

Initiative für Normdaten und Vernetzung: GND-Entwicklungsprogramm 2017-2021

Stand 06/2017

Inhalt:

1	Perspektiven und Herausforderungen	2
1.1	Normdaten als Rückgrat eines maschinenlesbaren, semantischen Netzes der Kultur und Wissenschaft	2
1.2	Herausforderung: Modernisierung der GND	2
2	Relevante Vorarbeiten	4
2.1	Generelles.....	4
2.2	Erneuerung der Organisation: die GND-Kooperative	4
2.3	CULTUREGRAPH – Infrastrukturelles „Ökosystem“ für die GND	6
3	Aktionsfelder, Ziele und Aktivitäten	8
3.0	AF-0 Strategie und Management	8
3.1	AF-1 Organisation und Kommunikation – Die Organisation für Museen, Archive und wissenschaftliche Einrichtungen öffnen	10
3.2	AF-2: Datenhaltung, Datenpflege und Standardisierung – Ein spartenübergreifendes Normdatensystem aufbauen.....	12
3.3	AF-3: Import und Datamining – Daten effizient analysieren, vernetzen, optimieren und integrieren	15
3.4	AF-4: Visualisierung und Endnutzeranwendungen – Den Zugang zum GND-Netzwerk verbessern.....	18
3.5	AF-5: Datenversorgung und Erschließungsprozesse – Die kooperative Informationsversorgung fördern	20
3.6	AF-6: Zusammenarbeit mit anderen Communities – Nutzerkreise und Anwendungskontexte erweitern.....	22

1 Perspektiven und Herausforderungen

1.1 Normdaten als Rückgrat eines maschinenlesbaren, semantischen Netzes der Kultur und Wissenschaft

Die semantische Vernetzung von Daten über Grenzen von Systemen und Domänen hinweg ist das große Anliegen der Linked-Data-Bewegung. Im Bibliothekssektor – insbesondere im deutschsprachigen Raum – hat dieser Gedanke eine lange Tradition. Kooperativ gepflegte Normdaten bilden hier schon lange Brücken zwischen Daten.

Zunächst vorwiegend als Instrument zu Arbeitersparnis und Standardisierung eingeführt, setzte man zunehmend auf die Möglichkeiten zur Verbesserung der Recherche in Bibliothekskatalogen. Im Zeitalter des Webs liegt aber das größte Potential der Normdaten darin, das Rückgrat eines maschinenlesbaren, semantischen Netzes der Kultur und Wissenschaft zu bilden: inhaltliche Verbindungen zwischen verzeichneten Kulturgütern, assoziierten Personen, Einrichtungen, Orten, Ereignissen, Projekten, Forschungsdaten, etc. auszuweisen und maschinell auswertbar zu machen, die sonst im Verborgenen bleiben würden.

Diese Möglichkeiten, inhaltliche Zusammenhänge zwischen Sammlungen verschiedener Kultureinrichtungen und Datenbanken auszuwerten, sind insbesondere für integrierende Kulturportale wie *die Deutsche Digitale Bibliothek* (DDB) und für Forschungsvorhaben (wie u.a. in einer Studie von *Knowledge Exchange* ausführlich beleuchtet¹) attraktiv. Aber auch jeder einzelnen beteiligten Kultur- und Wissenschaftseinrichtung bieten sich dadurch große Mehrwerte: Die eigenen Sammlungen, Projekte und Ergebnisse können in einen größeren Zusammenhang gestellt und es können neue thematische Zugänge angeboten werden. Für die Kulturdomäne insgesamt ergäben sich Synergien in der Datenpflege und Nutzung. Diese Einschätzung wird von Museen, Archiven, Mediatheken und Bibliotheken einhellig geteilt. Ein im Dezember 2015 von der DNB und DDB organisierter Workshop hat dies noch einmal klar bestätigt.

Um diese Entwicklung zu befördern, hat die Deutsche Nationalbibliothek gemeinsam mit ihren Partnern vor einigen Jahren die am stärksten verbreiteten deutschsprachigen Normdaten in ein System zusammengeführt: Die „Gemeinsame Normdatei“ (GND) beinhaltet einen kooperativ gepflegten Korpus an Datensätzen zu Personen, Körperschaften, Konferenzen, Geografika, Sachschlagwörtern und Werktiteln. Diese Entitäten verbinden die Datenbanken und Sammlungen des gesamten deutschsprachigen Bibliothekswesens. Darüber hinaus fungiert die GND als zentraler Anknüpfungspunkt für internationale Systeme, andere Domänen und Informationssysteme. Zunehmend wird sie daher auch von Archiven, Museen, Kultur- und Wissenschaftsportalen und im Rahmen von Forschungsprojekten genutzt.

1.2 Herausforderung: Modernisierung der GND

Mit der Ausweitung der GND-Nutzung steigen allerdings auch die Anforderungen an die Organisation, Kommunikation, Regelwerksarbeit, Werkzeuge und Infrastruktur. Das wachsende Netzwerk muss unter Ausnutzung der aktuellen technischen Möglichkeiten professioneller aufgestellt werden, um in der Flut der Daten und Anwender nicht seine größten Stärken zu verspielen: Verlässlichkeit, Transparenz, Eindeutigkeit und jene Aussagekraft, die durch einen hohen Anteil semantischer Verknüpfungen erreicht wird. Mit

¹ Maurits van der Graaf, Leo Waaijers. *Authority files: Breaking out of the library silo to become signposts for research information*. 2014. URL: http://repository.jisc.ac.uk/6224/1/Authority_files_-_Breaking_out_of_the_library_silo.pdf

einer wachsenden Zahl an nicht-standardisierten, unverbundenen redundanten Informationsknoten wäre niemandem gedient. Auch besteht zu Recht die Erwartung, dass endlich die Nutzungsmöglichkeiten der Datei in Form von attraktiven Fachanwendungen ausgeschöpft werden.

Die GND-Partner haben sich in einer aktuell geführten Grundsatzdebatte zur Strategie der GND einstimmig für eine konsequente Ausweitung und Modernisierung ausgesprochen. Wie groß der Bedarf hierfür ist, hat zuletzt ein von DNB und DDB gemeinsam organisierter Workshop mit Museen, Archiven und Mediatheken klar bestätigt. Dasselbe Echo ergab sich auch in der Zusammenarbeit mit wissenschaftlichen Einrichtungen und Netzwerken wie beispielsweise dem [Netzwerk Technikfolgenabschätzung](#). Für eine Öffnung der Redaktion im großen Stil sind allerdings noch organisatorische und technische Voraussetzungen zu schaffen. Weitere Herausforderungen bestehen in der Stärkung der Fähigkeiten zur semantischen Vernetzung, zur Integration in weltweite Datenerzeugungsprozesse, sowie zur Ausnutzung maschineller Möglichkeiten (Datamining). Zusammen mit den GND-Partnern plant die DNB diese in den nächsten Jahren zu erarbeiten. Bis 2021 soll die GND schrittweise modernisiert, re-organisiert und um neue Partner aus dem Museums- und Archivbereich erweitert werden.

Während es weiterhin möglich ist, den Betrieb der GND im bestehenden Verbund durch die Eigenanteile der DNB und der anderen größeren GND-Partner zu schultern, kann die nötige Weiterentwicklung nur über gut ausgestattete Anschubprojekte dargestellt werden. Zwar investiert die DNB seit Jahren als technischer Dienstleister und Koordinator in die Modernisierung der Organisation und Infrastruktur, und sie unterstützt interessierte Kultureinrichtungen beim Einstieg in die GND – unter anderem auch über ihr Engagement in der Deutschen Digitalen Bibliothek. Und auch die größten GND-Partner wie beispielsweise die Bibliotheksverbände unterstützen diese Entwicklung mit großem Engagement. Die von den großen Partnern für die allgemeinen Bedarfe des GND-Netzwerks abzweigbaren Eigenanteile reichen aber nicht, um die zunehmenden Anforderungen der vielen unterschiedlichen Parteien an die zentralen Infrastrukturalien und an den organisatorischen Rahmen abzudecken. Daher wirbt die DNB gemeinsam mit ihren Partnern um eine Förderung einer Modernisierungsinitiative durch die Deutsche Forschungsgemeinschaft. In diesem Rahmen sollen auch Finanzierungsmodelle entwickelt werden, wie künftige Entwicklungssprünge von den Partnern des GND-Netzwerks weitestgehend selbstständig getragen werden können.

Im Folgenden werden relevante Vorarbeiten erläutert und im darauffolgenden Abschnitt die strategischen Ziele und deren mögliche Umsetzung skizziert.

2 Relevante Vorarbeiten

2.1 Generelles

Seit ihrem Bestehen wurde von den Partnern und der Community regelmäßig in Projekte und Angebote rund um die GND investiert. In zahlreichen Kooperationsprojekten wie *IN2N*, *ORCID DE*, *CONTENTUS*, *CULTUREGRAPH*, *EntityFacts* und dem *GND-Webformular* zur Personenerfassung konnten fachliche und technische Grundlagen (Schnittstellen, Werkzeuge, Geschäftsprozesse) für die Modernisierung der GND gelegt werden. All diese Projekte haben zum anhaltenden Erfolg der GND beigetragen und Grundlagen für die in diesem Papier dargelegte Strategie gelegt. Auf die einzelnen relevanten Vorarbeiten wird an geeigneter Stelle im detaillierten Entwicklungsplan (Abschnitt 3) hingewiesen.

Der allgemeine Nachteil des bisherigen Vorgehens sind die teils zu geringe Nachhaltigkeit über den Projektzeitraum hinaus und die mangelnde allgemeine Nutzbarkeit. Gerade für grundlegende Modernisierungsvorhaben fehlten die technischen und organisatorischen Rahmenbedingungen. Im Folgenden werden daher die entscheidenden Vorarbeiten mit Blick auf die organisatorische und technische Nachhaltigkeit vorgestellt.

2.2 Erneuerung der Organisation: die GND-Kooperative

2.2.1 Motivation und Ausgangslage

Die GND wird von der Deutschen Nationalbibliothek, allen deutschsprachigen Bibliotheksverbänden mit den angeschlossenen Bibliotheken, der Zeitschriftendatenbank und zahlreichen weiteren Einrichtungen gemeinschaftlich geführt (im Folgenden kurz *GND-Partner* genannt). Die Mitarbeit in der GND erfolgt entweder über Verbände oder nach direkter Absprache mit der Deutschen Nationalbibliothek.

Die aktuelle Kooperation ist historisch gewachsen und funktioniert heute aufgrund der langen gemeinsamen Zusammenarbeit und der zahlreichen inhaltlichen Berührungspunkte der verschiedenen Partner. Abseits von Festlegungen in Protokollen des STA existieren keine formalen Vereinbarungen zur GND. Spätestens die gewünschten Erweiterungen und Entwicklungen machen allerdings eine verbindliche Klärung der gegenseitigen Erwartungen, der Rechte und der Pflichten rund um die GND nötig. Daher haben die Partner beschlossen, der GND einen neuen formalen Rahmen zu geben². In dieser Kooperationsvereinbarung zur *GND-Kooperative* werden die Grundprinzipien, die Voraussetzungen einer Mitarbeit, die Lastenverteilung und Finanzierung sowie die Organisation geklärt. Die Gründung der GND-Kooperative ist für Mitte 2017 geplant.

² siehe Protokoll der 27. Sitzung des Standardisierungsausschusses vom 24. Juni 2015 (<http://www.dnb.de/afsprotokolle>, TOP 5: „Strategische Überlegungen zur GND“)

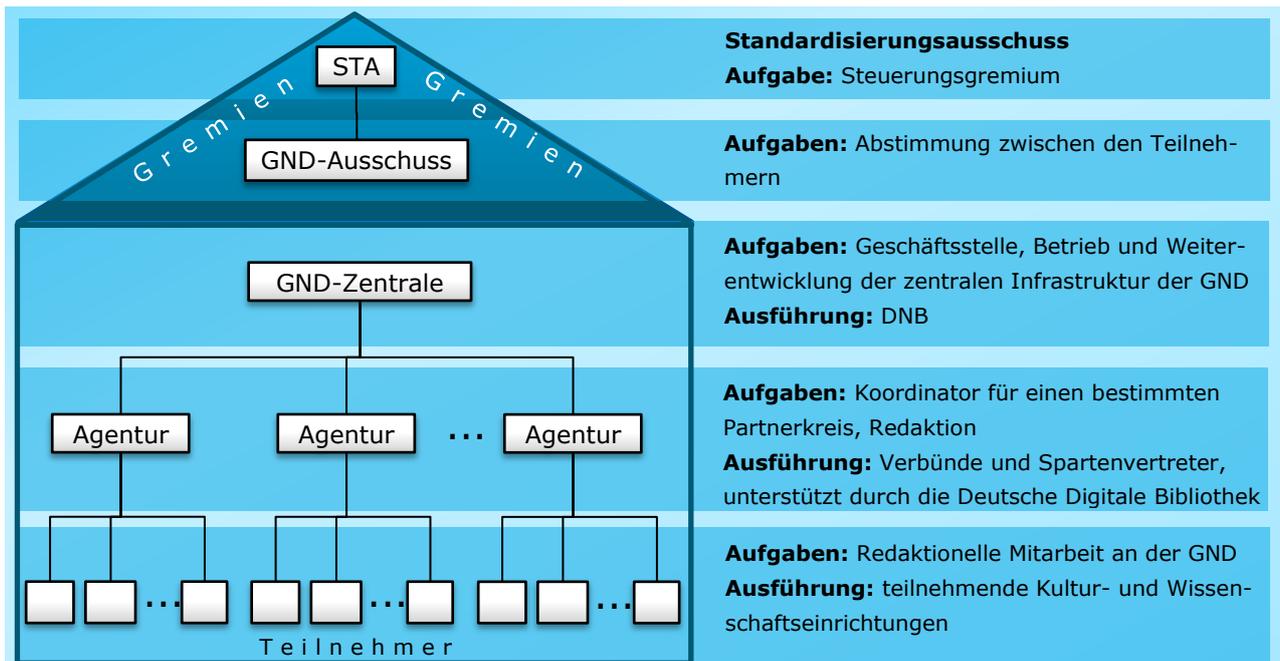


Abb. 1: Organisation der GND-Kooperative

Um die etablierten, im Grundsatz erfolgreichen Strukturen zu stützen, wurde ein mehrstufiges Organisationsmodell bestehend aus der GND-Zentrale, GND-Agenturen und Teilnehmern gewählt (vgl. Abb. 1). Ziel dieser Struktur ist es, die betriebliche Last zu verteilen und den speziellen Anforderungen der verschiedenen Interessengruppen besser gerecht zu werden. Die Arbeit wird wie in der Vergangenheit durch den STA gesteuert. Ein neu eingerichtetes dem STA unterstelltes Gremium, der GND-Ausschuss, ist für die Abstimmung zwischen den Partnern und für die Erarbeitung der Strategie zuständig. In diesem Gremium sind unter anderem die Zentrale und alle Agenturen vertreten. Diese Regelungen sollen allerdings nur einen groben formalen Rahmen für die weiteren Entwicklungen bilden, um die Integration nicht-bibliothekarischer Partner in diese Struktur vorzubereiten – beispielsweise durch den Aufbau neuer Agenturen und durch die Einbindung von Spartenvertretern in die genannten Gremien, sowie durch die Weiterentwicklung des Kooperationsmodells.

Damit eine solche Interessensvertretung ausgefüllt werden kann, müssen sich die Einrichtungen innerhalb ihrer Community organisieren. Hier kommt der Deutschen Digitalen Bibliothek (DDB) eine besondere Rolle zu: Fachliche und politische Kommunikationsstrukturen zwischen den verschiedenen Einrichtungen aus Kultur und Wissenschaft (KWEs) sowie Abstimmungsprozesse zu gemeinsamen Regeln, Formaten und Verfahren sind hier bereits etabliert und können von der GND-Kooperative nachgenutzt werden. Auch ist gerade für die DDB ein flächendeckender Einsatz von Normdaten das ideale Instrument, um die Zusammenhänge zwischen den höchst unterschiedlichen Sammlungen sichtbar und semantisch durchsuchbar zu machen. Die künftige spartenübergreifende Organisationsstruktur der GND und die Organisationsstrukturen des DDB-Netzwerks sollten daher aufeinander abgestimmt sein.

2.2.2 Herausforderungen

Für die bisher aktiven Teilnehmer kann weitestgehend auf die bestehenden Verbundstrukturen zurückgegriffen werden. Die Verbundzentralen agieren als Agentur für die Verbundteilnehmer. Anders stellt sich die Situation für neu hinzukommende Kultur- und Wissenschaftseinrichtungen dar, die bislang nicht über die Bibliotheksverbände oder über andere im Standardisierungsausschuss vertretenen Einrichtungen angeschlossen sind. Beispielsweise müssen für die meisten Museen und Archive noch Einrichtungen zu GND-Agenturen aufgebaut werden. Hierzu müssen auch die beiderseitigen Rahmenbedingungen geklärt und die speziellen Anforderungen der neu hinzukommenden Fachcommunities berücksichtigt werden.

Der operative Betrieb ist durch die etablierte Arbeitsteilung unter den bestehenden GND-Partnern weiterhin gesichert. Zu klären ist allerdings die langfristige Finanzierung der Weiterentwicklung der GND. Dieser Aspekt wurde zur Gründung der Kooperative aufgrund seiner Komplexität und der noch aufzubauenden Agenturen zurückgestellt. Hier müssen in den nächsten Jahren verschiedene Modelle ausgearbeitet und diskutiert werden.

2.3 CULTUREGRAPH – Infrastrukturelles „Ökosystem“ für die GND

2.3.1 Motivation und Ausgangslage

Die infrastrukturelle Weiterentwicklung der GND soll auf Grundlage der von der DNB betriebenen offenen Dienstplattform CULTUREGRAPH.org erfolgen. CULTUREGRAPH ist als Ökosystem bestehend aus vorhandenen bzw. neu zu entwickelnden sich ergänzenden Werkzeugen, Infrastrukturelementen und gemeinsamen zentralen Datenpools für eine kooperative Erschließung konzipiert. Die größten Vorteile des Ansatzes sind:

- In Projekten entwickelte Dienste können, wenn sie für die Plattform entwickelt wurden, leicht integriert und betrieben werden und sind mit den bestehenden Diensten kompatibel bzw. können sich zu komplexen Geschäftsgängen ergänzen
- Dienste können auf eine gemeinsame umfassende Datenbasis zugreifen und diese erweitern
- Normdaten, Titeldaten, Inhaltsdaten und andere Datenquellen sind nur konzeptionell, aber nicht technisch getrennte Datenbanken. Dies ist insbesondere für Datamingverfahren, Statistiken, Vorschlagssysteme und andere „smarte“ Verfahren rund um die GND-Pflege und Anreicherung wichtig.

In der Konzeption und Entwicklung ist der CULTUREGRAPH-Funktionsbereich, der sich mit Dataming beschäftigt, mit Abstand am weitesten fortgeschritten. Die DNB nutzte die Dataming-Module gemeinsam mit den Verbundsystemen, um die Bestände miteinander abzugleichen, sie in verschiedener Hinsicht statistisch auszuwerten und untereinander sowie mit externen Datenquellen zu vernetzen. Die Datenbasis umfasst annähernd alle Katalogdaten des deutschsprachigen Bibliothekswesens inklusive der Gemeinsamen Normdatei. Folgende Hauptanwendungsfälle wurden in Pilotanwendungen adressiert:

- Das Zusammenführen sich überschneidender Titeldatenbestände, d.h. die bibliografischen Nachweise zu denselben Publikationen: Diese werden in Bündeln zusammengeführt und mit einem gemeinsamen Identifier versehen
- Darauf aufbauend: die Identifikation von Werken als Kandidaten für neue Werkdatensätze
- Die statistische Auswertung der Nutzung von GND-Sätzen (Welche Normdatensätze werden wie häufig genutzt) als Grundlage für Bereinigungen und die Priorisierung von redaktionellen Eingriffen
- Der Aufbau eines leicht in Webanwendungen integrierbaren zentralen Einstiegspunkts für die GND („EntityFacts“). Ziel des Dienstes ist es, zu jedem GND-Element ein digitales „Faktenblatt“ anzubieten, das alle aggregierten Informationen und semantischen Querverbindungen zur jeweiligen Entität zusammenträgt. Unter anderem nutzt das Portal der Deutschen Digitalen Bibliothek diesen Dienst, um die GND in die eigene Recherche zu integrieren.

Weitere Nutzungsszenarien und Funktionalitäten (wie beispielsweise Datenmanagement, Nutzermanagement, Recherche und Visualisierung) sollten folgen, wurden aber mit Blick auf den Start des [CIB-Projektes](#) zunächst zurückgestellt. Für die DNB ist die Plattform von großer Bedeutung für ihre aktuellen strategischen Planungen (vgl. [Strategische Prioritäten 2017-2020](#)). Daher wird sie in den nächsten Jahren verstärkt in den Ausbau und die Ergänzung der erfolgreich getesteten Konzepte und Infrastrukturkomponenten investieren.

2.3.2 Herausforderungen

Die folgende Abbildung zeigt die für den Ausbau der GND benötigten Funktionsmodule. Zwar wird die DNB die wichtigsten Basiskomponenten, Programm- und Schnittstellenbibliotheken in Eigeninitiative bereistellen (dies gilt insbesondere für das Datamining, Teile der Datenhaltung, der Datenpflege, des Datasupply Frameworks und den Bereich „Statistic & Mappings“). Für die GND werden aber zahlreiche Anpassungen und dedizierte Entwicklungen nötig sein, die nur in Kooperationsprojekten geleistet werden können.

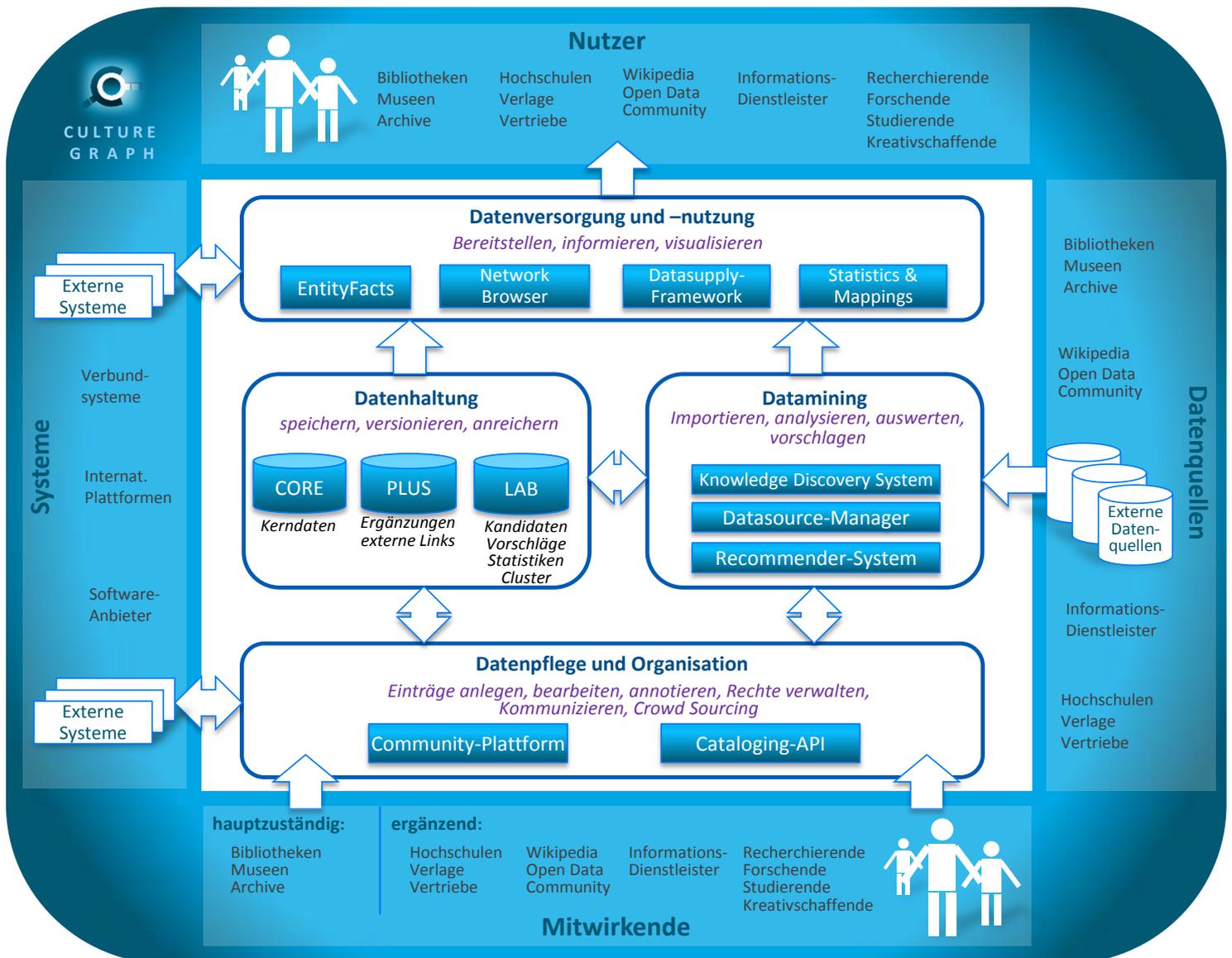


Abb. 2: Funktionale Zielarchitektur der GND - verwirklicht als Module der CULTUREGRAPH-Plattform

3 Aktionsfelder, Ziele und Aktivitäten

Nachfolgend werden die anvisierten Ziele und Aktivitäten gegliedert in Aktionsfelder aufgeführt:

- AF-0 Strategie und Management
- AF-1 Organisation und Kommunikation
- AF-2 Datenhaltung, Datenpflege und Standardisierung
- AF-3 Import und Datamining
- AF-4 Visualisierung und Endnutzeranwendungen
- AF-5 Datenversorgung und Erschließungsprozesse
- AF-6 Zusammenarbeit mit anderen Communities

Dieses Kapitel ist nicht als detaillierter Projektplan zu verstehen, sondern als grober Rahmen für die Durchführung sich ergänzender Aktivitäten. Diese werden Teils im Rahmen von Projektstrukturen umgesetzt, teils aber auch als Einzelmaßnahmen. Wo besondere Wechselbeziehungen zwischen Aktivitäten bestehen, ist dies explizit im Text vermerkt. Den Aktionsfeldern werden die Stakeholder und betroffenen Funktionsmodule (vgl. Abb. 2) zugeordnet. Eine grobe Planung der Aktivitäten und Projekte findet sich am Ende – hervorgehoben sind hierbei jene Teile, für die ein grundsätzlicher Förderbedarf besteht (selbstverständlich sind auch in diesen Bereichen erhebliche Eigenanteile vorgesehen). Nicht alle planungsrelevanten Faktoren können heute bereits abgesehen werden. Teilweise bauen die Pakete aufeinander auf, und in vielen Fällen müssen noch geeignete technische und fachliche Partner gefunden werden. Dies muss im Rahmen der Ausarbeitung der dedizierten Projektanträge erfolgen und ist nicht Gegenstand dieses Dokuments. Ganz generell müssen die Priorisierung und die Ausgestaltung einzelner Aspekte mit Blick auf die im Laufe des Umsetzungsprozesses gewonnenen Kenntnisse und Erfahrungen turnusmäßig nachjustiert werden.

3.0 AF-0 Strategie und Management

3.0.1 Inhalt des Aktionsfelds

Dieses Aktionsfeld dient der Verortung aller Aktivitäten zur Strategieentwicklung und des Managements von Entwicklungen. Hierzu zählt das Management und Koordination der im Folgenden aufgeführten Aktivitäten und Projekte, u.a.:

- Ausarbeiten und Festlegen der strategischen Leitlinien und Ziele der GND
- Gesamtkoordination der Projekte und Aktivitäten
- Turnusmäßige Aktualisierung des Entwicklungsprogramms
- Ausarbeiten von Projektanträgen
- Steuerung / Begleiten von Kooperationsprojekten (s.u.)

3.0.2 Übersicht über die zu koordinierenden Projektvorhaben

In der folgenden Tabelle werden die aktuell vorgesehenen Entwicklungsprojekte aufgelistet, die im Rahmen des Aktionsfelds vorbereitet, koordiniert und begleitet werden müssen. Projekte, für die eine Förderung durch die DFG beantragt werden soll, sind farblich unterlegt. Die Projekte werden in den Arbeitsplänen der jeweils zugehörigen fachlichen Aktionsfelder wieder aufgegriffen. Einige der Projekte erstrecken sich inhaltlich über mehrere Aktionsfelder.

Projekte (bei absehbarem Förderbedarf farblich markiert)									
Kürzel	Projekte	AF	Konsortium	Status	17	18	19	20	21
ORC	„ORCID DE“ (läuft bereits) - Verzahnung der GND mit dem ORCID-System	AF6	- Helmholtz Open Science Koordinationsbüro - am Deutschen GeoForschungsZentrum GFZ - Universitätsbibliothek Bielefeld - DNB	läuft					
ARAC1	„Culturegraph Arachne“ - Planung einer Umgebung für Datamining und Vernetzung von Daten	AF3	- DNB	Beantragt					
DDUP	„D-DUP“: Projekt zur großflächigen Bereinigung von Dubletten und Inkonsistenzen	AF3	- GND-Ausschuss (Planung Steuerung) - Alle Partner der GND-Kooperative (Durchführung)	noch auszuarbeiten					
TEMA	„TeMa“: Verwirklichung von Recommenderfunktionen für GND-Sachbegriffe (zunächst für DNB-interne Erschließungsprozesse)	AF3	- DNB	Genehmigt					
4C-1	„GND4C“ - Öffnung der GND für Kulturdaten 1. Projektphase: Konzeption, Prototyping, Pilotbetrieb	AF1-4	- DNB (Projektleitung) - DDB - 2 Spartenvertreter als künftige GND-Agenturen [Landesarchiv Baden-Württemberg, Bildarchiv Foto-Marburg] - 1 bestehende GND-Agentur [Bibliotheksservice-Zentrum Baden-Württemberg] - 1 technisch-konzeptioneller Partner im Bereich kollaborative Datenpflege [digiCULT e.G.]	Antrag in Vorbereitung					
ARAC2	„Culturegraph Arachne 2“ - Umsetzung einer Umgebung für Datamining und Vernetzung von Daten		- DNB (Projektleitung, betriebliche Anforderungen und technische Integration) - Ggf. technisch konzeptioneller Partner - GND-Ausschuss (Evaluation, Test)	noch auszuarbeiten					
CRAT	„Culturegraph Cratera“ – Verwirklichung einer spartenübergreifenden Kollaborationsplattform zum Registrieren, Erfassen und Teilen von Kulturdaten	AF2	- DNB (Projektleitung, betriebliche Anforderungen und technische Integration) - 1 fachlich-konzeptioneller Partner - 1 technischer Partner - GND-Agenturen (für den Aufbau spartenspezifischer Sichten, als Anforderungsgeber und für Tests) - Ggf. Wikimedia	noch auszuarbeiten					
SYNC	„Culturegraph InSync“ – Entwicklung von und Werkzeugen und Verfahren zur Unterstützung kooperativer Erschließungsprozesse	AF5	- DNB / GND-Agenturen - Ggf. Anbieter von Bibliothekssystemen	noch auszuarbeiten					
EXP	„Culturegraph EntityExplorer“ - Aufbau eines browser-basierten Endnutzerdienstes für die Suche in der GND als zentraler Einstiegspunkt ins GND-Netzwerk	AF4	- DNB / GND-Agenturen - Technologiepartner / Partner im Bereich Visualisierung und Endnutzeroberflächen	noch auszuarbeiten					
4C-2	„GND4C“ - Öffnung der GND für Kulturdaten 2. Projektphase: Verstetigung, Überführung der Ergebnisse aus der 1. Phase in Produktion, Ausweitung	AF1-4	- Siehe 4C-1	noch auszuarbeiten					
HUG	„Hugin“ – Aufbau von Endnutzerdiensten für eine Normdatenzentrische Recherche in kulturellen Sammlungen	AF4	- GND-Kooperative, DDB (Anforderungsgeber, Test und Referenznutzer) - Forschungspartner - ggf. technisch-konzeptioneller Partner für das Backend und Indexierungsinfrastruktur	noch auszuarbeiten					

3.1 AF-1 Organisation und Kommunikation – Die Organisation für Museen, Archive und wissenschaftliche Einrichtungen öffnen

3.1.1 Inhalt des Aktionsfelds

AF-1 konzentriert sich auf allgemeine organisatorische und fachliche Aspekte. Es soll unter anderem die Grundlage für den reibungslosen Betrieb der GND und die fachliche Zusammenarbeit zwischen beteiligten Institutionen und den Sparten legen. Die zentrale Gremium für diese Arbeit ist der GND-Ausschuss und der ihm übergeordnete Standardisierungsausschuss (STA). Zu den regulären Aufgaben des GND-Ausschusses in AF-1 zählen:

- Organisation der Zusammenarbeit und Kommunikation zwischen den Partnern
- Organisation der Zusammenarbeit mit anderen Gremien im Bereich der Standardisierung (z.B. Fachgruppen des STA)
- Organisation des Betriebs und der Weiterentwicklung der GND
- Organisation und Durchführung Community-Arbeit und allgemeiner Öffentlichkeitsarbeit

3.1.2 Entwicklungsziele

Die bestehenden GND-Partner sind mit der konstituierenden Sitzung des GND-Ausschusses im Januar 2017 den ersten praktischen Schritt in Richtung einer neuen, in die Zukunft gerichteten Organisationsstruktur gegangen (vgl. Abschnitt 2.2). Diese Umorganisation muss nun weiter mit Leben gefüllt werden. Darauf aufbauend gilt es, die Aufnahme neuer Partner aus nicht-bibliothekarischen Bereichen vorzubereiten. Eine Schlüsselrolle kommt hier dem Aufbau von zentralen Organisationseinheiten für nicht-bibliothekarische Kultur- und Wissenschaftseinrichtungen (KWEs) wie beispielsweise Museen, Archive, Mediatheken, Universitäten und Hochschulen zu: Diese sogenannten „GND-Agenturen“ sollen als erste Ansprechpartner für die neuen Partner aus den jeweiligen Sparten dienen, diese in den Gremien und Arbeitsgruppen der GND vertreten und spartenspezifische Fachkonzepte erarbeiten. Auch im Bereich der Kommunikation sind die Angebote und Strukturen noch auf die Bibliothekswelt ausgerichtet. Hier gilt es parallel zur fachlichen und organisatorischen Öffnung, neue Kommunikationskanäle in Richtung der anderen Sparten und der Wissenschaft zu etablieren.

- **Ziel 1.1:** Praktische Umsetzung der neuen Organisationstruktur
- **Ziel 1.2:** Schrittweise Erweiterung der Organisation für die Mitwirkung durch Museen und Archive (beinhaltet eine tragfähige und nachhaltige Regelung zur Arbeitsteilung und Finanzierung).
- **Ziel 1.3:** Intensivierung der Kommunikation mit den verschiedenen Interessengruppen über verschiedene Kommunikationskanäle
- **Ziel 1.5:** Erhöhung der Sichtbarkeit der GND, der GND-Kooperative und seiner Ziele durch den Aufbau eines ansprechenden Web-Auftritts, allgemeine Öffentlichkeitsarbeit und Publikationstätigkeit

3.1.3 Wichtigste Stakeholder

Mitwirkende		Nutzer		Systeme / Datenquellen	
1.	KWEs (künftige Agenturen und Teilnehmer, GND-Ausschuss)	2.	KWEs	4.	Anbieter von Systemen für Museen, Archive und Mediatheken
		3.	Wikipedia, Open Data Community Recherchierende, Forschende, Studierende, Kreativschaffende Informationsdienstleister		
Gremien / Arbeitsgruppen					
Standardisierungsausschuss (STA)					
Ausschuss der GND (AGND)					

Fachgruppen des STA zu Regelwerks- und Formatfragen
DINI AG KIM
Konferenz der Leiterinnen und Leiter der Archivverwaltungen des Bundes und der Länder (KLA)
Diverse DDB-Gremien und Arbeitsgruppen

3.1.4 Betroffene Funktionsbereiche

--

3.1.5 Zusammenhänge zu den Entwicklungen in anderen Aktionsfeldern

Das AF legt die organisatorische Grundlage für die Entwicklungen in allen anderen Aktionsfeldern. Es bietet den Rahmen für die nötigen Abstimmungsprozesse.

3.1.6 Relevante Vorarbeiten

Siehe Abschnitt 2.2 *Erneuerung der Organisation: die GND-Kooperative*

3.1.7 Entwicklungsaktivitäten

Herzstück dieses Aktionsfeldes bilden die beiden Projektphasen des Projekts *GND4C* (Kürzel: *4C-1* und *4C-2*). Es beinhaltet zwei Arbeitspakete, die dem Aktionsfeld zuzuordnen sind. Für das Projekt ist ein hoher Anteil von Eigenanteilen vorgesehen. Hinzu kommen die in Eigenanteil erbrachten Leistungen der beteiligten Gremien.

#	Entwicklungsaktivitäten (bei absehbarem Förderbedarf farblich markiert)	17	18	19	20	21
E.1.1	Abschluss einer Kooperationsvereinbarung mit allen bestehenden GND-Partnern					
E.1.2	Festlegen eines verbindlichen mittelfristigen Arbeitsplans für die Partner der GND-Kooperative mit konkreten Aufgabenzuweisungen					
E.1.3	Entwickeln eines Kommunikationskonzepts					
E.1.4	Aufbau eines neuen Webauftritts					
E.1.4 (4C-1.AP1)	Entwickeln eines Kooperationsmodells, das Kultur- und Wissenschaftseinrichtungen aller Art nachhaltig zur vollwertigen Teilnahme an der GND ertüchtigt und eine tragfähige Arbeitsteilung und Finanzierung der Kooperation sicherstellt.					
E.1.5 (4C-1.AP1)	Testweise Operationalisierung der neuen Konzepte durch den Pilotbetrieb erster GND-Agenturen für Museen und Archive.					
E.1.6 (4C-1.AP4)	Entwickeln und Durchführen von Dialogformaten zur Erhebung, Koordination und Dokumentation von community-spezifischen Anforderungen und zum Abstimmen gemeinsamer übergreifender Standards					
E.1.7 (4C-1.AP4)	Aufbau einer web-basierten Dialogplattform, die die Sichtbarkeit und Wirkung der Entwicklungen rund um die GND fördern soll					
E.1.8	Erweitern der Kooperationsvereinbarung					
E.1.X	Pilotbetrieb der neuen Agenturen / Übergang in den Dauerbetrieb					
E.1.X (4C-2)	Aufnahme weiterer Agenturen und Communities aus dem Bereich Kultur- und Wissenschaft					

3.2 AF-2: Datenhaltung, Datenpflege und Standardisierung – Ein spartenübergreifendes Normdatensystem aufbauen

3.2.1 Inhalt des Aktionsfelds

Das Aktionsfeld beinhaltet alle Aufgaben und Aktivitäten rund um die Datenhaltung, Datenpflege und Standardisierung. Es umfasst dabei sowohl fachliche (z.B. die Pflege des Datenformats und der Erfassungsregeln) als auch infrastrukturelle Aspekte (z.B. die technische Umgebung zur Datenhaltung –und Pflege).

Zu den regulären Aufgaben des Arbeitsfelds zählen:

- Anpassungen von Anwendungsrichtlinien, Erläuterungen, Erfassungsleitfaden, Erfassungshilfen (GND-Ausschuss in Zusammenarbeit mit der STA-Fachgruppe Erschließung).
- Pflege des Datenformats (GND-Ausschuss in Zusammenarbeit mit der STA-Fachgruppe Datenformate).

3.2.2 Entwicklungsziele des Aktionsfelds

Die Datenstrukturen, Geschäftsprozesse, Regeln und Schnittstellen sind im Wesentlichen auf die Bibliothekswelt ausgelegt. Um die GND universell einsetzbar und offen für Beiträge aus nicht-bibliothekarischen Bereichen zu machen und gleichzeitig die Nutzbarkeit für bibliothekarische Kernanwendungen zu erhalten, müssen neue Regelungen und Konzepte für das Normdatenangebot implementiert werden. Dabei sind einerseits die Standards und Praxisempfehlungen aus den verschiedenen Sparten zu berücksichtigen, andererseits müssen Anforderungen, die sich aus den Integrationsbestrebungen ergeben, zurück in die spartenspezifischen Standardisierungsgremien gespielt werden. Ein weiterer ungeklärter Bereich ist der Umgang mit automatisierter Datenerzeugung: Bei GND-Partnern kommen zunehmend auch maschinelle Verfahren zum Einsatz. Die GND muss sich zu dieser Entwicklung positionieren und klären, auf welche Weise sie automatisierte Prozesse unterstützen kann, ohne dabei die Verlässlichkeit und Qualität der GND zu verwässern.

Wichtigstes Konzept, um den unterschiedlichen Interessen gerecht zu werden, ist die Einführung von Anwendergruppen und gruppenspezifischer Bereiche und Sichten: Einige Datenbereiche sind Teil eines von allen Anwendern spartenübergreifend gepflegten und redaktionell bearbeiteten Datensets (GND-CORE), andere Teile werden von den verschiedenen Anwendergruppen gegenseitig als optionale Ergänzungen begriffen, die genutzt werden können, aber redaktionell unabhängig sind (GND-PLUS). Weiterhin müssen ausgefeilte Bearbeitungsrichtlinien und Workflows geschaffen werden.

Ein weiteres benötigtes Feature ist das detaillierte Tracking von Änderungen und Provenienzen (eine unabdingbare Voraussetzung für eine weitreichende Öffnung der Datenpflege, für die Integration automatischer Verfahren und Crowdsourcing-Projekte) sowie der Ausbau der Cataloging-API um weitere web-basierte Maschinenschnittstellen, welche die hinzugekommenen Features unterstützen.

- **Ziel 2.1** Strategie- und Konzeptentwicklung zur Unterstützung automatischer Erschließungsverfahren
- **Ziel 2.2:** Einigung auf einen gemeinsamen, spartenübergreifenden Mindeststandard (GND-CORE) für GND-Daten. Dieser soll sowohl für manuelle Eintragungen als auch für Massenimporte (siehe auch AF-3) verbindlich sein. Teilziel ist dabei auch die Klärung des Umgangs mit Datensätzen, die diesen Anforderungen nicht genügen (vgl. Abb. 2, Datenpool „LAB“)
- **Ziel 2.3:** Umsetzung einer modularen Datenhaltung mit gruppenspezifischen Sichten (GND-PLUS):
 - Aus- bzw. Aufbau der Infrastruktur
 - Migration des Datenbestands
- **Ziel 2.4:** Evaluation von Konzepten zur Datenmodellierung, von Arbeitsprozessen und von Werkzeugen aus der Softwareentwicklung als Lösungsansatz für die kollaborative und institutionsübergreifende Pflege von Normdaten

- **Ziel 2.5:** Verbesserung des Datensatzmanagements:
 - detailliertes Tracking von Änderungen und Provenienzen
 - Änderungsvorschläge können bei fehlender Berechtigung im Datensatz hinterlegt und eingereicht werden
 - Kommentierungsfunktion

3.2.3 Stakeholder

Mitwirkende		Nutzer		Systeme / Datenquellen	
1.	KWEs (GND-Teilnehmer und Interessierte)	4.	KWEs	3.	Anbieter von Systemen für Museen, Archive und Mediatheken
2.	Wikipedia, Open Data Community	5.	Wikipedia, Open Data Community		
	Verlage, Vertriebe		Verlage, Vertriebe		
	Recherchierende, Forschende, Studierende, Kreativschaffende		Recherchierende, Forschende, Studierende, Kreativschaffende		
			Informationsdienstleister		

3.2.4 Adressierte Funktionsbereiche (vgl. Abbildung 2: Funktionale Zielarchitektur)

- Datenhaltung / Datenpflege und Organisation:
 - Eine **Community-Plattform** und eine passende Erweiterung in der Datenhaltungsschicht (**GND PLUS**) sollen die GND u.a. um die nötigen Funktionalitäten erweitern. Wichtige Features:
 - Modulare Datenstruktur: Gemeinsamer Kern / community-spezifische Erweiterungen
 - Datenmanagement:
 - Tracking von Provenienz und Änderungen
 - Zusammenführen unterschiedlicher Zustände eines Datensatzes
 - Flexibles Rechtemanagement
 - Editieren von Normdaten über intuitive Webformulare
 - Werkzeuge zum Aufbau von Crowd-Sourcing-Projekten (→ AF-6)
 - Aufbauend auf den Ergebnissen des Projekts IN2N soll eine zeitgemäße **Cataloging-API** für die Normdatenpflege entwickelt werden, die sich für eine leichte Integration in Fremdsysteme und Geschäftsgänge aller Art eignet. Erweiterung der Cataloging-API um neue web-basierte Schnittstellen

3.2.5 Zusammenhänge zu den Entwicklungen in anderen Aktionsfeldern

- Die aus Formatänderungen resultierenden Anpassungen am Datenbestand und Importverfahren sind Teil von AF-3.
- Weiterhin werden in AF-2 wichtige Grundlagen für eine Beteiligung weiterer Communities (AF-6) gelegt: Ohne Fortschritte im Datenmanagement (Tracking von Provenienzen, flexible Rechtevergabe, modulare Erweiterungen im Datenmodell, etc.) wird eine breite Beteiligung Dritter nicht möglich sein.

3.2.6 Relevante Vorarbeiten

Projekt	Beschreibung	Projektbeteiligte
IN2N	Prototypische Entwicklung und Erprobung von Werkzeugen und Verfahren für eine domänenübergreifende Nachnutzung und Pflege der Gemeinsamen Normdatei (GND) am Beispiel des DIF	DNB Deutsches Filminstitut (DIF)
BEACON	Einfaches und sehr beliebtes Dateiformat mit dem Links auf Webseiten angegeben werden können, die Inhalte zu Normdaten bieten.	Wikipedia

3.2.7 Entwicklungsaktivitäten

#	Entwicklungsaktivitäten (bei absehbarem Förderbedarf farblich markiert)	17	18	19	20	21
E.2.1	Strategie- und Konzeptentwicklung zur Unterstützung automatischer Erschließungsverfahren					
E.2.2	Modernisierung und Aktualisierung des zentralen Datenpools von Katalog- und Sammlungsdaten aus dem GND-Netzwerk (Bibliotheksdaten)					

E.2.3 (4C-1.AP2)	<p>CORE vs. PLUS - Spezifikation eines gemeinsamen Mindeststandards und von Erweiterungsmodulen für das Datenmodell und die Erfassungsregeln</p> <ul style="list-style-type: none"> - Entwicklung eines abstrakten Basisdatenmodells für die GND, das den Aufbau fachlicher Sichten unterstützt. - Darauf basierende Spezifikation eines GND-CORE-Datenmodells und fachspezifischer Erweiterungen Sichten unter Berücksichtigung bestehender Standards. Dies soll anhand konkreter Fallbeispiele und unter anderem über die Analyse von Datensammlungen unterschiedlicher Kultureinrichtungen erfolgen. Dabei werden auch die bestehenden Anwendungsrichtlinien evaluiert. - Evaluation der entwickelten Modelle unter Zusammenarbeit mit Vertretern und Gremien der verschiedenen Communities mit Blick auf Funktionalität und Umsetzbarkeit. 	
E.2.4 (CRAT)	<p>Verwirklichung einer spartenübergreifenden Kollaborationsplattform zum Registrieren, Erfassen und Teilen von Kulturdaten:</p> <ul style="list-style-type: none"> - Evaluation von Arbeitsprozessen und Werkzeugen aus der Softwareentwicklung als Lösungsansatz für die kollaborative und institutionsübergreifende Pflege von Normdaten - Prototypischer Aufbau des CORE-Datenbestands - Prototypische technische Umsetzung der spartenspezifischen Sichten/Erweiterungen (Bibliotheken, Museen, Archive) - Integration in die bestehende Infrastruktur 	
E.2.X	<p>Überführung der Projektergebnisse in den operativen Betrieb:</p> <ul style="list-style-type: none"> - Überführen der Infrastruktur in Produktionsreife - Implementierung des Betriebskonzepts - Migration der Produktiv-Daten - Umstellen aller Produktionsprozesse - Erstellen von Dokumentation / Durchführung von Schulungen 	
E.2.X (4C-2)	<p>Erweiterung des zentralen Pools von Katalog- und Sammlungsdaten um Daten aus dem nicht-bibliothekarischen Bereich</p>	

3.3 AF-3: Import und Datamining – Daten effizient analysieren, vernetzen, optimieren und integrieren

3.3.1 Inhalt des Aktionsfelds

AF-3 bündelt alle Aktivitäten zur Analyse, Vernetzung, Normalisierung und Optimierung von von GND-Daten und potentiell zu integrierender Fremddatenquellen (Funktionsbereich Datamining). Das Arbeitsfeld umfasst

- die Erarbeitung fachlicher Vorgaben - z.B. Kriterien für den Abgleich von Datensätzen (Matching),
- die Konzeption und Implementierung von Werkzeugen – z.B. die Umsetzung eines Recommender-Systems zur Unterstützung der Datenvernetzung
- die Durchführung von Verarbeitungen - z.B. die Durchführung von Datenbereinigungen.

3.3.2 Entwicklungsziele

Normdaten sollen als Teil eines weltumspannenden Netzwerks verstanden werden. Daher sind externe Verknüpfungen integraler Bestandteil der Normdaten und genauso relevant wie interne Inhalte: Das Datenangebot beinhaltet nicht alle Informationen, verlinkt aber zu allen Stellen, an denen weitere Informationen vorhanden sind. Von den aktuellen Strukturen werden externe Verknüpfungen allerdings nicht gut unterstützt: Für die Erstellung, Wartung, Pflege und Aktualisierung externer Links gibt es bisher keine durchgängigen Regelungen und keine geeignete technische Infrastruktur. Auch gibt es bislang nur wenig Unterstützung für die effiziente Bereinigung von Dubletten, das Ergänzen von internen Beziehungen und für die Integration bestehender Datenquellen (z.B. existierende Datenbestände neuer GND-Partner). Ziel ist der Aufbau einer integrierten Lösung für diesen Bereich. Die bereits erfolgreich getesteten Ansätze der CULTUREGRAPH-Datamining-Module sollen in diesem Paket weiter verfolgt und mit Konzepten und Werkzeugen aus anderen Vorhaben kombiniert werden.

- **Ziel 3.1:** Verbesserung der Qualitätssicherung beim Import. Dies beinhaltet das Festlegen von Importkriterien (passend zum definierten Mindeststandard (vgl. Ziel 2.2)).
- **Ziel 3.2:** Reduzieren von Dubletten und Inkonsistenzen im Datenbestand
- **Ziel 3.3:** Erhöhen der Anzahl an validen Verknüpfungen zwischen Entitäten
- **Ziel 3.4:** Aufbau einer modernen Infrastruktur zur Pflege und Verwaltung der GND für den deutschsprachigen Raum, welche das Vernetzen von Informationen in den Mittelpunkt rückt und das Anlegen, die Pflege und das Angebot externer und interner Verknüpfungen optimal unterstützt
- **Ziel 3.5:** Reguläres Monitoring der Datenqualität des Bestands (z.B. hinsichtlich der Konsistenz der Datensätze / Auffälligkeiten / Dubletten)
- **Ziel 3.6:** Aufbau eines zentralen Datenpools von Katalog- und Sammlungsdaten aus dem GND-Netzwerk. Dieser dient als Datengrundlage für Matching-Verfahren und für den Ausbau der GND als zentraler „Wegweiser“ (Teil von AF-4) auf die verknüpften Kulturdaten³
- **Ziel 3.7:** Reguläres Verfahren zur Berechnung von Werk-Clustern

3.3.3 Stakeholder

Mitwirkende		Nutzer	Systeme / Datenquellen		
1.	KWEs (GND-Teilnehmer und	4.	KWEs (insbesondere Kulturportale,	3.	Anbieter von Bibliothekssystemen

³ Für einen großen Teil der Bibliotheksdaten existiert dieser als Teil der CULTURGRAPH-Umgebung bereits. Allerdings müssen die Architektur und Verfahren in diesem Bereich basierend auf den praktischen Erfahrungen der letzten Jahre angepasst werden. Weiterhin gilt es, den Datenpool um Sammlungsdaten aus dem nicht-bibliothekarischen Bereich anzureichern. Hier könnte es sinnvoll sein, auf dem Datenbestand und den Verfahren der DDB aufzusetzen.

	Interessierte)		Semantic Web/Digital Humanities-Projekte)	
2.	Wikipedia, Open Data Community	5.	Wikipedia, Open Data Community	Anbieter von Systemen für Museen, Archive und Mediatheken
			Verlage, Vertriebe	Wikipedia, Open Data Community
			Recherchierende, Forschende, Studierende, Kreativschaffende	Informationsdienstleister
			Informationsdienstleister	

3.3.4 Adressierte Funktionsbereiche (vgl. Abbildung 2: Funktionale Zielarchitektur)

- Ein Funktionsmodul **DataSource Manager** dient der Verwaltung externer Datenquellen. Hierüber wird das Einspeisen von „Fremddaten“ in die Geschäftsprozesse der GND gesteuert. Das Modul beinhaltet eine erweiterbare Umgebung mit Werkzeugen und Verfahren zur Abfrage von externen Datenquellen, zur Normalisierung von Daten und zum Scheduling – also zur Definition, in welchem Turnus Aktualisierungen abgefragt werden. Die Komplexität liegt in der Vielfalt der externen Datenquellen und Formate.
- Ein **Knowledge Discovery Modul** ermöglicht die Analyse, Extraktion und automatische Generierung von Daten und Statistiken (Datamining) sowie Clustern und Mappings (Konkordanzen). Ausgangsdaten sind dabei sowohl Fremddaten als auch die GND-Datenbasis („GND CORE“ und „GND PLUS“).
- Das **Knowledge Discovery Modul** bietet neben einer Basisumgebung auch eine Integrationsschicht für die Anbindung von Werkzeugen (z.B. von Tools zur automatischen Beschlagwortung).
- Die Berechnungsergebnisse werden in einem eigenen Datenpool („**GND LAB**“) abgelegt. Die Lab-Daten sind keineswegs Daten „zweiter Klasse“. Sie können beispielsweise eine wichtige Grundlage für Suchanwendungen und Erschließungsprozesse sein. Daher stehen die Daten in der Regel für Nutzer genauso offen abrufbar zur Verfügung wie die übrigen GND-Daten. Vom LAB-Datenpool aus können Teile der gewonnenen Informationen nach zu definierenden Regeln (über vollautomatische, semi-automatische und manuelle Prozesse) auch direkt in die GND-Datenbasis eingehen.
- Eines der Werkzeuge, die bei diesem Schritt helfen sollen, ist ein **Recommender System** (oder auch *Erschließungs-Assistent*). Dieses Modul soll die intellektuelle Erschließungsarbeit durch automatisch generierte Vorschläge unterstützen. Die Regeln, nach denen Vorschläge bestimmt werden, sind spezifisch für den jeweiligen Anwendungsfall konfigurierbar. Eine zentrale Funktion dieses Moduls ist die Unterstützung beim Setzen von Verknüpfungen und dem Zusammenführen von Daten. Dies ist mit Blick auf die Einbindung von Museen und Archiven insbesondere für die effiziente Integration bestehender Datenbestände in die GND wichtig.

3.3.5 Zusammenhänge zu den Entwicklungen in anderen Aktionsfeldern

- AF-3 bietet die Werkzeuge und Verfahren für die Anpassung des GND-Datenbestands auf die in AF-2 konzipierten Regelwerks- und Formatänderungen
- Bezüglich der Visualisierung von GND-Einträgen und deren Zusammenhängen sowie der Oberflächengestaltung kann idealerweise auf Entwicklungsergebnisse aus AF-4 zurückgegriffen werden: Recommender Systeme für die Erschließung decken sich in den Anforderungen an Visualisierung (Darstellen von Zusammenhängen, Navigieren in einem Beziehungsgraphen) und dem Unterbreiten von Vorschlägen („meinten Sie...“, Angebot ähnlicher Entitäten) stark mit jenen für eine Endnutzerrecherche.

3.3.6 Relevante Vorarbeiten

Projekt	Beschreibung	Projektbeteiligte
DA2	Computerunterstützte Sacherschließung mit dem Digitalen Assistenten	IZUS/Universitätsbibliothek Stuttgart
PETRUS	Verfahren zur automatische Inhaltserschließung	DNB
xTree	Mapping von Vokabularen	digiCULT-Verbund
NERD	Automatische Entitätenerkennung in elektronischen Dokumenten	DNB Max-Planck-Institut für Informatik (MPI)

3.3.7 Entwicklungsaktivitäten

#	Aktivitäten (bei absehbarem Förderbedarf farblich markiert)	17	18	19	20	21
E.3.1	Verbesserung der Qualitätssicherung beim Import - Definition und Verbesserung allgemeiner Q-Kriterien für Datenimporte (u.a. Matching-Kriterien) - Anpassen der Import-Verfahren (u.a. Abstellen von Tn7-Importen)					
E3.2 (ARAC1)	Planung einer Umgebung für Datamining und Vernetzung von Daten (Zielinfrastruktur des Bereichs Datamining)					
E.3.3 (DEDUP)	Großflächige Reduktion von Dubletten und Bereinigung von Inkonsistenzen					
E.3.4 (TEMA)	Terminologiemanagement: Verwirklichung von Recommenderfunktionen für GND-Sachbegriffe					
E.3.5 (4C-1.AP3)	Konzeption von Matching- und Mappingverfahren für nicht-bibliothekarische Datenquellen - Konzeption und experimentelle Implementierung semi-automatischer Verfahren (Matching, Clustering, Mapping), um existierende nicht-bibliothekarische Datenbestände effizient mit der GND zu verknüpfen oder sie in diese zu integrieren. - Evaluation der Prototypen und Verfahren anhand konkreter Referenzdatensets und anhand des DDB-Bestands.					
E.3.6	Konzeption und experimentelle Implementierung semi-automatischer Verfahren (Matching, Clustering, Mapping) zur Erzeugung bibliografischer Werksätze					
E.3.7 (ARAC2)	Aufbau der Zielinfrastruktur für den Bereich Datamining: - Evaluation bestehender Ansätze - Weiterentwicklung der Knowledge Discovery Module (Schwerpunkte Clustering und Datenabgleiche) - Aufbau von Geschäftsprozessen und Werkzeugen rund um das Datasource Management					
E.3.X (4C-2.AP3)	Implementierung der Matching- und Mappingverfahren auf der Zielarchitektur - Anpassung und Weiterentwicklung der Vernetzungs- und Analyseumgebung für die unterschiedlichen Zielgruppen					

3.4 AF-4: Visualisierung und Endnutzeranwendungen – Den Zugang zum GND-Netzwerk verbessern

3.4.1 Entwicklungsaktivitäten

Es fehlt ein attraktiver Zugang zum GND-Netzwerk – eine Möglichkeit, in den direkten und indirekten Zusammenhängen des GND-Netzwerks zu stöbern und diese zu erkunden. Dabei geht es nicht nur um die Beziehungen zwischen den Normdaten: Nutzer wie Teilnehmer wünschen sich einen zentralen Anlaufpunkt, über den darüber hinaus alle mit einem GND-Eintrag verknüpften Ressourcen und Webangebote gefunden werden können – die GND als Wegweiser. Zudem mangelt es an einer einfachen Möglichkeit, ein Normdaten-basiertes Stöbern für den eigenen Bestand zu verwirklichen – ein Aspekt, der gerade für kleinere Einrichtungen von entscheidender Bedeutung ist.

Der Service soll zum Gesicht der GND mit ansprechenden Visualisierungen und Navigationsfunktionen ausgebaut und so generalisiert werden, dass es sich als Fenster oder Schnittstelle leicht in beliebige andere Portale und Suchmaschinen integrieren lässt. Vorträge und Tutorials sollen die Verbreitung unterstützen.

- **Ziel 4.1:** Verwirklichung von Visualisierungs- und Navigationsmodulen für die verschiedenen Entitätstypen
- **Ziel 4.2:** Anzeige und Aufbau eines Angebots von EntityFacts als zentrale Referenz für GND-Einträge und für die damit verknüpften Sammlungen und Datenbanken („Wegweiser-Funktion“)
- **Ziel 4.3:** Konzeption und Umsetzung (Referenzimplementierung) eines neuartigen, integrierten semantischen Recherchesystems basierend auf dem GND-Netzwerk (Recherchestrategie: Zusammenstellen von Entitäten und Zusammenhängen).

3.4.2 Stakeholder

Mitwirkende	Nutzer	Systeme / Datenquellen
3. KWEs (GND-Teilnehmer und Interessierte)	1. Recherchierende, Forschende, Studierende, Kreativschaffende	5. Anbieter von Bibliothekssystemen
Wikipedia, Open Data Community	2. KWEs	Anbieter von Systemen für Museen, Archive und Mediatheken
Recherchierende, Forschende, Studierende, Kreativschaffende	4. Informationsdienstleister	Wikipedia, Open Linked Data
	Wikipedia, Open Data Community	
	Verlage, Vertriebe	

3.4.3 Adressierte Funktionsbereiche (vgl. Abbildung 2: Funktionale Zielarchitektur)

- Der Dienst **EntityFacts** bietet eine zentrale Übersicht zu jeder registrierten Entität. Die Entitäten werden über stabile Identifier adressiert. In der Übersicht werden sowohl die Kerndaten als auch alle ergänzenden Informationen angeboten. Dies beinhaltet auch aggregierte Fremddaten und Verweise auf verknüpfte externe Datenquellen. EntityFacts bietet eine Sicht für Endnutzer (HTML-Sicht) als auch eine Maschinen-Sicht in verschiedenen Formaten (REST, XML, RDF, ...), die eine leichte Integration in Drittangebote erlaubt
- Ein **Networkbrowser** unterstützt die Endnutzer wie auch Erschließenden beim Navigieren durch das semantische Netz der GND und beim Auffinden angeschlossener Angebote (z.B. Sammlungen/Kataloge) und Datenbestände. Entscheidend sind die übersichtliche Visualisierung der Zusammenhänge und ergonomische Navigations- und Filtermöglichkeiten.
- **Statistics & Mappings:** Der Zugriff auf im Datenlabor errechnete Statistiken (z.B. zu verwaisten Elementen, zu hochgradig verknüpften Elementen, zu vermeintlichen Dubletten, ...) bietet insbesondere den Mitwirkenden Einrichtungen wichtige Hinweise zur Optimierung der GND.

3.4.4 Zusammenhänge zu anderen Aktionsfeldern

- Der Ausbau des CULTUREGRAPH-Datenpools von Katalog- und Sammlungsdaten (vgl. AF-3) ist eine Vorarbeit für die in AF-4 geplante Anzeige der entsprechenden Referenzen („Wegweiser-Funktion“).
- Allgemein befördern die in AF-4 entwickelten Visualisierungen und Navigationsansätze alle anderen Aktionsfelder.

3.4.5 Relevante Vorarbeiten

Projekt/Dienst	Beschreibung	Beteiligte
EntityFacts	Entity Facts ist ein Datendienst, der maschinenlesbare "Faktenblätter" zu Entitäten der GND bereitstellt.	DNB
Online GND	Online-Angebot für die Suche in der GND	BSZ
Web GND	Online-Angebot für die Suche in der GND	Fa. Eurospider
lobid.org	LOD-Angebot rund um Normdaten und bibliografischen Daten mit darauf aufbauenden Diensten wie Recherche-Anwendungen und einer Abfrage-API	hbz
Wikipedia-Personensuche	Suchmöglichkeit für Personenartikel der deutschsprachigen Wikipedia	Wikipedia

3.4.6 Entwicklungsaktivitäten

#	Aktivitäten (bei absehbarem Förderbedarf farblich markiert)	17	18	19	20	21
E.4.1	Aufbau von Visualisierungs- und Navigationsmodulen (diverse Forschungs- und Entwicklungskooperationen/Auftragsarbeiten) ⁴					
E.4.2 (EXP)	EntityExplorer: Aufbau eines browser-basierten Endnutzerdienstes für die Suche in der GND als zentraler Einstiegspunkt ins GND-Netzwerk, beinhaltet - Wegweiser-Funktion - Integrierbarkeit in externe Webangebote					
E.4.X (4C-2)	Optimierung der Normdatenunterstützung in der Recherche der DDB - Indexanreicherungen mit GND-Verknüpfungen durch automatische Abgleichverfahren - Erweiterung der Suchunterstützung für die Recherche mit Normdaten					
E.4.X (HUG)	Projekt „Culturegraph Hugin“: Referenzimplementierung einer Normdaten-zentrierten Recherche					

⁴ Antragerstellung seitens des Forschungspartners denkbar

3.5 AF-5: Datenversorgung und Erschließungsprozesse – Die kooperative Informationsversorgung fördern

3.5.1 Inhalt des Aktionsfelds

Die Modernisierung der Datenversorgung wird mit der Öffnung der GND und dem Hinzukommen weiterer Partner, einem wachsenden Normdatenbestand und Hinzukommenden Erweiterungen im Datenmodell (AF-1, AF-2) zunehmend kritisch. Eine Öffnung im großen Stil wird sich ohne Fortschritte in diesem Bereich nicht umsetzen lassen. Sie würde auf Kosten der Qualität und Stabilität der Dienste und Systeme der aktuellen GND-Partner gehen.

3.5.2 Entwicklungsziele

Die unterschiedlichen Einsatzbereiche, die wachsende Zahl an Einträgen und Anwendern und die zunehmenden internationalen Verflechtungen erzeugen eine Komplexität, der die aktuelle Organisation und Infrastruktur der GND nicht gerecht werden. Die Auswirkungen durch die Einführung größerer Änderungen und die Integration neuer Partner müssen beherrschbar bleiben. Die aktuelle Organisation zwischen den Anwendern und Partnern der GND ist zur Durchführung solch grundlegender Veränderungen nicht ausgelegt. Veränderungen an der GND, zu große Datenlieferungs-Volumina und entsprechende Einspielungen bei GND-Partnern haben in der Vergangenheit wiederholt zu Problemen geführt. Es werden daher klare Regelungen und technische Freiräume für notwendige Weiterentwicklungen bei gleichzeitigem Schutz des Tagesgeschäfts benötigt.

Aktuell besteht eine enge Kopplung zwischen dem an der DNB betriebenen Master-System und den Empfängersystemen. In der Regel werden die Daten komplett gespiegelt vorgehalten. Aktualisierungen dieser Spiegel ziehen in der Regel eine ganze Kette von weiteren lokalen Änderungen (Indexierungen, Prüfungen, Weiterverteilung an untergeordnete Systeme) nach sich. Durch die sehr unterschiedlichen Architekturen und Konfigurationen der angebundenen Systeme kommt es dabei (insbesondere bei Massenänderungen) immer wieder zu spezifischen lokalen Problemen. Um die stark wachsende Heterogenität und Zahl der Änderungen, die eine Erweiterung der GND zur Folge hätte, in den Griff zu bekommen, muss die Datenversorgungsinfrastruktur der GND-Zentrale und bei den Bibliothekspartnern technisch und konzeptionell dringend überarbeitet werden.

Um den Anforderungen an eine zeitgemäße, kooperativ organisierte und international vernetzte Erschließung zu genügen, ist allerdings nicht nur die Aktualität der Empfänger-Systeme entscheidend, sondern noch mehr die Aktualität des zentralen Masters: Die Geschäftsprozesse der beteiligten Einrichtungen müssen auch bei Öffnung und Erweiterung der GND so gestaltet sein, dass lokale Änderungen mit Bezug zu Normdaten innerhalb derselben Transaktion auch am Master durchgeführt werden. Auch ist es für das Vermeiden von Dubletten und Doppelarbeit entscheidend, dass Neueintragungen stets in Kommunikation mit dem Master erfolgen. Im Idealfall melden die im Rahmen des Aktionsfelds 3 entwickelten Assistenzsysteme dem Katalogisierenden mögliche passende Kandidaten. Da solche Vorschläge unter Rückgriff auf Daten zu assoziierten Kulturgüter und Publikationen errechnet werden, gilt dieses Prinzip nicht nur für die Normdaten selbst, sondern auch **für Katalog- und Sammlungsdaten**. Es gilt also, den zentralen Datenpool zu Katalog- und Sammlungsdaten in GND-PLUS stets aktuell zu halten.

- **Ziel 5.1:** Modernisierung der Datenversorgungsinfrastruktur:
 - Organisationsübergreifende Vereinheitlichung der Datenversorgungsinfrastruktur
 - Vermeiden unnötiger Replikationen (Beschränkung auf spezifische Sichten⁵ und Teilbestände)
- **Ziel 5.2:** Verzahnung der Erschließungsgeschäftsgänge der angeschlossenen Bibliotheken:

⁵ meint u.a. das Ausblenden von Änderungen an sparten-spezifische Erweiterungen, die nicht für alle Partner relevant sind (vgl. AF-2)

- Integration der in AF-3 entwickelten Assistenzsysteme in die eigene Erschließungsumgebung
- Simultanes Übertragen von relevanten Aktualisierungen (bezogen auf Normdaten sowie relevante Titelinformationen) an die GND-Datenbasis

3.5.3 Stakeholder

Mitwirkende		Nutzer		Systeme / Datenquellen	
1.	KWEs (insbesondere GND-Agenturen für das Bibliothekswesen)	2.	KWEs (insbesondere GND-Agenturen für das Bibliothekswesen)	5.	Anbieter von Bibliothekssystemen
		3.	Verlage und Vertriebe Informationsdienstleister		Anbieter von Systemen für Museen, Archive und Mediatheken

3.5.4 Adressierte Funktionsbereiche (vgl. Abbildung 2: Funktionale Zielarchitektur)

- Ein *Data-Supply-Framework* bietet eine performante, empfängerspezifisch konfigurierbare Datenversorgung. Unter anderem ist es möglich, Änderungen an der Masterdatei auszublenden, die für das eigene Tagesgeschäft irrelevant sind. Dadurch werden unnötige Datentransfers und daraus resultierende teure lokale Updateprozesse verhindert. Denn nicht in allen Bereichen ist eine zeitnahe und vollständige Übernahme von Neueintragungen und Änderungen für die eigene Anwendung sinnvoll. Über dieses Framework sollte auch in Zukunft die Anbindung an Cloud-Lösungen gelöst werden.

3.5.5 Relevante Vorarbeiten

3.5.6 Entwicklungsaktivitäten

Besondere Herausforderung ist die Entwicklung von Adaptern für die verschiedenen Systemlösungen. Diese Arbeit können in weiten Teilen nur die Systemanbieter und/oder Datenempfänger selbst leisten. Seitens der GND-Zentrale (DNB) können nur die Abläufe, Protokolle und Schnittstellen erarbeitet werden und (mit Unterstützung durch technische Partner) ggf. Referenzimplementierungen bereitgestellt werden. Die Umsetzung auf Empfängerseite ist aber im Allgemeinen Sache des Partners und seines Systemanbieters. Daher sind auf diesem Feld Kooperationsprojekte mit den verschiedenen Partnern und Systemanbietern erforderlich.

#	Aktivitäten (bei absehbarem Förderbedarf farblich markiert)	17	18	19	20	21
E.5.X	- Modernisierung der Datenversorgungsinfrastruktur					
(SYNC)	- Verzahnung der Erschließungsgeschäftsgänge					

3.6 AF-6: Zusammenarbeit mit anderen Communities – Nutzerkreise und Anwendungskontexte erweitern

3.6.1 Inhalt des Entwicklungsfelds

Es besteht eine starke Abhängigkeit vom Erfolg im AF-2 (Tracking von Provenienzen, flexible Rechtvergabe, modulare Erweiterungen im Datenmodell, etc.). Auf Grundlage der aktuellen Umgebung können solche Angebote nur in reduziertem Maße entwickelt werden.

3.6.2 Entwicklungsziele

Die GND soll nicht nur Brücken zwischen einem Kreis kooperierender Kultur- und Wissenschaftseinrichtungen und deren Sammlungen und Datenbanken bauen, sondern auch zu anderen Communities. Aus solchen Verbindungen erwachsen für die GND-Teilnehmer und Nutzer vielfältige Vorteile:

- Die internationale Vernetzung der GND mit Norm- und Katalogdaten weltweit (wie z.B. über VIAF) erleichtert die internationale Zusammenarbeit zwischen Kultureinrichtungen und verbessert global/international-angelegte Rechercheangebote und Nachweissysteme
 - Daten aus allen Lebens- und Fachbereiche können in das eigene Rechercheangebot integriert werden; die eigene Sammlung wird in einen erweiterten, belebenden Kontext gesetzt, die einen neuen Zugang ermöglichen.
 - Über Crowdsourcing-Ansätze und Kooperationen mit Netzwerken anderer Communities (Wissenschaftler, Fachexperten, Autoren, Verlage, Studierende, ...) lässt sich die Qualität der Daten und Angebote verbessern
 - Über Forschungsk Kooperationen können die kulturellen Sammlungen und Daten in weitere Anwendungskontexte gebracht werden und die Formulierung und Bearbeitung neuer Forschungsfragen unterstützen.
 - Die Zusammenarbeit mit Informationsdienstleistern und der Wikipedia erhöht die Reichweite der Daten und Sichtbarkeit der Sammlungen. Alltägliche Internetrecherchen münden häufiger in Angebote des GND-Netzwerks und machen diese leichter zugänglich.
- **Ziel 6.1:** Weiterentwicklung des Virtual International Authority File (VIAF)
 - **Ziel 6.2:** Ausbau der erfolgreichen Zusammenarbeit mit Wikipedia
 - **Ziel 6.3:** Ausbau der Vernetzung mit externen Datenquellen unterschiedlicher Fachbereiche
 - **Ziel 6.4:** Intensivierung der Zusammenarbeit mit verwandten Netzwerken wie ORCID und ISNI
 - **Ziel 6.5:** Stufenweiser Aufbau von Angeboten, die weitere Adressaten wie Hochschulen, Verlage und Kulturschaffende und private Akteure (Citizen Science) nicht nur in der Rolle als Nutzer, sondern auch als aktiv Mitwirkende einbeziehen

3.6.3 Stakeholder

Mitwirkende		Nutzer	Systeme / Datenquellen
1.	Wikipedia, Open Data Community Recherchierende, Forschende, Studierende, Kreativschaffende Informationsdienstleister Verlage, Vertriebe KWEs	1. (identisch zu „Mitwirkende“)	2. (stark projektabhängig/kann für dieses AF nicht allgemein angegeben werden)

3.6.4 Adressierte Funktionsbereiche (vgl. Abbildung 2: Funktionale Zielarchitektur)

- Im Rahmen des AF wird die Umgebung zur Datenpflege (**Community-Plattform/Cataloging-API**) und Datennutzung bedarfsweise erweitert um

- Oberflächen, Schnittstellen und Werkzeuge für die Zusammenarbeit mit Drittsystemen anderer Communities (z.B. für die Integration in Verlagssoftware oder virtuelle Forschungsumgebungen)
- Oberflächen, Schnittstellen und Werkzeuge zum Aufbau von Crowd-Sourcing-Projekten

3.6.5 Relevante Vorarbeiten

3.6.6 Entwicklungsaktivitäten

#	Entwicklungsaktivitäten (bei absehbarem Förderbedarf farblich markiert)	17	18	19	20	21
E.6.1	Beteiligung an Weiterentwicklung des Virtual International Authority File (VIAF) und Intensivierung des Datenaustauschs mit anderen Normdateien	■	■	■	■	■
E.6.2 (ORC)	DFG-Projekt „ORCID DE“	■	■	■	■	■
?	Ggf. Projektanträge / Projektkooperationen mit verschiedenen Stakeholdern	■	■	■	■	■