

## Archive und Online-Portale.

### Thesenpapier des ARK-Ausschusses „Archivische Fachinformationssysteme“ (AFIS)

#### Einleitung

Seit rund eineinhalb Jahrzehnten hat sich die Idee entwickelt, dass Kultureinrichtungen, nicht zuletzt Archive, das Wissen, das sie bewahren und das mehr und mehr in digitalisierter Form vorliegt, nicht nur in eigenen, einrichtungsspezifischen Online-Informationssystemen anbieten sollten. Die Idee von Online-Portalen, in denen diese Informationen – Erschließungsdaten wie Digitalisate – auch institutionen-, sparten- und/oder regionen- und staatenübergreifend den Nutzern zur Verfügung gestellt werden, verbreitete sich und wurde in verschiedenen Projekten und Vorhaben umgesetzt. Ein nationales Archivportal für Deutschland beispielsweise wurde erstmals 2003 von einer durch die Deutschen Forschungsgemeinschaft (DFG) eingesetzten Arbeitsgruppe gefordert.<sup>1</sup> Mit dem Start des Archivportals-D auf dem Deutschen Archivtag 2014 in Magdeburg ist somit ein wichtiges Etappenziel erreicht. Daneben und teils damit verbunden sind aber bereits andere Initiativen umgesetzt. Auf regionaler Ebene seien „Archive in NRW“<sup>2</sup> oder die Archivportale Niedersachsen und Thüringen<sup>3</sup> genannt, auf nationaler Ebene die Deutsche Digitale Bibliothek (DDB)<sup>4</sup> und auf internationaler Ebene das Archivportal Europa oder die Europeana<sup>5</sup>. Aber auch themenspezifische Portale wie monasterium oder das SED-/FDGB-Portal sind hier zu sehen.<sup>6</sup> Dass diese Aufzählung nur einen kleinen Ausschnitt aus der aktuellen „Portallandschaft“ darstellt, weist darauf hin, wie bedeutend die Thematik „Online-Portale“ mittlerweile für die Archive geworden ist. Die erste Erfahrung der letzten Jahre mit Portalen zeigt deutlich, dass die oben genannte Idee eine gute war: Die Portale können das einlösen, was man sich von ihnen versprochen hat - so beispielsweise ein verbessertes und komfortableres Angebot an Information für die Nutzer, aber auch eine größere Sichtbarkeit für die Archive. Mit diesen ersten Erfahrungen zeichnen sich aber auch Schwierigkeiten und Herausforderungen ab, die es nun gilt anzugehen. Der ARK-Ausschuss „Archivische Fachinformationssysteme“ (AFIS) hat vor diesem Hintergrund eine Reihe von Thesen entwickelt, die sich an die Archive und die Archiv-Community richten, und die die Punkte benennen, die aus seiner Sicht nun zu beachten sind, um die Idee der Online-Portale für die Archive in Deutschland weiter auf der Erfolgspur zu belassen.

---

<sup>1</sup> Gefordert wurde damals der „Aufbau eines Archiv-Portals, das die gesamte deutsche Archivlandschaft abbildet und einen benutzerfreundlichen Einstieg ermöglicht“, DFG-Arbeitsgruppe Informationsmanagement der Archive, Die deutschen Archive in der Informationsgesellschaft – Standortbestimmung und Perspektiven, [http://www.dfg.de/download/pdf/foerderung/programme/lis/Strategiepapier\\_archive\\_informationsgesellschaft151103.pdf](http://www.dfg.de/download/pdf/foerderung/programme/lis/Strategiepapier_archive_informationsgesellschaft151103.pdf), S.21; vgl. dazu Gerald Maier/Christina Wolf, Aufbau eines Archivportals-D innerhalb der deutschen Digitalen Bibliothek. DFG-Projekt zur Realisierung hat begonnen, in: Archivar 65 (2012), H. 4, [http://www.archive.nrw.de/archivar/hefte/2012/ausgabe4/ARCHIVAR\\_04-12\\_internet.pdf](http://www.archive.nrw.de/archivar/hefte/2012/ausgabe4/ARCHIVAR_04-12_internet.pdf), S. 404-406, hier S. 404. Das BAM-Portal wurde bereits seit 2001 aufgebaut, das Portal „Archive in NRW“ ist bereits seit 1998 am Netz, vgl. <http://www.bam-portal.de/> und Frank M. Bischoff, Das Projekt „Archive in NRW“ im Internet. Nutzung und Fortschreibung, in: Archivpflege in Westfalen und Lippe 53 (2000), S. 14-20, <http://www.lwl.org/waa-download/archivpflege/heft53/bischoff.pdf>. (Dieser Link sowie alle nachfolgenden Hinweise auf Informationen im WWW beziehen sich auf den Stand vom 28.07.2014.).

<sup>2</sup> <http://www.archive.nrw.de/>.

<sup>3</sup> <http://www.archivportal.niedersachsen.de/>; <http://www.archive-in-thueringen.de/>.

<sup>4</sup> <https://www.deutsche-digitale-bibliothek.de/>.

<sup>5</sup> <https://www.archivesportaleurope.net/>; <http://www.europeana.eu/>.

<sup>6</sup> <http://www.monasterium.net/>; <http://www.bundesarchiv.de/sed-fdgb-netzwerk/>.

## **Mehr Content – der gesamte Content**

Kein Portal ist lebensfähig ohne den Content, also die Inhalte seiner Datenlieferanten. Gleichzeitig ist einer der großen Vorteile von Portalen die übergreifende Recherche über den Datenbestand der einzelnen Einrichtung hinaus für eine gesamte Region, deutschlandweit, europaweit. Eine solche übergreifende Recherche ergibt aber nur Sinn, wenn zunehmend mehr Einrichtungen in einem Portal vertreten sind und gleichzeitig alle diese Einrichtungen auch alle Daten veröffentlichen.

Aus Sicht des ARK-Ausschusses AFIS muss es deshalb Ziel sein, dass nach und nach alle Archive in Deutschland mit allen ihren Erschließungsinformationen und Archivalien, die digitalisiert vorliegen, in den jeweiligen regionalen Archivportalen, im Archivportal-D und im Archivportal Europa vertreten sind. Nur dadurch werden Portale sinnvolle Rechercheinstrumente mit einem echten Nutzwert.

## **Eine durchdachte und einfache Portallandschaft**

Online-Portale mit Informationen aus Kultureinrichtungen gibt es mittlerweile eine ganze Reihe. Es ist – historisch gewachsen – eine „Portallandschaft“ entstanden, in der die verschiedenen Portale unabhängig voneinander aufgebaut wurden und oftmals unterschiedliche Träger und Betreiber haben.

Wenn hier nun die Forderung nach mehr, nach dem gesamten Content aufgestellt wird, bedeutet das im Umkehrschluss, dass die Betreiber der Archivportale aufgefordert sind, die einzelnen Portale sinnvoll miteinander zu vernetzen. Erste Schritte hierbei sind die mit dem Launch des Archivportals-D automatisch mögliche Datenweitergabe von diesem an das Archivportal Europa sowie der geplante Aufbau einer Infrastruktur bei verschiedenen regionalen Archivportalen, so dass diese zukünftig als Aggregatoren für das Archivportal-D fungieren können. Allerdings haben nicht alle Bundesländer solche regionalen Archivportale, die dieses Netz vervollständigen könnten. Auch die Datenweitergabe aus der DDB und dem Archivportal-D einerseits sowie dem Archivportal Europa andererseits an die Europeana ist grundsätzlich abgestimmt. Insgesamt sind aber zwischen den verschiedenen Portalen noch Abstimmungen im Detail notwendig, um die Portallandschaft in der Praxis funktionieren zu lassen. Darüber hinaus muss über die Einbeziehung weiterer Portale nachgedacht werden.

Idealerweise sollten Archive zukünftig mit nur einer Datenlieferung und dem Abschluss eines einzigen Datenüberlassungsvertrags in allen archivisch relevanten Portalen vertreten sein können.

## **Verbesserungen bei der Datenlieferung**

Als Mittel der Datenlieferung bietet OAI-PMH<sup>7</sup> ein normiertes Protokoll für regelmäßigen, automatischen Abruf, bei dem auch eine inkrementelle Lieferung möglich ist. Gleichzeitig zeigt aber die Erfahrung, dass OAI-PMH in der Praxis noch nicht reibungslos funktioniert. Insofern ist die aufwändigere Datenlieferung über FTP-Server derzeit aus pragmatischen Gründen vorzuziehen. Um für die datenliefernden Archive möglichst unkomplizierte und effiziente Lieferwege zu erreichen, sollte allerdings die Entwicklung bei OAI-PMH im Auge behalten und sein Einsatz immer wieder, auch hinsichtlich dessen Zweckmäßigkeit und Adaptionmöglichkeiten für EAD, geprüft werden.

Im Hinblick auf die Häufigkeit der Datenlieferung und der Onlinestellung von neuen und aktualisierten Daten ist es für die Archive aus Sicht des ARK-Ausschusses AFIS wichtig, dass eine regelmäßige Nachlieferung und Aktualisierung in möglichst kleinen Zeitabständen möglich ist.

---

<sup>7</sup> Open Archives Initiative Protocol for Metadata Harvesting; vgl. <http://www.openarchives.org/pmh/>.

Wichtig ist hier aber vor allem, dass bei Daten, die aus (datenschutz- oder urheber-)rechtlichen Gründen aus Portalen entfernt werden müssen, eine kurzfristige Löschung möglich ist.

### **EAD-Schnittstellen und stabile IDs**

Der Ingest-Prozess, also die Entgegennahme und Einbindung neuer Daten, ist beim Betrieb eines Online-Portals immer ein neuralgischer Punkt. Die Daten der einzelnen Lieferanten „fit“ zu machen zur Verwaltung und Präsentation im Portal, erfordert einen Mapping-Prozess<sup>8</sup>, der von Seiten des Portalbetreibers wie des Datenlieferanten Aufwand bedeutet. Um diesen Aufwand zu minimieren, wurden Austauschformate entwickelt. Für archivische Daten ist dies EAD<sup>9</sup>, zur genaueren Spezifikation nochmals enger gefasst in sog. Profilen oder Subsets – z. B. EAD(DDB) für die DDB/das Archivportal-D und im weiteren Sinn als gemeinsames Profil der deutschen Archive. Zu konstatieren ist hier allerdings, dass die auf dem Markt befindlichen Erschließungssoftwares zu großen Teilen noch über keine validen EAD(DDB)-Schnittstellen verfügen.

Außerdem sind für den Betrieb eines Online-Portals – mehr noch als beim Betrieb eines Online-Informationssystems für ein einzelnes Archiv – eindeutige und feste Identifier für jede Verzeichnungseinheit notwendig.<sup>10</sup> Auch diese werden von den Erschließungssoftwares nicht immer angeboten bzw. beim Export aus den Softwares nicht berücksichtigt.

Aus der Archiv-Community heraus hat es in Zusammenarbeit mit der EAD-AG und ausgehend von den beiden DFG-Projekten zur Digitalisierung und zum Archivportal-D bereits Workshops gegeben, um mit den entsprechenden Softwareherstellern zu diesen Themen ins Gespräch zu kommen. Dieser Prozess wird auch fortgesetzt. Gleichzeitig sind aber auch die Archive selbst gefordert: Am besten können sie als Kunden der Softwarehersteller diesen gegenüber ihre Wünsche platzieren und darauf aufmerksam machen, dass eine valide EAD(DDB)-Schnittstelle sowie feste IDs für ihre Teilnahme an Portalen notwendig sind.

### **Erschließung unter Einbeziehung von Normdaten**

Portale bieten durch ihren übergreifenden Ansatz ideale Voraussetzungen dafür, dass Erschließungsdaten und die zugehörigen Digitalisate von Archivalien besser vernetzt werden. Damit dies automatisch und zuverlässig erfolgen kann, ist aber die Einbeziehung von Normdaten bei der Erschließung wichtig, also von genormten und eindeutigen Bezeichnungen beispielsweise für Orte und Personen, wie sie die Gemeinsame Normdatei anbietet.<sup>11</sup> So können Digitalisate und die zugehörigen Erschließungsinformationen, die aus völlig unterschiedlichen Archiven stammen, sich aber beide auf denselben Ort oder dieselbe Person beziehen, automatisiert miteinander verknüpft werden. Aber auch über die Portale hinaus stehen die Daten dann in inhaltlicher Verbindung und im Kontext zu anderen Informationen aus dem Internet, wie es die Idee von Linked Data und des semantic web vorsieht.<sup>12</sup> Aus Sicht des ARK-Ausschusses AFIS sollten Archive deshalb bei der Erschließung wieder verstärkt Indices bei der Erschließung miteinbeziehen und diese dann mehr und

---

<sup>8</sup> Mapping bezeichnet das Übertragen von Daten aus einem Datenformat in ein anderes, beispielsweise aus den verschiedenen Originalformaten der Ausgangssysteme in das Datenformat eines Portals.

<sup>9</sup> Encoded Archival Description; vgl. dazu <http://www.landesarchiv-bw.de/ead>.

<sup>10</sup> Diese sind eine wichtige Voraussetzung für sog. Persistent Identifier, die beispielsweise in der DDB und dem Archivportal-D zukünftig verpflichtend sein werden.

<sup>11</sup> [www.dnb.de/gnd](http://www.dnb.de/gnd).

<sup>12</sup> Vgl. dazu Woitas, Kathi/Umlauf, Konrad (Hgg.), Bibliografische Daten, Normdaten und Metadaten im Semantic Web, Berlin 2013, urn:nbn:de:kobv:11-100209272); Danowski, Patrick/Pohl, Adrian, (Open-)Linked Data in Bibliotheken, Berlin 2013, <http://www.univie.ac.at/voeb/blog/?p=29122>; [http://de.wikipedia.org/wiki/Semantisches\\_Web](http://de.wikipedia.org/wiki/Semantisches_Web).

mehr selbstverständlich mit Normdaten versehen. Hierfür müssen auch Verfahren entwickelt werden, z. B. (teil-)automatisierte Datenanreicherungen, um fehlende personelle Kapazitäten bei vielen Archiven zu kompensieren.

### **Verstärkte Förderung der Digitalisierung**

Erschließungsdaten sind mit ihrem oftmals hohen und inhaltsreichen Informationsgehalt ein wichtiger Beitrag für Kulturgutportale. Deshalb müssen die strukturierte Erschließung mittels entsprechender Software und die Retrokonversionsbestrebungen in den Archiven weitergeführt werden. Zunehmend steht darüber hinaus aber die Forderung der Archivnutzer nach digitalisiertem Archivgut im Raum. So ist es ebenfalls wichtig, dass der Content in Archivportalen verstärkt auch aus Digitalisaten besteht. Auch im Hinblick auf die Nutzung von Archivportalen und damit auf deren Etablierung fordert der ARK-Ausschuss AFIS daher alle Entscheidungsträger in der Archivcommunity und in der weiteren Kultur- und Wissenschaftspolitik dazu auf, sich für die Förderung der Digitalisierung einzusetzen, diese weiterzuführen und zu verstärken. Ein Augenmerk muss dabei darauf liegen, dass der DFG-Digitalisierungspilot in eine Regelförderlinie einmündet. Darüber hinaus ist es aber auch wichtig, dass Bund, Länder, Kommunen und andere Träger die Digitalisierung von Archivgut (weiterhin) fördern und vorantreiben.

### **Angebote zur Speicherung und Langzeitsicherung von Digitalisaten**

In (Archiv-)Portalen können und sollen, wie oben beschrieben, auch Digitalisate präsentiert werden. Allerdings bieten die Portale in der Regel keine Möglichkeit, die Digitalisate zu speichern oder gar in ihrer Infrastruktur im Sinne einer Langzeitsicherung vorzuhalten. Bei FIZ Karlsruhe, dem Betreiber der DDB und des Archivportals-D, gibt es derzeit Überlegungen einen solchen Service kostenpflichtig anzubieten. Momentan sind die Archive jedoch selbst in der Verantwortung für die Speicherung und Langzeitsicherung der Daten. Da dies vor allem für kleinere Archive aber schwierig ist, sieht der ARK-Ausschuss AFIS hier das Desiderat für entsprechende Angebote.<sup>13</sup>

Für den Ausschuss: Wolfgang Krauth

01.08.2014

---

<sup>13</sup> Der ARK-Ausschuss AFIS wird hier die Entwicklung gemeinsam mit dem ARK-Ausschuss „Digitale Archive“ weiter verfolgen.