



Essentials für fachspezifische Empfehlungen (R 3.2.1)

Version 10.10.2012 - V. 1.0

Arbeitspaket 3.2 Fachspezifische Standards und Empfehlungen

Verantwortlicher Partner DT/PB

DARIAH-DE Aufbau von Forschungsinfrastrukturen für die e-Humanities

Dieses Forschungs- und Entwicklungsprojekt wird / wurde mit Mitteln des Bundesministeriums für Bildung und Forschung (BMBF), Förderkennzeichen 01UG1110A bis M, gefördert und vom Projektträger im Deutschen Zentrum für Luft- und Raumfahrt (PT-DLR) betreut.

GEFÖRDERT VOM



Bundesministerium
für Bildung
und Forschung

Projekt: DARIAH-DE: Aufbau von Forschungsinfrastrukturen für die e-Humanities

BMBF Förderkennzeichen: 01UG1110A bis M

Laufzeit: März 2011 bis Februar 2014

Dokumentstatus: Final, Version 1.0

Verfügbarkeit: öffentlich

Kontakt: dariah-de-uap3.2@gwdg.de

Autoren:

Nikolaos Beer, DT/PB

Kristin Herold, DT/PB

Wibke Kolbmann, DAI

Thomas Kollatz, STI

Stefan Schmunk, IEG

Joachim Veit, DT/PB

Unter weiterer Mitarbeit von:

Johannes Kepper (Musikcodierung), Harald Lordick (externe Verwaltung von Metadaten), Kristina Richts (Musikcodierung), Sebastian Rose (Diplomatik), Philipp Vanscheidt (Kodikologie)

Revisionsverlauf:

Datum	Autor	Kommentare
17.04.2012	Wibke Kolbmann	Gliederung Erstfassung
25.07.2012	Kristin Herold, Wibke Kolbmann, Thomas Kollatz, Stefan Schmunk	Einleitung, Standards und Metadatenformate, Geisteswissenschaftliche Grundkonzepte
14.09.2012	Kristin Herold, Wibke Kolbmann, Thomas Kollatz, Stefan Schmunk, Joachim Veit	inhaltliche Ausarbeitung
25.09.2012	Nikolaos Beer, Kristin Herold	Layout
10.10.2012	Stefan Schmunk, Joachim Veit	Endredaktion

Inhaltsverzeichnis

Präambel	4
1. Einleitung	5
1.1 Ziel und Auswahl für die fachspezifischen Empfehlungen	5
1.2 Begriffsklärung	6
1.3. Metadaten in den Geisteswissenschaften – Vorteile der Verwendung von Standards und mögliche Nutzungsszenarien	7
2. Standards und Metadatenformate in den Fachdisziplinen	10
2.1 Archäologie	14
2.2 Musikwissenschaft	17
2.3 Geschichtswissenschaft	27
2.4 Judaistik und Hebraistik	31
2.5. Geisteswissenschaftliche Hilfswissenschaften	33
3. Geisteswissenschaftliche Grundkonzepte	34
3.1 Datierung	35
3.2 Ortsbeschreibung und Georeferenzierung	36
3.3 Personeninformationen	40
3.4 Objektinformationen	43
3.5 Quellennachweise	48
3.6 Ereignisbeschreibung	51
4. Metadaten zum Forschungskontext in den Geisteswissenschaften	53
4.1 Forschungsmethode und -verfahren	54
4.2 Dokumentationsmöglichkeiten von administrativen Metadaten am Beispiel des <code>teiHeader</code>	56
5. Schlussfolgerungen	60
6. Anhang	62
6.1 Literatur und URLs	62
6.2 Abkürzungen	67
6.3 Zusammenstellung geisteswissenschaftlicher Quellentypen	69
6.4 MEI Beispiel (Robert Schumann: Der Abendstern op. 79/1)	73

Präambel

In seinen *Empfehlungen zur Weiterentwicklung der wissenschaftlichen Informationsinfrastrukturen in Deutschland bis 2020*¹ hat der Wissenschaftsrat im Juli 2012 u.a. ausgeführt: “Die Fachgemeinschaften bzw. Akteure in interdisziplinären Forschungsfeldern sollten, soweit nicht bereits vorhanden, Qualitätskriterien für die Generierung von Forschungsdaten und Richtlinien für ein geeignetes Datenmanagement erarbeiten. Diese sollten insbesondere Konventionen für die Erstellung von Metadaten [...] enthalten” (S. 56). Genau in diesem Sinne möchten die nachfolgenden *Essentials für fachspezifische Empfehlungen* die Diskussion über die Qualität und Leistungsfähigkeit unterschiedlicher Metadatenformate befördern helfen, zugleich aber auch eine Orientierungshilfe für Wissenschaftler bieten, die neue Forschungs- oder Digitalisierungsprojekte in Angriff nehmen.

Zwei einschränkende Bemerkungen müssen dem nachfolgenden Report vorangestellt werden:

1. Es handelt sich um Empfehlungen, die zunächst aus dem Blickwinkel jener geisteswissenschaftlichen Fächer erarbeitet wurden, die sich zum jetzigen Zeitpunkt innerhalb von DARIAH-DE mit Datenstrukturen beschäftigen (d.h. Fächer mit teils sehr spezifischer thematischer und methodischer Ausrichtung; vgl. dazu Punkt 1.1 der Einleitung). Daher sind diese *Essentials* als ein Ausgangspunkt für weitere Diskussionen und Ergänzungen zu verstehen. Eine Erweiterung des Fächerspektrums ist geplant, so dass in den kommenden Monaten aus disziplinärer Perspektive umfangreichere und breiter gefächerte Empfehlungen gegeben werden können.

2. In diesem Papier sind zunächst die *Essentials* in grob strukturierter Form als fortlaufender Text dargestellt. Um den Charakter der Empfehlungen als Orientierungshilfe für die Einrichtung neuer Projekte bzw. zur Anregung der Diskussion über Metadaten in bereits bestehenden Projekten besser zur Geltung zu bringen, wird dieser Text in naher Zukunft in die webbasierte Wikistruktur von DARIAH-DE überführt. Dort soll dieser Report als ein *living document* nicht nur von den jetzigen Verfassern, sondern durch die Beteiligung der gesamten DARIAH-Community und auch weiterer Experten und Spezialisten aus verschiedenen geisteswissenschaftlichen Forschungs- und Digitalisierungsprojekten weiter gepflegt, überarbeitet und durch zusätzliche fachspezifische oder fächerübergreifende Beispiele angereichert werden.

Abschließend ist darauf hinzuweisen, dass dieses Papier bis Mitte 2013 um *Informationen und Essentials für fachspezifische Empfehlungen für die Verwendung von Daten* ergänzt wird. Einige der Themen und Aspekte, die diesbezüglich im hier

¹ <http://www.wissenschaftsrat.de/download/archiv/2359-12.pdf>

vorliegenden Report angesprochen wurden, werden dann weiter ausgeführt; gleiches gilt für die von DARIAH-DE AP 3.3 zu erarbeitenden fachübergreifenden Empfehlungen für die Verwendung von Daten- und Metadatenformaten. Allen Reports ist gemein, dass DARIAH-DE das Ziel verfolgt, den einzelnen geisteswissenschaftlichen Forschern Kriterien für einen nachhaltigen Umgang mit Daten und Metadaten an die Hand zu geben und zugleich disziplinären Experten und Fachwissenschaftlern die Möglichkeit einräumen will, ihre Erfahrungen und Expertise in diese *Knowledge-Living-Documents* einzupflegen.

1. Einleitung

1.1 Ziel und Auswahl für die fachspezifischen Empfehlungen

Die nachfolgenden Empfehlungen verstehen sich als eine erste Orientierungshilfe für die Dokumentation und Erfassung sogenannter Metadaten durch Wissenschaftler und Forschungseinrichtungen in der eigenen Institution oder in Forschungsprojekten. Sowohl Interessierten aus den *Digital Humanities* als auch klassisch arbeitenden Geisteswissenschaftlern soll gleichzeitig ein Leitfaden zum Einstieg in die Thematik der Metadaten in den Geisteswissenschaften geboten werden.

In diesem Report werden dabei einige fachspezifische Standards vorgestellt und thematisiert: So u.a. aus der Archäologie, der Geschichtswissenschaft, der Musikwissenschaft, der Epigraphik, der Judaistik und der Hebraistik. Der Fokus liegt hierbei auf der Vermittlung von Empfehlungen für ausgewählte Aspekte deskriptiver Metadaten. In einigen geisteswissenschaftlichen Disziplinen wurden in den letzten Jahren fachspezifische Standards erarbeitet, die häufig dieselben Konzepte behandeln, z.B. zur Datierung, der Beschreibung/Lokalisierung von Orten, Ereignissen o.ä. Abhängig vom jeweiligen Forschungsinteresse und disziplinären Konventionen (aber auch von der Tatsache, ob es sich um eher forschende oder bewahrende Institutionen wie etwa Museen oder Archive handelt) werden die mit diesen Konzepten verbundenen Inhalte oft unterschiedlich erfasst und lassen sich nicht ohne weiteres aufeinander abbilden. Aus diesem Grund werden nachfolgend auch einige daran anschließende Fragen thematisiert und diskutiert: Welche Anforderungen stellen Geisteswissenschaften an Metadatenmodelle? Welche Probleme ergeben sich bei der Datenerfassung? Wie kann – trotz unterschiedlicher methodischer Herangehensweise – die Interoperabilität dieser erfassten Metadaten gewährleistet werden? Damit geisteswissenschaftliche Forschung bei digitaler Erfassung von Forschungsdaten auch Raum für empirische Vorgehensweisen schafft, sollten Metadaten nicht nur zu den Objekten, sondern ebenso zum Entstehungsprozess der Forschungsdaten erhoben werden. Welchen Stellenwert haben diese Aspekte?

1.2 Begriffsklärung

Um die Frage nach der Verwendung von Metadaten im Zusammenhang mit Forschungsdaten in den Geisteswissenschaften klären zu können, bedarf es einer Definition dessen, was ein Standard ist, was Metadaten sind und wie Standards und Metadatenformate zusammenhängen.

Ein **Standard** ist gekennzeichnet durch ein einheitliches bzw. vereinheitlichtes Vorgehen, das dokumentiert, allgemein anerkannt ist und angewandt wird.

Hinweis: Standards, die nur von einer einzelnen Institution oder in einem einzelnen Forschungsprojekt genutzt werden, bieten zwar bereits den Vorteil der konsistenten Datenerfassung, jedoch ist zu empfehlen, vor der Konzeption eines eigenen Standards oder Metadatenmodells bereits existierende zu recherchieren und sich gegebenenfalls an diesen zu orientieren bzw. diese gemeinsam weiter zu entwickeln.

Metadaten werden gemeinhin als "Daten über Daten" beschrieben. Es handelt sich dabei um (strukturierte)² Daten zu (digitalen) Ressourcen oder Objekten. Metadaten entstehen auf verschiedenen Erfassungs- und Informationsebenen. Sie beschreiben Inhalt und Form von Objekten, aber auch weitere Merkmale, wie z.B. Zugangsbedingungen, die für die Wiedergabe einer Ressource und deren langfristige Speicherung erforderlich sind. Als dokumentarische Bezugseinheit können Sammlungen, Einzelobjekte oder auch ein einzelnes Faktum dienen. Je nach Funktion ist es sinnvoll, die unterschiedlichen Metadatentypen noch weiter zu unterteilen. Die Gruppierung in deskriptive, strukturelle und administrative Metadaten bildet dabei in der Fachdiskussion eine geläufige Einteilungsvariante. Deskriptive Metadaten beinhalten inhaltsbezogene Informationen, die zum Auffinden und für die Deutung einer Ressource benötigt werden. Ihre Inhaltsstruktur wird mit Hilfe struktureller Metadaten beschrieben. Administrative Metadaten, zu denen etwa Urheber- und Verwertungsnachweise zählen, dienen schließlich dazu, die Verwaltung und weitere Verarbeitung einer Ressource zu dokumentieren.

In der dokumentarischen Praxis entstehen Metadaten bei der Katalogisierung von Ressourcen. Diese erfolgt nach formalen und sachlichen Aspekten unter Heranziehung spezifischer Regelwerke und (un)kontrollierter Vokabularien, wie z.B. Klassifikationen, Thesauri oder Normdateien.

Metadatenformate sind in der Regel an Standards gebunden und spezifizieren Datenelemente und Datenwerte, um die strukturierte Erfassung von Metadaten zu

² Von strukturierten Metadaten kann gesprochen werden, wenn diese nach Vorgaben eines Standards erfasst bzw. erhoben wurden.

ermöglichen, damit sie maschineninterpretierbar sind. Der Grad der Formalisierung ist entscheidend für die Klassifizierung der Metadatenformate, sei es als einfache Taxonomie mit lediglich hierarchischer Gliederung von Informationen, sei es als Ontologie³, die auch komplexere Beziehungen von Informationen darstellbar macht.

1.3. Metadaten in den Geisteswissenschaften – Vorteile der Verwendung von Standards und mögliche Nutzungsszenarien

In der geisteswissenschaftlichen Forschung ist die Erstellung von Metadaten Teil des Forschungsprozesses, um z.B. Forschungsquellen zu beschreiben, die als Quellengrundlage für die Beantwortung der Forschungsfragen dienen. Die Quellenkritik erzeugt Metadaten, die Auskunft über Provenienz, Glaubwürdigkeit und Aussagekraft einer Quelle geben. Diese im Forschungsprozess anfallenden Metadaten strukturiert und systematisch bereits beim Entstehen zu erfassen, bildet die Grundlage für eine Offenlegung für weitere Nutzergruppen. Die Nachnutzbarkeit von Forschungsdaten wird maßgeblich dadurch beeinflusst, ob diese auffindbar sind und die Bedingungen zur Nutzung der Forschungsdaten kenntlich gemacht sind. Erst die grobe Klassifizierung von Daten/Quellen über Metadaten ermöglicht es, in der Vielfalt verfügbarer Objekte auf vergleichbares Material zu stoßen, das der Beweisführung der eigenen Forschungsthese dienen kann. Um die redundante Erfassung dieser Metadaten zu minimieren, ist zu empfehlen, die jeweils erzeugten Metadaten über die eigene Institution hinaus ortsunabhängig, ohne Zugangsbeschränkungen, in digitalisierter Form und maschinenauswertbar zur Verfügung zu stellen. Dabei stellt aufgrund der Vielzahl verwendbarer Formate die Interoperabilität von Metadaten zum jetzigen Zeitpunkt eine (technische) Herausforderung dar, die durch den Einsatz von internationalen Standards und Austauschformaten bisher am vielversprechendsten zu lösen scheint.

Jedoch zeigt die Konvertierung der eigenen Metadaten in diese Standards und Austauschformate einige Konsequenzen. Die Anpassung der Metadaten (Crosswalk, Harmonisierung) ist mit einem arbeitstechnischen Mehraufwand verbunden. In jedem Falle müssen Inkonsistenzen in den Altdaten zuvor beseitigt werden. Geht der Weg hin zu einem datenärmeren Format, beispielsweise Dublin Core⁴, muss abgesichert werden, dass nicht für den Gegenstand wichtige Angaben verloren gehen. Wird umgekehrt in ein datenreicheres Format konvertiert, beispielsweise TEI, sind gegebenenfalls fehlende Pflichtangaben vorab zu ergänzen. Codebeispiel 1 zeigt die relativ begrenzten Möglichkeiten der Metadatenvergabe mit Hilfe des Dublin Core Metadata Element Set⁵, in

³ Siehe auch DARIAH-DE Report R 2.2.1

⁴ <http://dublincore.org/>

⁵ für weiterführende Informationen zu den einzelnen Elementen des DC Standards siehe: <http://dublincore.org/documents/dces/>

dem 15 verschiedene Metadaten-elemente angegeben werden können.

```
<html xmlns="http://www.w3.org/1999/xhtml">
  <head>
    <meta content="text/html; charset=utf-8" http-equiv="Content-Type" />
    <title>Carl Maria von Weber für Elise Vigitill [Stammbuchblatt].
    Nürnberg, 10. September 1792
    </title>
    <link rel="schema.DC" href="http://purl.org/dc/elements/1.1/" />
    <link rel="schema.DCTERMS" href="http://purl.org/dc/terms/" />
    <meta name="DC.creator" content="Peter Stadler" />
    <meta name="DC.date" content="2011-05-31T16:53:00+01:00"
    scheme="DCTERMS.W3CDTF" />
    <meta name="DC.publisher" content="Carl-Maria-von-Weber-Gesamtausgabe" />
    <meta name="DC.type" content="Text" scheme="DCTERMS.DCMIType" />
    <meta name="DC.format" content="text/html" scheme="DCTERMS.IMT" />
    <meta name="DC.contributor" content="Weber, Carl Maria von" />
    <meta name="DC.contributor" content="Veit, Joachim" />
    <meta name="DC.identifier"
    content="https://euryanthe.de/wega/de/A002068/Korrespondenz/A040043"
    scheme="DCTERMS.URI" />
    <meta name="DC.title" content="Carl Maria von Weber für Elise Vigitill
    [Stammbuchblatt]. Nürnberg, 10. September 1792" />
    <meta name="description" content="Stammbucheintrag für Elise Vigitill, &#34;im
    6ten Jahr seines Alters&#34;; (Blatt mit Bleistift vorliniert);" lang="de" />
    <meta name="DC.subject" content="ego-document" />
    <meta name="DC.language" content="de" scheme="DCTERMS.RFC3066" />
    <meta name="DC.rights" content="All rights reserved" />
  </head>
  <body/>
</html>
```

Codebeispiel 1: Metadatenangabe mit Dublin Core (mit weitgehend selbsterklärenden Attributnamen)

Mitarbeiter müssen zugleich in der Anwendung des neuen Standards geschult werden. Gerade dieser Aspekt ist nicht zu unterschätzen, denn primär sind Wissenschaftler darauf fokussiert, Daten für die *eigene* Forschung zu erheben. Eine Nachnutzung bereits existierender Forschungsdaten ist in den Geisteswissenschaften bisher nicht üblich. Gegenwärtig arbeiten vorrangig Museen, Bibliotheken und Archive an der Entwicklung bzw. Weiterentwicklung von Standards. Durch die Standardisierung sollen hier Informationen für externe Nutzer bereit gestellt werden, die Gegenstände leichter wiederauffindbar machen, wodurch vor allem dem Servicegedanken Rechnung getragen wird, Zugänge zu vereinfachen. Forschungsrelevante Informationen werden in diesen Standards zwar auch erfasst, jedoch wird die Verfeinerung der Datenstrukturen nicht gezielt nach Forschungsinteressen fortgeführt.

Ein weiteres Problem für den Nutzer stellt die laufende Aktualisierung von Standards und

Formaten dar, die durch veränderte Forschungsmethoden nötig wird. Veränderungen in Standards führen immer auch zur Notwendigkeit der Überarbeitung von Altdaten, damit diese mit neu erzeugten vergleichbar bleiben.

Standards werden in der Regel anwendungsorientiert entwickelt. Erstrebenswert ist deshalb meist nur der gegenseitige Austausch jener Ausschnitte aus Datenstrukturen, bei denen Überschneidungen im Arbeitsprozess bzw. der Forschungsmethode auftreten, d.h. allein das Filtern und Mappen von **relevanten** Daten auf den eigenen Standard bzw. das Metadatenformat als Arbeitsgrundlage ist von Interesse. DARIAH-DE entwickelt eine sog. *Schema Registry*, die es Forschern ermöglichen wird, mittels *Crosswalks* Daten aus verschiedenen Quellformaten in ein gewünschtes neues Format zu konvertieren. Voraussetzung für solche Crosswalks sind jedoch definierte Standards für Forschungsdatensammlungen.

Die Erfassung und Erschließung von geisteswissenschaftlichen Forschungsdaten und deren digitale Bereitstellung und Anreicherung mit Norm- und Metadaten – die technische Grundbedingung einer interoperablen Nachnutzung – spielten in der Vergangenheit für geisteswissenschaftliche Forschungsprojekte nur eine periphere Rolle. Insbesondere geisteswissenschaftliche Forschungseinrichtungen – ob außeruniversitär oder universitär –, aber auch größere Forschungsprojekte und -verbände, wie sie beispielsweise in den letzten Jahren durch die Exzellenzinitiative⁶ oder auch durch die Sonderforschungsbereiche (SFB) der DFG in Deutschland gefördert wurden, zeigten in diesem Bereich der Standardisierung kaum nennenswerte Aktivitäten. Dies lag und liegt einerseits daran, dass oftmals die digitale Erschließung und Metadatenverzeichnung von Forschungsdaten nicht Bestandteil von Forschungsbewilligungen waren und forschungszentrierte Projekte fast ausschließlich an ihren unmittelbaren Forschungsergebnissen gemessen wurden und werden. Zudem – und dies ist ein zentraler Unterschied zu den Sozialwissenschaften – werden geisteswissenschaftliche Forschungsdaten entweder nur in bestimmten Disziplinen, wie z.B. der Archäologie, durch Wissenschaftler selbst erhoben, oder im Rahmen von Langzeitprojekten – diese sind fast ausschließlich an Akademien angesiedelt –, die sich auf ausgewählte Quellenkorpora und Forschungsdaten-Collections fokussieren. Darüber hinaus wird die Erfassung von Forschungsdaten per se nicht als integraler Bestandteil des wissenschaftlichen Forschungsprozesses betrachtet und es besteht in verschiedenen geisteswissenschaftlichen Disziplinen – in unterschiedlicher Dimension – zurzeit nur eine geringe Akzeptanz, solche Arbeiten und Leistungen, die gerade im Rahmen von klassischen Forschungsprojekten entstehen könnten, als wissenschaftliche Leistung von Einzelforschern anzuerkennen. Dies sind Aspekte, die sich in den kommenden Jahren rasant verändern müssen, da die Forschungsförderer in der Bundesrepublik Deutschland,

⁶ <http://www.dfg.de/foerderung/programme/exzellenzinitiative/>

aber auch z.B. die Europäische Kommission, durch eine sich verändernde Forschungspolitik hierauf bereits jetzt reagieren.⁷ Gerade die Erfassung von Forschungsdaten, deren Anreicherung mit Metadaten und der interoperable und nicht auf einzelne Forschungsfragen und -kontexte fokussierte Zweitzugriff und die Nachnutzung durch andere Geisteswissenschaftler bilden letztlich eine der zentralen Säulen von Verfahren und Methoden der Digital Humanities.⁸

Wie bereits zuvor angesprochen, hat die hier vorliegende Studie nicht das Ziel, *eine* verbindliche und unbedingt einzuhaltende Nutzungsempfehlung bzgl. der Verwendung von Metadatenstandards bzw. von Normdaten zu geben. Sie versteht sich vielmehr als Evaluation bestehender Entwicklungen der vergangenen Jahre in den unterschiedlichen geisteswissenschaftlichen Fachdisziplinen und möchte zugleich geisteswissenschaftlich Forschende bei der Wahl von spezifischen disziplinären oder generisch interdisziplinären Metadatenformaten unterstützen.

2. Standards und Metadatenformate

In einigen geisteswissenschaftlichen Fachdisziplinen haben sich in den vergangenen Jahren fachspezifische nationale und international angewandte Standards herausgebildet, mit denen bestimmte Kategorien von Forschungsdaten, Forschungsobjekten oder auch Forschungsprozessen beschrieben werden können. Hierzu zählen beispielsweise die Archäologie und die Musikwissenschaft. Auch ist festzustellen, dass sich bei einer Vielzahl von Editionsprojekten – unabhängig von ihrer disziplinären Verortung – solche Standardisierungsprozesse herausgebildet haben. So sind hier an erster Stelle die Text Encoding Initiative (TEI)⁹ mit einem internationalen Spektrum oder auch das deutsche Projekt TextGrid¹⁰ zu nennen, die maßgeblich zu einheitlicheren Verfahren beitragen. In anderen geisteswissenschaftlichen Disziplinen, so unter anderem der Geschichtswissenschaft und der Judaistik, die überwiegend text- und quellenorientiert arbeiten und eine Vielzahl höchst divergierender methodischer Zugriffe besitzen, lassen sich solche Entwicklungen für die Gesamtheit einer einzelnen Disziplin nicht feststellen.

⁷ Vgl. hierzu *Horizon 2020*, http://ec.europa.eu/research/horizon2020/index_en.cfm?pg=home&video=none oder auch die *Empfehlung zur Weiterentwicklung der Forschungsinfrastruktur in Deutschland bis 2020*, die im Juli 2012 vom Wissenschaftsrat herausgegeben wurde: <http://www.wissenschaftsrat.de/download/archiv/2359-12.pdf>

⁸ Vgl. hierzu den DARIAH-DE-Report R 2.2.3 vom 31. August 2012. *Erster Bericht über die Verwendung von Verfahren der Digital Humanities in den Geistes- und Kulturwissenschaften*. https://dev2.dariah.eu/wiki/download/attachments/2295542/M223_DH-Verfahren.pdf?version=1&modificationDate=1347378358954

⁹ <http://www.tei-c.org/index.xml>

¹⁰ <http://www.textgrid.de/>

Dies hat verschiedene Gründe. Einerseits verwenden diese Disziplinen oftmals Quellen und Daten, die von Bibliotheken und Archiven bereits mit Metadaten erschlossen wurden, so dass eine spezifisch disziplinäre Entwicklung von Metadatenstandards obsolet erscheint, und andererseits lassen deren vielfältige textorientierte Forschungsfragen und die projektspezifischen Erkenntnisinteressen oftmals eine ausschließliche Erfassung in normierten Metadaten schemata nicht zu. Als weiterer Aspekt ist zu nennen, dass in der Vergangenheit von Seiten der Drittmittelförderer für ausschließlich inhaltlich orientierte und ausgerichtete Forschungsprojekte die Erfassung und Erschließung des verwendeten Quellenmaterials mit standardisierten Metadaten nicht explizit gefördert wurden. Aber auch hier sind in den letzten Jahren beispielsweise durch das Förderprogramm für wissenschaftliche Literaturversorgungs- und Informationssysteme (LIS)¹¹ der Deutschen Forschungsgemeinschaft Empfehlungen entwickelt worden, die maßgeblich zu Standardisierungen und einer interdisziplinären Interoperabilität beigetragen haben. Trotz dieser Einschränkungen ist erkennbar, dass auch bei den zuletzt genannten Disziplinen eine stärkere Verwendung von Normdaten, wie beispielsweise des *Thesaurus of Geographic Names* oder der *Gemeinsamen Normdatei*¹², erkennbar ist und diese Entwicklungen auch auf Dauer unterstützt werden sollten.

Zur Notwendigkeit semantischer Auszeichnungen

Eine Maschine dürfte Schwierigkeiten haben, im folgenden Satz „Hans Hamburger genießt in Paris einen Berliner“ Ortsnamen von Personennamen und populären Speisen zu unterscheiden. Beißt Hans Hamburger in Paris/Frankreich oder in Paris/Texas in die Süßspeise? Hier könnte eine Spezifizierung über den *Getty Thesaurus of Geographic Names* für geographische Präzisierung sorgen, gleiches gilt für Hans Hamburger, denn immerhin gibt es drei Personen dieses Namens mit je eigener ID in der Personennamendatei (PND)¹³ und diese gilt es dann zu spezifizieren. Und um auszuschließen, dass mit „Berliner“ Kennedys „Ik bin ein Berliner“ assoziiert wird, könnte durch Hinweis auf Kategorie 642 „Mahlzeiten und Tischkultur“ der *Deweyschen Decimal Classification* auf die Backware hingewiesen werden.

Eine einfache Suche bei Wikipedia¹⁴ kann ebenfalls die Unklarheiten bzw. Probleme veranschaulichen, die beim Text-Mining und dem Matching auf bestimmte Vokabulare entstehen – zum Beispiel bei der Erkennung von Homonymen; siehe dazu etwa die Begriffsklärung zu „London“ in der deutschsprachigen Wikipedia.¹⁵

¹¹ <http://www.dfg.de/foerderung/programme/infrastruktur/lis/>

¹² http://www.dnb.de/DE/Standardisierung/Normdaten/GND/gnd_node.html

¹³ <http://d-nb.info/gnd/116422556>, <http://d-nb.info/gnd/116422521>,
<http://d-nb.info/gnd/125041055>

¹⁴ <http://de.wikipedia.org>

¹⁵ http://de.wikipedia.org/wiki/London_%28Begriffskl%C3%A4rung%29

Interne und externe Verwaltung von Metadaten

Bevor im Folgenden einzelne disziplinspezifische Aspekte im Umgang mit der Erfassung von Metadaten aufgezeigt werden, sei hier noch auf zwei grundsätzliche Speicherungs- bzw. Verwaltungsmöglichkeiten von Metadaten hingewiesen:

1. Metadaten können **intern** verwaltet werden, wenn sie im Kontext der regulären Daten einer Quelle mit erfasst werden. Ein separater Bereich innerhalb der Datei ist dabei den Metadaten vorbehalten – in der Regel der Kopfbereich (*header*) einer Datei. So verwendet beispielsweise der Musikcodierungsstandard MEI einen vom eigentlichen *body* (der hier den codierten Notentext enthält, vgl. die Bezeichnung *music*) abgesetzten Metadaten-Kopfbereich *meiHead*. Im folgenden Codebeispiel sind die durch entsprechende Angaben noch aufzufüllenden Unterrubriken dieses *meiHead* mit sichtbar, die sich an die fast identischen Strukturen von TEI anlehnen.

```
<mei>
  <meiHead>
    <!-- intern verwaltete Angabe der Metadaten -->
    <altId/>
    <fileDesc/>
    <encodingDesc/>
    <workDesc/>
    <revisionDesc/>
  </meiHead>
  <music> <!-- Notentext --> </music>
</mei>
```

Codebeispiel 2: Verwaltung von Metadaten im MEI-Standard

2. Bei der **externen** Verwaltung von Metadaten wird eine externe Datei mit Metadaten (hier: Codebeispiel 3) parallel zu der Datei (hier: Codebeispiel 4), die die eigentliche Quelle beschreibt bzw. enthält, angelegt.

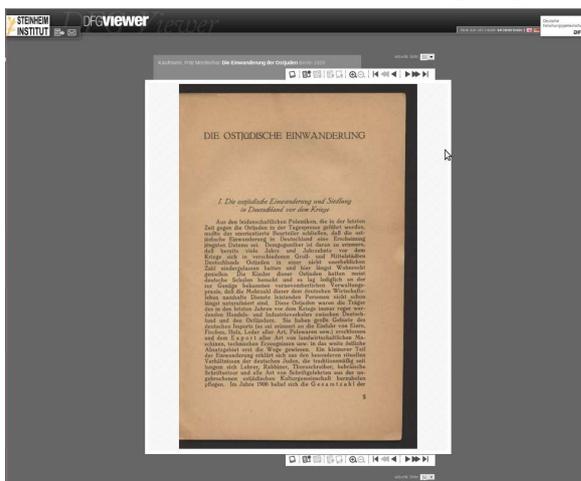


Abbildung 1: Kaufmann, Fritz Mordechai: Die Einwanderung der Ostjuden, Berlin 1920; retrodigitalisierte Printpublikation des STI, angezeigt im DFG-Viewer

Um die Korrespondenz der beiden physikalisch voneinander getrennten Dateien sicher zu stellen, wird in der externen Metadatenfile ein URI angegeben, der auf die eigentliche Quelldatei verweist.

```
<mets:mets xsi:schemaLocation="http://www.loc.gov/mets/mets.xsd">
  <mets:dmdSec ID="md245233015823">
    <mets:mdWrap MIMETYPE="text/xml" MDTYPE="MODS">
      <mets:xmlData>
        <mods version="3.0" xsi:schemaLocation="http://www.loc.gov/mods/v3
          http://www.loc.gov/standards/mods/v3/mods-3-0.xsd">
          <titleInfo>
            <title>Die Einwanderung der Ostjuden</title>
          </titleInfo>
          <name type="personal">
            <displayForm>Kaufmann, Fritz Mordechai</displayForm>
          </name>
          <originInfo>
            <place>
              <placeTerm type="text">Berlin</placeTerm>
            </place>
            <dateIssued>1920</dateIssued>
          </originInfo>
        </mods>
      </mets:xmlData>
    </mets:mdWrap>
  </mets:dmdSec>
  <mets:amdSec ID="amd245233015823">
    <mets:rightsMD ID="rights245233015823">
      <mets:mdWrap MIMETYPE="text/xml" MDTYPE="OTHER" OTHERMDTYPE="DVRIGHTS">
        <mets:xmlData>
          <dv:rights>
            <dv:owner>Steinheim-Institut</dv:owner>
            <dv:ownerLogo>http://www.steinheim-institut.
              org/ressourcen/test/steinheim-institut-logo.png</dv:ownerLogo>
            <dv:ownerSiteURL>http://www.steinheim-institut.de</dv:ownerSiteURL>
          </dv:rights>
        </mets:xmlData>
      </mets:mdWrap>
    </mets:rightsMD>
    <mets:digiprovMD ID="digiprov245233015823">
      <mets:mdWrap MIMETYPE="text/xml" MDTYPE="OTHER" OTHERMDTYPE="DVLINKS">
        <mets:xmlData>
          <dv:links>
            <dv:reference/>
          </dv:links>
        </mets:xmlData>
      </mets:mdWrap>
    </mets:digiprovMD>
  </mets:amdSec>
</mets:mets>
```

Codebeispiel 3: Ausschnitt der zugrundeliegenden METS/MODS-Datei für die Angabe der Metadaten

In obigem Codebeispiel 3 für die externe Angabe der Metadaten hat jede Seite des digitalisierten Buches einen eigenen URI, hier sichtbar im Element `<mets:dmdSec ID="md245233015823">`. Dieser URI verweist auf die eigentliche Quelldatei, in diesem Beispiel also einen Scan der Seite, die oben dargestellt ist. Codebeispiel 4 zeigt diese eigentliche Quelle als Sourcecode, einen Scan im Format jpg in der Auflösung 600.

```

<mets:fileGrp USE="MIN">
  <mets:file MIMETYPE="image/jpeg" ID="245233015823-00000_600">
    <mets:FLocat LOCTYPE="URL" xlink:href="http://www.steinheim-institut.org/
      ressourcen/test/245233015823/600/245233015823-00000_600.jpg"/>
  </mets:file>
  <mets:file MIMETYPE="image/jpeg" ID="245233015823-00001_600">
    <mets:FLocat LOCTYPE="URL"
xlink:href="http://www.steinheim-institut.org/ressourcen/test/245233015823/600/24523301
5823-00001_600.jpg"/>
  </mets:file><mets:file MIMETYPE="image/jpeg" ID="245233015823-00002_600">
    <mets:FLocat LOCTYPE="URL"
xlink:href="http://www.steinheim-institut.org/ressourcen/test/245233015823/600/24523301
5823-00002_600.jpg"/>
  </mets:file>
</mets:fileGrp>

```

Codebeispiel 4: Die eigentliche Quelldatei; Metadatenangabe erfolgt extern (siehe Codebeispiel 3)

Darüberhinaus ist auch die Verwaltung dieser beiden Dateien innerhalb eines Archives möglich. Damit kann also auch die Entscheidung über die Speicherungs- bzw. Verwaltungsmöglichkeit von Metadaten die Auswahl eines Formates beeinflussen.

2.1 Archäologische Standards und Metadatenformate

Die Zuordnung der Archäologie entweder zum kunstwissenschaftlichen oder dem geschichts- und kulturwissenschaftlichen Fachbereich an den Universitäten deutet bereits darauf hin, dass das methodische Vorgehen bzw. die Forschungsschwerpunkte in dieser Wissenschaft breit gefächert sind¹⁶. Man kann zudem nicht von *der* Archäologie sprechen, sondern nur von *den* Archäologien, da eine Vielzahl von Kulturkreisen mit archäologischen Methoden untersucht werden, aber aufgrund ihrer Verschiedenheit jeweils andere Herangehensweisen hierfür erforderlich sind. Beeinflusst durch die Entwicklung der "New Archaeology" in der amerikanischen und britischen prähistorischen Forschung entstanden hochspezialisierte, verwandte Wissenschaften wie die Archäobotanik, Archäozoologie, Bauforschung, Unterwasserarchäologie oder Spatial archaeology, des weiteren fanden Methoden aus der Anthropologie und Soziologie Eingang in den Methodenkanon der Archäologie. Wenn man also über Standardisierung in den Archäologien diskutiert, wird deutlich, dass die Archäologien sich aus Standards anderer Disziplinen bedienen und diese auf ihre Bedürfnisse hin modifizieren.

¹⁶ Einführende Literatur zur Notwendigkeit von Standards in der Archäologie, siehe: Richards, J. D., *From anarchy to good practice: the evolution of standards in archaeological computing*, Archeologia e Calcolatori, 2009, <http://eprints.whiterose.ac.uk/10707/> sowie: Schloen, J., *Archaeological Data Models and Web Publication Using XML*, Computers and the Humanities, 2001, <http://www.springerlink.com/content/jk1q91115506k6j5/abstract/>

In den Archäologien wurden deshalb verschiedene Metadatenstandards entwickelt, die für bestimmte Aufgaben optimiert sind, z. B. für die Grabungsdokumentation (Befund, Funde, Vermessungsdaten etc.) oder Objektbeschreibung. Häufig orientieren sich die zugrundeliegenden Datenmodelle für die Dokumentation an projekt- bzw. institutionenspezifischen Forschungs- und Aufgabenschwerpunkten.

Aufgrund der Öffnung der Archive für die überregionale Forschung und Nachnutzung der Forschungsdaten initiiert durch Förderrichtlinien, die die Digitalisierung und den Austausch von Forschungsdaten fördern, und des Weiteren durch das Outsourcing von Grabungsaufträgen an private Grabungsfirmen, entstand in den letzten Jahren zunehmend der Druck, nationale Lösungen für die Standardisierung bei der Datenerfassung archäologischer Forschung zu finden. Ergebnisse dieser Entwicklung sind Standardisierungsaktivitäten für Austauschformate von Daten und Grabungsberichten wie **ADex**¹⁷ und dem **MIDAS Heritage framework**¹⁸ oder auch der **Dutch Archaeology Quality Standard**¹⁹, der Richtlinien für das Berichtswesen für Grabungsfirmen definiert. Die Landesdenkmalpflegeämter in Deutschland haben vergleichbare Richtlinien für die Dokumentation von Grabungen herausgegeben²⁰:

- LVR-Amt für Bodendenkmalpflege im Rheinland, Prospektions- und Grabungsrichtlinien für drittfinanzierte archäologische Maßnahmen²¹
- Vorgaben zur Dokumentation archäologischer Ausgrabungen in Bayern²²
- Ausgrabungen und Prospektion, Durchführung und Dokumentation, Verband der Landesarchäologen in der Bundesrepublik Deutschland²³
- Brandenburgisches Landesamt für Denkmalpflege und Archäologisches Landesmuseum, Abt. Bodendenkmalpflege, Richtlinien Grabungsdokumentation²⁴
- Landesamt für Denkmalpflege Hessen, Abt. Archäologie und Paläontologie, Richtlinien zur Grabungsdokumentation²⁵
- Landesarchäologie Sachsen, Dokumentation²⁶

¹⁷ <http://www.landesarchaeologen.de/verband/kommissionen/archaeologie-und-informationssysteme/projektarbeitsgruppen/adex/>

¹⁸ <http://www.english-heritage.org.uk/professional/archives-and-collections/nmr/heritage-data/midas-heritage/>

¹⁹ http://www.dans.knaw.nl/sites/default/files/file/archief/KNA_uk.pdf

²⁰ Als grundlegende Literatur für sämtliche Richtlinien in der Grabungsdokumentation, siehe: Jörg Biel, Dieter Klonk (Hrsg.): *Handbuch der Grabungstechnik*, 1994

²¹ http://www.bodendenkmalpflege.lvr.de/denkmalerschutz_prakt_bodendenkmalpflege/richtlinien_2011.pdf

²² http://www.blfd.bayern.de/medien/vorg_doku_arch_ausg.pdf

²³ http://www.landesarchaeologen.de/fileadmin/Dokumente/Dokumente_Kommissionen/Dokumente_Grabungstechniker/grabungsstandards_april_06.pdf

²⁴ <http://www.bldam-brandenburg.de/images/stories/PDF/richtlinien%20zur%20grabungsdokumentation.pdf>

²⁵ <http://hessen-archaeologie.de/Download/LfDHessenGrabungsrichtlinien2005.pdf>

²⁶ <http://www.archaeologie.sachsen.de/59.htm>

Damit liefern diese Richtlinien bereits Standards für den Datenerhebungsprozess der Grabung in den Archäologien. Ihren Niederschlag finden diese Daten bisher in Grabungsberichten, Grabungstagebüchern, Karten von Plana, fotografischer und zeichnerischer Dokumentation, Karten zu geomagnetischer Prospektion u.a., die aber bisher nicht in digitaler Form und zum Schutz der Fundstätten auch nicht ohne Zugriffskontrolle zugänglich gemacht werden können. Weiterhin wird deutlich, dass aufgrund der föderalen Struktur jedes Bundesland eigene Richtlinien herausgibt.

In den USA entwarf David Schloen die Ontologie OCHRE (Online Cultural Research Environment), die für verschiedene altertumswissenschaftliche Projekte zur Datenmodellierung genutzt wird. Ein Teilausschnitt aus dessen Datenstruktur wird im Metadatenformat ArchaeoML²⁷ für das Projekt Open Context²⁸ genutzt. Open Context entwickelt eine webbasierte Open-Access-Plattform für die Organisation, den Zugang zu und die Anreicherung archäologischer Daten, bisher mit dem Fokus auf amerikanische Projekte.

Viele Fundobjekte sind in die Sammlungen von Museen überführt worden und unterliegen hier den Dokumentationsstandards aus der Museologie. In Deutschland erarbeitete die Fachgruppe Dokumentation im Deutschen Museumsbund das Harvestingformat museumdat²⁹, angelehnt an das in den USA vom J. Paul Getty Trust mit ARTstor entwickelte CDWA lite³⁰. In Großbritannien hat sich SPECTRUM³¹ durchgesetzt, das sowohl Richtlinien für die Prozesse im Museumsmanagement bietet, als auch dazu passend ein Metadatenformat aufweist. Unter Leitung von Collection Trust³² wird SPECTRUM weiterentwickelt.

Im Zuge der Internationalisierung von Forschungsprojekten sind nun weitere Standards für Austauschformate auf europäischer und internationaler Ebene definiert worden. Im Projekt CARARE (Connecting ARchaeology and ARchitecture in Europeana) entstand der gleichnamige Standard CARARE³³ für den Datenaustausch zu archäologischen Objekten und Denkmälern mit Europeana. Das Metadatenformat LIDO³⁴ (Lightweight Information Describing Objects) führt auf internationaler Ebene Standards der Museumsdokumentation

²⁷ http://ochre.lib.uchicago.edu/index_files/Page845.htm

²⁸ <http://opencontext.org/>

²⁹ <http://www.museumdat.org/>

³⁰ http://www.getty.edu/research/publications/electronic_publications/cdwa/index.html

³¹ <http://www.collectionslink.org.uk/programmes/spectrum>

³² <http://www.collectionstrust.org.uk/>

³³ <http://carare.eu/ger/Materialien>

³⁴ <http://network.icom.museum/cidoc/working-groups/data-harvesting-and-interchange/what-is-lido.html>

zusammen und ist konform zur Ontologie CIDOC-CRM³⁵ (CIDOC object-oriented Conceptual Reference Model) für Kulturelle Überlieferung, die von der Fachgruppe Dokumentation des Komitees des internationalen Museumsverbandes³⁶ (ICOM) entwickelt wurde.

Vielfach sind in der Archäologie Standards bzw. Richtlinien zur Erfassung von Forschungsdaten und -prozessen bereits vorhanden, diese sind aber noch nicht digital auswertbar, da sie bisher keinen Eingang in Metadatenformate (siehe Grabungsdokumentationsrichtlinien) gefunden haben.

2.2 Musikwissenschaftliche Standards und Metadatenformate

In der Musikwissenschaft werden bislang nur in Teilbereichen computerbasierte Methoden eingesetzt, um aktuellen Forschungsfragen nachzugehen. Die größte Verbreitung erfahren reine Datenbank-Angebote wie die *Virtuelle Fachbibliothek Musikwissenschaft*³⁷, die Informationen aus Bibliothekskatalogen, dem *Internationalen Quellenlexikon der Musik* (RISM, s.u.) und anderen online zugänglichen Quellen bündelt. Während derartige Angebote unzweifelhaft eine große Arbeitserleichterung für Musikwissenschaftler darstellen, bieten sie kaum eigenständige computerbasierte Forschungsperspektiven. Solche werden dagegen vor allem in zwei Bereichen adressiert: Einerseits im weitläufigen Feld des *Music Information Retrieval*, also der automatisierten Analyse von Audiodaten oder digitalen Notentexten, andererseits im Bereich digitaler Musikedition. Der erste dieser Bereiche arbeitet häufig mit relativ detailarmen Daten, etwa im MIDI-Format (s.u.), die eine Beschränkung auf die wesentlichsten Parameter erleichtern. In der Musikphilologie und angrenzenden Bereichen wird dagegen in der Regel mit möglichst detailreichen Daten gearbeitet; grundsätzlich ist hier eine größere Nähe zu anderen digital arbeitenden Geisteswissenschaften bzw. den sogenannten *Digital Humanities* festzustellen. Für musikwissenschaftliche Forschungsarbeiten sind dabei unter anderem folgende Datenobjekte von besonderem Interesse:³⁸

- Notentexte/Notate
- Klang- und/oder Videoaufzeichnungen (Audio/Video)
- Klangerzeuger (Musikinstrumente, Stimme) und Klangerzeugung (Akustik)
- Interpreten, Komponisten, Kopisten, Drucker, Mäzene, Institutionen, Organisationen, Kulturen
- Ikonographien
- Texte zur Geschichtsschreibung, Musiktheorie, -ästhetik, -psychologie, -soziologie

³⁵ <http://www.cidoc-crm.org/>

³⁶ <http://icom.museum/what-we-do/professional-standards/standards-guidelines.html>

³⁷ <http://www.vifamusik.de>

³⁸ Eine detaillierte Auflistung von Quellentypen findet sich im Abschnitt 6.3 dieses Reports.

Während in vielen Bereichen, die sich vom Objekttyp her mit Gegenständen anderer Wissenschaften überschneiden, keine gesonderten Beschreibungssprachen notwendig sind bzw. objektabhängig gängige Systeme entsprechend erweitert werden können, müssen bei Notentexten aufgrund der objektspezifischen Anforderungen für die Digitalisierung und Dokumentation der entsprechenden Metadaten eigene fachspezifische Standards angewendet werden.

Etlliche Teildisziplinen sind dagegen geprägt vom Umgang mit Texten, die sich trotz sehr fachspezifischer Inhalte oder Strukturen gut mit Hilfe des **TEI**-Standards darstellen lassen, der auch in anderen Disziplinen als erste Wahl für die wissenschaftliche Erfassung von Textmaterialien gilt und meist in Kombination mit entsprechenden Normdaten genutzt wird. Nur eine bislang verschwindende, jedoch kontinuierlich wachsende Zahl von Musikwissenschaftlern arbeitet bislang mit TEI; als Regel muss dagegen noch die Verwendung üblicher Textverarbeitungs- oder Datenbankprogramme angesehen werden, öffentlich verwertbare Metadaten fallen dabei meist nicht an bzw. werden nicht bereitgestellt. Eine spezielle Herausforderung stellen in diesem Kontext gemischte Texte (etwa Musiktraktate mit Notenanteilen o. ä.) dar – hier müssen ggf. mehrere Standards kombiniert werden, um die Gegenstände vollständig erfassen zu können.

Musikinstrumente, die meist im Kontext von Museen digital erfasst werden, können ebenfalls mit interdisziplinären Standards beschrieben werden. Das 2009 gegründete MIMO³⁹ (Musical Instrument Museums Online) verwendet zur Auszeichnung von Musikinstrumenten beispielsweise den interdisziplinären Standard **LIDO** (Lightweight Information Describing Objects)⁴⁰, eine XML-Darstellung zur Erfassung verschiedenster musealer Objekte.

Standardisierte Daten spielen eine zentrale Rolle in dem seit 1952 weltweit tätigen musikwissenschaftlichen Katalogisierungsprojekt **RISM** (Répertoire Internationale des Sources Musicales). Die inzwischen online zugänglichen über 700.000 Datensätze⁴¹ werden immer häufiger als Metadaten für die damit verbundenen Digitalisate der Musikalien aus den jeweiligen Bibliotheken benutzt, die dann oft direkt aus dem RISM-Katalog abrufbar sind.⁴²

Für den Austausch und die Analyse akustischer Daten hat sich seit den 1980er Jahren das Binärformat **MIDI** (Musical Instrument Digital Interface) etabliert, das lediglich akustisch relevante Informationen speichert (und damit u.a. eine Berücksichtigung von Enharmonik,

³⁹ <http://www.mimo-international.com/>

⁴⁰ <http://network.icom.museum/cidoc/working-groups/data-harvesting-and-interchange/what-is-lido>

⁴¹ <http://opac.rism.info/index.php>

⁴² Zu Verfahren, die diese Daten zunächst in eigene Strukturen überführen vgl. Anne Graham u. Deborah Pierce, "RISM Data as Metadata for Digital Collections", Paper bei der RISM-Konferenz *Music Documentation in Libraries, Scholarship, and Practice* in Mainz, 4. Juni 2012, zugänglich unter <http://www.rism.info/de/publikationen/konferenz-2012.html>.

die für weiterführende musikanalytische Fragestellungen von zentraler Bedeutung ist, ausschließt). Aufgrund dieser Einschränkungen, die auch den Bereich der Metadaten betreffen, wird MIDI im wissenschaftlichen Bereich beinahe ausschließlich zur eher in die Breite ausgerichteten Analyse großer Datenbestände verwendet. Als Austauschformat digitaler Notensatzdaten wurde es inzwischen weitgehend vom diesbezüglich deutlich mächtigeren MusicXML (s.u.) verdrängt. Ähnliche Einschränkungen wie bei MIDI gelten dagegen auch für **MP3**-Dateien, für die mit **ID3** (Identify an MP3) allerdings ein komplementäres, wenn auch weiterhin rudimentäres Metadatenformat geschaffen wurde. Ausdrücklich für die Beschreibung von Metadaten für jegliche Form von multimedialen Anwendungen wurde dagegen in jüngerer Zeit **MPEG-7** (Multimedia Content Description Interface) entwickelt⁴³. Angesichts der komplexen Situation bei Audiodaten wird hier zunächst auf eine detailliertere Beschreibung verzichtet.

Bei der Verwendung von interdisziplinären Standards im Textbereich müssen für musikwissenschaftliche Vorhaben meist zusätzlich Angaben von (teils) fachspezifischen Informationen in den Schemata-Angaben hinzugefügt werden, wie Verweise auf musikalische Werke, Rollen, bestimmte Quellen, Aufführungen, Aufführungsorte, Aufführungsbesprechungen, Tonarten usw.

Sofern eine solche disziplinspezifische Anpassung erfolgt, können also – abgesehen von Notentexten oder akustischen Daten – viele der im Fokus musikwissenschaftlicher Fragestellungen stehenden Forschungsobjekte in ausreichender Weise mit Hilfe vorhandener interdisziplinärer Standards dargestellt werden. In diesen Bereichen stellen sich daher kaum zusätzliche Anforderungen an die Interoperabilität von Forschungsdaten.

Musikphilologie – eigene fachspezifische Standards für Notentexte⁴⁴

Bevor im Folgenden die Möglichkeiten der Metadatenauszeichnung von Notenmaterial vorgestellt wird, sei hier der explizite Hinweis erlaubt, dass diese Standards primär für die Auszeichnung des eigentlichen Notentextes entwickelt wurden, nicht um eine qualitativ hochwertige und möglichst vollständige Erfassung der zugehörigen Metadaten zu erreichen. Dieser Sachverhalt wird durch die in der Regel knappe Form und die ausschliesslich interne Verwaltung aller erhobenen Metadaten in der gleichen Datei bestärkt. Die Vorstellung und Besprechung der einzelnen Standards für die Erfassung des eigentlichen Notentextes wird in einem Folgebericht⁴⁵ ausführlicher besprochen werden.

Grundsätzlich sind bei den nachfolgend genannten Notationsstandards (die nicht mit der

⁴³ <http://mpeg.chiariglione.org/>

⁴⁴ Zu diesem Thema siehe auch: Eleanor Selfridge-Field (Hrsg.), *Beyond Midi: The Handbook of Musical Codes*, Cambridge 1997; Johannes Kepper, *Musikedition im Zeichen neuer Medien: Historische Entwicklung und gegenwärtige Perspektiven musikalischer Gesamtausgaben*, Norderstedt 2011.

⁴⁵ Siehe R3.2.2, der bis zum Frühjahr 2014 erstellt wird.

Nutzung von Notationsprogrammen verwechselt werden dürfen) zwei Formen zu unterscheiden: ältere, ASCII-basierte Repräsentationsformen, sowie jüngere XML-basierte Auszeichnungssprachen. Viele der älteren Formate wurden gezielt nur für die Erfassung und maschinelle Verarbeitung einstimmiger Phänomene (z.B. für Volkslied-Analyse, Themensuche, Aufspüren thematischer Korrespondenzen usw.) entwickelt; Mehrstimmigkeit wurde zumeist erst in späteren Formaten thematisiert. Allerdings spielt die Unterscheidung der Datengrundlage gerade bei der Beurteilung der Leistungsfähigkeit der Formate im Metadatenbereich eine eher untergeordnete Rolle, lag doch – wie erwähnt – der Schwerpunkt in allen Fällen eindeutig bei der Erfassung des eigentlichen Notentextes.

Um eine bessere Vergleichbarkeit der einzelnen Codierungsbeispiele zu ermöglichen, wurden in jedem Standard die Metadaten von Robert Schumanns Lied *Der Abendstern*, op. 79/1 ausgezeichnet.

MuseData

MuseData⁴⁶ wurde von Walter P. Hewlett als eines der ersten Projekte am 1983 gegründeten Center for Computer Assisted Research in the Humanities, Stanford (CCARH) entwickelt. Die Auszeichnung von Metadaten erfolgt dabei in den ersten zwölf Zeilen des *header* (Record 1–12). Diese sehr begrenzte Möglichkeit der Angabe von Metadaten (die ggf. durch weitere, frei definierbare *Records* ergänzt werden kann) ist in den beiden folgenden Codebeispielen durch Benennung der entsprechenden Elemente erkennbar. Die hier nicht genutzten ersten drei Zeilen/Records können für Angaben des Copyright und weitere identifizierende Informationen genutzt werden. Nachfolgend ist eine Beispielcodierung der Metadaten des Lieds *Der Abendstern* von Robert Schumann abgebildet.

```
Record 1
Record 2

Record 3    09/04/12 K. Herold
Record 5    WK#:79 MV#:1
Record 6    Reprint der Erstausgabe Leipzig 1849
Record 7    Liederalbum für die Jugend
Record 8    Der Abendstern
Record 9    Pianoforte
Record 10   Lied
Record 11   group memberships: score
Record 12   score: part 1 of 28
Record 13   [frei definierbar]
```

Codebeispiel 5: Beispielhafte Auszeichnung der Metadaten in MuseData

⁴⁶ <http://musedata.org/>; nähere Beschreibung in *Beyond Midi* (a.a.O.), S. 402–442, speziell zum *header* lediglich S. 406–408.

Zusätzlich zu den obligatorischen zwölf Feldern des o.g. Beispiels stehen noch weitere Auszeichnungsfelder optional zur Verfügung. Diese sollten auch genutzt werden, denn es fehlt unter anderem die obligatorische Belegung eines Autor/Komponisten-Feldes. Bei den Codierungen des CCARH werden diese Daten in der Regel in den drei ersten Zeilen (im obigen Beispiel leer) abgelegt. Da diese Zeilen, ebenso wie die frei definierbaren ab Record 13 jedoch nicht standardisiert sind, erschwert dies den Austausch zwischen verschiedenen Institutionen, die diese Felder in projektspezifischer Weise nutzen.

Humdrum

Das von David Huron am CCARH entwickelte Humdrum⁴⁷ dient nicht nur der Auszeichnung des Notentextes, sondern bietet mit dem darauf aufsetzenden *Humdrum Toolkit*, einer Sammlung Auswertungs- und Bearbeitungsskripte, auch vielfältige Möglichkeiten musikalischer Analyse. Diese steht konsequent im Fokus des Formats, während die Unterstützung satzspezifischer Fragestellungen bestenfalls rudimentär ist. Das *Humdrum Toolkit* beinhaltet darüber hinaus eine Reihe von Skripten, die Noten aus anderen Dateiformaten (u.a. MuseData, MIDI, Finale, Guido, MusicXML) nach Humdrum konvertieren bzw. sie teils auch in diese Formate exportieren können, so dass das Format zu den bislang attraktivsten im musikwissenschaftlichen Umfeld gehört.

Humdrum unterscheidet für die Angabe zusätzlicher, über die reine Codierung des Notentextes hinausgehender Daten zwischen *Comment Records* und *Reference Records*. *Comment Records* werden mit ein oder zwei Ausrufezeichen gekennzeichnet; ein Ausrufezeichen leitet lokale Kommentare ein, zwei Ausrufezeichen werden dagegen für globale Kommentare genutzt, die beispielsweise die Quelle und den Titel des Stücks beschreiben. Die mit drei vorangestellten Ausrufezeichen markierten Einträge, die als *Reference Records* bezeichnet werden, beschreiben laut Humdrum Website Informationen wie sie üblicherweise in bibliothekarischen Angaben bzw. bei der Katalogisierung begegnen. Als Beispiel folgt wiederum eine Auszeichnung der Metadaten des Lieds *Der Abendstern* von Robert Schumann.

⁴⁷ <http://musiccog.ohio-state.edu/Humdrum/guide02.html>; einen Überblick über die Reference Records gibt: <http://musiccog.ohio-state.edu/Humdrum/guide.append1.html#AIN>; eine kritische Bewertung bei Kepper, a.a.O., S. 324–337.

```

!! Liederalbum für die Jugend
!! Transcribed by Kristin Herold
!! No. 79/1 "Der Abendstern"
**kern      **lyrics
!           in german
...         ...
!!!COM:     Schumann, Robert
!!!CDT:     1810/6/8/-1856/7/29
!!!ODT:     1849
!!!OTL@@GER: Der Abendstern
!!!PUB:     Breitkopf & Härtel
!!!YEC:     Copyright 1991 by Breitkopf & Härtel
!!!AGN:     Lied
!!!AST:     Romantic
!!!AMT:     simple
!!!AIN:     Ipiano Ivovx

```

Codebeispiel 6: Codierung der Metadaten in Humdrum; die Akronyme sind in einer begleitenden Liste festgehalten (hier ab Zeile 6: Composer, Composer's Dates, Date of Composition, Title in German, Publication Status, Date and owner of Electronic Copyright, Genre Designation, Style-period, Metric Classification, Instrumentation)

Humdrum bietet über die im vorstehenden Beispiel sichtbaren einzelnen Zeilen (*spines*) hinaus eine Reihe zusätzlicher Metadatenzeilen für Unterrubriken der Kategorien Authorship, Recording, Performance, Work Identification, Group Information, Imprint, Copyright, Analytic Information, Representation und Electronic Citation. Dies geht weit über die Möglichkeiten von MuseData hinaus.

MusicXML

In MusicXML⁴⁸ sind fünf top-level Elemente für die Auszeichnung von Metadaten vorhanden: *work*, *movement-number*, *movement-title*, *identification*, und *part-list*. Nur das letztgenannte Element *part-list* ist obligatorisch, da es für eine Ordnung und Benennung der nachfolgenden Codierung des musikalischen Inhalts der Notenzeilen verantwortlich ist. Alle anderen Elemente sind optional.

Verglichen mit der Detailliertheit der Codierungen des eigentlichen Notentextes wirkt der header-Bereich von MusicXML sehr eingeschränkt. Dies resultiert aus der primären Zielsetzung des Codierungsformats: Es ist vornehmlich als Austauschformat zwischen verschiedenen proprietären und freien Notensatzprogrammen im Einsatz, nicht jedoch als langfristiges Speicherformat.

⁴⁸ <http://www.makemusic.com/musicxml>

```

<score-partwise version="3.0">
  <work>
    <work-number>op. 79</work-number>
    <work-title>Liederalbum für die Jugend</work-title>
  </work>
  <movement-number>1</movement-number>
  <movement-title>Der Abendstern</movement-title>
  <identification>
    <creator type="composer">Robert Schumann</creator>
    <creator type="lyricist">August Heinrich Hoffmann von Fallersleben</creator>
    <rights>Copyright 2012 Music Encoding Initiative</rights>
    <encoding>
      <encoding-date>2012-08-15</encoding-date>
      <encoder>Kristina Richts</encoder>
      <encoding-description>MusicXML 1.0 example</encoding-description>
    </encoding>
    <source>Based on the reprint of the First Printing Leipzig 1849</source>
  </identification>
  <credit page="1">
    <credit-type>page number</credit-type>
    <credit-words>3</credit-words>
  </credit>
  <credit page="1">
    <credit-type>title</credit-type>
    <credit-words>Kinderlieder. (Hoffmann von Fallersleben.)</credit-words>
  </credit>
  <credit page="1">
    <credit-type>title</credit-type>
    <credit-words>1. Der Abendstern.</credit-words>
  </credit>
  <credit page="1">
    <credit-type>plate number</credit-type>
    <credit-words>8062</credit-words>
  </credit>
  <part-list>
    <score-part id="P1">
      <part-name>Singstimme.</part-name>
    </score-part>
    <score-part id="P2">
      <part-name>Pianoforte.</part-name>
    </score-part>
  </part-list>
  <part id="P1">
    <measure number="1"/>
  </part>
</score-partwise>

```

Codebeispiel 7: Metadaten in MusicXML

MEI⁴⁹

Zur Entstehungsgeschichte des Standards MEI heisst es auf der Homepage der Music Encoding Initiative:

“Seeing the need for a comprehensive markup language in the academic music community, in 1999 Perry Roland of the University of Virginia set about creating an XML schema (DTD) for the representation of music notation. Eventually this DTD became known as MEI because it was influenced by the same principles that guided the creation of the Text Encoding Initiative (TEI).”⁵⁰

Seit 2007 wird die Entwicklung des Formats von einer internationalen Gruppe von Musikwissenschaftlern, -bibliothekaren, Editoren und Informatikern, dem sogenannten MEI Council, gesteuert. Dieses hat im August 2012 eine neue, durch Nutzung des gleichen technischen Unterbaus nun erstmals mit TEI unmittelbar kombinierbare Version **MEI 2012** vorgelegt. Auch daran wird die ausdrückliche (aber nicht ausschließliche) Ausrichtung des Codierungsformats auf den wissenschaftlichen Bereich hin deutlich; zugleich soll mit der modularen Struktur des Formats ein breites Spektrum des schriftlich überlieferten Repertoires sowie wissenschaftlicher Einsatzszenarien abgedeckt werden. Neben dem starken Fokus auf digitale Editionen und entsprechenden Auszeichnungsmöglichkeiten zur formalisierten Erfassung editorischer Sachverhalte ist MEI ausdrücklich auch auf eine musikbibliothekarische Nutzung hin konzipiert. Einen ersten Eindruck der auch in diesem Bereich sehr umfangreichen Möglichkeiten der Metadatenauszeichnung in MEI kann das folgende, aus Gründen besserer Übersichtlichkeit stark gekürzte Beispiel, wiederum zu Schumanns Lied, vermitteln (eine vollständige Version des headers findet sich im Anhang).⁵¹

Die Metadaten werden dabei in einem eigenen Bereich, dem `<meiHead/>` ausgezeichnet, während die Notation der Musik im anschließenden `<music/>` Bereich erfolgt.

⁴⁹ <http://music-encoding.org>

⁵⁰ <http://music-encoding.org/about/history>

⁵¹ Das Beispiel wurde freundlicherweise von Kristina Richts vom DFG/NEH-Projekt *Digital Music Notation Data Model and Prototype Delivery System* zur Verfügung gestellt; zur vollständigeren Wiedergabe vgl. Anhang 6.4.

```

<mei http://www.music-encoding.org/ns/mei" meiversion="2012">
  <meiHead>
    <fileDesc>
      <titleStmt>
        <title>Der Abendstern : an electronic transcription</title>
        <respStmt>
          <resp>Composed by:</resp>
          <persName role="composer" authority="GND" authURI="http://d-nb.info/gnd"
            dbkey="118611666">Robert Schumann</persName>
          <resp>Machine-readable transcription by:</resp>
          <persName role="encoder">Kristina Richts</persName>
        </respStmt>
      </titleStmt>
      <pubStmt>
        <unpub/>
      </pubStmt>
      <sourceDesc>
        <source n="1" xml:id="op79">
          <titleStmt>
            <title type="main">Lieder-Album für die Jugend für Singstimme(n) und
              Klavier op. 79</title>
            <title type="uniform">Lieder für die Jugend</title>
            <respStmt>
              <persName role="composer" authority="GND"
                authURI="http://d-nb.info/gnd"
                dbkey="118611666">Robert Schumann</persName>
              <resp>Herausgegeben von</resp>
              <persName role="editor" authURI="http://d-nb.info/gnd"
                authority="GND" dbkey="10400097X">Ulrich Mahlert</persName>
            </respStmt>
          </titleStmt>
          <editionStmt>
            <edition>Reprint der Erstausgabe Leipzig 1849</edition>
          </editionStmt>
          <pubStmt>
            <respStmt/>
            <address/>
            <identifier type="ordernumber">Breitkopf & Härtel 8307</identifier>
            <availability>
              <useRestrict>© 1991 by Breitkopf & Härtel, Wiesbaden</useRestrict>
            </availability>
          </pubStmt>
          <physDesc/>
          <seriesStmt/>
          <classification/>
          <contents/>
        </source>
      </sourceDesc>
    </fileDesc>
    <workDesc/>
  </meiHead>
  <music>
    <!-- Codierung des eigentlichen Notentextes -->
  </music>
</mei>

```

Codebeispiel 8: Beispiel für die Struktur der Metadaten in MEI

Fazit

Hinsichtlich der Möglichkeiten zur Angabe musikalischer Metadaten bieten nahezu alle genannten Formate nur sehr eingeschränkte Möglichkeiten, die zumeist wenig mehr als eine eindeutige Identifizierung der Dateien erlauben. Dies gilt in unterschiedlichem Maß für die folgenden Formate:

- MuseData
- Humdrum
- MusicXML

Als Standard mit umfassenden Möglichkeiten zur Integration von Metadaten lässt sich bisher nur ein einziges Format benennen:

- MEI

Einige wesentliche Unterschiede zwischen den genannten Formaten im Hinblick auf die Angabe von Metadaten sind in der folgenden Tabelle zusammengetragen. Die Angaben beziehen sich auf die obligatorischen bzw. in den Empfehlungen der jeweiligen Standards zu verwendenden Elementfelder. Zusätzliche frei verfügbare Felder, die mit den jeweils fehlenden Informationen individuell angereichert werden müssten, werden aufgrund der bedingten Übertragbarkeit dieser Informationen hier nicht berücksichtigt.

Kriterium	Muse Data	Hum drum	MusicXML	MEI
Personendaten und Funktionsbezeichnungen	Nennung in spine 1-3 möglich	Nennung	nur Nennung ohne Referenzierungsmöglichkeit <creator/> <encoder/>	mit GND-URI oder anderen Normdateien verlinkbar; allen Personen kann z.B. durch die MARC Code List for Relators eine Funktion zugewiesen werden. <persName role="composer"/> etc.
Publikationsangaben	-	Nennung	wenig detailliert <source/>	detaillierte Angaben möglich <pubStmt/>
Einbettung in den Kontext	-	-	-	möglich, z.B. Inhaltsangaben der Quelle(n), zu dem ein Einzelstück gehört <contents/> oder Reihenangaben <seriesStmt/>

Incipitangabe	-	-	-	Codierung des Incipits von Musik und Text möglich
Physische Beschreibung der Quelle	-	-	sehr geringe Möglichkeiten <credits/>; Spezifizierung über <credit-type/>- Elemente als reine Textfelder, z.B. <credit-type>plate number</credit-type>	sehr ausführlich möglich in der <physDesc/>
Klassifikationsvermerk	-	-	-	Angabe von Klassifikationen, Thesauri etc. und einzelnen darin enthaltenen Termen möglich <classification/>
differenzierte Angabe von Einzelstück und Quellen	Nennung	Nennung	nicht möglich, Angaben beziehen sich ausschliesslich auf Einzelstück	beide Beschreibungen detailliert möglich <workDesc/> (Einzelstück) und <sourceDesc/> (Quellen)
Beschreibungselemente des Werkes	sehr gering	sehr gering	sehr gering <partname/>	sehr umfangreich <workDesc/>

Tabelle 1 : Vergleich der Möglichkeiten musikwissenschaftlicher Metadatenauszeichnung notierter Musik

2.3 Standards und Metadatenformate in der Geschichtswissenschaft

Die Geschichtswissenschaft gehört zu den geisteswissenschaftlichen Fachdisziplinen in denen bislang kaum explizit disziplinäre Metadatenstandards entwickelt wurden. Hierfür sind sicherlich zwei Aspekte von Bedeutung. Einerseits handelt es sich bei der Geschichtswissenschaft um eine Disziplin, die keinen festen und standardisierten Kanon an Forschungsfragen und Erkenntnisinteressen besitzt. Dies hängt in einem starken Maße damit zusammen, dass diese Disziplin oft fachübergreifend und interdisziplinär forscht, um beispielsweise Entwicklungen, Ereignisse und Geschehnisse der Vergangenheit in gänzlich unterschiedlichen Themenkomplexen zu erarbeiten. Das Fach Geschichte und die heutigen Historiker-Communities beschäftigen sich schon lange nicht mehr ausschließlich mit den Fragen, "wer-wann-was-ge-tan" hat bzw. "wann-was-warum-geschah", sondern interdisziplinäre und auch transnationale bzw. interkulturelle Fragestellungen stehen im

Mittelpunkt der derzeitigen Forschungsaktivitäten.⁵²

Andererseits werden für die Forschung höchst unterschiedliche Forschungsdaten und Quellen verwendet. Neben Primär- und Sekundärliteratur verwenden Historiker eine Vielzahl gänzlich unterschiedlicher schriftlicher Quellen; von mittelalterlichen Urkunden und frühneuzeitlichen Druckschriften über Tageszeitungen des 19. Jahrhunderts und Verwaltungsschriftgut bis hin zu Korrespondenz und persönlichen individuellen Schriftüberlieferungen. Zudem werden in einem historischen Forschungsprozess aber auch vielfältige nicht text-basierte Quellen und Forschungsdaten verwendet – die folgende Aufzählung soll hier als Illustration dienen: Statistiken, Karten, Gemälde, Aquarelle, Fotografien, Siegel, Münzen, Plakate, Oral-History-Interviews, Radiosendungen, Noten, Wetter- und Klimadaten, Denkmäler, Fassadenabwicklungen, Inschriften, sonstige bauliche Überreste usw.⁵³ Dieser Facettenreichtum und diese Vielfalt an potentiell Quellen- und Datenmaterial sind der zentrale Grund, weshalb sich disziplinäre Metadatenformate für die Geschichtswissenschaft bislang nicht herausbilden konnten und weshalb es auch zukünftig unwahrscheinlich sein dürfte, dass sich explizit historische Metadatenformate, die mehrere Forschungsdatentypen bzw. -kategorien umfassen, entwickeln werden.

Aus diesen Gründen werden in der Geschichtswissenschaft zum jetzigen Zeitpunkt eine Vielzahl der in diesem Report genannten Formate, Standards und Normdatentypologierungen verwendet, oftmals auch in Mischformen oder auszugsweise. Gebräuchlich ist augenscheinlich am meisten **Dublin Core** (DC) und darüber hinaus das im Kontext des DFG-Viewers⁵⁴ verwendete **METS**-Anwendungsprofil.⁵⁵ Dies liegt sicherlich vor allem an den *Praxisregeln Digitalisierung*⁵⁶, die von der DFG im Rahmen des Förderprogramms *Wissenschaftliche Literaturversorgungs- und Informationssysteme* (LIS) als verbindliche Empfehlung für Digitalisierungsprojekte im Jahre 2009 herausgegeben wurden.⁵⁷ Die Rezeption dieser Empfehlung zeigt zudem sehr deutlich, dass Empfehlungen von Seiten der Forschungsförderer in Deutschland, die auf

⁵² Das Fach Geschichtswissenschaft durchlief nach Ende des Zweiten Weltkrieges einen ungeheuren Wandlungsprozess, so dass sich eine Vielzahl von Sub- und Teildisziplinen entwickelten, zu denen beispielsweise die Medizingeschichte, Technikgeschichte und Mediengeschichte zählen – diese Aufzählung könnte beliebig erweitert werden –, aber auch Teildisziplinen, die sich stark der Methoden und Forschungsverfahren der Sozialwissenschaften bedienen.

⁵³ Für eine detaillierte exemplarische Aufstellung siehe Kap. 6.3 “Zusammenstellung geisteswissenschaftlicher Quellentypen”.

⁵⁴ <http://dfg-viewer.de/ueber-das-projekt/>

⁵⁵ http://dfg-viewer.de/fileadmin/groups/dfgviewer/METS_Anwendungsprofil_2.0.pdf

⁵⁶ http://www.dfg.de/download/pdf/foerderung/programme/lis/praxisregeln_digitalisierung.pdf

⁵⁷ <http://www.dfg.de/foerderung/programme/infrastruktur/lis/>

“Best-Practice-Erfahrungen” und der wissenschaftlichen Expertise von geisteswissenschaftlichen Forschern und wissenschaftlichen Bibliothekaren fußen, zu einem generischen und unmittelbaren Impact bei Standardisierungsbewegungen führen können. Gerade hinsichtlich zukünftiger notwendiger Entwicklungen wären weitere Empfehlungen und Praxisregeln, speziell für Quellen- und Datensammlungen, die sich aus heterogenen Datentypen zusammensetzen, wünschenswert.

Im Bereich der historisch-kritischen Editionen hat sich in den letzten Jahren die Verwendung von XML-ausgezeichnetem Datenmaterial auf Basis von **TEI** durchgesetzt. Eine Entwicklung wie sie in anderen geisteswissenschaftlichen Disziplinen ebenfalls stattfand.⁵⁸ Die Anreicherung der Texte erfolgt zudem in zunehmenden Maße mit Normdaten: Personen-, Orts- und Sachinformationen, meist basierend auf nationalen bzw. internationalen Referenzsystemen wie z.B. der GND und des TGN.

Im Jahre 2004 wurde auf einer Tagung an der LMU München von einer internationalen Gruppe von Wissenschaftlern beschlossen, die **Charters Encoding Initiative (CEI)** zu initiieren, die das Ziel verfolgt, mittelalterliche und frühneuzeitliche Urkunden elektronisch unter Zuhilfenahme von XML auszuzeichnen.⁵⁹ Hieraus hat sich der CEI-Standard entwickelt, der ein exzellentes Beispiel dafür ist, dass Metadatenstandards, die in der Geschichtswissenschaft verwendet werden, nicht ausschließlich einen disziplinären Charakter aufweisen, sondern vielmehr auf spezifische Quellen- und Datentypen angewandt werden. Das nachfolgende Codebeispiel zeigt eine Codierung im CEI-Standard. Weitere Beispiele für mit CEI ausgezeichnete Dokumente sind auf der CEI-Website verfügbar⁶⁰ und nähere Ausführungen finden sich bei dem Unterkapitel zur Diplomatik.

```
<?xml version="1.0" encoding="UTF-8"?>
<cei:text xmlns="http://www.monasterium.net/NS/cei"
  xmlns:cei="http://www.monasterium.net/NS/cei" b_name="St. Dorothea, CanReg"
  id="222137" n="St. Dorothea, CanReg$892" type="charter">
  <cei:front>
    <cei:sourceDesc>
      <cei:sourceDescRegest>
        <cei:bibl>Kartei Stiftsarchiv Klosterneuburg</cei:bibl>
      </cei:sourceDescRegest>
    </cei:sourceDesc>
  </cei:front>
```

⁵⁸ Auf Code-Beispiele wird an dieser Stelle verzichtet, da dies innerhalb des Reports an mehreren Stellen ausführlich geschieht und keine strukturellen, sondern nur graduelle Unterschiede hierzu festzustellen wären.

⁵⁹ <http://www.cei.lmu.de/>

⁶⁰ <http://www.cei.lmu.de/examples/>

```

<cei:body>
  <cei:idno id="1775_IV_08" old="222137">1775 IV 08</cei:idno>
  <cei:chDesc>
    <cei:abstract>Salvatorkirche, Wien (Messenstiftung). Josef Georg Hörl
      Bürgermeister der Stadt Wien, bestätigt, dass, da das Stift St. Dorothea die
      Stiftung der Maria Claudia Demuth nicht angenommen habe, das Stiftungskapital
      von 2000 Gulden auf Wunsch der Universalerbin Maria Anna Demuth der St.
      Salvatorkapelle übergeben wurde.</cei:abstract>
    <cei:issued>
      <cei:placeName>Wien</cei:placeName>
      <cei:dateRange to="17750408" from="17750408">8. April 1775</cei:dateRange>
    </cei:issued>
    <cei:witnessOrig>
      <cei:traditioForm>orig.</cei:traditioForm>
      <cei:archIdentifier>
        <cei:arch>Stiftsarchiv Klosterneuburg
          (http://www.stift-klosterneuburg.at)</cei:arch>
        </cei:archIdentifier>
      <cei:physicalDesc>
        <cei:material>Pergament</cei:material>
      </cei:physicalDesc>
      <cei:auth>
        <cei:sealDesc>
          <cei:seal>Ein anhängendes Siegel.</cei:seal>
          <cei:seal>Siegl'er: Der Bürgermeister.</cei:seal>
        </cei:sealDesc>
      </cei:auth>
    </cei:witnessOrig>
    <cei:witListPar>
      <cei:witness>
        <cei:traditioForm>cop.</cei:traditioForm>
        Handschrift D 80, D, Nr. 34
      </cei:witness>
    </cei:witListPar>
    <cei:diplomaticAnalysis>
      <cei:quoteOriginaldatierung>Den 8. April
        monaths</cei:quoteOriginaldatierung>
      <cei:listBiblEdition>
        <cei:bibl>Quellen zur Geschichte der Stadt Wien, I. Abt., III. Bd.
          (Regesten), 2607, S. 80</cei:bibl>
        </cei:listBiblEdition>
      </cei:diplomaticAnalysis>
      <cei:lang_MOM>Deutsch</cei:lang_MOM>
    </cei:chDesc>
  </cei:body>
  <cei:back>
    <cei:divNotes/>
  </cei:back>
</cei:text>

```

Codebeispiel 9: CEI Beispielcodierung Monasterium

Während die bisher genannten Beispiele vor allem dazu benutzt werden können, um Dokumente und Quellen zu erfassen und zu beschreiben, wird die **Encoded Archival**

Description (EAD) auf Basis von XML für die Beschreibung von Findbüchern aus Archiven, Museen und Bibliotheken verwendet.⁶¹ EAD wurde speziell für die elektronische Erfassung von Findmitteln entwickelt und entstand bereits 1995 an der Universität of California, Berkeley. Daneben existiert für die Beschreibung von Findmitteln noch das **“Standard-Austauschformat” (SAFT)**, das zwischen 2001 und 2004 im Rahmen des DFG-Projektes “Entwicklung von Werkzeugen zur Retrokonversion archivischer Findmittel” entwickelt wurde.⁶²

Festzuhalten bleibt, dass spezifisch disziplinäre geschichtswissenschaftliche Metadatenformate nicht existieren, sondern die Erfassung von Forschungsdaten und Quellen gerade aus geschichtswissenschaftlicher Perspektive stärker Objekttyp- bzw. informationsbezogen betrachtet werden muss und gerade dieser Umstand bei der Verwendung von Metadatenformaten berücksichtigt werden sollte.

2.4 Standards und Metadatenformate in der Judaistik und Hebraistik

Wie in der Geschichtswissenschaft gibt es auch in den Fächern Judaistik/Jüdische Studien sowie der Hebraistik keine fachspezifischen Metadatenstandards. Doch können Judaistik und Jüdische Studien als interdisziplinär angelegte Fächer, die sich sowohl philologischen als auch historischen Forschungsfragen widmen, allgemeingültige Metadatenstandards verwenden.

XML und rechtsläufige Sprachen (hebräisch, arabisch) harmonieren in den gängigen XML-Editoren, wenn überhaupt, nicht optimal. Links/rechts orientiertes XML-Tagset fragmentiert rechtsläufigen Text und führt dazu, dass eine Datei zwar maschinenlesbar ist, der kodierte Text sich aber dem menschlichen Leser kaum mehr erschließt.

Solange dieses Problem nicht grundsätzlich gelöst ist, wird die Akzeptanz von **TEI** und/oder **XML** in Hebraistik und Arabistik gering sein. Gleiches gilt auch für geschichtswissenschaftliche frühneuzeitliche Editionen – so z.B. bei vielfältigen Akademieprojekten –, in denen im zu edierenden Text auch hebräische Begriffe zu finden sind, wie z.B. *Controversia et Confessio*, wengleich die rechts/links-XML-Problematik bei solchen Projekten eher eine untergeordnete Problematik darstellt.

Darüberhinaus unterstützen gängige DH-Tools – etwa XML-Editoren⁶³, Juxta⁶⁴, CollateX⁶⁵

⁶¹ <http://www.loc.gov/ead/>

⁶² <http://www.archivschule.de/content/462.html>

Die Saft XML-tag Library findet sich unter: <https://internet.archivschule.uni-marburg.de/SAFT/doku.php>

⁶³ Am besten visualisiert hier bislang Oxygen <http://www.oxygenxml.com>

⁶⁴ <http://www.juxtasoftware.org>

und Voyant⁶⁶ – right to left Languages gar nicht, oder ungenügend. Diese Problematik soll folgendes Beispiel (Codebeispiel 10) verdeutlichen:

Codiert ist in dem Beispiel zweimal der erste Satz der hebräischen Bibel Gen 1,1: Im ersten Absatz `<p>...</p>` stört lediglich, dass der hebräische Satz hinter dem schließenden `</p>` beginnt und vor dem eröffnenden `<p>` endet.

Unübersichtlich wird es im zweiten Paragraph, in dem das erste Wort und darin der erste Buchstabe ausgezeichnet wurden, hier wird der Lesefluss völlig zerstört. Wie gesagt, das XML ist valide, der Text noch maschinenlesbar, einem menschlichen Leser jedoch wird hier ein hohes Mass an Konzentration abverlangt.

```
<text>
  <body>
    <p>בראשית ברא אלהים את השמים ואת הארץ</p>
    <p><w n="first word of the hebrew bible"><note n="first letter">ב</note>ראשית</w>
      ברא אלהים את השמים ואת הארץ</p>
  </body>
</text>
```

Codebeispiel 10: left-to-right orientiertes xml-Tagset

Etwas besser ist es in der folgenden Darstellung, die das Team der Oxygen-Entwickler für die rtl-community ermöglicht hat ("change-text-orientation"). Nun ist zwar der Text gut zu verfolgen, aber die Tags und deren Attribute kaum mehr.

```
<body>
  <p>בראשית ברא אלהים את השמים ואת הארץ</p>
  <w n="first word of the hebrew bible"><note n="first letter">
    ברא אלהים את השמים ואת הארץ</note></w>
</body>
<text/>
```

Codebeispiel 11: right-to-left orientiertes xml-Tagset

⁶⁵ <http://collatex.sourceforge.net>

⁶⁶ <http://voyant-tools.org>

2.5. Standards und Metadatenformate in geisteswissenschaftlichen Hilfswissenschaften

Epigraphik

In der Epigraphik und Papyrologie liegt mit **EpiDoc**⁶⁷ TEI XML für epigraphische Daten seit geraumer Zeit ein auf TEI P5 beruhender Standard zur sachgerechten Beschreibung epigraphischer und papyrologischer Objekte vor. Dieser Standard wird kontinuierlich weiterentwickelt und gepflegt. Auf der "Markup-list"⁶⁸ diskutiert die Fachcommunity über Probleme epigraphischer Auszeichnung und mögliche Erweiterungen.

Die gültige Version des EpiDoc Schema kann zum Beispiel im Oxygen-XML-Editor über ein RNG-File direkt in die Editionsarbeit eingebunden werden.

```
<?xml version=1.0" encoding="UTF-8"?>  
<?oxygen RNGSchema="http://www.stoa.org/epidoc/schema/8.11/tei-epidoc.rng" type=2xml"?>
```

Codebeispiel 12: EpiDoc Schema Einbindung

Diplomatik

Für mittelalterliche und frühneuzeitliche Urkunden erarbeitet die **Charters Encoding Initiative (CEI)**⁶⁹ ein Vokabular, das Konzepte der Diplomatik für die XML-Erschließung von mittelalterlichen und frühneuzeitlichen Urkunden bereitstellt. Die Tag Library⁷⁰ beinhaltet eine vollständige Liste der verfügbaren Elemente. Das darauf aufbauende XML-Schema⁷¹ baut noch auf TEI P4 auf und kann direkt in allen XML-Editoren verwendet werden. Die CEI diskutiert ihre Arbeit in der Mailingliste cei-l@lists.lrz-muenchen.de. Eine technische Integration mit der Semantic-Web-Version des *Vocabulaire Internationale de Diplomatique*⁷² und eine TEI P5-Implementation des Vokabulars sind in Arbeit.

Kodikologie

Für Handschriften sind im deutschen Raum die Richtlinien der Deutschen Forschungsgemeinschaft (DFG) einschlägig.⁷³ Auf der Grundlage dieser Richtlinien wurde für die Darstellung in XML **Manuscriptum XML (MXML)** entwickelt.⁷⁴ Über den

⁶⁷ <http://epidoc.sourceforge.net/>

⁶⁸ <http://lsv.uky.edu/archives/markup.html>

⁶⁹ <http://www.cei.lmu.de/>

⁷⁰ <http://www.cei.lmu.de/taglib>

⁷¹ <http://www.cei.lmu.de/schema/cei060122.xsd>

⁷² <http://www.cei.lmu.de/VID>

⁷³ Vergleiche: <http://www.manuscripta-mediaevalia.de/hs/kataloge/HSKRICH.htm>

⁷⁴ <http://www.manuscripta-mediaevalia.de/hs/handbuch.pdf>

deutschsprachigen Raum hinaus ist ferner das Modul *Manuscript Description* der Text Encoding Initiative zu berücksichtigen.⁷⁵

3. Geisteswissenschaftliche Grundkonzepte

Um Interoperabilität von Forschungsdaten diskutieren zu können, für die die Standardisierung Voraussetzung ist, soll im folgenden auf einzelne Grundkonzepte eingegangen werden, die in den meisten Geisteswissenschaften gleichermaßen anzutreffen sind und die bereits in unterschiedlichen Forschungskontexten verwendet werden. Sie definieren eine gemeinsame "Schnittmenge" aller Geisteswissenschaften und damit auch Teilbereiche, in denen gemeinsame Forschungsinteressen disziplinübergreifend entstehen können. Zentrale Fragen sind dabei folgende: Wie verwenden die unterschiedlichen geisteswissenschaftlichen Disziplinen Konzepte zur Datierung, Ortsbeschreibung, Objektbeschreibung oder Informationen zu Personen und Institutionen, welche interdisziplinären Schnittmengen lassen sich erkennen?

Zudem ist festzustellen, dass ein- und dieselbe Quelle bzw. die gleiche Forschungsdatensammlung aus der Perspektive unterschiedlicher geisteswissenschaftlicher Fachdisziplinen unterschiedlich beschrieben werden können und daher zu einer unterschiedlichen Anwendung identischer Metadaten schemata führen können. Das nachfolgende Beispiel soll deshalb verdeutlichen, dass die inhaltliche Erfassung von Quellen und Forschungsdaten oftmals nicht einheitlich erfolgen kann, sondern sowohl forschungsfragengesteuert als auch abhängig vom geisteswissenschaftlichen Forschenden und dessen Forschungsperspektive ist. Zudem ist eine starke Abhängigkeit zwischen dem individuellen Forschungskontext und der Erfassung der Forschungsdaten erkennbar, da der Fokus des Bearbeiters nicht per se auf der Erschließung des Datenmaterials liegt, sondern vielmehr dies nur einen einzelnen Schritt der Analyse darstellt. Dieses Beispiel illustriert damit, dass die Vergabe von Metadaten immer im Kontext eines wissenschaftlichen Forschungs-Workflows zu sehen ist.

Die folgende Übersicht zeigt die unterschiedliche Schwerpunktsetzung verschiedener Disziplinen – hier Geschichtswissenschaft/Kirchengeschichte und Judaistik – in Bezug auf das gleiche zu erfassende Ausgangsmaterial. Die differierenden Forschungsinteressen- und -thematiken können sich dann auch in der Vergabe der Metadaten widerspiegeln, was in den folgenden Unterabschnitten dieses Gliederungspunktes weiter ausgeführt wird.

⁷⁵ <http://www.tei-c.org/release/doc/tei-p5-doc/en/html/MS.html>;

Vergleiche auch: http://www.hab.de/bibliothek/wdb/master/doku/doku_msDesc.htm

Ausgangsmaterial	Themen und Forschungsfragen der Geschichte/Kirchengeschichte	Themen und Forschungsfragen der Judaistik
Universal Kirchenzeitung ⁷⁶	<ol style="list-style-type: none"> 1. Kirchenasyl, Ev. Gefängnisgesellschaften und Strafvollzug im 19. Jahrhundert. 2. "Europa aus der Sicht von Außen"– Die Wahrnehmung des christlichen Europagedankens aus Sicht von Mission, Missionaren und Missionierten. 3. Kirchliches Publikationswesen zwischen Vormoderne und Moderne 	<ol style="list-style-type: none"> 1. Jüdische Gelehrte im 18. Jh. 2. Strömungen im europäischen Judentum des 18. Jhs. 3. Missionierung der Juden im 19. Jh.
Fotografie 20. Jh: Nationalsozialismus/ Antisemitismus/ Reichspogromnacht 9. November 1938 ⁷⁷	<ol style="list-style-type: none"> 1. Die Reichspogromnacht 1938 in Südwestdeutschland und die Enteignung der deutschen jüdischen Gemeinden 2. Verfolgung, Antisemitismus, Holocaust und Völkermord im Blick der Volksgemeinschaft, 1930 und 1945. 3. Von Kodak zur Leica - Fotografie, Fototechnik und Fotografen zwischen Kaiserreich und Bundesrepublik 	<ol style="list-style-type: none"> 1. Folgen der Reichspogromnacht für Juden in Baden-Württemberg 2. Öffentliche Reaktion auf die öffentliche Verfolgung jüdischer Bürger 3. Rekonstruktion individueller jüdischer Schicksale

Tabelle 2: Themen und Forschungsfragen von Geschichte/Theologie und Judaistik in Bezug auf das selbe Forschungsobjekt

3.1 Datierung

Die Datierung von Forschungsdaten erfordert, dass neben den Datumsangaben noch weitere Informationen angegeben werden, damit eine eindeutige Zuordnung erfolgen kann. Quellen und Forschungsdaten, die im Rahmen von geisteswissenschaftlichen Forschungsprojekten analysiert werden, können zu einem Zeitpunkt entstanden sein, bei dem ein anderes Kalenderformat als Referenzsystem galt. Insofern müssen neben den eigentlichen Datumsangaben auch Informationen über das Kalenderformat hinterlegt werden, wie z.B. ob es sich um einen Julianischen- oder Gregorianischen-Kalender handelte, oder die Daten in einer Epoche entstanden, in der der Französische

⁷⁶ https://textgridlab.org/dariahdav/wiki/AP3-Forschungsdaten/AP3_4/Daten_STI/Uni_kir_zei_076_bilddateien.pdf

⁷⁷ https://textgridlab.org/dariahdav/wiki/AP3-Forschungsdaten/AP3_4/Daten_IEG/Bundesarchiv_Bild_183-86686-0008,_Baden-Baden,_Festnahme_von_Juden.jpg

Revolutionskalender oder auch der Russische Revolutionskalender galt.⁷⁸ Nur wenn diese Informationen mit angegeben werden, können hierbei auch eindeutige und kalendarisch richtige Vernetzungen und Kontextualisierungen erreicht werden, so dass eine Vergleichbarkeit in der Analyse des Datenmaterials gewährleistet ist. Dies sind Aspekte, die bislang vielfach in vorhandenen Metadatenstandards bzw. -schemata nicht abgebildet werden können.

Vielfältige Formen zur Erfassung bestimmter Zeitabschnitte (“von bis”) oder für in einer Richtung der Zeitachse unbestimmte Zeiträume (“vor”, “nach”, “nicht vor”, “nicht nach”) sind innerhalb der verschiedenen Codierungsformate entwickelt worden. Dagegen erfolgt die simple Datumsangabe inzwischen meist formal einheitlich nach dem etablierten **ISO 8601** Standard (“Date and Time Formats”; der Standard wurde so 1992 auch in Europa in der EN 28601:1922 übernommen und folgt bei Einzeldaten dem Schema: JJJJ-MM-TT hh:mm:ss)⁷⁹. Diesen Standard hat die Library of Congress jüngst durch das Extended Date/Time Format (**EDTF**) 1.0⁸⁰ erweitert und dabei andererseits einige redundante Angaben des ISO 8601 beseitigt.

Auch der Umgang mit und die Benennung von Epochenzuweisungen kann sich als sehr problematisch herausstellen. Das Bibliotheksservice-Zentrum Baden-Württemberg⁸¹ hat beispielsweise Thesauri für Stilepochen⁸² und Zeiträume entwickelt.

3.2 Ortsbeschreibung und Georeferenzierung

Die Standardisierung von Ortsbeschreibungen in den Geisteswissenschaften unterliegt mehreren Problemstellungen. So verändern sich Ortsnamen über die Zeit, durch Sprachwandel oder Umbenennung. Orte werden unterschiedlichen administrativ-politischen Verbänden zugewiesen, aufgrund von Verschiebungen der Kulturgrenzen oder Änderungen in der Staatsform und -organisation. Eindeutig kann eine Ortsbeschreibung nur über ein Zuordnungssystem mit Geokoordinaten erfolgen. Aus Sicht der Quellengrundlage stellen verschiedene Schreibweisen und Sprachvarianten von Ortsbezeichnungen in historischen Quellen eine Herausforderung für die maschinelle Auswertung da. Gerade auch in der Archäologie muss ein Fundort keine spezifische Ortsbezeichnung in Form eines Namens aufweisen, sondern lässt sich mitunter nur über Geokoordinaten erfassen. Die Darstellung von Georeferenzen durch topografische und

⁷⁸ Siehe für einen Überblick zu den unterschiedlichen Kalenderformaten:

http://de.wikipedia.org/wiki/Liste_der_Kalendersysteme

⁷⁹ Siehe dazu z.B. die Zeitangaben in den TEI-Beispielen in Codebeispiel 20 und 21

⁸⁰ <http://www.loc.gov/standards/datetime/pre-submission.html>

⁸¹ <https://wiki.bsz-bw.de/doku.php?id=mare-team:museums-archivsysteme:regelwerke:start>

⁸² [https://wiki.bsz-bw.de/lib/exe/fetch.php?media=m-team:musis:museumsvokabular:](https://wiki.bsz-bw.de/lib/exe/fetch.php?media=m-team:musis:museumsvokabular:infosheet_stilepochen_mai2011.pdf)

[infosheet_stilepochen_mai2011.pdf](https://wiki.bsz-bw.de/lib/exe/fetch.php?media=m-team:musis:museumsvokabular:infosheet_stilepochen_mai2011.pdf)

thematische Karten ist eine weitere Variante, die gerade bei historischen Karten keiner standardisierten Form unterliegt. Exaktes Georeferenzieren in historischen Karten ist häufig weniger von Bedeutung, vielmehr ist die Veranschaulichung von kulturellen, religiösen, politischen, wirtschaftlich relevanten Bezügen bezogen auf einen bestimmten Forschungskontext zentraler Punkt für die Gestaltung. Häufig lässt sich an der einfachen Ortsbeschreibung nicht erkennen, in welchem Kontext die Ortsangabe zu lesen ist, ob als Aufbewahrungsort eines Objektes, Fundort, historischer Ort o.ä. Dies erschwert es, Daten zu Ortsnamen und Georeferenzen aus verschiedenen Quellsystemen und -formaten in Beziehung zueinander zu setzen und sie in automatisierter Weise zu analysieren⁸³.

Wenn Geokoordinaten die genaueste Erfassung von Ortsbezeichnungen darstellen, welche Koordinatensysteme gibt es dann? Das **Gauß-Krüger-Koordinatensystem** ist ein kartesisches Koordinatensystem und wurde im 19. Jhd. von Carl Friedrich Gauß und Johann Heinrich Louis Krüger entwickelt. Diesem Standard unterliegen die meisten Kastasterpläne. Heute weit verbreitet ist die Nutzung des **WGS84 – World Geodetic System 1984**, das jedoch zunehmend durch **UTM – Universal Transverse Mercator** als Standard ersetzt wird, beide Standards basieren auf einem Rasterdatenmodell.

```
<history>
  <origin>
    <date notBefore="1790-05-21">1790-05-21</date>
    <country type="ISO_3166" key="XA-DE-BY">Germany
      <region>Bavaria</region>
    </country>
    <settlement type="TGN" key="7004326">Bayreuth
      <geogName>Cemetery, Jewish
        <geo>49.934064 11.594224</geo>
      </geogName>
    </settlement>
  </origin>
</history>
```

Codebeispiel 13: Auszug aus einem TEI-Header mit Referenz aus ISO 3166, TGN und WGS84 (zu den Abkürzungen vgl. nachfolgende Erläuterungen)

Normdaten und Thesauri bieten eine Möglichkeit verschiedene für die Ortsbeschreibung und Georeferenzierung benötigte Informationen miteinander zu verknüpfen. Welche Vokabulare stehen zur Verfügung bzw. sind besonders häufig in Gebrauch?

⁸³ Exemplarisch für den Aufwand, der für die Zusammenführung aller benötigten Daten für eine teilautomatische Analyse notwendig ist, sei hier auf das Kapitel 3.2 "Collection and mapping of archaeological data" hingewiesen in: Mayke Wagnera, Pavel Tarasovb, Dominic Hosnera, Andreas Flecka, Richard Ehricha, Xiaocheng Chena, Christian Leipeb, "Mapping of the spatial and temporal distribution of archaeological sites of northern China during the Neolithic and Bronze Age", in: *Quaternary International*, available online 2 July 2012, <http://dx.doi.org/10.1016/j.quaint.2012.06.039>

ISO 3166⁸⁴ ist eine internationale Norm von zwei- und dreibuchstabigen Länderkürzeln, die aus drei Teilen besteht. Der erste Teil ISO 3166 -1 ist für aktuelle Staaten, der zweite Teil ISO 3166 -2 für Untereinheiten von Staaten und der dritte Teil für auslaufende Länderkürzel von historischen Staaten seit 1977.

Getty Thesaurus of Geographic Names Online (TGN)⁸⁵ ist ein Thesaurus des J. Paul Getty Trust, der sowohl aktuelle als auch historische Bezeichnungen in verschiedenen Sprachen hierarchisch abbildet und für die Anwendung in der Museumsdokumentation entwickelt wurde. Er enthält ca. 1.106.000 Datensätze, die 91.200 Orte beschreiben. Des Weiteren beinhaltet der Thesaurus zusätzliche Informationen zu den erfassten Orten wie ungefähre Geokoordinaten, eine Kategorisierung nach Ortstyp und unterscheidet zwischen geophysischen und geopolitischen Datensätzen. Der Thesaurus ist jedoch lizenzpflichtig.

DARIAH-DE bietet mit dem *Getty Thesaurus of Geographical Names REST Service*⁸⁶ ein Werkzeug an, mit dem sich TGN-IDs schnell ermitteln lassen.

Beispiel: Der Parameter **ac** (autocompletion) bietet eine Liste sämtlicher Ortsnamen, die wie im folgenden Beispiel mit "goet" beginnen: <http://ref.dariah.eu/tgnsearch/tgnquery.xml?ac=goet>.

Der Parameter ***In-** long – usage: liefert ausführliche Informationen zu einem Ort: <http://ref.dariah.eu/tgnsearch/tndquery.xml?In=Berlin>.

Ebenso können mit einer bekannten TGN-ID über den Parameter **id** sämtliche Informationen zu dem betreffenden Ort abgerufen werden: <http://ref.dariah.eu/tgnsearch/tgnquery.xml?id=7004325>

Virtual International Authority File (VIAF)⁸⁷ wird vom OCLC (Online Computer Library Center) gehostet und ist ein Gemeinschaftsprojekt einiger Nationalbibliotheken zur Zusammenführung ihrer Normdateien. Geonormdaten werden hier ebenfalls miteinander verknüpft und sind somit in mehrsprachiger Übersetzung vorhanden. VIAF unterstützt auch nichtlateinische Schriften wie Hebräisch, Kyrillisch und Arabisch.

2007 wurde die Direktive **Infrastructure for Spatial Information in the European Community (INSPIRE)**⁸⁸ durch die Europäische Kommission in Kraft gesetzt. Ziel ist es, eine Geodateninfrastruktur zu schaffen, die Zugang zu sämtlichen Geodaten staatlich öffentlicher Einrichtungen gewährleistet. Auf europäischer Ebene erarbeitet die Koordinierungsstelle Europäische Geodateninfrastruktur ESDI Richtlinien, die von nationalen Koordinierungsstellen in den jeweiligen Ländern umgesetzt werden, für

⁸⁴ http://www.iso.org/iso/country_codes/iso_3166_code_lists.htm

⁸⁵ <http://www.getty.edu/research/tools/vocabularies/tgn/index.html>

⁸⁶ <https://dev2.dariah.eu/wiki/display/DARIAHDE/DNB+Personennamendatei+REST+Service>

⁸⁷ <http://viaf.org/>

⁸⁸ <http://inspire.jrc.ec.europa.eu/>

Deutschland die Geodateninfrastruktur -DE (GDI-DE, <http://www.geoportal.de>). INSPIRE unterstützt den offenen Standard GML vom **Open Geospatial Consortium (OGC)**⁸⁹. Die INSPIRE Datenspezifikationen definieren auch ein application profile für Protected sites⁹⁰. Protected sites sind unter anderem „*menschengemachte Objekte wie Architektur und archäologische Stätten oder kulturelle Objekte*“. Bis 2019 sollen die INSPIRE Richtlinien implementiert sein und Geodaten, die diesen Richtlinien entsprechen, stehen dann Forschern für die Nachnutzung zur Verfügung.

KML – Keyhole Markup Language, ist eine Auszeichnungssprache für Geodaten. KML ist ein Standard des OGC und wird u.a. von GoogleEarth und GoogleMaps sowie dem DARIAH-Geobrowser unterstützt.

```
<?xml version="1.0" encoding="UTF-8" standalone="yes"?>
<kml xmlns="http://www.opengis.net/kml/2.2"
  xmlns:gx="http://www.google.com/kml/ext/2.2"
  xmlns:atom="http://www.w3.org/2005/Atom"
  xmlns:kal="urn:oasis:names:tc:ciq:xsd:schema:xAL:2.0">
  <Folder>
    <Placemark>
      <name>Jüdischer Friedhof Adelebsen, In Steilhang</name>
      <address>Adelebsen</address>
      <TimeStamp>
        <when>1733</when>
      </TimeStamp>
      <Point>
        <coordinates>9.74058,51.575771</coordinates>
      </Point>
    </Placemark>
  </Folder>
</kml>
```

Codebeispiel 14: Ausschnitt einer aufbereiteten KML-Datei für das DARIAH spatio-temporal interface (Geobrowser)

GeoNames ist eine Datenbank mit derzeit 10 Millionen geografischen Namen in verschiedenen Sprachen, mit Angabe von Längen- und Breitengrad, Höhe über Meer, Einwohnerzahl, administrative Unterteilung und Postleitzahlen. Alle Koordinaten verwenden WGS84. Die Daten von GeoNames sind unter der Creative Commons Lizenz Namensnennung veröffentlicht und können somit frei nachgenutzt werden.

Iconclass ist eine Klassifikation für die Erfassung von Bildinhalten, auch

⁸⁹ <http://www.opengeospatial.org/>

⁹⁰ http://inspire.jrc.ec.europa.eu/documents/Data_Specifications/INSPIRE_DataSpecification_PS_v3.1.pdf

Ideallandschaften, mythologischen Orte etc., siehe Abschnitt Objektbeschreibungen.

Im *Getty Thesaurus of Geographic Names* werden nur ungefähre Angaben für die Referenzierung mit Koordinaten vorgenommen. Die meisten Thesauri gehen zudem nicht bis auf die Ebene von Gebäuden, Straßen oder gar Punkten. Lediglich INSPIRE wird dies zukünftig zumindest für *Protected Sites*-Inhalte bereitstellen. Um diese Lücke zu schließen, entwickeln Projekte, die im Umfeld der Digital Humanities entstehen, neue Formen von Gazetteers (Ortsverzeichnissen). Ein Beispiel für einen solchen Gazetteer ist **Pleiades**⁹¹. Pleiades ermöglicht es Forschern weltweit, ihre historischen geographischen Informationen auf dieser Plattform zu veröffentlichen und in digitaler Form zu teilen. Derzeit liegt der Fokus auf Ortsinformationen der griechisch-römischen Antike, jedoch weitet sich das Angebot zunehmend auch auf Informationen zu vorderasiatischer, byzantinischer, keltischer und mittelalterlicher Geographie aus und umfasst auch Angaben zu Gebäuden, Denkmälern und Fundplätzen menschlicher Aktivität, die keinen eigenen Ortsnamen haben, sowie mythologischen Orten. Pleiades ist spezialisiert und damit nur für bestimmte Geisteswissenschaften im Umfeld der Klassischen Archäologie und Alten Geschichte interessant, deckt aber dafür die Bedürfnisse dieser Wissenschaften besser als Universalklassifikationen ab.

Auch im Bereich der Georeferenzierung von historischen Karten entstehen neue Werkzeuge⁹², die es ermöglichen, Bilddaten zu georeferenzieren und Punktdaten zu gewinnen.

Durch die weite Verbreitung und Übersetzung in unterschiedliche Sprachen bieten internationale Standards und Universalklassifikationen zur Ortsbeschreibung gerade im Hinblick auf multilinguale Interoperabilität Vorteile, jedoch sind sie häufig nicht feingranular genug, um alle Anforderungen in der Forschung abzudecken.

3.3 Personeninformationen

Neben Orts- und geographischen Informationen stellen biographische- und personenbezogene Informationen Daten dar, die in allen geisteswissenschaftlichen

⁹¹ <http://pleiades.stoa.org/Members/sgillies/news-items/linked-data-for-pleiades-places>

⁹² Im Palantir Blog wird eine Methode basierend auf Quantum GIS beschrieben, <http://www.netpalantir.it/news/index/how-to-georeference-an-image>. Quantum GIS ist ein open source Geoinformationssystem (<http://www.qgis.org/>). Der webbasierte Georeferencer.org (<http://www.georeferencer.org/>) erlaubt das Bookmarken von historischen Karten aus dem Netz, in denen dann manuell einige Punkte markiert und mit Punkten an demselben Ort in Google-Maps verknüpft werden. Anschließend wird automatisiert ein KML-File erzeugt, mit dem weitere Funktionen von Georeferencer.org arbeiten (Genauigkeitsgradbemessung von historischen Karten etc.)

Disziplinen erhoben und verarbeitet werden, da sie per se einen interdisziplinären Charakter aufweisen und auf vielfältige Weisen in unterschiedlichen Forschungskontexten verwendet werden können. Zugleich sind Personeninformationen eindeutig referenzierbar. Das grundsätzliche Problem besteht allerdings darin, dass oftmals Personennamen unterschiedlich geschrieben wurden und werden – abhängig vom historischen, sprachlichen, geographischen und zeitlichen Kontext – und dadurch z.T. höchst unterschiedliche synonyme Namensschreibungen vorzufinden sind. Als Beispiel soll hierfür Johann Wolfgang von Goethe dienen, für den z.B. in der PND derzeit 67 synonyme Schreibweisen angegeben sind. Hier ist generell die Empfehlung auszusprechen, dass Personeninformationen prinzipiell auf Normdatensätze bezogen werden sollten. Nur dadurch ist gewährleistet, dass sie eindeutig referenzierbar werden und die beschriebenen Forschungsdaten interoperabel nachgenutzt werden können. Zudem wird die Datenqualität der Normdaten durch eine redaktionelle Betreuung sichergestellt. Folgende exemplarische Aufzählung zeigt, welche Dienste für die Personennormdatenvergabe derzeit vorhanden sind:

Die **Personennamendatei** der Deutschen Nationalbibliothek – abgekürzt **PND** (bzw. seit 19. April 2012 Bestandteil der umfassenderen Gemeinsamen Normdatei **GND**⁹³) – wurde seit Mitte der 1990er Jahre entwickelt und diente zunächst vor allem “zur Vereinheitlichung der Ansetzungsformen” von Autorennamen aus dem Kontext der Ermittlung von Titeldaten und deren Verfasser in Bibliotheken. Dabei stand die Erstellung “individualisierter Datensätze”, die z. B. Personen des gleichen Namens eindeutig identifizierten, im Mittelpunkt. Jedem Datensatz wurde eine neunstellige Zahlenkombination als “Identifizier” (IDs) zugeordnet, die sogenannte PND-Nummer.

Im Zuge der Integration der PND in die GND wurden auf nunmehr 10-stellige IDs umgestellt.

Die DNB legt beim Einrichten eines individualisierten Normdatensatzes auch die Ansetzungsformen von Namen fest; übliche unterschiedliche Schreibweisen (oder Pseudonyme) werden jedoch mit festgehalten. Personen-Normdaten können mithilfe von **PND Beacon** vernetzt werden. Die IDs der Personen werden dabei mit einem einfachen Protokoll gefunden. Um biografische Daten einzelner Personen weiter anzureichern, wird PND Beacon beispielsweise in der Carl-Maria-von-Weber- Gesamtausgabe eingesetzt.⁹⁴

⁹³ Die Gemeinsame Normdatei (GND) führt die zuvor einzeln bestehenden Normdateien – die Gemeinsame Körperschaftsdatei GKD, die Personennamendatei PND und die Schlagwortnormdatei SWD – in einem kooperativ erarbeiteten eindeutigen neuen Bezugssystem zusammen. Das verwendete Dateiformat ist dabei MARC 21. Vgl. dazu http://www.dnb.de/DE/Standardisierung/Normdaten/GND/gnd_node.htm

⁹⁴ <http://www.weber-gesamtausgabe.de/de/A002068>

DARIAH-DE bietet mit dem *DNB Personennamendatei REST Service*⁹⁵ ein Werkzeug an, mit dem sich PND-IDs schnell ermitteln lassen.

Beispiel: Der Parameter **ac** (autocompletion) bietet eine Liste sämtlicher Personen, die wie im folgenden Beispiel mit "goet" beginnen. <http://ref.dariah.eu/pndsearch/pndquery.xql?ac=goet>

Der Parameter ***In-** long - usage: liefert ausführliche Informationen zu einer Person:

<http://ref.dariah.eu/pndsearch/pndquery.xql?In=Ringeln>

Ebenso können mit einer bekannten PND-ID über den Parameter **id** sämtliche Informationen zu der betreffenden Person abgerufen werden: <http://ref.dariah.eu/pndsearch/pndquery.xql?id=118601121>

Diese Vollinformation lässt sich auch im RDF-Format ausgeben: <http://d-nb.info/gnd/118601121/about>

VIAF (Virtual International Authority File) ist ein Verbundprojekt, das die Personennormdateien der teilnehmenden Institutionen virtuell in einen einzigen "name authority service" überführt.

```
<li id="mediabar_linkaditemlviaf">
  <a href="http://viaf.org/viaf/106964661"
    title="Napoleon Bonaparte (1769&ndash;1821)&nbsp;- VIAF"
    class="external-link viaf" target="_blank">Napoleon Bonaparte
    (1769&ndash;1821)&nbsp;- VIAF&nbsp;&nbsp;&nbsp;</a>
  <a id="mediabar_linkaditemldnb" href="http://d-nb.info/gnd/118586408"
    title="Napoleon Bonaparte (1769&ndash;1821)&nbsp;- DNB"
    class="external-link dnb" target="_blank">DNB&nbsp;</a>
  <a href="http://www.deutsche-biographie.de/pnd118586408.html"
    title="Napoleon Bonaparte (1769&ndash;1821)&nbsp;- ADB/NDB"
    class="external-link adbndb" target="_blank">ADB/NDB&nbsp;</a>
</li>
```

Codebeispiel 15: *VIAF- und GND-Verwendung: Die Französische Revolution als europäisches Medienereignis (IEG)*⁹⁶

Alle Forschungsprojekte, die nicht nur auf Personen aus dem deutschsprachigen Raum referieren wollen, können einheitlich auf die VIAF-ID verweisen, da hier internationale Personeninformationen zu finden sind, andererseits sind die nationalen Nummerierungssysteme aber ohnehin in der VIAF mit abgelegt (diese ist also eine virtuelle Normdatei, die aus dem Abgleich von z. Zt. 25 Normsystemen von 20 Partnern entsteht), so dass es auch völlig legitim ist, weiterhin die PND (bzw. GND) als ID zu benutzen.

⁹⁵ <https://dev2.dariah.eu/wiki/display/DARIAHDE/DNB+Personennamendatei+REST+Service>

⁹⁶ <http://www.ieg-ego.eu/reichardtr-2010-de>
URN: urn:nbn:de:0159-2010101173_[2012-09-27].

In den Kunstwissenschaften ist die Verwendung der PND ebenfalls nicht besonders geeignet, da viele Künstlernamen fehlen und die Fachvokabulare hier ausführlichere Informationen bieten. Als empfehlenswerter Thesaurus bzw. Normdatei sind hier die *Union List of Artist Names*⁹⁷ (**ULAN**) und die Prometheus Künstlernamenansatzungsdatei⁹⁸ (**PKND**) zu nennen. Während viele Forschungs- bzw. Editionsprojekte aufgrund dieser Lücken der PND interne Namensdateien führen, ist mit dem Personendaten-Repository der Berlin-Brandenburgischen Akademie der Wissenschaften ein erster Schritt zur Zusammenführung solcher Informationen auch im geisteswissenschaftlichen Kontext gestartet worden.

Neben Einzelpersonen sind in vielen Fällen auch Personengruppen (Ehepaare, Familien, soziale Gruppen) zu referenzieren. Problemfälle stellen häufig auch Herrscherhäuser oder die Darstellung verwandtschaftlicher Beziehungen zwischen Personen dar. Dies geht aber über die hier primär genannte Aufgabe der eindeutigen Referenzierung von Personen hinaus und lässt sich nur mit Modellen erfassen, die Objektbeziehungen behandeln.

3.4 Objektinformationen

Bei Objekten kann zwischen mobilen und immobilen Objekten unterschieden werden. Für die geisteswissenschaftliche Forschung ist die besonders detaillierte Erfassung der Eigenschaften von Objekten relevant. Angaben zu Material, Größe, Form, Verwendungszweck, Erhaltungszustand, Aufbewahrung, Ausstellung und Dekor sind durch Beobachtung zu ermitteln und gehören zur vorikonographischen Beschreibung⁹⁹. Die Zuweisung von Objekten zu einem bestimmten Hersteller, Künstler oder einer Werkstatt, oder auch zu einem Stil bzw. einem Kulturkreis erfordert hingegen vertiefte Kenntnisse zu Kulturbedingungen, Einstellungen und Philosophie von Epochen und Kulturen und ist Teil der ikonographischen Analyse und ikonologischen Interpretation im Forschungsprozess.

The Cultural Objects Name Authority (CONA)¹⁰⁰ und der **Art & Architecture Thesaurus (AAT)**¹⁰¹ sind Produkte des J. Paul Getty Trust. CONA wird als Normdatei für Kunst, Archäologische Artefakte und Gebrauchsgegenstände in Museen, Universitätssammlungen und privaten Sammlungen dienen, die Aufnahme von naturwissenschaftlichen Sammlungen ist jedoch nicht geplant. Über CONA werden

⁹⁷ <http://www.getty.edu/research/tools/vocabularies/ulan/about.html>

⁹⁸ <http://prometheus-bildarchiv.de/de/tools/pknd>

⁹⁹ Die Teilung des Forschungsprozesses in der Kunstwissenschaft in Vorikonographische Beschreibung, Ikonographische Analyse und Ikonologische Interpretation stammt von Erwin Panofsky. Panofsky, E.: "Ikonographie und Ikonologie", in: Kaemmerling, E. (Hg.): *Bildende Kunst als Zeichensystem. Ikonographie und Ikonologie, Bd. 1. Theorien – Entwicklung – Probleme*, Köln 1994.

¹⁰⁰ <http://www.getty.edu/research/tools/vocabularies/cona/about.html>

¹⁰¹ <http://www.getty.edu/research/tools/vocabularies/aat/about.html>

Persistente Identifikatoren zu Objekten verwaltet werden. Der Art & Architecture Thesaurus unterscheidet sich von CONA dadurch, dass es sich hier um einen Thesaurus handelt, der auch zur Klassifizierung von Objekten in CONA genutzt werden kann. Derzeit ist er nur in englischer und spanischer Sprache erhältlich, jedoch arbeiten das Institut für Museumsforschung in Berlin¹⁰², das National Digital Archives Program in Taiwan¹⁰³, das Rijksbureau voor Kunsthistorische Documentatie¹⁰⁴, das Istituto Centrale per il Catalogo e la Documentazione in Rom¹⁰⁵ und das Canadian Heritage Information Network¹⁰⁶ an einer Übersetzung bzw. Integration ihres Vokabulars in den AAT. Der AAT weist sieben Hauptfacetten auf:

Assoziierte Begriffe – abstrakte Begriffe, wie Schönheit, Kennerschaft, Metapher, Freiheit
Physische Attribute – empfindbare und messbare Eigenschaften, wie Größe, Form, Textur
Stile und Epochen – stilistische und Epochen-Begriffe, wie Tang-Dynastie oder Chippendale
Agens – Personen oder Organisationen, wie Zeichner, Drucker, Architekt, Werkstatt des ...
Aktivitäten – Begriffe, wie Korrosion, Ausstellung, Malen
Material – Stoffe/Substanzen, wie Holz, Eisen, Ton, Nylon
Objekttypen – Begriffe, wie Gemälde, Amphore, Garten

Dewey-Dezimalklassifikation (DDC)¹⁰⁷ ist eine Universalklassifikation, die mittlerweile von OCLC (Online Computer Library Center) weiterentwickelt wird. Es gibt Übersetzungen in mehr als 30 Sprachen. DDC wird vor allem von Bibliotheken genutzt, ist aber auch bei anderen Einrichtungen weit verbreitet und international gebräuchlich. Diese Universalklassifikation eignet sich aufgrund ihres Verbreitungsgrades, um auf einer weniger feingranularen Ebene über heterogene Bestände eine multilinguale Interoperabilität der Metadaten zu schaffen. Seit 2003 gibt es eine deutsche Übersetzung, die in einem DFG-geförderten Projekt 2002–2005 von der FH Köln realisiert wurde. Die zehn Hauptfacetten orientieren sich am Kanon der Wissenschaften:

Informatik, Informationswissenschaft, allgemeine Werke
Philosophie & Psychologie
Religion
Sozialwissenschaften
Sprache
Naturwissenschaften
Technik, Medizin, angewandte Wissenschaften
Künste und Unterhaltung
Literatur
Geschichte und Geografie

¹⁰² <http://www.smb.museum/ifm/index.php?ls=10&topic=Aufgaben&lang=de&te=ja&tf=ja>

¹⁰³ http://www.ndap.org.tw/1_org_en/introduction.php

¹⁰⁴ <http://english.rkd.nl/about-the-rkd>

¹⁰⁵ <http://www.iccd.beniculturali.it/index.php?en/115/cataloguing-standards>

¹⁰⁶ <http://www.rcip-chin.gc.ca/index-eng.jsp>

¹⁰⁷ http://www.ddc-deutsch.de/Subsites/ddcdeutsch/DE/Home/home_node.html

Iconclass¹⁰⁸ wird vom Rijksbureau voor Kunsthistorische Documentatie in Den Haag im Projekt *Mnemosyme* mit internationaler Beteiligung weiterentwickelt. Es bietet eine Klassifikation insbesondere für die ikonographische Beschreibung von Bildinhalten und wird deshalb meist für die inhaltliche Erschließung von kunsthistorischen Sammlungen eingesetzt. Als Beispiel für die Verwendung in der Musikikonographie sei hier auf das *Répertoire International d'Iconographie Musicale (RidIM)*¹⁰⁹ verwiesen, das für die ikonographische Beschreibung bzw. die Erfassung der allgemeinen Bildinhalte sowie der musikalischen Kontexte dem internationalen ICONCLASS-System folgt. Iconclass ist nicht in allen Sprachen vollständig übersetzt, umfasst aber Terme in Englisch, Deutsch, Französisch und Italienisch. Der Thesaurus weist zehn Hauptfacetten auf¹¹⁰:

abstrakte, ungegenständliche Kunst

Religion und Magie

Natur

der Mensch (allgemein)

Gesellschaft, Zivilisation, Kultur

abstrakte Ideen und Konzeptionen

Geschichte

Bibel

Literatur

Klassische Mythologie und Antike Geschichte

Schlagwortnormdatei/Gemeinsame Normdatei (SWD/GND)¹¹¹, **Library of Congress Subject Headings (LCSH)**¹¹², **Répertoire d'autorité-matière encyclopédique et alphabétique unifié (RAMEAU)**¹¹³, **Virtual International Authority File (VIAF)**¹¹⁴ – In den Nationalbibliotheken verschiedener Länder wird seit Jahrzehnten an Vokabularen zur Sacherschließung gearbeitet. In der Regel ist dieses Vokabular aufgrund des breiten Spektrums an Themen, die im Bezug auf die gesammelte Literatur erfasst werden sollen, universell angelegt und nicht besonders ausdifferenziert. Interessant ist jedoch das Projekt zur Zusammenführung der nationalen Normdateien der Nationalbibliotheken zu einer internationalen Normdatei, dem *Virtual International Authority File*, das es dann aufgrund der Konkordanzen ermöglichen wird, zu einem gewissen Grad multilinguale Interoperabilität in der Verschlagwortung zu erreichen.

¹⁰⁸ <http://www.iconclass.nl/about-iconclass/what-is-iconclass>

¹⁰⁹ <http://www.ridim-deutschland.de/ridim/index.php?pcontent=startseite>

¹¹⁰ Die deutsche Übersetzung wird vom Bildarchiv Marburg und der Herzog August Bibliothek in Wolfenbüttel betreut: <http://www.hab.de/bibliothek/kataloge/image-browser.htm>

¹¹¹ http://www.dnb.de/DE/Standardisierung/Normdaten/GND/gnd_node.html

¹¹² <http://www.loc.gov/aba/cataloging/subject/>

¹¹³ <http://rameau.bnf.fr/>

¹¹⁴ <http://viaf.org/>

Einsprachiges Vokabular

Die **Hessische Systematik**¹¹⁵ ist eine deutschsprachige Systematik, die von und für kulturhistorische Museen erstellt wurde und vom Hessischen Museumsverband gepflegt wird. Die Dokumentation ist online abrufbar¹¹⁶. Sie besteht aus 18 Hauptgruppen:

Architektur
Bildwerke
Brauch und Fest
Forstwirtschaft/Jagd/Fischerei
Freizeit/Unterhaltung/Genuss
Gesundheit
Handwerk/Industrie/Handel
Hauswirtschaft
Kleidung
Kommunikation
Landwirtschaft
Messen und Wiegen
Öffentlichkeit und Gemeinwesen
Religion und Glaube
Schriftgut
Spielzeug/Spiel/Sport
Transport und Verkehr
Wohnen

Die **Hornbostel-Sachs-Klassifikation** wird in der Musikwissenschaft für Musik-Ikonographien eingesetzt. Es handelt sich um ein Klassifikationssystem für Musikinstrumente, das 1914 entwickelt wurde. Diese Klassifikation wird beispielsweise im *Répertoire International d'Iconographie Musicale (RIdIM)*¹¹⁷ genutzt. Aber auch jüngere Projekte greifen auf diese Klassifikation zurück, jedoch indem sie (teilweise sehr stark) modifiziert¹¹⁸ wird, so im *Musical Instrument Museums Online (MIMO)*¹¹⁹.

Die **Oberbegriffsdatei**¹²⁰ (**OBG**) ist in 45 Hauptgruppen und darunter liegende Untergruppen aufgeteilt, von *Aneignende Wirtschaft* über *Architektur*, *Behältnis*, *Beleuchtung*, *Besteck*, *Bildwerk*, *Elektrotechnisches Gerät*, *Gefäß*, *Graphik/Photographie*, *Hausrat*, *Gesundheitswesen*, *Körperpflege*, *Hygiene*, *Kleidung*, *Landwirtschaftliches Gerät*, *Landwirtschaftliche Maschine*, *Maschinenwesen*, *Meßgerät*,

¹¹⁵ <http://museum.zib.de/museumsvokabular/documents/systematik-hessen-original-2003.pdf>

¹¹⁶ <http://museum.zib.de/museumsvokabular/index.php?main=download&ls=9&co=we&ln=de>

¹¹⁷ <http://www.ridim-deutschland.de/ridim/index.php?pcontent=startseite>

¹¹⁸ http://www.mimo-international.com/documents/results/MIMO_Deliverable_3.3.pdf

¹¹⁹ <http://www.mimo-international.com/>

¹²⁰ <http://museum.zib.de/museumsvokabular/documents/obg.pdf>

Möbel, Münze, Medaille, Zahlungsmittel, Musikinstrument, Klangkörper, Schmuck, Spiel, Spielzeug, Textilie, Waffe, Werkzeug, Gerät bis Zunftzeichen. Diese Systematik ist derzeit bei über 150 Museen in Deutschland in Gebrauch. Seit 2008 wird die OBG im webbasierten Vokabularverwaltungstool xTree geführt. Die redaktionelle Überarbeitung erfolgt in einer überregionalen Arbeitsgruppe, bestehend aus:

- Landesstelle für die nichtstaatlichen Museen in Bayern, München
- Institut für Museumsforschung, Berlin
- digiCULT-Verbund eG, Kiel
- Stiftung Historische Museen Hamburg – Museum der Arbeit
- LWL-Freilichtmuseum Hagen – Landesmuseum für Handwerk und Technik

Object Names Thesaurus¹²¹ und Materials Thesaurus¹²² des British Museum – 1980 begann das British Museum seine Terminologie zur Verschlagwortung zu strukturieren und erarbeitete einen Thesaurus für Objekte und einen Thesaurus für Materialien. Der Object Names Thesaurus besteht aus 33 Hauptgruppen und der Materials Thesaurus gliedert sich in organische, anorganische und Werkstoffe.

agricultural/subsistence

animal equipment

animal/vegetal remains

architecture

arms/armour

artefact

childcare equipment

communications equipment

component

container

costume

culinary equipment

currency

fire/fire-lighting equipment

food product

furniture/furnishings

games/sporting equipment

instrument

lighting equipment

miscellaneous equipment

performance arts

personal ornament

production/replication

¹²¹ http://www.collectionslink.org.uk/assets/thesaurus_bmon/Objintro.htm

¹²² http://www.collectionslink.org.uk/assets/thesaurus_bmm/matintro.htm?phpMyAdmin=OYNyINPdn3sQmoXugKH1gcCLSW0

religious/ritual equipment
sample
science/medicine
sculpture
social control equipment
stimulant/narcotic equipment
textile
toilet/cosmetic equipment
transport/carrier
visual representation

Thesaurus of Monument Types¹²³, English Heritage¹²⁴ – English Heritage ist das britische Äquivalent auf nationaler Ebene zu den Landesdenkmalpflegeämtern in Deutschland und veröffentlicht Standards und Richtlinien für die Katalogisierung von Denkmälern. Der Thesaurus of Monument Types umfasst 18 Hauptgruppen zu immobilien Objekten:

agriculture and subsistence
civil
commemorative
commercial
communications
defence
domestic
education
gardens parks and urban spaces
health and welfare
industrial
maritime
monument <by form>
recreational
religious ritual and funerary
transport
unassigned
water supply and drainage

3.5 Quellennachweise

Die Provenienz geisteswissenschaftlicher Daten und Quellen ist für die Forschung und Analyse von zentraler Bedeutung. Nur wenn der einzelne Wissenschaftler erkennen kann, woher die Forschungsdaten stammen, in welchem größeren Datenzusammenhang sie einzuordnen sind und vom wem sie erstellt wurden, ist die Kontextualität und Kausalität von Forschungsergebnissen sichergestellt. So ist beispielsweise die Information bedeutsam,

¹²³ http://thesaurus.english-heritage.org.uk/thesaurus.asp?thes_no=1

¹²⁴ <http://www.english-heritage.org.uk/publications/guidelines-and-standards/>

in welchem Datenbestand ein Brief gefunden wurde, unabhängig davon, wer der Adressat und wer der Schreiber war. Diese Information ermöglicht z.B. Rückschlüsse darauf, wer einen Briefinhalt noch kannte oder über den Briefinhalt informiert war. Wie mit solchen wichtigen, über die Beschreibung der eigentlichen Daten hinausgehenden Informationen verfahren werden kann, soll das folgenden Beispiel anhand der TEI-Codierung verdeutlichen.

Angaben zu Quellen, ihren Standorten, nähere Beschreibungen ihrer Zusammensetzung und Provenienz werden in TEI gebündelt als *Source-description* (<sourceDesc/>) innerhalb des übergeordneten Abschnitts *File Description* (<fileDesc/> der Metadaten (siehe oben, z. Zt. S. 32) dokumentiert.

Gewöhnlich besteht eine <sourceDesc/> in TEI P5 aus folgenden (teils alternativ zu verwendenden) Unterabschnitten:

- <bibl/> = bibliographische Angaben, falls es sich bei der Quelle um Printmedien handelt
- <msDesc/> = *manuscript description*, falls es sich um handschriftliche Quellen handelt

Bibliographische Angaben können in freier oder strukturierter Weise erfolgen, sind aber nicht spezifisch für eine <sourceDesc/>, so dass hier auf den entsprechenden separaten Abschnitt der TEI-Guidelines verwiesen werden kann.¹²⁵ Für die Handschriftenbeschreibung steht ein umfangreiches Arsenal an Möglichkeiten zur Verfügung, dessen Erörterung den vorliegenden Rahmen sprengen würde. Daher seien hier nur einige grundlegende Merkmale (Besitzerangaben, Kurzbeschreibung, Provenienzvermerke usw.) angeführt (weitere Beispiele werden in die Wiki-Version dieses Dokuments integriert). Zur Illustration dient ein etwas komplizierterer Fall von Quellenüberlieferung, um zu demonstrieren, dass auch solche Fälle problemlos darstellbar sind.

Im nachfolgenden Beispiel handelt es sich um ein in drei Fragmente aufgelöstes Briefautograph, das heute in drei unterschiedlichen Bibliotheken aufbewahrt wird. Statt eines einzelnen Textzeugen (<witness/>) sind hier also gleich drei vorhanden, die in einer entsprechenden Liste (<listWit/>) zusammengestellt und durch die Attribute @n durchnummeriert sind. Jedes einzelne Fragment wird danach mit Besitzerangaben, Signaturen, einer kurzen Beschreibung sowie ggf. Angaben zur Provenienz oder in diesem Falle auch Hinweisen auf Veröffentlichungen des Briefes versehen. Die Einträge insgesamt werden durch die hier aus Platzgründen unterdrückten Containerelemente <sourceDesc/> bzw. innerhalb desselben <msDesc/> umschlossen.

¹²⁵ <http://www.tei-c.org/release/doc/tei-p5-doc/en/html/CO.html#COBITY>

```

<listWit>
  <witness n="1" xml:id="fragment1">
    <msDesc>
      <msIdentifier>
        <country>D</country>
        <settlement>Berlin</settlement>
        <repository n="D-B">Staatsbibliothek zu Berlin - Preußischer
          Kulturbesitz</repository>
        <idno>Mus.ep. Weber, C. M. v. 76</idno>
      </msIdentifier>
      <physDesc>
        <p>1 Bl. (2 b.S. o.Adr.)</p>
        <p>einige Wasserflecken im Text; rote Bleistiftmarkierungen
          (vermutlich von MMW) </p>
      </physDesc>
      <additional>
        <listBibl>
          <bibl n="firstPrint">(bislang keine Veröffentlichung
            nachzuweisen)</bibl>
        </listBibl>
      </additional>
    </msDesc>
  </witness>

```

Codebeispiel 16: Beginn des Nachweises eines aus drei Fragmenten bestehenden Briefes in TEI P5, <witness n="1">

```

<witness n="3" xml:id="fragment3">
  <msDesc>
    <msIdentifier>
      <country>US</country>
      <settlement>New York</settlement>
      <repository>Pierpont Morgan Library</repository>
      <collection>Mary Flagler Cary Music Collection</collection>
    </msIdentifier>
    <physDesc>
      <p>1 Bl. (2 b.S.einschl. Adr.), Fragment: Nachschrift vom 27.
        Januar</p>
    </physDesc>
    <history>
      <provenance>Madigan, Thomas F. (New York): Kat. 54 (1929), Nr.
        54-199</provenance>
      <provenance>Charavay, Maison Gabriel: Cat. 14. Mai 1881 (Slg. Johann
        Kafka), Nr. 63</provenance>
    </history>
  </msDesc>
</witness>
</listWit>

```

Codebeispiel 17: Fortsetzung mit Teil 3 desselben Briefes, <witness n="3">

Umfangreiche Möglichkeiten zum Hinterlegen von Informationen zu den benutzten Quellen bietet auch MEI (vgl. Punkt 2.2 und das Codierungsbeispiel im Anhang 6.4). Bedingt durch die Struktur musikalischer Quellen kann hier eine so komplexe Beschreibung entstehen, dass eigene graphische Hilfsmittel zu deren Eingabe entwickelt wurden (vgl. den Editor MerMEId am Danish Center for Music Publication der Königlichen Bibliothek Kopenhagen¹²⁶).

Zusammenfassend ist zu empfehlen, dass gerade bei Datensammlungen – wie z.B. einem Briefkonvolut – ein Metadaten-Schema verwendet wird, bei dem detaillierte Angaben zur Provenienz, zum Kontext und zu weiteren Korrelationen der Einzeldaten angegeben werden können.

3.6 Ereignisbeschreibung

Ereignisse spielen in der geisteswissenschaftlichen Forschung insbesondere in den geschichtswissenschaftlich arbeitenden Wissenschaften eine zentrale Rolle und können als historischer Kontext, in dem Objekte, Personen und Ideen in einem Ereignisraum aufeinandertreffen, definiert werden¹²⁷. Als Beispiel sei hier die Dokumentation der Stratigraphie bei archäologischen Grabungen zu nennen. Zerstörungshorizonte verweisen auf Brände, Kriege, Überflutungen, die sich zum Teil auf bekannte historische Ereignisse zurückführen lassen, damit geben sie in der Archäologie Anhaltspunkte für die Datierung und die Ursache der darunter- und darüberliegenden Schichten (post quem/ante quem) und der Funde und Befunde in diesen Schichten an einem bestimmten Ort¹²⁸.

Die ereignisfokussierte Dokumentation führt zu einer genaueren Erfassung von Objekten, da die Relationen von Informationen beschrieben werden. Anders als die reinen Fakten bieten ereignisbasierte Datenmodelle die Möglichkeit, Beteiligungen, Teil-Ganzes-Beziehungen und Klassifikationen von Objekten in Abhängigkeit von Statusveränderungen zu beschreiben.

Innerhalb von TEI ist für diese Zwecke ein eigenes Element `<event/>` eingeführt worden, das mit Personen, Orten oder Organisationen assoziiert werden kann. So lassen sich beispielsweise in der `<sourceDesc/>` Ereignisse innerhalb der Biographie einer Person strukturiert beschreiben.

¹²⁶ Dieser von Axel Teich Geertinger entwickelte Editor wird voraussichtlich Ende 2012 als open-source-Software verfügbar gemacht; vgl. <http://www.kb.dk/en/kb/nb/mta/dcm/projekter/mermeid.html>

¹²⁷ M. Doerr, A. Kritsotaki, "Documenting Events in Metadata", in: *The 7th International Symposium on Virtual*

Reality, Archaeology and Cultural Heritage VAST (2006),
<http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.79.6127>

¹²⁸ Martin Doerr, Dimitris Plexousakis, Katerina Kopaka, Chryssoula Bekiari, "Supporting Chronological Reasoning in Archaeology", in: *Computer Applications and Quantitative Methods in Archaeology Conference, CAA2004*, <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.81.434>

```

<tei:TEI xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xmlns:xhtml="http://www.w3.org/1999/xhtml" xmlns:tei="http://www.tei-c.org/ns/1.0"
  xsi:schemaLocation="http://www.tei-c.org/ns/1.0">
  <tei:teiHeader>
    <tei:fileDesc>
      <tei:sourceDesc>
        <tei:listPerson>
          <tei:person>
            <tei:persName>
              Margarethe Ulrike Auguste Marie Caroline Elsa Reger
            </tei:persName>
            <tei:birth when="1870-10-25">1870</tei:birth>
            <tei:event type="marriage" when="1891-02">
              <tei:label>Heirat mit Generalmajor Franz von Bercken, 1891
            </tei:label>
            </tei:event>
            <tei:event type="divorce" when="1899-04">
              <tei:label>Scheidung, 1899</tei:label>
            </tei:event>
            <tei:event type="marriage" subtype="legal" when="1902-10-25">
              <tei:label>Standesamtliche Heirat mit Max Reger, 1902
            </tei:label>
            </tei:event>
            <tei:event type="marriage" subtype="ecclesiastic"
              when="1902-12-07">
              <tei:label>Kirchliche Heirat mit Max Reger, 1902
            </tei:label>
            </tei:event>
            <tei:event type="adoption" when="1908-10-30">
              <tei:label>Christa Reger adoptiert</tei:label>
            </tei:event>
            <tei:event type="adoption" when="1910-04-29">
              <tei:label>Selma Charlotte »Lotti« Reger adoptiert
            </tei:label>
            </tei:event>
            <tei:event type="foundation" when="1947-10-25">
              <tei:label>
                Max-Regger-Institut/Elsa-Regger-Stiftung in Bonn gegründet
              </tei:label>
            </tei:event>
            <tei:death when="1951-05-03">1951</tei:death>
          </tei:person>
        </tei:listPerson>
      </tei:sourceDesc>
    </tei:fileDesc>
  </tei:teiHeader>
</tei:TEI>

```

Codebeispiel 18: Auszug der *sourceDesc* der Biographie Elsa Regers aus dem enzyklopädischen Teil der Reger-Werkausgabe. Mit freundlicher Genehmigung: Max-Regger-Institut/Elsa-Regger-Stiftung, Karlsruhe.

Das oben dargestellte Beispiel zeigt einen vereinfachten Abschnitt der TEI-Datei zur Biographie Elsa Regers, wie er im enzyklopädischen Teil der Reger-Werkausgabe Verwendung findet.

4. Metadaten zum Forschungskontext in den Geisteswissenschaften

Im Vergleich zu Standardisierungsaktivitäten im Bereich der beschreibenden Metadaten für geisteswissenschaftliche Daten und Quellen (siehe dazu Abschnitte 2 und 3 dieses Reports), lässt sich bei administrativen Metadaten ein grundlegender Bedarf, für den zurzeit in den meisten Codierungsformaten noch keine befriedigende Umsetzungsmöglichkeit geschaffen wurde, erkennen.

Im Sinne des klassischen Wissensmanagements können **administrative Metadaten** bzw. Verwaltungsmetadaten folgende ressourcen- oder objektbezogene Informationen enthalten hinsichtlich:¹²⁹

- Herstellung
- Speicherung
- Verwahrung
- Rechtemanagement
- Verwaltung der Zugriffsrechte

Allgemeiner formuliert können alle Informationen, die für die “Verwaltung” von Daten benötigt werden, unter diesen Begriff gefasst werden.

Die Beschreibung und Erfassung von geisteswissenschaftlichen Forschungsdaten mit Metadaten geschieht oftmals, sieht man von thematisch ausgerichteten Editionsprojekten ab, als ein Arbeitsschritt eines individuellen Forschungsprozesses. Dieser Forschungsprozess, der sich durch ein Erkenntnisinteresse und durch spezielle Forschungsfragen und damit verbundenen Methoden und Verfahren strukturiert, sollte auch als (Meta-)Information abrufbar sein, damit nachvollziehbar bleibt, wie die entsprechenden Forschungsdaten erhoben, erfasst und aufbereitet wurden, also nachnutzende Forscher den “Herstellungsprozess” rekonstruieren können.

¹²⁹ Vgl. dazu etwa das Online-Tutorial von Ulrike Spree (HAW Hamburg) unter <http://www.bui.haw-hamburg.de/pers/ulrike.spree/remind/ueberblick.htm>

Der folgende Abschnitt *4.1 Forschungsmethode und -verfahren* zeigt die Notwendigkeit einer stärkeren Fokussierung auf administrative Metadaten in diesem Zusammenhang auf. Anschließend werden am Beispiel von TEI in Abschnitt *4.2 Dokumentationsmöglichkeiten von administrativen Metadaten* auf dem gegenwärtig möglichen Stand demonstriert. Dabei sei aber nochmals darauf hingewiesen, dass TEI einer der wenigen Standards ist, der diese Möglichkeiten überhaupt anbietet.

4.1 Forschungsmethode und -verfahren

Forschungsmethoden und Forschungsverfahren spielen in allen geisteswissenschaftlichen Disziplinen eine, wenn nicht sogar, die zentrale Rolle. Anhand dieser lassen sich wissenschaftliche Ergebnisse überprüfen und unter anderem feststellen, ob die für die Analysen und die Beantwortung des forschungsbezogenen Erkenntnisinteresses und der individuellen Forschungsfragen verwendeten Quellen bzw. erhobenen Forschungsdaten nach wissenschaftlichen Kriterien ausgewählt wurden und diese Auswahl nachvollziehbar und für Dritte überprüfbar ist. Nur so lässt sich die Validität und auch Variabilität von geisteswissenschaftlichen Ergebnissen überprüfen. DARIAH-DE verwendet die im Folgenden näher ausgeführten Definitionen, um zwischen *Forschungsmethoden* und *Forschungsverfahren* zu unterscheiden und um zugleich die Komplexität dieser Informationen zu verdeutlichen, die bislang in vorhandenen Metadatenstandards und -schemata nur unzulänglich abgebildet werden.

Mit dem Begriff einer wissenschaftlichen **Forschungsmethode** wird ein erkenntnis- und forschungsfragenorientiertes, reflexiv kontrolliertes, theoretisch begründetes, auf einen Forschungsgegenstand zentriertes Vorgehen bezeichnet, das eine Arbeitshypothese und einen interpretativen Horizont beinhaltet. **Forschungsmethoden der Digital Humanities** werden durch die digital vorliegenden Repräsentationen des Forschungsgegenstands und durch die computergestützte Manipulation dieser Repräsentationen entweder erleichtert und erweitert – oder aber überhaupt erst in dieser Form möglich.

Ein **Forschungsverfahren** ist ein planvoller, systematischer und praktischer Umgang mit Forschungsgegenständen, dem die Forschungsmethode insbesondere mit ihrem interpretativen Horizont stets übergeordnet bleibt. Eine Forschungsmethode wird in der Regel durch mehrere Verfahren umgesetzt oder ermöglicht, umgekehrt können bestimmte DH-Verfahren im Kontext mehrerer Methoden eingesetzt werden. Diese besitzen sozusagen per se einen interoperablen Zugang bzw. interdisziplinäre Ausrichtung. **Verfahren der Digital Humanities** werden durch die digital vorliegenden Repräsentationen des Forschungsgegenstands und durch die computergestützte

Manipulation dieser Repräsentationen entweder erleichtert und erweitert – oder aber überhaupt erst in dieser Form möglich.

Während die in Forschungsprojekten ausgewählten und verwendeten Forschungsmethoden oft einen sehr starken disziplinären Charakter besitzen, können Forschungsverfahren eher generisch verwendet werden. So sind die Verfahren der Georeferenzierung oder des Textmining für unterschiedliche Disziplinen und Forschungsfragen anwendbar. Dies verdeutlicht den starken interoperablen Charakter von DH-Verfahren und verdeutlicht zugleich die Möglichkeit, dass mit solchen Verfahren gewonnene Forschungsdaten auch in anderen Forschungskontexten und für die Analyse anderer Forschungsfragen in anderen geisteswissenschaftlichen Fachdisziplinen nachgenutzt werden können.

Sowohl in den verwendeten Forschungsmethoden als auch in den daraus resultierenden Forschungsverfahren spiegeln sich die erkenntnisleitenden Fragestellungen der geisteswissenschaftlichen Forscher wider. Diese haben zudem einen direkten Einfluss auf die erhobenen Forschungsdaten. Nur *die* Daten und Quellen werden von dem einzelnen Forscher erfasst, erhoben oder auch analysiert, die einen unmittelbaren Einfluss auf die Beantwortung der eigenen individuellen Forschungsfrage besitzen. Diese Aspekte werden allerdings bislang durch Metadatenstandards und -schemata nicht erfasst, d. h. bei der digitalen Nachnutzung – und auf nichts anderes soll in diesem Kontext fokussiert werden – werden administrative Daten mit diesem Informationsgehalt kaum angegeben, so dass geisteswissenschaftliche Nachnutzer zurzeit nicht nachvollziehen können, in welchem Kontext, von wem und zur Beantwortung welcher Forschungsfragen die Forschungsdaten oder Datencollections erhoben wurden. Diese Problematik wird noch dadurch verschärft, dass oftmals die nicht verwendeten und dementsprechend nicht digitalisiert vorliegenden Daten der Originärdatenbestände nicht mit angegeben werden. Kurzum: Es besteht die Gefahr, dass der entstehungshistorische Kontext von Daten und Quellen bei einer digitalen Nachnutzung verschwindet. Diese administrativen Informationen sollten, so eine Empfehlung dieses Reports, zukünftig mit erhoben und durch standardisierte und somit maschinenlesbare Informationen erfasst werden. Nur dadurch wird sichergestellt, dass gerade unter der Perspektive der Interoperabilität Wissenschaftler den Entstehungskontext und die Provenienz der Daten erfassen können.

Wie in den bisherigen Ausführungen deutlich wurde, werden die Forschungsfragen, -methoden und -verfahren zwar meist in Projektanträgen etc. ausformuliert, jedoch fehlt bisher ein Metadatenformat für die Maschineninterpretierbarkeit solcher Informationen.

4.2 Dokumentationsmöglichkeiten von administrativen Metadaten am Beispiel des *teiHeader*

Die Text Encoding Initiative (TEI) stellt innerhalb des sog. *teiHeader* eine mehrstufige, modular erweiterbare Struktur zur Erfassung unterschiedlicher administrativer Metadaten zur Verfügung:

1. `<fileDesc/>`
Beschreibung der einzelnen elektronischen Ressource (als Computerfile)
2. `<encodingDesc/>`
Dokumentation der Beziehungen zwischen Quelle(n) u. elektron. File
3. `<profileDesc/>`
Angaben zum "Text-Profil" (Sprache, Urheber, Klassifikationen usw.)
4. `<revisionDesc/>`
Beschreibung der Revisionsgeschichte einer Datei

Im engeren Sinne fallen nur die Punkte 2 und 4 in den Bereich der administrativen Metadaten. Sowohl `<encodingDesc/>` als auch `<revisionDesc/>` erlauben Untergliederungen in teils freier, aber auch stärker strukturierter Weise (letztere empfiehlt sich stets im Hinblick auf die leichtere maschinelle Auswertung von Angaben).

Die File Description (deskriptive Metadaten)

Die `<fileDesc/>` bezieht sich rein deskriptiv auf den vorliegenden elektronischen Datensatz mit Angaben zum Titel, zur Art der vorliegenden Datei, ihrem Umfang und ihrer Publikation, aber auch zu den dabei benutzten Quellen (zu letzterem Punkt vgl. Abschnitt 3.5) und wird daher an dieser Stelle nicht genauer behandelt.

Die Elemente einer Encoding Description in TEI

Im Container-Element `<encodingDesc/>` (*encoding description*) können allgemeine Aspekte zur Art der Codierung eines Objekts untergebracht werden. Die beiden wichtigsten formalisierten Bestandteile einer solchen Beschreibung sind:

- `<projectDesc/>`, eine Kurzcharakteristik des digitalen Projekts (Projektbeschreibung)
- `<editorialDecl/>`, Erklärungen zu den zugrundeliegenden Editionsprinzipien

Damit ist lediglich ein Minimal-Set an Informationen abgedeckt, die zudem meist knapp gehalten sind, weil sie in der Regel nicht spezifisch für ein einzelnes File, sondern für ganze Sammlungen sind, so dass man mit einem Verweis auf separat abgelegte ausführlichere Daten (etwa separate Editionsrichtlinien) auskommt.

```

<?xml version="1.0" encoding="UTF-8"?>
<teiHeader>
  <!-- andere Teile des teiHeader -->
  <encodingDesc>
    <projectDesc>
      <p>Carl-Maria-von-Weber-Gesamtausgabe, Sämtliche Briefe, Digitale
        Edition</p>
    </projectDesc>
    <editorialDecl>
      <p>Die Übertragung und Texteinrichtung folgt den Editionsrichtlinien der
        WeGA, vgl. http://www.weber-gesamtausgabe.de/de/Editionsrichtlinien</p>
      <p>Der Text des vorliegenden Files folgt ausnahmsweise dem späteren
        Druck D-2, da dieser möglicherweise auf das Original zurückgeht.</p>
    </editorialDecl>
  </encodingDesc>
</teiHeader>

```

Codebeispiel 19: Beispiel einer *encodingDesc* der Carl-Maria-von-Weber-Gesamtausgabe

Es ist jedoch – speziell bei einem umfangreichen Einzeltext – möglich, auch umfangreichere Angaben zu bestimmten Aspekten des editorischen Vorgehens strukturiert abzulegen. Das betrifft etwa Angaben zu Korrekturen (*<correction/>*), zum Grad der Normalisierung von Orthographie (*<normalization/>*), zum Umgang mit Silbentrennungen (*<hyphenation/>*), zur Kennzeichnung zitierter Abschnitte (*<citation/>*) u.v.a.m.

Mit einer Reihe weiterer Deklarationen lassen sich auch forschungsmethodische Details festhalten:¹³⁰

- Angaben zu den Leitlinien bei der Anlage eines Korpus oder einer Kollektion von Texten (*<samplingDecl/>*)
- Erklärung zu der Verwendung der Markup-tags (also der Elemente) im Dokument (*<tagsDecl/>*)
- Erklärung zur Struktur von Verweisen im Dokument (*<refsDecl/>*)
- Verwendete Taxonomien im Text (*<classDecl/>*)

In der Regel werden diese Angaben frei innerhalb eines *<p/>*-tags formuliert. Beispiele hierzu wird die Wiki-Version des vorliegenden Dokuments enthalten.

Angaben zum Textprofil

Die umfangreichen Möglichkeiten zur Angabe von Aspekten des “Textprofils” in der *<profileDesc/>* gehören zur inhaltlich erschließenden Beschreibung eines Files – also nicht im engeren Sinne zu den administrativen Metadaten (vgl. hierzu auch die

¹³⁰ Es sind hier nicht alle möglichen Elemente genannt, zu weiteren Details vgl. <http://www.tei-c.org>.

Getty-Klassifikation bzw. den Abschnitt 3.4 Objektinformationen). Die fakultative `<profileDesc/>` beschreibt den jeweils in der Datei wiedergegebenen Text als historischen, z.B. mit Angaben zu seiner Entstehung, zur verwendeten Sprache, Textsorte usw.

Die Beschreibung der Revisionsgeschichte einer Datei

Innerhalb der `<revisionDesc/>` (revision description) des TEI-Headers besteht die Möglichkeit, alle Veränderungen an einer vorliegenden Datei zu protokollieren (die Ausführlichkeit der Einträge hängt dabei selbstverständlich von den jeweiligen Bedürfnissen in einem kollaborativen Arbeitsumfeld ab). Jede Änderung (`<change/>`) wird datiert und einem Verantwortlichen zugewiesen (vgl. die Attribute `@when` und `@who` im folgenden Beispiel). Die Granularität der Beschreibung der Eingriffe sollte dabei in einem Projekt mit zahlreichen Mitarbeitern feiner sein als in leicht überschaubaren Kontexten. Die beispielhafte Revisions-Historie innerhalb eines Briefdokuments, wie sie innerhalb der digitalen Carl-Maria-von-Weber-Gesamtausgabe¹³¹ im `<teiHeader/>` angelegt wird, zeigt das folgende Codebeispiel:

```
<?xml version="1.0" encoding="UTF-8"?>
<teiHeader>
  <!-- andere Teile des teiHeader -->
  <revisionDesc status="candidate">
    <change when="2012-06-19" who="#AS">alte ID's und @keys ersetzt</change>
    <change when="2012-04-18" who="#JV">Personen ausgezeichnet, Korrekturlesung
      Haupttext</change>
    <change when="2011-04-03" who="#PS">Text erstmals eingerichtet, Kommentare
      eingearbeitet</change>
  </revisionDesc>
</teiHeader>
```

Codebeispiel 20: Beispiel für eine `revisionDesc` innerhalb eines Briefdokuments

Das generell bestehende Defizit an maschinenlesbar hinterlegten Metadaten, die sowohl Herstellung als auch spätere Änderungen an digitalen Ressourcen dokumentieren, wird in dieser Codierungsform sichtbar: Die einzelnen Herstellungs- und Änderungsprozesse sind lediglich innerhalb des Elements `<change/>` als Fließtext ausgezeichnet und können dadurch nicht maschinell ausgewertet werden. Dies ist hier ausschliesslich für die angegebenen Attribute der Zeit (`when=" "`) und Person (`who=" "`) möglich.

Um einen (idealisierten) Workflow maschinell erfassbar zu machen, müssten die bisher im Fließtext ausformulierten Beschreibungen stärker formalisiert werden. Das kann einerseits in einfachster Weise durch ein kontrolliertes Vokabular erfolgen, das nur festgelegte Werte bzw. das Vorkommen von Standardbegriffen in den eingetragenen Texten erlaubt.

¹³¹ <http://weber-gesamtausgabe.de/de/Index>

Denkbar ist z. B., für ein bestimmtes Textopus festzulegen, dass folgende Begriffe in einer Revisionshistorie auftauchen dürfen:

- initiale Transformation
- formale Auszeichnung
- semantische Auszeichnung (ggf. differenziert: Auszeichnung Personen/Werke)
- Erstkorrektur
- Kommentierung
- Endkorrektur

In diesem Falle könnte dann gezielt nach bestimmten Stufen der Veränderung gesucht werden.

Eine weitere Möglichkeit ist die Überführung einer solchen Begrifflichkeit in Attributwerte. Neben den bislang in TEI im Element `<change/>` vorwiegend verwendeten Attributen `@when`, `@who` und `@n` (letzteres für eine ggf. vorzunehmende Nummerierung oder Vergabe auslesbarer Zahlenwerte) kann neuerdings auch das Attribut `@type` für die Kennzeichnung und spätere Ausgabe definierter Werte genutzt werden (etwa `type="initTrans"`, d.h. initiale Transformation).

Beispiel für eine entsprechend individuell angepasste, stärker formalisierte Revisionsbeschreibung:

```
<teiHeader>
  <!-- andere Teile des teiHeader -->
  <revisionDesc>
    <change type="rel01Corr" when="2011-04-08" who="#PS">Abschließende Korrektur
      für Release 1</change>
    <change type="semMarkup" when="2011-04-03" who="#PS">Veranstaltungsorte
      ausgezeichnet</change>
    <change type="semMarkup" when="2012-05-03" who="#PS">Ortsnamen
      ausgezeichnet</change>
    <change type="formalMarkup" when="2012-04-18" who="#JV">Lateinschreibung
      und Konsonantengemination ausgezeichnet</change>
    <change type="semMarkup" when="2012-03-03" who="#PS">Personen und Werke
      ausgezeichnet</change>
    <change type="initCorr" when="2011-06-20" who="#AS">Korrekturgelesen bis
      Ende Kapitel 5</change>
    <change type="initCorr" when="2011-06-19" who="#AS">Korrekturgelesen bis
      Ende Kapitel 2</change>
    <change type="initTrans" when="2011-01-01" who="#KO">Initiale
      Transformation, grobe Fehler beseitigt</change>
  </revisionDesc>
</teiHeader>
```

Codebeispiel 21: Modifizierte `revisionDesc` innerhalb eines Briefdokuments (TEI P5)

Alle vorstehenden Beispiele bleiben aber bei eher formalen Beschreibungen des

Arbeitsprozesses und der Voraussetzungen, auf denen eher beruht, stehen – sie demonstrieren daher die Notwendigkeit einer detaillierteren Berücksichtigung des Forschungskontexts auch im dafür bereits am besten geeigneten TEI-Schema.

5. Schlussfolgerungen

Die vorstehenden Beschreibungen, Analysen und die dabei erst in Ansätzen verwirklichte Evaluation der Formate zeigen, dass sich die Geisteswissenschaften und damit verbunden die Digital Humanities erst in jüngster Zeit mit dem Thema Standardisierung und der Verwendung von interdisziplinären Austauschformaten auseinanderzusetzen beginnen. Es entsteht der Eindruck, dass die textbasierten Geisteswissenschaften weniger mit den für ihre Zwecke meist nicht ausreichenden Standards der Bibliotheken und Archive arbeiten können, sondern effektiver mit einem Format wie TEI; jedoch gibt es auch hier bereits Entwicklungen hin zu spezialisierten, auf die Forschungsinteressen an bestimmten Textquellen zugeschnittenen Standards, wie EpiDoc, CEI und MXML. Bei den Objektivwissenschaften, wie den Archäologien und Kunstwissenschaften, wird deutlich, dass sie in der Standardisierung sehr viel stärker auf die Verwendung verschiedener Metadatenformate setzen, die jeweils für bestimmte Forschungsmethoden besonders geeignet sind. Allein die enorme Heterogenität der in den Geisteswissenschaften behandelten Forschungsdaten, Objekte und Quellen – und damit auch die Heterogenität der darauf bezogenen Forschungsfragen – erklärt bereits die Notwendigkeit dieser Vielfalt. In allen Geisteswissenschaften fehlt jedoch bisher ein Standard zur Erfassung des Forschungskontextes selbst. Hier muss zukünftig verstärkt darauf hingewirkt werden, dass diese für den wissenschaftlichen Forschungsprozess unerlässlichen Informationen zur Nutzung digital vorliegender Forschungsdaten standardisiert erfasst werden können.

Die Erarbeitung von gemeinsamen Normdatensätzen oder Thesauri bietet für die Erschließung in allen Geisteswissenschaften einen ersten Ansatz, um über Disziplinengrenzen hinweg Nutzen und Mehrwerte durch die Möglichkeit der Interoperabilität zu schaffen. Dies sind Aspekte, die verstärkt an geisteswissenschaftliche Forschende vermittelt und zugleich aber auch eingefordert werden sollten, für die andererseits aber gerade im Bereich der Museen, Bibliotheken und Archive hervorragende Vorarbeiten existieren, die in die Arbeiten der Geisteswissenschaften mit einbezogen und ggf. disziplinspezifisch erweitert werden sollten. Hier ist weniger Neu-, denn Weiterentwicklung des Bestehenden und eine stärkere Kooperation zwischen Wissenschaft und Museen, Bibliotheken und Archiven zu fordern.

Zugleich dürfte durch den vorliegenden Report deutlich werden, dass von Seiten der Forschungsförderer die Berücksichtigung von Standards und Normdaten stärker als bislang eingefordert werden muss, wenn größere Interoperabilität als ein in absehbarer Zeit zu erreichendes Ziel gelten soll. Dabei bleibt stets behutsam zwischen generell zu

berücksichtigenden Normen und für disziplin- oder projektspezifische Fragestellungen notwendigen Modifikationen, Ergänzungen oder gar Neuentwicklungen abzuwägen. In allen Fällen müssen die für einen solchen Umbau in den Geisteswissenschaften notwendigen Ressourcen aber in den Zeitbudgets und der Personalausstattung von Forschungsprojekten stärker berücksichtigt werden – bis hin zum verstärkten Schulungsbedarf, der für diesen Wandel so lange essentiell bleibt, bis die universitäre Ausbildung diese Aufgaben übernommen hat. Forschungsförderer sind am ehesten in der Lage, in ihren Leitlinien für die Beantragung von Projekten die künftige Notwendigkeit der Dokumentation des Forschungskontextes als eine neue Aufgabe interdisziplinärer bzw. Ergebnisse nachnutzender Forschung zu verdeutlichen.

Wenn der vorliegende Report den Titel *Essentials für fachspezifische Empfehlungen* trägt, so dürfte nach den vorausgehenden Beschreibungen und Analysen deutlich geworden sein, dass die wichtigste dieser Empfehlungen nicht lauten kann, sich auf ein bestimmtes Format zu einigen, da die Wahl eines Formats letztlich in Abhängigkeit von Forschungsgegenstand und Fragestellung bestimmt wird. Vielmehr ist die kritische Evaluierung der Formate bezüglich ihrer jeweiligen Leistungsfähigkeit zur Dokumentation von Metadaten ins Zentrum der Überlegungen vor einer Festlegung auf ein bestimmtes Format zu rücken. Aus der Sicht der hier beteiligten Einzeldisziplinen war ein direkter Vergleich aufgrund der beschränkt zur Verfügung stehenden Beispiele nur an einigen Stellen möglich. Es ist zu hoffen, dass mit der Öffnung der Diskussion über den Horizont der jetzt Beteiligten hinaus rasch eine Sammlung von Beispielen entsteht, die auf umfassenderer Grundlage – und in der dann übersichtlicheren Darstellung innerhalb des DARIAH-Wiki – direkte Vergleiche der Leistungsfähigkeit der Formate (und damit ihrer ggf. noch zu behebenden Defizite) erlaubt und damit zugleich dazu beiträgt, Orientierungshilfen zu bieten und innerhalb des DARIAH-DE-Arbeitspakets 3.4 gezielte Mapping-Verfahren zu entwickeln, die zu einer formatübergreifenden Interoperabilität der Daten beitragen. Bei all diesen Diskussionen sollte künftig auch darauf geachtet werden, dass Partner aus Museen, Bibliotheken und Archiven hinzugezogen werden, denn diese Institutionen werden wesentlich oder zumindest mit dafür zuständig sein, die von der Forschung erzeugten Daten langfristig vorzuhalten. Für die in beiden Bereichen verwendeten, oft noch getrennten Standards werden Mapping-Verfahren oder bereichsabhängige Erweiterungen verstärkt notwendig sein, wenn hier eine stärkere Durchdringung möglich werden soll.

In diesem Sinne trägt die vorliegende Dokumentation ausdrücklich eine Versionsnummer und ist als ein sich nun zunächst für die Beteiligung der Projektpartner der DARIAH-Community öffnendes *living document* zu verstehen, das sich durch eine anhaltende Diskussionskultur zu einem Kompendium nachhaltiger und nachvollziehbarer “Empfehlungen” weiterentwickeln soll.

6. Anhang

6.1 Literatur und URLs

Art & Architecture Thesaurus (AAT)

<http://www.getty.edu/research/tools/vocabularies/aat/about.html>

Bayerisches Landesamt für Denkmalpflege, *Vorgaben zur Dokumentation archäologischer Ausgrabungen in Bayern*, 2010

http://www.blfd.bayern.de/medien/vorg_doku_arch_ausg.pdf (aufgerufen, September 2012)

Brandenburgisches Landesamt für Denkmalpflege und Archäologisches Landesmuseum, *Richtlinien zur Grabungsdokumentation*, 2006

<http://www.bldam-brandenburg.de/images/stories/PDF/richtlinien%20zur%20grabungsdokumentation.pdf>

(aufgerufen, September 2012)

Biel, J., Klonk, D. (Hrsg.), *Handbuch der Grabungstechnik*, Gesellschaft für Vor- und Frühgeschichte, Stuttgart, 1994

BSZ, Datierungsthesaurus "Stilepochen"

https://wiki.bsz-bw.de/lib/exe/fetch.php?media=m-team:musis:museumsvokabular:infosheet_stilepochen_mai2011.pdf

Carl-Maria-von-Weber Gesamtausgabe

<http://www.weber-gesamtausgabe.de/de/A002068>

CEI

<http://www.cei.lmu.de/> (aufgerufen, August 2012)

Country Codes – ISO 3166

http://www.iso.org/iso/country_codes/iso_3166_code_lists.htm

Cultural Objects Name Authority (CONA)

<http://www.getty.edu/research/tools/vocabularies/cona/about.html>

DARIAH-DE-Report R 2.2.3, *Erster Bericht über die Verwendung von Verfahren der Digital Humanities in den Geistes- und Kulturwissenschaften*, V1.0, 31. August 2012.

https://dev2.dariah.eu/wiki/download/attachments/2295542/M223_DH-Verfahren.pdf?version=1&modificationDate=1347378358954 (aufgerufen, September 2012)

Deutschen Nationalbibliothek, Gemeinsame Normdatei (GND)

http://www.dnb.de/DE/Standardisierung/Normdaten/GND/gnd_node.html (aufgerufen, September 2012)

Dewey-Dezimalklassifikation

http://www.ddc-deutsch.de/Subsites/ddcdeutsch/DE/Home/home_node.html

Martin Doerr, Dimitris Plexousakis, Katerina Kopaka, Chryssoula Bekiari, Supporting Chronological Reasoning in Archaeology, In Computer Applications and Quantitative Methods in Archaeology Conference, CAA2004, <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.81.434>

Doerr, Martin / Kritsotaki, A., Documenting Events in Metadata, The 7th International Symposium on Virtual Reality, Archaeology and Cultural Heritage VAST (2006),
<http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.79.6127>

English heritage

<http://www.english-heritage.org.uk> (aufgerufen, September 2012)

EpiDoc

<http://epidoc.sourceforge.net/> (aufgerufen, August 2012)

Fachgruppe Dokumentation im Deutschen Museumsbund, museumdat, 2007

<http://museum.zib.de/museumdat/museumdat-v1.0.pdf> (aufgerufen, September 2012)

Gemeinsame Normdatei (GND)

http://www.dnb.de/DE/Standardisierung/Normdaten/GND/gnd_node.html

Getty Thesaurus of Geographic Names (TGN)

<http://www.getty.edu/research/tools/vocabularies/tgn/index.html>

Hessische Systematik

<http://museum.zib.de/museumsvokabular/documents/systematik-hessen-original-2003.pdf>

<http://museum.zib.de/museumsvokabular/index.php?main=download&ls=9&co=we&ln=de>

Horizon 2020 – the framework programme for research and innovation

http://ec.europa.eu/research/horizon2020/index_en.cfm (aufgerufen, September 2012)

Humdrum

<http://musiccog.ohio-state.edu/Humdrum/guide.toc.html> (aufgerufen, September 2012)

Iconclass

<http://www.iconclass.nl/about-iconclass/what-is-iconclass>

Infrastructure for Spatial Information in the European Community (INSPIRE)

<http://inspire.jrc.ec.europa.eu/>

INSPIRE – Application Profile for Protected Sites

http://inspire.jrc.ec.europa.eu/documents/Data_Specifications/INSPIRE_DataSpecification_PS_v3.1.pdf

Kalendersysteme, Übersicht:

http://de.wikipedia.org/wiki/Liste_der_Kalendersysteme (aufgerufen, September 2012)

Kepper, J., *Musikedition im Zeichen neuer Medien: Historische Entwicklung und gegenwärtige Perspektiven*

musikalischer Gesamtausgaben, 2011.

Landesamt für Denkmalpflege Hessen, *Richtlinien zur Grabungsdokumentation*, 2005
<http://hessen-archaeologie.de/Download/LfDHessenGrabungsrichtlinien2005.pdf> (aufgerufen, September 2012)

Landesarchäologie Sachsen-Anhalt, *Dokumentation*
<http://www.archaeologie.sachsen.de/59.htm> (aufgerufen, September 2012)

Library of Congress, Extended Date / Time Format (EDTF)
<http://www.loc.gov/standards/datetime/pre-submission.html>

Library of Congress Subject Headings (LCSH)
<http://www.loc.gov/aba/cataloging/subject/>

LIDO
<http://network.icom.museum/cidoc/working-groups/data-harvesting-and-interchange/what-is-lido> (aufgerufen, September 2012)

LVR-Amt für Bodendenkmalpflege im Rheinland, *Prospektions- und Grabungsrichtlinien für drittfinanzierte archäologische Maßnahmen*, 2011
http://www.bodendenkmalpflege.lvr.de/denkmalenschutz_prakt_bodendenkmalpflege/richtlinien_2011.pdf
(aufgerufen, September 2012)

Materials Thesaurus – British Museums
http://www.collectionslink.org.uk/assets/thesaurus_bmm/matintro.htm?phpMyAdmin=OYNyINPdn3sQmoXugKH1qcCLSW0
http://www.collectionslink.org.uk/assets/thesaurus_bmm/matintro.htm?phpMyAdmin=OYNyINPdn3sQmoXugKH1qcCLSW0

Mayke Wagnera, Pavel Tarasovb, Dominic Hosnera, Andreas Flecka, Richard Ehricha, Xiaocheng Chena, Christian Leipeb, Mapping of the spatial and temporal distribution of archaeological sites of northern China during the Neolithic and Bronze Age, Quaternary International, Available online 2 July 2012,
<http://dx.doi.org/10.1016/j.quaint.2012.06.039>.

MuseData – an electronic library of classical music scores
<http://musedata.org/> (aufgerufen, September 2012)

Musical Instrument Museums Online
<http://www.mimo-international.com/> (aufgerufen, September 2012)

Music Encoding Initiative
<http://music-encoding.org> (aufgerufen, August 2012)

MusicXML
<http://www.makemusic.com/musicxml> (aufgerufen, September 2012)

Object Names Thesaurus – British Museums
http://www.collectionslink.org.uk/assets/thesaurus_bmon/Objintro.htm

Open Geospatial Consortium
<http://www.opengeospatial.org/>

Online Cultural Research environment (OCHRE)
http://ochre.lib.uchicago.edu/index_files/Page845.htm (aufgerufen, September 2012)

Online Georeferencing Tool for scanned Maps
<http://www.georeferencer.org/>

Panofsky, Erwin: Ikonographie und Ikonologie, in: Kaemmerling, E. (Hg.): Bildende Kunst als Zeichensystem. Ikonographie und Ikonologie, Bd. 1. Theorien – Entwicklung – Probleme, Köln 1994.

Pleiades
<http://pleiades.stoa.org/Members/sgillies/news-items/linked-data-for-pleiades-places>

Prometheus Künstlernamenansetzungsdatei
<http://prometheus-bildarchiv.de/de/tools/pknd>

Quantum GIS
<http://www.qgis.org/>

Répertoire d'autorité-matière encyclopédique et alphabétique unifié (RAMEAU)
<http://rameau.bnf.fr/>

Répertoire International d'Iconographie Musicale
<http://www.ridim-deutschland.de/ridim/index.php?pcontent=startseite>

Richards, J. D., *From anarchy to good practice: the evolution of standards in archaeological computing*, Archeologia e Calcolatori, 2009,
<http://eprints.whiterose.ac.uk/10707/> (aufgerufen, August 2012)

Schlagwortnormdatei
http://www.dnb.de/DE/Standardisierung/Normdaten/GND/gnd_node.html

Schloen, J., *Archaeological Data Models and Web Publication Using XML, Computers and the Humanities*, 2001, <http://www.springerlink.com/content/jk1q91115506k6j5/abstract/> (aufgerufen, August 2012)

Selfridge-Field, E. (Hrsg.), *Beyond Midi: The Handbook of Musical Codes*, 1997

SPECTRUM: the UK Museum Collections Management Standard
<http://www.collectionslink.org.uk/programmes/spectrum> (aufgerufen, September 2012)

Spree, Ulrike, Online-Tutorial
<http://www.bui.haw-hamburg.de/pers/ulrike.spree/remind/ueberblick.htm>

Thesaurus of Monument Types – English Heritage

http://thesaurus.english-heritage.org.uk/thesaurus.asp?thes_no=1

Text Encoding Initiative

<http://www.tei-c.org>

UK Data Archive, Study-level data documentation

<http://www.data-archive.ac.uk/create-manage/document/study-level> (aufgerufen, September 2012)

Union List of Artist Names

<http://www.getty.edu/research/tools/vocabularies/ulan/about.html>

Verband der Landesarchäologen in der Bundesrepublik Deutschland

<http://www.landesarchaeologen.de/verband/kommissionen/archaeologie-und-informationssysteme/projektarbeiten/gruppen/adex/> (aufgerufen, August 2012)

Verband der Landesarchäologen in der Bundesrepublik Deutschland, *Ausgrabungen und Prospektion - Durchführung und Dokumentation*, 2006

http://www.landesarchaeologen.de/fileadmin/Dokumente/Dokumente_Kommissionen/Dokumente_Grabungstechniker/grabungsstandards_april_06.pdf (aufgerufen, August 2012)

Virtual International Authority File (VIAF)

<http://viaf.org/>

Willems, W.J.H., Brandt, R.W., *Dutch Archaeology Quality Standard*, Den Haag, 2004

http://www.dans.knaw.nl/sites/default/files/file/archief/KNA_uk.pdf (aufgerufen, September 2012)

Wissenschaftsrat, *Empfehlungen zur Weiterentwicklung der wissenschaftlichen Informationsinfrastrukturen in Deutschland bis 2020*, 2012

<http://www.wissenschaftsrat.de/download/archiv/2359-12.pdf> (aufgerufen, September 2012)

6.2 Abkürzungen

AAT	Art & Architecture Thesuarus
ADeX	Archäologischer DateneXport-Standard
BSZ	Bibliotheksservice-Zentrum Baden-Württemberg
CARARE	Connecting ARchaeology and ARchitecture in Europeana
CCARH	Center for Computer Assisted Research in the Humanities
CDWA	Categories for the Description of Works of Arts
CEI	Charters Encoding Initiative
CONA	Cultural Objects Name Authority
DC	Dublin Core
DDC	Dewey-Dezimalklassifikation
DFG	Deutsche Forschungsgemeinschaft
EAD	Encoded Archival Description
DNB	Deutsche Nationalbibliothek
EDTF	Extended Date / Time Format
EpiDoc	Epigraphic Documents in TEI XML
FRBR	Functional Requirements for Bibliographic Records
GIS	Geoinformationssystem
GND	Gemeinsame Normdatei
INSPIRE	Infrastructure for Spatial Information in the European Community
KML	Keyhole Markup Language
LCSH	Library of Congress Subject HEadings
LIDO	Lightweight Information Describing Objects
LVR	Landschaftsverband Rheinland
MEI	Music Encoding Initiative
METS	Metadata Encoding and Transmission Standard
MIDAS	A Manual and Data Standard for Monument Inventories
MODS	Metadata Object Description Schema
MXML	Manuscriptum XML
OCHRE	Online Cultural Research environment
OGC	Open Geospatial Consortium
PKND	Prometheus Künstlernamenansetzungsdatei
RAMEAU	Répertoire d'autorité-matière encyclopédique et alphabétique unifié
RDF	Resource Description Framework
RidM	<i>Répertoire International d'Iconographie Musicale</i>
SAFT	Standard-Austauschformat
SWD	Schlagwortnormdatei

TEI	Text Encoding Initiative
TGN	Getty Thesaurus of Geographic Names
ULAN	Union List of Artist Names
URI	Uniform Resource Identifier
UTM	Universal Transverse Mercator
VIAF	Virtual International Authority File
XML	Extensible Markup Language

6.3 Zusammenstellung geisteswissenschaftlicher Quellentypen

Die folgende Auflistung soll einen ersten Überblick über die vielfältig verwendeten Quellen- und Datentypen in den Geisteswissenschaften aufzeigen. Diese Auflistung hat keinen Anspruch auf Vollständigkeit, sondern soll vielmehr die vielfältigen Ausprägungen aufzeigen, wie sie beispielsweise in der Judaistik, Theologie, Geschichtswissenschaft, Archäologie, Musikwissenschaft bekannt sind.

„Schriftliche“ Quellen

- Analysen
- Aufführungsberichte
- Bekenntnisschriften
- Bibelkommentare, Exegetische Abhandlungen
- Biographie (Autobiographie)
- Superkommentare (Kommentar zum Kommentar)
- Briefe
- Chroniken
- Findbücher
- Flugblätter
- Gebetbücher
- Gelegenheitsschriften ("Diss"-Projekt)
- Genizot ("Schriftgräber")
- Gesetze, Verordnungen, Anordnungen
- Grabungstagebücher
- Grundbücher
- Inschriften (Grabsteine, Baudenkmäler usw.)
- Editionsgeschichte: frühere Abschriften/Abklatsch
- Kalender
- Kataloge (Werkkataloge, Ausstellungskataloge)
- Katechismen
- Kirchenbücher
- Kirchenordnungen
- Kirchliches und staatliches Schrift- und Verwaltungsgut
- Konzertanzeigen
- Gemeindeakten
- Akten von öffentlichen und privaten Wohlfahrtseinrichtungen
- Vereinsakten
- Korrespondenz
- Libretto
- Liturgie
- Memorbücher
- Notizen

- Päpstliche Bullen und sonst. Urkunden
- Plakat (Veranstaltungsankündigung)
- Poesie (mit Untergattungen)
- Predigten, Leichenpredigten usw.
- Programmzettel
- Protokolle
- Publikationen/Bücher
- Reiseberichte
- Rezensionen
- Rollenheft
- Soufflierbuch / Regiebuch
- Streitschriften
- Subskribentenlisten
- Synodale oder parlamentarische Beschlüsse und Handlungsanweisungen
- Tagebücher
- Tauf- und Geburtsregister
- Textbuch
- Tonsatzlehren
- Sterberegister
- Übersetzungen
- Urbar
- Urkunden
- Urteile
- Verfassungen & Gesetze
- Statuten
- Verträge
- Vorreden, Vorworte, etc.
- Vorlesungsverzeichnisse und Matrikel
- Wörterbücher, Lexika, Enzyklopädien

Forschungsliteratur

- Monographien
- Handbücher
- Aufsätze
- Artikel
- Enzyklopädien
- Wörterbücher
- Bibliografien

Visuelle Quellen

- Gedruckte Reproduktionen
- (Kupfer)-Stiche
- Bauzeichnungen

- Filme
- Fotografien
- Glasplatten
- Gemälde
- Malereien
- Plakate
- Sigel
- Technische Skizzen
- Zeichnungen

Audiodokumente

- Interviews
- Musik: Mitschnitt
- Oral-History Interviews
- Radiosendungen
- Reden und Vorträge

Musik (notierte)

- Albumblatt
- Arrangement
- Klavierauszug
- Particell
- Partitur
- Skizze u. Entwurf
- Stimmen

„Empirisches“ Datenmaterial

- Haushaltspläne
- Kassenbücher
- Kollektenbücher
- Produktionszahlen
- Rechnungen
- Statistiken
- Statistisches Jahrbuch des Deutsch-Israelitischen Gemeindebundes
- Subskribentenlisten
- Warenlisten und -aufstellungen
- Wetter- und Klimadaten

3D-Objekte

- Denkmäler, Architektur, Aufnahmen von Steindenkmälern
- Gebrauchsgegenstände
- Globus

- Häuser
- Interieur
- Kirchenfenster
- Kunst
- Ornamente - Symbolik
- Mobiliar
- Münzen
- Musikinstrumente
- Reliquien
- Schmuck und Kunstwerke
- Plastik
 - Idealplastik (Götter; Mythologische Gestalten; Nicht-mythologische Statuen; Tiere; Fabeltiere / Mischwesen)
 - Porträtplastik (Griechisch; Etruskisch-italisch; Römisch)
 - Sarkophage (Griechisch; Etruskisch; Republikanisch-frühkaiserzeitlich; Kaiserzeitlich; Christlich)
 - Reliefs (Griechisch; Griechisch-unteritalisch; Etruskisch; Römisch)
 - Figürliche Bauplastik (Griechenland; Grossgriechenland; Italien; Stützfiguren; Theaterdekorationen; Bauteile; Historische Reliefs)
 - Inschriften (Griechisch; Lateinisch)
 - Stuck (Etruskisch; Römisch)
- Malerei (Etruskisch; Italisch / Grossgriechisch; Pompejanisch; Römisch)
- Mosaik (Schwerpunkt: Italien; Römische Provinzen)
- Vasen (Protogeometrisch; Geometrisch; Protokorinthisch; Korinthisch; Ostgriechisch; Westgriechisch; Lakonisch; Chalkidisch; Boiotisch; Attisch; Italisch-Etruskisch; Etruskisch; Unteritalisch; Hellenistisch; Römisch)
- Terrakotta (Schwerpunkt Italien: Figürlich; Architektonisch)
- Bronze (Statuetten; Spiegel; Waffen; Geräte; Gefässe)
- Kleinkunst (Glas; Edelstein; Gemmenabdrücke)

Karten

- Flurkarten
- Georeferenzierbare Orte, z.B. Plätze, Städte
- Byzanz - Konstantinopel - Istanbul
- nicht georeferenzierbar, z.B. El Dorado / Eden / Zion
- Landkarten
- Planskizzen (z.B. PDF, SVG, Auto-Cad)
- Navigationskarten
- Wetterkarten

6.4 MEI Beispiel (Robert Schumann: *Der Abendstern* op. 79/1)¹³²

```
<?xml version="1.0" encoding="UTF-8"?>
<?xml-model href="mei-all.rng" type="application/xml"
schematypens="http://relaxng.org/ns/structure/1.0"?>
<mei xmlns:xlink="http://www.w3.org/1999/xlink"
xmlns="http://www.music-encoding.org/ns/mei"
meiversion="2012">
  <meiHead>
    <fileDesc>
      <titleStmt>
        <title>Der Abendstern : an electronic transcription</title>
        <respStmt>
          <resp>Composed by:</resp>
          <persName role="composer" authority="GND"
authURI="http://d-nb.info/gnd"
dbkey="118611666">Robert Schumann</persName>
          <resp>Machine-readable transcription by:</resp>
          <persName role="encoder">Kristina Richts</persName>
        </respStmt>
      </titleStmt>
      <pubStmt>
        <unpub/>
      </pubStmt>
      <sourceDesc>
        <source n="1" xml:id="op79">
          <titleStmt>
            <title type="main">Lieder-Album für die Jugend für
Singstimme(n) und Klavier
op. 79</title>
            <title type="uniform">Lieder für die Jugend</title>
          </titleStmt>
          <respStmt>
            <persName role="composer" authority="GND"
authURI="http://d-nb.info/gnd"
dbkey="118611666">Robert Schumann</persName>
            <resp>Herausgegeben von</resp>
            <persName role="editor" authURI="http://d-nb.info/gnd"
authority="GND"
dbkey="10400097X">Ulrich Mahlert</persName>
          </respStmt>
          <editionStmt>
            <edition>Reprint der Erstausgabe Leipzig 1849</edition>
          </editionStmt>
        </source>
      </sourceDesc>
    </fileDesc>
  </meiHead>

```

¹³² Das nachfolgende Beispiel stammt von Kristina Richts, Mitarbeiterin des DFG/NEH-Projekts *Digital Music Notation Data Model and Prototype Delivery System*.

```

<respStmt>
  <corpName role="publisher" authority="GND"
    authURI="http://d-nb.info/gnd"
    dbkey="330001418">Breitkopf und Härtel</corpName>
  <resp>Druck:</resp>
  <corpName role="printer">PIROL Notendruckerei,
    Minden</corpName>
</respStmt>
<address>
  <addrLine>Wiesbaden</addrLine>
  <addrLine>Germany</addrLine>
</address>
<identifrier type="ordernumber">Breitkopf & Härtel
  8307</identifrier>
<availability>
  <useRestrict>© 1991 by Breitkopf & Härtel,
    Wiesbaden</useRestrict>
</availability>
</pubStmt>
<physDesc>
  <dimensions>30 cm</dimensions>
  <extent>57 S.</extent>
  <plateNum>8062</plateNum>
  <titlePage n="1">
    <p>Breitkopf Archiv</p>Frühdrucke und bibliophile Ausgaben
    von hohem Quellenwert</p>Early printings and
    bibliophiles' editions of a high documentary value</p>
    <p>herausgegeben von / edited by</p>Joachim Draheim</p>
    <p>Robert Schumann</p>(1810-1856)</p>
    <p>Lieder-Album für die Jugend</p>für Singstimme(n) und
    Klavier</p>for Voice(s) and Piano</p>op. 79</p>
    <p>Reprint der Erstausgabe Leipzig 1849</p>Reprint of the
    First Printing Leipzig 1849</p>
    <p>herausgegeben von / edited by</p>Ulrich Mahlert</p>
    <p>Breitkopf & Härtel, Wiesbaden</p>Edition Breitkopf
    8307</p>Printed in Germany</p>
  </titlePage>
  <titlePage n="2">
    <p>Lieder</p>für die </p>Jugend</p>von</p>Robert
    Schumann</p>op. 79</p>
    <p>Leipzig, Breitkopf & Härtel.</p>
    <p>C.Hahn.lith.</p>
  </titlePage>
</physDesc>
<seriesStmt>
  <title>Breitkopf-Archiv</title>
  <respStmt>
    <resp>Herausgegeben von:</resp>
    <persName role="editor" authURI="http://d-nb.info/gnd"

```

```

        authority="GND" dbkey="123921228">Joachim
        Draheim</persName>
    </respStmt>
</seriesStmt>
<classification>
    <termList>
        <term classcode="#DNB_Sachgruppen">
            <identifizier>02a</identifizier>
            <term>Werke für Einzelstimmen</term>
        </term>
    </termList>
    <classCode xml:id="DNB_Sachgruppen"/>
</classification>
<contents>
    <head>INHALT.</head>
    <label>Kinderlieder von Hoffmann von Fallersleben.</label>
    <contentItem n="1">1. Der Abendstern.</contentItem>
    <contentItem n="2">2. Schmetterling.</contentItem>
    <contentItem n="3">3. Frühlingsbotschaft.</contentItem>
    <contentItem n="4">4. Frühlingsgruss.</contentItem>
    <contentItem n="5">5. Schlaraffenland.</contentItem>
    <contentItem n="6">6. Sonntag.</contentItem>
    <contentItem n="7">Zwei Zigeunerliedchen, aus dem Spanischen
        von E. Geibel.</contentItem>
    <contentItem n="8">Des Knaben Berglied, v. L.
        Umland.</contentItem>
    <contentItem n="9">Mailied.</contentItem>
    <contentItem n="10">Käuzlein, aus des Knaben
        Wunderhorn.</contentItem>
    <contentItem n="11">Hinaus in's Freie, von Hoffmann von
        Fallersleben.</contentItem>
    <contentItem n="12">Der Sandmann, von H. Kletke.</contentItem>
    <contentItem n="13">Marienwürmchen, aus des Knaben
        Wunderhorn.</contentItem>
    <contentItem n="14">Die Waise, von Hoffmann von
        Fallersleben.</contentItem>
    <contentItem n="15">Das Glück, von F. Hebbel.
        Zweistimmig.</contentItem>
    <contentItem n="16">Weihnachtlied, von H. C.
        Andersen</contentItem>
    <contentItem n="17">Die wandelnde Glocke, von
        Goethe.</contentItem>
    <contentItem n="18">Frühlingslied, von Hoffmann von
        Fallersleben.</contentItem>
    <contentItem n="19">Frühlings Ankunft, von Hoffmann von
        Fallersleben.</contentItem>
    <contentItem n="20">Die Schwalben. Fliegendes Blatt.
        Zweistimmig.</contentItem>
    <contentItem n="21">Kinderwacht. Fliegendes

```

```

        Blatt.</contentItem>
    <contentItem n="22">Des Sennen Abschied, von
        Schiller.</contentItem>
    <contentItem n="23">Er ist's, von E. Mörike.</contentItem>
    <contentItem n="24">Spinnelied. Fliegendes Blatt.
        Dreistimmig.</contentItem>
    <contentItem n="25">Des Buben Schützenlied, von F.
        Rückert.</contentItem>
    <contentItem n="26">Schneeglöckchen, von F.
        Rückert.</contentItem>
    <contentItem n="27">Lied Lynceus des Thürmers, aus Goethe's
        Faust.</contentItem>
    <contentItem n="28">Mignon, von Goethe.</contentItem>
</contents>
</source>
</sourceDesc>
</fileDesc>
<workDesc>
    <work xml:id="op.79_1">
        <titleStmt>
            <title type="main">Der Abendstern</title>
            <title type="uniform">Lieder für die Jugend &lt;Der
                Abendstern&gt;</title>
        <respStmt>
            <persName role="composer" authority="GND"
                authURI="http://d-nb.info/gnd"
                dbkey="118611666">Robert Schumann</persName>
            <persName role="lyricist" authority="GND"
                authURI="http://d-nb.info/gnd"
                dbkey="118552589">August Heinrich Hoffmann von
                Fallersleben</persName>
        </respStmt>
    </titleStmt>
    <key>A major</key>
    <tempo>Langsam.</tempo>
    <meter meter.count="2" meter.unit="4"/>
    <perfMedium>
        <performer n="1">
            <instrVoice>Singstimme.</instrVoice>
        </performer>
        <performer n="2">
            <instrVoice>Pianoforte.</instrVoice>
        </performer>
    </perfMedium>
    <classification>
        <termList>
            <term classcode="#DNB_Sachgruppen">
                <identifier>02a</identifier>
                <term>Werke für Einzelstimmen</term>
            </term>
        </termList>
    </classification>

```

```

        </term>
    </termList>
    <termList>
        <term classcode="#OCLC_DDC">
            <identifier>782.42</identifier>
            <term>Song</term>
        </term>
    </termList>
    <classCode xml:id="OCLC_DDC" authority="OCLC"
        authURI="http://www.oclc.org/dewey/resources/
        summaries/default.htm#700"/>
</classification>
<incip>
    <incipText>
        <p>Du lieblicher Stern</p>
    </incipText>
    <score>
        <scoreDef>
            <staffGrp>
                <staffDef n="1" label="Singstimme" lines="5"
                    meter.count="2"
                    meter.unit="4" clef.shape="G" clef.line="2"
                    key.sig="3s"/>
                <staffGrp label="Pianoforte" symbol="brace">
                    <staffDef n="2" lines="5" meter.count="2"
                        meter.unit="4"
                        clef.shape="G" clef.line="2" key.sig="3s"/>
                    <staffDef n="3" lines="5" meter.count="2"
                        meter.unit="4"
                        clef.shape="F" clef.line="4" key.sig="3s"/>
                </staffGrp>
            </staffGrp>
        </scoreDef>
        <section>
            <measure n="0" metcon="false">
                <staff n="1">
                    <layer>
                        <note pname="a" oct="4" dur="8">
                            <verse n="1">
                                <syl>V. 1. Du</syl>
                            </verse>
                            <verse n="2">
                                <syl>V. 2. Wie</syl>
                            </verse>
                            <verse n="3">
                                <syl>V. 3. So</syl>
                            </verse>
                            <verse n="4">
                                <syl>V. 4. Wie</syl>
                            </verse>
                        </note>
                    </layer>
                </staff>
            </measure>
        </section>
    </score>
</incip>

```

```

        </verse>
        </note>
        <dynam place="below" tstamp="0.5">p</dynam>
        <dir place="above">Langsam.</dir>
    </layer>
</staff>
<staff n="2">
    <layer>
        <note pname="a" oct="4" dur="8"/>
        <dynam place="below" tstamp="0.5">p</dynam>
    </layer>
</staff>
<staff n="3">
    <layer>
        <note pname="a" oct="3" dur="8"/>
    </layer>
</staff>
<slur tstamp="0.5" curvedir="below" dur="1m+2.5"
    staff="2"/>
</measure>
<measure n="1">
    <staff n="1">
        <layer>
            <note pname="c" oct="5" dur="4">
                <verse n="1">
                    <syl con="d">lieb</syl>
                </verse>
                <verse n="2">
                    <syl>lieb'</syl>
                </verse>
                <verse n="3">
                    <syl>blick'</syl>
                </verse>
                <verse n="4">
                    <syl>nickst</syl>
                </verse>
            </note>
            <note pname="b" oct="4" dur="8">
                <verse n="1">
                    <syl con="d">li</syl>
                </verse>
                <verse n="2">
                    <syl>ich</syl>
                </verse>
                <verse n="3">
                    <syl>ich</syl>
                </verse>
                <verse n="4">
                    <syl>du</syl>
            </note>
        </layer>
    </staff>

```

```

        </verse>
    </note>
    <note pname="a" oct="4" dur="8">
        <verse n="1">
            <syl>cher</syl>
        </verse>
        <verse n="2">
            <syl>doch</syl>
        </verse>
        <verse n="3">
            <syl>nach</syl>
        </verse>
        <verse n="4">
            <syl>mir</syl>
        </verse>
    </note>
</layer>
</staff>
<staff n="2">
    <layer>
        <chord dur="4">
            <note pname="c" oct="4"/>
            <note pname="e" oct="4"/>
            <note pname="a" oct="4"/>
            <note pname="c" oct="5"/>
        </chord>
        <beam>
            <chord dur="8">
                <note pname="d" oct="4"/>
                <note pname="e" oct="4"/>
                <note pname="g" oct="4"/>
                <note pname="b" oct="4"/>
            </chord>
            <chord dur="8">
                <note pname="e" oct="4"/>
                <note pname="a" oct="4"/>
            </chord>
        </beam>
    </layer>
</staff>
<staff n="3">
    <layer>
        <chord dur="4">
            <note pname="a" oct="2"/>
            <note pname="e" oct="3"/>
            <note pname="a" oct="3"/>
        </chord>
        <beam>
            <chord dur="8">

```

```

        <note pname="b" oct="2"/>
        <note pname="e" oct="3"/>
    </chord>
    <chord dur="8">
        <note pname="c" oct="3"/>
        <note pname="e" oct="3"/>
        <note pname="a" oct="3"/>
    </chord>
    </beam>
</layer>
</staff>
</measure>
<measure n="2">
    <staff n="1">
        <layer>
            <note pname="f" oct="4" dur="4" dots="1">
                <verse n="1">
                    <syl>Stern,</syl>
                </verse>
                <verse n="2">
                    <syl>dich</syl>
                </verse>
                <verse n="3">
                    <syl>dir,</syl>
                </verse>
                <verse n="4">
                    <syl>zu</syl>
                </verse>
            </note>
        </layer>
    </staff>
    <staff n="2">
        <layer>
            <chord dur="2">
                <note pname="d" oct="4"/>
                <note pname="f" oct="4"/>
            </chord>
        </layer>
    </staff>
    <staff n="3">
        <layer>
            <chord dur="2">
                <note pname="d" oct="3"/>
                <note pname="a" oct="3"/>
            </chord>
        </layer>
    </staff>
</measure>
</section>

```

```
        </score>
      </incip>
    </work>
  </workDesc>
</meiHead>
<music/>
</mei>
```