

E-Papers der Archivschule Marburg

Hochschule für Archivwissenschaft

Nr. 3

Bernhard Homa

Voraussetzungen und Kriterien für den Einsatz von Normdaten im Landesarchiv NRW

Transferarbeit des 50. wissenschaftlichen Lehrgangs
an der Archivschule Marburg

Betreuer der Archivschule Marburg:
Betreuerin des Ausbildungsarchivs:

Dr. Dominik Haffer
Dr. Martina Wiech

Marburg/Lahn 2019

HESSEN



Die Archivschule Marburg ist eine Einrichtung des Landes Hessen.

Voraussetzungen und Kriterien für den Einsatz von Normdaten im Landesarchiv NRW

von Bernhard Homa

1. Einleitung: Ziel, Methodik, Motivation

Ziel nachfolgender Untersuchung* ist es:

- 1.) Voraussetzungen für den Einsatz von „Normdaten“ für die archivischen Fachaufgaben, insbesondere die Erschließung, im LAV NRW zu prüfen;
- 2.) Kriterien für die Auswahl geeigneter Bestände und Findmittel für die Anreicherung mit Normdaten zu entwickeln, insbesondere als Vorbereitung zur empirischen Überprüfung des erwartbaren Mehrwertes von Normdaten;¹
- 3.) Abgeleitet aus 1.) und 2.) einige Handlungsempfehlungen aufzustellen.

Im LAV NRW bestanden bei Abfassung dieser Arbeit keine praktischen Vorerfahrungen in der Arbeit mit Normdaten. Zugleich gibt es hinsichtlich der organisatorischen und technischen Voraussetzungen bestimmte nicht oder nur langsam veränderbare Faktoren. Beides beeinflusste die Vorgehensweise und die zugrundeliegenden Mittel in der Beantwortung der Fragestellung.

Folgende Punkte werden behandelt:

- 1.) Generalia zur Normdaten-Thematik: Definition, Typologie; Vorteile und Nachteile; Notwendigkeit der Befassung mit der Thematik (Kap. 2);
- 2.) Die GND als derzeit einzig praktikables Normdatenrepositorium für Archive (Kap. 3);
- 3.) Voraussetzungen und Probleme bei der Anwendung von Normdaten; Entwicklung einer Bewertungsskala zur Ermittlung der besonders geeigneten Bestände/Findmittel² (Kap. 4);
- 4.) Fazit: Handlungsempfehlungen, insbesondere organisatorische und technische Aspekte im Hinblick auf die benannten Bedingungen und Kriterien (Kap. 5).

Diese Arbeit behandelt *nicht*:

- 1.) Konkrete betriebswirtschaftliche Kosten-Nutzen-Analysen für den Einsatz von Normdaten;
- 2.) Konkrete Fragen der technischen Umsetzung.

Folgende Mittel wurden für die Erarbeitung angewendet:

- 1.) Literaturrecherche;
- 2.) Umfragen und Besprechungen:
 - a) intern: zu geeigneten Beständen/Findmitteln
 - b) extern: zu Vorerfahrungen mit Normdaten als auch zur Bestimmung geeigneter Bestände/Findmittel;
- 3.) Eigene Überprüfung innerhalb des Archivinformationssystems (AIS) des LAV.

* Diese Studie hätte ohne die freundliche Mitwirkung zahlreicher Ansprechpartner und Hinweisgeber nicht entstehen können. Allen namentlich erwähnten als auch den nicht zusätzlich erwähnten Personen sei daher summarisch an dieser Stelle nochmals herzlich gedankt.

¹ Die Notwendigkeit einer solchen Priorisierung und darauf aufbauenden Prüfung unter den Bedingungen im LAV NRW wird näher in Kapitel 2.3. und 4 erläutert.

² Im LAV NRW herrschte bei Abfassung dieser Studie teilweise noch eine n:n-Beziehung zwischen Beständen und Findmitteln, sodass ein Findmittel auch mehrere Bestände umfassen konnte. Dies soll perspektivisch geändert werden, hat aber die Folge, dass die Frage der Eignung sich je nach konkretem Fall auf Bestand, Findmittel oder beides beziehen kann.

Die Gründe, warum es für das LAV NRW sinnvoll ist, sich eingehender mit der Normdaten-Thematik zu befassen, werden im Kap. 2.3 erläutert – ohne einige Vorbemerkungen in Kap. 2.1. und 2.2. wäre die Motivation nicht verständlich.

2. Normdaten: Generalia und Anwendung im Archivwesen

2.1. Definition, Standards, Typen

Für Normdaten gibt es verschiedene Definitionen,³ zentral sind jeweils drei Punkte:

- 1.) Identifizierung beziehungsweise Individualisierung von Entitäten;
- 2.) Damit verbunden die Möglichkeit der Indizierung und Verknüpfung zwischen verschiedenen Informationsressourcen im Semantic Web;
- 3.) Die Ansetzung dieser Normdaten nach klar festgelegten Regeln (Ontologien, normierte Vokabulare).

Das Moment der Normierung trifft nun auf Normdateien selbst gerade nicht zu: Es gibt inzwischen eine Vielzahl unterschiedlicher Normdatensystematiken, die untereinander – erklärlich durch unterschiedliche Entstehungskontexte und Zielsetzung – keineswegs immer kompatibel sind. Diese diversen Systematiken betreffen zum Beispiel Personen, Orte, künstlerische Motive, Berufe, Sachschlagworte und vieles mehr.⁴ Dominierend sind Normdaten bekanntermaßen im Bibliotheksbereich, wo es neben den nationalen beziehungsweise nationalsprachlichen Normdatenbeständen wie der GND⁵ auch die Aggregation im internationalen Repository VIAF gibt,⁶ sowie mit RDA einen eigenen Standard für die Ansetzung.⁷ Auch der ICA hat mit ISAAR(CPF) und dem zugehörigen Format EAC einen eigenen Normdatenstandard entwickelt.

2.2. Anwendung im Archivwesen: Vorteile und Nachteile

Die Vorteile einer Verwendung von Normdaten⁸ sind auch im deutschen Archivwesen mittlerweile grundsätzlich anerkannt⁹. Hauptsächlich – teilweise miteinander verbunden – zu nennen sind:

- 1.) Sichtbarkeit: Die eigenen Erschließungsdaten und Informationsressourcen werden von einem breiteren Personenkreis wahrgenommen.
- 2.) Recherche: Bestands-, archiv- und spartenübergreifende Recherche ist möglich, diese ist zudem durch die spezifische Indizierung viel präziser als bei einfachen Volltextrecherchen.¹⁰ Dieser Aspekt hängt mit den folgenden beiden zusammen.

³ Cf. exemplarisch VOSS u.a. 2014, S. 11; AUMANN 2016; ZEDLITZ 2017, bes. S. 22; WETTLAUFER / WESTPHAL 2014, S. 274 f.

⁴ Ohne Anspruch auf Vollständigkeit (cf. auch die in voriger Anm. genannte Literatur): für Berufe: HISCO (socialhistory.org/en/projects/hisco-history-work) [letzter Zugriff 16.2.2017]; für künstlerische Motive: Iconclass (iconclass.nl/home) [letzter Zugriff 16.02.2017]; Orte: GeoNames (geonames.org/) [letzter Zugriff 16.02.2017]; für Personen: CERL Thesaurus (thesaurus.cerl.org/cgi-bin/search.pl) [letzter Zugriff 16.2.2017]; für die Wikidata: Voss u.a. 2014. Zur leichten Einsichtnahme und direkten Nutzung sehe man bitte die Link-Sammlung auf S. 28–31.

⁵ Als Einstieg statt vieler: wiki.dnb.de/display/ILTIS/Informationsseite+zur+GND [letzter Zugriff 16.2.2017]; ZEDLITZ 2017, S. 23 f. Zu den Abkürzungen s. S. 24 f.

⁶ www.viaf.org/ [letzter Zugriff 16.2.2017].

⁷ Als Einstieg statt vieler: wiki.dnb.de/display/RDAINFO/RDA-Info [16.2.2017].

⁸ Cf. aus bibliothekarischer Sicht statt vieler etwa SCHEVEN 2015.

⁹ Cf. exemplarisch SPERL 2016; BRÜBACH 2015A, S. 166–168; DERS. 2015B; MÜLLER 2015, bes. 13 f. und öfter; KRAUTH 2015, S. 8; KLUTTIG 2014; für das Bundesarchiv SPRAU 2014, S. 25–30; diverse Beiträge im Blog „Archivalia“, exemplarisch GRAF 2015 und DERS. 2013A (dort bes. auch die Diskussion); in mehrfacher Hinsicht instruktiv ist ZIWES 2014; aus bibliothekarischer Sicht: ZUNSTEIN 2012. Die folgenden Ausführungen sind auch Quintessenz aller Gesprächsvermerke und ausgewerteten Fragebögen.

- 3.) Befriedigung von Nutzerinteressen: Gerade Personen und Orte sind klassische Einstiegspunkte in die Recherche durch bestimmte Nutzergruppen; zudem besteht vielfach der Wunsch nach Zusammenführung und Kontextualisierung von Entitäten, die in unterschiedlichen Schreibweisen vorliegen.¹¹
- 4.) Verknüpfung: Informationsressourcen ganz unterschiedlicher Provenienz können über eine Entität als Sucheinstieg zusammengeführt werden. Auf diese Weise können Nutzer in wenigen Schritten auf Angebote gelangen, die ihnen zuvor nie bekannt waren und die sie bei klassischen Recherchetechniken auch nicht finden würden – denn im Gegensatz etwa zu relevanzbasierten Suchmaschinen werden über Entitäten verknüpfte Ressourcen grundsätzlich gleichwertig behandelt.
- 5.) Verbesserte maschinelle Auswertungsmöglichkeiten: Normdaten und die mit diesen verknüpften Ressourcen eignen sich dank der eindeutigen Referenzierung über Identifikatoren hervorragend für Forschungsansätze der digitalen Geisteswissenschaften beziehungsweise als „Linked Open Data“.¹²

Diese Vorteile betreffen grundsätzlich sowohl die archivische Anbieter- als auch die Nutzerseite, wie aus den Umfragen des Verfassers deutlich wird.¹³ Innerhalb des Archivwesens wäre zusätzlich noch die Verwendung für mehrere Fachaufgaben – Erschließung, Überlieferungsbildung, Benutzung – zu nennen.¹⁴

Hauptargument gegen die Verwendung von Normdaten ist – wenig überraschend – die Ressourcenproblematik und hier insbesondere die Frage, ob Aufwand und Ertrag in angemessenem Verhältnis zueinander stehen.¹⁵ Letztlich wiederholt sich dabei die ältere Diskussion um Sinn und Unsinn von Tiefenerschließung sowie die Indizierung als Teil von letzterer. Hierbei wurde und wird von den Kritikern ins Feld geführt, dass sich der Mehrwert für die Recherche nicht bemessen lasse, die Nutzung von Beständen nur punktuell sei und die Nutzer etwa Sachschlagworte kaum oder nur unvollkommen für die Recherche einsetzen.¹⁶ Eine quantifizierende Kos-

¹⁰ Dieser Mehrwert übersteigt hinsichtlich der indizierten Begriffe deutlich denjenigen eines Enthältvermerks, dessen Erstellung zudem deutlich länger dauert und demnach teurer ist. Sofern Sachschlagworte verwendet werden, könnten starke organisatorische Veränderungen, wie sie für moderne Verwaltungen typisch sind, über funktionsbezogene Sachschlagworte abgefangen werden.

¹¹ Cf. speziell zu dieser für ‚Familienforscher‘ typischen Erwartung RIEDEL 2013, S. 52; unabhängig von der Realisierbarkeit ist dieses Interesse jedenfalls quantitativ ernst zu nehmen.

¹² Cf. NEUDECKER / REHM 2016, S. 111 ff. (mit Beispielen); JAGODZINSKI 2016, S. 94–96; MÜLLER 2015, bes. S. 8–10, 14–16 (mit Beispielen, etwa Graphen im Portal Kalliope); WETTLAUER / WESTPHAL 2014, S. 275 f.; SANDNER 2013, S. 31 f. und öfter. Die Verbreitung von LOD ist im Archivwesen noch gering, was teilweise mit Rechtsfragen (Schutzfristen, Verpflichtung zur Nennung der Provenienz) zu tun hat; „Open Data“, aber nicht LOD bieten etwa das Niederländische Nationalarchiv (<gahetna.nl/opedata> [letzter Zugriff 16.2.2017]) und das Bundesarchiv (<www.bundesarchiv.de/fachinformationen/05230/index.html.de> [letzter Zugriff 16.2.2017]), wobei bei den meisten Ressourcen die faktisch Gemeinfreiheit bedeutende CC0-Lizenz verwendet wird.

¹³ So bewerteten von neun Teilnehmern vier den Vorteil von Normdaten für die Fachaufgabe Erschließung als hoch, vier weitere zumindest als mittig; für die Benutzung sahen sogar sieben von neun einen hohen Vorteil. Für die anderen vorgegebenen Fachaufgaben sahen dagegen hinsichtlich der Überlieferungsbildung vier und hinsichtlich der Öffentlichkeitsarbeit fünf keinerlei Vorteil, nur jeweils einer bzw. zwei einen hohen, cf. Fragebogen Teil B), Nr. 6.).

¹⁴ Eine Nutzung von Normdaten für die Überlieferungsbildung scheint bisher nur im Landesarchiv Baden-Württemberg praktisch erprobt worden zu sein, cf. ZIWES 2014; DERS. 2011; entsprechend wurde in der Umfrage der Wert von Normdaten für diese Fachaufgabe nur mittelmäßig eingestuft.

¹⁵ Dies war auch der am häufigsten und zudem höchsten bewertete Nachteil in der Umfrage (Fragebogen Teil B), Nr. 7.): acht von neun sahen hier Nachteile, vier bewerteten diesen als hoch.

¹⁶ HOEN / GESSMANN 2008, S. 70 (ohne empirische Nachweise); MÜLLER 2007, S. 61 (unter Verweis auf eine Studie aus dem Bibliotheksbereich). Allerdings räumt Müller zumindest Orts- und Personenindizes in der archiv- und bestandsübergreifenden Recherche einen Mehrwert ein, doch würden dafür Generalindices benötigt, „deren Pflege allenfalls bei Ortsnamen praktikabel erscheint, bei Personennamen aber rasch an Grenzen stößt. Hier erscheint es ausreichend [...], zusätzlich zur Volltextsuche automatisch generierte Stichwortlisten anzubieten, auf eine Individualisierung und Identifizierung der einzelnen Einträge aber zu verzichten“ (ebd. S. 62). Diese Einschätzung ist mittlerweile durch die Entwicklung im Bereich der Bi-

ten-Nutzen-Analyse fehlt auch im Bibliotheksbereich weitgehend: Wie viele Nutzerinteressen durch welchen Ressourceneinsatz beim Einsatz von Normdaten beziehungsweise überhaupt der Tiefenerschließung in den (Online-)Findmitteln ohne eine zusätzliche Anfrage an die verwahrenden Institutionen befriedigt werden können, lässt sich nicht sicher beantworten.¹⁷

Ein Nebenaspekt der Ressourcenthematik ist zudem der zumindest für die Einführungsphase hohe Schulungsbedarf,¹⁸ da der Einsatz von Normdaten im Archivwesen bisher wenig verbreitet war und ist.¹⁹

Bisher nicht diskutiert wurde das prinzipielle Problem der „Filterblase“: Informationsressourcen mit Normdatenbezug, die aber nicht entsprechend indiziert sind, werden bei einer gezielten Suche über Indizes potenziell leicht übersehen.

2.3. Notwendigkeit der Relevanzprüfung durch das LAV NRW

Die Kostenproblematik wird als Faktor dauerhaft erhalten bleiben. Die eingehende Befassung mit der Thematik „Normdaten“ seitens der Archive und so auch des LAV NRW ist dennoch aus mindestens vier Gründen angezeigt:

- 1.) Die Fachaufgabe der Bereitstellung von Archivgut für die Nutzung über Erschließungsinformationen, speziell durch entsprechende Informationssysteme und Onlineportale, ist in den letzten Jahren ein wieder zunehmend diskutiertes Thema; Grundtenor dabei ist, dass jenseits etablierter Standards wie Tektonik und Volltextrecherche Archive die Suchgewohnheiten der Nutzer stärker überprüfen und ihre Angebote entsprechend anpassen müssen.²⁰
- 2.) Die Zahl der Kulturinstitutionen und -angebote, die Normdaten zur Verknüpfung mit den eigenen Ressourcen nutzen, wächst seit Jahren.²¹

bibliotheken und des Semantic Web teilweise überholt (cf. im Folgenden). Auf die langfristig negativen Folgen qualitativ minderwertiger Erschließungsdaten gerade im Onlinebereich verweist MÜLLER 2015, S. 8.

¹⁷ Laut Besprechung mit Armin Kühn, BSZ-BW, erhebt auch das BSZ keine spezifischen Daten zur Nutzung von Normdaten. Ausgewertet werden allerdings die Zugriffe auf die Online GND (OGND): diese rangierte mit 10 Mio. Zugriffen im Jahr 2016 direkt hinter dem Verbund-OPAC.

¹⁸ Dies benannte etwa die Hälfte der externen Antwortenden als Problem mittlerer Wichtigkeit.

¹⁹ Trotz intensiver Recherche konnte der Verfasser nur die im Quellen- und Literaturverzeichnis unter den Einsendern der Fragebögen genannten ausmachen; nicht mehr angefragt wurde allerdings das Archiv für Christlich-Soziale Politik. Explizit ausgeklammert blieben zudem weitere am KOOP-LITERA-Verbund beteiligte Archive (cf. die Liste unter onb.ac.at/koop-litera/explore/explore-deutschland.html) [letzter Zugriff 16.2.2017]. Da es sich hier um die spezifische Gattung der Nachlässe handelt, hätte dies keinen Mehrwert zu den schon erzielten Umfrageergebnissen erbracht.

²⁰ Cf. summarisch statt vieler die Beiträge in: Neue Wege ins Archiv 2016, und hier bes. den Erfahrungsbericht zur Nutzung der Onlinerecherche von FRICKE / SCHLUDI 2016, S. 150–153; weiter WAHRMANN 2014, bes. S. 25–30; die Frage nach dem Rechercheverhalten der Nutzer erwähnt auch schon MÜLLER 2007, S. 44. Quantifizierende Nachweise lieferte zuletzt die Nutzerstudie von REINHARDT / MUNDT 2016 für das Hessische Landesarchiv, in welcher sich gut 20 % der Nutzer unzufrieden mit der Aussagekraft der Erschließungsdaten und gut 50 % zufrieden oder sehr zufrieden zeigten. Die Bibliotheken haben ähnliche Problemlagen, cf. MÜLLER 2015 mit weiteren Hinweisen.

²¹ Es gibt diesbezüglich keine zentrale und umfassende Auswertung – dies ist angesichts der dem Semantic Web inhärenten Charakteristik verteilter Ressourcen auch gar nicht möglich. Hinweise zu Personennormdaten in der GND bieten aber etwa de.wikipedia.org/wiki/Wikipedia:BEACON, beacon.findbuch.de/seealso/pnd-aks und old.datahub.io/dataset/pnd-beacon [alle Seiten: letzter Zugriff 16.2.2017]; die Listen sind teilweise nicht aktuell, die Zahl der Anbieter muss noch erheblich höher geschätzt werden. Als Beispiele für die Nutzung von EAC-CPF cf. etwa das Archivportal Europa (archives-portaleurope.net/) [letzter Zugriff 16.2.2017]), zu weiteren Projekten cf. EITO-BRUN 2016, S. 152 f. Dass solche Überlegungen auch in anderen Kultursparten angestellt werden, belegt der Beitrag von LORDICK 2016.

- 3.) Normdaten sind Ihrem Ursprung nach bibliothekarisch, doch ist etwa die DNB grundsätzlich an einer spartenübergreifenden Kooperation interessiert. Es ist strategisch unklug, nicht mindestens Aussagen über die Anwendbarkeit im eigenen Archiv treffen zu können.²²

Dennoch ist die Verankerung des Themas in der archivischen Fachwelt im Vergleich etwa mit den Bibliotheken, aber auch anderen Kulturinstitutionen noch schwach: Meist sind nur kleinere Archive und daher Einzelpersonen involviert,²³ spezifische Projektgruppen gibt es kaum.²⁴ Von den staatlichen Archivverwaltungen setzen offenbar nur das Bundesarchiv und das Landesarchiv Baden-Württemberg Normdaten standardmäßig ein, wobei allerdings jeweils Sonderbedingungen zu vermerken sind.²⁵ Das Hessische Landesarchiv befindet sich derzeit (Stand März 2017) noch in der Planungsphase.

Die Notwendigkeit der Befassung kann für das LAV NRW als gegeben betrachtet werden. Da noch keine Vorerfahrungen bestehen, ist es zwingend, für einen Einstieg, wie er hier konzipiert wird, folgende Fragen zu beantworten:

- 1.) Welcher Normdatenstandard wird ausgewählt?
- 2.) Welche Normdatentypen und welche Bestände/Findmittel eignen sich für eine empirische Pilotstudie zur Überprüfung des eventuellen Mehrwertes?
- 3.) Welche weiteren Voraussetzungen müssen geschaffen werden?

Da parallele Erfahrungen staatlicher Archivverwaltungen entweder nicht vorhanden oder wegen der genannten Sonderbedingungen nur eingeschränkt als Muster heranzuziehen sind, wird eine diesbezügliche Überprüfung die Besonderheiten des LAV NRW weitgehend vorrangig zu berücksichtigen haben.

3. Die GND als für Archive geeignetstes Normdatenrepositorium

Von den oben genannten unterschiedlichen Normdatenstandards kommt für eine Pilotstudie im LAV NRW momentan allein die GND infrage, weil

- 1.) diese der mittlerweile dominierende Normdatenstandard im deutschsprachigen Raum ist;
- 2.) die für das Archivwesen einzig noch denkbare Alternative EAC-CPF aus prinzipiellen wie aus praktischen Erwägungen zumindest zurzeit ungeeignet ist.

Zu 1.) Die GND hat sich seit ihrer Einführung 2012 zum beherrschenden Repositorium im deutschsprachigen Raum entwickelt. Dies gilt nicht nur im Hinblick auf die verschiedenen Typen von Normdaten,²⁶ die bei der GND quasi aus einer Hand angeboten werden, sondern auch für die Anzahl der Datensätze: Gab es zum Beispiel im April 2012 ca. sieben Millionen Personendatensätze (davon gut 2,6 Millionen individualisierte sogenannte Tp-

²² Kurz: wer in diesem Feld nicht ‚sprechfähig‘ ist, wird auch in anderen nicht gehört, wofür *pars pro toto* die Webarchivierung zu nennen wäre. Dass es solche Kooperationswünsche gibt, bestätigen neben den vom Verfasser geführten Besprechungen auch das Projekt IN2N, cf. dazu WIECHMANN 2015, die allerdings auch auf die weitgehend fehlende Befassung seitens der Archive mit Normdaten (in diesem Fall EAC/XML) hinweist, cf. ebd. S. 153, sowie weiter ECKES 2014; cf. auch EITBRUN 2016, S. 152 f.

²³ So das mehrheitliche Ergebnis der Umfrage, cf. Fragebogen Teil A), Nr. 3.).

²⁴ Mit der offenbar einzigen Ausnahme des GStA PK, das freilich durch die enge Anbindung an Normdaten nutzende Institutionen wie die Staatsbibliothek Berlin auch einen Sonderfall darstellt.

²⁵ So nutzt das Bundesarchiv momentan [Stand Januar 2017] Normdaten offenbar nur für die Filmerschließung und hier für Personen und Sachschlagworte aus der GND. Das Landesarchiv Baden-Württemberg nutzt die Normdaten aus der GND bereits im Regelbetrieb in der Erschließung – in diesem Fall gibt es einerseits ohnehin schon seit langem einen Thesaurus mit Orts- und Personendescriptoren, andererseits eine eigene Ortsnormdatenbank.

²⁶ Personen, Orte, Körperschaften, Sachschlagworte et cetera.

Sätze), so waren es 2015 bereits gut 8,7 Millionen.²⁷ Für die übrigen Datentypen sind ähnliche Zuwachsraten zu verbuchen.²⁸

Es spricht nicht gegen die Verwendung der GND, dass für spezifische Normdatentypen und spezielle Bestände andere Normdatenrepositorien umfangreicher sind und daher bei einem Abgleich mehr Treffer zu erzielen vermögen.²⁹ Denn der große Vorteil der GND ist zusätzlich, dass diese über BEACON-Dateien die sparten- und anbieterübergreifende Vernetzung von Informationsressourcen ermöglicht.³⁰

Die GND hat für Archive zweifellos den schweren Nachteil einer starken ‚Publikationslastigkeit‘: Erfasst sind vor allem Orte und Personen, von denen, an denen oder über die publiziert wurde. ‚Prominente‘ oder ‚bedeutende‘ Orte in Archivgut müssen in der GND keineswegs erfasst sein.³¹ Problematisch ist außerdem der Umgang mit dem Typus der ‚Körperschaften‘ (Tb-Sätze): So führt laut den Regeln der RDA eine Namensänderung einer Körperschaft zu einem neuen Datensatz.³² Zwar können auch Geografika (Tg-Sätze) Gebietskörperschaften umfassen, doch werden Gebietskörperschaften, die als ‚Urheber‘ agieren, nach RDA wiederum als ‚Körperschaften‘ aufgefasst. Auch hier führt zudem eine Namensänderung meist zur Erstellung eines neuen Datensatzes.³³ Gerade für die in archivischen Beständen üblichen Namenswechsel von Körperschaften im historischen Verlauf erschwert dies eine eindeutige Zuordnung. Denn es ist quasi vom Verständnis von ‚Werk‘ und ‚Urheberschaft‘ abhängig, ob ein Geografikum als ‚Ort‘ oder als eine ‚Körperschaft‘ einzustufen ist.³⁴ Diese Mehrdeutigkeit dürfte insbesondere bei einem gewünschten automatisierten Abgleichsverfahren zu massiven Problemen führen – erst recht, wenn Archive aktiv eigene Neuerfassungen in die GND einbringen wollen. Davon abgesehen ist unsicher, welchen Mehrwert spezielle Körperschaftsnormdaten für die archivische Erschließung und Benutzung haben: Wegen des Sprengelbezugs der Archive werden zum Beispiel die meisten Behörden nur in einem bestimmten Archiv auftauchen, sodass höchstens ein bestands-, aber kein archiv- oder gar spartenübergreifender Nutzen gegeben ist.

Ein weiteres Problem sind die unterschiedlichen archivischen und bibliothekarischen Verzeichnungsstrategien. So ist trotz der umfangreichen Arten von Relationen, die in GND-Sätzen vorgesehen sind,³⁵ eine direkte Verknüpfung daraus mit Archivgut, wenn es als Entität in der GND erfasst werden soll, kaum möglich. Denn

²⁷ [DNB] Jahresbericht 2015, S. 49 (aus dem Jahresbericht wird leider nicht erkennbar, wie viele der Personendatensätze individualisiert sind, doch die Tendenz ist deutlich): [DNB] 2012.

²⁸ Bei den Geografika gab es im Vergleich von 2012 bis 2015 nur einen geringen Zuwachs, was allerdings hauptsächlich daran lag, dass sich zwischen 2014 und 2015 der Datenbestand der Tg-Sätze um gut 20.000 Einheiten reduzierte; woran dies lag, ist unklar, tatsächlich gab es auch 2015 gut 10.000 Neuaufnahmen an Tg-Sätzen, cf. [DNB] Jahresbericht 2015, S. 49.

²⁹ Cf. als Beispiel die Diskussion um Ortsnormdaten bei GeoNames und in der GND in GRAF 2013B.

³⁰ Zu dieser Funktion cf. <de.wikipedia.org/wiki/Wikipedia:BEACON> [letzter Zugriff 8.3.2017]; zum Format <[gbv.github.io/beaconspec/beacon.html](https://github.io/beaconspec/beacon.html)> [letzter Zugriff 23.3.2017]; ZIWES 2014.

³¹ So die Hinweise mehrerer Antworten im Fragebogen sowie in verschiedenen Besprechungen.

³² Cf. [DNB] EH-K 21 sowie [BSZ-BW] GND-Schulung 1, S. 177.

³³ [DNB] EH-G 06 sowie [BSZ-BW] GND-Schulung 2, S. 175 f.

³⁴ Zu dieser Problematik cf. auch BECKER 2013. Eine ähnliche Problematik ergibt sich beim Standard ISAAR(CPF), der Geografika nur unter den Körperschaften (als administrative Einheiten) verbucht, ansonsten aber auf Ortsnormdaten verzichtet, cf. weiter unten.

³⁵ Dies ist das Feld 500 (Beziehungen), cf. dazu und zum Folgenden insgesamt [BSZ-BW] Katalogisierungshandbuch 2016, S. 22 f., 30 f., 45–47; [DNB] GND-Codes im Unterfeld \$4 2016; exemplarisch für die Codes in den Beziehungsfeldern von Personennormdaten [DNB] EH-P 17.

hierfür kämen nur die Relationen zu „Werken“³⁶ in Frage, die schon ihrer Definition und aufgrund der fehlenden Möglichkeit, hier Hierarchien einer Tektonik abzubilden, ungeeignet sind.³⁷

Diese Problematik ist jedoch im Falle der GND mittels BEACON-Dateien zu umgehen: Die GND wird hier als Identifikator in der Indizierung und als Hinweisgeber dafür, dass zu einem Normdatensatz Informationsressourcen vorhanden sind, genutzt. Verlinkt werden dann aber nicht GND-Entitäten, sondern das jeweilige Onlineangebot des Anbieters.³⁸

Zu 2.): Der dem Normdatenformat EAC-CPF zugrundeliegende Standard ISAAR-CPF ist stark auf Provenienzbildner hin orientiert und daher nicht ohne weiteres mit der GND kompatibel.³⁹ Aus unter 1.) genannten Gründen verschärft sich dieses Problem, sobald von Normdaten für Körperschaften gehandelt wird. Bei einer Vielzahl von Beständen wäre daher eine Anreicherung der Erschließung mit ISAAR(CPF)-konformen Normdaten kaum möglich, eben weil es sich bei den betreffenden Entitäten nicht um Provenienzbildner handelt. Zudem sieht ISAAR(CPF) offenbar geografische Entitäten beziehungsweise Normdaten als solche nicht vor,⁴⁰ sondern nur als geografische administrative Einheiten unter den „Corporate Bodies“.⁴¹

Es gibt nun durchaus Projekte, die EAC-CPF und GND verbinden, wie etwa das Nachlassportal Kalliope.⁴² Im Nachlassbereich ist jedoch die Fokussierung auf Provenienzbildner weitgehend unproblematisch, da damit ohnehin die Mehrheit der Personen-Entitäten abgebildet wird.⁴³ Alternativ können – wie bei EAD – nach Institu-

³⁶ Feld 530 in den GND-Sätzen.

³⁷ Cf. zu dieser Problematik BECKER 2013; IMHOF 2008.

³⁸ Also eine Verzeichnungseinheit, ein Findbuch, Digitalisat, Katalogisat et cetera.

³⁹ Cf. [ICA] ISAAR(CPF) 2007, S. 12 (Nr. 1.9 und 1.10). Cf. auch hinsichtlich der Verwendung auch die Zusammenfassung des Bundesarchivs: bundesarchiv.de/archivgut_online/standards_werkzeuge/02254/index.html [letzter Zugriff 8.3. 2017]. Zum Standard insgesamt cf. exemplarisch SPERL 2016, S. 46 f.; OTTOSSON 2007, S. 162–164. Wenn Ottosson allerdings behauptet, „The EAC is as ISAAR(CPF) designed to comply with library standards for authority data in order to facilitate the exchange of such data between archives, libraries and museums“ (ebd. S. 165), so ist das im Ergebnis teilweise fragwürdig, cf. im Folgenden.

⁴⁰ Diese werden nur als Teil der Beschreibung des Provenienzbildners erwähnt, cf. [ICA] ISAAR(CPF) 2007, S. 27 ff. (Nr. 5.2.3). Eine Referenzierung auf Entitäten ist dabei offenbar nicht intendiert. Cf. auch die entsprechende Darstellung im EAC-CPF-XML-Schema auf der Seite der Staatsbibliothek Berlin: eac.staatsbibliothek-berlin.de/Diagram/cpf.html [letzter Zugriff 8.3.2017].

⁴¹ So wurde dies im schwedischen Nationalarchiv praktiziert (OTTOSSON 2007, S. 172 f.), eine Suche nach ortsbezogenen Entitäten scheint im Onlineinformationssystem allerdings nicht möglich zu sein.

⁴² Selbst hier ist aber die GND die ‚führende‘ Normdatei, cf. die Erläuterungen zu den Datenformaten in: kalliope.staatsbibliothek-berlin.de/de/standards/datenformate.html [letzter Zugriff 23.3.2017].

⁴³ Ähnliches belegt das Projekt „Social Networks and Archival Context“ (SNAC), das ebenfalls EAC-CPF als Grundlage verwendet (socialarchive.iath.virginia.edu/ [letzter Zugriff 8.3.2017]). Dort sind – zumindest in der vom Verfasser eingesehenen Prototyp-Version – nur Personen als Entitäten erfasst; die Relationen beinhalten bezeichnenderweise etwa Links zu Personen, Familien und Körperschaften, aber eben nicht zu Geografika. Zwar wird auch auf bibliothekarische Normdaten als externe Referenz verlinkt, doch dies betrifft offenbar nur die Normdaten für die angesetzte Person. Die referenzierten archivischen Bestände beinhalten – kaum überraschend – weitgehend nur Nachlässe/Autographen, für die im EAC-CPF-Element „ResourceRelation“ für das Attribut „ResourceRelationType“ der Datentyp bzw. Wert „creatorOf“ verwendet wird (cf. auch die EAC-Tag-Library unter eac.staatsbibliothek-berlin.de/fileadmin/user_upload/schema/cpfTagLibrary.html [letzter Zugriff 8.3. 2017]). Freilich gäbe es auch den Wert „subjectOf“, der aber dafür nicht genutzt wurde. Ohnehin ist dies angesichts der Orientierung von EAC auf Provenienzbildner ein zweifelhafter Datentyp. Diese starre Fixierung auf den Provenienzbildner führt offenbar dazu, dass teilweise Ressourcen verknüpft werden, die nicht der strengen Definition im EAC-CPF-Schema entsprechen: So werden im SNAC-Prototyp (cf. vorige Anm.) teilweise unter dem genannten Wert „creatorOf“ Briefe an die – und nicht von der – Person gelistet (exemplarisch cf. socialarchive.iath.virginia.edu/ark:/99166/w6r516c5 [letzter Zugriff 8.3.2017]), was kaum korrekt ist – doch offenkundig wurden die Alternativen „subjectOf“ und „other“ als noch schlechter eingestuft.

tion und Anforderungen spezielle EAC-CPF-Schemata entwickelt werden.⁴⁴ Dem Ziel eines spartenübergreifenden und standardisierten Zugangspunktes sind derartige Speziallösungen kaum dienlich.

Auch praktische Gesichtspunkte sprechen unter aktuellen Bedingungen gegen eine Verwendung von EAC-CPF: Jedes Normdatenrepositorium benötigt einen zentralen Anbieter, der für die Speicherung und Pflege sowie für die Vergabe eines Identifikators zuständig ist.⁴⁵ Dies gilt auch für die GND, für welche diese Aufgabe bei der DNB liegt. Für EAC-CPF muss es eine entsprechende zentrale Institution geben, entweder über ein Portal oder ein Nationalarchiv wie etwa in Schweden.⁴⁶ Abgesehen von der geringen Verbreitung von EAC im deutschsprachigen Raum⁴⁷ ist kein solcher zentraler Anbieter in Sicht.⁴⁸ Zudem ist zweifelhaft, ob überhaupt schon innerhalb des Archivwesens eine Einigung auf ein gemeinsames und einheitliches EAC-CPF-Profil erzielt werden könnte. Die Erfahrungen mit den abweichenden Profilen schon für EAD machen dies unwahrscheinlich.⁴⁹

Der Verzicht auf EAC-CPF bedeutet auch, eine Anbindung von eventuell aus der Erschließung generierten Normdaten an das ebenfalls EAC-CPF verwendende Archivportal Europa vorläufig zurückzustellen.

Es ist noch darauf zu verweisen, dass alle ICA-Standards, also ISAD(G), ISAAR(CPF), ISDIAH und ISDF derzeit überarbeitet werden und zu einem einheitlichen Gesamtstandard „Records in Context“ (RiC) verschmolzen werden sollen. Ein erster Entwurf dazu wurde im September 2016 vom ICA veröffentlicht.⁵⁰ Grund für diese umfassende Neubearbeitung sind einerseits technische Fortentwicklungen (Expansion relationaler Datenbanken, XML-Technologien, Semantic Web) und damit verbunden Veränderungen in den Möglichkeiten und Anforderungen an die archivistische Erschließung, andererseits auch die Erkenntnis, dass die vier bisherigen Standards nur sehr ungleichgewichtig Anwendung in der Fachwelt gefunden haben.⁵¹ Die Fokussierung auf Provenienzbildner als Entitäten ist, wenngleich abgeschwächt, auch im RiC-Modell weiterhin vorhanden,⁵² dafür sind nun Orte als Entitäten vorgesehen, die in der Definition zumindest Ähnlichkeiten mit den Geografika in der GND aufweisen.⁵³ Die vorgesehenen Relationen lassen zumindest andere Beziehungen zwischen Ressource (Ar-

⁴⁴ Dies wurde offenbar auch in Schweden so verwirklicht, und zwar explizit auch mit dem Anspruch, die Austauschbarkeit von Daten mit den Bibliotheken zu befördern, cf. OTTOSSON 2007, S. 175.

⁴⁵ Cf. insgesamt zur Problematik aus österreichischer Perspektive SPERL 2016, S. 51–57. Zu den notwendigen Pflichtangaben in EAC-CPF, zu denen u.a. ein eindeutiger Identifikator gehört, cf. [ICA] ISAAR(CPF) 2007, S. 18 (Nr. 4.7).

⁴⁶ OTTOSSON 2007, S. 161 f., 166–171.

⁴⁷ Keine der befragten externen Archivverwaltungen setzte EAC-CPF ein, mit der Ausnahme des Bundesarchivs in einem Pilotprojekt zu den SED- und FDGB-Beständen, cf. dazu IMHOF 2008, der freilich Stand 2008 eine schwache Verbreitung von Normdaten im Archivwesen konstatierte.

⁴⁸ Das Archivportal-D ist – zumal vor dem Hintergrund seiner Entstehung aus der Deutschen Digitalen Bibliothek – dazu weder willens noch in der Lage. Und aufgrund der föderalen Struktur hat das Bundesarchiv nicht dieselbe Stellung wie etwa das schwedische oder die angloamerikanischen Nationalarchive und kommt daher als derartiger zentraler Host ebenfalls nicht in Betracht.

⁴⁹ Cf. den Überblick auf der Seite des Bundesarchivs: <bundesarchiv.de/archivgut_online/standards_werkzeuge/02251/> [letzter Zugriff 09.03.2017]. EAD ist zudem exemplarisch für die Problematik, im Archivwesen einheitlich angewendete Datenprofile für eine maschinelle Auswertbarkeit zu schaffen.

⁵⁰ [ICA] 2016.

⁵¹ Ebd. bes. S. 7–11.

⁵² Die RiC-E4 wird sinnigerweise erneut als „agent“ bezeichnet, wenngleich die Definition („A person or group, or an entity created by a person or group, that is responsible for actions taken and their effects“) diesbezüglich weniger eindeutig ist (ebd. S. 14 f.).

⁵³ RiC-E13 Place: „Jurisdictions and any geographic or administrative point or area“ (ebd. S. 18 f.). Die Problematik der Abgrenzung von (Gebiets-)Körperschaften als Orte und Körperschaften als „Agents“ besteht auch im RiC-Modell.

chivgut) und Entitäten als diejenige eines Registraturbildners möglich erscheinen.⁵⁴ Doch trotz der extrem umfangreichen Beziehungsarten⁵⁵ bleibt abzuwarten, wohin die Entwicklung gehen wird. Entscheidend aber ist: „Records in Context“ ist bis dato nur ein Entwurf, das dazu gehörige Normdaten-Schema existiert noch nicht, und die genannten Hindernisse bei EAC-CPF würden teilweise auch für RiC fortgelten.

Aus all diesen Gründen kommt für das LAV NRW bis auf Weiteres nur die GND als Normdatenrepositorium in Betracht.⁵⁶ Die gleichen Gründe bedingen allerdings, dass beim momentanen Stand der Dinge nur Personen- und Ortsnormdaten in der GND für einen Abgleich infrage kommen, nicht dagegen die Körperschaftsnormdaten oder Sachschlagworte.⁵⁷ Die ausgewerteten Fragebögen als auch die Besprechungen sowohl intern im LAV NRW als auch extern befestigen diese Einschätzung ganz deutlich.⁵⁸

4. Auswahl geeigneter Bestände/Findmittel

4.1. Voraussetzungen

Die Wahl der Vorgehensweise für eine Pilotstudie zur Normdatenanreicherung im LAV NRW muss sich den dort herrschenden Bedingungen anpassen. Diese beinhalten im Einzelnen:

- 1.) Fehlende Thesauri: Es gibt keine archivinterne Deskriptorendatenbank,⁵⁹ ebenso wenig existiert eine eigene Ortsnormdatenbank.
- 2.) Fehlende Indizierung: Im LAV NRW wird nicht mehr indiziert, die Erschließungsrichtlinien sehen dies auch nicht vor, da Indices Teil der nicht mehr durch Richtlinien normierten Tiefenerschließung wären.⁶⁰
- 3.) Die Antworten nahezu aller im LAV NRW befragten Experten gingen dahin,⁶¹ dass eine manuelle Anreicherung – unabhängig davon, ob für bereits erschlossene oder noch unerschlossene Bestände – aus Kostengründen ausgeschlossen sei.⁶²

⁵⁴ Zum Beispiel RiC-R207 „Agent is associated with record“ oder RiC-R210 „Agent is subject of record“, cf. insgesamt die „Agent Relations“ und „Place Relations“ ebd. S. 52–61 und 84–88.

⁵⁵ Diese gehen mengenmäßig deutlich über die GND hinaus und umfassen insgesamt 792 Relationen, cf. ebd. S. 39–92. Allerdings handelt es sich faktisch um redundante Relationen, da jede Relation jeweils einmal aus der Perspektive der beiden verknüpften Entitäten betrachtet wird.

⁵⁶ In Österreich erwägt der VÖA offenbar, archivische Normdateien nach ISAAR(CPF) über ein Wiki oder ein noch aufzusetzendes Archivportal zu implementieren. Wie tragfähig ein solches Unterfangen im Hinblick auf die nötige Infrastruktur und das Ziel spartenübergreifender Vernetzung sein wird, bleibt abzuwarten.

⁵⁷ Die GND-Typen „Konferenzen“ und „Werktitel“ dürften für Archive keine praktische Relevanz haben.

⁵⁸ Fragebogen Teil B), Nr. 2.) sowie Besprechungsvermerke. Praktisch keine der befragten externen Archivverwaltungen, die die GND verwenden, arbeitete mit anderen Datentypen als Personen oder Geografika; eine Ausnahme ist hier lediglich das Bundesarchiv, das in der Filmerschließung Sachschlagworte indiziert.

⁵⁹ Allerdings gibt es eine interne ‚Prominentendatenbank‘, die zur Überlieferungsbildung eingesetzt wird, nicht jedoch zur Indizierung.

⁶⁰ Dies betrifft Erschließungsrichtlinien zu den Bereichen Justiz, Finanzen, Schule und Weiterbildung, nichtstaatliches Archivgut sowie Personalunterlagen. Es handelt sich in allen Fällen um interne, nicht publizierte Verwaltungsvorschriften. Festgelegt sind folgende Stufen: Stufe 0 = Akzession; Stufe 1 = Ersterschließung; Stufe 2 = Standarderschließung (diese kennt allerdings auch „optionale Angaben“, die schon einen Übergangsbereich zur Tiefenerschließung bilden); Stufe 3 = Tiefenerschließung (hierfür gibt es keine Richtlinien mehr!). Cf. auch nochmals HOEN / GESSMANN 2008, S. 73.

⁶¹ Cf. Fragebogen Teil B), Nr. 11.), sowie die Besprechungsvermerke mit den Mitarbeitern des LAV.

⁶² Die von einem Mitarbeiter des LAV ins Spiel gebrachte Möglichkeit von Crowdsourcing-Projekten ist interessant und tatsächlich auch der einzige Fall, in dem momentan im LAV durch Externe indiziert wird, nämlich bei den Personenstandsregistern, cf. <genealogy.net/vereine/wggf/?Westfalen_seit_1815:Westfaelische_Juden-_und_Dissidentenregister_1815-1874> [letzter Zugriff 8.3.2017]. Doch kommt eben diese Form der manuellen Indizierung nur bei spezifischen und ‚prominenten‘ Beständen mit interessierten Gruppen von Archivnutzern infrage, cf. nochmals RIEDEL 2013.

Aus diesen Vorbedingungen ergibt sich, dass die Möglichkeit des automatisierten Abgleichs mit den GND-Sätzen ein starkes Kriterium für die Auswahl der Bestände beziehungsweise Findmittel⁶³ darstellt; als einziges ist dies aber unzureichend. Die einzelnen als sinnvoll eruierten Kriterien sollen im Folgenden erläutert werden.

4.2. Priorisierungskriterien und Bewertungsmatrix

4.2.1. Methodik

Die folgenden Kriterien wurden teils aus den mittels Recherche, Fragebögen und Besprechungen gewonnenen Erkenntnissen geschöpft, teils basieren sie auf verwandten Kriterienkatalogen des LAV NRW sowie aus der Literatur zu den Themen Digitalisierung und Bestandserhaltung. Die Priorisierungskriterien sind außerdem tendenziell an den in Fragebögen und Besprechungen dem Verfasser von den Mitarbeitern der Fachdezernate des LAV benannten geeigneten Bestände- beziehungsweise Archivalientypen ausgerichtet.⁶⁴ Diese „Eignung“ ist die nötige Vorauswahl an Beständen, die überhaupt erwartbar nennenswerte Mengen an Personen- und Ortsnormdaten enthalten.

Aus den Kriterien wurde eine Bewertungsmatrix erstellt (cf. Kap. 4.2.8). Bei deren Erstellung und der Gewichtung der einzelnen Kriterien wurde ebenfalls auf methodisch ähnliche Vorlagen für die Priorisierung in der Bestandserhaltung, insbesondere auf eine kürzlich vom Hessischen Landesarchiv publizierte Modellstudie zu diesem Thema zurückgegriffen.⁶⁵ Wie in den Vorlagen gibt es für die im Folgenden ausführlicher zu erläutern- den Priorisierungskriterien die vier Eignungskategorien⁶⁶ „hoch“, „mittel“, „niedrig“ und „kein“, denen jeweils Punkte – in der Abfolge 3, 2, 1, 0 Punkt(e) – zugeordnet wurden. Hauptschwierigkeit ist natürlich, welche Schwellenwerte für die jeweilige Einstufung angesetzt werden – diese Entscheidung ist letztlich immer in gewissem Maß subjektiv. Im Unterschied zu den Vorlagen wird die Gewichtung der einzelnen Priorisierungskriterien nicht über prozentuale Anteile an der Gesamtmenge hergestellt,⁶⁷ sondern über eine zusätzliche dreistufige Faktorbildung (dreifacher, zweifacher, einfacher Faktor). Hauptgrund ist, dass wegen der heterogenen Kriterien eine in den Verhältnissen zueinander angemessene prozentuale Gewichtung sehr schwierig schien.

Es wurden keine Ausschlusskriterien definiert, auch nicht für die Kriterien „Strukturierte Erschließungsdaten“ oder „Archivrechtlicher Status“. Dies hätte bestimmte Kriterien über die Gewichtung hinaus unzulässig überhöht.

Ebenso spielte die Frage regressives (Retrokversion) oder progressives (noch unerschlossene Bestände) Vorgehen keine Rolle als Kriterium. Denn ein derartiges Kriterium wäre mit der Frage des archivrechtlichen Sta-

⁶³ Auf die dringende Notwendigkeit solcher Verfahren angesichts der Datenmenge weist zu Recht ZIWES 2014 hin, der auch ein praktisches Beispiel in Form des Abgleichs von Daten mit der wikipedia-Personensuche (cf. tools.wmflabs.org/person-data/) [letzter Zugriff 15.3.2017]) und der Landesbibliographie Baden-Württemberg bietet – die GND wird dabei allerdings nur indirekt abgeprüft, während es in vorliegender Studie um den direkten Abgleich geht.

⁶⁴ Cf. Fragebogen Teil B), Nr. 1.) und 4.). Im Einzelnen wurden genannt (nicht alles davon findet sich im LAV NRW): Personenstandsregister, Personalakten, Berufungsakten, Promotionsakten, Prüfungsakten, Prozessunterlagen, Einbürgerungsunterlagen, NS-Unterlagen (Gestapo, Entnazifizierung, Wiedergutmachung, Rückerstattung/Entschädigung), Nachlässe, Kriminalakten, Testamente, Todeserklärungen, Adoptionen, Gefangenenpersonalakten, Urkunden, Topographische Daten, Karten, Pläne, Grundbücher/Grundakten, Bildbestände (insbesondere Fotos und Luftbildbestände/Schrägluftaufnahmen), Aufschwörungstafeln.

⁶⁵ [Hessisches Landesarchiv] 2016, hier bes. S. 12–15; Matrix Beständepriorisierung [Bestandserhaltung] LAV NRW [ca. 2011, unveröffentlicht]; WIECH 2013; HAFFER 2013; Martina WIECH, Auswahlmerkmale und Priorisierung archivalischer Quellen für die Digitalisierung. Fachliche und organisatorische Aspekte [Präsentation, ca. 2011, unveröffentlicht].

⁶⁶ Die Schadensmatrix des LAV NRW kennt allerdings insgesamt fünf.

⁶⁷ So das Vorgehen von [Hessisches Landesarchiv] 2016, S. 14, wo allerdings die Punktwerte mit den vollen Zahlen der als Gewichtung festgelegten Prozentanteile multipliziert werden. Die Gewichtung seitens des Hessischen Landesarchivs war laut Besprechung mit Johannes Kistenich-Zerfaß eine rein facharchivarische Entscheidung ohne direkte Vorlagen aus der Literatur.

tus in Kollision geraten und wegen der präferierten Automatisierung ohnehin derartig problematisch gewesen, dass keine aussagekräftige Auswertung mehr möglich gewesen wäre.

Die folgenden Kriterien beanspruchen keineswegs Vollständigkeit oder eine Form höherer Objektivität – letztere wird nur insofern angestrebt, als dass die Auswahl der Kriterien, die Grade der Eignung und ihre Gewichtung reflektiert und zu begründen versucht werden.

4.2.2. Kriterium 1: Strukturierte Erschließungsdaten

Unter diesem Kriterium werden Erschließungsdaten verstanden, die durch entsprechende Felder oder Auszeichnung der Daten im AIS maschinenlesbar sind und damit automatisiert weiterverarbeitet werden können. Dies gilt natürlich auch für einfache beziehungsweise Feldtypen der Formalerschließung wie Signatur, Laufzeit et cetera. Gemeint sind im vorliegenden Fall jedoch Daten der Sacherschließung, die in strukturierter Form wiedergegeben werden.

Die Bedeutung strukturierter Erschließungsdaten ergibt sich aus dem Format der GND, einem strukturierten XML-Format.⁶⁸ Für die Identifikation von Entitäten aus einem Erschließungs-Datensatz existieren momentan im Wesentlichen drei Verfahren:⁶⁹

- 1.) Einfaches „Match&Merge“: Dieses besteht im Prinzip aus dem Abgleich externer Erschließungsdaten mit den strukturierten Datenfeldern der GND – das bei der DNB übliche Verfahren.⁷⁰
- 2.) Named-Entity-Recognition (NER): Dieses Verfahren ist bisher hauptsächlich bei Volltexten üblich.
- 3.) Phonetische Verfahren: Diese Verfahren sind besonders für Bestände/Zeiträume mit großer Variationsbreite in der Namensschreibweise interessant.

Nr. 1.) ist bereits Standard und vom Hessischen Landesarchiv und der DNB exemplarisch an Beständen überprüft worden. Dabei hat sich gezeigt, dass ein Abgleich nur bei gut strukturierten Daten funktioniert.⁷¹ Die Verfahren 2.) und 3.) sind, wiewohl im Informatikbereich bereits gängig, für Archive wohl im Moment noch Zukunftsmusik. Zudem beinhalten sie hinsichtlich der Normdaten-Thematik im Grunde zwei Schritte: Im ersten wird überhaupt das Vorliegen einer Entität (bei der NER) oder möglicher Namensvarianten zu einer Entität (bei den phonetischen Verfahren) festgestellt, im zweiten dann ein Abgleich mit der GND durchgeführt – in beiden Fällen benötigt es zudem weitere Merkmale für die Disambiguierung.⁷² Dennoch sollten diese Möglichkeiten der Indizierung beziehungsweise erweiterten Recherchefunktion seitens der Archive im Auge behalten werden.

Strukturierte Daten sind jedoch nicht nur wegen der ebenfalls strukturierten Erfassung in der GND notwendig, um die jeweiligen Felder sofort automatisch zuordnen zu können und nicht erst einmal die relevanten Daten extrahieren zu müssen. Ebenso wichtig ist, dass durch strukturierte Daten ein gezielter Abgleich verschiedener Merkmale möglich wird, was die Trefferquote stark verbessert: beispielsweise bei Personen neben dem vollen Namen exakte Lebensdaten, Beruf, Geburts- oder Sterbeort et cetera. Entsprechende Pflichtangaben schreibt

⁶⁸ Innerhalb der Bibliotheksverbände wird MARC21/XML verwendet, als LOD auch RDF/XML, cf. zu den Formaten die Hinweise und weiterführenden Links in dnb.de/DE/Service/DigitaleDienste/SRU/sru_node.html und https://dnb.de/DE/Service/DigitaleDienste/LinkedData/linkeddata_node.html [für beide letzter Zugriff 8.3.2017].

⁶⁹ Zum Folgenden wird insbesondere auf die Fragebögen von und Besprechungen mit Kristina Starkloff, Daniel Fähle und Stefan Aumann rekurriert, sowie weiterhin WETTLAUER / WESTPHAL 2014, S. 274 ff.; SNAE 2007. Auf die technischen Details kann hier nicht weiter eingegangen werden.

⁷⁰ Es handelt sich um ein Verfahren mit drei Ergebnisarten: Treffer, kein Treffer, Zweifelsfälle – die letzte Gruppe ist die problematischste, weil sie manuell nachbearbeitet werden muss.

⁷¹ Die Trefferquote betrug maximal ca. 15 %, wobei hauptsächlich Personendaten abgeglichen wurden.

⁷² Kaum zufällig werden die Verfahren auch kombiniert, cf. exemplarisch rosette.com/function/name-matching/ [letzter Zugriff 8.3.2017]. Solche kommerziellen Lösungen dürften für Archive freilich mittelfristig utopisch bleiben.

die Individualisierungsrichtlinie der DNB auch vor;⁷³ bei Geografika sind diese zumindest bei Namensgleichheit vorgesehen.⁷⁴

Das AIS V.E.R.A. des LAV NRW⁷⁵ hält neben der einfachen Sachakten-Erschließungsmaske auch mehrere Erschließungsmasken mit derartig strukturierten Eintragungsmöglichkeiten in der Sacherschließung vor.⁷⁶ Von diesen kommen wegen der Einschränkung auf Personen und Geografika in der GND sowie den Bedingungen des Match&Merge-Verfahrens der DNB bei einem automatisierten Abgleich nur folgende Erschließungsmasken mit abgleichsrelevanten Feldtypen in Frage:

- Urkunden: „Ort“
- Karten/Pläne: „Ort“, „Region/Land“, „TK-NR [Nummer der Topographischen Karte]“, „Längengrad“, „Breitengrad“
- Plakate: „Druckort“
- Bild/Foto: „Fotograf“
- Film: „Name“, „Anschrift“⁷⁷
- Luftbilder: A) Schrägluftaufnahmen: „Ort“, „OKZ [Ortskennzahl]“, „Land“ B) Reihenmessbilder: „Land“, „Hersteller“ C) Luftbildpläne: „Land“, „Gauß-Krüger [Hochwert, Rechtswert]“
- Grundbücher/Grundakten: „Gemeinde/Gemarkung“
- Personenbezogene Akten: „Name“, „Vorname“, „Geburtsdatum“, „Todesdatum“, „Geburtsort“, „Geschlecht“, „Amt“/Beruf“
- Prozessverfahrensakten: „Beteiligte“ (untergliedert in „Name“, „Vorname“) und dazu gehörig „Geburtsdatum“, „Todesdatum“
- Personenstandsüberlieferung: A) Personenstandsregister – hier Hinweismittelungen: „Vorname“, „Familiennamen“, „Geburtsname“, „Geburtsort“ B) Namensverzeichnis: „Vorname“, „Familiennamen“, „Geburtsname“, „Datum des Ereignisses“.

Insgesamt liegen also vor allem für Personen mehrere strukturiert erfassbare Merkmale in der Erschließung vor, weniger dagegen für Geografika. Welche Ergebnisse dies beim Match&Merge-Verfahren erbringt, lässt sich ohne empirische Prüfung nicht abschätzen.

Zusammengefasst: Aufgrund des Wunsches/der Notwendigkeit einer weitgehenden Automatisierung des Abgleichs sowie den technischen Bedingungen des bei Abfassung der Arbeit üblichen Abgleichsverfahrens erhält das Kriterium „Strukturierte Erschließungsdaten“ den höchsten Faktor in der Gewichtung (dreifacher Faktor).

Angemerkt sei noch, dass das Vorliegen gut strukturierter Erschließungsdaten neben dem technischen auch einen strategischen Aspekt hat: Die DNB ist zwar sehr an Kooperationen, aber ebenso an einer guten Datenqualität interessiert.⁷⁸ Und diese Qualität bemisst sich, zumindest aus Sicht der DNB, vor allem in einer klar gegliederten und einheitlichen Strukturierung. Auch wenn bei den momentanen Bedingungen im LAV NRW –

⁷³ [DNB] EH-P 16 schreibt je nach Stufe des Tp-Satzes mindestens drei Individualisierungsmerkmale aus unterschiedlichen Merkmalsgruppen vor; der Personennamen zählt nicht zu den Merkmalen, sondern zu den automatischen Pflichtangaben.

⁷⁴ [DNB] EH-G 2. Hauptsächlich geschieht dies über weitere geografische Angaben (Bundesland etc.).

⁷⁵ Zur grundsätzlichen Struktur der Masken und Feldtypen cf. [startext] 2007, S. 14–18.

⁷⁶ Zum Folgenden cf. insgesamt [startext] 2007, S. 66–102.

⁷⁷ Diese Angaben werden jeweils für die übergeordnete Ebene „Hersteller“, „künstlerische Gestaltung“, „Auftraggeber“, „Inhaber der Verwertungsrechte“, „Überlieferung in anderen Archiven“ erfasst.

⁷⁸ Besprechung mit Peter Sandner (HHStAW).

und wohl bei den meisten Archiven – nur der passive Abgleich auf der Tagesordnung steht: Wollen Archive zukünftig die GND mit archivrelevanten Daten befüllen,⁷⁹ so gewinnt die Thematik Datenstruktur massiv an Gewicht.

4.2.3. Kriterium 2: Archivrechtlicher Status

Bei der Verknüpfung zwischen Normdaten und Archivgut und/oder Erschließungsdaten handelt es sich grundsätzlich um eine Veröffentlichung im Sinne von § 8 ArchivG NRW. Da der Nutzerkreis potenziell unbegrenzt und zugleich anonym ist, können keine Auflagen in der Nutzung erteilt werden: Was immer im Internet verknüpft wird, muss archivrechtlich ohne Einschränkung frei publiziert werden können.⁸⁰

Dieses Problem betrifft natürlich ausschließlich Personennormdaten, die freilich den ganz überwiegenden Teil der GND ausmachen.⁸¹ Gerade hier stellt zudem § 7 Abs. 2 ArchivG NRW hinsichtlich der Verknüpfung personenbezogener Daten strenge Voraussetzungen, um die schutzwürdigen Belange Betroffener zu wahren. Zu beachten sind weiter die üblichen personenbezogenen Schutzfristen sowie zusätzlich die auf schutzwürdige Belange bezogenen Nutzungsversagungsgründe.⁸²

Generell sind Bestände ab dem frühen 20. Jahrhundert (Stichjahr 1917⁸³) problematisch. Zugleich aber ist Verknüpfung und Verkettung über Bestände und Institutionen hinweg das Wesen der GND und des Semantic Web. Gleichwohl sind Fälle denkbar, in denen zwar nicht eine einzelne Verzeichnungseinheit, dafür aber summarisch das Findmittel – in diesem Zusammenhang verstanden als mindestens Eintrag des / der betreffenden Bestandes / Bestände in einer Tektonik – verknüpft werden kann. Daher wird in der Bewertungsmatrix auch der Faktor 2 bei Verknüpfbarkeit des Findmittels und der Faktor 3 bei Verknüpfbarkeit der VZE angesetzt.⁸⁴

Bestände beziehungsweise Findmittel, die den verlängerten Schutzfristen des BArchG aufgrund bundesrechtlicher Geheimhaltungsvorschriften unterliegen,⁸⁵ sind weitgehend ungeeignet. Denn hier wäre schon die Verknüpfung des Findmittels ein Verstoß gegen das zu schützende Geheimnis. Die entsprechenden Findbücher sind denn auch vor Ablauf der Schutzfristen nicht frei im Netz verfügbar. Allerdings kann es natürlich auch in solchen Findmitteln VZE geben, bei denen alle Schutzfristen abgelaufen sind. Doch müsste selbst dann geprüft werden, ob das zu schützende Geheimnis nicht weiterhin einen Nutzungsversagungsgrund darstellt.⁸⁶ Für einen effizienten, insbesondere automatisierten Ablauf der Verknüpfung ist dies offenkundig untauglich, und insgesamt gilt dies wohl für die meisten derartigen Bestände nach 1949. Auch wenn die verlängerten Schutzfristen

⁷⁹ Dies wird, wenn die Archive sich bei der DNB Gehör verschaffen wollen, langfristig unumgänglich sein, so die Einschätzung mehrerer Gesprächspartner: bloßes passives Konsumieren reicht nicht aus.

⁸⁰ Dieser Problematik, die u.a. auch den Umgang mit dem Thema „virtueller Lesesaal“ betrifft, widmen sich mehrere Beiträge in: Neue Wege ins Archiv 2016. Weiter wird auf Gespräche mit rechtskundigen Mitarbeitern des LAV NRW rekurriert. Fragen des UrhG und KunstUrhG werden an dieser Stelle nicht behandelt, da diese nach Auskunft auch der externen Gesprächspartner für die Anwendung von Normdaten praktisch weitgehend irrelevant sind. Ebenso kann die im April 2016 bekannt gemachte und ab Mai 2018 gültige EU-Datenschutz-Grundverordnung (Verordnung (EU) 2016/679) unbeachtet bleiben, da die dortigen Regelungen die Verarbeitung historischer personenbezogener Daten zu Archivzwecken grundsätzlich ermöglichen (cf. insbesondere ebd. den Artikel 89).

⁸¹ Cf. nochmals für 2015 [DNB] 2016, S. 49: ca. 8,67 Mio. von ca. 11,27 Mio. Datensätzen.

⁸² § 6 Abs. 2 Satz 1 Nr. 2 bis 4, § 7 Abs. 1 Satz 2 Nr. 1 bis 3 ArchivG NRW.

⁸³ Durch das am 10.03.2017 in Kraft getretene neue BArchG gleichen sich nun die personenbezogenen Schutzfristen, unabhängig davon, ob der Zugang zu Unterlagen nach ArchivG NRW oder BArchG erfolgt.

⁸⁴ Die dritte theoretische Alternative – nämlich lediglich anzuzeigen, dass im Archiv etwas zu dieser Person vorhanden ist, aber nicht was – mag zwar archivrechtlich einwandfrei, dürfte aber für Benutzung und Recherche kaum sinnvoll sein.

⁸⁵ § 11 Abs. 3 BArchG.

⁸⁶ § 13 Abs. 1 Nr. 3 BArchG sowie § 6 Abs. 3 Satz 1 Nr. 4 ArchivG NRW, doch kommt bei diesen Geheimnissen auch ebd. die Nr. 3 in Betracht, da Angehörige durch eine Verknüpfung der Daten betroffen sein können.

nach BArchG abgelaufen sind, können daher einzelne VZE aus eben genannten Gründen bei einer automatischen Verknüpfung problematisch bleiben: Was bei einer ‚normalen‘ Nutzung im Lesesaal durch Auflagen eingehengt werden kann, ist in der anonymen Nutzung im Internet nicht mehr kontrollierbar.

Etwas anders gestaltet sich die Sachlage bei den personenbezogenen Schutzfristen: Zwar wird auch hier in den meisten Fällen die Verknüpfung von Normdatei und Erschließungsdaten den oben genannten Rechtsvorschriften des ArchivG NRW beziehungsweise BArchG zuwiderlaufen. Dies gilt aber mit Sicherheit nicht für Unterlagen, die Amtsträger in Ausübung ihres Amtes betreffen.⁸⁷ In diesen Fällen kann mindestens das Findmittel, gegebenenfalls sogar die einzelne VZE verknüpft werden, sofern darin nicht Dritte (Ehepartner oder andere Personen) genannt werden, für die die personenbezogenen Schutzfristen sehr wohl noch Geltung besitzen. Für den großen Block der Prozessverfahrensunterlagen wiederum wird dagegen weitgehend nur nach Ablauf der einschlägigen Schutzfristen eine Verknüpfung des Findmittels ohne Angabe der konkreten VZE – für deren Einsicht dann ein individueller Nutzungsantrag gestellt werden müsste – möglich sein.⁸⁸ Denn in den Erschließungsdaten der einzelnen VZE können schutzwürdige Belange Betroffener oder Dritter enthalten sein. Grundsätzlich problematisch wäre allerdings auch in diesem Fall noch, ob und in welchen Fällen schon aus dem Namen des Findmittels auf einen Prozessgegenstand geschlossen werden dürfte: Insbesondere bei Prozessgegenständen, die unter die genannten besonderen Geheimhaltungsvorschriften des BArchG fallen, wird dies kaum zulässig sein.

Zusammengefasst: Soll die Problematik des Persönlichkeits- und Geheimnisschutzes komplett umgangen werden, bedeutet dies faktisch einen Ausschluss aller personenbezogenen Unterlagen aus der Zeit nach dem jeweiligen Stichjahr – wie sinnvoll eine solche Radikallösung angesichts der gerade für das 20. Jahrhundert großen Menge an Unterlagen mit potenziellem Normdatenbezug⁸⁹ wäre, sei dahingestellt. Dies auch vor dem Hintergrund, dass gerade diese gleichförmig aufgebauten „Massenakten“ noch am ehesten in strukturierter Form⁹⁰ erschlossen sind.

Die Möglichkeit einer Verknüpfung hängt zugleich von der Online-Verfügbarkeit der Findmittel und daher von der archivrechtlichen Einschätzung ab, welche Erschließungsdaten ganz oder teilweise veröffentlichungs-fähig sind.⁹¹ Wo aufgrund der Gestalt der Unterlagen individuelle Prüfungen auf schutzwürdige Belange auch nach Ablauf von Schutzfristen notwendig erscheinen, ist die arbeitsökonomisch sinnvolle automatisierte Verknüpfung kaum möglich.

Für den Fall, dass in einem Bestand/Findmittel teilweise nur das Findmittel, teilweise die VZE verknüpfbar wären, sollte aus Gründen der Vereinfachung dem jeweils mehrheitlich vorhandenen Status bei der Gewichtung der Vorrang gegeben werden.

⁸⁷ § 7 Abs. 3 Satz 2 ArchivG NRW, ebenso § 11 Abs. 4 BArchG.

⁸⁸ Der Öffentlichkeitsgrundsatz nach § 169 GVG gilt nur für die persönliche Anwesenheit im Prozess und ist gerade nicht auf die daraus entstehenden Unterlagen übertragbar (etwa im Sinne eines Analogieschlusses zum IFG). Ohnehin kennt dieser Grundsatz zahlreiche Einschränkungen, cf. ebd. ab § 170.

⁸⁹ Dies ist auch das Ergebnis der Umfrage: Die meisten Befragten sahen einen deutlichen Schwerpunkt der Normdateneignung in der Zeit 1800 bis 1945 und erst recht nach 1945 (teilweise bis zu 90 %).

⁹⁰ Cf. nochmals Kap. 4.2.2.

⁹¹ Auf die Sonderproblematik, dass das Einstellen einzelner Punkte in der Tektonik bzw. Findmittel ohne weitere VZE im Onlineinformationssystem von <www.archive.nrw.de/> [letzter Zugriff 8.3.2017] erheblich einfacher ist als im Archivportal-D, kann an dieser Stelle nicht eingegangen werden.

4.2.4. Kriterium 3: Umfang

Das Kriterium bezieht sich definitorisch auf die Anzahl der VZE und nicht etwa der laufenden oder Regalmeter: auch wenige VZE können viel Archivgut beinhalten.

Dennoch erscheint dieser Faktor zunächst widersinnig: Bestände/Findmittel können eine vierstellige Anzahl von VZE besitzen, aber wenn es sich um unstrukturierte Daten handelt, ist die Trefferquote womöglich dennoch geringer als bei deutlich kleineren Beständen/Findmitteln. Der Grund wird aber deutlich, wenn nochmals die Ergebnisse der hessischen Pilotstudie mit der DNB rekapituliert werden: Da die Erfahrungen mit dem automatisierten Abgleich darauf hindeuten, dass nur Trefferquoten bis durchschnittlich maximal 15 % erzielt werden, sind schon aus statistischen Gründen Bestände mit weniger als hundert VZE kaum sinnvoll. Hier zeigt sich im Übrigen erneut das Problem fehlender archivrelevanter Daten in der GND, was Personen und Orte gleichermaßen betrifft.⁹² Unter diesen Umständen begünstigt das Kriterium Umfang eindeutig Bestände des 20. Jahrhunderts.⁹³

Methodisch fraglich ist, wie die Grenzen zwischen hoher, mittlerer, niedriger oder keiner (= zu geringer) Eignung des Umfangs gezogen werden sollen. Dies wäre je nach Archiv sehr unterschiedlich, ist von der Verteilung der Extrema abhängig. Im Falle des LAV NRW beruhen die vom Verfasser vorgenommenen Abstufungen letztlich auf eigenen Recherchen im AIS V.E.R.A. und einer daraus vorgenommenen Abschätzung sowie auf Besprechungen mit den Mitarbeitern an den jeweiligen Standorten. Eine umfassende Auswertung über die durchschnittliche Größe von Beständen und Findmitteln existiert nicht, wobei zwischen beidem noch zu unterscheiden wäre: Es gibt Bestände mit über 100.000 VZE, aufgeteilt in mehrere Findbücher, während Findbücher wiederum bis zu ca. 50.000 VZE beinhalten können. Daneben gibt es aber Bestände mit nur einer VZE. Doch selbst bei Anlegung einer präziseren statistischen Auswertung des Umfangs aller Bestände bliebe die Entscheidung über die Geeignetheit in gewissem Maße subjektiv.

Der Umfang an VZE ist zudem auf die Zukunft gerichtet: Werden Verfahren wie NER oder phonetischer Abgleich für Archive anwendbar,⁹⁴ gewinnt das Kriterium stark an Gewicht. Unter den momentanen Bedingungen wird es mit dem Faktor 2 angesetzt.

Es leuchtet ein, dass das Kriterium „Umfang“ auch massiv an Gewicht gewinnt, wenn es mit Kriterium 1 „Strukturierte Erschließungsdaten“ kombiniert wird: Bestände und Findmittel mit großem Umfang und hohem Anteil an strukturierten VZE besitzen eine potenziell hohe Eignung zur Normdatenverknüpfung.

4.2.5. Kriterium 4: Benutzungsfrequenz

Die Häufigkeit der Benutzung ist ein klassisches Kriterium in der Priorisierung von Beständen für die Bestandserhaltung und/oder Digitalisierung. Auch hier ergibt sich allerdings das Problem, dass die Frage, wann ein Bestand als „häufig“ benutzt einzustufen ist, von den jeweiligen spezifischen Gegebenheiten abhängt. Für das LAV

⁹² Die Ortsnormdatenbanken des Landesarchivs Baden-Württemberg, aber etwa auch des Hessischen Landesamtes für geschichtliche Landeskunde gehen deutlich über die in diesen Datenbanken ebenfalls referenzierten Geografika der GND hinaus. Das Landesarchiv Baden-Württemberg hat aber darüber hinaus genau aus diesem Grund die zusätzlichen Ortsnormdaten der eigenen Datenbank in die GND importiert.

⁹³ Natürlich werden gerade die „Massenakten“ des 20. Jahrhunderts mittels Archivierungsmodellen stark komprimiert. Doch sowohl in den Beständen/Findmitteln zu Behörden/Gerichten mit langer Existenzzeit und vielen Aussonderungen als auch zu Spezialfällen wie etwa NS-Unterlagen – bei denen weit weniger oder gar nicht kassiert wurde – sind teils fünfstelligen Zahlen von VZE zu finden.

⁹⁴ Dafür gibt es Indizien, cf. <semantic-media-web.de/referenten/?detail=31> [letzter Zugriff 8.3.2017].

NRW können die Angaben über die Benutzungen aus dem AIS erhoben werden, zumindest für zwei Standorte lagen dem Verfasser Daten aus den letzten Jahren vor.⁹⁵

Die methodische Problematik des Einsatzes von Benutzungsfrequenzen als Grundlage von Priorisierungen besteht in mindestens drei Punkten:⁹⁶

- 1.) Absolute Frequenzen sind nicht aussagekräftig, die Nutzungshäufigkeit muss immer ins Verhältnis zur Bestandsgröße gesetzt werden.
- 2.) Hohe Nutzungsfrequenzen sind oft nicht von Dauer und daher als Grundlage für langfristige Maßnahmen fragwürdig – bei Durchschnittswerten müssen daher Extrema berücksichtigt werden.
- 3.) Die Bestände haben unterschiedliche ‚Startbedingungen‘: Bestände ohne Schutzfristen – also meist solche vor dem 20. Jahrhundert – können leichter recherchiert und genutzt werden als solche mit laufenden Schutzfristen.⁹⁷

Dennoch macht es aus drei Gründen Sinn, Benutzungsfrequenzen als Kriterium für den Normdateneinsatz anzuwenden:

- 1.) Hohe Nutzungsfrequenzen bedeuten gewöhnlich auch entsprechende Rechercheintensität, und zwar gerade über Onlinefindmittel – mit einer Indizierung lässt sich diese Recherche unterstützen, auch wenn nur Teile der Bestände/Findmittel indiziert werden können.
- 2.) Umgekehrt sind niedrige Nutzungsfrequenzen bei ansonsten günstigen Voraussetzungen bei den übrigen Kriterien ein Anreiz, mittels Normdaten die Wahrnehmung des Bestands zu steigern – die Punkteskala wäre dann gegebenenfalls umzukehren.⁹⁸
- 3.) Da die Benutzung schon schutzfristenfreier Bestände potenziell höher ist, schafft dies einen gewissen Ausgleich zum Kriterium des Umfangs, das die ‚Massenbestände‘ des 20. Jahrhunderts begünstigt.

Die bei Abfassung dieser Arbeit verfügbaren Zahlen über die Nutzungsintensität für das LAV NRW sind nicht langfristig genug, um daraus quantifizierende Abstufungen der Eignung nach hoch, mittel, gering und keine zuzulassen. Dieser Punkt muss daher offenbleiben, zumal dafür aus oben genannten Gründen hierfür auch das Verhältnis zwischen Bestandsgröße und Nutzungshäufigkeit miteinbezogen hätte werden müssen. Als Faktor wurde die doppelte Gewichtung angesetzt.

Zusammengefasst: Der methodischen Probleme eingedenk kann Nutzungsintensität aufgrund der Korrelation mit der Verfügbarkeit von Erschließungsdaten im Netz ein Kriterium für die Eignung für den Normdateneinsatz sein. Entschieden werden muss, ob eine hohe oder eine niedrige Frequenz das Eignungskriterium sein sollen.

4.2.6. Kriterium 5: Vorhandene Digitalisate

Das Kriterium zielt zunächst auf den Aspekt der Benutzerfreundlichkeit: Da Normdaten Teil des Semantic Web sind, besteht das grundsätzliche Ideal der medienbruchfreien Verlinkung der zugehörigen Informationsressourcen. Weiter geht das Kriterium in mehrerer Hinsicht mit den archivischen Digitalisierungsstrategien konform: So

⁹⁵ Für die Abteilung W in den Jahren 2007 bis 2012: CRABUS 2013 (für die nötigen Verbesserungen zum Zwecke einer raschen und automatisierten Erhebung der Nutzungsfrequenz in V.E.R.A. cf. ebd. S. 4, 28); für die Abteilung OWL im Jahr 2016: Unpublizierte Excel-Tabelle vom 06.02.2017.

⁹⁶ Zu den Problemen cf. teilweise CRABUS 2013, S. 14–19 und 23–27.

⁹⁷ Natürlich gibt es Ausnahmen, wie etwa Kartenbestände, cf. dazu die umfangreiche Liste bei CRABUS 2013, Anhang.

⁹⁸ Also eine hohe Eignung bei niedriger Nutzungsintensität statt umgekehrt. Eine ähnliche strategische Überlegung kann es in der Priorisierung für Digitalisierungsprojekte geben, indem gerade nicht die am häufigsten benutzten Bestände als besonders geeignet bewertet werden.

sind bei der Digitalisierung Benutzungsfrequenz, archivrechtlicher Status, aber auch (eher subjektive) Punkte wie „historischer Wert“ oder „Rückgratbestände“ zentrale Priorisierungskriterien. Es handelt sich also *per se* schon um herausgehobene Bestände, bei denen damit die Wahrscheinlichkeit eines Treffers in der GND höher anzusetzen ist. Das Ziel der besseren Bekannt- und Zugänglichmachung von Archivgut mittels Digitalisierung wird durch den Einsatz von Normdaten unterstützt.

Allerdings bedeutet die Digitalisierung von Archivgut nicht automatisch, dass die entsprechenden Archivalien schon komplett ohne Nutzungsbeschränkungen im Netz verfügbar sind.⁹⁹ Soll darauf abgehoben werden, müsste das Kriterium auf „vorhandene und veröffentlichungsfähige / veröffentlichte Digitalisate“ umgemünzt werden.

Insgesamt ist dieses Kriterium vor allem indirekt sinnvoll, indem es eine Konformität mit parallelen Digitalisierungsstrategien unterstützt und die Kriterien 2 (Archivrechtlicher Status) und 4 (Benutzungsfrequenz) unterfüttert. Daneben zeigt das Kriterium einen zusätzlichen Nutzerkomfort bei Recherche und Benutzung an. Es wird mit einem nur einfachen Faktor in der Gewichtung gewertet.¹⁰⁰

4.2.7. Kriterium 6: „Rückgratbestände“ / „Historischer Wert“

„Historischer Wert“ oder „Rückgratbestände“ lösen als Priorisierungskriterien mittlerweile bestenfalls Gelächter aus, gelten sie doch als subjektiv, nicht quantifizierbar und angesichts einer ohnehin schon komprimierten Überlieferung als fragwürdig.¹⁰¹

Dennoch macht gerade ein solches subjektives Kriterium als Gegengewicht zu den weitgehend von externen oder technischen Gegebenheiten bestimmten Kriterien 1 bis 4,¹⁰² teilweise auch 5, Sinn.¹⁰³ Denn hier kommt die historische Kenntnis der Archivare über ihre Bestände – auf diese muss für die Vorauswahl der für eine Anreicherung potenziell geeigneten Bestände- und Archivalientypen ohnehin rekuriert werden – noch am ehesten zum Tragen. Um aber nicht bloße Willkür walten zu lassen, sind folgende Leitfragen bei der Einschätzung anzuwenden:

- 1.) Welche Bestände/Findmittel haben das Potenzial, dass darin enthaltene Entitäten in mehreren Beständen vorkommen?
- 2.) Welche Bestände/Findmittel haben das Potenzial, dass darin enthaltene Entitäten archiv- oder gar spartenübergreifend vorkommen?

Beides beruht auf dem Grundsatz des Semantic Web, verschiedene Informationsressourcen unterschiedlicher Provenienz zu bündeln.

Die methodische Problematik des Kriteriums ist dennoch unverkennbar. Daher wurde nur der einfache Faktor in der Gewichtung angesetzt.

⁹⁹ So werden etwa die Entnazifizierungsunterlagen im LAV NRW komplett digitalisiert, aber wegen noch laufender Nutzungsbeschränkungen nicht sofort online gestellt.

¹⁰⁰ Zumal auch mittelfristig nur 5 bis 10 % des gesamten Archivguts digitalisiert werden können, cf. dazu die letzten Empfehlungen der KLA.

¹⁰¹ Dass auch in den anderen Kriterien letztlich subjektives Ermessen steckt, wurde schon in den Vorbemerkungen betont.

¹⁰² Da das Archivrecht stark von anderen Rechtsvorschriften geprägt wird, ist es ebenfalls darunter einzustufen.

¹⁰³ Dieses Problem trifft im Übrigen auch auf Priorisierungskriterien in der Digitalisierung und der Bestandserhaltung teilweise zu.

4.2.8. Bewertungsmatrix und Beispielberechnungen

Die aus den Kriterien erstellte Matrix gestaltet sich wie folgt

Kriterien	Eignung				Faktor	Begründung
	hoch (3 P.)	mittel (2 P.)	niedrig (1 P.)	kein (0 P.)		
1.) Strukturierte Erschließungsdaten (Masken) vorhanden und befüllt	> 50 % der VZE	25–50 % der VZE	5–25 % der VZE	< 5 % der VZE	3	weitgehend nur automatisierter Abgleich möglich
2.) Archivrechtlicher Status						
- a) VZE direkt verknüpfbar (Anteil)	> 50 % der VZE	25–50 % der VZE	5–25 % der VZE	< 5 % der VZE	3	direkte Verlinkung der VZE möglich
- b) FB verknüpfbar (Anteil)	> 50 % der VZE	25–50 % der VZE	5–25 % der VZE	< 5 % der VZE	2	direkte Verlinkung des FB möglich
3.) Ausreichender Umfang an normdaten-geeigneten Archivalien	> 1000 VZE	500–1000 VZE	100–500 VZE	< 100 VZE	2	Aufwand bei nur geringem Umfang lohnt nicht: keine Aussagekraft bei Pilotierung; Anzahl der VZE aussagekräftiger als Regalmeter
4.) Nutzungsintensität 1.) hoch 2.) niedrig	Schätzung oder nach Auswertung des AIS	Schätzung oder nach Auswertung des AIS	Schätzung oder nach Auswertung des AIS	Schätzung oder nach Auswertung des AIS	2	1.) Mehrwert wahrscheinlich auch zukünftig hoch 2.) bessere Sichtbarkeit noch wenig genutzten Archivguts
5.) Digitalisate vorhanden	> 50 % der VZE	25–50 % der VZE	5–25 % der VZE	< 5 % der VZE	1	Stärkere Reichweite bei Verknüpfung; Nachteil: nur für wenige Bestände gültig
6.) Rückgratbestände / Historischer Wert	Wertung des zuständigen Fachdezernats	Wertung des zuständigen Fachdezernats	Wertung des zuständigen Fachdezernats	Wertung des zuständigen Fachdezernats	1	Historischer Wert, Aussagekraft über mehrere Bestände hinweg; Nachteil: subjektiv

Maximal können also 36 Punkte erreicht werden. Ein tatsächlich geeigneter Bestand sollte mindestens über 50 % der möglichen Punkte erzielen.

Für eine Vorauswahl wäre es am günstigsten, die quantifizierbaren Kriterien 1, 3, 4 und 6¹⁰⁴ automatisch aus dem AIS erheben zu können. Diese Möglichkeit bestand bei Abfassung dieser Arbeit im Falle von V.E.R.A. nur teilweise. Daher war der Verfasser auf die unter Kap. 4.2.1. genannten Hinweise der Fachdezernate angewiesen. Beispielrechnungen könnten wie folgt aussehen:

Beispiel 1: LAV NRW, Abt. R., Best. Sammelbestand Personalakten (PS)

	Kriterium 1	Kriterium 2	Kriterium 3	Kriterium 4	Kriterium 5	Kriterium 6	Gesamt
Eignung x Faktor	$3^{105} \times 3 =$ 9 P.	$2 \times 2^{106} =$ 4 P.	$3^{107} \times 2 =$ 6 P.	$2^{108} \times 2 =$ 4 P.	$0 \times 1 = 0$ P.	$2^{109} \times 1 =$ 2 P.	25 P.

Beispiel 2: LAV NRW, Abt. W, Best. Karten A (Katasterkarten)

	Kriterium 1	Kriterium 2	Kriterium 3	Kriterium 4	Kriterium 5	Kriterium 6	Gesamt
Eignung x Faktor	$0^{110} \times 3 =$ 0 P.	$3^{111} \times 2 =$ 6 P.	$3^{112} \times 2 =$ 6 P.	$3^{113} \times 2 =$ 6 P.	$3 \times 1 =$ 1 P.	$2^{114} \times 1 =$ = 2 P.	21 P.

¹⁰⁴ Im Falle des archivrechtlichen Status spielen noch andere Erwägungen als die reinen Laufzeiten eine Rolle, cf. weiter oben.

¹⁰⁵ Maske „Personenbezogene Akte“, für 100 % der VZE genutzt und Datenfelder befüllt.

¹⁰⁶ Laufzeit 1888–2015, teils personenbezogene Schutzfristen, aufgrund § 7 Abs. 3 Satz 2 ArchivG NRW aber für über 50 % mindestens FB verknüpfbar.

¹⁰⁷ Über 1000 VZE.

¹⁰⁸ Genaue Daten lagen nicht vor, eine Abschätzung erfolgte aufgrund der vorliegenden Daten für die Abteilungen W und OWL. Hohe Benutzungsintensität wird hier mit hoher Eignung gleichgesetzt.

¹⁰⁹ Das Vorkommen der erfassten Personen in anderen Beständen ist unsicher; wegen des Archivierungsmodells Personalverwaltung im LAV befinden sich allerdings wahrscheinlich etliche ‚Prominente‘ darunter, die spartenübergreifend relevant sein könnten.

¹¹⁰ Die Maske „Karten/Pläne“ ist zwar vorhanden, aber die strukturierten Felder wurden nicht genutzt, sondern die Ortsangaben über die Klassifikation abgebildet; inwiefern ein automatischer Abgleich demnach überhaupt Treffer erbringen würde, kann ohne empirische Prüfung nicht geklärt werden.

¹¹¹ In geringem Umfang besteht theoretisch noch eine allgemeine Schutzfrist, diese ist faktisch aber zu vernachlässigen, daher sind die VZE direkt verknüpfbar.

¹¹² Über 1000 VZE.

¹¹³ Am häufigsten genutzter Bestand nach CRABUS 2013. Hohe Benutzungsintensität wird hier mit hoher Eignung gleichgesetzt.

¹¹⁴ Einschätzung des Fachdezernats.

Beispiel 3: LAV Abt. W, Best. Q 923 (Justizvollzugsanstalt Münster)

	Kriterium 1	Kriterium 2	Kriterium 3	Kriterium 4	Kriterium 5	Kriterium6	Gesamt
Eignung x Faktor	$3^{115} \times 3 =$ 9 P.	$1 \times 2^{116} =$ 2 P.	$3^{117} \times 2 =$ 6 P.	$1^{118} \times 2 =$ 2 P.	$0 \times 1 =$ 0 P.	$1^{119} \times 1$ $= 1$ P.	20 P.

Für die Gesamtpunktzahlen werden an dieser Stelle keine Eignungen nach „hoch“, „mittel“, „niedrig“ oder „kein“ vergeben. Eine solche Einschätzung ist erst bei Vorliegen einer ausreichenden Anzahl geprüfter Bestände sinnvoll, wenn die Verteilung der Extrema wie der Durchschnittswerte erkennbar wird.

5. Fazit: Handlungsempfehlungen und Ausblick

Zum Abschluss und zugleich als Fazit soll noch eine Handlungsempfehlung formuliert werden, was sowohl für das LAV NRW als auch generell für Archive beim Einstieg in die Thematik „Normdaten“ zu beachten ist. Dies umfasst im Einzelnen folgende, aufeinander aufbauende Punkte:¹²⁰

A) Vorbereitungsphase

- 1.) Organisation sowie Festlegung der Ziele und prioritären Kriterien
- 2.) Auswahl der bevorzugten Normdatenstandards und -typen
- 3.) Auswahl potenziell geeigneter Bestände/Findmittel und Überprüfung der tatsächlichen Eignung der Anreicherung mit Normdaten

Grundsatzentscheidung → wenn negativ: Abbruch

B) Einführungsphase

- 4.) Schaffung der nötigen Voraussetzungen
- 5.) Einführung und kontinuierliche Überprüfung

Zu den einzelnen Phasen sind noch einige Erläuterungen anzubringen:

Phase A)

Zu 1.) Organisatorisch dürfte die Einrichtung einer Projektgruppe am zweckmäßigsten sein. Zunächst muss geklärt werden, was überhaupt mit dem Normdateneinsatz erreicht werden soll.¹²¹ Dazu gehört auch die Frage, ob nur passiv oder auch aktiv Normdaten genutzt beziehungsweise eingebracht werden sollen. Ebenso ist zu erörtern, ob nur die noch unerschlossenen Bestände oder auch bereits erschlossene angereichert werden sollen.

¹¹⁵ Maske „Personenbezogene Akte“, für 100 % der VZE genutzt und Datenfelder befüllt.

¹¹⁶ Laufzeit 1810–1995, vielfach personenbezogene Schutzfristen, geschätzt maximal 25 % verknüpfbar.

¹¹⁷ Über 1000 VZE.

¹¹⁸ Nach CRABUS 2013 nicht unter die intensiv genutzten Bestände zu rechnen. Hohe Benutzungsintensität wird hier mit hoher Eignung gleichgesetzt.

¹¹⁹ Das Vorkommen der erfassten Personen in anderen Beständen ist möglich (Justizunterlagen), archiv- und spartenübergreifend aber unwahrscheinlich.

¹²⁰ Eine erschöpfender Ablaufplan mit allen Optionen kann an dieser Stelle nicht erfolgen.

¹²¹ Cf. dazu die weiter oben in Kap. 2.2. genannten Ziele: Sichtbarkeit, Recherche, Nutzerinteressen, Verknüpfung von Informationsressourcen, maschinelle Auswertbarkeit. Institutionenabhängig sind noch weitere denkbar, etwa Nachnutzung für die Überlieferungsbildung.

Aus Sicht des Verfassers kann dies aus in Kap. 4.2. genannten Gründen kein Kriterium für die Priorisierung sein, aber dies muss jede Archivverwaltung selbst entscheiden.

Weiter müssen die Voraussetzungen und Kriterien abgesteckt werden: etwa, ob schon genutzte Deskriptoren/Thesauri vorhanden sind, sowie die Grundsatzfrage, ob automatisiert oder auch manuell angereichert werden soll, sowie die Frage, welche Kriterien für die Eignung angesetzt werden und wie letztere definiert wird (Punktesystem, Prozentzahlen, Schwellenwerte).¹²²

Diese Fragen stehen in engem Zusammenhang, wie das Beispiel des LAV NRW zeigt: Die Erschließungsrichtlinien des LAV sehen, wie erwähnt, keine Indizierung mehr vor; diese wäre bei manueller Anreicherung Teil der kaum noch praktizierten Stufe 3 (Tiefenerschließung) – im Falle jedoch einer automatisierten Anreicherung würde diese Anreicherung maximal der Stufe 2 (Standardserschließung) zuzuordnen sein.¹²³

Zu 2.) Zu klären ist, welche Normdatenstandards (EAC-CPF, GND, gegebenenfalls weitere) und welche Normdatentypen (Personen, Orte, Körperschaften, Sachschlagworte) genutzt werden sollen. Dies ist zwingend abhängig von den zuvor definierten Zielen.

Zu 3.) Abhängig von Nr. 2.) und 3.) müssen Bestände-, Findmittel- und Archivalientypen identifiziert werden, die sich für eine Anreicherung eignen *könnten*.¹²⁴ Als weiterer Schritt muss abhängig von den unter 1.) definierten Kriterien zumindest stichprobenartig empirisch geprüft werden, wie viele Bestände sich *tatsächlich* eignen. Überprüft werden sollte nicht nur die prozentuale Trefferquote insgesamt, sondern auch die spezielle Trefferquote über verschiedene Bestände-/Findmittelarten hinweg sowie wenn möglich, wie viel Prozent der ermittelten Normdaten in mehreren Beständen/Findmitteln auftauchen.

Zu 4.) Die Entscheidung läuft darauf hinaus, ab welchen Schwellenwerten von der grundsätzlichen Eignung des eigenen Archivguts ausgegangen und die Normdatenanreicherung als nützlich und notwendig erachtet wird. Dies kann letztlich nur individuell – gemäß den Zielen, zur Verfügung stehenden Mitteln et cetera – entschieden werden.¹²⁵

Phase B)

Zu 1.) Die erforderlichen administrativen Voraussetzungen sind zu schaffen. Dies betrifft zum Beispiel die Fragen des wirtschaftlichen Aufwandes, der Anpassung der Erschließungsrichtlinien und der Benennung von zuständigen Fachleuten beziehungsweise Organisationseinheiten¹²⁶ sowie mögliche Schulungsbedarfe für die übrigen Mitarbeiter. Erst an dieser Stelle sollte die Frage der technischen Umsetzung gestellt werden: Denn ob schon selbst komplex, muss die IT-Anwendung den konzeptionellen Vorgaben angepasst werden und nicht umgekehrt.¹²⁷

¹²² Es ist je nach Zielvorgabe denkbar, im Gegensatz zum in vorliegender Studie angewendeten Modell Ausschlusskriterien (etwa Strukturierte Erschließungsdaten, Archivrecht) zu definieren.

¹²³ Sofern schon erschlossene Bestände damit angereichert würden, handelte es sich quasi um optionale Angaben in der Standardserschließung; bei einer Ersterschließung dagegen wäre der automatische Abgleich im Grunde dem Import externer Daten aus einer Registratur zuzuordnen und daher eher noch Teil der Stufe 1 (Ersterschließung).

¹²⁴ Dies kann nur individuell je nach Archivverwaltung geschehen. Für größere Archive bieten allerdings die unter Kap. 4.2.1. erwähnten, von den Mitarbeitern des LAV NRW genannten Typen schon gute Hinweise.

¹²⁵ Rücksprache mit bereits Normdaten einsetzenden Archiven ist zu empfehlen; cf. auch die deutlich auf über 50 % Eignung hinweisenden Ergebnisse von Fragebogen Teil B), Nr. 3.).

¹²⁶ Die Organisationsfrage sollte angesichts der spartenübergreifenden Relevanz und des notwendigen dauerhaften Umgangs mit einer Normdaten vorhaltenden Institution nicht unterschätzt werden.

¹²⁷ Zu den technischen Möglichkeiten für die GND cf. Fragebogen Teil B), Nr. 10). Wegen der bisher weitgehend nur manuell durchgeführten Anreicherung nannten die meisten Archive die Webanwendung, für automatisierte Verfahren dürfte allerdings SRU zukünftig bedeutsamer sein, wohingegen ein eigener Client mit den gängigen AIS kaum zu kombinieren sein

Zu 2.) Die Evaluation des ‚Erfolgs‘ sollte sich ebenfalls an quantifizierbaren Kriterien orientieren. Denkbar sind etwa: Zugriffe auf angereicherte Bestände und Onlinefindmittel, Verbreitung der eigenen Angebote bei anderen Anbietern, Zahl der Anfragen bei angereicherten Beständen, Nutzerbefragungen. Zur kontinuierlichen Überprüfung gehört im Übrigen auch, sowohl die generellen Entwicklungen im Bereich des Semantic Web zu beobachten als auch den Kontakt zu den fachlich einschlägigen Institutionen und Ansprechpartnern zu halten. Denn die Entwicklungsdynamik ist hier derartig hoch, dass mittelfristig wahrscheinlich wieder Veränderungen vorgenommen werden müssen, was ohne entsprechende externe Expertise nicht zu leisten sein wird.¹²⁸ Sofern nur passiv Normdaten genutzt werden, ist die Frage aktiver Beiträge aus weiter oben genannten Gründen im Auge zu behalten.

Noch drei weitere Handlungsempfehlungen seien angefügt, die sich aus der Normdaten-Thematik ergeben, teilweise aber darüber hinausgehen:

- 1.) Datenqualität: Bei der schon jetzt enormen Masse an Erschließungsdaten, die zwangsläufig weiter anwachsen wird, sollte strukturierten Erschließungsdaten der Vorzug vor unstrukturierten Daten in Titelaufnahmen oder Enthältvermerken gegeben werden. Eine von Beginn an schlechte Datenqualität ist im Grunde irreparabel und hinsichtlich zukünftiger Nutzererwartungen wie Nutzungsmöglichkeiten durch maschinelle Auswertungs- und Anreicherungsverfahren äußerst kritisch zu betrachten. Sowohl die Fachanforderungen an die Archivsoftware als auch die eigene Erschließungspraxis müssen dies berücksichtigen.
- 2.) Kooperation und Kommunikation: Bei derartigen Themen, die archiv- oder gar spartenübergreifende Relevanz und zugleich eine starke Entwicklungsdynamik besitzen, müssen Archive kontinuierlich und öffentlich den Arbeitsstand austauschen und gemeinsame Anforderungen festlegen.¹²⁹ Denkbar wären entsprechende Arbeitsgruppen über die Fachverbände oder die KLA. Denn abgesehen von der schon arbeitsökonomisch gebotenen Nachnutzung der Erfahrungen anderer: Ohne gemeinsame Stimme ist es unmöglich, gegenüber großen Institutionen wie der DNB den eigenen Anforderungen Gehör zu verschaffen. Und letzteres wird zudem ohne eigene aktive Beiträge kaum möglich sein.
- 3.) Archive als Teil der Informationsinfrastruktur: Die unter 1.) und 2.) genannten Anforderungen sind auch deshalb zu beachten, weil sie einen Teil des seitens der großen Forschungsförderungsinstitutionen wie auch der öffentlichen Hand (Bund, Länder) seit etlichen Jahren forcierten Modells einer vernetzten und verteilten Informationsinfrastruktur darstellen.¹³⁰ Auch hier sollten Archive nicht bloß passiv rezipieren, sondern aktiv beitragen, was nur mit entsprechenden konzeptionellen Grundlagen und archivübergreifenden Anforderungsprofilen funktionieren wird.

Zusammengefasst: Archive werden sich dem Thema Normdaten als Teil des Semantic Web langfristig nicht entziehen können. Mindestens sollte daher gegenüber Nutzern und externen Einrichtungen begründet werden können, warum sich unter den aktuellen Voraussetzungen eine Normdatenanreicherung nicht lohnt. In weiterer Perspektive ist das Thema allerdings auch geeignet, sich kritisch mit den eigenen Arbeitsmethoden in den verschiedenen Fachaufgaben auseinanderzusetzen.

dürfte. Doch selbst bei der Webanwendung bedarf es oft zwischengeschalteter Systeme, wie sie etwa das Landesarchiv Baden-Württemberg einsetzt.

¹²⁸ Dies ist zudem ein weiteres Argument, schon jetzt eher auf automatisierte Verfahren zu setzen.

¹²⁹ Dass dafür womöglich andere Formen als die Literatur, Archivtage oder die übliche Mund-zu-Mund-Propaganda nötig sind, steht auf einem anderen Blatt.

¹³⁰ Die inzwischen zahlreichen Strategiepapiere und Konzeptionen dazu sind in den Webauftritten der DFG, des Wissenschaftsrates, des Rates für Informationsinfrastrukturen (RfII) oder der Bund-Länder-Kommission offengelegt.

Zusammenfassung

Normdaten, ursprünglich aus der bibliothekarischen Erschließung erwachsen, werden als Teil des Semantic Web von immer mehr Kulturinstitutionen für ihre Bestände genutzt, sodass sich auch Archive damit befassen müssen. Den Vorteilen in Recherche und Breitenwirkung steht als Hauptnachteil der hohe Aufwand insbesondere bei manueller Anreicherung des heterogenen Archivguts gegenüber. Geprüft werden muss daher, welcher Normdaten-Standard und welche Normdatentypen genutzt werden sollen, welche Bestände/Findmittel in Frage kommen und ob der möglichst empirisch nachzuweisende Mehrwert (Trefferquote) den Aufwand übersteigt. Die Arbeit plädiert für eine Nutzung der GND durch Archive und liefert ein Prüfraster mit sechs Kriterien, um die Eignung archivischer Bestände/Findmittel messen zu können. Abschließend werden einige Handlungsempfehlungen gegeben, die sowohl den Umgang mit Normdaten als auch daraus resultierend zukünftige Anforderungen an die archivischen Fachaufgaben betreffen.

Abkürzungsverzeichnis

AIS	=	Archivinformationssystem
ArchivG NRW	=	Archivgesetz Nordrhein-Westfalen (in der Fassung vom 30.09.2014)
BArchG	=	Bundesarchivgesetz (in der Fassung vom 10.03.2017)
BSZ-BW	=	Bibliotheksservice-Zentrum Baden-Württemberg
DFG	=	Deutsche Forschungsgemeinschaft
DNB	=	Deutsche Nationalbibliothek
EAC	=	Encoded Archival Context
EAC-CPF	=	Encoded Archival Context – Corporate Bodies, Persons and Families
EAD	=	Encoded Archival Description
GND	=	Gemeinsame Normdatei
GStA PK	=	Geheimes Staatsarchiv Preußischer Kulturbesitz
GVG	=	Gerichtsverfassungsgesetz (in der Fassung vom 22.12.2016)
HStAW	=	Hessisches Hauptstadtarchiv Wiesbaden
ICA	=	International Council on Archives
IFG	=	Informationsfreiheitsgesetz
ISAAR(CPF)	=	International Standard Archival Authority Record for Corporate Bodies, Persons and Families [2nd Edition]
ISAD(G)	=	General International Standard Archival Description
ISDF	=	International Standard for Describing Functions
ISDIAH	=	International Standard for Describing Institutions with Archival Holdings
KLA	=	Konferenz der Leiterinnen und Leiter der Archivverwaltungen des Bundes und der Länder
KunstUrhG	=	Gesetz betreffend das Urheberrecht an Werken der bildenden Künste und der Photographie (in der Fassung vom 16.02.2001)
LAV NRW	=	Landesarchiv Nordrhein-Westfalen
LOD	=	Linked Open Data
MARC	=	Machine-Readable Cataloging
OWL	=	Ostwestfalen-Lippe
P.	=	Punkte
R	=	Rheinland

RDA	=	Resource Description and Access
RDF	=	Resource Description Framework
UrhG	=	Urheberrechtsgesetz (in der Fassung vom 20.12.2016)
VIAF	=	Virtual International Authority File
VÖA	=	Verband Österreichischer Archivarinnen und Archivare
VZE	=	Verzeichnungseinheit(en)
W	=	Westfalen
XML	=	Extensible Markup Language

Quellen- und Literaturverzeichnis

Literatur

- [BSZ-BW] GND-Schulung 1 = [BSZ-BW], GND-Schulung Tag 1 (online: wiki.bsz-bw.de/lib/exe/fetch.php?media=v-team:katalogisierung:normdaten:gnd:einfuehrung:gnd_rda_tag1.pdf) [8.3.2017]).
- [BSZ-BW] GND-Schulung 2 = [BSZ-BW], GND-Schulung Tag 2 (online: wiki.bsz-bw.de/lib/exe/fetch.php?media=v-team:katalogisierung:normdaten:gnd:einfuehrung:gnd_rda_tag2.pdf) [8.3.2017]).
- [BSZ-BW] Katalogisierungshandbuch 2016 = [BSZ-BW], Katalogisierungshandbuch. Normdaten der GND für Formal- und Sacherschließung. Stand 19.10.2016 (online: verbund-swop.bsz-bw.de/files/171/kathb_gnd.pdf) [9.3.2017]).
- [DNB] 2012 = [DNB], Bereitstellung des ersten GND-Grundbestandes. April 2012 (online: www.dnb.de/SharedDocs/Downloads/DE/DNB/service/rundschreibenBereitstellungGNDGrundbestand.pdf?_blob=publicationFile) [2.3.2017]).
- [DNB] EH-G 2 = [DNB] EH-G. Gleichnamigkeit bei geografischen Einheiten (online: wiki.dnb.de/download/attachments/90411357/EH-G-02.pdf?version=5&modificationDate=1445945288000&api=v2) [13.3.2017]).
- [DNB] EH-G 06 = [DNB] EH-G 06. Namensänderungen bei Geografika (online: wiki.dnb.de/download/attachments/90411357/EH-G-06.pdf?version=2&modificationDate=1443430195000&api=v2) [8.3.2017]).
- [DNB] EH-K 21 = [DNB] EH-K 21. Namensänderung bei Körperschaften und Konferenzen (online: wiki.dnb.de/download/attachments/90411359/EH-K-21.pdf) [8.3.2017]).
- [DNB] EH-P 16 = [DNB] EH-P 16. Individualisierungsrichtlinie (online: wiki.dnb.de/download/attachments/90411361/EH-P-16.pdf) [13.03.2017]).
- [DNB] EH-P 17 = [DNB] EH-P 17. Codes (online: wiki.dnb.de/download/attachments/90411361/EH-P-17.pdf) [9.3.2017]).
- [DNB] GND-Codes im Unterfeld \$4 2016 = [DNB], GND-Codes für Beziehungen im Unterfeld \$4 alphabetisch nach Code. Stand 19.11.2016 (online: www.dnb.de/SharedDocs/Downloads/DE/DNB/standardisierung/inhaltsschliessung/gndCodes.pdf?_blob=publicationFile) [09.03.2017]).
- [DNB] Jahresbericht 2015 = [Deutsche Nationalbibliothek], Zwei Null Eins Fünf. Jahresbericht 2015. 2016 (URN: [urn:nbn:de:101-2016-052300](http://nbn-resolving.org/urn:nbn:de:101-2016-052300)).
- [Hessisches Landesarchiv] 2016 = [Hessisches Landesarchiv¹³¹], Abschlussbericht zum Modellprojekt 13/2015: Risikoanalyse durch Schadenserfassung und Beständepriorisierung. Praxistest einer Methode zur Erstellung eines Arbeitsprogramms für die Erhaltung des schriftlichen Kulturguts am Beispiel des Hessischen Staatsarchivs Darmstadt. 2016 (online: landesarchiv.hessen.de/sites/landesarchiv.hessen.de/files/content-downloads/Schadenserfassung%20und%20Best%C3%A4nde%20priorisierung_2016.pdf) [26.9.2016]).
- [ICA] 2016 = [ICA], Records in Context. A Conceptual Model for Archival Description. Consultation Draft v. 1. September 2016 (online: www.ica.org/sites/default/files/RiC-CM-0.1.pdf) [8.2.2017]).
- [ICA] ISAAR(CPF) 2007 = [ICA] ISAAR(CPF) – Internationaler Standard für archivische Normdaten (Körperschaften, Personen, Familien). Zweite Ausgabe. Herausgegeben und bearbeitet vom Komitee für Erschließungsstandards des Internationalen Archivrates (ICA-CDS) Wien 2004. Deutsche Fassung übersetzt und bearbeitet von Nils Brübach / Jenny Kotte / Paul Vogt. Dresden/Hamburg/Vaduz 2007 (online: www.ica.org/sites/default/files/CBPS_Guidelines_ISAAR_Second-edition_DE.pdf) [16.2.2017]).

¹³¹ Verfasser: Lars Adler / Johannes Kistenich.

- [startext] 2007 = [startext GmbH], Handbuch V.E.R.A. Verzeichnung. Stand 29.10.2007 [unveröffentlicht, internes Dokument des LAV NRW].
- Becker 2013 = Irmgard Christa Becker, Stellungnahme der Archivschule Marburg zur Nutzung von RDA (Resource Description and Access) in Archiven (online: <www.dnb.de/SharedDocs/Downloads/DE/DNB/standardisierung/rdaKultur2013Becker.pdf?__blob=publicationFile> [10.01.2016])
- Becker 2015 = Irmgard Christa Becker (Hg.), Netz werken. Das Archivportal-D und andere Portale als Chance für Archive und Nutzung. Beiträge zum 19. Archivwissenschaftlichen Kolloquium der Archivschule Marburg 2015 (Veröffentlichungen der Archivschule Marburg 61).
- Bischoff 2007 = Frank M. Bischoff (Hg.), Benutzerfreundlich – rationell – standardisiert. Aktuelle Anforderungen an archivierte Erschließung und Findmittel. Beiträge zum 11. Archivwissenschaftlichen Kolloquium der Archivschule Marburg. Marburg 2007 (Veröffentlichungen der Archivschule Marburg 46).
- Brübach 2015A = Nils Brübach, Austauschformate und Normierung – welche Anforderungen stellen Portale an Archive?, in: Becker 2015, S. 157–182.
- Brübach 2015B = Nils Brübach, Erschließung im 21. Jahrhundert, in: Brandenburgische Archive 32 (2015), S. 3–7.
- Crabus 2013 = Mirko Crabus, Nutzerstruktur und Beständenutzung im LAV NRW W (2007–2012). Transferarbeit im Rahmen des Referendariats für den höheren Archivdienst. [Marburg] 2013 (unveröffentlicht, masch.).
- Eckes 2014 = Georg Eckes, Institutionenübergreifende Integration von Normdaten. Kooperativ gepflegte Normdaten: das Kooperationsprojekt IN2N, in: Info7 29/3 (2014), S. 38–39.
- Eito-Brun 2016 = Ricardo Eito-Brun, Remote Access to EAC-CPF Context and Authority Records for Metadata Indexing: a Solution Based on Open Information Retrieval Standards, in: Archival Science 16/2 (2016), S. 149–165.
- Fricke / Schludi 2016 = Thomas Fricke / Ulrich Schludi, Nutzerberatung im Internet. Von der Fragestellung zum Bestand, in: Neue Wege ins Archiv 2016, S. 147–155.
- Haffer 2013 = Dominik Haffer, Digitalisierung und Internetpräsentation von Archivbeständen. Empfehlungen für die Priorisierung und Online-Stellung am Beispiel des Sächsischen Staatsarchivs, in: Karsten Uhde (Hg.), Von A(mtsdruckschriften) bis Z(eitgeschichtliche Sammlungen) – Vielfalt im Archiv. Ausgewählte Transferarbeiten des 43. und 44. Wissenschaftlichen Kurses an der Archivschule Marburg. Marburg 2013 (Veröffentlichungen der Archivschule Marburg 56), S. 109–140.
- Hoen / Gessmann 2008 = Barbara Hoen / Sebastian Geßmann, Die Schönheit der Chance. Erschließungsstandards im Landesarchiv NRW, in: Archivar 61/1 (2008), S. 69–73.
- Imhof 2008 = Andres Imhof, Using International Standards to Develop a Union Catalogue for Archives in Germany. Aspects to Consider Regarding Interoperability between Libraries and Archives, in: D-Lib Magazine 14 (2008) (online: <dlib.org/dlib/september08/imhof/09imhof.html> [8.3.2017]).
- Jagodzinski 2016 = Silke Jagodzinski, Europäische Archive und das Web 2.0, in: Neue Wege ins Archiv 2016, S. 87–97.
- Kluttig 2014 = Thekla Kluttig, Gemeinsame Normdatei und Archive, in: Clemens Rehm / Monika Storm / Andrea Wettmann (Hg.), Nachlässe – Neue Wege der Überlieferung im Verbund. Gemeinsame Frühjahrstagung FG 1 und FG 6. Halle a. d. Saale 2014 (Veröffentlichungen des sächsischen Staatsarchivs. Reihe A: Archivverzeichnisse, Editionen und Fachbeiträge. Bd. 17), S. 85–88.
- Krauth 2015 = Wolfgang Krauth, Archive und Onlineportale. Thesen für den weiteren Erfolg, in: Archivar 68/1 (2015), S. 6–9.
- Müller 2015 = Gerhard Müller, Die Normdatei als ein Mittel in der Erschließung von Archivbeständen, in: Brandenburgische Archive 32 (2015), S. 8–16.
- Müller 2007 = Peter Müller, Schnell zum Ziel – Erschließungspraxis und Benutzererwartungen im Internetzeitalter, in: Bischoff 2007, S. 37–63.
- Neudecker / Rehm 2016 = Clemens Neudecker / Georg Rehm, Digitale Kuratierungstechnologien für Bibliotheken, in: 027.7. Zeitschrift für Bibliothekskultur 4 (2016), S. 104–116.

- Neue Wege ins Archiv 2016 = Neue Wege ins Archiv. Nutzer. Nutzung. Nutzen. 84. Deutscher Archivtag in Magdeburg. Redaktion: Monika Storm in Verbindung mit Irmgard Christa Becker. Fulda 2016 (Tagungsdokumentationen zum Deutschen Archivtag 19).
- Ottosson 2007 = Per-Gunnar Ottosson, Implementing ISAAR(CPF) and EAC in the National Archival Database of Sweden (NAD), in: Bischoff 2007, S. 161–176.
- Reinhardt / Mundt 2016 = Christian Reinhardt / Sebastian Mundt, Nutzerstudie 2016 des Hessischen Landesarchivs. Erste Ergebnisse. 2016 (online: landesarchiv.hessen.de/sites/landesarchiv.hessen.de/files/content-downloads/Nutzerstudie%202016%20Ergebnisse_final-20160818_0.pdf) [27.9.2016]).
- Riedel 2013 = Julia Anna Riedel, Online-Angebote von Archiven. Auswertung einer Nutzerbefragung, in: Archivar 66/1 (2013), S. 51–53.
- Sandner 2013 = Peter Sandner, Linked Open Data in archivischen Fachinformationssystemen, in: Arbido 2013/4 (2013), S. 30–33.
- Scheven 2015 = Esther Scheven, Zukunft und Weiterentwicklung der verbalen Inhaltserschließung – aktueller Stand der Diskussion. Präsentation 27.5.2015 (online: opus4.kobv.de/opus4-bib-info/files/1980/Vortrag-Zukunft-Inhaltserschlie%C3%9Fung.pdf) [23.3.2017]).
- Snae 2007 = Chakkrit Snae, A Comparison and Analysis of Name Matching Algorithms (online: waset.org/publications/8664/a-comparison-and-analysis-of-name-matching-algorithms) [13.3.2007]).
- Sperl 2016 = Karin Sperrl, Die archivischen Erschließungsstandards ISDIAH, ISAD(G), ISAAR(CPF) in der Praxis – Möglichkeiten der Umsetzung, in: Scrinium 70 (2016), S. 43–57.
- Sprau 2014 = Mirjam Sprau, Die Akten des Generalbundesanwalts beim Bundesgerichtshof. Ein Erschließungsmodell. Transferarbeit im Rahmen der Laufbahnprüfung für den höheren Archivdienst. [Marburg] 2014 (unveröffentlicht, masch.).
- Voß u.a. 2014 = Jakob Voß / Susanna Bausch / Julian Schmitt / Jasmin Bogner / Viktoria Berkelmann / Franziska Ludemann / Oliver Löffel / Janna Kitroschat / Maiia Bartoshevskaja / Katharina Seljuzki, Normdaten in Wikidata. Version 1.0. 2014 (online: hshdb.github.io/normdaten-in-wikidata/normdaten-in-wikidata.html) [9.1.2017]).
- Wahrmann 2014 = Carl Christian Wahrmann, Im Portal und lokal. Überlegungen für den Umgang mit Erschließungsinformationen und Digitalisaten im Archivportal-D/DDB und dem Onlineinformationssystem des Sächsischen Staatsarchivs. Transferarbeit im Rahmen der Laufbahnprüfung für den höheren Archivdienst an der Archivschule Marburg – Hochschule für Archivwissenschaft. [Marburg] 2014 (unveröffentlicht, masch.).
- Wettlaufer / Westphal 2014 = Jörg Wettlaufer / Sina Westphal, Digital Humanities, in: Archivar 67/3 (2014), S. 270–277.
- Wiech 2013 = Martina Wiech, Der Kriterienkatalog des Landesarchivs Nordrhein-Westfalen für die inhaltliche Auswahl von Archivgut zur Digitalisierung, in: Unsere Archive 58 (2013) S. 12–16.
- Wiechmann 2015 = Brigitte Wiechmann, Spartenübergreifende Nutzung der Gemeinsamen Normdatei (GND) am Beispiel des Projekts „Institutionenübergreifende Integration von Normdaten (IN2N)“, in: o-bib 2/4 (2015), S. 151–160.
- Zedlitz 2017 = Jesper Zedlitz, Biographische Normdaten – ein Überblick, in: Archivar 70/1 (2017), S. 22–25.
- Ziwes 2014 = Franz-Josef Ziwes, Archive als Leuchttürme. Die Erschließung mit Normdaten als Aufgabe und Chance, in: Archive ohne Grenzen. Erschließung und Zugang im europäischen und internationalen Kontext. 83. Deutscher Archivtag in Saarbrücken. Hg. vom VdA. Red. Monika Storm in Verbindung mit Irmgard Christa Becker. Fulda 2014 (Tagungsdokumentationen zum Deutschen Archivtag 18), S. 79–87.
- Ziwes 2011 = Franz-Josef Ziwes, Überlieferungsbildung und die Intelligenz im Web. Digitale Hilfsmittel bei der Bewertung personenbezogener Unterlagen [Präsentation beim Workshop „Ziele und Methoden archivischer Bewertung“ im Landesarchiv Baden-Württemberg am 1.12.2010] [2011] (online: www.landesarchiv-bw.de/sixcms/media.php/120/52526/Workshop_Ziwes_Beitrag.pdf) [29.3.2017]).
- Zunstein 2012 = Philipp Zunstein, Die Rolle des Semantic Web für Bibliotheken: Linked Open Data und mehr: Welche Strategien können hier Bibliotheken in die Zukunft führen?, in: Perspektive Bibliothek 1/1 (2012), S. 81–102 (online: journals.ub.uni-heidelberg.de/index.php/bibliothek/article/download/9398/3275) [15.3.2017]).

Internetressourcen

- Aumann 2016 = Stefan Aumann, Vernetzung heterogener Informationsressourcen über Normdaten [Vortrag auf der Tagung Genealogie PLUS: Partnership, Open Access und maschinengestützte Indizierung. Hessisches Staatsarchiv Marburg, 24./25.11.2016]:
<[youtube.com/watch?v=oYKftzkrY4U&index=13&list=PLElrWLCQvZaSEXe7_YoESBYszdqJSAH0T](https://www.youtube.com/watch?v=oYKftzkrY4U&index=13&list=PLElrWLCQvZaSEXe7_YoESBYszdqJSAH0T)> [15.03.2017].
- Graf 2015 = Klaus Graf, GND (richtig) nutzen! (Blogbeitrag vom 7.8.2015: <archiv.twoday.net/stories/1022465461/> [15.3.2017]).
- Graf 2013A = Klaus Graf, Die Archive und die GND (Blogbeitrag vom 25.3.2013: <archivalia.hypotheses.org/7331> [15.3.2017]).
- Graf 2013B = Klaus Graf, Vernetzung ortsbezogener historischer Daten (Blogbeitrag vom 24.11.2013: <archivalia.hypotheses.org/5792#comments> [15.03.2017]).
- Lordick 2016 = Harald Lordick, Fachspezifische und nutzerzentrierte Perspektiven – Quellen vernetzen mit der Gemeinsamen Normdatei (Blogbeitrag vom 27.11.2016: <djjd.hypotheses.org/1181> [15.3.2017]).
- Online GND (OGND): <swb.bsz-bw.de/DB=2.104/START_WELCOME> [15.3.2019]).

Besprechungsvermerke

- Besprechung mit Stefan Aumann und Lutz Vogt, Hessisches Landesamt für geschichtliche Landeskunde, 15.02.2017
- Besprechung mit Mechthild Black-Veldtrup, Landesarchiv NRW, Abt. W, 01.02.2017
- Besprechung mit Beate Dördelmann, Landesarchiv NRW, Abt. W, 28.02.2017
- Besprechung mit Daniel Fähle, Landesarchiv Baden-Württemberg / Archivportal-D, 18.01.2017
- Besprechung mit Daniel Fähle und Andreas Neuburger, Landesarchiv Baden-Württemberg, 13.02.2017
- Besprechung mit Jörg Franzkowiak, Landesarchiv NRW, Abt. R, 19.01.2017
- Besprechung mit Martin Früh, Landesarchiv NRW, Abt. R, 20.01.2017
- Besprechung mit Bastian Gillner und Peter Klefisch, Landesarchiv NRW, Abt. R, 20.01.2017
- Besprechungen mit Ralf Maria Guntermann, Landesarchiv NRW, Abt. W, 23.02. und 14.03.2017
- Besprechung mit Johannes Kistenich-Zerfaß, Staatsarchiv Darmstadt, 20.03.2017
- Besprechung mit Benjamin Kram, Landesarchiv NRW, Fachbereich Grundsätze, 21.03.2017
- Besprechung mit Armin Kühn, Bibliotheksservice-Zentrum Baden-Württemberg (BSZ-BW), 02.02.2017
- Besprechung mit Matthias Meusch, Landesarchiv NRW, Abt. R, 19.01.2017
- Besprechung mit Peter Sandner, Hessisches Hauptstaatsarchiv Wiesbaden, 08.02.2017
- Besprechung mit Klaus Schleiter, Archivschule Marburg, Systemadministrator, 08.03.2017
- Besprechung mit Sina Westphal, Bundesarchiv, Referat AT 2, 03.02.2017
- Besprechungen mit Martina Wiech, Landesarchiv NRW, Abt. R, 19.01. und 15.03.2017
- Besprechung mit Uwe Zuber, Landesarchiv NRW, Abt. R, 20.01.2017

Eingesandte Fragebögen und weitere schriftliche Mitteilungen

Anmerkung: Bis auf das Sächsische Staatsarchiv setzten alle hier aufgeführten Archive zum Zeitpunkt der Abfassung Normdaten entweder schon ein oder waren in entsprechenden Planungen.

- Holger Berwinkel, Politisches Archiv des Auswärtigen Amtes, 07.02.2017
- Daniel Fähle, Landesarchiv Baden-Württemberg, / Archivportal-D, 02.02.2017
- Sebastian Gleixner, Bundesarchiv, Referat AT 3, 07.02.2017
- Wolfgang Hegewaldt, Archiv der Akademie der Bildenden Künste, 14.02.2017

Linda Krause, Sächsisches Staatsarchiv, 03.02.2017

Franziska Mücke, Geheimes Staatsarchiv Preußischer Kulturbesitz, 28.02.2017

Klaus Nippert, KIT-Archiv, 16.01.2017

Kristina Starkloff, Archiv der Max-Planck-Gesellschaft, 19.01. und 24.01.2017

Sina Westphal, Bundesarchiv, Referat AT 2, 30.01.2017

Franz-Josef Ziwes, Landesarchiv Baden-Württemberg, 13.01.2017

Anhang: Formular Fragebogen

NB: Die Fragen unter A) wurden nur den externen Ansprechpartnern gestellt, bei den Befragten innerhalb des Landesarchivs NRW aber ausgelassen. Zudem wurden intern bei Teil B) nur die Fragen Nr. 1.) bis 4.) sowie 11.) [als Nr. 5.)] gestellt.

A) Grundsätzliches/Organisation

1.) Werden in Ihrer Einrichtung Normdaten genutzt?

- ja nein noch nicht, aber geplant

Erläuterungen (optional):

2.) Wer ist mit der Thematik Normdaten fachlich befasst?

Werden in Ihrer Einrichtung in der Erschließung standardmäßig auch Deskriptoren im Archivinformationssystem angelegt?

- ja nein noch nicht, aber geplant

3.) Bei welchen archivischen Fachaufgaben werden Normdaten eingesetzt?

- Einzelpersonen
 Projektgruppe (bitte Projekt/e benennen):
 Bestimmten Referaten / Dezernaten / Abteilungen zugewiesen (bitte benennen):

Erläuterungen (optional):

4.)

- Erschließung
 Überlieferungsbildung
 Sonstiges (bitte benennen):

B) Archivfachliche Fragen

1.) Welche Beständearten beziehungsweise Archivaliengattungen eignen sich Ihrer Meinung für den Einsatz von Normdaten? (vgl. auch Frage B2)

2.) Welche Arten von Normdaten eignen sich Ihrer Meinung nach für die archivische Erschließung? (vgl. auch Frage B1)

- Personen
 Orte
 Körperschaften
 Sonstige (bitte benennen)

3.) Wie viele der Bestände und/oder Verzeichnungseinheiten Ihrer Einrichtung eignen sich für Einsatz von Normdaten in der Erschließung (Schätzung)?

- unter 25 %
- zwischen 25 und 50 %
- über 50 %

4.) Wie verteilen sich die für Normdaten geeigneten Bestände/Verzeichnungseinheiten chronologisch (Schätzung in Prozent?)

Erläuterung: Die gewählten zeitlichen Grenzen haben verwaltungs- bzw. verfassungsgeschichtliche Hintergründe, insofern für die jeweiligen Phasen von spezifischen Arten von Unterlagen, etwa hinsichtlich der Quantitäten und formalen Gestaltungsmerkmale, ausgegangen wird.

- vor 1800:
- 1800 bis 1945:
- 1945 bis heute:

5.) Nutzen Sie Normdaten nur passiv oder auch aktiv?

- Passiv: Nur für die eigenen Fachaufgaben
- Passiv und Aktiv: Für die eigenen Fachaufgaben, zusätzlich wird aktiv eigenes Wissen in die Normdateien eingebracht (Neuanlage; Korrektur/Ergänzung)
- Passiv, aber aktive Befüllung geplant

6.) Welche Vorteile bringt der Einsatz von Normdaten für die einzelnen Fachaufgaben (bitte bewerten)?

Fachaufgabe	hoher Vorteil	mittlerer Vorteil	kein Vorteil
Erschließung			
Überlieferungsbildung			
Benutzung			
Öffentlichkeitsarbeit			
Sonstiges (bitte benennen)			

7.) Bringt der Einsatz von Normdaten auch Nachteile (bitte bewerten)?

- Ja, und zwar:

Art des Nachteils	hoher Nachteil	mittlerer Nachteil	kein Nachteil
Zeit- und Personalkosten			
hoher Schulungsbedarf			
nur für wenige Bestände-/ Findbucharten geeignet			
Sonstiges:			

Momentan keine erkennbar

8.) Wie stufen Sie die Gewichtung von Vor- und Nachteilen ein?

- Vorteile und Nachteile sind etwa gleich groß
- Vorteile überwiegen die Nachteile
- Nachteile überwiegen die Vorteile

9.) Welche technischen Instrumente setzen Sie zur Nutzung von Normdaten ein (Mehrfachnennungen möglich)?

- Eigener Client (über die Normdaten speichernde Institution)
- Webanwendung
- SRU-Schnittstelle
- BEACON-Dateien
- Sonstiges (bitte benennen)

10.) Was müsste aus Ihrer Sicht geschehen, um den Einsatz von Normdaten im Archivbereich noch attraktiver und/oder sinnvoller erscheinen zu lassen?

11.) Weitere Bemerkungen zur Thematik