

Susanne Blumesberger

Metadaten als Mehrwerte

Abstract: Der Beitrag gibt zunächst einen allgemeinen Überblick über die Bedeutung von Metadaten und erläutert den Wandel des Begriffs durch die Digitalisierung. Waren früher mit Metadaten vor allem Daten über Daten gemeint, zum Beispiel in Form von bibliografischen Angaben zu Büchern, hat sich der Begriff heute ausgeweitet; zugleich übernehmen Metadaten heute mehr Funktionen als früher. Darüber hinaus werden verschiedene Aspekte von Metadaten vorgestellt: Die Vorteile von hochqualitativen Metadaten werden ebenso aufgezeigt wie die heutigen und zukünftigen Herausforderungen – insbesondere auch im technischen Bereich. Beleuchtet wird zudem das Verständnis von Metadaten in den unterschiedlichen Fachdisziplinen. Im Fokus steht die Frage, wie Metadaten Daten im weitesten Sinne – also Bilder, Texte, Video- oder Audiofiles – so anreichern können, dass diese schnell auffindbar sind und möglichst langfristig, breit sowie darüber hinaus auch nachhaltig genutzt werden können. Ebenso wird dargestellt, welche Auswirkungen es hat, wenn nicht mehr – so wie früher – nur wenige hochqualifizierte Experten Metadaten vergeben.

Einleitung

Bis vor wenigen Jahren hatten Metadaten ausschließlich die Aufgabe, physisch vorhandene Objekte, wie Bücher, Texte oder auch audiovisuelle Materialien, zu beschreiben. Heute benötigt man darüber hinaus Metadaten auch für die Beschreibung von sogenannten *Digital-born-Materialien*, also von Texten oder Bildern, die nur digital vorhanden sind, bzw. von digitalisierten Objekten wie beispielsweise Büchern oder Karten, die nach wie vor physisch vorhanden sind, von denen aber auch digitale Versionen in Repositorien angeboten werden. Metadaten sind potenzielle Sucheinstiege (Suchbegriffe) in den verschiedenen Nachweissystemen wie Datenbanken, Katalogen, Discovery-Systemen oder auch in Suchmaschinen im Internet. Sie sollen Kunden eine effiziente Suche ermöglichen, zu relevanten Suchergebnissen führen und die Relevanzbeurteilung unterstützen.

1 Die Entwicklung der Metadaten

Bibliothekare entwickelten über Jahrhunderte hinweg komplexe Systeme, um die wachsende Fülle an zu archivierenden Materialien – vor allem Bücher, aber auch Handschriften, Nachlässe und Karten – in Katalogen zu repräsentieren, sie dadurch handhabbar und den Lesern zugänglich zu machen.

1.1 Metadaten sind Daten

Heute werden Metadaten durch die Online-Verfügbarkeit der Daten nicht mehr nur im bibliothekarischen Umfeld verwendet und auch nicht mehr ausschließlich von ausgewiesenen Informationsexperten erstellt und verwaltet. Damit kam es gewissermaßen zu einem Paradigmenwechsel und zu einer Neudefinition von Metadaten. Metadaten weiterhin nur als Daten über Daten zu beschreiben, würde unter den aktuellen Bedingungen zu kurz greifen. Metadaten können je nach Blickwinkel sehr unterschiedlich definiert werden. Man könnte darunter beispielsweise auch HTML-Seiten verstehen, die die Auffindbarkeit von Webseiten unterstützen. Metadaten können in verschiedenen Fachrichtungen auch zu Daten, zu Forschungsdaten werden. Denkt man beispielsweise an die Literaturgeschichte, sind bibliografische Angaben in einem Bibliothekskatalog oft wichtige Informationen, manchmal sogar Primärquellen der jeweiligen Forschung.

1.2 Zahlreiche Vorteile durch hochqualitative Metadaten

Ausführliche Metadaten steigern die Sichtbarkeit im Netz. Für Wissenschaftler, die sich z. B. mit auf Tonträgern festgehaltenen Interviews beschäftigen, geben Metadaten Auskunft, wie sie zu der Information auf den Tonträgern kommen und wie diese zu interpretieren sind. Für Archivare geben Metadaten Hinweise darauf, wie sie Informationen verlustfrei für die Zukunft speichern können. Außerdem erlauben standardisierte Metadaten den Austausch von Informationen zwischen Gedächtnisinstitutionen. Metadaten unterstützen Open Science¹, vor allem verlinkte offene Daten (Linked Open Data)² erlauben den Zugriff und den Verweis auf offene Ressourcen und eröffnen damit im Bildungsbereich und in der Wissenschaft neue Möglichkeiten. Im Sinne der „FAIR Data Principles“ sollen Daten *findable* (auffindbar), *accessible* (zugänglich), *interoperable* (austauschbar) und *reusable* (wiederverwendbar sein).³ Metadaten liefern hier einen sehr wichtigen Beitrag. Im Umgang mit personenbezogenen Daten garantieren Metadaten Sicherheit und Datenschutz. Das sind aber nur einige Beispiele. Metadaten sind auch aus unserem Alltag nicht wegzudenken: Sie beschrei-

¹ Offene oder öffentliche Wissenschaft hat zum Ziel, Wissenschaft möglichst ohne Barrieren einer großen Anzahl von Menschen zugänglich zu machen; dazu zählen der freie Zugang zu wissenschaftlicher Literatur (Open Access), offene Daten (Open Data), offene Bildungsressourcen, offene Quellen usw.

² Dabei handelt es sich um frei verfügbare Daten, die mit einem permanenten Identifier versehen sind, per http abgerufen werden können und wieder auf andere Daten verweisen.

³ Vgl. <https://www.force11.org/group/fairgroup/fairprinciples> (Abruf: 2018.03.12).

ben Produkte, die wir über das Internet bestellen wollen, man findet sie auf Büchern, in Beschreibungen von Geräten und vielem mehr.

1.3 Metadaten und Repositorien

Ein breites Aufgabenspektrum übernehmen Metadaten in digitalen Repositorien.⁴ Dort präsentieren sie nach wie vor das archivierte digitale Objekt, können aber auch den Zugang zu Objekten barrierefrei gestalten, indem sie beispielsweise Bilder mit einer Beschreibung versehen, sie in mehreren Sprachen verfügbar machen oder zusätzliche Hinweise für die Nutzung des digitalen Objekts anbieten. Metadaten können aber noch viel mehr. Forschungsdaten, Sammlungsobjekte usw. werden erst durch eine umfassende Beschreibung verständlich und so auch für andere Disziplinen und für spätere Generationen wiederverwendbar.

Daraus entstehen neue Herausforderungen. Metadaten müssen – wie die archivierten Daten und Informationsressourcen selbst – stets gepflegt und langzeitarchiviert werden. Um hochqualitative Metadaten – wenn möglich – automatisch zu generieren und sie optimal zu visualisieren, ist sowohl bibliothekarisches als auch fachspezifisches Wissen notwendig. Im Bereich der *Digital Humanities*⁵ werden qualitativ hochwertige Metadaten dazu genutzt, Forschungsergebnisse zu visualisieren und spannende Projekte umzusetzen. Eine Herausforderung der Zukunft wird sein, Forschern den Wert der Metadaten bewusst zu machen, denn das Erstellen und Pflegen sowie die langfristige Verfügbarkeit sind ressourcenintensiv. Bibliotheken können hier eine zentrale Rolle einnehmen, indem sie Wissen einbringen und die Forscher unterstützen. Schließlich entstehen daraus auch neue Tätigkeitsfelder wie beispielsweise Metadaten-Manager.⁶

Aufgrund der vielfältigen Funktionen, die Metadaten heute übernehmen, kann man sie als strukturierte Informationen über Objekte definieren.⁷

Die Archivierung von digitalen Objekten aller Art, beispielsweise Digitalisaten von Büchern, Karten, Texten, Videos, Audiodateien und multimedialen Objekten, hat dazu geführt, dass diese Objekte mit weiteren Informationen versehen werden müs-

4 Weiterführende Informationen in Blumesberger, Susanne: Die Welt der Metadaten im Universum von Repositorien. In: Mitteilungen der Vereinigung österreichischer Bibliothekarinnen und Bibliothekare 68/3, 2015, S. 515–528.

5 Das interdisziplinär ausgerichtete Fach Digital Humanities (digitale Geisteswissenschaften) bezieht sich auf die Anwendung von computergestützten Verfahren und die Verwendung von digitalen Ressourcen in den Geistes- und Kulturwissenschaften.

6 Weitere Informationen in Krier, Laura; Strasser, Carly A.: Data Management for Libraries. An Imprint of the Library Association. Chicago: ALA TechSource, 2014.

7 Vgl. Rühle o. J.

sen, um eine sinnvolle, nachhaltige und zum Teil auch globale Nutzung unterschiedlicher Zielgruppen zu ermöglichen. Dazu gehört zum Beispiel die Verwendung von standardisiertem Vokabular, von *Linked Data*⁸, aber auch die mehrsprachige Beschreibung der Daten. Gleichzeitig werden die Nutzer der Daten oft auch zu jenen, die die Daten mit weiteren Informationen anreichern. Ersteller und Nutzer dieser Objekte sprechen oft von *Tags*, Schlagwörtern, die von Nutzern vergeben werden, *Folksonomies*, die die Sammlung dieser Schlagwörter darstellen, oder *Social Bookmarks*.⁹ Die Verantwortung, Objekte zu beschreiben und diese Beschreibungen auch zu pflegen, liegt nicht mehr nur in den Händen von wenigen hochspezialisierten Experten, sondern auch oder vor allem bei jenen, die diese Objekte erstellen und nutzen. Mitte der 1990er Jahre wurde der Begriff *Metadaten* vor allem für Geodaten¹⁰ verwendet, die man für die Dokumentation, den Austausch und das Archivieren von Informationen über räumliche Positionierungen benötigte.

Der Aufbau von Repositorien in wissenschaftlichen Institutionen – zumeist Hochschul- und anderen wissenschaftlichen Bibliotheken – hat zu einer intensiven Auseinandersetzung mit Metadaten und damit zu vielen weiteren Fragen geführt. Unter anderem wurde und wird diskutiert, welches Metadatenschema man nutzen sollte, um die Inhalte möglichst rasch und ohne Informationsverlust mit anderen Systemen austauschen zu können. Daraus ergibt sich wiederum die Diskussion, ob ein Schema für alle Arten von Objekten passt und wenn nicht, welches vor allem für die jeweilige Fachdisziplin passend ist.¹¹ Mit der Etablierung von digitalen Archivierungssystemen stellte sich auch die Frage, wer für die Erstellung der Metadaten zuständig ist. Waren früher selbstverständlich die Bibliothekare unangefochtene Experten auf diesem Gebiet, reicht dieses rein bibliothekarische Wissen heute nicht mehr aus, um den Gesamtkomplex der Metadaten abzudecken, der nicht nur die reine Beschreibung von Daten umfasst, sondern auch (technisches) Wissen über Metadatenstandards, Schnittstellen, die die Kommunikation zwischen Systemen ermöglichen, Mapping, also das Übersetzen von einem Standard zum anderen, und vor allem fachspezifisches Wissen über die Objekte selbst enthält. Vor allem Letzteres verlangt die Einbeziehung von Fachwissenschaftlern der verschiedenen Disziplinen in den Beschreibungsprozess der Daten, also

8 Weitere Informationen in Neubauer, Gregor: Visualisierung von typisierten Links in Linked Data. In: Mitteilungen der Vereinigung österreichischer Bibliothekarinnen und Bibliothekare 70/2, 2017, S. 179–199.

9 S. auch <https://www.e-teaching.org/technik/kommunikation/socialbookmarking> (Abruf: 2018.03.12).

10 Geodaten beschreiben Objekte, die durch eine Position im Raum referenzierbar sind. Sie sind in Georeferenzdatenbanken abrufbar und stellen wichtige Basisdaten für die Stadtplanung, Archäologie, Kartografie, aber auch für das Gesundheitssystem dar. Mithilfe dieser Daten können beispielsweise Evakuierungspläne erstellt werden.

11 Einen Überblick findet man unter https://en.wikibooks.org/wiki/Open_Metadata_Handbook/Metadata_Standards (Abruf: 2018.03.12).

in die Erstellung von Metadaten. Dies führt einerseits zu einer Mehrbelastung der Forscher, die im Rahmen ihrer Forschungsarbeit oft nicht über ausreichende Ressourcen verfügen, um die generierten Daten adäquat zu beschreiben, andererseits gibt es den Wissenschaftlern zugleich die Möglichkeit, selbst zu bestimmen, wie ihre Forschungsergebnisse zugänglich und gut auffindbar gemacht werden. Autoren von Büchern oder Zeitschriftenartikeln hatten bisher kaum bis gar keinen Einfluss darauf, wie ihre Publikationen in Bibliotheken auffindbar gemacht wurden. Die Einbindung der Forscher in die Metadatenerstellung führt zu einer engen Zusammenarbeit der beiden Gruppen; Forschungsdatenmanagement¹², wie man es heute versteht, ist ohne diese enge Kooperation schwierig, da nur die Wissenschaftler selbst wissen, was ihre Daten konkret beinhalten und aussagen, und nur sie diese Daten korrekt beschreiben können. Dabei kommt den Bibliotheken die Aufgabe zu, Forscher bei der Erstellung und Archivierung ihrer Objekte und der Metadaten möglichst effektiv zu unterstützen und vor allem auch bewusstmachend zu wirken, da vielen Wissenschaftlern das Thema Metadaten bisher wenig bis gar nicht vertraut ist.

Daten werden im Forschungsprozess verwendet und in großen Mengen generiert, oft jedoch nur lokal gespeichert und auch nur unzulänglich beschrieben. Das führt dazu, dass diese oft aufwendig und teuer erstellten Daten weder von den Forschern selbst noch von anderen nachgenutzt werden können. Drittmittelgeber haben zum Teil bereits darauf reagiert und fordern sogenannte *Datenmanagementpläne* ein, die unter anderem eine Vorausschau für den Umgang mit Finanzen, Verantwortlichkeiten, Zeitbudget, Sicherheits- und Ethikfragen, Langzeitverfügbarkeit und auch die Handhabung der Metadaten beinhalten müssen. Damit sind all jene, die in Drittmittelprojekten arbeiten, verpflichtet, sich mit dem Thema Metadaten auseinanderzusetzen. Dies hat zur Konsequenz, dass Bibliothekare jener Bibliotheken, die Forschungsunterstützung betreiben, sich verstärkt mit den unterschiedlichen Fachrichtungen auseinandersetzen müssen, um die Forscher möglichst gut in ihrem Forschungsprozess begleiten zu können.

Zahlreiche Wissenschaftler kennen sich mittlerweile gut mit verschiedenen Metadatenschemata wie *Dublin Core*¹³ – einem standardisierten Vokabular für die Erzeugung von Metadaten –, „Metadata Objects Description Schema“ (MODS)¹⁴, dem XML-Format für bibliografische Metadaten, oder „Metadata Encoding and Transmission Standard“ (METS)¹⁵ aus. Aber viele spezielle Metadatenstandards, etwa der *Darwin*

12 „Forschungsdatenmanagement bezeichnet den Prozess, der alle Methoden und Verfahren umfasst, die zur Sicherung der langfristigen Nutzbarkeit von Forschungsdaten angewendet werden: die Generierung, die Bearbeitung, die Anreicherung, die Archivierung und die Veröffentlichung.“ <http://www.forschungsdaten.org/index.php/Forschungsdatenmanagement> (Abruf: 2018.03.12).

13 S. <http://dublincore.org/> (Abruf: 2018.03.12).

14 S. <http://www.loc.gov/standards/mods/> (Abruf: 2018.03.12).

15 S. <http://www.loc.gov/standards/mets/> (Abruf: 2018.03.12).

*Core*¹⁶ für die Biologie oder die „Text Encoding Initiative“ (TEI)¹⁷ für Digital-Humanities-Projekte erfordern spezielles Wissen. In Zukunft wäre es zielführend, wenn sich auch Bibliothekare mit unterschiedlichen Metadatenschemata beschäftigen würden und die Möglichkeit hätten, sich auf diesem Gebiet weiterzubilden. Das Verwenden geeigneter Schemata ermöglicht erst den Austausch der Daten in den jeweiligen Communitys und macht die Interoperabilität, also die Zusammenarbeit und den Austausch von Metadaten zwischen den Systemen möglich.¹⁸

Verwenden die Forscher strukturierte Metadatenschemata, garantieren sie damit gleichzeitig eine konsistente Beschreibung ihrer Daten, die auch von künftigen Wissenschaftlern genutzt werden können. Wichtig sind strukturierte und hochqualitative Metadaten auch für die Auffindbarkeit der Daten und somit für die Sichtbarkeit der Forschung, was aufgrund der heutigen immensen Fülle an Forschungsbetrieb immer wichtiger wird. Eine gute Beschreibung der Daten ermöglicht es auch den Repositorien-Managern, für eine langfristige Verfügbarkeit der Daten und die Möglichkeit eines Austausches zu sorgen.

2 Arten von Metadaten

Metadaten beziehen sich auf folgende zentrale Aspekte: auf den Inhalt, auf den Kontext und auf die Struktur. Demnach werden unterschieden:

- beschreibende Metadaten;
- administrative Metadaten;
- strukturelle Metadaten und
- technische Metadaten.¹⁹

2.1 Beschreibende Metadaten

Beschreibende Metadaten umfassen Informationen über die Objekte. So wäre ein Bild „Boot auf einem Teich“ mehr oder weniger wertlos, würde nicht die Information mitgeliefert, dass es sich dabei um eine Fotografie handelt, die die Weltreise dokumentiert, die Ludwig Lorenz von Liburnau 1892/93 als wissenschaftlicher Begleiter im Gefolge des Erzherzogs Franz Ferdinand von Österreich-Este unternommen hat.²⁰

¹⁶ S. <http://rs.tdwg.org/dwc/> (Abruf: 2018.03.12).

¹⁷ S. <http://www.tei-c.org/index.xml> (Abruf: 2018.03.12).

¹⁸ Weitere Informationen in Baca, Mutha: *Introduction to Metadata*. 2. Auflage. Los Angeles: Getty Research Institute, 2008.

¹⁹ Vgl. bspw. <https://wiki.dnb.de/pages/viewpage.action?pageId=95651769> (Abruf: 2018.03.12).

²⁰ Vgl. <http://phaidra.univie.ac.at/o:576159> (Abruf: 2018.03.12), <http://dublincore.org/> (Abruf: 2018.03.12), <http://www.loc.gov/standards/mods/> (Abruf: 2018.03.12), <http://www.loc.gov/standards/mets/>

Diese Metadaten können auch Informationen zu unterschiedlichen Versionen der Objekte beinhalten sowie Links zu anderen Objekten und Annotationen umfassen. Bei Forschungsdaten kann man zwei Varianten unterscheiden: einerseits die Metadaten, die die Datensets beschreiben, und andererseits die Metadaten, die die einzelnen Daten beschreiben. Metadaten, die Datensets beschreiben, umfassen Informationen wie den Urheber, das Entstehungsdatum und auf welche Weise die Daten entstanden sind – zum Beispiel, welche Tools für die Datengewinnung verwendet wurden oder welche Messverfahren verwendet wurden. Damit erhalten die Daten relevante Zusatzinformationen, um auch in der Zukunft verstanden und besser nachvollzogen werden zu können.

Forscher sollten sich möglichst früh mit den geplanten Repositorien vertraut machen, um die verlangten Pflichtfelder adäquat befüllen und somit die Daten möglichst genau beschreiben zu können. Meist ist es sinnvoll, sich auf *Dublin Core* zu stützen, da dieses Schema kein Expertenwissen voraussetzt und 15 überschaubare Elemente enthält: Beitragende, Ort und Zeit, Ersteller, Datum, Beschreibung, Format, Identifizierer, Sprache, Verlag/Herausgeber, Beziehungen, Rechte, Quelle, Stichworte, Titel, Dokumententyp.²¹ Benötigt man weitere Informationen, steht *Qualified Dublin Core*²², das unter anderem auch Hinweise auf Provenienz und Verhältnisse zwischen den Objekten enthält, zur Verfügung. In letzter Zeit wird auch oft das „DataCite Metadata Schema“²³ herangezogen, vor allem um „Digital Object Identifiers“ (DOI)²⁴ registrieren zu können. Bei der Beschreibung der einzelnen Daten sollten sich die Forschenden bewusst sein, dass es für den wissenschaftlichen Austausch und die Wiederverwendbarkeit der Daten wichtig ist, immer gleiche Begriffe, möglichst aus Thesauri, zu verwenden, die das Verstehen der Daten für alle erleichtern.²⁵ Hier sollten die Bibliothekare unterstützend eingreifen. Selbstverständlich muss auch darauf geachtet werden, dass auch die Metadaten vor Verlust und unberechtigtem Zugriff sicher gespeichert werden, und dies bereits vor der Archivierung in einem Repository. Für Bibliotheken ergeben sich damit neue Aufgaben im Bereich der Forschungsunterstützung. Das traditionelle Wissen der Bibliothekare über Metadaten kann genutzt werden, um Forschern bewusst zu machen, wie wertvoll Metadaten sind.

(Abruf: 2018.03.12), <http://rs.tdwg.org/dwc/> (Abruf: 2018.03.12), <http://www.tei-c.org/index.xml> (Abruf: 2018.03.12).

21 Vgl. https://glossar.hs-augsburg.de/Dublin_Core_Metadata (Abruf: 2018.03.12).

22 Vgl. <http://dublincore.org/documents/2000/07/11/dcmes-qualifiers/> (Abruf: 2018.03.12).

23 Vgl. <https://www.datacite.org/> (Abruf: 2018.03.12).

24 Vgl. <https://www.doi.org/> (Abruf: 2018.03.12).

25 Weitere Informationen in Kopácsi, Sándor; Hudak, Rastislav; Ganguly, Raman: Implementierung eines Klassifikationsservers für Metadatenorganisation für Langzeitarchivierungssysteme. In: Mitteilungen der Vereinigung österreichischer Bibliothekarinnen und Bibliothekare 70/2, 2017, S. 225–243.

2.2 Administrative Metadaten

Administrative Metadaten geben Auskunft darüber, um welchen Ressourcentyp und welches Format es sich handelt, über Identifier sowie über die Auswahlkriterien für die Digitalisierung. Metadaten beschreiben die rechtliche Situation der Daten, informieren über Zugriffsrechte bzw. darüber, ob es sich um kein Born-digital-Objekt handelt, sowie über den Aufbewahrungsort des analogen Objekts. Diese Angaben sind grundlegend, um Daten langfristig verfügbar zu halten. Das von der Digital Library Federation (DLF) entwickelte METS-Schema²⁶, das von der Library of Congress (LoC) gepflegt wird, gibt in sieben Bereichen Auskunft über technische Aspekte, Formate, die rechtliche Situation, über den Inhaber der Daten, Rechte, über die zugrunde liegenden Quellen und die Provenienz der Daten und wird vor allem für digitalisierte Fotografien und Karten eingesetzt. All dies hilft dabei, während des gesamten Lebenszyklus der Daten Veränderungen in den Daten wahrzunehmen und nachvollziehen zu können. Mitprotokolliert werden sollte auch, wer wann welche Änderungen an den Metadaten vorgenommen hat. Bei den rechtlichen Angaben haben sich vor allem die Creative-Commons-Lizenzen durchgesetzt. Die gemeinnützige Organisation Creative Commons (CC) bietet Lizenzen zur Veröffentlichung von Texten, Bildern, Musik oder Videos an, um Urhebern die Möglichkeit zu geben, anderen Personen die Nutzung ihrer Werke unter bestimmten Bedingungen zu erlauben.²⁷

2.3 Strukturelle Metadaten

Strukturelle Metadaten, die man zu den administrativen Metadaten zählen kann, helfen ebenfalls dabei, die Daten langfristig zu sichern und nachnutzbar zu machen.

²⁶ Der Kopfteil enthält Metadaten, die das jeweilige METS-Dokument selbst beschreiben, einschließlich der Angaben zum Bearbeiter oder Herausgeber des METS-Dokuments. Der Abschnitt für die Erschließungsangaben kann sowohl Verweise auf ein externes Dokument wie auch in das METS-Dokument eingebettete Angaben oder beides enthalten. Der Abschnitt für die Verwaltungsangaben liefert Informationen über die Herstellung und Speicherung von Dateien, über Urheberrechte und über die digitalisierte Vorlage. Im Dateienabschnitt werden alle Dateien mit Inhalten, aus denen das digitale Objekt besteht, aufgelistet. Die Strukturbeschreibung bildet den inneren Aufbau des digitalen Objekts ab und verknüpft die Elemente der Struktur mit den Dateien, aus denen der Inhalt des digitalen Objekts besteht, sowie mit deren Metadaten. Der Abschnitt mit den Strukturverknüpfungen erlaubt es den Erstellern von METS-Dokumenten, das Vorhandensein von Hyperlinks zwischen einzelnen Knoten des im Strukturabschnitt dargestellten hierarchischen Aufbaus des digitalen Objekts zu beschreiben. Ein Abschnitt über das Verhalten des digitalen Objekts kann verwendet werden, um ausführbare Anweisungen für das Verhalten mit den Inhalten in METS-Objekten zu verknüpfen. S. http://www.loc.gov/standards/mets/METSOverview.v2_de.html (Abruf: 2018.03.12).

²⁷ Vgl. <https://creativecommons.org/licenses/?lang=de> (Abruf: 2018.03.12).

Sie beschreiben die physische und logische Struktur des digitalen Objekts, was vor allem dann entscheidend ist, wenn es sich um Daten handelt, die sich aufeinander beziehen, wie z.B. digitale Bücher, die in Kapitel unterteilt sind, die wiederum Seiten beinhalten. Hier ist es wichtig, dass diese Struktur gut beschrieben ist. Auch bei im Repositorium miteinander verknüpften Daten ist es wichtig zu dokumentieren, auf welche Art und Weise diese Daten miteinander korrespondieren. Sammlungen von Objekten können durch Unterordner klar strukturiert werden und erleichtern damit die Suche nach ähnlichen Daten. Bilder, Texte und Videos können somit in einen thematischen Zusammenhang gebracht werden.

2.4 Technische Metadaten

Technische Metadaten enthalten eine Dokumentation über Soft- und Hardware, Information zu den Formaten, zu Komprimierung oder Skalierungsmöglichkeiten. Ein Beispiel dafür ist das „Exchangeable image file format“ (Exif), das bei der digitalen Fotografie Standard ist.²⁸ Metadaten, die sich auf die langfristige Sicherung der Daten beziehen, enthalten Angaben zu Migrationen der Daten, also z.B. zum Wechsel von einem Format zu einem anderen, und sämtliche Änderungen, die an den Daten vorgenommen wurden. Zusätzlich sind auch Metadaten zu nennen, die statistische Daten enthalten und sich auf die Nutzung und Nutzungshäufigkeit der Objekte beziehen.

Metadaten in Repositorien erlauben auch neue Suchmöglichkeiten durch die unterschiedlichen Vernetzungen der einzelnen Beschreibungen und erleichtern damit auch das Auffinden von Informationen durch unterschiedliche Zielgruppen, die bei der Suche individuell vorgehen und anders als früher unterschiedliche Zugänge haben, die nicht mehr voraussetzen, dass bestimmte Suchstrategien eingehalten werden. Sammlungen an Informationen – beispielsweise digitalisierte Nachlässe – können damit auch von einem einzelnen Objekt ausgehend durchsucht und entdeckt werden, der Nutzer muss nicht mehr die oberste Ebene kennen und von dort aus seine Suche starten. Die Hierarchie bzw. die Struktur muss also nicht schon von vornherein bei Beginn der Suche bekannt sein.

Metadaten ermöglichen virtuelle Ausstellungen, die Repräsentation von Sammlungen, sind aber auch beispielsweise für den Einsatz in der Lehre unverzichtbar. Mithilfe des „Learning Object Metadata“-Schemas (LOM)²⁹, das 2002 vom „Institute of Electrical and Electronics Engineers“ (IEEE)³⁰ entwickelt wurde, können Angaben wie Alter

28 Herstellerunabhängiges Metadatenformat für den Austausch von Bild- und Fotodateien für die Kommunikation zwischen Digitalkamera und Drucker, Notebook, Tablet, Smartphone etc.

29 S. z.B. http://edutechwiki.unige.ch/en/Learning_Object_Metadata_Standard (Abruf: 2018.03.12).

30 <https://www.ieee.org/index.html> (Abruf: 2018.03.12)

der Zielgruppe, Art der Lernressource (z.B. Experiment, Text, Übung, Grafik) und voraussichtlicher Zeitaufwand, aber auch rechtliche Bedingungen usw. festgehalten werden.³¹ Durch diese Zuschreibungen werden Daten zu Lernressourcen, die in Schulen und Universitäten genutzt werden können.

3 Das Generieren von Metadaten

Das Erzeugen und Managen von Metadaten ist zu einem hochkomplexen Vorgang geworden, der sowohl inhaltliches Fachwissen als auch technisches Spezialwissen erfordert. Ein wichtiger Punkt ist auch, dass Metadaten und Daten fix miteinander verbunden sein sollen, dass die Metadaten ein fixer Bestandteil der Daten sind, denn sie begleiten die Daten während ihres gesamten Lebenszyklus. Vorteilhaft ist es daher, wenn nicht nur eine Person für das Erstellen von Metadaten zuständig ist, damit Arbeitsabläufe in Institutionen klar definiert und eingehalten werden können.

Metadaten können auch durch Crowdsourcing erzeugt werden, indem man das Wissen vieler Menschen, vor allem auch Externer, zu einem Thema/Sachverhalt nutzt und sie einlädt, beispielsweise Personen oder Orte auf Fotos zu identifizieren bzw. schon vorhandene Beschreibungen anzureichern oder auch zu korrigieren. Ein Fotoarchiv einer Gemeinde könnte beispielsweise die Einwohner bitten, abgebildete Personen zu benennen, damit würde man das Archiv aufwerten. Dies erfordert eine gewisse Qualitätskontrolle, aber es bietet den Bibliotheken die Möglichkeit, auf diese Weise Bestände verfügbar zu machen, die ansonsten der Öffentlichkeit verwehrt wären, da die Bibliothek weder über das Wissen noch über die Kapazitäten verfügt, um z. B. große Sammlungen an Fotos zu erschließen.³²

Fazit

Metadaten in Repositorien eröffnen neue Suchmöglichkeiten durch die unterschiedlichen Vernetzungen der einzelnen Beschreibungen und erleichtern damit das Auffinden von Informationen durch unterschiedliche Zielgruppen, die bei der Suche anders als früher (wo bestimmte Suchstrategien eingehalten werden mussten) indivi-

31 S. z. B. IEEE Standards Association: IEEE Learning Technology Standards Committee (LTSC) Systems Interoperability in Education and Training, 2017, <https://ieee-sa.imeetcentral.com/ltsc/> (Abruf: 2018.03.12). Einen Kurseditor findet man unter <ftp://ftp.kom.e-technik.tu-darmstadt.de/papers/HFM+01-1-paper.pdf> (Abruf: 2018.03.12). Das Leibniz-Institut für Wissensmedien bietet unter <https://www.e-teaching.org> (Abruf: 2018.03.12) zahlreiche Informationen für die Gestaltung der Hochschullehre mit digitalen Medien an.

32 Vgl. den Beitrag „Open Innovation und Crowdsourcing: Das Management von Offenheit“ von Ursula Georgy in diesem Handbuch.

duell unterschiedliche Zugänge haben. Sammlungen an Informationen, beispielsweise digitalisierte Nachlässe, können damit von einem einzelnen Objekt ausgehend durchsucht und entdeckt werden, der Nutzer muss nicht mehr die oberste Ebene kennen und von dort aus seine Suche starten. Die Hierarchie bzw. die Struktur muss also nicht bei Beginn der Suche bekannt sein.

Zusammenfassend lassen sich unter anderem folgende Vorteile von Metadaten auflisten:

- Sie ermöglichen den Zugang zu den Daten. Keine Bibliothek kommt ohne Metadaten – in diesem Fall Katalogen, Discovery-Systemen und Suchmaschinen – aus. Vor allem in der „digitalen“ Welt sind Metadaten unverzichtbar, will man Objekte auffinden.
- Daten werden erst durch Metadaten sichtbar. Ähnlich wie Bücher in einer Bibliothek durch den Katalog erschlossen werden, informieren Metadaten über die archivierten Daten.
- Die Weitergabe von Metadaten ermöglicht den Aufbau von Portalen und Datenbanken wie beispielsweise der Europeana³³.
- Forschungsdaten ohne (weitere) Erklärungen sind kaum nutzbar.
- Metadaten reduzieren die Komplexität von Systemen.
- Metadaten erlauben den internationalen und fächerübergreifenden Austausch der Daten.
- Metadaten ermöglichen die globale Nutzung der Daten.
- Durch die mehrsprachige Beschreibung und die Verwendung von Thesauri, aber auch natürlichsprachlichen Texten sowie Verlinkungen auf andere Ressourcen werden Daten für eine breite Nutzerschicht zugänglich.
- Metadaten erleichtern den Zugang zu Informationen für Personen mit körperlichen Einschränkungen. Beschreibungen von Bildern oder Grafiken erleichtern sehbehinderten oder blinden Personen den Zugang zu Daten, ebenso verhelfen Videos mit Gebärdensprache gehörlosen Personen zur Nutzung von digitalen Inhalten.
- Metadaten bewahren den Kontext der Daten auf.
- Metadaten machen die Daten für die Lehre zugänglich.
- Metadaten dokumentieren die Nutzung von Daten.
- Sie verwalten unterschiedliche Versionen von Daten.
- Metadaten geben Auskunft über rechtliche Aspekte. Sie zeigen, wer die Rechte an den jeweiligen Objekten besitzt, welche Lizenzen vergeben wurden und wie die Zugriffsrechte gestaltet sind.
- Sie enthalten Informationen darüber, wie die digitalen Objekte auch zukünftig am Leben erhalten und weiter genutzt werden können.

³³ S. <https://www.europeana.eu/portal/de> (Abruf: 2018.03.12).

- Metadaten geben Auskunft über unterschiedliche kulturell oder politisch bedingte Benennungen von Personen oder Orten.
- Sie geben Hinweise auf eventuell analog nicht mehr vorhandene Objekte.
- Durch die Angabe des Urhebers, dessen Lebensdaten und der Entstehungsdaten des Werks sind wertvolle Hinweise auf die jeweilige rechtliche Situation vorhanden. Die Angabe von Lizenzen erleichtert die Nachnutzung, ebenso die Information, ob ein Text bereits publiziert ist oder nicht.
- Kontrolliertes und standardisiertes Vokabular ermöglicht ein rasches Auffinden der Daten, da unterschiedliche Schreibweisen bzw. historische Benennungen miteinander verlinkt sind bzw. Homonyme kenntlich gemacht werden.
- Persistente Identifier, wie beispielsweise die Digital Object Identifiers (DOI), garantieren die Zitierung der Objekte, die Open Researcher and Contributor ID (ORCID) wird als einmalige „Adresse“ für Forscher eingesetzt; Namensgleichheiten führen nicht mehr zu Verwechslungen, Namensänderungen sind bei der Nachverfolgung wissenschaftlicher Karrieren kein Hindernis mehr.
- Metadaten ermöglichen virtuelle Ausstellungen sowie die Repräsentation von Sammlungen.

Die Metadaten selbst, aber auch der Umgang mit ihnen hat sich in der letzten Zeit stark gewandelt. Durch die zunehmende Digitalisierung von Objekten stehen zwar immer mehr Informationen zur Verfügung, die Auffindbarkeit ist allerdings nur mehr dann wirklich gegeben, wenn die Informationen gut beschrieben sind und von Mensch und Maschine gelesen werden können. Wie Metadaten möglichst rasch, zeitsparend und wenn möglich maschinell gewonnen werden können, wird nur eine der vielen Fragen sein, die uns in dieser Hinsicht noch beschäftigen werden.³⁴

Literatur

Letztes Abrufdatum der folgenden Internet-Dokumente ist der 2018.03.12.

- Baca, Mutha: Introduction to Metadata. 2. Auflage. Los Angeles: Getty Research Institute, 2008.
- Blumesberger, Susanne: Die Welt der Metadaten im Universum von Repositorien. In: Mitteilungen der Vereinigung österreichischer Bibliothekarinnen und Bibliothekare 68/3, 2015, S. 515–528.
- Fensel, Anna: Auf dem Weg zu semantischen APIs für Forschungsdatendienste. In: Mitteilungen der Vereinigung österreichischer Bibliothekarinnen und Bibliothekare 70/2, 2017, S. 157–169.
- Kopácsi, Sándor; Hudak, Rastislav; Ganguly, Raman: Implementierung eines Klassifikationservers für Metadatenorganisation für Langzeitarchivierungssysteme. In: Mitteilungen der Vereinigung österreichischer Bibliothekarinnen und Bibliothekare 70/2, 2017, S. 225–243.

³⁴ Weitere Informationen in Fensel, Anna: Auf dem Weg zu semantischen APIs für Forschungsdatendienste. In: Mitteilungen der Vereinigung österreichischer Bibliothekarinnen und Bibliothekare 70/2, 2017, S. 157–169.

Krier, Laura; Strasser, Carly A.: Data Management for Libraries. An Imprint of the Library Association. Chicago: ALA TechSource, 2014.

Neubauer, Gregor: Visualisierung von typisierten Links in Linked Data. In: Mitteilungen der Vereinigung österreichischer Bibliothekarinnen und Bibliothekare 70/2, 2017, S. 179–199.

Pomerantz, Jeffrey: Metadata. Cambridge, London: The MIT Press, 2015.

Rühle, Stefanie: Kleines Handbuch Metadaten. O.J. http://www.kim-forum.org/Subsites/kim/SharedDocs/Downloads/DE/Handbuch/metadaten.pdf?__blob=publicationFile