

# ■ UMGANG MIT METADATEN IN REPOSITORIEN – EINE ÖSTERREICHWEITE UMFRAGE

von *Susanne Blumesberger* & *Alexander Zartl*

**Zusammenfassung:** Dieser Artikel enthält eine Auswertung der Telefonumfrage zum Thema „Metadaten“, die im Sommer 2015 unter den am Projekt e-Infrastructures Austria teilnehmenden Institutionen durchgeführt wurde. Ziel der Befragung war eine Bestandsaufnahme der an den einzelnen Universitäten im Einsatz befindlichen Repositorien, um einen Überblick über die wichtigsten Softwaresysteme und deren spezifische Einsatzzwecke, die am häufigsten verwendeten Metadatenstandards und die Gepflogenheiten bzw. Erwartungen der Nutzer und Nutzerinnen zu gewinnen. Zukünftige Planungen und Entwicklungen waren ebenfalls Gegenstand des Fragebogens.

**Schlüsselwörter:** Metadaten; Repositorien; Österreich

## HANDLING METADATA IN REPOSITORIES – AN AUSTRIA-WIDE SURVEY

**Abstract:** This article discusses the results of a telephone survey on the subject of metadata that was carried out in summer 2015 among the institutions participating in the project e-Infrastructures Austria. The aim of the survey was to provide an overview of the most important software systems and their specific uses, the most commonly used metadata standards and the habits and expectations of the users. Future plannings and developments were also included in the questionnaire.

**Keywords:** Metadata; repositories; Austria



Dieses Werk ist lizenziert unter einer

[Creative-Commons-Lizenz Namensnennung 4.0 International](https://creativecommons.org/licenses/by/4.0/)

### **Inhalt**

1. Vorbemerkung
2. Liste der befragten Institutionen
3. Ergebnisse

## 1. Vorbemerkung

Ziel der Befragung war, festzustellen, wie man möglichst rasch und effizient herausfinden könnte, wie die Partnerinstitutionen am Projekt *e-Infrastructures Austria* derzeit mit Metadaten in ihren Repositorien umgehen, bzw. wie ihre Planung diesbezüglich aussieht. Da zum Zeitpunkt der Fragebogenerstellung (April 2015) noch nicht alle Partnerinstitutionen über ein Repositoryum verfügten und somit auch keine Erfahrungen mit Metadaten in diesem Bereich hatten, musste die Befragung ziemlich niederschwellig angesetzt werden. Die Gruppe einigte sich auf einen schriftlichen Fragenbogen, der zwar an die AnsprechpartnerInnen geschickt werden sollte, von ihnen jedoch nicht alleine ausgefüllt, sondern in Form eines angekündigten telefonischen Gesprächs abgehandelt werden sollte. Der Vorteil dabei ist, dass im persönlichen Gespräch nachgehakt werden kann und sollte etwas nicht oder falsch verstanden worden sein, die Möglichkeit zur Nachfrage gegeben ist.

In einem ersten Schritt wurden die Fragen überlegt und möglichst klar und deutlich formuliert. Danach wurden die jeweiligen AnsprechpartnerInnen an den einzelnen Institutionen gesucht. In einer vorbereiteten Mail wurde um einen Gesprächstermin ersucht, der Fragebogen wurde bereits beim ersten Kontakt mitgeschickt, um den GesprächspartnerInnen die Möglichkeit zu geben, sich vorab Antworten überlegen zu können, bzw. auch um ihnen die Gelegenheit zu bieten, weitere ExpertInnen an der eigenen Institution in die Befragung miteinzubeziehen.

Während des telefonischen Interviews wurde das Gespräch in Stichworten niedergeschrieben und gleich danach ergänzt. Die personalisierten Fragebögen wurden archiviert, aber – so ist es mit den InterviewpartnerInnen abgesprochen – nicht an die Öffentlichkeit weitergegeben, sondern nur anonym weiter verarbeitet.

Zu den Umfrageergebnissen ist zu ergänzen, dass die Ergebnisse natürlich nur eine Momentaufnahme sein können. Es zeigt sich jedoch bereits deutlich, dass der Umgang mit Metadaten in einem Repositoryum noch keineswegs selbstverständlich bzw. erprobt ist. Für viele Institutionen sind Repositorien eine neue Aufgabe. Metadaten sind zwar aus dem bibliothekarischen Alltag bekannt, die Umsetzung dieses Wissens auf die Beschreibung digitaler Objekte ist jedoch noch relativ ungewohnt und steht ziemlich am Anfang. Hier ist nicht nur technisches Know-How wichtig, sondern vor allem auch das Bewusstsein, wie wichtig Metadaten für ForscherInnen bzw. für künftige NutzerInnen der Objekte sind.

Eine Anmerkung zu den gewählten Diagrammtypen: Balkendiagramme kommen stets dort zum Einsatz, wo Mehrfachantworten möglich waren

und die Summe der Prozentzahlen daher größer als 100 sein kann. Entscheidungsfragen mit nur einer Antwortmöglichkeit werden durch Kreisdiagramme veranschaulicht.

Wir bedanken uns bei allen GesprächspartnerInnen für die freundliche Unterstützung des Projekts.

## 2. Liste der befragten Institutionen

Mit Ausnahme der Österreichischen Bibliothekenverbund und Service Ges.m.b.H., die aufgrund ihrer speziellen Aufgabenstellung mit Bezug auf den gesamten Bibliotheksbereich eine Sonderstellung einnimmt und daher nicht vergleichbar ist, konnten im Zeitraum von Juni bis September 2015 alle am Projekt *e-Infrastructures Austria* teilnehmenden Organisationen in die Befragung einbezogen werden.

- Akademie der bildenden Künste Wien
- Arbeiterkammer Wien
- Institute of Science and Technology Austria (IST)
- Medizinische Universität Graz
- Medizinische Universität Wien
- Montanuniversität Leoben
- Österreichische Akademie der Wissenschaften
- Österreichische Nationalbibliothek
- Technische Universität Graz
- Technische Universität Wien
- Universität für angewandte Kunst Wien
- Universität für Bodenkultur Wien
- Universität für künstlerische und industrielle Gestaltung Linz
- Universität für Musik und darstellende Kunst Graz
- Universität für Musik und darstellende Kunst Wien
- Universität Graz
- Universität Innsbruck
- Universität Klagenfurt
- Universität Linz
- Universität Mozarteum Salzburg
- Universität Salzburg
- Universität Wien
- Veterinärmedizinische Universität Wien
- Wirtschaftsuniversität Wien

## 3. Ergebnisse

### 3.1. Allgemeine Fragen

#### 3.1.1. Welche Arten von Daten werden an Ihrer Einrichtung gesammelt?

Bei dieser Frage gab es kein vorgegebenes Antwortvokabular, sondern es wurden die von den InterviewpartnerInnen gewählten Bezeichnungen übernommen. Daraus ergibt sich eine unscharfe Kategorienbildung mit Überlappungen zwischen den Begriffen „Bilder“ und „Grafiken“ oder „Texte“ und „Publikationen“ etc. Diese Vorgangsweise wurde gewählt, um zu verhindern, dass bestimmte Arten von Daten von vornherein ausgeschlossen werden.

Hier wird ersichtlich, dass an erster Stelle Hochschulschriften archiviert werden (sollen), in weiterer Folge Bilder, Publikationen, Texte, Videos, und an 6. Stelle Forschungsdaten.

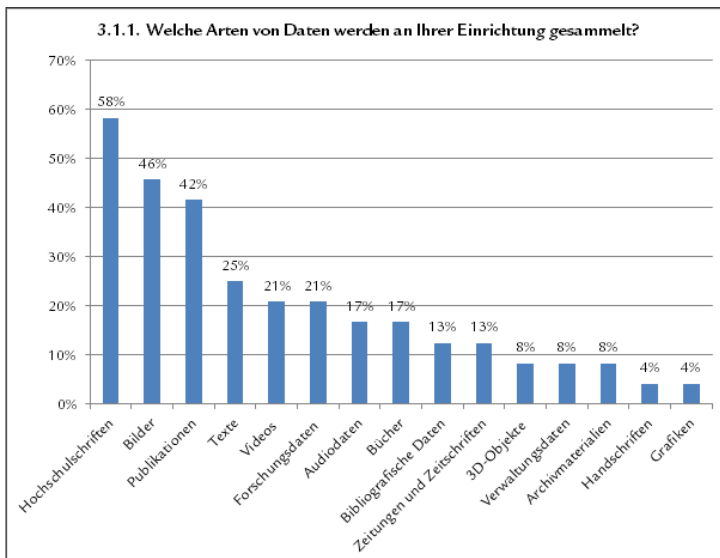
Speziell zum Begriff „Forschungsdaten“ ist anzumerken, dass diese Bezeichnung weiter ausgelegt wurde, als es die üblichen Definition – Messprotokolle, Umfragedaten, sonstige Rohdaten – nahelegt. So wurden z.B. in einem Fall explizit multimediale (Video-)Daten darunter verstanden. Insgesamt gilt damit das Prinzip, dass der Kontext bestimmt, was Forschungsdaten sind und nicht der Datentyp.

Das entspricht auch der Definition, die für die Umfrage, die sich an Forschende an der Universität Wien richtete, verwendet wurde:

„Unter dem Begriff ‚Forschungsdaten‘ sind alle Daten zu verstehen, die im Zuge wissenschaftlicher Forschungs- und künstlerischer Schaffensprozesse entstehen (z.B. Text, Tabellen, Video, Audio, Grafik etc.) und/oder auf deren Grundlage Ihre Forschungsergebnisse und/oder Kunstwerke basieren – z.B. durch Experimente, Quellenforschungen, Messungen, Erhebungen, Digitalisate oder Entwürfe.“

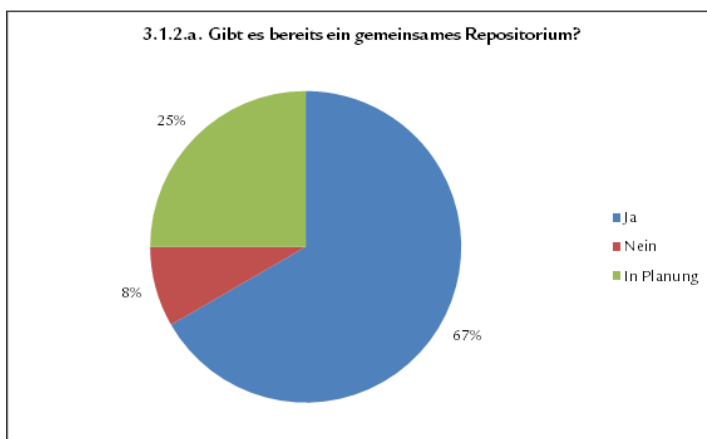
Bei der Interpretation der Grafik ist zu beachten, dass manche Einrichtungen über mehrere Repositorien verfügen und darin unterschiedliche Arten von Daten archiviert werden, wie etwa auch an der Universität Wien, wo der Hochschulschriftenserver wissenschaftliche Abschlussarbeiten enthält, das Repositorium Phaidra für alle Dateiformate offen ist und grundsätzlich alle genannten Arten von Daten archivieren kann und zusätzlich Unidam vor allem, aber nicht nur für Bildmaterial zur Verfügung steht.



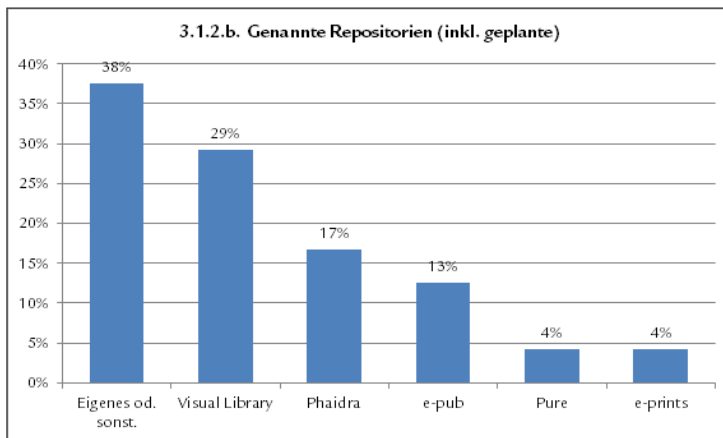


### 3.1.2. Gibt es bereits ein gemeinsames Archiv bzw. Repository für die Angehörigen Ihrer Einrichtung?

Nur zwei der TeilnehmerInnen am Projekt e-Infrastructures Austria planen zum Zeitpunkt der Befragung noch kein Repository; bei den in Planung befindlichen Repositorien war noch nicht überall bekannt, welches System gewählt werden würde.



Bei den bereits vorhandenen oder fix geplanten Systemen sah die Verteilung folgendermaßen aus:



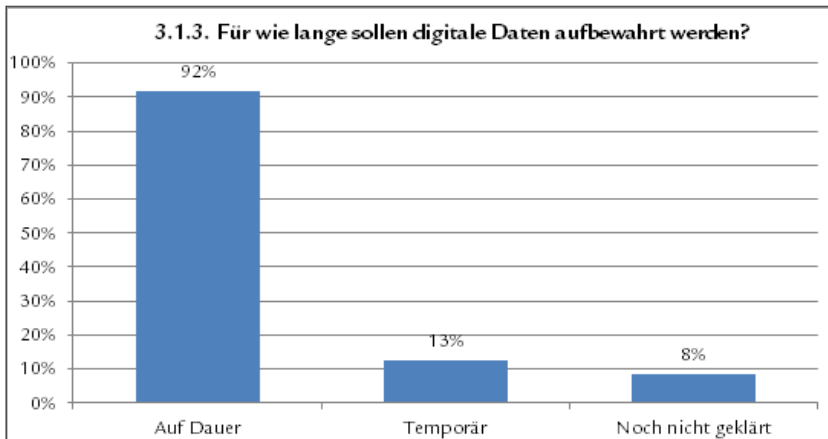
Abgesehen von den Spitzenreitern Visual Library, Phaidra und e-pub, die an mehr als zehn Prozent der Einrichtungen in Gebrauch stehen, zeigt sich eine sehr uneinheitliche Repositorienlandschaft mit einem hohen Anteil an Eigenentwicklungen.

Das verbreitetste Produkt mit einem Anteil von 29% ist Visual Library der semantics Kommunikationsmanagement GmbH. Dem Einsatz dieser Software liegt ein Konsortialmodell der Österreichischen Bibliothekenverbund- und Service-GmbH zu Grunde, sodass die nötige Infrastruktur (Hardware, Software, Datensicherung) zentralisiert ist und nicht von den Einrichtungen selbst getragen werden muss. Visual Library ist zudem in die bestehenden Workflows des Bibliothekenverbundes integriert und kann durch die Suchmaschine Primo erschlossen werden.

Ebenfalls auf Visual Library basieren die Publikationsserver BOKU:ePUB (Universität für Bodenkultur Wien), uni-pub (Universität Graz) etc. Eine Entwicklung der TU Graz ist hingegen das Portal epub.oaew der Österreichischen Akademie der Wissenschaften.

Phaidra (Akronym für: Permanent Hosting, Archiving and Indexing of Digital Resources and Assets) steht bei insgesamt 14 Einrichtungen in Österreich und dem europäischen Ausland im Einsatz. Technisch basiert das System auf der Open-Source-Software Fedora, federführend bei der Entwicklung ist die Universität Wien. Im Gegensatz zu Visual Library bilden die Phaidra-Anwender keinen technischen Verbund, sondern betreiben jeweils eigene (z.T. auch ausgelagerte) Instanzen.

### 3.1.3. Für wie lange sollen digitale Daten aufbewahrt werden? Auf Dauer oder nur temporär?

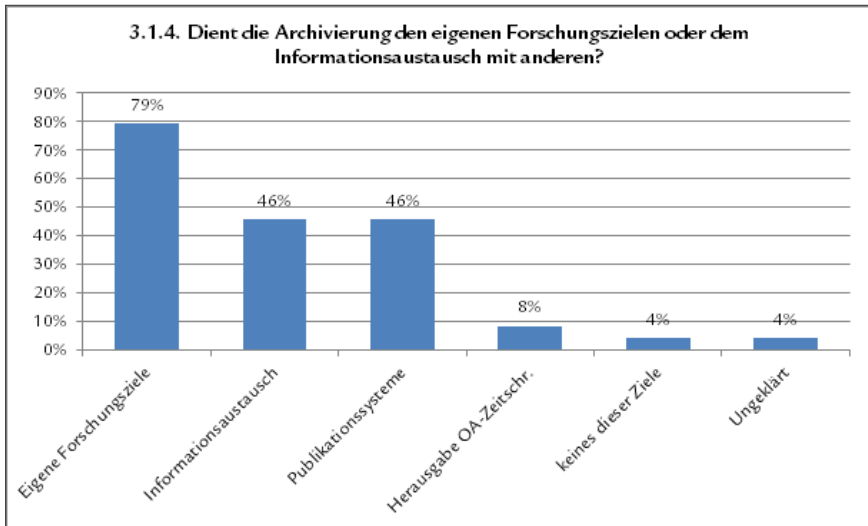


Die Grafik zeigt, dass alle befragten Institutionen, die diese Frage schon geklärt haben, ihre Daten langzeitarchivieren möchten. In drei Fällen ist zusätzlich auch die Möglichkeit der temporären Speicherung wichtig. Der Wunsch nach rein temporärer Archivierung wurde nicht genannt.

Der Wunsch nach zeitlich begrenzter Speicherung ergibt sich vorwiegend aus dem Bedarf nach einer sicheren Ablage für vorläufige Arbeitsdokumente, Rohfassungen und sonstige Zwischenergebnisse mit Ablaufdatum – nicht nur aus dem wissenschaftlichen Bereich, sondern vor allem auch in administrativem Zusammenhang. Ein Feature, welches die Entfernung nicht mehr benötigter Dokumente aus dem System grundsätzlich ermöglicht, wurde von einigen der Interviewpartner als grundlegende Funktion eingestuft.

### 3.1.4. Dient die Archivierung den eigenen Forschungszielen der Einrichtung oder dem Informationsaustausch mit anderen?

Bei dieser Frage waren mehrere Antworten möglich. Die Grafik zeigt, dass die Archivierung hauptsächlich der eigenen Forschung, bzw. der Darstellung des eigenen Forschungoutputs dient (Publikationsserver, z.T. auch mit PR- und Repräsentationsfunktion). An der zweiten Stelle steht der Informationsaustausch.



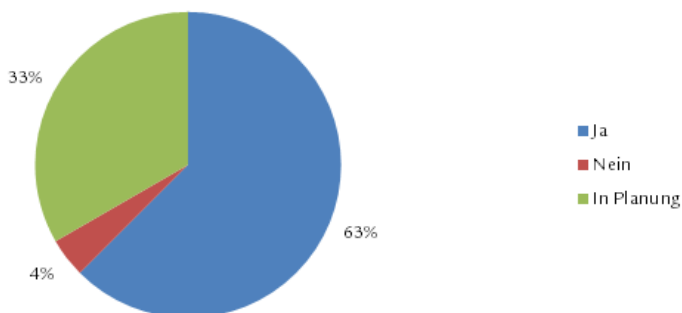
### 3.1.5. Gibt es eine Schnittstelle zwischen Bibliothekskatalog und Repository bzw. ist eine solche in Planung?

Die Schnittstelle zwischen Bibliothekskatalog und Repository ist für die meisten Partner bereits realisiert oder in Planung.

Am besten vernetzt sind die verschiedenen Repositorien nach heutigem Stand mit dem Bibliothekssystem ALEPH, was sich aus dem Anspruch auf eine Erfassung der digitalen Ressourcen in den etablierten Bibliothekskatalogen und -suchmaschinen ergibt: die zentrale Nachweisstelle soll im Normalfall der Katalog bleiben. Eine ausgereifte Lösung in diesem Zusammenhang bietet Visual Library, das einerseits durch die Suchmaschine Primo erschlossen werden kann (vgl. 3.1.2.) und andererseits die Möglichkeit bietet, die in ALEPH erzeugten Katalogdaten automatisch zu übernehmen. In Fällen, wo eine Übernahme von Metadaten aus dem Repository nach ALEPH per Datenschnittstelle bisher nicht möglich ist, spielt die manuelle Katalogisierung (mittels Eintrag in Kategorie 655) noch immer eine wichtige Rolle.

In Einzelfällen realisiert sind auch Schnittstellen zwischen lokal betriebenen Forschungsinformationssystemen und SCOPUS oder Web of Science (z.B. zur Anreicherung des Katalogisats mit Metadaten zu Zeitschriftenveröffentlichungen) oder SAP (als Unterstützung für das Rechnungswesen).

### 3.1.5. Gibt es eine Schnittstelle zwischen Bibliothekskatalog und Repository bzw. ist eine solche in Planung?



## 3.2. Fragen mit Bezug auf Metadaten

### 3.2.1. Gibt es für die Metadaten eine spezifische Policy?

Wie ersichtlich, haben erst sehr wenige Institutionen eine spezifisch ausformulierte Policy für Metadaten. Jene, die bereits eine Policy haben, verwenden i.d.R. internationale Standards.

Die University of Bath richtet sich in ihrer „Research Data Policy Guidance“ an alle Forschenden, egal ob sie intern oder extern finanziert sind und gibt folgendes vor: „Researchers must provide sufficient metadata (descriptive information) and explanatory documentation about their research data to ensure that data are discoverable, understandable and re-usable“.

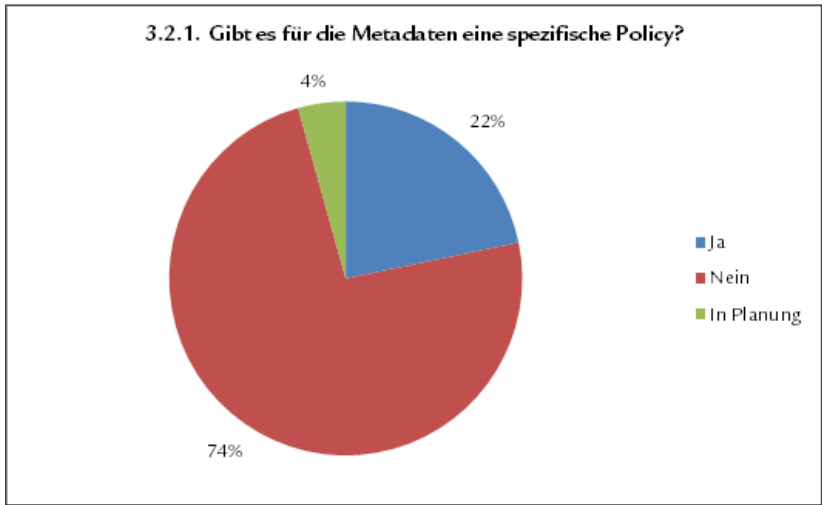
Die Universität Bath stellt folgende Forderungen:

1. Documentation and descriptive information must be recorded as early as possible after data creation.
2. Relevant metadata standards or minimal information sets should be used where applicable.
3. Descriptive information should enable other researchers to understand the research data sufficiently to be able to re-use it for any reasonably foreseeable purpose, such as those outlined in the original project proposal or impact plan.<sup>1</sup>

Europeana und z.B. die Aalto University empfehlen, für die Metadaten die Creative Common Lizenz CC0 1.0<sup>2</sup> zu vergeben.<sup>3</sup> Dies kann in manchen

Ländern, wie zum Beispiel in Österreich, aufgrund des Urheberrechtsgesetzes problematisch sein.

Insgesamt gibt es derzeit zwar wenige Policies, die sich direkt auf Metadaten beziehen, indirekt werden die Metadaten in den Policies aber oft erwähnt.<sup>4</sup>

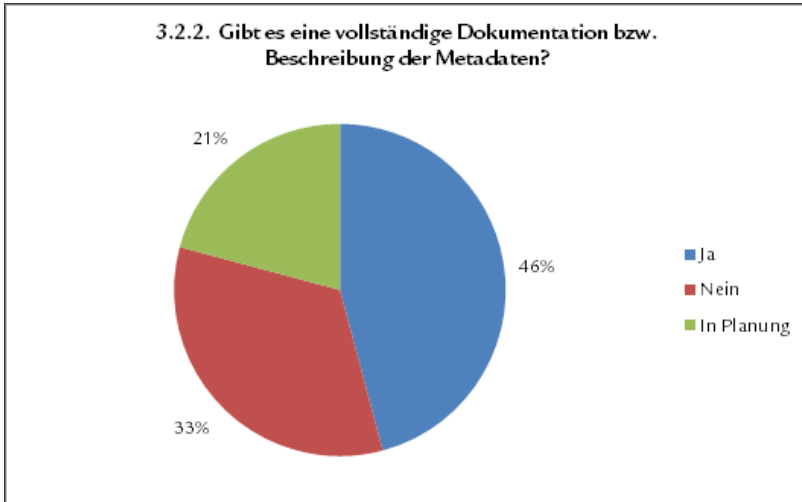


### 3.2.2. Gibt es eine vollständige Dokumentation bzw. Beschreibung der Metadaten?

Bei der Zählung wurde die Antwort „teilweise“ als „Ja“ gewertet, im Aufbau befindliche Dokumentationen wurden dem Punkt „in Planung“ zugeschlagen.

Inwieweit eine Beschreibung der Metadaten notwendig ist, hängt unter anderem davon ab, ob das Repositorium auch Metadatenfelder anbietet, die über Angaben zu AutorInnen, zum Titel, zum Datum oder über eine Beschreibung hinausgehen. In Phaidra, das ja der gesamten Universität Wien, also sehr vielen Disziplinen zur Verfügung steht, werden zum Teil sehr detaillierte Felder angeboten, die nicht selbsterklärend sind. Eine Besonderheit ist, dass man in Phaidra auch Angaben zu analogen Objekten machen kann, beispielsweise kann man die Abmessungen des Originalobjekts eintragen, was für Sammlungen bedeutend sein kann, es können aber auch zwei weitere Daten für analoge Objekte eingetragen werden. So besteht die Möglichkeit z.B. für die Entstehungszeit eines Bauwerkes ein „Datum

von“ und ein „Datum bis“ einzutragen, bzw. eine zeitliche Abdeckung einzutragen. Diese Angaben zur Provenienz, bzw. die kontextuellen Angaben beinhalten auch mehrere freie Felder, die mit forschungsrelevanten Informationen gefüllt werden können. Ohne zusätzliche Angaben würde dies die User verwirren.



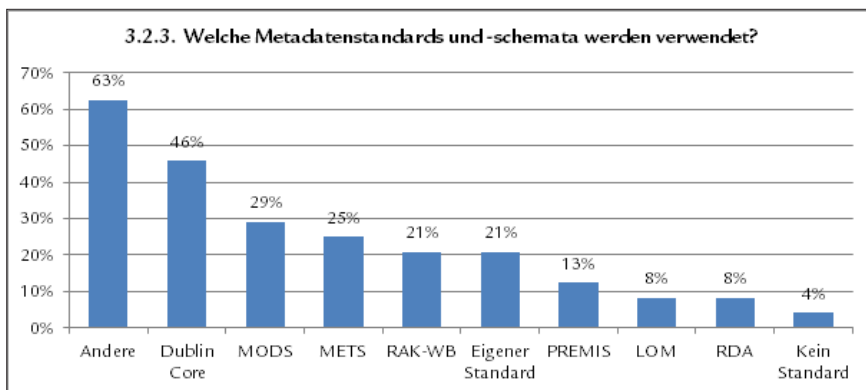
### 3.2.3. Welche Metadatenstandards und -Schemata werden innerhalb Ihrer Institution verwendet?

Hier werden nur die am häufigsten genannten Schemata eigens berücksichtigt, die übrigen werden unter „Andere“ zusammengefasst. Es gab kein vorgegebenes Antwortvokabular, was dazu führt, dass auch Metadatenstandards genannt wurden, die keine Standards für Repositorien sind. Zugelassen wurde auch die Nennung der Regelwerke RAK bzw. RDA statt der inhaltlich besser passenden Datenformate MAB oder MARC. Die Wahl der Standards hängt natürlich stark vom jeweiligen Langzeitarchivierungssystem ab. Dublin Core ist Spitzenreiter.

Kurzbeschreibungen der am häufigsten genannten Standards:

- Dublin Core: ein von auf Initiative des OCLC seit Mitte der 1990er Jahre entwickelter Standard. 15 Kernfelder (Core elements), später erweitert um sogenannte qualified terms. Es gibt keine Pflichtfelder, alle Felder sind mehrfach verwendbar. Ausführliche Dokumentation und Anwendungsbeispiele unter <http://dublincore.org/>.

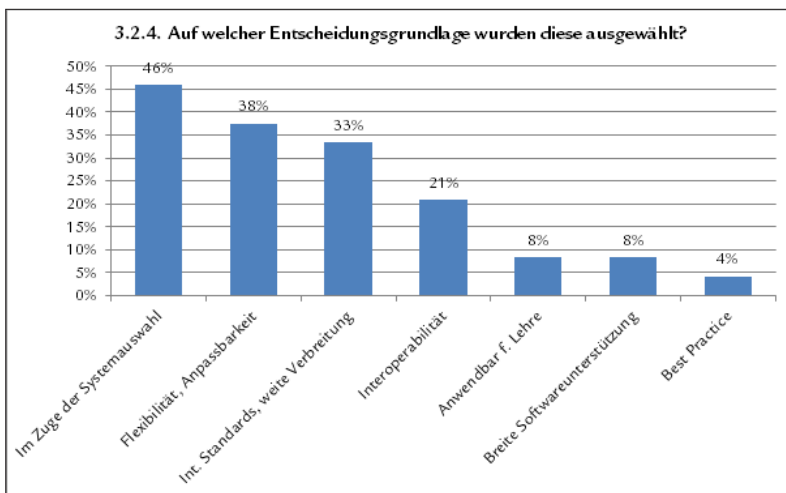
- MODS / METS (kurz für *Metadata Object Description Schema* bzw. *Metadata Encoding & Transmission Standard*): von der Library of Congress entwickelte und betreute XML-Schemata. Während MODS (teilweise in Anlehnung an MARC) für die Erfassung bibliografischer Metadaten gemacht ist, dient METS konkret zur Beschreibung von digitalen Sammlungen. Dokumentation unter <http://www.loc.gov/standards/mods/> bzw. <http://www.loc.gov/standards/mets/>.
- PREMIS (kurz für *PREservation Metadata: Implementation Strategies*): ebenfalls ein von der LoC für die strukturierte Beschreibung langzeitarchivierter digitaler Objekte. PREMIS ist somit eine mögliche Alternative zu METS; Gegenüberstellungen finden sich in der von der LoC angebotenen Dokumentation unter <http://www.loc.gov/standards/premis/>.



### 3.2.4. Auf welcher Entscheidungsgrundlage wurden die verwendeten Schemata ausgewählt?

Die Antworten waren relativ breit gefächert. An erster Stelle steht „Im Zuge der Systemauswahl“, was wohl impliziert, dass bei der Systemauswahl auch andere Kriterien als die dort verwendeten Metadatenstandards im Vordergrund standen und diese gewissermaßen „mitgekauft“ wurden. Auch auf Flexibilität und die Verwendung internationaler Standards (vgl. Frage 3.2.1.) wurde Wert gelegt.

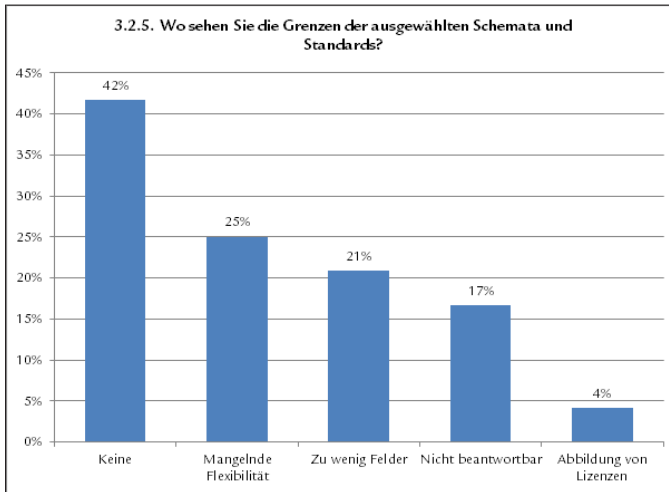




### 3.2.5. Wo sehen Sie die Grenzen der ausgewählten Schemata und Standards?

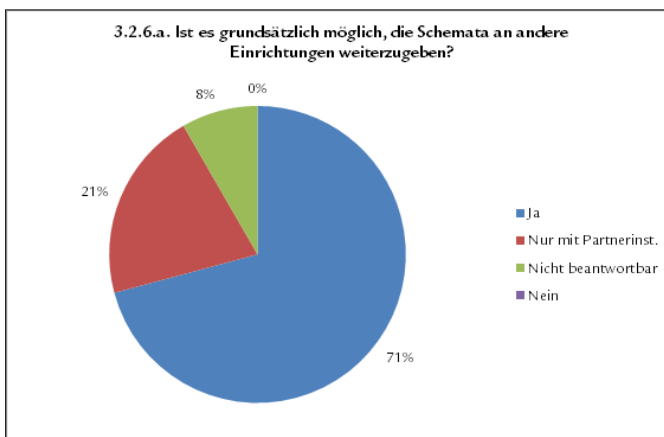
Die meisten InterviewpartnerInnen sahen (noch) keine Grenzen. Beim Beschreiben von sehr institutionsspezifischen Daten tritt gelegentlich der Wunsch nach zusätzlichen Metadatenfeldern auf. Andere Sonderwünsche wurden subsumiert unter dem Begriff „mangelnde Flexibilität“.

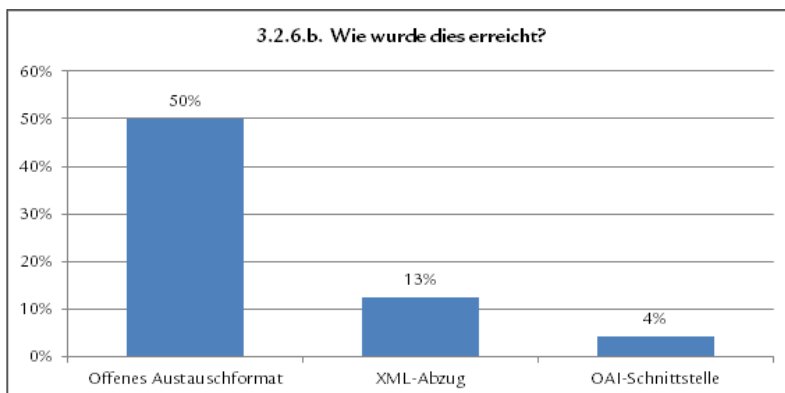
Wo die Grenzen von ausgewählten Standards und Schemata gesehen werden, hängt natürlich sehr stark vom Verwendungszweck des Repositoriums ab. Werden nur Hochschulschriften oder Zeitschriftenartikel archiviert, reichen die Angebote meist aus. Schwierig wird es erst dann, wenn man viele unterschiedliche Wünsche bedienen muss und das zu archivierende Material sehr heterogen ist. Am Beispiel der Universität Wien, wo sämtliche Nachlassmaterialien, historische Bücher, Bilder, Forschungsergebnisse in allen Formaten, audiovisuelle Materialien usw. langzeitarchiviert werden sollen, stößt man bei herkömmlichen Standards und Schemata recht schnell an die Grenzen, wenn man alle Wünsche der User erfüllen möchte. Erfahrungsgemäß genügen in den meisten Fällen eine eingehende Beratung und zum Teil ein kleiner Kompromiss, um nicht Spezialfelder anlegen zu müssen, die nicht kompatibel mit anderen Systemen sind, wodurch beim Austausch mit anderen evtl. Information verloren geht. Damit hängt auch die mangelnde Flexibilität, die von den Befragten genannt worden ist, zusammen. Je flexibler das System ist, desto schwieriger sind die Anpassung an andere Systeme und der Austausch der Daten.



### 3.2.6. Ist es grundsätzlich möglich, die Schemata an andere Einrichtungen weiterzugeben? Wie wurde dies erreicht?

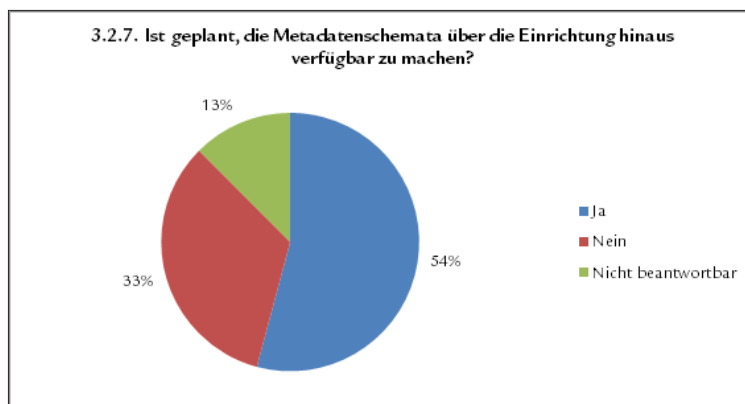
Bei dieser Frage waren sich bis auf jene Institutionen, die noch kein System haben und deshalb auch keine Antwort geben konnten, alle einig, dass die Schemata weitergegeben werden können. Abhängig vom gewählten System – wie z.B. bei Phaidra – geschieht dies durch den Austausch unter Partnerinstitutionen. Da fast alle ein offenes Format verwenden, ist eine prinzipielle Weitergabe zumeist von Vornherein möglich.





### 3.2.7. Ist geplant, die Metadatenschemata über die Einrichtung hinaus verfügbar zu machen?

Diese Frage wurde mehrheitlich mit „ja“ beantwortet, einige Institutionen hatten noch keine Entscheidung darüber getroffen. Das Open-Data-Konzept ist zurzeit noch wenig bekannt und wurde nur in wenigen Fällen explizit angesprochen.



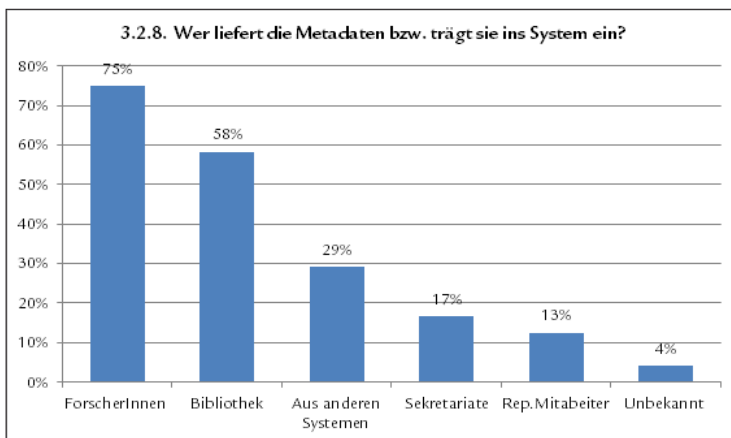
### 3.2.8. Wer trägt die Metadaten ins System ein?

Interessant ist der hohe Anteil jener Einrichtungen, bei denen die ForscherInnen und WissenschaftlerInnen die Metadaten selbst eintragen. Gleich

dahinter folgt das Modell, nach dem die Erfassung der Metadaten von der Bibliothek vorgenommen wird. In einigen Fällen werden die Metadaten (automatisiert) aus anderen Systemen übernommen, weniger oft durch die Verwaltungseinheiten auf Institutsebene (Sekretariate) oder speziell damit betraute RepositoriumsmitarbeiterInnen.

Tendenziell verhält es sich – erwartungsgemäß – so, dass Systeme, die der Forschungsdokumentation dienen, eher vom wissenschaftlichen Personal bzw. den Sekretariaten befüllt werden, während die Betreuung von Systemen im bibliothekarischen Umfeld durch BibliotheksmitarbeiterInnen erfolgt. Hier ergibt sich das Problem, dass beide Personengruppen i.d.R. nur nebenbei für die Repositorien arbeiten und keine spezielle Ausbildung dafür besitzen, was sich in z.T. wechselhafter Datenqualität niederschlägt. Der Beruf des „Content Managers“ mit einschlägigen spezifischen Kenntnissen ist bisher erst wenig verbreitet, wobei zumindest im Bibliotheksbereich ein wachsendes Bewusstsein für dieses Problemfeld zu konstatieren ist. Parallel dazu herrscht gemäß den Aussagen der BefragungsteilnehmerInnen an den Instituten überwiegend die Ansicht, dass (Meta-)Datenerfassung für digitalen Content keine wissenschaftliche, sondern eine bibliothekarische Aufgabe ist. Es handelt sich somit um ein Betätigungsfeld, das die Bibliotheken für sich neu erschließen und professionalisieren könnten.

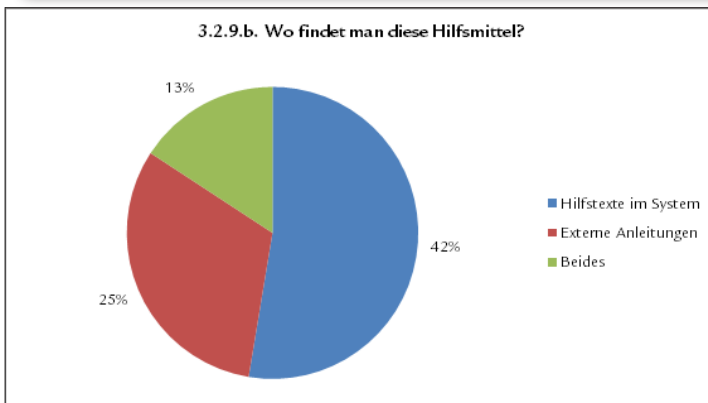
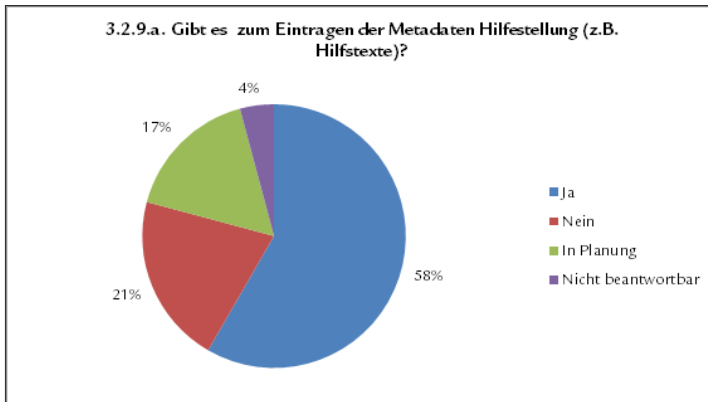
Auch hier hängt die Antwort stark vom Einsatz und der Policy des Repositoriums ab. So kann es z.B. aus rechtlichen Gründen notwendig sein, dass die User die Metadaten selbst ins System eintragen, weil sie für die Richtigkeit haften, zum Beispiel bei Hochschulschriften. Handelt es sich um Scans von historischen Archivbeständen oder Büchern, kann das Eintragen der Metadaten durchaus auch von BibliotheksmitarbeiterInnen übernommen werden.



### 3.2.9. Gibt es zum Eintragen der Metadaten Hilfestellung (z.B. Hilfstexte, wie die Metadatenfelder auszufüllen sind)? Wo findet man diese?

Die überwiegende Mehrheit der Institutionen bietet Hilfestellungen an, und zwar entweder in Form von Hilfstexten, die direkt im System zu finden sind, oder als externe schriftliche Anleitungen. Viele Einrichtungen verfügen über beides.

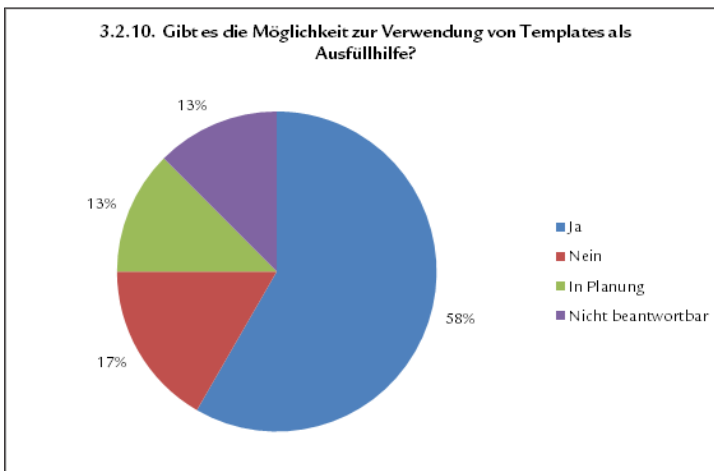
Vor allem bei Feldern, die über Autor, Titel, Datum und Beschreibung hinausgehen, ist es sinnvoll, Hilfstexte anzugeben, um den Usern eine Orientierung anzubieten. Im besten Fall sollten diese Hilfestellungen in mehreren Sprachen angeboten werden und für alle Personen, auch für jene, die über kein Hintergrundwissen verfügen, verständlich sein. Werden nur Standards und keine zusätzlichen Felder angeboten, sollte das Ausfüllen der Felder mehr oder weniger intuitiv erfolgen können.



### 3.2.10. Gibt es die Möglichkeit zur Verwendung von Templates als Ausfüllhilfe?

Das Anbieten von Templates ist ebenfalls überwiegend bereits realisiert bzw. in Planung.

Templates bieten zwei große Vorteile: Erstens kann man den Upload von ähnlichen Objekten wesentlich beschleunigen, wenn man nur diejenigen Eingaben ändern muss, die beim betreffenden Objekt abweichen und zweitens vermeidet man Fehler bzw. Inkonsistenz. Werden zum Beispiel im Rahmen eines Forschungsprojekts von mehreren Personen Daten hochgeladen und beschrieben, hilft ein gemeinsames Template, das gewünschte Vokabular zu verwenden, die passenden Begriffe im Thesaurus oder bei den Stichworten auszuwählen und es verringert sich damit auch die Anzahl der Tippfehler. Besonders komfortabel ist es, wenn man im System die Templates mit anderen Personen teilen kann, damit wird zum Beispiel der Projektleitung die Erstellung und Weitergabe des gewünschten Templates an die MitarbeiterInnen ermöglicht.

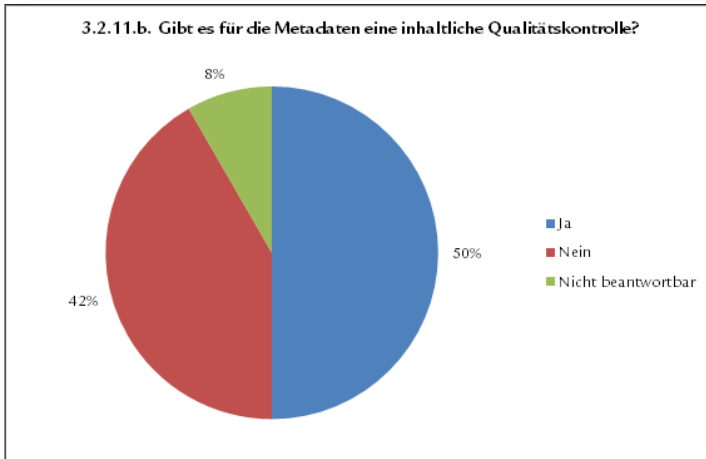
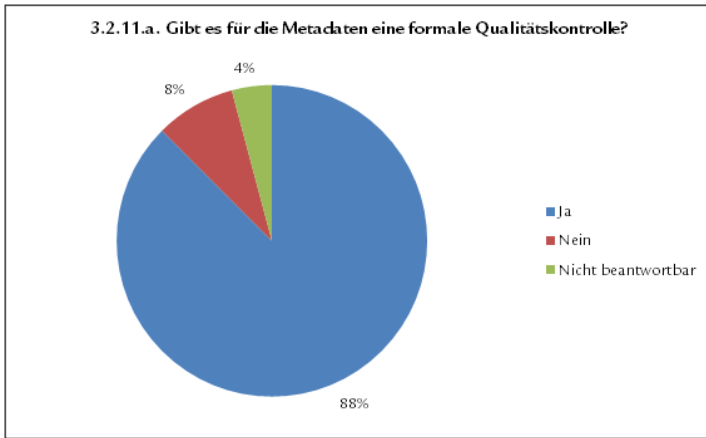


### 3.2.11. Gibt es für die Metadaten eine formale und / oder inhaltliche Qualitätskontrolle?

Nur wenige InterviewpartnerInnen gaben an, keine formale Qualitätskontrolle zu haben oder zu planen. Bei der inhaltlichen Qualitätskontrolle sieht es anders aus. Hier hat oder plant nur eine Minderheit der Institutionen keine oder nur eine sehr marginale Kontrolle. Viele konnten die Fra-

ge zu diesem Zeitpunkt aber auch noch nicht beantworten, da sie noch kein System hatten, bzw. es noch nicht im Regelbetrieb verankert war.

Die bevorzugte Methode zur formalen Kontrolle sind Pflichtfelder in der Eingabemaske, die bestimmte Mindestangaben bei der Datenerfassung erzwingen. Statt freier Textfelder werden – wenn möglich – vorgegebene Listen angeboten, die nur standardisierte Eingaben zulassen, z.B. bei der Verwendung von Thesauri (vgl. 3.2.12.). Weitere Kontrollmöglichkeiten: Datums-eingabe per Kalender-Steuer-element; automatische Übernahme von Anga-ben zum Erfasser aus der Personal-datenbank (verknüpft über die Anmelde-daten – keine Datenerfassung ohne Login im Repositorium); intellektuelle Prüfung bibliografischer Daten auf RAK-Konformität durch Fachkräfte.



### 3.2.12. Verwenden Sie einen Thesaurus oder ein sonstiges kontrolliertes Vokabular? Wenn ja, welches?

Die überwiegende Mehrheit verwendet Thesauri oder ein kontrolliertes Vokabular. Die Grafik 3.2.12.b. zeigt, dass die Auswahl an Thesauri und Klassifikationen recht groß ist, einige nannten auch spezielle Thesauri, die hier unter „andere“ subsumiert sind. Die Nennung freier Schlagwörter wurde trotz der Frage nach kontrolliertem Vokabular im Diagramm berücksichtigt, um diesen möglicherweise auch interessanten Aspekt nicht auszuklamern.

Die im österreichischen Bibliothekenverbund weit verbreitete Basisklassifikation (BK) steht an der Spitze. Überhaupt prägen Thesauri aus dem Bereich der traditionellen bibliografischen Kataloge (= ALEPH) die Antworten sehr stark, wie die weitere Nennung von Dewey (Dezimalklassifikation) und RSWK / Gemeinsamer Normdatei (GND) zeigt. Unter den nicht-universalen, fachspezifischen Klassifikationen ist nur der Getty-Thesaurus (Art & Architecture, Geographic Names) einigermaßen weit verbreitet.

### 3.2.13. Wie viele Felder für die Metadaten sind vorgesehen?

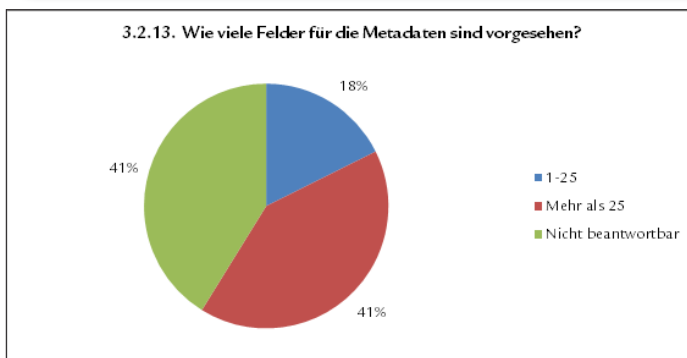
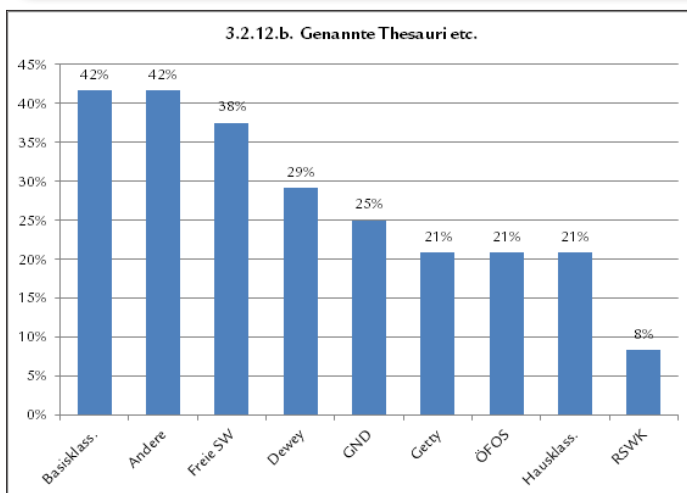
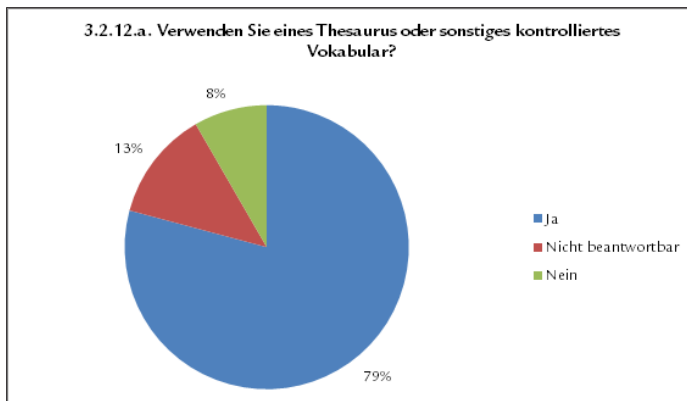
Hier ist ein leichter Trend zu mehr als 25 Metadatenfeldern zu erkennen. Einige InterviewpartnerInnen ergänzten jedoch, dass meist nicht alle dieser Felder verwendet werden. Der hohe Anteil von „nicht beantwortbar“ ergibt sich vor allem aus der Tatsache, dass einige der befragten Einrichtungen mehrere Repositorien (oder auch keines) betreiben und daher eine eindeutige Antwort nicht möglich ist.

### 3.2.14. Werden zurzeit auch Daten ohne Metadaten gespeichert?

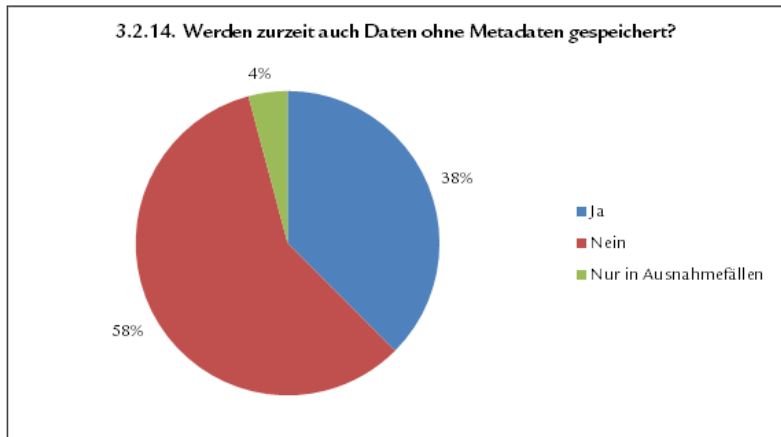
Daten ohne Metadaten werden zumeist außerhalb der offiziellen Repositorien (d.h. lokal auf Institutsebene etc.) gespeichert; die meisten Repositorien selbst lassen dies gar nicht zu. Die Antwort „Nein“ bedeutet i.d.R., dass über Praktiken der Datenerfassung außerhalb der Repositorien keine genauen Informationen vorliegen. Ziel ist in den allermeisten Fällen die Erfassung von Metadaten, die das Auffinden der beschriebenen Objekte erleichtern.

Denn, so wie Jeffrey Pomerantz schreibt: „Metadata is a map, Metadata is a means by which complexity of an object is represented in a simpler form.“<sup>6</sup>





Die Metadaten sollten auch schon vor dem geplanten Upload in ein Repository vorliegen, denn erfahrungsgemäß ist das nachträgliche Erstellen von Metadaten nicht nur sehr zeitaufwändig, sondern auch mit hohen Kosten verbunden. In vielen Fällen ist es auch schwierig, Metadaten nachträglich zu rekonstruieren, wenn beispielsweise nach Projektende die MitarbeiterInnen nicht mehr greifbar sind.



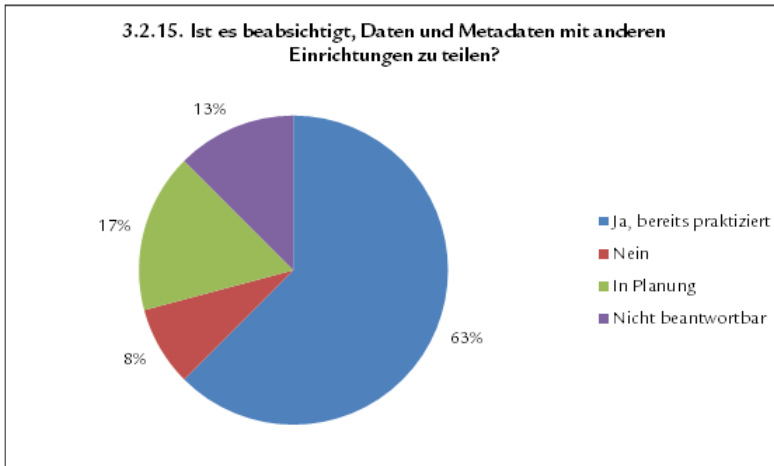
### 3.2.15. Ist es beabsichtigt, Daten und Metadaten mit anderen Einrichtungen zu teilen?

Eine deutliche Mehrheit der Befragten gibt an, die Daten und Metadaten mit anderen Einrichtungen teilen zu wollen und dies auch bereits zu praktizieren. Besonders für jene Einrichtungen, die sich dem Visual Library-Verbund angeschlossen haben, ist der Austauschgedanke ein wichtiger und wurde mehrfach als Grund für die Auswahl dieser Repositoryssoftware genannt.

Die konkreten Formen der Zusammenarbeit sind vielfältig: Datenlieferung an andere Repositorien und Datenbanken wie Europeana, Genios, UFORDAT oder arXiv; Öffnung des Repositoriums für das Harvesting (z.B. durch die Bielefeld Academic Search Engine [BASE]) über eine OAI-Schnittstelle; Lieferung von Forschungsdaten aus geförderten Projekten an den Fonds zur Förderung der wissenschaftlichen Forschung (FWF). Ein Thema für die nahe Zukunft ist der Ausbau von Linked Open Data-Schnittstellen.

In einem Fall wurden Probleme in urheber- und datenschutzrechtlichen Fragen angesprochen, die dem eigentlich gewünschten Umfang des Daten-

austausches Grenzen setzen. Hier wird die Rechtslage als insgesamt zu wenig transparent beurteilt und es besteht Klärungsbedarf bezüglich dessen, was erlaubt ist und was nicht.

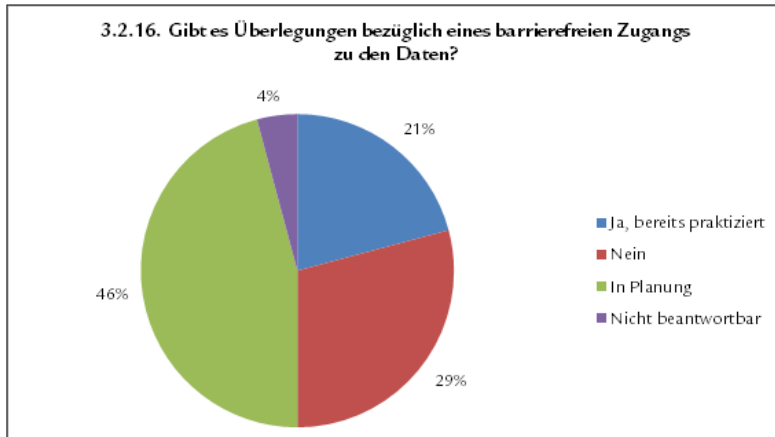


### 3.2.16. Gibt es Überlegungen bezüglich eines barrierefreien Zugangs zu den Daten?

Bemerkenswert ist, dass erst wenige Institutionen diese Frage bejahen konnten, was vorwiegend an den derzeit verwendeten, dafür nicht ausgelegten Systemen liegt. Die überwiegende Mehrheit strebt jedoch nach Barrierefreiheit und es gibt mehrere konkrete Projekte zur Umsetzung dieses Vorhabens.

Die barrierefreie Zurverfügungstellung von Daten in Informationsvermittlungssystemen wird mittlerweile auch im österreichischen Bundesbehindertengleichstellungsgesetz (BGStG) thematisiert<sup>7</sup>. Dennoch kann bisher kaum ein Repositoryum wirklich als barrierefrei gelten. Die Metadaten können hier in zweifacher Weise helfen, die Barrierefreiheit zu unterstützen. Einerseits dienen Beschreibungen dazu, z.B. Bilder für blinde oder sehbehinderte Personen zu erklären und erfahrbar zu machen, bzw. Videos mit Gebärdensprache können gehörlosen Personen, die die Gebärdensprache als Muttersprache erlernt haben und Deutsch als Fremdsprache erlernen mussten, Unterstützung beim Verstehen der Daten geben. Andererseits können aber auch Hilfstexte, die die Metadatenfelder erklären, in einfacher Sprache verfasst werden, sodass nicht nur Personen mit Lern-

schwierigkeiten, sondern auch Menschen, die eine andere Muttersprache haben, den Sinn besser erfassen können.<sup>8</sup>



Mag.<sup>a</sup> Dr.<sup>in</sup> Susanne Blumesberger, MSc  
ORCID: <http://orcid.org/0000-0001-9018-623X>  
Universität Wien, Bibliotheks- und Archivwesen  
E-Mail: [susanne.blumesberger@univie.ac.at](mailto:susanne.blumesberger@univie.ac.at)

Dr. Alexander Zartl  
Universität Wien, Bibliotheks- und Archivwesen  
E-Mail: [alexander.zartl@univie.ac.at](mailto:alexander.zartl@univie.ac.at)

- 1 Bauer, Bruno et al.: Forschende und ihre Daten. Ergebnisse einer österreichweiten Befragung – Report 2015. Version 1.2. DOI: <https://doi.org/10.5281/zenodo.32043>
- 2 <http://www.bath.ac.uk/research/data/policy/research-data-policy-guidance.html#policy-guidance-1> (letzter Aufruf am 31.10.2016)
- 3 <https://creativecommons.org/publicdomain/zero/1.0/deed.de>
- 4 [http://www.aalto.fi/en/midcom-serveattachmentguid-1e6480be-8be5bd0480b11e693ae8b27aad222542254/2016\\_02\\_10\\_datapolicy.pdf](http://www.aalto.fi/en/midcom-serveattachmentguid-1e6480be-8be5bd0480b11e693ae8b27aad222542254/2016_02_10_datapolicy.pdf) (letzter Aufruf am 31.10.2016)
- 5 An dieser Stelle vielen Dank an Imola Dora Riehle-Traub, die sich im Rahmen des H2020 Projekts LEARN (LEaders Activating Research Networks) intensiv mit Policies auseinandersetzte.

- 6 Pomerantz, Jeffrey: Metadata. Cambridge, London: The MIT Press 2015, p. 12.
- 7 Vgl. Bundesbehindertengleichstellungsgesetz (BGStG) § 6 Abs. 5: „Barrierefrei sind bauliche und sonstige Anlagen, Verkehrsmittel, technische Gebrauchsgegenstände, Systeme der Informationsverarbeitung sowie andere gestaltete Lebensbereiche, wenn sie für Menschen mit Behinderungen in der allgemein üblichen Weise, ohne besondere Erschwernis und grundsätzlich ohne fremde Hilfe zugänglich und nutzbar sind.“
- 8 Siehe auch: Blumesberger, Susanne; Andreas Jaitler: Metadata and Accessibility. <http://phaidra.univie.ac.at/o:459817>