

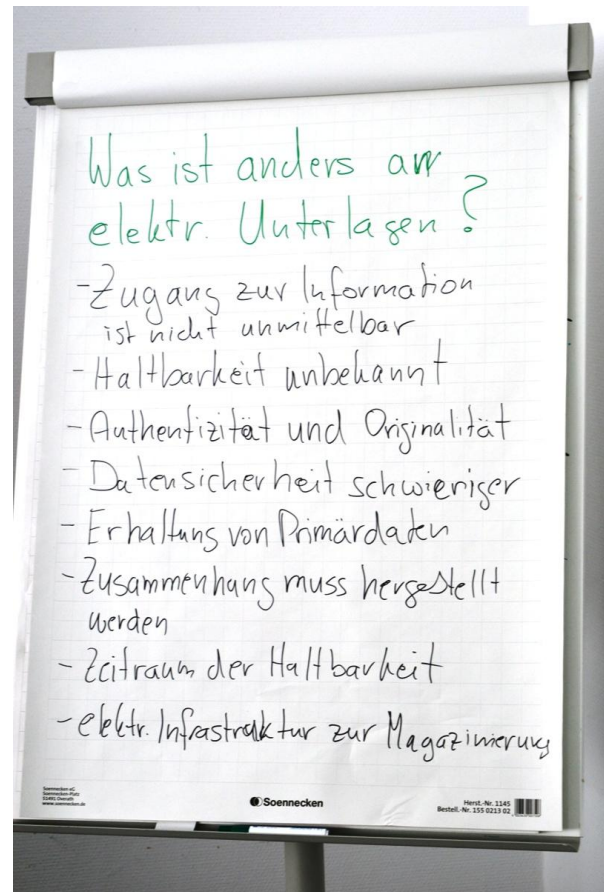
Das Digitale Archiv NRW – ein Seminar zur Einführung in die digitale Langzeitarchivierung

Vom 14. bis 16. März 2017 trafen sich Mitarbeiterinnen und Mitarbeiter aus Archiven und anderen Gedächtniseinrichtungen, um sich über den aktuellen Stand der digitalen Langzeitarchivierung im Allgemeinen und das Projekt „Digitales Archiv NRW“ im Besonderen zu informieren.

Der erste Tag widmete sich vor allem den Grundlagen der digitalen Langzeitarchivierung (LZA). Nach der Vorstellungsrunde und einer Gruppenarbeit zum Unterschied zwischen klassischen und elektronischen Unterlagen, stellten Dr. Benjamin Bussmann (Hist. Archiv der Stadt Köln) und Dr. Michael Habersack (LVR-AFZ, Pulheim-Brauweiler) die zurzeit diskutierten Strategien und die Herausforderungen der LZA vor. In einem weiteren gemeinsamen Vortrag vertieften sie die Theorie der Migrationsstrategie und stellten Normen und Konzepte der LZA, wie das OAIS-Referenzmodell vor. Dabei lernten die Teilnehmenden en passant das Fachvokabular kennen. Abgerundet wurde ihr Vortrag durch eine Präsentation von Jens Peters (LVR-Infokom; Köln), der sich den vielfältigen Dateiformaten und ihrer Eignung für die digitale LZA widmete. Er sensibilisierte die Zuhörer für die verschiedenen Ausprägungen gleich lautender Dateiformate und die damit verbundenen Schwierigkeiten, verlässliche Aussagen über deren Identität und Validität zu treffen.

Am Nachmittag griff Dr. Habersack die Gruppenarbeit des Vormittags wieder auf und stellte den Workflow bei digitalen Daten den gängigen Verfahren der analogen Archivarbeit gegenüber. Eine weitere Gruppenarbeit, die nach den Haupttypen der elektronischen Überlieferung fragte, bildete den Einstieg in den nächsten Theorieblock. Dr. Julia Krämer-Riedel (Hist. Archiv der Stadt Köln) und Tim Knoche (Amt f. Informationsverarbeitung der Stadt Köln) erläuterten die Archivierung strukturierter Daten, wie sie z. B. in Fachverfahren oder in E-Akten entstehen. Dabei wurden die Werkzeuge SIARD und Ingestlist zur Archivierung von Datenbanken ebenso vorgestellt wie diverse Fachverfahrenslisten, die den Einstieg in die Bestandsaufnahme in der eigenen Institution erleichtern. Für die Übernahme von un- oder schwach strukturierten Daten wie z. B. File-Ablagen, die spartenübergreifend in den Registraturen der meisten Einrichtungen vorkommen, empfahl Dr. Michael Habersack eine konsequente Strukturierung im Vorfeld. Er mahnte die Anwendung von Aktenplänen und Dateinamen-Konventionen an und verdeutlichte dies an Beispielen. Vor allem seine Ausführungen über Doppelablagen und unbrauchbare Datei- und Ordernamen ließ manchen Teilnehmenden Parallelen zur Lage im eigenen Haus entdecken.

Zum Abschluss des ersten und als Einstimmung für den folgenden Tag stellte Jürgen Plat-

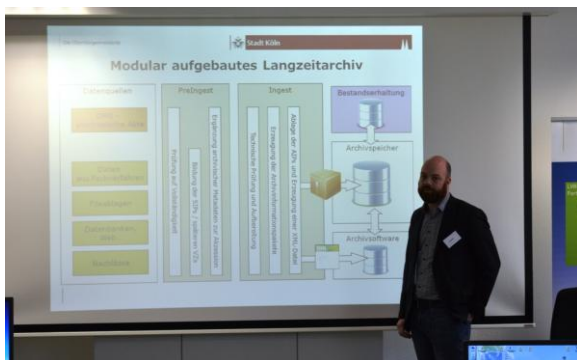


Annäherung an das Thema LZA mit klassischen Mitteln

te (Geschäftsstelle des DA NRW, Dortmund) den Lösungsverbund DA NRW vor. Er berichtete über den Auftrag der Landesregierung ein träger-, sparten- und institutionen-übergreifendes digitales Langzeitarchiv für digitale und digitalisierte Kulturgüter aus staatlichen und nicht-staatlichen Archiven, Bibliotheken, Museen aufzubauen. Aus den unterschiedlichen Anforderungen der Kulturinstitute entwickelten mehrere Dienstleister des Landes und der Kommunen technische Lösungen, die nun unter dem Dach des DA NRW bereit gestellt werden. Dies erlaube die Nutzung von Synergien für die beteiligten Institutionen und ein Set passender Lösungen für die unterschiedlichen Kulturinstitute.

Als Vorbereitung auf die Beschäftigung mit der Verbundlösung DiPS.kommunal vertiefte Dr. Bussmann zum Auftakt des zweiten Tages den Blick auf die digitale Bestandserhaltung und die signifikanten Eigenschaften. Deren Definition stellte sich den Teilnehmenden als eine der Bewertung von analogem Schriftgut vergleichbare zukünftige archivische Kernaufgabe dar. Thorsten Preuss (Amt für Informationsverarbeitung der Stadt Köln) fuhr mit einer Übersicht über aktuelle Kooperationen im Bereich der Langzeitarchivierung fort.

Im nachfolgenden Themenblock stellten Thorsten Preuss und Tim Knoche die Verbundlösung DiPS.kommunal vor und demonstrierten die Bearbeitung einer Test-Einlieferung mit dem PreIngest-Tool PIT Plus und die anschließende Weiterverarbeitung im Web Cube.



Tim Knoche erläutert den Teilnehmenden DiPS.kommunal in der Theorie.

Die Sicht des Amtes für Informationsverarbeitung der Stadt Köln als Anbieter der Software wurde anschließend durch einen Werkstattbericht von Dr. Habersack aus der Mandantensicht des LVR-AFZ ergänzt, der sich auf erste Tests beschränken musste, da die Software-Lösung erst seit sehr kurzer Zeit im AFZ zur Verfügung steht. Der Nachmittag stand den Teilnehmenden zur praktischen Erprobung des PreIngest mit PIT Plus zur Verfügung. Anhand von Testdaten verschiedenster Formate übten sie unter fachlicher Anleitung des Teams vom Historischen Archiv und des Amtes für Informationsverarbeitung der Stadt Köln die Strukturierung und Anreicherung der Daten mit Metadaten und die Überprüfung der technischen Verarbeitungsprozesse.

Zum Auftakt des dritten Tages sorgte Dr. Habersack zunächst für einen Überblick über Identifikatoren und erläuterte die Unterschiede zwischen URN, URL und URI. Die Ausführungen bildeten die Grundlage für einen Exkurs über das Zusammenspiel des entstehenden DA NRW und seinem geplanten Portal, das den Zugang der Bürgerinnen und Bürgern zum kulturellen Erbe Nordrhein-Westfalens sicher stellen soll, mit dem bereits seit über einem Jahrzehnt aktiven Landesarchivportal Archive.NRW. Dr. Habersack hob die unterschiedliche Zielsetzung und Funktion der Portale hervor. Während das Portal „Archive in NRW“ auf die strukturierte Recherche zielt und nicht für die Anzeige größerer Mengen von Digitalisaten konzipiert sei, setze das neue Portal des DA NRW genau darauf, Objekte des kulturellen Erbes NRW einer breiten Öffentlichkeit zu präsentieren, ohne den tektonischen Zusammenhang abzubilden.

Dann wurde es erneut konkret und Jens Peters und seine Kollegin Gaby Bender stellten die zweite Verbundlösung DNS Software Suite vor. Ähnlich wie am Vortag erhielten die Teilnehmenden zunächst einen Überblick über die Komponenten der LZA-Lösung und eine Demonstration des SIP-Builders anhand von Testdaten. Auch der SIP-Builder ermöglicht es den als archivwürdig eingestuften Dateien umfangreiche Metadaten hinzuzufügen. Ein Schwerpunkt liegt im Rechte-Management für die spätere Präsentation im DA NRW Portal.



Nach der Demonstration erhielten die Teilnehmenden die Gelegenheit, den SIP-Builder selbst auszuprobieren und einige Testdaten für die Einlieferung ins DA NRW vorzubereiten. Die an den SIP-Builder anschließenden Identifikations- und Migrationsprozesse im sogenannten Content Broker, die aus dem SIP ein Archival Information Package (AIP) bilden, erfolgen automatisiert, so dass der Nutzer nicht manuell eingreifen kann.

Die Teilnehmenden testen den Sip-Builder der DNS-Software Suite.

Im Anschluss berichtete Ralf Matalla (ULB Düsseldorf) als langjähriger Projektpartner der DNS-Lösung aus der Anwendersicht. Er machte deutlich, dass die ULB Düsseldorf mit ihren für die digitale LZA vorgesehenen digitalen Sammlungen über sehr homogenes Material verfüge, das fast ausschließlich im Tiff- und PDF-Format vorliege. Die Preingest-Prozesse sind mit der Bibliothekssoftware verknüpft, so dass die Datenübermittlung problemlos verlaufe. Aus seiner Sicht ergeben sich vor allem Schwierigkeiten im Bereich der sehr großen zu verarbeitenden Datenmengen bzw. der zu geringen Verarbeitungsgeschwindigkeit auf Seiten des Digitalen Archivs.

Bevor Michael Habersack den Teilnehmenden dann einen Überblick über die Langzeitarxivlösungen und den Stand der Umsetzung in Deutschland verschaffte, griff er noch einmal das durchgängig erörterte Thema der Datenformate auf. Was zuvor in der Theorie erörtert wurde, konnten die Teilnehmenden nun mit Hilfe des Programms „shotgun“ selbst erproben. Sie simulierten an gängigen Dateitypen den Verlust von bits (so genannten „bit rot“) in unterschiedlichem Ausmaß und konnten im Anschluss die Ergebnisse der zerstörerischen Eingriffe direkt beobachten. Dabei wurde die unterschiedliche Widerstandsfähigkeit der verschiedenen Datenformate augenscheinlich.

Zum Abschluss der drei informationsreichen Tage resümierte die Moderatorin Monika Marner M. A. (LVR-AFZ, Pulheim-Brauweiler) die Veranstaltung und leitete damit die Abschlussdiskussion ein. Es wurde deutlich, dass der Aufbau der digitalen LZA sich permanent in der Entwicklung befindet und belastbare Praxiserfahrungen weiterhin rar sind. Dennoch wurde dem Publikum deutlich, dass sie die Initiative ergreifen müssen, um notwendige erste Schritte wie die Identifizierung („Fachverfahrensliste“) und Strukturierung („elektronischer Aktenplan“) der entstehenden Daten in den Verwaltungen zu gehen und den Aufbau der LZA vorzubereiten und vor allem dort, wo das bisher nicht der Fall ist, die Kommunikation mit der jeweiligen IT-Abteilung in Gang zu setzen.

Text: Monika Marner M. A.

Fotos: Dr. Hanns-Peter Neuheuser-Christ