

Protokoll: 62. Sitzung der FAG TI

Ort: Göttingen

Datum: 09.03.2017

Teilnehmer: Herr Bollin, Dr. Blenkle, Dr. Busse, Herr Farrenkopf, Herr Hahn, Herr Leydolph, Dr. Lindemann, Herr Schmitz, Herr Tuschick, Herr v.d. Ahe

Entschuldigt: Herr Nord

VZG-Vertreter: Herr Diedrichs, Herr Dührkohp

Gäste: Herr Bähr, Frau Lindlar

Protokoll: K. Leydolph

Inhalt

Tagesordnung.....	1
TOP 1: Genehmigung des Protokolls, Abstimmung des Folgetermins	1
TOP 2: Genehmigung der Tagesordnung	1
TOP 3: Langzeitarchivierung als Service der VZG, Herr Dührkohp.....	2
TOP 5: Angebot LZA der TIB, Herr Bähr / Frau Lindlar	3
TOP 6: AG Discovery, konstituierende Sitzung, Dr. Blenkle	4
TOP 7: Kurzbericht aus der Verbundzentrale, Herr Diedrichs	4
Top 8: Verschiedenes	4

Tagesordnung

1. Genehmigung des Protokolls, Abstimmung des Folgetermins
2. Genehmigung der Tagesordnung
3. Langzeitarchivierung als Service der VZG, Herr Dührkohp
4. Mittag
5. Angebot LZA der TIB, Herr Bähr
6. AG Discovery, konstituierende Sitzung, Dr. Blenkle
7. Kurzbericht aus der Verbundzentrale, Herr Diedrichs
8. Verschiedenes

TOP 1: Genehmigung des Protokolls, Abstimmung des Folgetermins

- Keine Anmerkungen zum alten Protokoll
- Folgetermin: Di 20.06.2017

TOP 2: Genehmigung der Tagesordnung

- Tagesordnung wurde ohne Änderungen akzeptiert

TOP 3: Langzeitarchivierung als Service der VZG, Herr Dührkohp

- Durch die VZG wurden drei am Markt vertretene LZA-Lösungen evaluiert. Die Basis der Evaluation war, das Archive als „Dark Archive“ zu nutzen. In den Projekten sind sehr unterschiedliche Materialien zu berücksichtigen. Der Anspruch an die eigene Anwendung ist, diese langfristig anzubieten. Die Entwicklung von Exit-Strategien wird als wichtiger Bestandteil der Realisierung gesehen.
- Evaluiert wurden: Aredo (DNB), Digitales Archive NRW (Landschaftsverband Rheinland) und Rosetta der TIB
- Aredo, der Langzeitarchivierungsdienst der DNB, wurde mit dem Projekt „Deposit Hagen“ der FernUni Hagen getestet. Die Realisierung auf dieser Basis kommt aber zurzeit nicht Frage. Die Kosten belaufen sich auf 1,69 € pro GB und Jahr.
- DA NRW, eine Open Source Software Lösung der Uni zu Köln, wurde mit dem technischen Partner LVR-Infokom, getestet. Die Evaluierung ergab, dass 3 Entwickler für drei Jahre und für den Betrieb 2 Entwickler und ein Datenanalyst benötigt werden. Die Entwicklungskosten liegen ca. bei 700.000,00€ und die Betriebskosten bei ca. 250.000,00€ pro Jahr. Zum Zeitpunkt der Evaluierung gab es Unklarheiten zum Betrieb, eine vertragliche Regelung war offen. Bis Ostern 2017 soll ein Geschäftsmodell vorliegen. Eine Pilotimplementierung soll außerhalb von NRW in der Uni Erfurt und der Uni Jena erfolgen.
- Das Langzeitarchiv der TIB, welches auf Basis von Rosetta von Exlibris betrieben wird, wurde als mögliche Lösung durch den Bedarf der Verbundbibliotheken in Betracht gezogen. Im November 2015 wurde mit dem BSZ und dem HBZ die Entscheidung getroffen. Für den GBV entstehen keine Einstiegskosten. Schnittstellen zu Goobi (Taskmanager) und GBV sind vorhanden. Im Angebot sind Service Level mit Workflow definiert. So enthält der Service Level 2, bei dem sich die Kosten auf 1.500,00€ pro TB und Jahr belaufen, die Technology Watch und das Preservation Management.
- Die VZG begleitet weitere Projekte in deren Rahmen LZA eine Rolle spielt. So werden auf Basis easyDB das DFG Projekt LaZAR und ein Projekt mit dem Institut für Slawistik und Kaukasiologie der Uni Jena realisiert. Hier wird eine Schnittstelle auf der Basis der easyDB API entwickelt. Die MyCoRe-Projekte OpenAgrar (Bibliotheken des BMELV) und das BMBF-Projekt zur Bereitstellung von Forschungsdaten werden durch die MyCoRe eigene SIP-Generierung Daten an Rosetta liefern. Hier ist eine eventuelle Nachnutzung für Aredo geplant.
- Durch die Anfrage der Bibliotheken wurde ein Workflow für die digitale Sicherung durch die VZG erstellt. Die geplante Realisierung wird hauptsächlich vom Preis bestimmt. Das Ziel war eine verlässliche Backuplösung zum niedrigsten Preis zu schaffen. Die Anlieferung soll per SFTP erfolgen. Nach der Validierung wird eine UUID generiert und danach auf Tape gesichert (Bitstream-Preservation). Der Dienst kostet 0,50€ pro GB (vgl. BSZ 2,00€). Die Bereitstellung als Gebrauchskopie im Präsentationserver (inkl. Katalogeintrag) soll 5,00€ pro GB kosten.
- Herr Buse weist auf die Möglichkeiten der Speicherung bei kommerziellen Storage-Anbietern hin (Amazon S3 Angebot), was durch die Staatsbibliothek zu Berlin geprüft wurde. Die Staatsbibliothek sieht in diesem Dienst eine Möglichkeit je nach Bearbeitungsstand im Digitalisierungsprozess notwendige Zwischenspeicherungen kostengünstig vornehmen zu können.
- Herr Hahn hat darauf verwiesen, das die VZG die Bibliotheken in Bezug auf LZA beraten und unterstützen muss, da in den Bibliotheken die notwendige Expertise nicht erstellt werden kann. Die Erfahrungen in der TIB machen dies auch deutlich. In der TIB ist seit 2009

erhebliches Know How im Bereich der LZA aufgebaut wurden, welches jetzt in das Angebot einfließt. Die Rolle der VZG ist die eines Partners mit Erfahrungen in verschiedenen Präsentationssystemen. In den Projekten der Bibliotheken müssen die Kosten für den Betrieb des Archives und die Beratung eingeplant werden.

- Vortrag als Anhang beigelegt

TOP 5: Angebot LZA der TIB, Herr Bähr / Frau Lindlar

- Zur Evaluierung technologischer, organisatorischer und institutioneller Anforderungen an ein kooperativ betriebenes System zur digitalen Langzeitarchivierung in der TIB wurden während einer Pilotphase 2009 – 2011 verschiedene Systeme (DuraSpace, DIAS, LOCKSS, Portico, Tessella SDB, Ex Libris Rosetta) analysiert und ein Testsystem auf Basis von Ex Libris Rosetta implementiert.
- Der Produktivbetrieb startete 2012. Hierbei wird Rosetta als Systemkern des Goportis Langzeitarchivs eingesetzt. Das System wird von der TIB gehostet und beinhaltet aktuell ca. 20 Millionen Dateien (ca. 74 TB).
- Im kooperativen Betrieb des Langzeitarchivs im Verbund der drei zentralen Fachbibliotheken fungiert die TIB als Host, sie stellt Hardware, Software und Infrastruktur zur Verfügung. Es wird eine gemeinsame Strategie zur digitalen Langzeitarchivierung verfolgt. Mit den Partnern werden Dienstleistungsverträge abgeschlossen, personelle Ressourcen werden geteilt. Es wird eine gemeinsame Policy entwickelt und eine gemeinsame Zertifizierung angestrebt.
- Das Dienstleistungsangebot beinhaltet die Bereitstellung von Infrastruktur und Hardware, inklusive Konfiguration, Monitoring und Backup. Weiterhin gehören die Bereitstellung von Software inklusive kundenspezifischer Konfiguration, Qualitätskontrollen und optional die DOI-Vergabe zum Dienstleistungsangebot. Entsprechend des vereinbarten Servicelevel werden folgende Leistungen angeboten.

Level 1

- Beratung und Konzeptentwicklung zur Erstellung der SIP
- Pre-Ingest (Analyse und Strukturierung der Objekte und Metadaten)
- Ingest und AIP-Erstellung gemäß Konzept
- Bitstream-Preservation
- Erstellung von Berichten und Statistiken
- Auslieferung von AIP an den Kunden (trigger event)

Level 2 (wie Level 1 plus)

- Technologie Watch
- Preservation Management

Level 3 (wie Level 2 plus)

- Access (permanenter Zugriff auf DIP)

- Es werden textuelle und nichttextuelle Materialien (Bild, AV, 3D) in den entsprechenden Standardformaten archiviert. Derzeit werden keine nichttextuellen Forschungsdaten verarbeitet.
- Die Kosten belaufen sich netto auf 1.000,00€ bei Level 1, 1.500,00€ bei Level 2 und 2.000,00€ bei Level 3. Sie beinhalten die Kosten für Hardware, Infrastruktur, Rosetta-Lizenz und level-abhängig die Kosten für Preservation Management. Kosten für Beratung, Analyse, Datenaufbereitung, DOI-Vergabe und Preservation Action werden zusätzlich bzw. einmalig erhoben.

- Kostenreduzierungen sind möglich. So ist denkbar, dass die zu archivierenden Daten auf Speichersystemen (die gewisse Anforderungen erfüllen müssen) des Auftraggebers verbleiben und nach Verarbeitung in der TIB nur die Metadaten in der LZA gespeichert werden.
- Die Light-Version muss individuell verhandelt werden. Der Preis ist vom jeweiligen Aufwand abhängig. Wichtig ist auch hierbei, dass die Integrität der Daten gesichert werden muss.
- Einbindung von Cloud-Storage angeboten in die Lösung der TIB wäre aus Sicht der Teilnehmer wünschenswert.
- Eine Kurzbeschreibung des Dienstleistungsangebotes und die Vorträge sind als Anlage beigefügt.

TOP 6: AG Discovery, konstituierende Sitzung, Dr. Blenkle

- Am 26.01.2017 war die konstituierende Sitzung der AG. Das Protokoll befindet sich im GBV-Wiki unter GBV - Arbeitsgruppen (diverse) – AG Discovery.
- Die Teilnehmer verfügen über ein breites Spektrum an Erfahrungen mit unterschiedlichen Discovery-Lösungen. Als zentrales Thema der AG zeichnet sich der Ausbau des GBV-Zentral als offener Metadaten-Index und als Alternative zu kommerziellen Systemen ab.
- Bibliotheken sollen bei Lizenzierung die Verwendung der Metadaten und eventuell der Volltexte mit verhandeln. Ziel sollte es sein, dass diese im GBV-Zentral verwendet werden dürfen. Hierzu ist eine Normierung der Daten notwendig, was durch die Bibliotheken unterstützt werden soll (XML-Schema wird von der VZG vorgegeben).
- Die AG will sich auch mit dem Thema Metadatenqualität beschäftigen und Anforderungen an Metadaten beschreiben. Das nächste Treffen ist für Mai 2017 geplant.

TOP 7: Kurzbericht aus der Verbundzentrale, Herr Diedrichs

- Die heiße Phase der Katalogzusammenführung hat begonnen. Die Katalogisierer beider Verbünde treffen sich um die Formatzusammenlegung zu diskutieren. Die Vorgaben der lokalen Bibliothekssysteme und der Fernleihe müssen berücksichtigt werden. Ein erster Entwurf für ein gemeinsames Format soll Mitte März 2017 vorliegen. Danach beginnen die Tests mit einer Test-DB. Für das 4. Quartal ist ein erster Vollabzug geplant. Ab 1. Quartal 2018 sollen dann Schulungen beginnen.
- Aufbau einer gemeinsamen Datenbank bedeutet auch Dublettenbeseitigung und Suchschlüsselanpassung, welche noch diskutiert wird.
- Für die Updates der lokalen Bibliothekssysteme ist eine Schließzeit der Katalogisierung von ein bis zwei Tagen einzuplanen.

Top 8: Verschiedenes

- Informationen von Herrn Diedrich:
 - WIN 10 kann verwendet werden, WinIBW 3.7 läuft darunter
 - Neuer webbasierter Katalogisierungs-Klient
 - CC Web (Frankreich im Sommer)
 - Herbst marc21/PICA3

- Funktion geringer als nativer Client
- WinIBW 4 wird entwickelt