Stephan Bartholmei, Francesca Schulze

Die Weiterentwicklung der Deutschen Digitalen Bibliothek

Ein Blick hinter die Kulissen

Seit zehn Monaten steht die Beta-Version des Internetportals der Deutschen Digitalen Bibliothek (DDB) öffentlich zur Verfügung und bietet einen zentralen Zugang zu digitalen Inhalten aus mehr als 90 deutschen Kultur- und Wissenschaftseinrichtungen[1]. Über das Portal sind rund 5,7 Mio. digitale Objekte und Einträge zu noch nicht digitalisiertem Kulturgut kostenfrei recherchierbar. Rund 1.700 Besucherinnen und Besucher nutzen dieses Angebot täglich. Sie kommen überwiegend aus Deutschland, etwa 20 % von ihnen aus anderen Teilen der Welt.

Laut einer Vielzahl von positiven Rückmeldungen liegt die Attraktivität der DDB vor allem in der spartenübergreifenden und interdisziplinären Präsentation der Objekte. So sind sowohl im Datenbestand als auch in der Kultur- und Wissenschaftslandkarte alle Sparten repräsentiert – von Archiven, Bibliotheken, Denkmalpflege- und Forschungsinstitutionen bis hin zu Mediatheken und Museen. Fast 2.000 Einrichtungen haben sich bereits für die DDB registriert und sind in der Landkarte verzeichnet. Darüber hinaus erhält die Servicestelle DDB[2] als zentrale Anlaufstelle für neue Partner täglich Anfragen von Einrichtungen, die ihre digitalen Bestände einbringen und die DDB somit aktiv mitgestalten möchten. Dieses große Interesse ist nicht zuletzt auf die kooperative Mitarbeit der Fachstellen der DDB zurückzuführen, die als Experten und Vermittler in ihren jeweiligen Sparten und Netzwerken agieren.

DDB präsentiert ihre Objekte spartenübergreifend und interdisziplinär

Für den Anfang kann sich diese positive Resonanz sehen lassen. Sie zeigt, dass das Gemeinschaftsvorhaben DDB auf dem richtigen Weg ist, seinen öffentlichen Auftrag zu erfüllen. Es stellt sich die Frage, was die DDB unternimmt, um das Angebot noch attraktiver zu gestalten und damit die unterschiedlichen Ansprüche verschiedener Nutzergruppen – von der kulturinteressierten Öffentlichkeit über Informationsspezialisten aller Sparten bis hin zu IT-Entwicklern – zu bedienen? Dieser anspruchsvollen Aufgabe widmen sich die 13 Einrichtungen des Kompetenznetzwerks[3] gemeinsam mit dem technischen Betreiber, dem FIZ Karlsruhe – Leibniz-Institut für Infrastruktur, sowie Entwicklungspartnern wie dem Fraunhofer-Institut für Intelligente Analyse- und Informationssysteme IAIS. Seit dem Beta-Launch ist im Hintergrund des Portals viel passiert ...

DDB-Angebot soll noch attraktiver werden

Einführung des Europeana Data Models

Im Zentrum der Arbeiten stand die Umstellung des DDB-internen Datenformats auf das Europeana Data Model (EDM)[4] und damit verbunden die Optimierung der Datenverarbeitungs- und Einspielungsprozesse in die DDB. Vor dem sogenannten Ingestieren in das Repository der DDB werden die in einem der Eingangsformate[5] der DDB gelieferten Daten in dieses interne Datenformat transformiert.

EDM basiert auf Relationsprinzipien von Linked Data

Das Europeana Data Model ist ein einfaches und flexibles Modell, das auf den Relationsprinzipien von Linked Data basiert. Die Daten werden in der Syntax des Ressource Description Frameworks (RDF) modelliert. Das Modell besteht aus 15 Klassen und bietet die Besonderheit, eine Ressource aus drei verschiedenen Blickwinkeln zu beschreiben: als kulturwissenschaftliches Objekt edm:ProvidedCHO, dessen digitale Repräsentation(en) edm:Web Ressource sowie dessen Metadaten ore:Aggregation. Mit dieser Grundausstattung werden nicht nur die semantische Vernetzung der kulturwissenschaftlichen Objekte und ihrer zugehörigen Metadaten in der DDB verbessert, sondern auch neue, explorative Recherche- und Präsentationsmöglichkeiten über verschiedene Datenbestände hinweg den Nutzerinnen und Nutzern des Portals eröffnet.

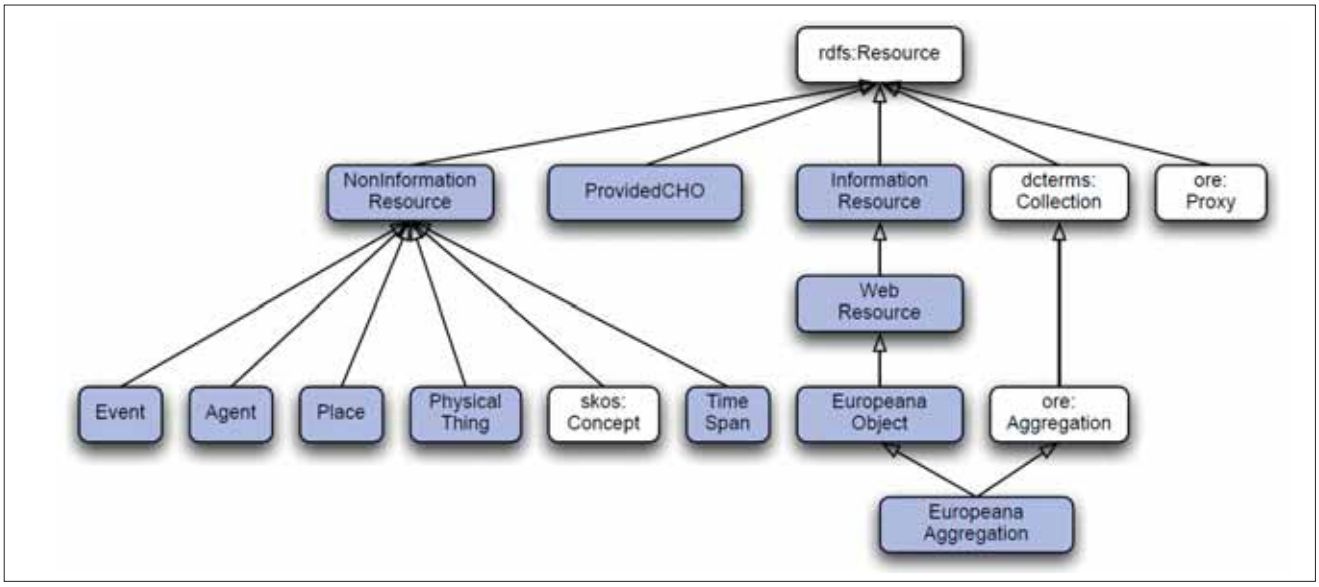
Neue, explorative Recherche- und Präsentationsmöglichkeiten

Neue Recherche- und Darstellungsformen

Im DDB-Portal wird das EDM vor allem für drei semantische Funktionen verwendet werden:

- Die sinnvolle Filterung von Suchergebnissen (Facetten).
- Die hierarchische Darstellung von zusammengehörigen Objekten eines Bestandes oder einer Dokumentationseinheit in der Detailansicht eines Treffers. Somit kann man beispielsweise von einem Artikel zum Zeitschriftenheft bis zur übergeordneten Zeitschrift navigieren (Hierarchien).
- Die Verlinkung von Objekten mit Kontextinformationen, wie beispielsweise Personen, Orten und Zeitangaben, die nach und nach aus extern gepflegten Normdatenbeständen an die DDB angebunden sein werden (z. B. Gemeinsame Normdatei – GND[6]).

In Grundzügen sind diese Funktionen bereits in der Beta-Version des Portals sichtbar. Ein Schwerpunkt der gemeinsamen Entwicklungsarbeiten besteht darin, das Potenzial des interoperablen Datenmodells EDM weiter auszunutzen und auf der Präsentationsoberfläche sichtbar zu machen. Dabei waren viele konzeptionelle Vorüberlegungen notwendig: Möchten Nutzerinnen und Nutzer beispielsweise gezielt nach einer Person suchen und in den Treffern danach filtern können, welche Werke



Das Europeana Data Model
Violett hinterlegte Klassen gehören zum Namensraum edm:

von oder über diese Person geschrieben wurden, auf welchen Fotos sie abgebildet oder in welchen Filmen sie zu sehen ist? Diesen Bedürfnissen möchte die DDB mit einer sinnvollen Ausdifferenzierung der bereits sichtbaren Basisfacetten in Kombination mit anderen Recherche-, Verlinkungs- und Darstellungsformen nach und nach Rechnung tragen, wobei hier die Event-basierte Modellierung und die verwendeten kontrollierten Vokabulare für die Event-Typen (z. B. »Creation« für die Unterscheidung nach »Werke von«) eine wichtige Rolle spielen.

Angebotserweiterung um Personenseiten

Als erste signifikante Neuerung wird das Angebot um Personenseiten erweitert werden, die bereits in einer internen prototypischen Implementierung vorliegen. Durch die Anbindung eines Normdatenservices der Deutschen Nationalbibliothek (DNB) werden GND-Personendaten sowie verlinkte Daten oder sogar Bilder aus anderen Quellen, wie beispielsweise der Wikipedia, dynamisch in der Portaloberfläche präsentiert werden. In der Personenseite fließen alle diese Datenbestände zusammen, wodurch eine zentrale, explorative Suche in der DDB und den anderen Datenquellen möglich wird. Ein weiteres Entwicklungsvorhaben ist die geokoordinatenbasierte Darstellung von Suchergebnissen in einer Landkarte, für das ein ähnliches Modellierungs- und Datenvorhaltungsprinzip zur Anwendung kommen wird.

Der Datenumfang in diesen Funktionalitäten hängt stark von der Verfügbarkeit standardisierter Informationen in den Lieferdaten ab. Die Servicestelle DDB gibt den Einrichtungen Hilfestellungen, die Qualität ihrer Daten zu verbessern und vermittelt Wege zur Nutzung der GND, aber auch anderer Vokabulare und Konkordanzen. Dies ist eine große Herausforderung, da die Datenqualität in den unterschiedlichen Sparten sehr heterogen ist. Entsprechende Dokumentationen, Vokabulare und Hilfestellungen sollen für die Datengeber in einem Serviceportal bereitgestellt werden. Auf diese Weise möchte die DDB dazu beitragen, dass Metadatenstandards in größerem Umfang verwendet werden und sich so die Qualität der Daten schrittweise verbessert.

Heterogene Datenqualität in unterschiedlichen Sparten

Datenlieferung an Europeana

Durch die Anwendung des Datenmodells EDM wird auch die Datenlieferung der DDB an das europäische Kulturportal Europeana[7] vereinfacht. Da beide Portale EDM als gemeinsamen Interoperabilitätsstandard nutzen, müssen die Daten der Kultur- und Wissenschaftseinrichtungen nur einmal von der DDB umgewandelt werden. Damit wird der gesamte Prozess des konzeptionellen und technischen Mappings wesentlich effizienter und eine zügige Weitergabe der Sammlungsdaten an Europeana gewährleistet. Aber auch die Nachnutzung eines verbreiteten Community-Standards birgt Herausforderungen. So mussten die Metadatenexperten des DDB-Kompetenznetzwerks mit der Situation umgehen, dass im Europeana-Portal einige der benötigten Aspekte des Datenmodells (z. B. die Klasse »Event«) noch nicht implementiert sind. In enger Zusammenarbeit mit der Europeana hat die DDB daher ein Profil des EDM definiert, das sowohl die Portalfunktionen der DDB als auch diejenigen der Europeana bedient.

Vereinfachung der Datenlieferung an Europeana durch EDM

Mittlerweile sind die Arbeiten weit fortgeschritten: Sowohl komplexe Bibliotheksdaten (Formate: MARCXML, METS-MODS) als auch Archivdaten (EAD-DDB) sind bereits vollständig in das EDM umgesetzt. In den kommenden Wochen wird die Umwandlung des gesamten Datenbestandes aller Sparten in der DDB, einschließlich Mediathek- (DC-Profile), Forschungs- (indexMeta), Denkmalpflege- (denkXWeb) und Museumsdaten (LIDO), abgeschlossen sein.

Die technische Infrastruktur wird aktuell in Abstimmung mit den Kollegen der Europeana implementiert. Danach können die DDB-Daten in regelmäßigen Abständen automatisiert über eine OAI-Schnittstelle von Europeana abgeholt werden. Die Servicestelle DDB trägt dafür Sorge, dass nur Datensätze oder Datenbestände, die mit der Lizenz »Creative Commons Zero Universal Public Domain Dedication«[8] ausgezeichnet sind, an Europeana weitergegeben werden und klärt die Frage der Rechteauszeichnungen mit den Datengebern.

Nur CC0-Daten werden weitergegeben

Aber auch für andere Anwendungen der DDB spielen standardisierte Rechteangaben eine wichtige Rolle: Um den Nutzern verlässliche Informationen darüber zu geben, was sie mit den gefundenen Inhalten oder Metadaten tun dürfen und was nicht, hat die DDB ein Lizenzmodell entwickelt, das sich in Bezug auf digitale Inhalte am Vorbild der Europeana orientiert. So wird in einer nächsten Version des DDB-Portals für jedes Einzelobjekt ein Rechtelabel angezeigt, das zu einer Auskunftsseite mit Lizenzinformationen weiterleitet.

Nach Abschluss dieser Arbeiten sind die organisatorischen und technischen Voraussetzungen gegeben, damit die DDB produktiv und effizient Daten an Europeana weitergeben kann. Die Datenlieferung an Europeana ist eine Serviceleistung der DDB, die insbesondere kleineren Einrichtungen hilft, die nicht über die technische Infrastruktur verfügen, ihre Bestände selbst in das europäische Kulturportal zu integrieren.

Zur Erfüllung des öffentlichen Auftrages, den deutschen Beitrag für die Europeana zu erbringen, gehört nicht nur die Datenlieferung an Europeana. Die DDB ist auf verschiedenen Ebenen im Netzwerk der Europeana aktiv und positioniert sich als nationaler, spartenübergreifender Aggregator. Im Herbst dieses Jahres wird ein von der DDB organisierter Workshop mit anderen nationalen Partnern der Europeana stattfinden, der den Teilnehmern ein gemeinsames Forum für den Austausch ihrer Erfahrungen und Best Practices bietet.

Gemeinsamer Workshop von DDB und Europeana geplant

Personalisierung

Um die DDB für ihre Benutzer als Recherchewerkzeug aufzuwerten, wurde ein personalisierter Bereich geschaffen. In der ersten Ausbaustufe dieser neuen Funktion kann der Benutzer einzelne Suchtreffer als Favoriten speichern. Diese sind dann im personalisierten Bereich unter »Meine Favoriten« in Listenform zugänglich, können dort verwaltet und per E-Mail versandt werden.

Zur Nutzung der Personalisierungsfunktion muss man sich entweder direkt bei der DDB registrieren oder sich via OpenID[9] mit seinem Google™- oder Yahoo™-Konto anmelden[10].

Bequeme 1-Click-Anmeldung mittels OpenID

In der nächsten Ausbaustufe wird der Benutzer eigene Favoritenlisten anlegen und selbst gewählte Namen für diese Listen vergeben können. Auf diese Weise lassen sich einschlägige Treffer für verschiedene Rechercheprojekte komfortabel verwalten. Weiterhin wird es die Möglichkeit geben, Suchanfragen in ähnlicher Form wie die Favoriten zu speichern, verwalten und verteilen zu können.

Anwendungsprogrammierschnittstelle – API

Mit der Veröffentlichung einer API[11] geht die DDB einen weiteren wichtigen Schritt auf ihrem Weg zu einer Plattform, die digitale Inhalte aus Kultur und Wissenschaft für Mensch und Maschine zugänglich und mit den Methoden des Semantischen Webs verknüpfbar macht.

Neben die Suchfunktion im Portal der DDB tritt die Möglichkeit, als Nutzer eigene Anwendungen zu entwickeln, die auf den Datenbestand der DDB zugreifen. Das können z. B. Webseiten oder mobile Apps sein, die über die API mit der DDB kommunizieren.

Nutzer können eigene Anwendungen entwickeln

Die DDB-API ist als REST[12]-Dienst ausgestaltet. Es werden Methoden bereitgestellt, die aus anderen Programmen mittels des HTTP-Protokolls über das Internet aufgerufen werden können. Zurzeit umfasst die API 25 Methoden, die die in der DDB befindlichen Daten in ganz unterschiedlicher Weise zur Verfügung stellen. Umfang und Format der Daten sind dabei durch den Verwendungszweck bestimmt, für den die jeweilige Methode gedacht ist. Beispielsweise sind für Trefferlisten wenige, kurze Informationen ausreichend, während der DFG-Viewer[13] einen unveränderten METS-Container aus den Originaldaten benötigt.

Per API werden nur die in der DDB gespeicherten Metadaten und digitalen Derivate ausgeliefert, nicht die bei den Datengebern vorgehaltenen qualitativ hochwertigeren Digitalisate.

Der Zugriff per API ist nicht dafür ausgelegt, in großem Umfang Daten aus der DDB zu »saugen«. Dem wird in den zugehörigen Nutzungsbedingungen Rechnung getragen. Darüber hinaus behält sich die DDB im Einzelfall eine Drosselung des

Datenabflusses auf von ihr festgelegte Geschwindigkeiten oder ein bestimmtes Volumen vor.

Anonyme Nutzung nicht zulässig

Deshalb wird es auch keine anonyme Nutzung der API geben[14]. Interessierte Entwickler können sich online registrieren und erhalten daraufhin einen Schlüsselcode, einen sogenannten »API Key«, der bei jedem Zugriff via API übertragen werden muss. In einem gestaffelten Beta-Programm werden sich nach und nach immer weitere Nutzerkreise als Entwickler registrieren können, bis um den Jahreswechsel 2013/14 die Registrierung für jedermann freigegeben werden wird. Die DDB plant für diesen Herbst einen oder mehrere Entwickler-Workshops. Interessenten aus den bei der DDB registrierten Einrichtungen, die über grundlegende Programmierkenntnisse verfügen, sind herzlich eingeladen, sich bei der Servicestelle der DDB zu melden[15], falls sie an einem solchen Workshop teilnehmen wollen.

Entwickler-Workshops im Herbst geplant

Eine erste, sehr anspruchsvolle Anwendung der DDB-API ist mit dem Archivportal-D bereits in der Entwicklung[16]. Weitere Anwendungen dieser Art, die einen besonderen Zugang oder eine spezielle Perspektive auf den Datenbestand der DDB bieten, sind denkbar. Daneben erscheinen ansprechende Visualisierungen der semantischen Beziehungen zwischen den in der DDB enthaltenen Objekten, weiterhin mobile Anwendungen, die eine haptische Navigation durch die DDB erlauben, sowie statistische Auswertungen der in der DDB gespeicherten Metadaten als attraktive Anwendungen, die auf Basis der API programmiert werden könnten.

Es ist zu erwarten, dass die überraschendsten und überzeugendsten Ideen für die Anwendung der API von den Nutzern und den Datengebern der DDB kommen werden. Die einen wissen am besten, wie die DDB für sie zu nutzen sein sollte, und die anderen wissen am besten, welche Schätze in den Datenbeständen der DDB zu entdecken sind.

Ausblick

Nachdem die erste Jahreshälfte von den vorstehend dargestellten internen, nicht im Portal sichtbaren

BIS-C 2000
4th. generation
Archiv- und Bibliotheks-InformationsSystem
DABIS.eu
Gesellschaft für Datenbank-InformationsSysteme
DABIS.eu - alle Aufgaben - ein Team
Synergien: WB-Qualität und ÖB-Kompetenz
Software: Innovation und Optimierung
Web - SSL - Warenkorb und Benutzeraccount
Web 2.0 und Cataloque enrichment
Verbundaufbau und Outsourcing-Betrieb
Archiv Bibliothek Dokumentation
singleUser System multiUser
Lokalsystem und Verbund
multiDatenbank multiServer
multiProcessing multiThreading
skalierbar stufenlos
Unicode multiLingual
Normdaten redundanzfrei
multiMedia Integration
Software - State of the art - flexible
Über 23 Jahre Erfahrung und Wissen
Leistung Sicherheit
Standards Offenheit
Stabilität Verläßlichkeit
Generierung Adaptierung
Service Erfahrenheit
Outsourcing Support
Dienstleistungen Zufriedenheit
GUI-Web-Wap-XML-Z39.50-OAI-METS
Portale mit weit über 17 Mio Beständen
http://Landesbibliothek.eu http://bmlf.at
http://OeNDV.org http://VThK.eu
http://VolksLiedWerk.org http://bmwfj.at
http://Behoerdenweb.net http://wkweb.at
DABIS GmbH
Heiligenstädter Straße 213, 1190 Wien, Austria
Tel. +43-1-318 9777-10 * Fax +43-1-318 9777-15
eMail: office@dabis.eu * http://www.dabis.eu
Zweigstellen: 61350 - Bad Homburg vdH, Germany / 1147 - Budapest, Hungary / 39042 - Brixen, Italy
Ihr Partner für Archiv-, Bibliotheks- und DokumentationsSysteme

Weiterentwicklungsarbeiten bestimmt war, wird der Schwerpunkt im zu Ende gehenden Jahr 2013 auf der Transformation der bestehenden Daten in das neue interne Datenformat, dem Re-Ingest dieser Daten und dem Einspielen der Bestände neuer Datengeber liegen. Gemäß der Content-Strategie der DDB wird die Servicestelle in enger Abstimmung mit dem Kompetenznetzwerk auf die Ausgewogenheit, Vielfalt und Güte der neuen Daten achten.

Die Deutsche Digitale Bibliothek auf der Frankfurter Buchmesse

Besuchen Sie die Deutsche Digitale Bibliothek auf der Frankfurter Buchmesse vom 9. bis 13. Oktober 2013. Sie finden uns in Halle 4.2 am Gemeinschaftsstand der Deutschen Nationalbibliothek und der DDB (Standnummer L73).

Anmerkungen

1 <http://www.deutsche-digitale-bibliothek.de>

2 S. Baumgart, Denise; Schulze, Francesca: Die Servicestelle der Deutschen Digitalen Bibliothek. In: Dialog mit Bibliotheken 24 (2012) 2, S. 8 - 12.

3 Über das Kompetenznetzwerk der Deutschen Digitalen Bibliothek informiert die Website: <http://www.deutsche-digitale-bibliothek.de/content/competence-network/>

4 Die Dokumentation des Europeana Data Model findet sich unter: <http://pro.europeana.eu/edm-documentation>

5 Über die Eingangsformate der DDB informiert die Website: <http://www.deutsche-digitale-bibliothek.de/content/ddb/data_delivery>

6 <http://www.dnb.de/DE/Standardisierung/GND/gnd_node.html>

7 <http://www.europeana.eu>

8 <http://creativecommons.org/publicdomain/zero/1.0>

9 OpenID ist ein Open-Source-Protokoll zur Nutzer-Authentifizierung im Internet, s. <http://openid.net/> Es hat den Vorteil, dass sich der Nutzer nur einmal bei einem OpenID-Provider mit allen seinen Daten registrieren muss. Zur Sicherheit von OpenID s. das SOA-Kompendium des BSI v2.0 <https://www.bsi.bund.de/DE/Themen/weitereThemen/SOA/StudienPublikationen/Studien_Publikationen.html> Abschnitt 6.4.4.4

10 Diese beiden OpenID-Provider wurden für eine Phase ausgewählt, in der Erfahrungen mit diesem Anmeldeverfahren gesammelt werden sollen. Die Auswahl ist weder abschließend noch unabänderbar.

11 englisch: Application Programming Interface, deutsch: Anwendungsprogrammierschnittstelle, statt »die API« wird im Deutschen häufig auch »das API« verwendet

12 s. z. B. <https://de.wikipedia.org/wiki/Representational_State_Transfer>

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

14 Für die Zwecke einer Linked-Open-Data-Vernetzung der DDB wird eine eigene technische Lösung geschaffen, die mit den Sicherheitsbedürfnissen und Rechtevorgaben der DDB-Datengeber vereinbar ist.

15 per E-Mail an: service@deutsche-digitale-bibliothek.de

16 <http://www.landesarchiv-bw.de/web/54267> und Verweise.