
Geodatenkatalog-DE – Architektur und Inhalte

Jürgen WALTHER

Zusammenfassung

Die Infrastructure for Spatial Information in Europe (INSPIRE) kann nicht alle Fragestellungen bezüglich Metadaten und deren Bereitstellung im Sinne der Nutzeranforderungen beantworten. Primäres Ziel ist dort der Aufbau der Infrastruktur. Wie kann ich als Nutzer über Internetanwendungen intuitiv und komfortabel nach Daten suchen, mir wichtige Informationen in verständlicher Form abholen und die beschriebenen Geodaten visualisieren? Welche Rolle spielt in diesem Zusammenhang das Publish-Find-Bind-Paradigma? Wie kann eine Daten-Service-Kopplung umgesetzt werden? Diese Fragen sind so alt wie das Thema Geodateninfrastruktur und es gibt verschiedene Antworten darauf. Eine einheitliche Sicht gibt es noch nicht.

Dieser Beitrag gibt mögliche Antworten auf o. g. Fragestellungen und die Frage, inwieweit eine einheitliche Sicht notwendig ist. Er stellt die Architektur und Organisation des Geodatenkatalog-DE vor und behandelt Fragen zu Metadatenfeldern im Kontext von INSPIRE und ISO. Es handelt sich also um einen Praxisbericht zur Umsetzung von Standards und Nutzeranforderungen.

Geodatenkatalog-DE

Organisation und Grundlagen

Das GDI-DE Modellprojekt Geodatenkatalog-DE startete mit dem Beschluss des Lenkungsremiums GDI-DE vom 15.05.2007 zur Erstellung eines Projekt- und Ablaufplanes. Nach Vorliegen des Planes wurde am 17.08.2007 die Umsetzung des Modellprojektes durch das Lenkungsremium GDI-DE beschlossen. Details zur Organisationsstruktur der GDI-DE finden sich unter: <http://www.gdi-de.org>. Unter Nutzung dieser Strukturen wurde eine Projektgruppe gebildet und alle deutschen Katalogbetreiber der Behörden und der Forschung kontaktiert. Die Koordinierung der regionalen Metadatenbestände liegt in der Verantwortung der Bundesländer. Somit ist aus Sicht des Geodatenkatalog-DE die Einbindung von Katalogen auf den Bund und die Länder beschränkt.

Der Geodatenkatalog-DE leistet mittels einer performanten, standardisierten, harmonisierten und konsolidierten Metadatenbereitstellung als zentrale Komponente einen weiteren Beitrag zum Aufbau der Geodateninfrastruktur Deutschland (GDI-DE). Der übergeordnete Plan, in den sich der Geodatenkatalog-DE einfügt, ist durch das Architekturkonzept der GDI-DE (GDI-DE 2010) vorgegeben. Die zugrundeliegenden Standards sind ebenfalls im Dokument zu den Standards und Architekturen für E-Government-Anwendungen (SAGA) aufgenommen (SAGA 2008). Im internationalen Umfeld wird der Geodatenkatalog-DE die zentrale Bereitstellung von Metadaten für INSPIRE und GEOSS realisieren.

Zielsetzungen

Metadaten sind für die Dokumentation und Recherche von Geoinformationen unerlässlich. Sie schaffen Transparenz über Geodatenbestände und ermöglichen eine effiziente Recherche.

Um einen Gesamtüberblick zu bekommen, ist es notwendig, Katalogdienste von Bund und Ländern in einem Verbund zusammenzuführen. Die Bereitstellung von Metadaten der kommunalen Dienste soll innerhalb der Bundesländer koordiniert werden. Thematisch stehen insbesondere die in der Nationalen Geodatenbasis (NGDB) und INSPIRE geforderten Inhalte im Fokus.

Die Zusammenführung der Metadaten im Geodatenkatalog-DE ermöglicht eine Konsolidierung sowie Überprüfung der Konformität und Aktualität, eine Indexierung und das fachübergreifende Ranking.

Ein Gesamtüberblick über alle Geodatenbestände in Deutschland ist für eine schnelle Internetrecherche, effizienten Datenbezug und Synergie Effekte von besonderer Bedeutung. Metadaten zu Geoinformationen zu einem bestimmten Thema werden aus verschiedenen Fachbereichen verfügbar gemacht.

INSPIRE schreibt die Bereitstellung von Metadaten über Katalogdienste (Suchdienste) vor (INSPIRE 2009). Metadaten müssen gemäß einem INSPIRE-Profil (INSPIRE 2008) bereitgestellt werden. Dies beinhaltet die Bereitstellung von Metadaten zu Geodaten und zu Geodiensten. Die Katalogdienste müssen der Durchführungsbestimmung Netzdienste (INSPIRE 2009) entsprechen.

Der Geodatenkatalog-DE muss folgende Anforderungen erfüllen:

- Erfüllung der Informationspflicht und Bereitstellung von Metadaten gemäß der INSPIRE-Richtlinie
- Anforderungen aus GDI-DE, insbesondere Architekturkonzept
- Informationsbereitstellung für Verwaltung, Wissenschaft, Wirtschaft und Bürger
- Verknüpfung mit Darstellungsdiensten im Sinne des Publish-Find-Bind Prinzips

Architektur und Aufbauprozess

Die Architektur des Geodatenkatalog-DE ist darauf ausgerichtet, über einen Katalogdienst einen zusammengeführten Metadatenbestand bereitzustellen. Der Geodatenkatalog-DE integriert die folgenden Komponenten:

- Dezentrale Kataloge der Projektpartner
- Crawling Komponente zum Einsammeln der dezentralen Daten
- Filter Mechanismus zum Aussortieren von Dubletten
- Index für die schnelle Suche
- XML-Dateiablage für die Bereitstellung der kompletten Metadaten
- Katalogdienst konform zu ISO Application Profile 1.0 (OGC 2007)
- Recherche Anwendung zur Suche auf dem zentralen Datenbestand
- Ranking Mechanismus zur Sortierung der Suchergebnisse
- Testumgebung für ISO Application Profile 1.0 (AP ISO 1.0)

Abbildung 1 veranschaulicht die Architektur. Auf der Ebene der Metadatenerezeuger befinden sich die Datenbestände, die von den Datenbereitstellern über AP ISO 1.0 konforme Schnittstellen zur Verfügung gestellt werden. Bereits an dieser Stelle wird im Sinne eines Aufbauprozesses eine Qualitätssicherungsebene eingezogen. Die Datenbereitsteller müssen ihren Suchdienst mithilfe eines AP ISO 1.0 Tests auf Konformität prüfen. Die Testumgebung wurde im Rahmen des Modellprojektes Geodatenkatalog-DE erstellt und basiert auf der Team Engine des OGC (www.gdi-de.org)

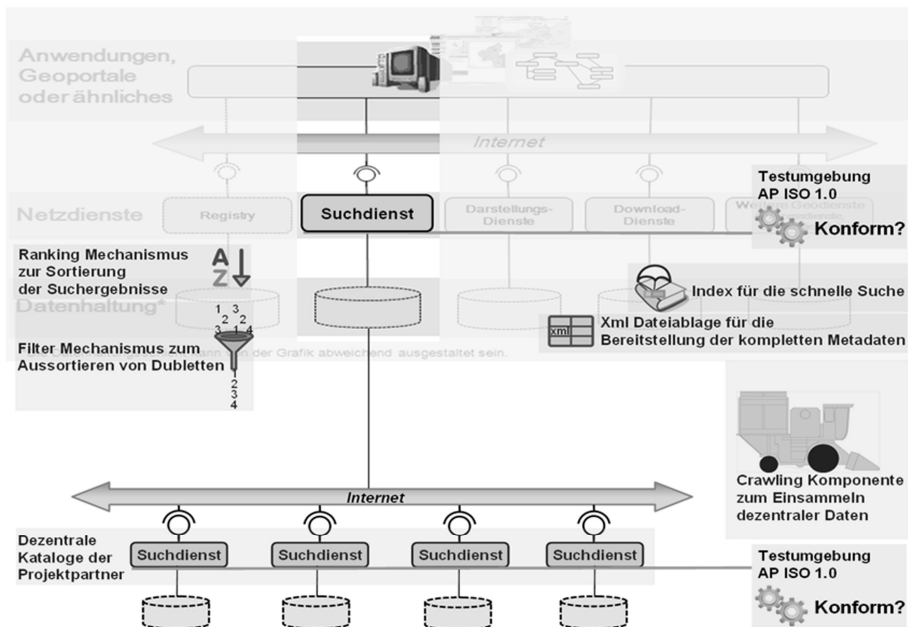


Abb. 1: Architektur des Geodatenkatalog-DE

Publish-Find-Bind

Die wichtigste Herausforderung der nächsten Jahre wird es sein, die inhaltliche Qualität der Daten und ihre Verknüpfung zu verbessern.

Dies sind die entscheidenden Voraussetzungen für eine nutzerfreundliche Bereitstellung von Geoinformationen im Sinne des Publish-Find-Bind, also des Weges von der Veröffentlichung über das auffindbar machen bis zur Verknüpfung zu i. d. R. Kartendiensten (und zurück). Die wesentliche Grundlage hierfür sind Metadaten.

Ein Blick auf die Daten in einer Geodateninfrastruktur zeigt folgende Metadatenbestände:

- Metadaten zu Geodaten
- Metadaten zu Netzdiensten
- GetCapabilities zu Netzdiensten

Diese Metadatenbestände liegen i. d. R. noch getrennt (also unverknüpft) vor. Sie werden oft von unterschiedlichen Stellen/Personen gepflegt und sind daher nur selten konsistent.

Um dem Nutzer eine Navigation von der Suche über die Metadaten zu den Kartendiensten und auch vom Kartendienst zu den zugehörigen Metadaten zu ermöglichen, sind bestimmte Übereinstimmungen innerhalb der o. g. Metadaten notwendig:

- Eindeutiger Resource Identifier innerhalb der Daten-Metadaten
- Vorhandensein dieses Resource Identifiers im operatesOn-Tag des Service-Metadaten-satzes (Verknüpfung)
- Angabe der Parameter für den Aufruf eines Kartendienstes
- Verlinkung innerhalb der GetCapabilities eines Kartendienstes (im Dienst- und Layerbereich) zu den zugehörigen Metadaten

Wenn diese Voraussetzungen einheitlich in allen Katalogen erfüllt sind, kann eine beliebige Anwendung/ein Geoportal das Publish-Find-Bind visualisieren (Abb. 2).

Wie exakt diese Voraussetzungen erfüllt werden, muss z. T. noch über Konventionen beschrieben werden, die durch die Standards von ISO 19115 (ISO 2003), ISO 19119 (ISO 2005), ISO 19139 (ISO 2007) und OGC AP ISO 1.0 (OGC 2007) sowie INSPIRE (s.o.) noch nicht definiert sind. Hier ist insbesondere der Aufruf von Layern eines Dienstes aus den Metadaten heraus zu nennen.

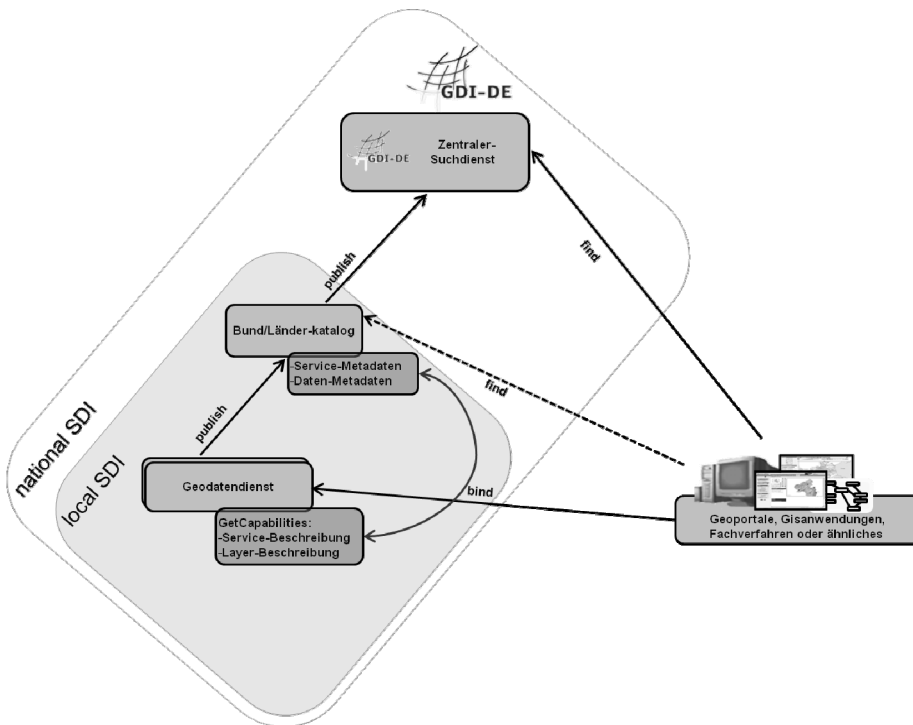


Abb. 2: Publish-Find-Bind

Fazit und Ausblick

Der Geodatenkatalog-DE stellt als zentrale Komponente der GDI-DE ein breites Spektrum an Metainformationen bereit. Unter den Gesichtspunkten der Nutzerorientierung ist es allerdings sinnvoll, alle existierenden Geo-Metadaten recherchierbar zu machen. Die Umsetzung dieser quantitativen Anforderung, wie auch der die Anforderungen an die Datenqualität wird weiterhin vor allem eine organisatorische Herausforderung bleiben.

Die technischen Grundlagen des Geodatenkatalog-DE – wie die vieler anderer Katalogdienste auch – sind geeignet, als Dreh- und Angelpunkt innerhalb des Publish-Find-Bind (also des Veröffentlichen, Auffinden und Visualisieren von Geodaten) zu agieren.

Literatur

- GDI-DE – GEODATENINFRASTRUKTUR DEUTSCHLAND (Hrsg.) (2010): Architektur der Geodateninfrastruktur Deutschland, Version 2.0. <http://www.gdi-de.org/>.
- INSPIRE (2008): Verordnung (EG) Nr. 1205/2008 der Kommission vom 3. Dezember 2008 zur Durchführung der Richtlinie 2007/2/EG des Europäischen Parlaments und des Rates hinsichtlich Metadaten. <http://inspire.jrc.ec.europa.eu/>.
- INSPIRE (2009): Verordnung (EG) Nr. 976/2009 der Kommission vom 19. Oktober 2009 zur Durchführung der Richtlinie 2007/2/EG des Europäischen Parlaments und des Rates hinsichtlich der Netzdienste. <http://inspire.jrc.ec.europa.eu/>.
- ISO/TC 211 (2003): International Standard ISO 19115 Geographic information – Metadata. <http://www.isotc211.org/>.
- ISO/TC 211 (2005): International Standard ISO 19119 Geographic information – Services. <http://www.isotc211.org/>.
- ISO/TC 211 (2007): International Standard ISO 19139 Geographic information – Metadata – Implementation specification. <http://www.isotc211.org/>.
- OGC (2007): OpenGIS Catalogue Service Specification 2.0.2 – ISO Metadata Application Profile. <http://www.opengeospatial.org/>.
- SAGA (2008): Standards und Architekturen für eGovernment-Anwendungen, Version 4.0, Schriftenreihe der KBSt, März 2008. http://cio.bund.de/DE/Standards/SAGA/saga_node.html.