

iPRES 2016 in Bern

13th International Conference on Digital Preservation

Angela Gastl, Franziska Geisser, Ana Sesartić, Roland Suri und Matthias Töwe^{1,2}

Die International Conference on Digital Preservation (iPRES) pendelt im Dreijahresrhythmus zwischen Europa, Amerika und Asien / Australien. Vom 3. bis zum 6. Oktober 2016 lud die Schweizerische Nationalbibliothek (NB) als Gastgeberin ins Kongresszentrum nach Bern und 340 Teilnehmerinnen und Teilnehmer aus 30 Ländern folgten der Einladung.

Die iPRES bietet Beiträge aus der Forschung zur Langzeitarchivierung und aus der Praxis archivierender Institutionen, so dass neben Kurz- und Langvorträgen, Podiumsdiskussionen und Postern auch drei Halbtage mit Workshops und Tutorials angeboten wurden. Zur inhaltlichen Qualität tragen die Spezialisierung der Konferenz sowie die Begutachtung der ausformulierten Beiträge bei. Der folgende Bericht gibt einen groben Überblick aus persönlichen Eindrücken der Autorinnen und Autoren. Allen Interessierten sei ein Blick in den Online-Konferenzband³ sehr empfohlen.

Eröffnung und Keynotes

Marie-Christine Doffey begrüßte als Direktorin der NB die Gäste. Die NB engagiert sich seit Jahren für die digitale Langzeitarchivierung und konnte nun auch die iPRES ins eigene Land holen. Es folgte ein Grußwort im Zeichen des digitalen Wandels von Christa Markwalder, Präsidentin des Nationalrates, der größeren der beiden Parlamentskammern der Schweiz. Barbara Signori von der NB führte als federführende Organisatorin nach jahrelanger Vorbereitung

mit spürbarer Freude in die Tagung ein.

Die drei ersten Konferenztage wurden jeweils mit einem Keynote-Vortrag eröffnet.

Digital Object Architecture

Den Anfang machte Robert E. Kahn, Gründer der Corporation for National Research Initiatives (CNRI), und einer der Väter des TCP/IP Protokolls. Er forderte ein Nachdenken über zukunftsfähige Strukturen und empfahl die Digital Object Architecture (DOA), um Objekte nicht nur aufzufinden, sondern auch sicherzustellen, dass sie gemäß der ursprünglichen Absicht interpretiert werden können. Voraussetzung sei die Maschinen- und Technologie-Unabhängigkeit des Datenmodells der DOA. Für eine weitgehende Interoperabilität müssten Systeme, User, Aktivitäten und Objekte jeweils mit eindeutigen Identitäten versehen werden.

Als Organisationsstruktur der DOA, zu der auch das Handle System gehört, wurde mit der DONA Foundation⁴ mit Sitz in Genf eine international akzeptierte Organisation bewusst außerhalb der USA geschaffen, um das Global Handle Registry durch „Multi-Primary Administrators“ zu verwalten. Der abstrakte und anspruchsvolle Vortrag schloss mit einigen Appellen: „Think more broadly! Be aware of new data structures. Interoperability is key. Security is important.“

Erhaltung digitaler Kunstwerke

In die Herausforderungen der Erhaltung digitaler Kunst führte der Keynote-Vortrag von Sabine Himmelsbach ein. Sie leitet das Haus der elektronischen Künste in Basel und veranschaulichte unter dem Titel „Net-based and Networked – Challenges for the Conservation of Digital Art“ die Herausforderungen und Lösungsansätze zur Erhaltung dieser Objekte (Abb. 1). Für einige Kunstwerke ist das Internet nicht nur technische Basis, sondern Teil des Werkes, indem zum Beispiel Inhalte aus dem Internet abgerufen und integriert werden. Sabine Himmelsbach betonte das sehr individuelle Vorgehen und die Herausforderung, die signifikanten Eigenschaften zu ermitteln. Notwendig sei es, die umgebenden Informationen und Beziehungen zu einem Werk zu erhalten.

Aus technischer Sicht erfordert die Erhaltung von Kunstwerken in der Regel die Emulation einer überholten Umgebung auf aktueller Hardware. Beispiele dazu wurden in weiteren Vorträgen präsentiert (s.u.).

Die Technik und wir

Der Vortrag von David Bosshart vom Gottlieb Duttweiler Institut⁵ diskutierte „Digital Memories, Digital Futures – The Interplay between Humans and Technology“. Viele Aussagen zur digitalen Durchdringung waren nicht neu („Everything is now“,

1 Das Autorenteam dankt Mohamed El-Saad für seine Durchsicht des Manuskripts.

2 Hinweis zu Interessenkonflikten: Alle Autorinnen und Autoren sind Angestellte der ETH-Bibliothek, die als Sponsor der Konferenz auftrat.

3 http://www.ipres2016.ch/frontend/organizers/media/iPRES2016/_PDF/IPR16.Proceedings_4_Web_Broschuere_Link.pdf [02.11.2016]

4 <https://dona.net/> [02.11.2016]

5 <http://www.gdi.ch/> [02.11.2016]

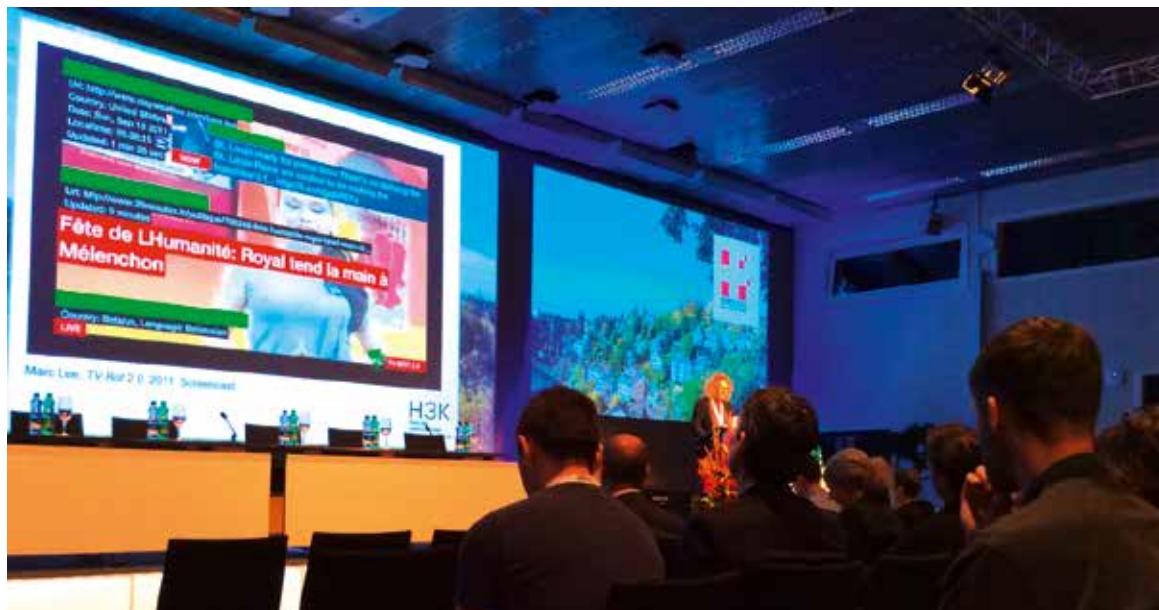


Abbildung 1:
Einen inspirierenden
Beitrag lieferte die
Keynote von Sabine
Himmelsbach vom
Haus der elektro-
nischen Künste in
Basel
(Foto: M. Töwe)

„Reality becomes negotiable“, „Record – Upload – Share – Create feedback loops“), die Verweise auf die Nebenwirkungen für das menschliche Verhalten aber durchaus bedenkenswert, etwa die latente Überforderung durch immer mehr Wahlmöglichkeiten sowie das Verschwinden klarer Grenzen zum Digitalen. Für Bosshart kann darin sogar eine gewisse Befreiung liegen, weil Menschen sich nicht mehr auf separate Geräte konzentrieren müssten, sondern wieder miteinander interagieren könnten. Er schloss mit Fragen: Wie stellen wir sicher, dass validiertes Wissen nicht einfach verschwindet? Wer hat die Verantwortung, wenn die Unterscheidung zwischen privat und öffentlich verschwimmt? Wie organisieren wir das Sammeln und Bewahren?

Aus dem Publikum wurde der Optimismus hinsichtlich der Verfügbarkeit des Internets kritisch hinterfragt, u.a. mit Hinweis auf die Angreifbarkeit des Internet of Things. Bosshart erwiderte, wir müssten und könnten lernen, mit dem Risiko umzugehen, und der Einzelne müsse ausreichend gut die Technologien verstehen, um zu realisieren, wenn er sich in Gefahr begeben. David Rosenthal (Stanford Univ. Libraries) fasste die Bedenken trocken zusammen: „Optimism is a very poor basis for digital preservation.“

Themen im Querschnitt

Als ein roter Faden kann die Herausforderung erwähnt werden, Daten mehr als „nur“ technisch zu erhalten: Entscheidend sind die Kontextinformationen, etwa für die Einschätzung, welche Eigenschaften digitaler Objekte erhalten werden sollten. Ebenso geben sie Hinweise zu Abhängigkeiten und schließlich bildet der Kontext die Voraussetzung dafür, dass digitale Objekte später auch außerhalb ihres Entstehungskontexts genutzt werden können.

Es gab mehrere Plädoyers dafür, statt retrospektiver Langzeitarchivierung stärker im Sinne einer „Digital Continuity“ zu denken und zu handeln. Durch vorausschauende Aktivitäten etwa bei der Softwareentwicklung sollen spätere Probleme vermieden oder zumindest die Möglichkeiten zu ihrer Behebung verbessert werden. Ein Beispiel hierfür sind die Ziele von UNESCO-PERSIST (Platform to Enhance the Sustainability of the Information Society Transglobally)⁶, in dessen Rahmen mit großen Softwarefirmen über Verbesserungen bei der Verfügbarkeit von obsoleter Software verhandelt wird (Natasia Milic-Frayling (Univ. of Nottingham)).

⁶ <http://en.unesco.org/news/persist-unesco-digital-strategy-information-sustainability?language=en> [03.11.2016]

OAIS – Standard oder „Schmusedecke“?

Die Einschätzung, dass die retrospektive Erhaltung zu spät greift, spielte auch bei der Diskussion zur Zukunft des OAIS-Referenzmodells eine Rolle. Gefordert wurden eine stärkere Berücksichtigung des Pre-Ingest und die Einführung von Representations und Intellectual Entities gemäß dem PREMIS-Standard. Es wurde sogar die Frage gestellt, ob die Mängel des OAIS-Referenzmodells sich überhaupt reparieren lassen. Als Referenzmodell sei OAIS trotz seines ISO-Status kein Standard, mit dem man klar validierbar konform sein könne, anders als etwa die konkretere Norm ISO 16363. Wie Steve Knight (National Library of New Zealand) überspitzt formulierte, könne man „alles“ damit machen, gefragt sei aber ein klarer Benchmark zur Frage: „Was muss ein digitales Archiv können?“ In der jetzigen Form erfülle der Standard für die Community eher die Funktion der Schmusedecke von Linus aus den Peanuts-Comics.

Werkzeuge spriessen – und dann?

Verschiedene Projekte stellten Werkzeuge zur Unterstützung und Automatisierung von Arbeitsschritten vor. In der Podiumsdiskussion zur Zusammenarbeit der nationalen Netzwerke kam allerdings zum Ausdruck, dass

sogar einige der intensiv genutzten Werkzeuge immer noch auf wackligen Füßen stehen und ihre Weiterentwicklung nicht gesichert ist. Ein Lichtblick ist die Übernahme der Verantwortung für JHOVE durch die Open Preservation Foundation (OPF), deren Mitglieder den Aufwand tragen, wie Joachim Jung von der OPF erläuterte. Ob dieses Modell auf andere Tools erweitert werden könnte, blieb offen. Steve Knight (National Library of New Zealand) stieß mit seiner Aufforderung zu einer Art Crowdfunding für Unterhalt und Entwicklung von Tools nicht auf grundsätzlichen Widerspruch, enthusiastische Begeisterung blieb allerdings auch aus. Hier spielt sicher eine Rolle, dass die wenigsten Institutionen Erfahrungen mit solchen Modellen haben. Im Folgenden werden ausgewählte Einzelvorträge der Konferenz vorgestellt.

Systeme und „Stewardship“

Sven Schlarb (Austrian Institute of Technology) stellte das System E-ARK Web⁷ und seine Referenzimplementierung im Projekt E-ARK (European Archival Records and Knowledge Preservation)⁸ vor. Neben den „typischen“ am OAIS orientierten Funktionen sind die Möglichkeiten zur Nutzung von Named Entity Recognition und anderen Data Mining Funktionen erwähnenswert, die sich auf eine Volltextindexierung stützen. Die Präsentation von Sibyl Schaefer (Univ. of California, San Diego) zur Migration des Chronopolis Preservation Systems zeigte, dass die Langzeitarchivierung bereits ihre eigenen „Altlasten“ erzeugt hat. Als Lehren aus der Erneuerung von Middleware und Anwendung zog Schaefer, möglichst unabhängig voneinander ersetzbare Microservices zu verwenden, jederzeit den direkten Zugriff auf die Daten sicherzustellen und möglichst „langweilige“ Technologien zu wählen, die

ohne sehr spezielles Know-how betretet werden können.

Dateiformat – das unbekannte Wesen, Teil I

Im Block „Quality & Risk Assessment“ sollte der Beitrag von Scott Witmer (Univ. of Illinois at Urbana-Champaign) eigentlich zeigen, welche Schwierigkeiten bei der Nutzung von Dateien aus einem Repositorium mit reiner Bitstrom-Erhaltung auftreten können (fehlende oder falsche Dateiendungen, als invalid erkannte Dateien, fehlende Software, kein Rendering in verfügbarer Software). Eher unfreiwillig kam bei Tests zur Wiedergabe von JPEG2000 mit Photoshop zum Ausdruck, wie sich die Wahl ungeeigneter Software auswirkt und wie schwierig eine geeignete Auswahl sein kann. Wann also ist das erfolgreiche Öffnen einer Datei ein hinreichendes Kriterium für ihre Validität? Nicht selten sind Viewer robuster bzw. unkritischer, als es die Spezifikation verlangt.

Peter May (British Library) verglich die Resultate verlustfreier Komprimierungen für TIFF-Dateien (LZW, ZIP, Group 4). Die Ergebnisse hängen von Farbtiefe, Anzahl vorkommender Farben und der „Entropie“ im Bild ab, zusätzlich erzielten die Tools LibTiff und ImageMagick nicht die gleichen Resultate. Dennoch gibt es für die Bedürfnisse der British Library eine Empfehlung: Group 4-Komprimierung für 1-bit Bilder und ZIP für andere Farbtiefen. Im Vergleich komprimiert ImageMagick meist stärker als LibTiff, obwohl ImageMagick die LibTiff Libraries verwendet. Insgesamt zeigt diese „einfache“ Fragestellung, dass die Abhängigkeiten kompliziert und teilweise komplex sind.

Schwerpunkt Digitale Kunst und Emulation

Klaus Rechert (Univ. Freiburg) zeigte exemplarisch das Vorgehen zur Erhaltung software-basierter Kunst auf: Zunächst werden Schnittstellen identifiziert und externe Abhängigkeiten ermittelt. Im Anschluss soll

die Zahl der Schnittstellen und Komponenten reduziert werden. Für das digitale Artefakt ist dann die zu emulierende Umgebung zu bestimmen. Falls sie noch nicht existieren sollte, muss sie unter Berücksichtigung der zu bedienenden Schnittstellen, der Hardware und der externen Abhängigkeiten entwickelt werden. Es leuchtet unmittelbar ein, dass sich hinter dieser abstrakten Beschreibung ein massiver Aufwand verbirgt! Dies bestätigte die Fallstudie von Philipp Tögel (Univ. der Künste Berlin) zu frühen Hypertextdokumenten von Vilém Flusser. In diesem Fall war die originale Hardware nützlich für die Analyse und den Ergebnisvergleich. Hier wurde auch deutlich, dass die Emulation selten alle Features der originalen Umgebung unterstützt. Es liegt daher in der Verantwortung der Bewertung durch die Kuratoren, welche Eigenschaften prioritär erhalten werden sollen. Daher ist es hilfreich, wenn die Kuratoren über ein gewisses technisches Verständnis verfügen.

Dragan Espenschied („Exhibiting Digital Art via Emulation – Boot-to-Emulator with the EMIL Kiosk System“) erläuterte anhand zweier Ausstellungsbeispiele die Nutzung der Emulationsplattform EMIL aus dem gleichnamigen Projekt⁹ als ausführbare und portable Version. Zentrale Voraussetzung war eine Trennung des eigentlichen Artefakts von seiner ausführbaren Umgebung. Der Beitrag von Espenschied (Rhizome¹⁰), Oleg Stobbe, Thomas Liebetaut und Klaus Rechert (alle Univ. Freiburg) wurde mit dem Best Paper Award der iPRES 2016 ausgezeichnet.

Auf einer anderen Ebene bewegte sich die Fallstudie zur Webarchäologie von Tjarda de Haan (Amsterdam Museum). Sie beschäftigt sich mit der Rekonstruktion der Website „De Digitale Stad“, die vor 22 Jahren als virtuelle Stadt aufgebaut wurde. Da

7 <https://github.com/eark-project/earkweb> [3.11.2016]

8 <http://eark-project.com/> [03.11.2016]

9 <http://www.multimedia-emulation.de/> [02.11.2016]

10 <http://rhizome.org/about/> [02.11.2016]

keine Daten vorhanden waren, starteten die Bemühungen mit einem Crowdsourcing-Aufruf, der verschiedenste Artefakte zu Tage förderte, u.a. einen eingemotteten Server sowie Magnetbänder mit einer frühen Momentaufnahme, die mit der hell-sichtigen Absicht angefertigt worden waren, zukünftige Archäologinnen damit arbeiten zu lassen! Neben den technischen Hürden stellen sich weitere Fragen: Wie kann das Medium präsentiert werden, wie können Anforderungen des Datenschutzes erfüllt werden?

Kunst und Forschung – wer entscheidet?

Die Vorträge zu digitaler Kunst boten Gelegenheit, über das Verhältnis des Datenproduzenten zu seinen Daten nachzudenken: Künstlerinnen sehen ihr digitales Kunstwerk nicht unbedingt als statisches Produkt an, das in genau dieser Form erhalten bleiben muss. Solange sie aktiv und an einem bestimmten Werk interessiert sind, entstehen vielmehr neue Versionen und Updates. Dabei hat der Künstler oft nicht alle Elemente selber unter Kontrolle: Verwendete APIs und Programme erfahren Updates, die die Funktion beeinträchtigen etc. Andere Künstler möchten ihr Kunstwerk in Bezug auf das „Look & Feel“ hin und wieder modernisieren. Und nicht zuletzt gibt es auch Netz-Künstlerinnen, die es durchaus in Ordnung finden, wenn ihr Werk aufgrund des veränderten Umfelds „stirbt“.

Es drängt sich ein Vergleich zu Forschungsdaten auf, deren Produzenten ähnliche Standpunkte vertreten: Die bereits publizierte Version eines Codes soll beispielsweise ohne weiteres ersetzt werden, wenn es eine neue gibt, während der Sinn einer dauerhaften Aufbewahrung in ihren Augen nicht gegeben ist.

Aber ist es auch wirklich richtig, dass Produzent / Künstlerin / Forschende abschließend über die Aufbewahrungsdauer entscheiden? Wer soll bewerten, was wie lange archiviert wird? So wie die Kunst-Kuratoren

Sammlungskriterien definieren, müssen Datenkuratoren Leitlinien für die Bewertung und die Aufbewahrungsdauer von Forschungsdaten finden und vertreten.

Erschließung und Automatisierung

Angesichts der Bedeutung der Kontextinformation stellt sich die Frage, wie diese effizient gewonnen werden kann. Einiges deutet darauf hin, dass digitale Archivarinnen immer weniger (digitale) Akten in aufwändiger Arbeit mit deskriptiven Metadaten werden erschließen können, weil ihr archaisches Know-how vermehrt für die anspruchsvoller werdende Bewertung gebraucht wird. Automatisierte Verfahren können helfen, nicht nur wie bisher vorhandene technische und deskriptive Metadaten und Strukturen zu extrahieren, sondern verstärkt Metadaten auf Basis der Inhalte zu erzeugen.

Greg Jansen und Richard Marciano (Digital Curation Innovation Center, Univ. of Maryland) stellten das Projekt „Brown Dog“¹¹ mit ersten Ergebnissen vor. Die Grundannahme des Projekts ist, dass Forschungsdaten in großer Menge und unstrukturiert vorliegen und nur nachnutzbar gemacht werden können, indem automatisierte Tools zur Extraktion von Eigenschaften und Daten eingesetzt werden. Brown Dog soll solche Tools sammeln und sie als Dienste zur Verfügung stellen. Ziel ist also nicht eine umfassende Anwendung, die „alles kann“, sondern eine Umgebung, die die Nutzung aller verfügbaren Hilfsmittel (Software, Werkzeuge, Dienste) unterstützt. Als Beispiel wurde das bereits bekannte Siegfried-Tool¹² erwähnt.

Andreas Rauber (TU Wien) präsentierte Ergebnisse der Data Citation Working Group der Research Data Alliance zur Zitierung dynamischer

Forschungsdaten¹³. Vorgeschlagen wird die Verwendung des Elements der „Query“ als genaue technische Beschreibung eines Datensatzes. Der Ausführungszeitpunkt der Abfrage wird als Zeitstempel verwendet und die so datierte Abfrage mit einer persistenten Identifikation (PID) versehen. Die Query ist somit ein beschreibendes Element sowie in Kombination mit dem Zeitstempel ein eindeutiger Schlüssel zu einem bestimmten Datensatz. Voraussetzung für die Anwendung einer Query-PID ist die langfristige Verfügbarkeit der Daten.

Heinz Werner Kramski (Deutsches Literaturarchiv Marbach) und Jürgen Enge (Hochschule für Gestaltung und Kunst Basel) zeigten die Arbeit am Nachlass von Friedrich Kittler mit funktionierenden Computern und Festplatten. Kittlers Programmier- und Experimentierfreude führte dazu, dass sogar die sichtbare Abnutzung des „Reset“-Knopfes des Computers als signifikante Eigenschaft zu erhalten wäre.

Verschiedene Tools kamen zum Einsatz, teilweise vor Startversuchen der Computer, da die Routineprüfungen beim Aufstarten alte Festplatten zerstören können. Anhand des Abgleichs von Prüfsummen mit der National Software Reference Library (NSRL)¹⁴ wurden ca. ein Drittel der Dateien als Teil von Software und Betriebssystem identifiziert, die nicht zu Kittlers eigenen Inhalten zählten.

Als Werkzeug dürfte insbesondere der „Indexer“ von Interesse sein (alias „Intelligent Read-Only Media Identification Engine“ oder „Intelligent Recursive Online Metadata and Indexing Engine“ oder „Ironmaidens“). Die Autoren hoben hervor, dass die Werkzeuge am wirkungsvollsten mit Wissen zum Sammlungskontext und zur intellektuellen Bearbeitung eingesetzt werden. Dies ist ein weiterer Hinweis darauf, dass archivarische

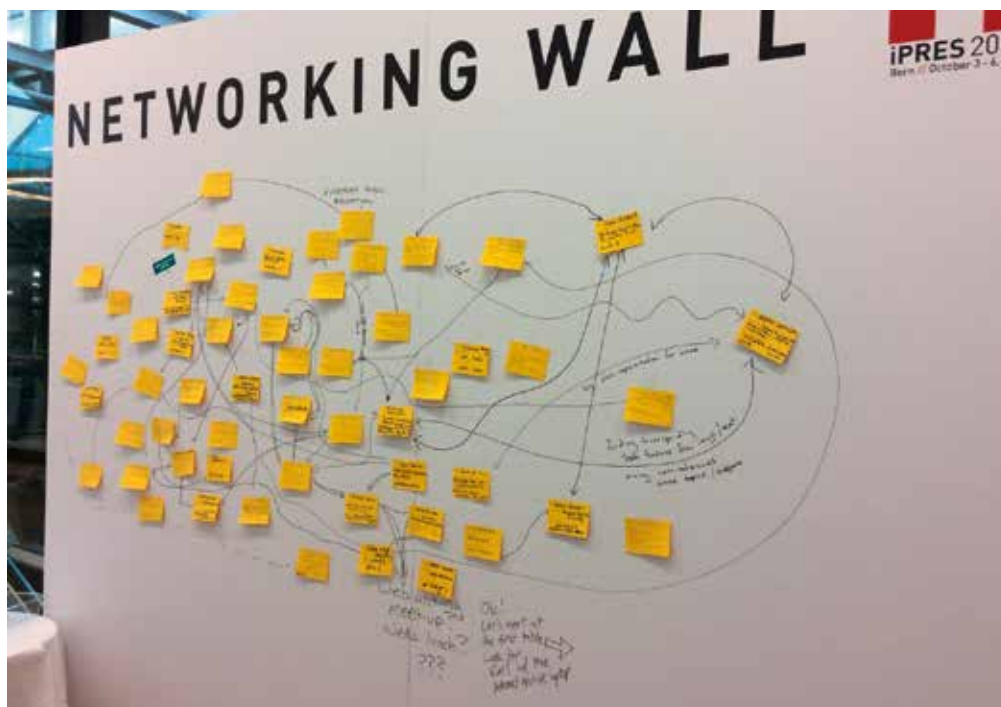
¹¹ <http://browndog.ncsa.illinois.edu/#home> [02.11.2016]

¹² <http://openpreservation.org/blog/2014/09/27/siegfried-pronom-based-file-format-identification-tool/> [02.11.2016]

¹³ <http://dx.doi.org/10.15497/RDA00016> [02.11.2016]

¹⁴ <http://www.nsrll.nist.gov/> [07.11.2016]

Abbildung 2:
Momentaufnahme der
„Networking
Wall“ an der
iPRES 2016
(Foto: M.
Töwe)



Kompetenzen trotz Teil-Automatisierung eher an Bedeutung gewinnen.

Was passiert, wenn...

Michelle Lindlar (Technische Informationsbibliothek) sprach ein unangenehmes Thema an: Welche Risiken sind zu bedenken, wenn ein Partner die Mitgliedschaft in einem Konsortium für die Langzeitarchivierung aufkündigt? Trotz funktionierender Zusammenarbeit kann sich zeigen, dass die Rollen im Konsortium nicht klar genug geregelt sind. Eine Empfehlung lautet daher zu prüfen, welche Rechtsinformationen sowohl im Langzeitarchiv als auch außerhalb des Systems in den Partnerinstitutionen gesichert werden müssen. Nützlich für die Klärung von Kostenfragen waren die Arbeiten des Projekts „4C Collaboration to Clarify the Costs of Curation“¹⁵.

Postersession

Die Postersession wurde mit „Minute-Madness“ Präsentationen der Posterinhalte eingeführt. Da sich die „Madness“ in Grenzen hielt, kam die bereitgestellte Kuhglocke kaum zum Einsatz. Die Bandbreite der Poster reichte von einer Kampagne der

National and State Libraries of Australasia zur Sensibilisierung der Öffentlichkeit für Fragen der digitalen Archivierung über Zertifizierung und Geodaten bis zur Bereitstellung von Emulationsumgebungen für Multimediaobjekte. Über den von nestor gestifteten Preis für das beste Poster¹⁶ konnten sich Franziska Geisser, Roland Suri und Matthias Töwe von der ETH-Bibliothek freuen, die Mitautoren dieses Berichts sind. Sie diskutierten Probleme bei der Identifizierung und Validierung von Formaten und wie vorläufig mit diesen umgegangen werden kann.

Workshops und Tutorials

An drei halben Tagen wurden 18 Workshops und Tutorials durchgeführt.

Das Verhältnis eines Produzenten zu seinen Daten wurde im Rahmen des Workshops „Personal Digital Archives“ (PDA) wieder angesprochen. In den USA wird seit ca. 2009 relativ intensiv zu diesem Thema publiziert. Die Library of Congress pflegt mehrere Websites¹⁷ zum Thema. In unserem Sprachraum gibt es seit die-

sem Jahr eine nestor-Arbeitsgruppe PDA¹⁸.

Während wir bei PDA an persönliche Fotosammlungen denken, die niemand bewerten oder mit Metadaten beschreiben kann, schlug Helen R. Tibbo (Univ. of North Carolina at Chapel Hill) den Bogen zu den Forschungsdaten: „Research Data can be personal, too! Owned / controlled by the researcher, often lovingly collected and crafted, often kept on laptop or phone.“

In der Tat kann die „Unternehmenskultur“ an Hochschulen dazu beitragen, dass Professoren ihre Unterlagen und Daten als persönliches Archiv betrachten. Insofern müssen Archivarinnen sie auch persönlich dazu motivieren, Ressourcen in eine sinnvolle Archivierung ihres Schatzes zu stecken. Aber wie lösen wir das Bewertungsproblem?

PREMIS 3.0

Ziel des Tutorials „Understanding and Implementing PREMIS“ war die Vermittlung von grundlegenden Kenntnissen zu PREMIS und zu Digital Preservation Metadata. Angela Dappert (British Library) gab zudem

¹⁵ <http://4cproject.eu/> [07.11.2016]

¹⁶ <http://dx.doi.org/10.3929/ethz-a-010735897> [17.11.2016]

¹⁷ <http://www.digitalpreservation.gov/personalarchiving/> [02.11.2016]

¹⁸ <https://wiki.dnb.de/display/NESTOR/AG+Personal+Digital+Archiving> [15.11.2016]

einen Überblick über wichtige Neuerungen in PREMIS 3.0. Dazu gehören die Behandlung der „Intellectual Entity“ (IE) als „Object“ mit Verbindung zu „Events“ und „Rights“ sowie mit der Möglichkeit, Kollektionen abzubilden und Versionierungsinformationen zu erfassen.

Neu werden so genannte „Environments“ als „Objects“ eingeführt. Ein „Environment“ kann Hard- oder Software sein – alles, was zur Wiedergabe oder Nutzung eines „Object“ benötigt wird. Ebenfalls ergänzt wurden „Physical Objects“, um den gesamten Lebenszyklus einschließlich eines physischen Originals abzudecken.

Dateiformat – das unbekannte Wesen, Teil II

Der Workshop „Sharing, Using and Re-using Format Assessments“ vertiefte Fragen zur Formatidentifizierung und -validierung. Wird ein Format wirklich durch eine (offizielle) Spezifikation definiert oder letztlich durch die Anwendung, die eine Datei erzeugt und wiedergeben kann? Die (Nicht-)Beantwortung dieser Frage hat Konsequenzen: Wenn die Spezifikation nicht eindeutig ist oder nicht sauber implementiert wird, werden Viewer mit der Zeit so „tolerant“ gemacht, dass sie die Mehrzahl der mit unterschiedlichen Tools erzeugten Dateien eines Formats wiedergeben können. Für die Langzeitarchivierung schaffen die „tolerierten“ Ungereimtheiten Probleme.

Als herausforderndste Formate wurden genannt: 3D-Objekte und Modelle, bitcoin (!), Forschungsdaten, PDF, proprietäre Formate. Vollständig obsolete Formate wurden in den Institutionen noch nicht angetroffen. Angeregt wurde eine gewisse Zentralisierung des Know-hows zur Formatbewertung, etwa zur gemeinsamen Pflege der verfügbaren Werkzeuge. Die verschiedenen Format Registries sollten zusammengeführt werden und es wäre gut zu wissen, welche Institution sich mit welchen Fragen und Formaten beschäftigt.

Sinnvoll wäre auch eine konkrete Information zu den Eigenschaften und zur Nutzung von Dateiformaten für die breitere Öffentlichkeit.

Weitere Workshops und Tutorials widmeten sich spezifischen Tools wie Fedora 4 als Plattform für Repositorien, Bitcurator für native digitale Offline-Inhalte oder SIARD 2 zur Überführung von Datenbanken ins SIARD-Format sowie den Validatoren veraPDF, MediaConch und DPF Manager aus dem PREFORMA-Projekt¹⁹. Daneben wurde in einem Workshop auch die PREMIS-Implementierung behandelt. Hinzu kamen Persistent Identifier, Speicherinfrastruktur, Geschäftsmodelle und Finanzierung, Forschungsdatenmanagement, neue Methoden zur Webarchivierung und ein Projektbericht von UNESCO-PERSIST (siehe oben) zu den Bemühungen, Softwareanbieter von der Bereitstellung überholter Software zu überzeugen.

Networking – analog und musikalisch

Weitere Veranstaltungen rundeten das Tagungsprogramm ab. Am Eröffnungabend empfing die NB die Gäste in ihren eigenen Räumen mit den Alphornbläsern des Swiss Alphorn Trio. Man muss nicht einmal Freund alpiner Folklore sein, um sich von diesen nicht alltäglichen Klängen berühren zu lassen. Etwas andere Klänge entlockte Mario Batkovic am Konferenzdinner in der Halle des historischen Berner Rathauses seinem Akkordeon und stellte so unter Beweis, was mit diesem Instrument möglich ist. Ein Grußwort von Isabelle Chasot, der Direktorin des Bundesamtes für Kultur, leitete den Abend ein, an dem auch die Preise für „Best Paper“ und „Best Poster“ überreicht wurden.

Ein Experiment war der Einsatz einer „Networking Wall“ während der Tagung, auf der Teilnehmerinnen ihre Kenntnisse vermerken und im Gegenzug gesuchte Kontakte und

Know-how notieren konnten (Abb. 2). Nach zaghaften Anfängen kamen zunehmend Querverbindungen zustande und das eine oder andere ad hoc-Treffen beim Mittagessen wurde auf diesem analogen sozialen Netzwerk verabredet.

Den Kolleginnen und Kollegen der NB gebührt der Dank der hochzufriedenen Gäste für eine erfolgreiche und hervorragend organisierte Konferenz in Bern, an deren Ende sich die Blicke bereits nach Kyoto richteten, wo die iPRES 2017 Station machen wird. ■



Angela Gastl

Data Archivist
ETH-Bibliothek, Digitaler
Datenerhalt
angela.gastl@library.ethz.ch



Dr. Franziska Geisser

Data Archivist
ETH-Bibliothek, Digitaler
Datenerhalt
franziska.geisser@library.ethz.ch



Dr. Ana Sesartić

Datenmanagementplanung
und Schulungen
ETH-Bibliothek, Digitaler
Datenerhalt
ana.sesartic@library.ethz.ch



Dr. Roland Suri

Fachreferat Maschinenbau
ETH-Bibliothek
roland.suri@library.ethz.ch



Dr. Matthias Töwe

Leitung Digitaler
Datenerhalt
ETH-Bibliothek
matthias.toewe@library.ethz.ch

Korrespondenzadresse:

ETH Zürich
ETH-Bibliothek
Dr. Matthias Töwe
Rämistrasse 101
CH-8092 Zürich
matthias.toewe@library.ethz.ch

¹⁹ <http://www.preforma-project.eu/> [15.11.2016]