Von der Theorie zur Praxis – Bestandserhaltung digitaler Unterlagen
Workshop des KLA-Ausschusses Digitale Archive im Bundesarchiv Koblenz

Im Bundesarchiv Koblenz trafen sich November 2018 auf Einladung des KLA-Ausschusses Digitale Archive einige Expertinnen und Experten aus den Bereichen Archiv, Bibliothek, Wirtschaft und Forschung, um eine Bestandsaufnahme der Bestandserhaltung digitaler Unterlagen im deutschsprachigen Raum mit Ausblicken in die internationale Diskussion zu wagen.

Im vergangenen Jahrzehnt sind digitale Bestandserhaltung und Nachhaltigkeit weltweit aus der Phase der Theorie und der Konzeption in die Phase der Umsetzung getreten. Während die Praxis in der deutschen Bibliotheks- und Archivwelt bis 2010 vor allem von Digitalisierung im Sinne von Scannen handelte, so stellt dieses Geschäft inzwischen eine Routinehandlung dar. Die heutige Herausforderung ist, die großen Mengen digitaler Objekte, die in den Archiven und in den Behörden entstanden sind, für eine dauerhafte Verwahrung zu ertüchtigen. Nicht zuletzt sind inzwischen Bücher erschienen, die eine Gesamtschau des Problems bieten, es historisch und gesellschaftlich einbetten und uns den Mut geben, einfach selbst die Verantwortung zu übernehmen.[1]

Warum digitale Nachhaltigkeit die gesamte Gesellschaft betrifft, erklärte William Kilbride (Digital Preservation Coalition DPC, Glasgow) mit seinem Vortrag über die „Bit List“ der digital bedrohten Arten.[2] Das Projekt führte laut Kilbride zu drei Kern-Thesen:

- Wir reden über undeutliche Gefahren für Daten oder Software, aber die Verluste sind echt und betreffen uns alle im Alltagsleben.
- Dies ist eine soziale und technische Herausforderung. Wenn wir nur technische Lösungen anbieten, haben wir keinen Erfolg.
- Digitale Bestandserhaltung handelt nicht von einem “dunklen Zeitalter”. Es geht darum, dem Versprechen des digitalen Zeitalters gerecht zu werden.

Bei der gemeinsamen Aufgabe hilft inzwischen eine größere Zahl von Softwareprodukten, deren Nutzer im Rahmen des Workshops zusammenkamen. Dabei kristallisierte sich heraus, dass die zentralen Systemkomponenten meist kostenpflichtig sind, während sich für viele Peripherieprodukte eine gemeinsame Nutzung anbietet. Ein Beispiel sind Erkenner und Validierer wie DROID, JHOVE oder veraPDF, die aber bei weitem noch nicht ausreichen und gemeinsam ergänzt werden müssen. Dies zeigten auch die Referate von Michelle Lindlar (TIB Hannover) und Yvonne Tunnat (ZBW Kiel).

Als neues Produkt dieser Art stellte Martin Wrigley (Open Preservation Foundation, OPF) das Preservation Action Registry (PAR) vor, welches auf einer gemeinsamen Datenbank (ähnlich wie PRONOM für DROID), die ein einheitliches Format für Formatprüfung und Erhaltungsaktionen vorsieht, beruht. Gründungspartner sind die Unternehmen hinter den Produkten Archivematica und Preservica. Die OPF will Projekte und Zwischenziele für das PAR genauer definieren und zeitnah Anwendungsfälle mit gut sichtbarem Nutzen bekannt machen.

Ein weiteres Novum war das Referat von Tobias Steinke und Nathalie Küchler von der Deutschen Nationalbibliothek (DNB), das ein Praxisproblem der Emulation betraf. Wer veraltete Computerumgebungen emulieren will, kommt nicht ohne die Erlaubnis der Firmen

---

[1] Christian Keitel: Zwölf Wege ins Archiv. Umrisse einer offenen und praktischen Archivwissenschaft, Stuttgart 2018; Trevor Owens: The Theory and Craft of Digital Preservation, Baltimore 2018; Heather Ryan, Walker Sampson: The No-nonsense Guide to Born-digital Content, London 2018. Zur Archivaliengattung der Dateisammlung auch Kai Naumann, Michael Puchta: Kreative digitale Ablagen und die Archive, München 2017.

[2] BitList: https://www.dpconline.org/our-work/bit-list (abgerufen 25.2.2019).

aus, die das Urheberrecht an dem Betriebssystem und anderen Programmen besitzen. Weder der nationale noch der EU-Gesetzgeber hat den Gedächtnisinstitutionen ausreichende Befreiungen von diesen Pflichten geliefert. Die Erlaubnis zur Nutzung konnten aber für die DNB überraschend leicht durch Anschreiben der IT-Industrie eingeholt werden. Einige Lücken ließen sich durch Open-Source-Ersatzprodukte schließen. Die einzige wichtige Ausnahme war Apple – von dieser Firma hatte die DNB bis dato trotz hartnäckiger Nachfragen noch keine Antwort bekommen.
In der Diskussion wurde dargelegt, dass bei Formaten, die ISO-Standards darstellen, mit dem Auslaufen der mit dem Standard verbundenen Patente (nach 20 Jahren) keine Entgeltforderungen durch die am Standard beteiligten Parteien mehr möglich sind.
Wie sich digitale Bestandserhaltung inzwischen verwirklichen lässt, zeigte der Werkstattbericht von Michael Tobegen (Staatsarchiv Hamburg) über ein Konversionstool für DIMAG. Das Bestandserhaltungsmodul BEM sucht im Digitalen Magazin nach Dateien in unerwünschten Formaten, konvertiert diese und spielt sie als neue Repräsentation ins Magazin zurück. Einen Lösungsansatz für das Problem, wie eingebettete Dateien archivierbar bleiben, stellte Rainer Jacobs (Bundesarchiv) vor. Über eine in Python geschriebene Anwendung ist es möglich, die Dokumente einzeln darzustellen, ohne den ursprünglichen Zusammenhang dabei aufzugeben.
Besondere Aufmerksamkeit verdienen auch die Vorträge, die von außerhalb der Archiv- und Bibliothekswelt kamen. Andreas Trautheim-Hofmann (PROSTEP AG, Darmstadt) berichtete aus der Automobil- und Luftfahrtindustrie. Dort sind Schadensersatz und Haftungsfragen auch nach Jahrzehnten noch bedeutsam, weshalb Produktionsdaten (Zeichnungen, Pläne, 3D-Modelle) langfristig verfügbar bleiben müssen. So sind einige Standards entstanden, die teils als ISO, teils als Europäische Norm (EN) verbindlich geworden sind. In ähnlicher Weise schaffen sich auch die Bauindustrie, die Filmindustrie und die Geo-Informatik verbindliche Standards, wie Michelle Lindlar (TIB Hannover), Siegfried Fößel (Fraunhofer IIS, Erlangen) und Wolfgang Kresse (Hochschule Neubrandenburg) zeigten.
Festzuhalten ist, dass Archive und Bibliotheken in gemeinsamen Bestrebungen zur digitalen Langzeitarchivierung näher zusammenzurücken müssen. Bei nestor wird der Erfahrungsaustausch bereits gepflegt, doch lässt sich die gegenseitige Beteiligung von Teilnehmern aus der jeweils anderen Community noch ausbauen.
Außer den genannten Personen referierten Christian Keitel (Überblick), Jenny Mitcham (Officedateien der 90er Jahre) und Martin Hoppenheit (Datenbankarchivierung). Wer im Detail die Inhalte der Vorträge nachlesen möchte, kann eine ausführliche Mitschrift und die Folien auf den Internetseiten des Bundesarchivs einsehen:
https://www.bundesarchiv.de/DE/Content/Artikel/Ueber-uns/Partner/KLA/kla-ausschuss-digit-arch.html