

Spartenübergreifende Öffnung und Weiterentwicklung der Gemeinsamen Normdatei: Das Projekt *GND für Kulturdaten (GND4C)*

Angela Kailus

Deutsches Dokumentationszentrum für Kunstgeschichte – Bildarchiv Foto Marburg
Philipps-Universität Marburg, Deutschland

kailus@fotomarburg.de

KURZDARSTELLUNG: Das DFG-Projekt GND für Kulturdaten (GND4C) will die Weichen für eine Weiterentwicklung und verbesserte Nutzbarkeit der Gemeinsamen Normdatei (GND) durch nicht-bibliothekarische Institutionen wie Museen, Archive, Forschungseinrichtungen, Denkmalbehörden und Mediatheken stellen. Dazu ist nicht nur der nachhaltige Aufbau einer spartenübergreifenden Organisationsform notwendig, sondern auch die Überprüfung und Weiterentwicklung des GND-Datenmodells und die Bereitstellung von Schnittstellen und Software-Tools zur Unterstützung des erweiterten Anwenderkreises. Der Beitrag stellt die Ziele und Strategien des Projekts vor, das im Mai 2018 begonnen hat.

1. EINFÜHRUNG

Im Web der Daten sind Normdaten der ideale Ansatz, um Kulturgüter, Bibliotheksbestände und Forschungsdaten fächer- und spartenübergreifend miteinander zu verbinden. Sie sind der Stoff, aus dem das Rückgrat eines globalen maschinenlesbaren, semantischen Netzes der Kultur und Wissenschaft aufgebaut werden kann. Auf dieser Grundlage können inhaltliche Verbindungen zwischen Kulturgütern, Bibliotheksbeständen, Forschungsdaten und assoziierten Personen, Körperschaften, Orten, Objekten, Ereignissen und Projekten ausgewiesen und maschinell auswertbar werden, die sonst im Verborgenen bleiben würden. Das Potential semantischer Vernetzung wurde im deutschsprachigen Raum besonders früh erkannt - 2012 legten die Deutsche Nationalbibliothek und ihre Partner mit dem Aufbau der global vernetzten Gemeinsamen Normdatei (GND) den Grundstein für einen fächer- und spartenübergreifenden Normdaten-Hub. Sie ist kostenfrei unter der Lizenz CC0 1.0 zugänglich und wird zunehmend auch von nicht-bibliothekarischen Kultur- und Wissenseinrichtungen genutzt.

Allerdings ist die GND organisatorisch, fachlich und technisch noch immer stark auf die Nutzung durch Bibliotheken ausgerichtet. Den Bedarfen anderer Kultur- und Wissenseinrichtungen wie Archiven, Museen, Mediatheken, Denkmalbehörden,

Universitäten und Wissenschaftsnetzwerken wird sie nur eingeschränkt gerecht. Vor diesem Hintergrund haben die GND-Partner sich einstimmig für eine konsequente Öffnung des Angebots für nicht-bibliothekarische Anwendungskontexte ausgesprochen. Aus dem bisher auf eine bibliothekarische Nutzung ausgerichteten Angebot soll ein spartenübergreifendes Produkt werden, das in diesem erweiterten Rahmen seine größten Stärken zum Tragen bringt: Verlässlichkeit, Transparenz, Eindeutigkeit und ein reichhaltiges Angebot semantischer Verknüpfungen.

2. DAS PROJEKT

Das von der Deutschen Forschungsgemeinschaft (DFG) im Rahmen des Programms *e-Research-Technologien* geförderte Projekt *GND für Kulturdaten (GND4C)* soll hierfür die Grundlagen legen. Die erste, auf 30 Monate ausgelegte Projektphase hat im Mai 2018 begonnen. Das Projekt wird federführend von der Deutschen Nationalbibliothek (DNB) geleitet. Weitere Projektpartner sind die Deutsche Digitale Bibliothek (DDB), das Bibliotheksservice-Zentrum Baden-Württemberg (BSZ), das Deutsche Dokumentationszentrum für Kunstgeschichte – Bildarchiv Foto Marburg (DDK), die Schleswig-Holsteinische Landesbibliothek mit der digiCULT-Verbund eG und das Landesarchiv Baden-Württemberg (LABW). Zahlreiche weitere Einrichtungen und Gremien aller Sparten haben den

Projektantrag durch einen Letter of Intent unterstützt.

Das Projekt verfolgt folgende Ziele:

2.1 NACHHALTIGER AUFBAU EINER SPARTEN- UND FÄCHER-ÜBERGREIFENDEN ORGANISATION

Angesichts der neuen Herausforderungen an die GND und der vielfältiger werdenden Anwenderkreise soll mit dem Konzept einer GND-Kooperative die Zusammenarbeit der Teilnehmer neu organisiert werden: Ziel ist zunächst die Festlegung einer gemeinsamen Strategie für die Weiterentwicklung der GND und die verbindliche Regelung und Organisation der Zusammenarbeit. Der Standardisierungsausschuss unter Leitung der DNB erarbeitet dazu Entwürfe für eine Kooperationsvereinbarung, die Leitlinien und die strategische Ausrichtung der Kooperative und eine Geschäftsordnung. Perspektivisch sollen alle an der GND teilnehmenden Einrichtungen für die Kooperative gewonnen werden, die Zielgruppe umfasst daher Bibliotheken, Archive, Museen und andere Kultureinrichtungen.

Das mehrstufige Organisationsmodell besteht aus der GND-Zentrale, GND-Agenturen und Teilnehmern mit jeweils definierten Rollen. Die Agenturen sind Ansprechpartner und Koordinatoren für eine Gruppe von GND-Partnern. Sie vertreten diese in den Gremien und Arbeitsgruppen der GND und bringen deren Anforderungen und Impulse ein.

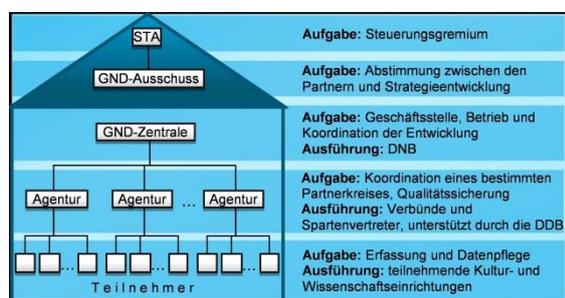


Abb. 1: Organisation der GND-Kooperative

Ein Ziel des Projektes ist es, das Organisationsmodell weiterzuentwickeln und die Integration nicht-bibliothekarischer Partner in diese Struktur vorzubereiten. Dazu sollen sich Einrichtungen innerhalb ihrer Community zu Interessengruppen organisieren, sich bestehenden Agenturen anschließen oder neue Agenturen aufbauen. Die GND4C-

Projektpartner beabsichtigen, in der sich formierenden Kooperative eine tragende Rolle für ihre Fachgemeinschaften zu übernehmen.

2.2 WEITERENTWICKLUNG DES DATENMODELLS

Um die GND universell einsetzbar und geeignet für Beiträge aus nicht-bibliothekarischen Bereichen zu machen, aber gleichzeitig die Nutzbarkeit für bibliothekarische Kernanwendungen zu erhalten, ist es notwendig, die bestehenden Regelungen und Konzepte für den Ausbau und die Pflege der GND weiterzuentwickeln. Außerdem muss das Datenmodell zur Unterstützung von Entitäten angepasst werden, die bisher noch nicht oder ungenügend durch die GND abgedeckt sind – beispielsweise Geografika, Bauwerke mit ihren Ausstattungsobjekten, Archivalien, Museumsobjekte, druckgraphische Werke und Handschriften. Von entscheidender Bedeutung ist zudem die geplante Einführung gruppenspezifischer Bereiche und Sichten: Einige Eigenschaften sind Teil eines spartenübergreifend gepflegten und redaktionell bearbeiteten Datenbereichs (GND-CORE), andere Teile werden von den verschiedenen Anwendungsgruppen als optionale gegenseitige Ergänzungen begriffen (GND-PLUS).

Ausgehend von einem im Projekt zu erarbeitenden abstrakten Basisdatenmodell werden das CORE-Datenmodell sowie die notwendigen fachspezifischen Erweiterungen spezifiziert. Dies erfolgt anhand konkreter Fallbeispiele und über die Analyse von Datensets unterschiedlicher Kultureinrichtungen. Die Ergebnisse werden im Standardisierungsausschuss der Bibliotheken hinsichtlich Funktionalität und Umsetzbarkeit evaluiert.

2.3 BEREITSTELLUNG VON SCHNITTSTELLEN UND WERKZEUGEN ZUR ANBINDUNG NEUER DATENQUELLEN

Insbesondere für das effiziente Vernetzen vorhandener nicht-bibliothekarischer Datenbestände mit der GND existiert bis dato keine ausreichende Unterstützung. Um diese in verteilten Fachdatenbanken mit der GND zu verbinden oder sie ganz oder teilweise in die GND zu integrieren, ist ein effizientes und gleichzeitig zuverlässiges Abgleichverfahren

nötig. Es gilt, bereits in der GND vorhandene Entsprechungen der benannten Entitäten zu identifizieren, bislang fehlende Referenzierungen herzustellen oder auch Informationen zu ergänzen. Dies ist nur mittels geeigneter Werkzeugunterstützung und ausgefeilter Qualitätssicherungsprozesse zur Vermeidung von Datensatz-Dubletten möglich.

Im Projekt ist daher die Entwicklung und prototypische Implementierung semi-automatischer Verfahren (Matching, Clustering, Mapping) vorgesehen, um existierende Datenbestände effizient mit der GND zu verknüpfen oder die Integration von extern generierten Daten in die GND zu ermöglichen.

Die Konzeption und Entwicklung der Softwaretools liegt in der Hand von digiCULT. Im Projektlabor werden unter Berücksichtigung der Anforderungsdefinitionen aller Projektbeteiligten lauffähige Prototypen zu Test- und Demonstrationszwecken erstellt und auf die Daten der Fallbeispiele angewendet. Die Ergebnisse sollen im Rahmen von Workshops einem breiteren Anwenderkreis vorgestellt und evaluiert werden.

2.4 KOMMUNIKATION MIT DEN VERSCHIEDENEN INTERESSEN-GRUPPEN

Neben den Kommunikationsstrukturen der GND-Kooperative müssen auch weitere Kommunikationskanäle in die Communities hinein aufgebaut werden. Nur so kann eine bedarfsgerechte Entwicklung und Verankerung der GND in der täglichen Praxis der verschiedenen Sparten und Fächer sichergestellt werden. Darüber hinaus gilt es, die GND in der allgemeinen Öffentlichkeit in ihrer neuen Rolle als Kulturdatennetzwerk sichtbar zu machen. Im Projekt sollen daher Dialogformate zur Erhebung, Koordination und Dokumentation von community-spezifischen Anforderungen und zur Abstimmung gemeinsamer übergreifender Standards entwickelt und durchgeführt werden (z. B. Veranstaltungen, Webplattform).

3. ANFORDERUNGSSZENARIEN: FALLBEISPIELE

Um praxistaugliche und belastbare Ergebnisse zu erhalten, erfolgt die Konzeptentwicklung

und Evaluation anhand repräsentativer und gleichzeitig besonders gewinnbringender Fallbeispiele. Dadurch wird einerseits die Unterstützung des jeweiligen Anwendungsfalls durch die GND konzeptionell vorbereitet. Andererseits sollen auf diesem Wege generelle Erkenntnisse und Erfahrungen im Umgang mit community-spezifischen Anforderungen gesammelt werden.

Alle Fallbeispiele adressieren primär jeweils einen Entitätstyp: Geografika, Werke (Bau- und Kunstwerke), Sachbegriffe, Personen. Sie berühren aber implizit – über die für die Modellierung benötigten Merkmale – auch andere Arten von Entitäten.

3.1. BAUWERKE UND ANDERE ORTSFESTE KUNSTWERKE

In Forschungskontexten und Dokumentationszusammenhängen der Kunstgeschichte, der Archäologie und verwandter Wissenschaften werden Bauwerke und ihre Ausstattungsobjekte sowie andere ortsfeste Monumente wissenschaftlich beschrieben und untersucht. Hierbei spielen auch graphische Bildquellen wie Architekturzeichnungen sowie fotografische Aufnahmen eine wichtige Rolle. Ressourcen zum überlieferten Denkmälerbestand sollten über die GND identifiziert werden können, um ihre eindeutige Referenzierbarkeit und somit die wissenschaftliche Praxis wesentlich zu verbessern.

Die deutsche Denkmalpflege insbesondere in den Landesdenkmalämtern erhebt und publiziert systematisch hochwertige Daten zu Bau- und Kunstdenkmälern in sämtlichen Regionen. Die GND ist in der Lage, diese forschungsrelevanten, fast ausnahmslos georeferenzierten Daten über die Entitäten der Bauwerk-Geografika (Entitätencode gib) und der Werke (Entitätencode wit) abzudecken. Bisher ist die quantitative Abdeckung der GND hier jedoch völlig unzureichend, die Binnenstrukturierung dieser Daten ist heterogen. URIs aus diesen Facetten der GND werden daher noch kaum in Linked-Data-Szenarien eingesetzt.

Die Bearbeitung dieses Fallbeispiels erfolgt durch das DDK. Referenzdatensets werden vom Landesamt für Denkmalpflege Bremen und vom DDK selbst bereit gestellt.

3.2. ORTE UND ANDERE GEOGRAFIKA

Für Kulturgut ist die Recherche nach Ortsbezügen ein zentrales Anliegen. Insbesondere die historische Forschung formuliert Fragestellungen regelmäßig unter Einbeziehung räumlich-geografischer Kriterien. Ein Großteil des Kulturguts kann lokal oder regional verortet werden. Bei der Erschließung kommen oft kontrollierte Vokabulare in Form regionaler geographischer Thesauri zum Einsatz. Die GND ist in diesem Bereich bislang nicht als übergreifende Referenz etabliert, da sie die benötigten Geografika bislang nicht im erforderlichen Umfang und in der nötigen Granularität (Ortsteile, kleinere Siedlungseinheiten) bereit stellt, die historische Dimension nur unzureichend abdeckt und nur partiell die etwa für GIS-Anwendungen notwendigen Geokoordinaten aufweist. Historische Fragestellungen erfordern zudem die persistente Identifikation einer geographischen Entität auch bei wechselnder Verwaltungszugehörigkeit.

Die Geografika werden insbesondere mit Blick auf die archivischen Anforderungen vom LABW bearbeitet, unter Hinzuziehung beispielhafter Datenbestände des LABW und des Hessischen Landesarchivs.

3.3. OBJEKTBSCHREIBENDE SACHBEGRIFFE DER MUSEUMS-DOKUMENTATION

In deutschen Museen wird für die wissenschaftliche Erschließung der Sammlungsbestände eine Vielzahl unterschiedlicher Erfassungssysteme verwendet, die etablierte Datenmodelle und Metadatenstandards oft nur ungenügend berücksichtigen. Voraussetzung für eine zukunftsfähige Nutzung der Bestandsdaten und Desiderat ist daher vor allem die einheitliche Erschließung mit verbreiteten, qualitativ hochwertigen kontrollierten Vokabularen, um die Voraussetzungen für eine qualifizierte bestandsübergreifende Recherchierbarkeit in Katalogverbänden und Publikumsportalen zu erfüllen. Divergierende Benennungen müssen durch gemeinsam genutzte Vokabulare aufgefangen werden.

Zur Anreicherung mit Normdaten bietet sich die GND als meistgenutztes Vokabular im deutschsprachigen Raum an. Ein großer Teil

der von Museen verwendeten Schlagwörter ist aber noch nicht in der GND-Facette der Sachschlagwörter verfügbar. Die Einbindung dieser Sachbegriffe in die GND wäre ein wichtiger Schritt zu einer besseren Metadatenqualität und Voraussetzung für ein gutes (semantisches) Retrieval in Portalen wie der DDB oder Europeana.

DigiCULT und die DDB werden anhand eines umfangreichen Referenzdatensets des Deutschen Historischen Museums Lösungsstrategien für die Integration der Sachbegriffe aus der GND in Datensätze aus dem Museumsbereich und für die Integration der objektbeschreibenden Daten aus dem Museumsbereich in die GND erarbeiten.

3.4. KÜNSTLERINNEN UND KÜNSTLER FÜR DIE MUSEALE FORSCHUNG

Für die zeitliche, geografische oder kontextuelle Zuordnung von Kulturerbe-Objekten nehmen Personen-Informationen eine zentrale Stellung ein. Eine GND-Referenzierung ermöglicht die eindeutige Identifikation von Personen und deren semantische Vernetzung. Sie erleichtert außerdem die Datenanreicherung durch Synchronisation mit anderen Ressourcen.

Gleichwohl ist die GND in der Objektdokumentation von Museen als Quelle und Referenz für biografische Normdaten bislang nicht etabliert. Obwohl Personendatensätze den überwiegenden Teil der GND ausmachen, lässt der Abdeckungsgrad von im Museumsbereich relevanten Personen zu wünschen übrig. Zeitgenössische Künstler, abgebildete Personen auf Bildwerken, Münzherren, Münzmeister, Kunsthandwerker, Auftraggeber, Voreigentümer im Rahmen der Provenienzforschung sind in bibliothekarischen Kontexten häufig noch nicht in Erscheinung getreten und daher in der GND nicht erfasst.

Über die Ergänzung der GND hinaus wird angestrebt, auch bereits verbreitet genutzte biografische Korpora wie das Allgemeine Künstlerlexikon (AKL) oder das World Biographical Information System (WBIS) mit den GND-Personen zu synchronisieren.

Das BSZ wird die Anforderungen von Museen in diesem Bereich anhand eines Personendatenbestands aus neun Kunstmuseen des MuSIS-Verbunds untersuchen.

4. AUSBLICK

Die erste Projektphase widmet sich der detaillierten Klärung von Anforderungen und Konzepten für den Ausbau der GND und endet mit einem Pilotbetrieb der erweiterten Organisationstruktur. Die zweite Projektphase (Monate 30–60) hat die Verstetigung und Erweiterung der Organisation sowie die Implementierung der entwickelten Konzepte und Werkzeuge für den produktiven Einsatz zu Ziel.

Die Strategie trifft hoffentlich auf eine große Akzeptanz in den angesprochenen Communities, so dass die GND in ihrem kontinuierlichen Ausbau, ihrer Pflege und Anwendung zu einem echten spartenübergreifenden Gemeinschaftsunternehmen wird.

5. LITERATURHINWEIS

[1] *GND für Kulturdaten (GND4C)* (2018) [online], Online im Internet: <https://wiki.dnb.de/pages/viewpage.action?pageId=13455796> (20.10.2018)

[2] GND-Kooperative, Kett, Jürgen, Deutsche Nationalbibliothek (2017): *Initiative für Normdaten und Vernetzung GND-Entwicklungsprogramm 2017-2021* [online], Online im Internet: <https://wiki.dnb.de/x/npnpBw> (20.10.2018)

[3] Van der Graaf, Maurits & Waaijers, Leo (2014): *Authority files: Breaking out of the library silo to become signposts for research information* [online], Online im Internet: http://repository.jisc.ac.uk/6224/1/Authority_files_-_Breaking_out_of_the_library_silo.pdf (20.10.2018)