten aus elektronischen Fachverfahren. In *Archivar* 2016: 69 (3) S. 252.

9 Das entlastet längerfristig die Erschliessungsarbeit und hilft, versehentliche Lücken zu vermeiden.

dokumentiert, um mithilfe der Notationen, aber eben auch von Videoaufzeichnungen, Interviews, Proben- und Bildmaterial spätere Wiederaufführungen zu ermöglichen. Die Videodokumentation erfolgte durch Angehörige des Instituts Kunst unter der Leitung von Reinhard Manz.<sup>10</sup>

- Die Online-Publikation eines kuratierten Datenauszugs des Forschungsprojektes *Grenzgang - Künstlerische Untersuchungen zur Wahrnehmung und Vermittlung von Raum im trinationalen Grenzgebiet* (2014-2015), das vom Schweizerischen Nationalfond gefördert wurde und am Institut Lehrberufe für Gestaltung und Kunst realisiert wurde.
- die Resultate der *Videowochen im Wenkenpark* (1984, 1986, 1988),<sup>11</sup> einem frühen Videokunstkurs, der von René Pulfer (kuratorisches Konzept) und Reinhard Manz (Produktion) in der Nähe von Basel organisiert wurde und an dem heute weltbekannte Videokünstler\*innen aus unterschiedlichen Ländern und der Schweiz teilgenommen und währenddessen Videowerke produziert haben.

Unter den in Kooperation aufgebauten, digitalen Sammlungen finden sich

- jene Video- (und Film-)arbeiten, die nach dem Ausstellungsprojekt *Σ Summe 2017* eine Berechtigung zum Verbleib in der Mediathek erhielten.
- die an der Musikakademie Basel FHNW zwischen 1974 und 2017 aufgebaute *Basler Bibliographie zur Historischen Musikpraxis*, die als ausgesprochen hochwertige Quellensammlung 2017 eingestellt wurde und die als historisches Dokument nun zur weiteren Bearbeitung z. B. via Zotero zugänglich ist.<sup>12</sup>
- Teile der *Digitalen See*, einem massgeblich in Basel verorteten digitalen Sammlungsbestand, der vornehmlich videografische Videodokumentationen zur Performancekunst, verzeichnet und aufbewahrt. Wie die Betreiberinnen Muda Mathis, Videokünstlerin und ehemalige Dozentin im Institut Kunst, Sus Zwick und Chris Regn schreiben, ist die *Digitale See* «kein Ort, sondern eine Idee und Haltung», die «zum Ziel [hat], Kunst, in diesem Fall Performance Dokumentationen, der freien ungebundenen Szene im Internet zugänglich zu machen».<sup>13</sup> Aus diesem wachsenden Bestand sind als abgeschlossene Teilbestände erschlossen: die *Performance Chronik Basel 1987–2006*,<sup>14</sup> das *Feministische\* Improvisatorium* (2018)<sup>15</sup> und die sogenannten *Radiosendungen* (2018)<sup>16</sup> des Kasko. Aktuell läuft ferner die Integration eines Grossteils der videografischen Dokumentationssammlung des Kasko - Kaskadenkonsensator: Raum für aktuelle Kunst und Performance (1994–2019).
- die Podcasts des Kulturprojekts *Digital Brainstorming* des Migros Kulturprozent, die der Journalist, Autor, Medien- und Kulturschaffende Do-

10 To Perform: [https://www.fhnw.ch/de/forschung-und-dienstleistungen/musik/hochschule-fuer-musik/projekte/to\\_perform](https://www.fhnw.ch/de/forschung-und-dienstleistungen/musik/hochschule-fuer-musik/projekte/to_perform) [Abruf 4. 10. 2019]. Die Archivierung des Projekts ist derzeit sistiert.

11 Wenkenpark: <http://mediathek.hgk.fhnw.ch/wenkenpark.php> [Abruf 4. 10. 2019].

12 Basler Bibliographie zur Historischen Musikpraxis: [https://www.zotero.org/groups/2171465/basle\\_bibliography\\_for\\_historical\\_performance\\_practice](https://www.zotero.org/groups/2171465/basle_bibliography_for_historical_performance_practice) [Abruf 4. 11. 2019].

13 Performance Chronik Basel: <https://mediathek.hgk.fhnw.ch/pcb.php> [Abruf 4. 10. 2019].

14 Ebd.

15 Feministisches Improvisatorium: [http://mediathek.hgk.fhnw.ch/feministisches\\_impro.php](http://mediathek.hgk.fhnw.ch/feministisches_impro.php) [Abruf 4. 10. 2019].

16 Radiosendungen: <http://mediathek.hgk.fhnw.ch/radiosendungen.php> [Abruf 4. 10. 2019].

minik Landwehr zwischen 1998 und 2019 realisiert hat und prominente Vertreter\*innen der aktuellen Szene der Medienkultur und -kunst zu Wort kommen lässt.

Zu den wichtigsten externen Quellen des Katalogsystems, deren Metadaten in den SOLR-Volltextindex zwar aufgenommen und dort aktualisiert, nicht aber dauerhaft gespeichert werden, gehören

- die bibliographischen (Meta-)Daten der *Swissbib* (inklusive Nationallizenzen), der Open Access Repositorien für Bücher (DOAB) und Zeitschriften (DOAJ), Springer und DeGruyter Open, die lizenzierten Zeitschriften der HGK, der FHNW und von Swissuniversities, die in der elektronischen Zeitschriftendatenbank EZB verwaltet werden, die Nationallizenzen und weitere, frei zugängliche, relevante Sammlungen z. B. der ETH Zürich.
- die Metadatensammlungen von *videokunst.ch* und von *EuropeanArtNet*, wobei letztgenanntes vornehmlich Künstler\*innen mit Referenzen auf Ausstellungen und Rezensionen benennt.<sup>17</sup>

Auch wenn die Teilbestände hier gemäss der zeitlichen Abfolge ihrer Erschliessung aufgelistet wurden, greift eine chronologische Betrachtung der Bestandsgeschichte zu kurz. Zum einen wurden in allen vier Bereichen gleichzeitig gesammelt. Zum andern wurde auf eine systematische Erschliessung entlang der dossiers-bildenden Bestandscluster, also z. B. zuerst alle Abschlussarbeiten/Thesen etc., verzichtet, um die unterschiedlichen Mentalitäten und Geschwindigkeiten der Institute zu respektieren. So konnten individuelle Übergabeformen gefunden werden. Während einige Institute derzeit noch entspannt beobachten, wie die Prozesse ablaufen, welche Aufwände effektiv beim Zusammenstellen der Daten anfallen oder wie sich die relativ einfachen, stets niederschweligen Workflows gegen die Entlastung der institutsinternen Aufbewahrung und der Regelung von Zugriffskonventionen z. B. bei der Herausgabe von Thesen rechnen, haben andere die Chancen der Zusammenarbeit mit dem Archiv erkannt und die Übergabe z. B. der Thesen direkt in die semesterbasierten Abläufe ihrer Sekretariate integriert. Wieder andere waren experimentell unterwegs und haben die Gelegenheit der strukturierten und dauerhaften Aufbewahrung genutzt, um Aspekte ihrer Archivkonzepte umzusetzen oder eigene Archivierungshaltungen zu etablieren, die in künftigen Projekten von Anfang an berücksichtigt werden. Nicht zuletzt gab es Einzelfälle, in denen produktbasiert aufgrund des technischen Wandels (System- und Software-Obsoleszenz) die Nähe zum Archiv gesucht wurde.<sup>18</sup>

### Aus- und Weiterentwicklung

Da längst nicht alle archivrelevanten, hochschuleigenen und strategisch spannenden Konvolute erschlossen sind, wächst das digitale Archiv in unter-

.....  
17 Derzeit wird geprüft, ob die Integration der Metadaten von Europeana Collections auch möglich ist, da diese Inhalte die Suchbedürfnisse unserer Nutzer\*innen sehr gut befriedigen und den Reflexionsradius erweitern könnten.

18 Eine klassische Notsicherung besteht in der Emulation oder Virtualisierung einer Out-of-Service gehenden Datenbankinfrastruktur, wie sie z.B. das Institut Industrial Design seit (2017) auf der Basis einer Filemaker-Applikation nutzte. Parallel zur mittelfristigen Verstetigung in der virtuellen Maschine (VM) wurden die Daten auch ins Katalogsystem der Mediathek und das digitale Archiv der HGK testweise migriert.

schiedliche Richtungen. Die kontinuierliche (Weiter-)Entwicklung folgt einem modularen Ansatz, der sich sowohl dynamisch an neue Softwarearchitekturen, Werkzeuge, Prozesse oder aktuell(st)e technologische Entwicklungen anforderungsbasiert anpasst, als auch die wandelbaren hochschulpolitischen oder strategischen Zielsetzungen umsetzen kann, ohne dabei die Dauerhaftigkeit der Inhalte zu gefährden. Diese grundlegende Flexibilität erscheint aufgrund der aktuellen Rahmenbedingungen unabdingbar und durchzieht sowohl die administrativ-strategische als auch die technologische Ausrichtung des Archivs.

Auf die Bestandsebene heruntergebrochen, bedeutet das beispielsweise, dass die Frage, wie repräsentativ die ggf. neue Struktur einer Sammlung für die bestehenden oder künftige Konvolute ist, darüber entscheidet, ob die zu übernehmenden Daten auf die bestehenden Datenmodelle gemappt oder eine neue Datenstruktur integriert wird. Gleiches gilt für die Bearbeitungstools: so wurde die Erkennungskaskade des Werkzeugs *Indexer* (s. u.) Ende 2018 um die Softwarebibliothek Siegfried<sup>19</sup> erweitert, weil erstmals grössere Mengen an Blu-Ray-Videodaten ins Archiv gelangten, die von den bestehenden Werkzeugen nicht korrekt erkannt wurden. 2019 wurde die *metadata-extractor*-Bibliothek eingebunden,<sup>20</sup> um Metadatenformate wie z. B. Exif, IPTC, XMP, JFIF und andere aus digitalen Fotografien extrahieren zu können.<sup>21</sup>

## Anforderungen und Zielsetzung

Eine der wichtigsten Ressourcen des digitalen Archivs der HGK sind die Datenlieferant\*innen. Sie haben die Inhalte über einen längeren Zeitraum produziert oder kumulierend zusammengetragen und kennen die Hintergründe, Rechtlage und Kontexte. Sie haben in der Regel individuelle Ordnungen oder Organisationsformen entwickelt, von denen aus inhaltlichen (Authentizität, Wahrung von Entstehungs- und Überlieferungskontexten) aber auch arbeitsökonomischen Gründen so viel wie möglich ins digitale Archiv der HGK übernommen werden soll.

Das Anforderungsportfolio sollte ferner folgende Wünsche berücksichtigen:

- *Ein System / eine Software für unterschiedliche Datentypen:* Trotz der fachlichen, thematischen und medialen Spannbreite sowie der variierenden Erschliessungstiefe sollten Katalogisierung und Aufbereitung über ein einziges Werkzeug (Frontend) funktionieren und sich quasi out-of-the-box den verschiedenen Bedürfnissen und Denktraditionen anpassen. Damit können ferner die Datenlieferant\*innen ggf. direkt in die Erschliessung oder Erfassungskontrolle eingebunden werden.
- *System- und Ortsunabhängigkeit:* Das Erschliessungssystem sollte systemunabhängig (Mac, PC, Linux) einsetzbar sein und das Editieren der Inhalte auch dann erlauben, wenn die Datenlieferant\*innen nicht (mehr) der Hochschule affiliert oder vor Ort sind.

19 <https://www.itforarchivists.com/siegfried/> [Abruf 4. 10. 2019].

20 <https://github.com/drewnoakes/metadata-extractor> [Abruf 4. 10. 2019].

21 Speziell hat uns die firmenspezifische MakerNote (=Kamera-ID) interessiert, welche Fotografien häufig eindeutig identifiziert und bei der Übernahme der Dokumentationsfotos der HGK Kommunikation anfiel.

- *Dauerhafter Zugriff:* Neben der Option des verteilten Arbeitens sollten die Datenlieferant\*innen ggf. auch nach der Übergabe die Möglichkeit haben, auf die strukturiert aufbereiteten Inhalte ihrer Sammlungen – jenseits des Recherche/Wiedergabe Interfaces im Internet – zuzugreifen.
- *Open Access:* Potentiell sollte die Möglichkeit zur Nachnutzung und Erweiterung der Archivkonvolute eingeräumt werden und zwar auf der Ebene der (Meta-)Daten. Wenn die Dateneigner\*innen dies wünschen, wie beispielsweise die Basler Bibliographie (s. o.). Gerade mit Blick auf die Archivierung von Forschungsdaten ist dies ein durchaus relevanter Aspekt.

Neben der niederschweligen Datenübergabe und (anhaltenden) Datenpflege sprachen eine ganze Reihe an Gründen für die Nutzung von Zotero.<sup>22</sup> Die open source Software wurde von der Corporation for Digital Scholarship und dem Roy Rosenzweig Center for History and New Media für die Verwaltung von (wissenschaftlicher) Literatur entwickelt und unterstützt mittlerweile 34 unterschiedliche Medientypen, darunter die Standard Archivformate. In der Mediathek spielt Zotero als Zwischensystem in der Archivvorstufe, bei der Aggregation und bei der fließenden Erschließung eine zentrale Rolle.

Zu den übergeordneten Zielen gehörten:

- *Rasche Sichtbarkeit der Resultate:* Da unsere Datenlieferant\*innen in der Regel noch leben, möchten sie ihre Inhalte möglichst schnell online sehen. Es braucht also einfache Workflows, welche die im Zotero verwalteten Inhalte leicht ins Rechercheinterface des Integrierten Katalogs ausspielen.
- *Offene Schnittstellen:* Damit die hauseigene Software in den Vereinnahmungs- und Wiedergabeprozessen auf die Daten in Zwischensystemen zugreifen kann, sind offene, dokumentierte Schnittstellen gefordert.
- *Statistische Auswertung:* Optional sollen statistische Auswertungsmechanismen bedient werden können, wobei hier das technologisch Machbare von der datenschutztechnischen Legalität reguliert wird.

## Workflows

Aus dem Anforderungsprofil haben sich ferner die Workflows ergeben, deren Ausgestaltung an die Quelle angepasst und bei Bedarf optimiert wird. Bestimmend sind hierbei Aspekte wie:

- *Materiallage:* Welche Daten liegen in welcher Form vor, woher stammen sie und wie sollen sie später zugänglich sein?
- *Metadatenlage:* Welche Beschreibungsdaten oder Kontextinformationen sind wie greifbar?
- *Anvisierte Erschließungsprozesse und -tiefe:* Wer soll die Inhalte erschliessen und welche Vorkenntnisse liegen vor?
- *Wiedergabezwecke/Nutzungsform:* Für welche Zielgruppe sollen die Inhalte erschlossen werden und für wen sollen sie später zugänglich sein?

Mit Blick auf die Verarbeitungsworkflows lassen sich die Schritte der Archivvorstufe, der Erschließung, der Medienverarbeitung/Zusammenführung,

.....  
 22 Corporation for Digital Scholarship; Roy Rosenzweig Center for History and New Media: Zotero Documentation [de:start]. Massachusetts 2019. Online: <https://www.zotero.org/support/de/start> [Abruf 4. 10. 2019].

der Wiedergabe sowie schliesslich der digitalen Archivierung unterschieden (vgl. Abbildung 1, unterste Zeile, Seite 114).

### Archivvorstufe

Auf die Anfrage zur Übergabe oder Übernahme von Beständen folgt die *Sichtung* des Bestands.<sup>23</sup> Dabei werden neben rechtlichen Aspekten die Prozesse und Erschliessungsworkflows definiert (vgl. Abbildung 1). Bei der anschliessenden *Datenlieferung* werden die Daten – zumeist mittels Memorysticks, Hard-Disks oder in Einzelfällen gerippt – auf die FHNW-Infrastruktur überführt (kopiert) oder über die Serverlaufwerke direkt freigegeben.<sup>24</sup>

Es folgt die inhaltliche *Erschliessung*, die fast immer *Zotero* nutzt. Daher werden die Quelldaten häufig in einem mehrstufigen Verfahren für dieses System vorbereitet. Unterschieden wird dann zwischen der Aufbereitung von Daten für den *Zotero*-Workflow und den sog. *Indexer*-Workflow<sup>25</sup>, der ebenfalls ins *Zotero* mündet, aber keine vorherigen, inhaltlichen Aufbereitungsleistungen erfordert.

Im Unterschied zu vollautomatisierten Ingestprozessen wird die Vollständigkeit der übergebenen Daten ebenso wie die Metadatenqualität händisch geprüft. Weil zumeist diverse kleinere Anpassungen durchgeführt werden müssen, sprechen wir von *Rohdaten*.

### Aufbereitung von (Roh-)Daten

Je nach Quelle werden die Metadaten teilweise oder vollständig durch die Mediathek erstellt. Sofern Metadaten vorhanden sind, korrigiert/harmonisiert die Mediathek diese.

Der Zugang zu den jeweiligen Daten und die Einbettung der Datenbasis variiert zwischen:

- *Files im Dateisystem*: Die einzelnen Dateien liegen in strukturierten Foldern vor, sodass die Kontexte, die Erschliessungssystematik, häufig Ortsangaben etc. ersichtlich sind. Gerade bei Daten, die wiederkehrende Metadateninformationen haben, wie z. B. bei Abschlussarbeiten der Dokumenttyp, Sammlungszugehörigkeit (Institut/Fachbereich, Jahrgang), bietet sich ein Zwischenschritt via Excel an, in dem die Datensätze verzeichnet werden (vgl. Abbildung 2 und 4, Seite 115, 117).<sup>26</sup>
- *Daten plus Tabelle*: Liegen Metadaten-Informationen bereits in Tabellenform (z. B. Excel, CSV) vor, muss lediglich die Konsistenz der Einträge überprüft,

23 Während bei der Übergabe die Datenlieferanten das Angebot offerieren, fragt die Mediathek bei der Übernahme die Dateninhaber aktiv an.

24 Die Speicherumgebung der FHNW besteht derzeit aus einem NetApp-Speicher. Die Datensicherung erfolgt als Block-Backup von Snapshots (tägliches inkrementelles Backup (5/7) und monatliches Full-Backup). Das Backup wird hybrid auf Tape und Disk gesichert

25 Die Kernfunktionalität des Indexers ist die Erstellung eines umfassenden Indexes, also «eines Hilfsmittels zur Auffindung von Sachbetreffen, Personenbetreffen oder Toponymen» Westphal, Sina: Indizieren. In: Terminologie der Archivwissenschaft, Archivschule Marbach. 2019. Online: <http://www.archivschule.de/uploads/Forschung/ArchivwissenschaftlicheTerminologie/Terminologie.html> [Abruf 4. 10. 2019]. Zur Funktionalität des Indexers vgl. Enge, Jürgen; Kramski, Heinz Werner: Exploring Friedrich Kittler's Digital Legacy on Different Levels: Tools to Equip the Future Archivist. In IRPES 2016, Bern 2016, S. 229–236.

26 In vielen Datenbank-Systemen werden derartige Inhalte über (Satz-)Schablonen erfasst.



ggf. angepasst und das Endresultat so auf die Ziel-Schemata von Zotero gemappt werden, dass der Import möglich wird.

- *Datenbank:* Beim Export von Metadaten aus Datenbanken besteht die Herausforderung darin, alle Inhalte bzw. Tabellen und die ursprünglichen Verknüpfungen sauber abzubilden. Bisher konnten alle Inhalte direkt abgebildet werden; sollten dies nicht mehr möglich sein, weil die Einträge nur als Ergebnisse (z. B. einer komplexen Suche aus einem Datenbanksystem) vorliegen, wird vermutlich das *Performance*-Prinzip angewandt, das Harvey und Weatherburn mit Blick auf prozessbezogene Unterlagen als weit verbreitet und anerkannt beschrieben haben.<sup>27</sup> Das anschließende Mapping erfolgt analog zum Vorgehen bei Tabellen.
- *(inkludierte) Objektmetadaten:* Die in Einzelmedien vorhandenen, semantischen und technischen Metadaten können z. T. automatisiert erkannt und extrahiert werden. Bisher findet neben den im *Indexer*-Workflow hinterlegten Routinen primär die Extraktion fotografischer Bildinformationen mittels *metadata-extractor* Anwendung.

### *Indexer-Workflow*

Wenn viele Dubletten, unterschiedliche auch historisch variierende Medienformate und eine erhöhte Dauerhaftigkeit erwartet werden, nutzen wir den *Indexer*-Workflow. Der *Indexer* extrahiert jene Objektmetadaten, die im OAIS-Modell als Repräsentationsmetadaten bezeichnet werden.<sup>28</sup> Dabei durchlaufen die Daten nach dem Kopieren auf die hauseigenen Systeme folgende Schritte:

- *Indexierung:* Die vereinnahmten Daten werden automatisiert indexiert. Zunächst werden alle Dateien in einer Index-Datenbank erfasst, die auch den Dateinamen und Speicherort dokumentiert.<sup>29</sup> Die Erkennungskaskade erstellt dann mithilfe diverser Softwarebibliotheken technische Metadaten (Fileinformationen), generiert Volltexte sowie bei bekannten Bild- und Video-Formaten Vorschaubilder und sie erkennt Dubletten (auf der Basis von Checksummen). Dabei finden u. A. folgende Softwarebibliotheken Anwendung: Apache Tika; ffmpeg; gvfs-info; imageMagick; Libmagic; OpenLayers; Siegfried; Sonogram (selbstentwickeltes C++ Programm). Auch wenn die Daten damit noch lang nicht semantisch erschlossen sind, werden wichtige Informationen sichtbar und können für die weitere Bearbeitung genutzt werden.
- *Inventarisierung:* Mithilfe von selbst-generierten Inventarnummern werden die *relevanten* Inhalte aus dem Datenpool des *Indexers* gefischt und die Informationsobjekte geschnürt.<sup>30</sup> Dabei wird zwischen Einzel-Inventarnummern (ein Objekt = eine Inventarnummer) und Gruppen (n-Objekte bekommen als Bandel je die gleiche Inventarnummer; die Files werden hochge-

27 Harvey, Ross; Weatherburn, Jaye: Preserving Digital Materials. Rowman & Littlefield 2018, S. 87.

28 Vgl. nestor-Arbeitsgruppe OAIS-Übersetzung / Terminologie: Referenzmodell für ein Offenes Archiv-Informations-System – Deutsche Übersetzung, Version 2, Frankfurt am Main 2013, S. 60, urn:nbn:de:0008-2013082706. <https://d-nb.info/104761314X/34> [Abruf 4. 10. 2019].

29 Dateinamen und Speicherorte geben wichtige Hinweise auf den Kontext und die ursprüngliche Ablagesystematik, die häufig für die Dossierbildung hilfreich ist.

30 Ein Informationsobjekt kann mehrere Dateien (Datenobjekte) enthalten und wird durch Repräsentationsinformationen erweitert. Vgl. nestor 2013, S. 19.

zählt) unterschieden (vgl. Abbildung 3, Seite 116). Pro Inventarnummer wird ein Handle vergeben.<sup>31</sup>

### Zotero-Import

Nach der Aufbereitung der (*Roh*-)Daten werden die Metadaten ins Zotero übertragen oder erscheinen dort, im Falle des *Indexer*-Workflows, vollautomatisch. Hierbei werden zwei Arten des Übertrags unterschieden:

- (*Tabellen*-)Import via *DIGMA*: Nach einer formalen Prüfung der Feldbezeichnungen werden die zuvor erstellten Tabellen via Skript eingelesen und so ins Zotero übertragen, dass die Datensätze im sog. Importbereich (= *Zotero*-Gruppe) als Einzeleintrag erscheinen.<sup>32</sup> Nun sind die in der Tabelle erschlossenen Inhalte ausgefüllt.
  - *Sortierung der Zotero-Einträge*: Es folgt der Übertrag vom Importbereich in die Zielsammlungen, die gemäss Aktenplan strukturiert sind. So können ganze Datencluster kopiert werden.<sup>33</sup> Die Dossierposition im Aktenplan wird in der *Zotero*-Struktur abgebildet, wobei die Sammlung der *Zotero*-Gruppe die jeweiligen Dossiers benennt. Die Provenienz der Datensammlung ist in einem separaten Dokument vermerkt.
  - *Erweiterung der Zotero-Einträge*: Sodann werden die Medien, die Zwischenzeitlich vereinnahmt wurden (s. u.) angehängt, indem die Links als Anhang in den Datensatz eingetragen werden. Eine Liste wird vom *DIGMA* bereitgestellt.

*Metadatenquelle Indexer*: Im Falle des *Indexer*-Workflows werden die *Zotero*-Einträge automatisch in dem Moment erzeugt, wo der Datensatz die Inventarnummer erhält. Die jeweiligen Medien/Ressourcen hängen als Link (Anhang) bereits an den Datensätzen und sind nummerierten. Zudem sind die Einzelfiles mit einem Auszug der technischen Metadaten aus der Erkennungskaskade des *Indexers* ausgestattet (vgl. Abbildung 4, Seite 117). Semantischen Informationen wie Autor\*in, Titel, Datum etc. müssen hingegen händisch eingetragen werden.

- *Sortierung der Zotero-Einträge*: Das Verschieben der Einträge in Unterordner erfolgt ebenfalls händisch, zumal die Verzeichnisstruktur der *Zotero*-Subfolder durch die erschliessende Person erstellt wird.

### Erschliessung

Als Referenz für die Erschliessung dient der *Dublin Core*,<sup>34</sup> auch wenn die Erschliessungstiefe und -qualität variiert. Zu den Kernmetadaten gehören neben Informationen zu Sammlungskonzentrat und dem jeweiligen Erschliessungskonzentrat Angaben zu Autor\*in, Titel, Jahr, wenn möglich Ort sowie weitere Metadaten wie etwa Schlagwörter.

31 <https://www.handle.net/> [Abruf 4. 10. 2019].

32 Die Metadaten verbleiben in einer MySQL-Datenbank, die als Gelenk fungiert.

33 Nachdem alle Datensätze in Zielsammlungen kopiert wurden, werden die Inhalte des Importbereichs gelöscht.

34 DCMI – Dublin Core Metadata Initiative: DCMI Metadata Terms. <https://www.dublincore.org/specifications/dublin-core/dcmi-terms/> [Abruf 4. 11. 2019].

Zugriffsrechte werden in der Regel abhängig von der jeweiligen Sammlung (Dossier) geregelt.<sup>35</sup> So definieren z. B. die Institute als originäre Sammlungseigner den Vertraulichkeitsstatus einer Dokumentengruppe/Dossiers. Die Entscheidung über die Zugänglichkeit wird i. d. R. durch die Institutsleitung oder ihr Sekretariat an die Mediathek kommuniziert, die diese in Absprache mit dem DIGMA umsetzt. Zur Auswahl stehen hierbei die Stufen:

- Freie Nutzung im Internet
- Nutzung innerhalb des Schweizer Hochschulnetzes (Shibboleth-Zugriff)
- Nutzung innerhalb der FHNW
- Nutzung für HGK Angehörige
- Nutzung für Institutsangehörige
- Nutzung für einzelne Jahrgänge eines Instituts plus z. B. Leitung
- Zugriff durch selektierte Einzelpersonen.

Längerfristig besteht das Ziel darin, alle semantischen Beschreibungsdaten frei im Internet zugänglich zu machen, auch wenn die eigentlichen (Meiden-)Daten, also PDFs, Video-, Audio-, Bilddaten nur definierten Nutzer\*innengruppen zugänglich sind oder (zeitliche) Embargos gelten. Von einigen Einträgen und Beständen dürfen allerdings auch die Metadaten nur im internen Bereich sichtbar sein.

### *Medienverarbeitung / Zusammenführung*

Nach der Erschliessung spaltet sich die Bearbeitung von Meta- und Medien-daten auf, sodass zwischen dem Import der semantischen Zotero-Metadaten in den SOLR-Index (Katalog-Workflow), den technischen Workflows des Media-Servers (Transkodierungs-Workflows) und den Archivierungs-Workflows unterschieden wird.

### *Katalog-Workflow*

Um die Recherche und eine einheitliche Wiedergabe über die unterschiedlichen Bestände und Sammlungen zu ermöglichen, werden die technischen und semantischen Metadaten (inkl. Handle) im *Integrierten Katalog* auf sogenannten Detailseiten sammlungsspezifisch ausgegeben (vgl. Abbildung 5 und 6, Seite 118–119).<sup>36</sup> Hierzu werden die Metadaten zunächst in den SOLR-Volltextindex importiert. Es folgt ein Mapping zwischen den spezifischen Metadaten der neu importierten Sammlung und den bereits existierende Wiedergabefolien, so dass ein wiedererkennbares Erscheinungsbild auch dann entsteht, wenn sehr heterogene Quellen vorliegen.

Zumeist erfolgen nach der ersten Bereitstellung der Resultate im Webinterface weitere Anpassungszyklen, in denen Fehler oder Sonderwünsche auf allen Ebenen behoben respektive implementiert werden. Zudem können Inhalte weiterhin erschlossen oder die Granularität der Sichtungcluster modifiziert werden.<sup>37</sup>

35 Neu werden die Nutzungsmetadaten (Copyright-Holder, Lizenz, Zugangsform) der Medien zudem direkt ins File geschrieben.

36 Weitere Ansichten finden sich in den Folien des zugehörigen Vortrags der AUDS 2019 online unter: <https://irf.fhnw.ch/handle/11654/27477> [Abruf 4. 10. 2019].

37 Bisher wurden Modifikationen lediglich in der gleichen Erschliessungsperiode eingebracht. Für Veränderungen, die zeitlich oder semantisch versetzt sind und damit als spätere/an-

## Transkodierungs-Workflow

Beim Einlesen der Mediendateien in den Mediaserver werden die Dateinamen in die Datenbank des Mediaservers eingetragen und eine Kopie bzw. ein Verweis auf die Mediendatei erzeugt. Zudem werden die technischen Metadaten mit FFmpeg\* extrahiert, sodass die Länge eines Videos, das Seitenverhältnis, die Auflösung, der Farbraum, der Container und der Codec etc. bekannt sind und als technische Metadaten in der Datenbank gespeichert werden können. Es folgt die Transkodierung, bei der web- und archivtaugliche Zielformate erzeugt werden:

- *Bilder* werden im Mediaserver in beliebige Grössen in alle gängigen Formate ausgeliefert und intern transkodiert. Sehr grosse Bilder können über einen IIIF-Server ausgeliefert (*Cantaloupe*) werden.<sup>38</sup>
- *Videodaten* werden aktuell in Originalauflösung beibehalten und in ein webtaugliches H.264 Video im MPEG4 Container transkodiert,<sup>39</sup> aus dem 25 äquidistante Screenshots im PNG-Format erzeugt werden.
- Von *Audiodaten* werden Sonogramme im SVG-Format erzeugt und ein webtauglicher MPEG4 Container mit AAC Audio (384 kbit, 44100 Hz) wird erzeugt. Ferner wird eine im OPUS-Codec kodierte Audiodatei mit 384 kbit, variabler Bitrate und Kompressionslevel 5 im WebM Container vorgehalten.<sup>40</sup>

Alle Streaming-Dateien werden als Derivat (d. h. mit einem Verweis auf das Eltern-Element der Quelle) vom Mediaserver erneut vereinnahmt, um dann Wiedergabepakete (DIP)<sup>41</sup> je nach Zugriffsart in Echtzeit in den jeweils gewünschten Auflösungen, Grössen und Medienformaten ausliefern zu können.

- PDFs, Audio- und Video-Daten werden (analog zu den grossen Plattformen Youtube, Vimeo etc.) via iFrame mit Player ausgeliefert.
- Alle weiteren Formate werden je nach Projektverlauf abhängig vom Inhalt und der Frage der Zugänglichkeit bearbeitet. Betroffen sind hier häufig raw- oder Produktionsformate (z. B. PSD, INDD, ProRes....) aber auch Spezialformate wie etwa GPX (z. B. beim *Grenzgang*-Projekt), 3D-Daten oder obsoletere Formate.

.....  
dersartige Zutat wie in einer weiteren Metadaten-schale kenntlich gemacht werden sollten, sind wir am Erarbeiten entsprechender Werkzeuge. Denkbar wäre eine Zusammenarbeit mit der UB-Freiburg, welche mit CiTAR bereits Konzepte zur stabilen Referenzierung emulierter Daten erarbeitet hat, vgl. Rechert, Klaus: Preserving Virtual Research Environments – Introducing CiTAR (Part 1) – Open Preservation Foundation, 2018. Online available <https://openpreservation.org/blog/2018/12/12/preserving-virtual-research-environments-introducing-citar-part-1/> [Abruf 4. 10. 2019].

38 Cantaloupe Image Server: <https://cantaloupe-project.github.io/> [Abruf 4. 10. 2019].

39 Videoencoding: Container: MP4 (isom/512), Metadata at start of file (faststart); Video: Encoder: x264; Codec: H264 / Profile: High / Level: 3.1; Preset: Slow; Scene detection: no; Decoder Buffer size: 2200k; GOP Size: 100; Framerate: 25fps; Colorspace: yuv420p; Bitrate: ca. 1100kb/s; Optimierungen: Deblocking (-1:-1), psychovisuelle (1.0:0.15).

40 Audio: Encoder: fdk\_aac; Codec: aac; Kanäle: Stereo; Abtastrate: 48000Hz; Bitrate: ca. 196kb/s.

41 Gemäss OAIS, vgl. nestor-Arbeitsgruppe OAIS-Übersetzung / Terminologie, und Deutsche Nationalbibliothek. 2013. „Referenzmodell für ein Offenes Archiv-Informationssystem – Deutsche Übersetzung“, 2013, Abschn. Version 2.

## Wiedergabe

Je nach Bestand wird eine sammlungsspezifische Startseite generiert, welche die Inhalte thematisch beschreibt. Ferner zeigt die Seite zumeist eine für den Bestand relevante Abbildung, listet sogenannte Einstiegslinks auf spezifische Suchergebnisse oder Detailseiten der Sammlung (z. B. die jeweiligen Künstler\*innen) auf und enthält häufig einen Suchschlitz für die gefilterte Suche.

## Archivierungs-Workflow

Im Mediaserver können die Medien (Originale) nicht nur transkodiert werden, sondern die Quelldaten können mit den transkodierten Medien sowie den Metadaten zu Archival Information Packages (AIPs) gebaut werden. Hierzu werden die Originaldateien, die Derivate sowie sämtliche technische und semantische Metadaten (als XML-File) in einen BagIT Archivcontainer geschrieben. Es ist vorgesehen, dass die technischen und die semantischen Metadaten ab 2020 in einem METS\*-Container dauerhaft aufbewahrt werden.<sup>42</sup>

Hinzu kommt für diese AIPs ein regelmässiges Reporting. Als Sonderwerkzeug des Archivierungs-Workflows wurde daher *TBBS – The Boring BagIT Store* entwickelt,<sup>43</sup> das AIPs oder Gruppen von AIPs an verteilten Orten speichern kann. Aufgrund der potentiell variierenden Sicherheit dieser Speicherorte können die Daten vor dem Transfer verschlüsselt werden. Auch können die Speicherorte bezüglich ihres Sicherheitslevels klassifiziert werden, sodass Inhalte, welche beispielsweise an weniger vertrauenswürdigen Speicherorten aufbewahrt werden, häufiger abgelegt werden. Das ermöglicht eine modulare Kosten-Sicherheits-Politik. Mit Blick auf die Datensicherheit wird zwischen der Sicherheit in der Datenspeicherung und der Sicherheit bezüglich des Datenzugriffs unterschieden. Alle Speicherschritte sowie die Informationen zu den gespeicherten Inhalten werden in umfassenden, automatisch generierten Reports dokumentiert.

## Fazit – oder: was bleibt

Niederschwelliges, verteiltes Arbeiten, Nutzerzentrierung und Kosteneffizienz sind grundlegende Aspekte, die den System- und Software-Entscheidungen des digitalen Archivs der HGK zugrunde liegen. Die ausgesprochen schlanke Erschliessungslösung mit *Zotero* wird durch das DIGMA und seine Programmierleistungen ebenso möglich, wie die Einrichtung spezifischer Workflows oder die fließende Anpassung an sich wandelnde Kontexte. So können alle Softwarebibliotheken und Systeme fallspezifisch eingerichtet oder ausgebaut werden. Nachhaltig wird das digitale Archiv durch die Speichersysteme, die Schlichtheit und Überprüfbarkeit der Ablage sowie das *TBBS*-System mit

42 Der automatisierte METS-Workflow soll voraussichtlich ab Frühjahr 2020 zur Verfügung stehen. Sofern dies nicht möglich sein sollte, werden die semantischen Metadaten als XML-File exportiert, sodass eine Datenbank-unabhängige Zusammenführung von technischen, semantischen Metadaten und Medien auch im Havariefall oder im Falle eines Systemtransfers möglich sein wird.

43 Enge, Jürgen; Lurk, Tabea: *TBBS – THE BORING BAGIT STORE: Storing digital art and cultural assets in a distributed storage infrastructure*, in Bearbeitung.

seinen Integritätsprüfungen, die explizit für den Archivierungszusammenhang entwickelt wurden.

Die Kulturwissenschaftlerin Simone Osthoff hat vor nicht allzu langer Zeit mit Blick auf aktuelle Veränderungen in Archiven der Gegenwartskunst beobachtet: «when the past refuses to settle and technology designed to upgrade and perpetuate itself further keeps the archive alive; it seemed to me that if not alive, the archive in contemporary art is in a state of flux, undead, ghostly».<sup>44</sup> So bedrohlich das *Im-Fluss-Bleiben* im klassischen Archivkontext erst einmal klingen mag, weil ja Stabilität und Dauerhaftigkeit installiert werden sollen, scheint hier doch ein wahrer Kern enthalten zu sein, der sich im Hinblick auf unsere Arbeit durchaus konstruktiv lesen lässt: So versuchen wir tatsächlich, dem *Im-Fluss-Bleiben* (flux) der digitalen Daten und Entwicklungen mit einer flexiblen Archivstruktur und einem für Inspirationen offenen Denken zu begegnen. Unser Ziel ist es dabei, Nachhaltigkeit zu schaffen ohne die vorhandenen und für Gestaltung und Kunst so wichtigen und einzigartigen Dynamik allzu massiv einzugrenzen. Untot (undead) oder lebendig zu halten, bleibt gerade mit Blick z. B. auf die Formatvielfalt eine stetige Herausforderung, die wir im vorliegenden Text ebenso ausgeklammert haben, wie das Vermeiden von verwaisten Daten, die quasi *geisterhaft* (ghostly) und quellenlos durchs Archiv spuken und in fast jedem, zumindest in natürlich gewachsenen, digitalen Beständen vorhanden sind. Beide Themen sind uns bewusst, was sich unter anderem im Kernmetadatensatz von Neuzugängen zeigt, der nur noch Rechte-, Lizenz, ggf. Embargo und Copyright-Angaben erhalten muss.

Das Preservation Management wird im vorgestellten System vornehmlich durch den Mediaserver ausgeführt, der Dank *TBBS* dokumentierte Speicher- und Kontrollroutinen ermöglicht. Derzeit wird zudem die gesamte Systemarchitektur inkl. *Indexer*-Routinen auf *Micoservices* umgestellt, wodurch ein zukunftsfähiges, einfach skalierbares und in den technologischen Teilkomponenten wandelbares System entsteht.

Dass Sammlungsbestände auch oder gerade nach der Archivierung potentiell im wissenschaftlichen Kontext nachgenutzt, erweitert oder (in Derivaten) modifiziert werden können, zeigt die Nähe zu Forschungsdatenzyklen, deren Natur im kontinuierlichen Weiterschreiben und in der Ermöglichung des Informationskontinuums liegt.<sup>45</sup> Hier greift ferner nach wie vor das von Andrew Treloar et al. entworfene Konzept der *Curation Boudaries*, das den Statuswechsel von Dokumenten beim Schwellenübertritt (Publikation) beschreibt, also in dem Moment, wenn z. B. neue Nutzer\*innengruppen Zugriff erhalten.<sup>46</sup> Dabei kommt es zu einer Festschreibung der bestehenden Metadaten, sodass Änderungen und Erweiterungen möglich werden, ohne dass der Quellenstatus gefährdet wird.<sup>47</sup>

44 Ludovico, Alessandro: Interview [with] Simone Osthoff, in: *Neural* No. 58 (2017: hacktivism), S. 7. Das Gespräch bezieht sich auf Ostoffs ebenfalls 2017 erschienenes Buch *Performing the Archive. The Transformation of the Archive in Contemporary Art* der Interviewten.

45 Digital Curation Centre: DCC Curation Lifecycle Model. 2019. <http://www.dcc.ac.uk/resources/curation-lifecycle-model> [Abruf 4. 10. 2019].

46 Treloar, Andrew, David Groenewegen, und Cathrine Harboe-Ree: *The Data Curation Continuum: Managing Data Objects in Institutional Repositories*. *D-Lib Magazine* 2007. 13 (9/10). <https://doi.org/10.1045/september2007-treloar> [Abruf 4. 10. 2019].

47 Vgl. Treloar, Andrew: *Presentations: Diagrams*. 2017. Online available <https://andrew.treloar.net/research/diagrams/> [Abruf 4. 10. 2019].



Das Archiv der HGK befindet sich daher in einer immerwährenden Weiterentwicklung. Es ist kritisch in seiner Selbstreflexion und beabsichtigt, sich durch den Austausch mit der Fachcommunity stets zu verbessern. Die Grundhaltung orientiert sich ein Stück weit an Konzepte des *verteilten Wissens* oder des *Networks of Care*, wie es im Kontext der Konservierung, Restaurierung und Archivierung von digitalen Kunstwerken seit geraumer Zeit diskutiert wird.<sup>48</sup>

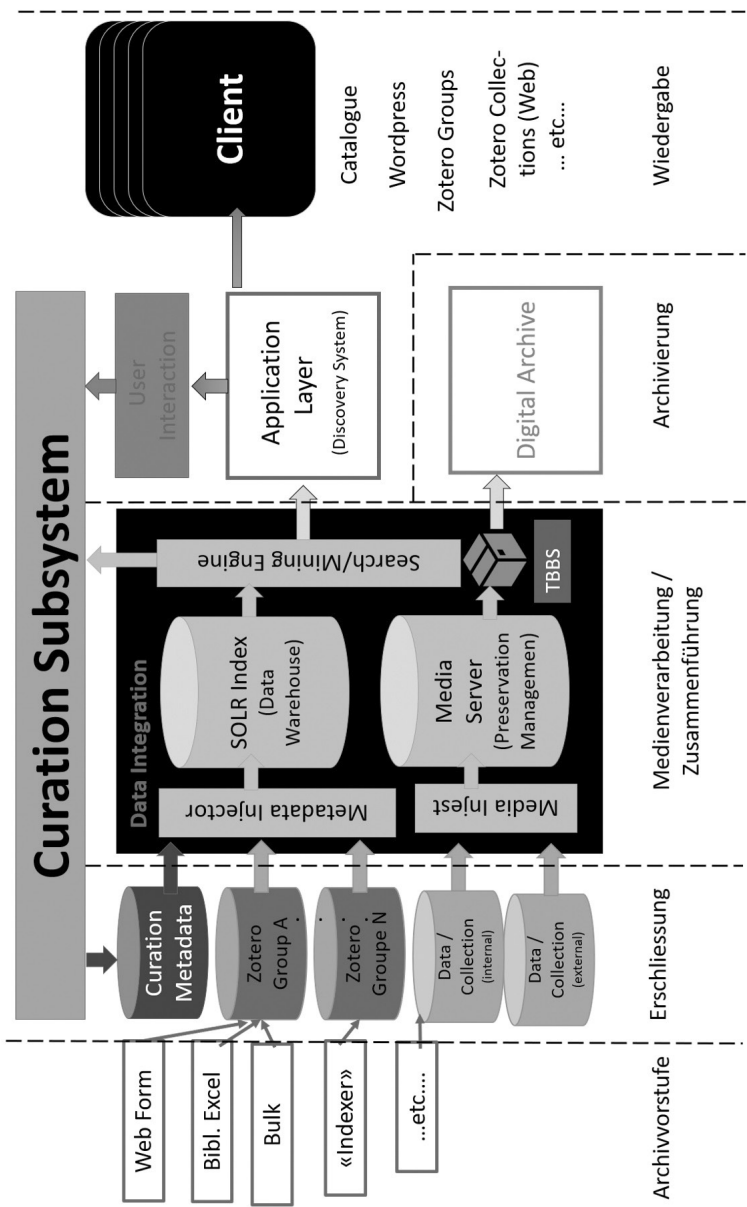
## Summary

The article presents the most important indexing workflows and tools of the digital archive of the Academy of Art and Design HGK FHNW Basel. Since the data are very heterogeneous in terms of media technology, structure, source, and quality (in depth of indexing), the collection and indexing processes adapt to the respective collections and contexts rather than vice versa. The renunciation of strict archive specifications or a singular filter requires a broad range of different tools, because despite the preservation of semantic, technical or cultural-aesthetic conventions and preferences, search and access should later be able to take place systematically and comprehensibly.

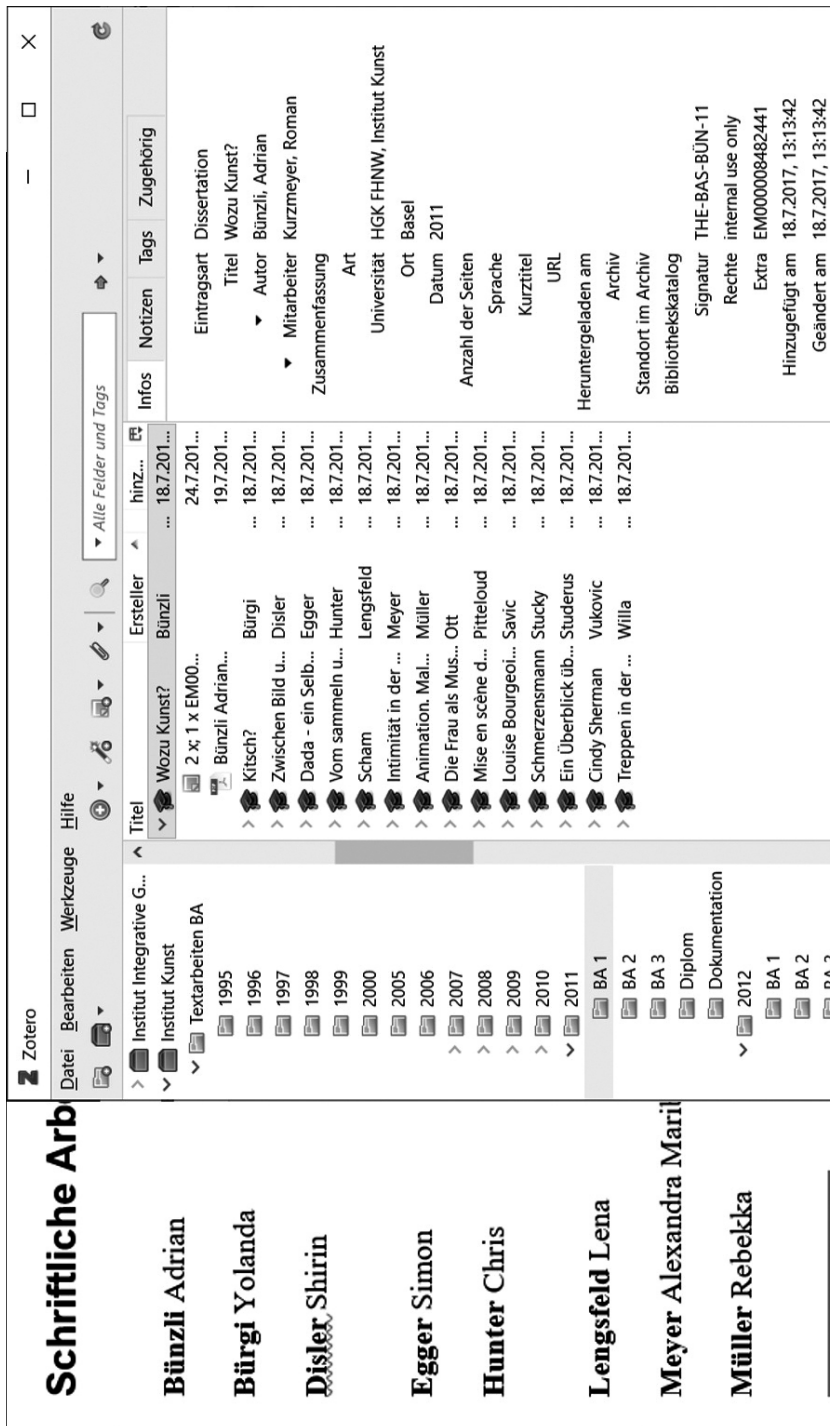
The text begins with a brief overview of the archival collections, which makes the reciprocal requirements of the data suppliers and the archives transparent. The process design and the sequence of the indexing routines can be derived from this relationship. Since all workflows are linked to digital tools, these are then presented. The outlook reflects on aspects that might be transferred to other respective archives.

48 Vgl. hierzu Dekker, Annet: Enabling the Future, or How to Survive FOREVER. In: A Companion to Digital Art (Wiley Blackwell Companions to Art History, 9), hrsg. von Christiane Paul, Chichester: Wiley Blackwell, 2016, S. 553–574.

Abbildungen



**Abb. 1:** Schematische Darstellung der Systemarchitektur. Zu unterst sind die unterschiedlichen Archivierungsschritte im zeitlichen Verlauf modellhaft gekennzeichnet.



**Abb. 2:** Zotero-Workflow: Ganz links eine Fileliste als Word-Liste; die Inhalte werden in ein Excel eingetragen (fehlt in Abb.). Mittig und rechts ist dann der Zotero-Eintrag zu sehen. Die Inventarnummer hier (Feld: Extra) entspricht dem Bibliotheks-Strichcode auf der Archivbox dieser spezifischen Sammlung.

Home

Globe Window

# Intelligent Read-Only Media Identification Engine

## Iron Maiden - The Indexer

11 Ergebnisse (0 Sekunden)

[file inventory "B00000077928"]

Expertensuche

Prev

1

2

Next

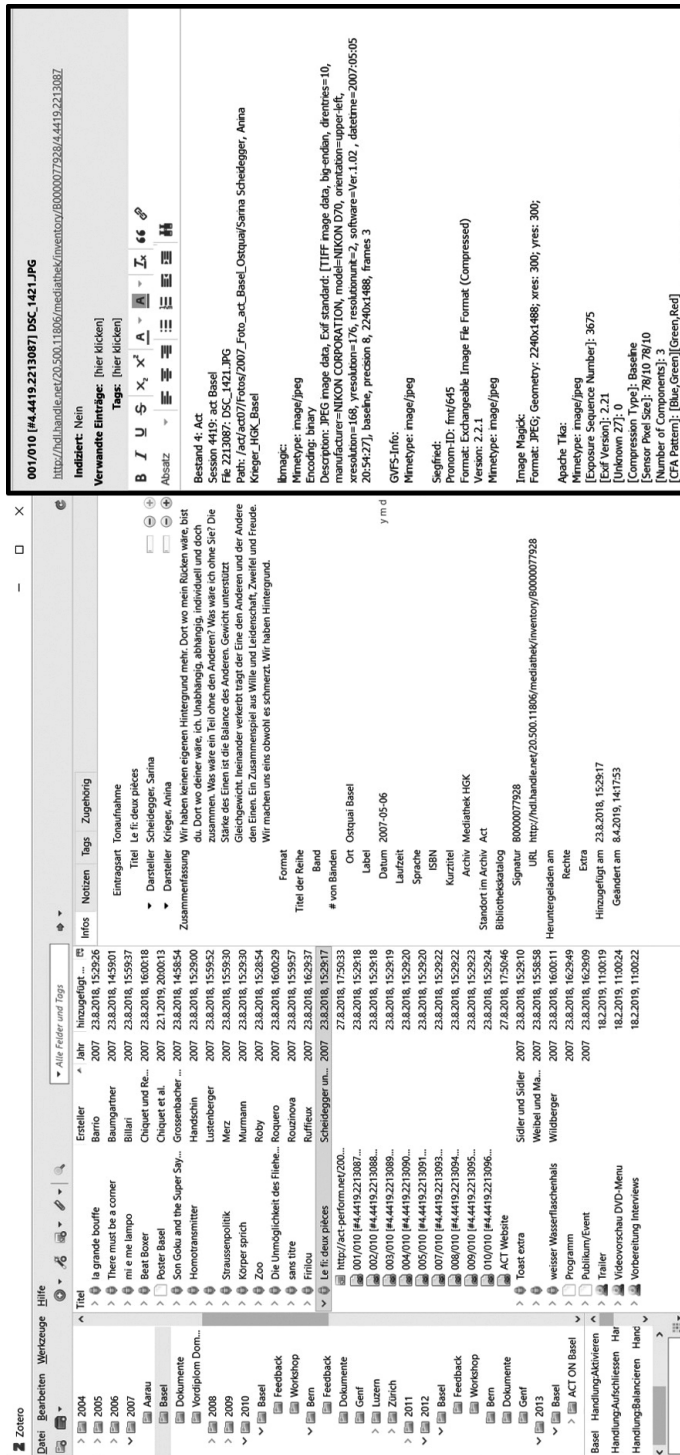
Page

	act	act_Basel	#4.4419.2213086	2018-08-20T14:45:52						#B00000077928	Exchangeable Image File Format (Compressed)
/act/acd07/Fotos/2007_Foto_act_Basel_Ostqual/Sarina Scheidegger, Anna Krieger_HGK_Basel											
	act	act_Basel	#4.4419.2213086	2018-08-20T14:45:52						#B00000077928	Exchangeable Image File Format (Compressed)
/act/acd07/Fotos/2007_Foto_act_Basel_Ostqual/Sarina Scheidegger, Anna Krieger_HGK_Basel											
content file litrmagic gvlfs-info tika imagick sieghied cite this item											
<p> <b>mimetype:</b> image/jpeg  <b>ion=</b> solutionunit=2, software=QuickTime 7.1.3, datetime=2007:05:24 14:36:40, hostcomputer=Mac OS X 10.4.9j, comment: "AppleLink", baseline, precision 8, 320x 240, frames 3         </p>											
	act	act_Basel	#4.4419.2101204	2018-05-25T17:57:41Z						#B00000077928	Exchangeable Image File Format (Compressed)
/act/acd07/Fotos/2007_Foto_act_Basel_Ostqual											
<p>            Sarina Scheidegger, Anna Krieger_HGK_Basel 4 KiB 2018-08-20T08:44:58Z         </p>											
	act	act_Basel	#4.4419.2213087	2018-08-20T14:45:52						#B00000077928	Exchangeable Image File Format (Compressed)
/act/acd07/Fotos/2007_Foto_act_Basel_Ostqual/Sarina Scheidegger, Anna Krieger_HGK_Basel											
<p>            DSC_1421.JPG image/jpeg 804.65 KiB 2018-08-20T08:44:58Z fmf0545         </p>											

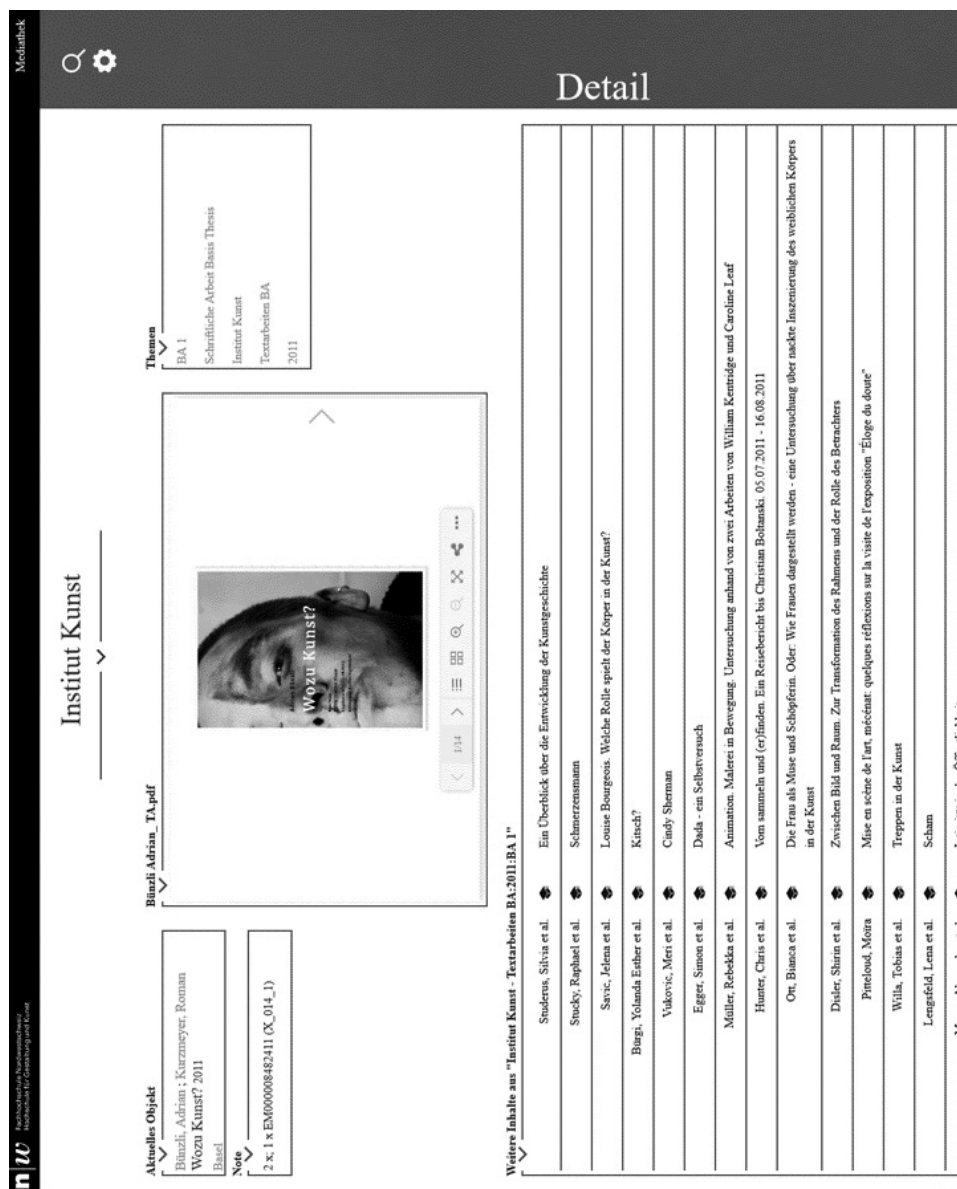
deutsches literatur archiv marbach

University of Applied Sciences and Arts Northwestern Switzerland Academy of Art and Design

n w



**Abb. 4:** Zotero-Eintrag eines Indexer-Objektes: Links die Sammlung, Mitte die semantischen Metadaten, Rechts die Indexer-Daten des Files 001/009 der Inventarnummer B0000078035



**Abb. 5:** Detailseite einer Theoriearbeit des Instituts Kunst; korreliert mit Abbildung 2.



www.musee-lu.com

© 2007 Kunsthaus Zürich  
Bildrechte für Gestaltung und Text

Museum

# ACT-Basel

## Aktuelle Objekte

Schidegger, Anna : Krieger, Anna  
Le fi, deux pièces 2007  
Ölfarbe auf Holz

## Note

<http://act-perform.net/2007/basel/sanna-schidegger-anna-krieger>

## Gallery

## ACT Website

<http://act-perform.net/2007/basel/sanna-schidegger-anna-krieger>

## Abstract

Wir haben keinen eigenen Hintergrund mehr. Dort wo mein Rücken war, bist du. Dort wo deiner war, ich. Unabhängig, abhängig, individuell und doch zusammen. Was wäre ein Teil ohne den Anderen? Was wäre ich ohne Sie? Die Stärke des Einen ist die Balance des Anderen. Gewicht unterteilt Gleichgewicht. Beinahe verkehrt trägt der Eine den Anderen und der Andere den Einen. Ein Zusammenspiel aus Wille und Leidenschaft. Zweifel und Freude. Wir machen uns eins obwohl es schmerzt. Wir haben Hintergrund.

## Weitere Inhalte aus "ACT-Basel - 2007-Basel"

Routinova, Anna	⌵	sans titre
	▶	Trailer
Baumgartner, Nino	⌵	There must be a corner
Barrio, Aude	⌵	la grande bouffe
Chiquet, Fabian et al.	⌵	Beat Boxer
Grossenbacher, Denise et al.	⌵	Son Goku and the Super Sayajin
Merz, Jan	⌵	Straussenspolitik
	▶	Programm

## Themen

Balancieren  
Berühren  
Schreiben  
Hochschule für Gestaltung und Kunst FH/NU Basel  
Abhängigkeit  
Balance  
Erfahrung  
Gegensätzlichkeit  
Gleichgewicht  
Sprache  
Unabhängigkeit  
Körper  
Stoff  
Wandelment  
ACT-Basel  
2007  
Basel

## Verweise

22066003 CIVIZZFKP  
22066003 T3JCBS8

## Detail

**Abb. 6:** Detailseite eines ACT-Eintrags als Resultat des Indexer-Zotero-Workflows; korreliert mit Abbildung 3 und 4.

## **23. zasedání pracovní skupiny Archivace záznamů z digitálních systémů (AUdS), Praha, 12. a 13. března 2019**

Editor: Karolína Šimůnková, Milan Vojáček

*Kolektivní monografie vznikla na základě konference Archivierung von Unterlagen aus digitalen Systemen (Archivace záznamů z digitálních systémů) a souvisejících seminářů, které se konaly ve dnech 12. a 13. března 2019 v Národním archivu v Praze.*

Vydavatel: Národní archiv, Archivní 2257/4, 149 00, Praha 4

Vydání: první

Tisk: Národní archiv, Archivní 2257/4, 149 00, Praha 4

Odborná redakce: Zbyšek Stodůlka, Karolína Šimůnková, Milan Vojáček

Jazyková redakce: autoři příspěvků

Překlad: Stefan Scholz (čl. Indexace audiovizuálních archivů s využitím metod automatického rozpoznávání řeči a obrazu)

Grafický návrh a sazba: © Eva Tamchynová

© Národní archiv, 2019

[www.nacr.cz](http://www.nacr.cz)

ISBN 978- 80-7469-100-3 (pdf)

## **23. Tagung des Arbeitskreises Archivierung von Unterlagen aus digitalen Systemen am 12. und 13. März 2019 in Prag**

Herausgeber: Karolína Šimůnková, Milan Vojáček, Národní archiv

Postanschrift: Archivní 2257/4, 149 01, Praha 4, Tschechische Republik

Herausgabe: 1.

Druck: Národní archiv, Archivní 2257/4, 149 01, Praha 4

Redaktion: Zbyšek Stodůlka, Karolína Šimůnková, Milan Vojáček

Korrektorat: Autorinnen und Autoren

Übersetzung: Stefan Scholz (Artikel: Die Indexierung von audiovisuellen Archiven unter Benutzung von Methoden der automatischen Sprach- und Bilderkennung)

Grafik, Satz: © Eva Tamchynová

© Národní archiv, 2019

(Alle Rechte für Beiträge liegen bei den jeweiligen Autorinnen und Autoren)

[www.nacr.cz](http://www.nacr.cz)

ISBN 978- 80-7469-100-3 (pdf)

