

Aufbau zentraler Komponenten der GDI-DE mit freier Software

Olaf Knopp, WhereGroup GmbH & Co. KG

Einführung

Im Auftrag des Bundesamt für Kartographie und Geodäsie (BKG) ist die WhereGroup zur Zeit mit der Implementierung zweier zentraler Komponenten der Geodateninfrastruktur Deutschland (GDI-DE) beschäftigt. Dies sind der „Geodatenkatalog-DE“ sowie die Migration der Mapserver-Komponenten des GeoPortal.Bund.

Beide Projekte sind zentrale Bausteine, die an das Geoportal.Bund als auch das Geoportal-DE angeschlossen werden sollen.

Ziel des Geodatenkatalog-DE ist der Aufbau eines zentralen Dienstes, der Katalogdienste bzw. deren Daten zusammenführt, an andere Strukturen abgibt und Recherchen erlaubt. Die Software soll Transparenz über die in Deutschland vorhandenen Geodatenbestände schaffen und die Weiterleitung zu entsprechenden Geodiensten ermöglichen. Das Modul umfasst Software zur Zusammenführung, Konsolidierung und Abgabe von Metadaten, ein entsprechendes Recherchewerkzeug sowie einen INSPIRE-konformen Metadateneditor zur Erfassung eigener Datensätze. Die Zusammenführung der Katalogdienste erfolgt über vorhandene Standard-Schnittstellen (OGC CSW).

Ziel dieser Konsolidierung ist es im Wesentlichen, Mehrfacheinträge zu bereinigen und eine zentrale Schnittstelle zu den Metadatenbeständen der GDI-DE zu schaffen. Die Abgabe erfolgt konform zu den Vorgaben der europäischen Initiative zur Schaffung einer Geodateninfrastruktur in der Gemeinschaft (INSPIRE) und der GDI-DE.

Die Mapserver-Komponente stellt als OGC-kompatibler Internet Kartenserver, welcher den Anforderungen von INSPIRE und GDI-DE genügt, dar und soll ebenfalls an Geoportal.Bund und später an das Geoportal-DE angeschlossen werden.

Die Hauptfunktionen dieses Moduls umfassen im Wesentlichen die Integration von Geodatendiensten und deren weitere Bereitstellung für Visualisierungskomponenten oder als INSPIRE-/GDI-DE-konforme Geodatendienste. Zu diesem Zweck können verschiedene WMS oder einzelne Layer zusammengefasst und konfiguriert werden, um sie in Form kaskadierender Dienste als „Virtuellen WMS“ zu veröffentlichen.

Die Umsetzung der beiden Module erfolgt ausnahmslos mit freier Software:

- Geodatenkatalog-DE
 - GeoNetwork opensource
 - Mapbender inkl. MeTaDor-Metadateneditor
 - PostgreSQL/PostGIS
- Mapserver-Komponente
 - MapProxy
 - Mapbender
 - PostgreSQL/PostGIS

Der Abschluss der beiden Projekte ist für das zweite Quartal 2011 geplant.

Technische Umsetzung

Da beide Module mit Mapbender eine gemeinsame Client-Komponente einsetzen und zudem ein identisches Rechte- und Rollenkonzept verwenden, werden Teile der Infrastruktur zentral bereitgestellt. Während für jedes System ein separater Dienste-Server aufgesetzt wurde, erfolgt die Speicherung der Daten (Datenbank Geonetwork, gemeinsame Datenbank Mapbender, Tilecache) auf einem zentralen Datenbank-Server, auf den beide Komponenten zugreifen.

Folgende Komponenten sind für die einzelnen Server vorgesehen:

Server Geodatenkatalog:

- Geonetwork [1] (ohne DB)
- Mapbender [2] (ohne DB)

Server Mapserver:

- Mapbender (ohne DB)
- MapProxy [3]

Datenbankserver

- PostgreSQL [4]
- PostGIS [5]

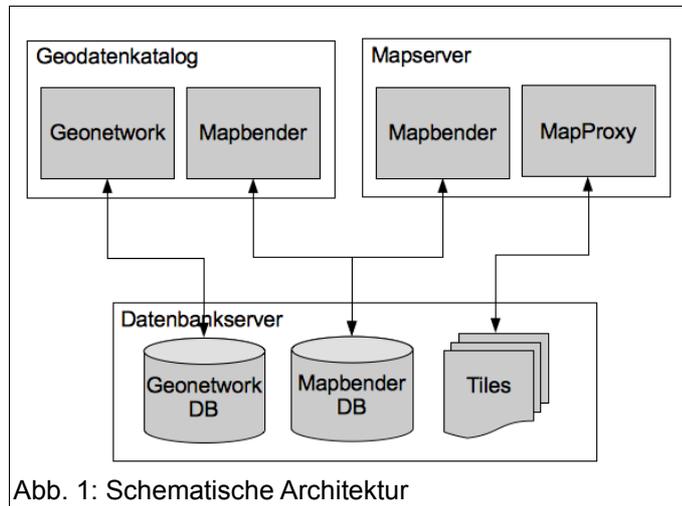


Abb. 1: Schematische Architektur

Die Serverarchitektur wird in Abbildung 1 schematisch dargestellt.

Geodatenkatalog-DE

Die Aufgabe des Geodatenkatalog-DE ist die Zusammenführung, Konsolidierung und Bereitstellung aller Metadatenbestände der GDI-DE. Für die ersten 3 Jahre wird von etwa 50 Katalogdiensten und rund 300.000 Metadatenätzen ausgegangen. Zur Zeit stehen im System rund 70.000 Metadatenätze aus ca. 20 Diensten zur Verfügung.

Zentrale Komponente des Katalogs ist GeoNetwork opensource als leistungsstarker Metadatenkatalog. Daneben werden folgende Anforderungen umgesetzt:

Harvesting Crawler:

Aufgabe des Crawlers ist die zentrale Konfiguration und Administration der einzelnen Harvestingaufträge. Die Aufgabe übernimmt das Harvesting Management des GeoNetwork, das zu diesem Zweck erweitert wurde. War es bisher nur möglich, das Intervall der einzelnen Harvestinganfragen anzugeben (in Tagen/Minuten/Stunden) können nun konkrete Termine, Uhrzeiten und Intervalle angegeben werden (z.B. „alle zwei Tage um 14.30 Uhr“).

XML Import

Neben der Übernahme von Metadaten aus OGC-kompatiblen Diensten (OGC CSW, OGC WMS/WFS Capabilities) soll auch der Import einzelner dateibasierter Metadatenätze möglich sein. Diese Funktionalität wird über MeTaDor, den in den Mapbender integrierten Metadateneditor, bereitgestellt.

Olaf Knopp: Aufbau zentraler Komponenten der GDI-DE mit freier Software

MeTaDor ermöglicht den Import, die Erzeugung und die Bearbeitung INSPIRE-konformer Metadatenätze sowie deren Veröffentlichung mittels CSW. Dabei fungiert die Software lediglich als Import- und Bearbeitungsclient für den eigentlichen Datenbestand, der in GeoNetwork gespeichert und über CSW abgegeben wird.

Die Metadatenlösung der WhereGroup

MeTaDor

Metadateneditorstartseite | Bearbeitung von Metadatenätzen zu Geodaten | Metadaten editieren - Ändern

Michel aus Lönneberga

intern Vor Speichern validieren? Speichern Validieren XML-Export

- Metadatenatz
- Identifikation
- Inhaltsbeschreibung
- Kategorisierung
- Schlüsselwörter
- Räumliche Angaben
- Zeitliche Angaben
- Dokumentation und Visualisierung
- Datenqualität
- Evaluierung
- Beschränkungen
- Vertriebsangaben
- Kontaktangaben

Basisinformationen zum Metadatenatz

Identifikator des Metadatenatzes:

Identifikator des übergeordneten Metadatenatzes:

Metadatenatzsprache:

Metadatenzeichensatz:

Letzte Überarbeitung der Metadaten:

Angaben zur Metadatenangabe

Bezeichnung der Spezifikation:

Version der Spezifikation:

Abb. 2: Erfassungsoberfläche von MeTaDor

Dublettenfilter

Durch die zentrale Konsolidierung der Metadatenbestände, wie die der Geodatenkatalog-DE vornimmt, kann es dazu kommen, dass ein Datensatz (eine UUID) über verschiedene Kataloge gefunden und geharvestet wird. Dies ist vor allem dann der Fall, wenn sich durch den Geodatenkatalog-DE geharvestete Dienste auch gegenseitig harvesten und so über redundante Datenbestände verfügen. Das System muss also über einen geeigneten Dublettenfilter verfügen, der identische Datensätze anhand der UUID identifiziert. GeoNetwork verfügt über einen solchen Mechanismus und aktualisiert einen Datensatz nur dann, wenn neue Daten mit identischer UUID und einem aktuelleren Bearbeitungsdatum identifiziert werden.

Rankingmechanismus

Zur effizienten Suche auf dem zentralen Metadatenbestand (s.u.) ist ein optimierter Rankingmechanismus vorgesehen. Da GeoNetwork die Metadaten als XML in einem Textfeld speichert ist der Einsatz von Datenbankmechanismen an dieser Stelle nicht zielführend. GeoNetwork setzt daher den Lucene-Index ein, ein Indizierungsmechanismus für Textdocumente. Lucene ermöglicht es, einzelne Felder des Dokumentes zu gewichten („boosten“), um aus Suchanfragen eine Relevanz (zwischen 0 und 100%) zu ermitteln. So wird bei einer Suche nach „Wasser“ das Vorkommen des Begriffs in den Abstracts oder dem Titel des Datensatzes höher bewertet, als das Vorkommen in Adress- oder Namensfeldern.

Rechercheanwendung

Zentrale Zugriffskomponente für den Metadatenbestand des Geodatenkatalog-DE ist die Web-Anwendung zur Recherche nach Metadaten. Sie ermöglicht die intuitive Suche nach Daten für Geodienste und Geodaten, die Detailansicht einzelner Datensätze, deren Download als XML sowie die Ansicht von Diensten in einem WebGIS-Client.

Der Suchclient wird wie der Metadateneditor mit Mapbender umgesetzt, um ein einheitliches Client-Framework zu verwenden.

Verschiedene Dienste wie der SNS (Semantischer Netzwerk Service) des Umweltbundesamtes [6] sowie der General Multilingual Environmental Thesaurus (GEMET) der Europäischen Umweltagentur [7] erleichtern die räumliche und Schlagwortsuche.

Die Suche erfolgt auf der CSW-Schnittstelle des GeoNetwork. Dieser Weg ermöglicht es, den in GeoNetwork konfigurierbaren Lucene-Index zu nutzen und macht eine zusätzliche Indizierung unnötig. Hinzu kommt, dass das optimierte Ranking (s.o.) auch für andere Clients nutzbar ist.

Die schematische Architektur des Geodatenkatalog-DE sowie die einzelnen Komponenten sind der Abbildung 4 zu entnehmen.



Abb. 3: Client für die einfache Suche

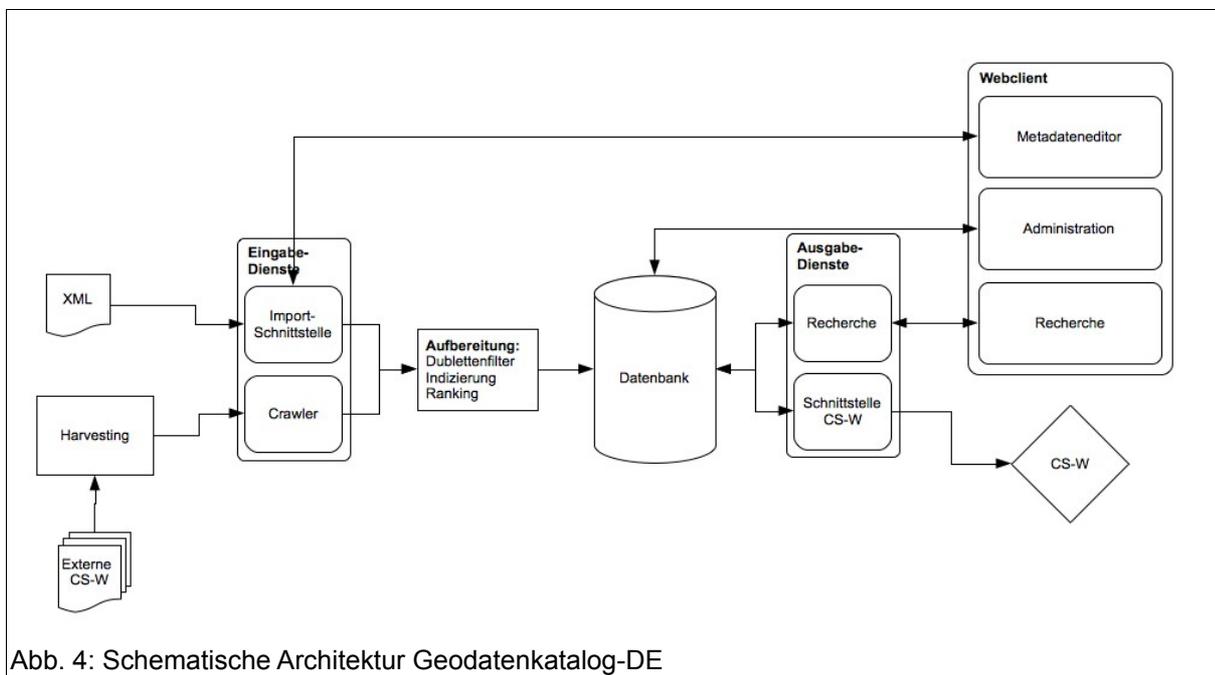


Abb. 4: Schematische Architektur Geodatenkatalog-DE

Mapserver-Komponente

Die Hauptfunktion des Mapservers ist die Integration von Geodatendiensten zur Bereitstellung im Geoportal.Bund und Geoportal-DE. Hierbei geht es in erster Linie darum, verschiedene Dienste in einem zentralen Repository zu verwalten und konfiguriert weitergeben zu können. Zur Verdeutlichung sei ein kurzes Beispiel genannt.

Viele Daten (z.B. Schutzgebiete) werden als OGC WMS von den zuständigen Behörden der Landesverwaltungen zur Verfügung gestellt. Ein Anwender, der deutschlandweite Informationen benötigt, ist also gezwungen 10 bis 20 einzelne Dienste mit unterschiedlichem Styling, abweichenden Maßstabsbereichen und gegebenenfalls uneinheitlichen Nutzungs- und Zugriffsbedingungen zu verwalten. Der zu entwickelnde Mapserver ermöglicht es, verschiedene Dienste zu konfigurieren und zu einem „Virtuellen WMS“ zusammenzufassen. Im vorliegenden Beispiel stehen dann die deutschlandweiten Daten in einem WMS mit vereinheitlichtem Styling zu Verfügung.

Zunächst war es angedacht, diese Funktionalität über den UMN MapServer in Form von kaskadierten WMS umzusetzen. Die Anforderungen für die GDI-DE gingen allerdings deutlich über die Möglichkeiten des UMN hinaus (z.B. Durchreichen der FeatureInfo-Requests), so dass sich für eine Umsetzung mit MapProxy entschieden wurde.

MapProxy-Komponente

Die MapProxy ist vor allem als Tile-Cache bekannt. Seine bisherigen Stärken lagen im Caching von WMS-Diensten mit einer deutlichen Performance-Verbesserung. Die Position innerhalb einer GDI zwischen Dienst und Client prädestiniert die Software allerdings dazu in Form eines klassischen Proxies weitere Funktionen zu übernehmen.

So wurden im Zuge der Entwicklung des Mapservers folgende Funktionen und Features im MapProxy implementiert:

- MultiMapProxy
 - Verwaltung beliebig vieler MapProxy-Konfigurationen.
 - Automatisches laden/neuladen
 - Verbesserung des Speicherverbrauchs/Ladezeit bei mehreren Konfigurationen
- WMS Handling
 - Verschachtelte WMS Layer (Gruppenlayer/Layertree)
 - Abfrage von Gruppenlayern
 - Zusammenfassen von kaskadierten Quell-WMS-Anfragen (eine Anfrage pro WMS-Server, auch bei mehreren Layern)
 - SLD Unterstützung für Quell WMS: sld (URL), sld_body (Text) oder lokale Datei
 - HTTP POST Anfragen für WMS Quellen mit SLD
 - statische LegendURLs pro Quell WMS Layer und pro MapProxy WMS Layer

Olaf Knopp: Aufbau zentraler Komponenten der GDI-DE mit freier Software

- Individuelle Timeouts für jede WMS Quelle
- Einblendung fehlgeschlagener Layer in WMS Antwort
- XSL Transformationen für eingehende und ausgehende FeatureInformationen
- *opacity* für jeden WMS Quelle zum Überlagern nicht-transparenter Layer
- *transparent_color* für jede WMS Quelle zum Umwandeln von konfigurierten Farbwerten in Transparenz

Die Erweiterungen stehen bereits in der aktuellen MapProxy-Version 1.0.0 zur Verfügung. Eine detaillierte Beschreibung findet man in der MapProxy Dokumentation [8].

Konfigurationsclient

Als Konfigurationsclient findet auch hier der Mapbender Verwendung. Die Konfiguration des MapProxy bzw. einzelner Virtueller WMS soll auf Basis der Anwendung „WMS-GUI Einstellungen“ des Mapbender entwickelt werden, die bereits heute die Möglichkeit bietet, die Darstellungsoptionen von WMS und einzelnen Layern individuell zu konfigurieren. Die Anwendung befindet sich zur Zeit im Planungsstadium.

Die schematische Architektur der Mapserver-Komponente ist der Abbildung 5 zu entnehmen.

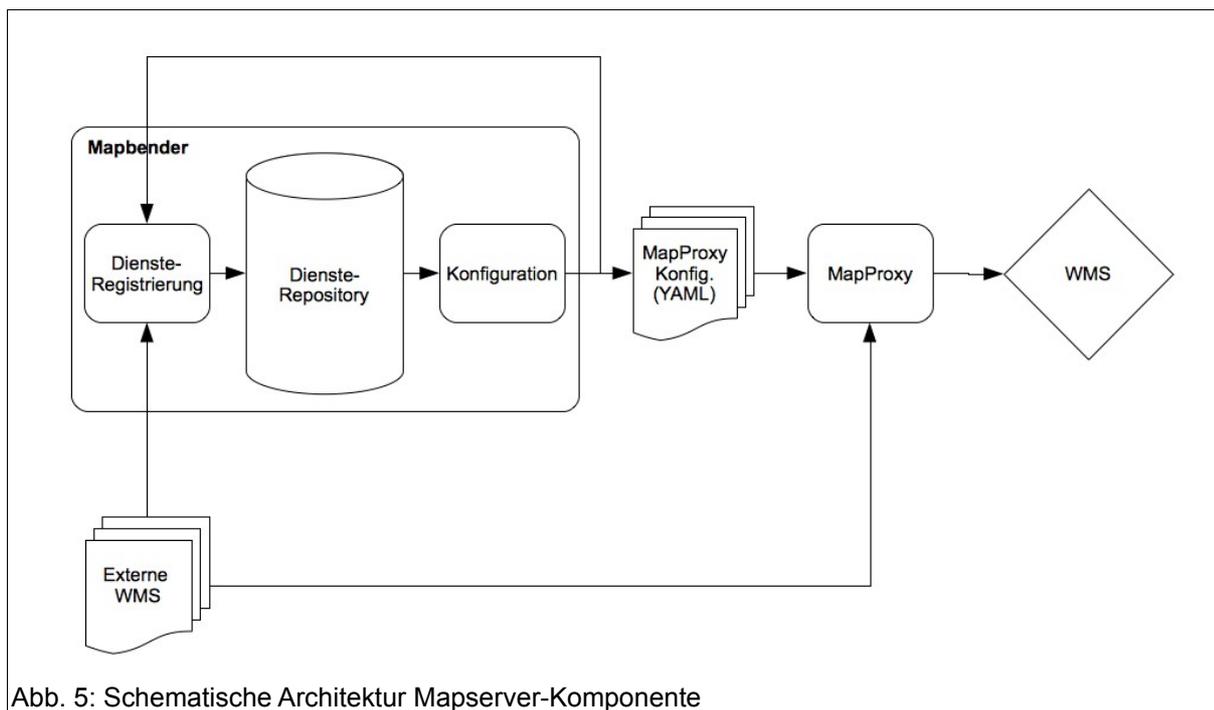


Abb. 5: Schematische Architektur Mapserver-Komponente

Fazit

Mit dem Geodatenkatalog-DE und den Mapserver-Komponenten werden zwei wichtige Bausteine der GDI-DE entwickelt, die einen zeitgemäßen und intuitiven Zugang zu den Meta- und Geodatenbeständen der GDI-DE bieten. Der Einsatz von freier Software ermöglicht die

Olaf Knopp: Aufbau zentraler Komponenten der GDI-DE mit freier Software

einfache Anpassung an die Anforderungen und Wünsche der Koordinierungsstelle und bietet größtmögliche Flexibilität bei der Einbindung der neuen Dienste in eine Vielzahl von Anwendungen und Portale

Hinzu kommen umfangreiche Erweiterungen der eingesetzten Softwarekomponenten (in erster Linie MapProxy), die zukünftig allen Nutzern zur Verfügung stehen.

Die beiden Entwicklungen Projekte sind daher gelungene Beispiele für Projekte mit einem Mehrwert für den Kunden und die Open Source Community.

Kontakt zum Autor:

Olaf Knopp
WhereGroup GmbH & Co. KG
Eifelstraße 8
53119 Bonn
0228/909038-27
olaf.knopp@wherogroup.com

Literatur

[1] <http://geonetwork-opensource.org>

[2] <http://www.mapbender.org>

[3] <http://www.mapproxy.org>

[4] <http://www.postgresql.org>

[5] <http://www.postgis.org>

[6] <http://www.semantic-network.de>

[7] <http://www.eionet.europa.eu/gemet>

[8] <http://mapproxy.org/docs/1.0.0/>