

Grenzübergreifende Web-GIS-Lösungen

Zusammenfassung

Der vorliegende Beitrag beschäftigt sich mit aktuellen Forschungsfragen im Bereich der Interoperabilität verteilter, heterogener GI Systeme. Die Heterogenität auf der Ebene der konzeptionellen Datenmodelle steht dabei im Mittelpunkt. Dieses Problem wird im Kontext der grenzübergreifenden, web-basierten Nutzung von Geodaten erörtert. Bestehende Ansätze zum Aufbau solcher grenzüberschreitender GIS-Anwendungen sowohl unter Verwendung von Raster- als auch Vektordaten werden bzgl. ihrer Leistungsfähigkeit und Praktikabilität verglichen. Vorgestellt wird ein neues Konzept für einen modellbasierten Ansatz für den web-basierten Zugriff auf verteilte, heterogene Geodaten, das derzeit in einem gemeinsamen Forschungsprojekt von der Eidgenössischen Technischen Hochschule Zürich (ETHZ) und der Technischen Universität München (TUM) entwickelt wird, basierend auf den Grundlagen der Service Oriented Architecture (SOA) und des Model Driven Architecture (MDA). Die Bodenseeregion dient als Testgebiet für das Projekt, in dem topographische Vektordaten aus Deutschland und der Schweiz verwendet werden.

Abstract

The paper addresses current research issues in the field of interoperability of heterogeneous GI systems. Special emphasis is placed on heterogeneity at the level of conceptual data models. This problem is discussed in the context of cross-border web-based GIS applications. Existing approaches for setting up such cross-border web-based GIS applications using both raster and vector data are compared, assessing their capabilities and practicability. A new concept of a Model Driven Architecture for accessing distributed, heterogeneous spatial data using web services currently being developed in a joint research project of the swiss Federal Institute of Technology Zürich and the Technische Universität München is presented, based on the fundamentals of Service Oriented Architectures (SOA) and the Model Driven Architecture (MDA). The Lake Constance region serves as cross-border test area for this project using topographic vector data from Germany and Switzerland.

aus Quellen in verschiedenen Ländern harmonisiert werden müssen, um sie nahtlos und konsistent kombinieren zu können, tritt dieses Problem besonders augenscheinlich zu Tage.

Die Harmonisierung von Daten ist zwar auch ein Haupt Gesichtspunkt von INSPIRE, aber „von den Mitgliedstaaten wird (...) nicht verlangt, das Format ihrer Raumdaten zu ändern; sie können stattdessen Schnittstellen bereitstellen, die heterogene Daten anhand eines einheitlichen Modells umwandeln“ (KOM 2004). Der Fokus liegt also nicht auf der vollständigen Harmonisierung der Daten zugrunde liegenden Datenmodelle, sondern eher auf Interoperabilität in einer Service Oriented Architecture (SOA), beispielsweise durch den Einsatz von Web Services, die Transformationen zwischen unterschiedlichen Datenmodellen ermöglichen.

In den bestehenden Standards des Open Geospatial Consortium (OGC) für Web Services (z.B. Web Feature Service), wird die syntaktische Heterogenität durch die Kapselung der internen Strukturen heterogener GI-Systeme unter Verwendung standardisierter Schnittstellen bewältigt. Die internen Strukturen, also auch die Datenmodelle bleiben dem Nutzer verborgen. Dies ermöglicht den einfachen Ad-hoc-Zugriff auf Geoinformationen aus heterogenen Quellen. Das Problem der semantischen Heterogenität wird von den bestehenden Web Services jedoch nicht gelöst. Um semantische Interoperabilität durch die Transformation zwischen verschiedenen Datenmodellen in OGC Web Services (OWS) zu ermöglichen, müssen letztere erweitert werden um Konzepte der Schematransformation / semantischen Transformation, die auf dem sog. Modellbasierten Ansatz (engl. Model Driven Architecture, MDA) beruhen. Genau dies ist Gegenstand des laufenden Forschungsprojektes „mdWFS“, in dem Know-how der ETHZ auf den Gebieten modellbasierter Datentransfer und semantische Interoperabilität sowie der TUM auf dem Gebiet OGC Web Services gebündelt wird, mit dem Ziel, die grenzüberschreitende Nutzung vor-

AUTOREN

Dr.-Ing. Andreas Donaubaue, Dipl.-Geogr. Astrid Fichtinger, Univ. Prof. Dr. Ing. Matthäus Schilcher, M. Sc. Florian Straub (TUM), Prof. Dr.-Ing. Alessandro Carosio, Dipl.-Math. Hans Rudolf Gnägi, Dipl.-Ing. Andreas Morf, Dipl.-Ing. Peter Staub (ETHZ)

1. Einführung

Geodateninfrastrukturen (GDI) auf regionaler, nationaler und internationaler Ebene stehen derzeit ganz oben auf der GI Agenda. Ein schneller, einfacher Zugang sowie die effiziente und nachhaltige Nutzung und Kombination von verteilten, heterogenen Geoinformatio-

nen über die Grenzen von GI-Systemen, Disziplinen, Branchen und Ländern hinweg sind heute wichtiger denn je und werden beispielsweise auch in dem Vorschlag für eine „Richtlinie des Europäischen Parlaments und des Rates zur Schaffung einer Raumdateninfrastruktur in der Gemeinschaft (INSPIRE)“ gefordert (KOM 2004). Um diese Ziele zu erreichen, müssen verschiedene Hindernisse überwunden werden. Eines der größten Hindernisse ist die Heterogenität der Daten auf der Ebene der Datenmodelle und der Semantik, ein Problem, das derzeit noch nicht genügend in der GI Standardisierung behandelt wird. In grenzübergreifenden, web-basierten GIS Anwendungen, für die heterogene Daten

Land	Dateneigentümer	topographische Vektordaten im M 1:25000	Beschreibung des Datenmodells in einer konzeptionellen Schemasprache
D	Landesamt für Vermessung und Geoinformation (Bayern)	ATKIS Basis-DLM (AAA-Modell)	in UML
D	Landesvermessungsamt Baden-Württemberg	ATKIS Basis-DLM (AAA-Modell)	in UML
CH	Bundesamt für Landestopografie (swisstopo)	VECTOR25 / GG25	in UML und INTERLIS
A	Bundesamt für Eich- und Vermessungswesen	DLM	nein

Tab. 1: Topographische Vektordaten im M 1:25000 in der Bodenseeregion

handener, verteilter Geodaten zu optimieren. Topographische Vektordaten im Maßstab 1:25000 aus der Bodenseeregion dienen als Testdaten für das Projekt.

Mögliche Anwendungsszenarien für den grenzübergreifenden, web-basierten Zugriff auf Geodaten sind beispielsweise grenzüberschreitende Standortanalysen (vgl. Huber 2005), grenzüberschreitende Raumentwicklung, Gewässerschutz, Hochwassermanagement, Einsatzsysteme für Rettungskräfte oder Tourismusinformationssysteme (vgl. Gläsel et al. 2006).

2. Bestehende Ansätze für die grenzüberschreitende, web-basierte Nutzung von Geodaten

2.1 Klassifikationsmerkmale

Ansätze für die grenzüberschreitende, web-basierte Nutzung von Geodaten können hinsichtlich möglicher Systemkonfigurationen beispielsweise nach den in Tab. 2 aufgeführten Kriterien klassifiziert werden. Dabei gibt es für jedes Kriterium verschiedene mögliche Aus-

prägungen. Im Folgenden werden unter „Quellsystemen“ die Geodatenbanken der Datenanbieter verstanden, unter „zentrales System“ eine Geodatenbank, in die Daten aus den Quellsystemen integriert werden und unter „Zielsystemen“ die Systeme, mit denen der Endanwender die Daten nutzt (z.B. Web-Client, Desktop-GIS, High-End-GIS, etc.).

Die Kombinationsmöglichkeiten der verschiedenen Ausprägungen ergeben eine Vielzahl von unterschiedlichen Systemarchitekturen. In Kapitel 2.2 werden fünf typische Kombinationen als Fallbeispiele ausgewählt und näher beschrieben. Für alle beschriebenen Fallbeispiele gelten die Rahmenbedingungen, dass Daten aus verschiedenen Ländern über die gleiche Art von Realwelt-Objekten (z.B. Topographie) auf der gleichen Maßstabsebene gemeinsam genutzt werden sollen.

2.2 Fallbeispiele

2.2.1 Integration von Rasterdaten in ein zentrales System

Die Integration von Rasterdaten in ein

zentrales System kann als der am wenigsten komplexe Fall zum Aufbau einer grenzübergreifenden Web-GIS-Anwendung angesehen werden. Als Beispiel kann hier der Bodensee Geodatenpool, ein Projekt der Vermessungsverwaltungen der Schweiz, Österreichs, Bayerns und Baden-Württembergs, herangezogen werden. In der ersten Projektphase wurden Rasterdaten (digitale Topographische Karten im Maßstab 1:50000) aus der Schweiz, Österreich und Deutschland in einen nahtlosen Rasterdatensatz integriert (vgl. Fig. 1). Die Datensätze der unterschiedlichen Vermessungsverwaltungen ähnelten sich bezüglich ihrer Struktur (thematische Layer) und ihrer Kartengrafik / Symbole / Legenden, basierten aber auf unterschiedlichen Raumbezugssystemen. Zur Erzeugung eines nahtlosen grenzübergreifenden Datensatzes wurden die Ausgangsdaten durch eine Transformation der unterschiedlichen Raumbezugssysteme nach ETRS89 / UTM 32 sowie durch eine Retuschierung der Kartengrafik / Symbole / Legenden in den überlappenden Gebieten an den Landesgrenzen harmonisiert. Ein Web-Client wurde entwickelt, der auf die zentrale Geodatenbank zugreift (Gläsel et al. 2006; Steudle 2006).

2.2.2 Integration von Vektordaten in ein zentrales System

Die Integration von Vektordaten in ein zentrales System kann als der „traditionelle“ Ansatz zur Kombination verteilter, heterogener Datenbestände gesehen werden. Der generelle Datenfluss ist vergleichbar mit dem bei Rasterdaten (s. Kap. 2.2.1), der Integrationsprozess ist bei Vektordaten jedoch komplexer. Im Wesentlichen können drei Komplexitätsstufen unterschieden werden, die im Folgenden am Beispiel von Daten über Verwaltungsgrenzen aus der Schweiz und Deutschland veranschaulicht werden.

■ Grafikorientierte Integration mit einer Harmonisierung von Datenformaten: Falls die Vektordatensätze der Schweizer und deutschen Verwaltungsgrenzen in grafikorientierten Vektorformaten wie DXF oder SVG vorliegen, können sie über entsprechende Import-Filter in ein zentrales System integriert werden. Da die Datensätze keine Informationen über die zugrunde liegenden Datenmodelle (z.B. über

Klassifikationsmerkmal	Ausprägungen
Datentyp (Quellsysteme, zentrales System, Zielsysteme)	■ Raster ■ Vektor (grafikorientiert, objektstrukturiert, ...)
Datenintegration aus Quellsystemen in ein zentrales System	■ Ja, zentrales System vorhanden ■ Nein, direkte Kopplung zwischen Quell- und Zielsystemen
Datenharmonisierung	■ Keine Harmonisierung ■ Harmonisierung der Raumbezugssysteme ■ Harmonisierung der Geometrie / Topologie ■ Harmonisierung der Kartengrafik / Symbole / Legenden ■ Harmonisierung des Datenmodelle
Herstellerabhängigkeit	■ Herstellerunabhängig (basierend auf Standards) ■ Herstellerspezifisch (proprietär)

Tab. 2: Klassifikationsmerkmale für die grenzüberschreitende, web-basierte Nutzung von Geodaten

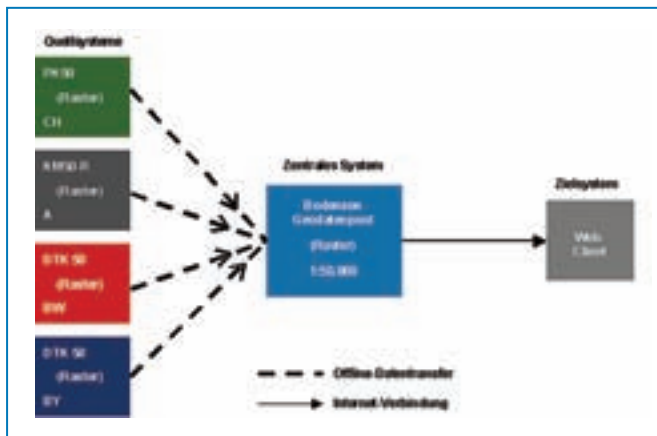


Fig. 1: Integration von Rasterdaten in ein zentrales System am Beispiel des Bodensee Geodatenpools

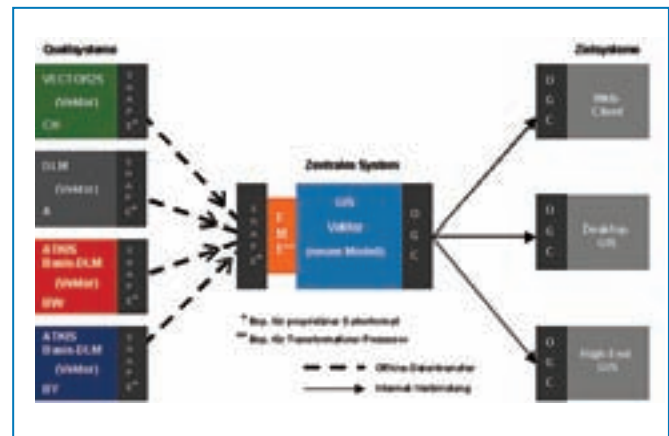


Fig. 2: Integration von Vektordaten in ein zentrales System mit Harmonisierung von Datenmodellen

semantische Unterschiede) beinhalten, kann der Nutzer die Schweizer und deutschen Daten nur visuell interpretieren. Er kann nur eine Harmonisierung von Raumbezugssystemen und / oder Kartengrafik / Symbolen / Legenden vornehmen, aber kaum komplexere Aufgaben durchführen. Im Prinzip sind die Möglichkeiten, die der grafikorientierte Vektordaten-Ansatz bietet, vergleichbar mit den Möglichkeiten des Rasterdaten-Ansatzes aus Kapitel 2.2.1.

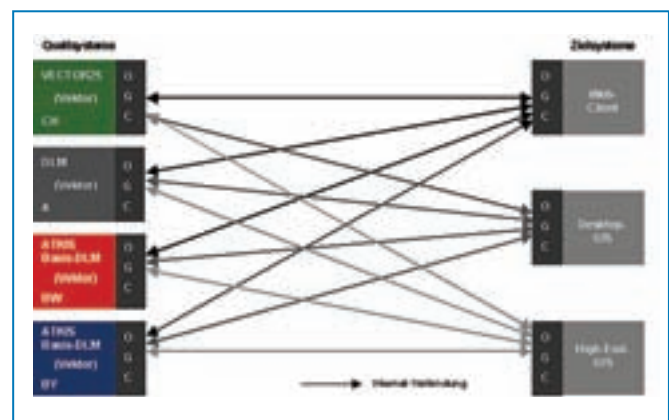
■ Objektstrukturierte Integration mit einer Harmonisierung von Datenformaten: Falls die Schweizer und deutschen Ausgangsdaten in objektstrukturierten Datenformaten zur Verfügung stehen (z.B. ESRI Shape files, OGC GML), können sie in ein zentrales System integriert werden, falls entsprechende Import-Filter vorhanden sind. Wenn die Daten ohne eine Harmonisierung der zugrunde liegenden Datenmodelle integriert werden, kann der Nutzer - wie im vorhergehenden Fall - die Daten visuell interpretieren. Zusätzlich kann der Nutzer aber i.d.R. auf mit der Geometrie verknüpfte alphanumerische Informationen zugreifen, die ihm die Durchführung von Analysen und die Interpretation der Semantik der Daten ermöglichen. Der Nutzer kann beispielsweise sehen, dass es in den Schweizer Daten vier Typen von Verwaltungsgebieten gibt („Land“, „Kanton“, „Bezirk“ und „Gemeinde“), wohingegen es in den deutschen Daten fünf Typen gibt („Nationalstaat“, „Bundesland“, „Regierungsbezirk“,

„Kreis“ und „Kommunales Gebiet“). Da die Datenmodelle in diesem Fall aber nicht harmonisiert sind, ist es umständlich, Abfragen über beide Ausgangsdatensätze zu formulieren. Wenn ein Nutzer beispielsweise Gemeinden in Deutschland finden will, die an Schweizer Gemeinden grenzen, muss er in der Abfrage die Ausdrücke „Gemeinde“ und „Kommunales Gebiet“ explizit verwenden, um als Resultat Geoobjekte von beiderseits der Landesgrenze zu erhalten.

■ Objektstrukturierte Integration mit einer Harmonisierung von Datenmodellen: Um eine Datenbasis zu erzeugen, die bezüglich des Dateninhaltes konsistent und nahtlos ist, muss der Nutzer die zugrunde liegenden Datenmodelle der Ausgangsdaten harmonisieren, beispielsweise durch die Definition von Beziehungen zwischen äquivalenten Konstrukten in den verschiedenen Modellen (z.B. zwischen der Schweizer „Gemeinde“ und dem deutschen „Kommunales Gebiet“). Dies kann auf unterschiedliche Weise erreicht werden, beispielsweise durch den Ein-

satz von Transformations-Prozessoren wie FME oder den sog. modellbasierten Datentransfer, der Konzepte der Model Driven Architecture (MDA) für den Datentransfer adaptiert. Das Prinzip des MDA in der Softwareentwicklung sieht den Entwurf eines exakten und formalen Modells der zu entwickelnden Software vor, das verwendet werden kann, um die eigentlichen ausführbaren Softwarekomponenten durch automatische Transformation zur generieren (vgl. Straub 2005). Beim modellbasierten Datentransfer können zusammen mit den Daten zusätzliche Informationen über das konzeptionelle Datenmodell übertragen werden, während beim formatbasierten Datentransfer nur das Transferformat beschrieben wird. Voraussetzung für einen modellbasierten Datentransfer ist, dass die konzeptionellen Datenmodelle auf exakte, formale und maschinenlesbare Weise in einer konzeptionellen Schemasprache (engl.: conceptual schema language) ►

Fig. 3: Direkte Kopplung zwischen Quell- und Zielsystemen mit OGC Web Services



Klassifikationsmerkmal	Integration von Rasterdaten	Integration von Vektordaten mit Harmonisierung von Datenformaten	Integration von Vektordaten mit Harmonisierung von Datenmodellen	Kombination von verteilten Systemen mit OGC Web Services
Datentyp im Quellsystem	Raster	Vektor (objektstrukturiert)	Vektor(objektstrukturiert)	Raster,Vektor (grafikorientiert, objektstrukturiert, ...)
Datentyp im zentralen System	Raster	Vektor (objektstrukturiert)	Vektor(objektstrukturiert)	
Datentyp im Zielsystem	Raster	Raster,Vektor (grafikorientiert, objektstrukturiert)	Raster,Vektor (grafikorientiert, objektstrukturiert)	Raster,Vektor (grafikorientiert, objektstrukturiert, ...)
Datenintegration aus Quellsystemen in ein zentrales System	Ja	Ja	Ja	Nein, direkte Kopplung zwischen Quell- u. Zielsystemen
Möglichkeit zur Harmonisierung von	Raumbezugssystemen, Kartengrafik / Symbole / Legenden	Raumbezugssystemen, Geometrie / Topologie Kartengrafik / Symbole / Legenden	Raumbezugssystemen, Geometrie / Topologie Kartengrafik / Symbole / Legenden, Datenmodellen	Raumbezugssystemen, Kartengrafik / Symbole / Legenden
Hersteller-abhängigkeit	Herstellerspezifisch	Herstellerspezifisch	Herstellerspezifisch	Herstellerunabhängig

Tab. 3: Vergleich der existierenden Ansätze für die grenzüberschreitende, web-basierte Nutzung von Geodaten

ge, CSL, z.B. UML) beschrieben sind. Ferner müssen Regeln zur automatisierten Ableitung eines Transferformates aus dem konzeptionellen Schema sowie Prozessoren, die das Transferformat lesen und schreiben können, vorhanden sein (vgl. Gnägi und Plabst 2004).

2.2.3 Direkte Kopplung zwischen Quell- und Zielsystemen mit OGC Web Services

Bei diesem Ansatz werden die Daten nicht in ein zentrales System integriert, sondern verbleiben in den Quellsystemen, in denen sie auch erhoben und gepflegt werden. Die Systeme können „ad hoc“ über das Internet/Intranet in einer SOA gekoppelt werden. Der Begriff SOA beschreibt das Konzept einer Systemarchitektur der allgemeinen IT, in der Funktionalität durch Services in einem Netzwerk, z.B. dem Internet, bereit gestellt wird. Die Komponenten einer SOA liegen in den meisten Fällen verteilt vor. Sie können lose gekoppelt werden, was sie wieder verwendbar macht. Auf die von den Services zur Verfügung gestellte Funktionalität kann über standardisierte Schnittstellen zugegriffen werden, die die internen Strukturen der zugrunde liegenden Systeme kapseln und so Interoperabilität zwischen den Systemen ermöglichen (Straub 2005). Im GIS-Bereich spezifiziert das OGC derartige Schnittstellen. Services, die diese Schnittstellen verwenden, werden als OGC Web Services bezeichnet. Beispiele sind der grafikorientierte Web

Map Service (WMS), der die Online-Darstellung von Karten ermöglicht oder der Web Feature Service (WFS), der den Zugriff auf objektstrukturierte Geodaten und (innerhalb bestimmter Grenzen) deren Analyse und Manipulation erlaubt.

In einem verteilten System auf Basis von OGC Web Services können heterogene Daten aus den verschiedenen Quellsystemen – bis zu einem gewissen Grad – bezüglich ihres Raumbezugssystems und ihrer Kartengrafik / Symbole / Legenden harmonisiert werden. Funktionalität zur Koordinatentransformation können entweder durch WMS / WFS Instanzen oder durch einen Web Coordinate Transformation Service (WCTS) zur Verfügung gestellt werden. Kartengrafik / Symbole / Legenden können mit Hilfe der Styled Layer Descriptor (SLD) Spezifikation harmonisiert werden. Eine Harmonisierung von Datenmodellen wird von den derzeitigen OGC Spezifikationen nicht unterstützt. Was grenzüberschreitende Abfragen betrifft, bietet die Verwendung eines WFS daher vergleichbare Möglichkeiten und Grenzen wie der Ansatz der objektstrukturierten Vektordatenintegration mit einer Harmonisierung von Datenformaten, der in Kapitel 2.2.2 beschrieben wird.

Wie mit Hilfe von OGC Web Services grenzüberschreitende GIS-Anwendungen aufgebaut werden können, illustrieren Jaenicke (2003) für den Nationalpark Bayerischer Wald und den angrenzenden tschechischen Sumava Nationalpark, sowie Riecken et al. (2003)

für die grenzüberschreitende Region Nordrhein-Westfalen und Niederlande. In diesen Fallbeispielen werden Daten von beiderseits der Grenze über WMS zur Verfügung gestellt.

2.3 Bewertung der existierenden Ansätze

Die Integration von Daten – seien es Vektor- oder Rasterdaten – aus unterschiedlichen Quellsystemen in ein zentrales System ist oft ein zeitaufwändiger und technisch anspruchsvoller Prozess (vgl. Huber 2005), der ein gewisses Maß an Expertenwissen voraussetzt. Die Daten müssen ggf. aus verschiedenen Quelldatenformaten in ein für das zentrale System lesbares Format konvertiert werden. Im Fallbeispiel der Integration von Vektordaten ist eine Formatkonvertierung meist komplexer als bei der Integration von Rasterdaten. Formatkonvertierungen sind oft auch nicht verlustfrei durchzuführen. Die Integration von Daten aus unterschiedlichen Quellsystemen in ein zentrales System impliziert überdies in den meisten Fällen eine redundante Datenhaltung. Der Datennutzer kann im Prinzip nur auf eine im zentralen System gespeicherte Kopie der Originaldaten und nicht auf die Originaldaten selbst zugreifen. Um ein gewisses Maß an Aktualität der Daten zu gewährleisten, muss der Prozess der Datenintegration in bestimmten Zeitintervallen wiederholt werden. Gleichwohl kann Datenintegration bei Daten, die sich nur innerhalb langer Zeitspannen verändern, ein praktikabler Ansatz sein.

Wenn die Datenintegration nur formatbasiert erfolgt, und der Betreiber des zentralen Systems keinen Zugriff auf eine maschinenlesbare Beschreibung der Datenmodelle hat, kann er nur „manuell“ einige Aspekte der Daten harmonisieren, wie beispielsweise die Kartengrafik / Symbole / Legenden, die Geometrie / Topologie (bei Vektordaten) und ggf. die Gruppierung in unterschiedliche Layer. Werden die Daten mit Hilfe des modellbasierten Datentransfers integriert, können auch die Datenmodelle und die Semantik harmonisiert werden. Dieser Ansatz ist leistungsfähiger als die formatbasierte Integration, aber auch komplexer in der Realisierung.

Der Ansatz „Direkte Kopplung zwischen Quell- und Zielsystemen mit OGC Web Services“ bietet Vorteile, indem er beispielsweise den Zeitaufwand für die Datenintegration erspart und einen Zugriff auf die Originaldaten anstatt möglicherweise veralteter Kopien ermöglicht. Dennoch weist er immer noch eine Reihe von Defiziten in Bezug auf Praktikabilität und Funktionalität auf (Donaubauer 2004). Bei einer Kombination von unterschiedlichen OGC Web Services können Konsistenzprobleme zwischen verschiedenen Typen von OGC Web Services oder verschiedenen Versionen einer OGC Spezifikation auftreten (z.B. WFS 1.0 und WFS 1.1). An der Technischen Universität München durchgeführte Tests zeigten überdies die Grenzen der Verteiltheit und Modularität auf. Die Beantwortung einer räumlichen Fragestellung (z.B. „Welche Flurstücke grenzen an Wasserflächen an und gehören Gemeinde X“) auf Basis verteilter Datenhaltung und verteilter Anfrageverarbeitung kann u.U. weniger effektiv sein als die Beantwortung derselben Frage aus einer integrierten Datenbasis. Die heutigen OGC Spezifikationen ermöglichen mehr oder weniger einfache, meist read-only Web-Anwendungen, an Analysefunktionalitäten (z.B. metrische Abfragen oder Polygon-Verschneidungen, die aus einer Anzahl bestehender Features neue Features erzeugen) besteht weiterhin ein Mangel. Die Harmonisierung von Datenmodellen wird in den derzeitigen OGC Web Services Spezifikationen nicht unterstützt. OGC Web Services ermöglichen syntaktische Interoperabilität zwischen Systemen durch

Kapselung der internen Strukturen der Systeme. Die konzeptionellen Modelle der von den OGC Web Services gekapselten Daten werden vor dem Nutzer verborgen. Diese Informationen werden aber benötigt, um semantische Interoperabilität zu ermöglichen, beispielsweise durch die Transformation zwischen unterschiedlichen Modellen. Eine Kombination der Ansätze des modellbasierten Datentransfers und der OGC Web Services erscheint also im Hinblick auf die grenzüberschreitende Nutzung von verteilten Geodaten sinnvoll.

3. Modellbasierter Ansatz für den Web-Zugriff auf verteilte Geodaten

3.1 Bestehende Ansätze zur Schematransformation in Web Services

Prozesse und Techniken, die benötigt werden, um unterschiedliche Datenmodelle zu harmonisieren, bzw. sie aufeinander abzubilden, werden in der deutschsprachigen Literatur als „Schematransformation“ oder „Modelltransformation“ bezeichnet.

Ansätze zur Schematransformation können anhand der folgenden Kriterien klassifiziert werden:

- **Abstraktionsebene:** Transformation auf konzeptioneller, logischer oder physikalische Ebene. Schematransformation auf konzeptioneller Ebene ist plattformunabhängig, wohingegen Ansätze auf der logischen und physikalischen Ebene plattformspezifisch sind.
- **Richtung:** Schematransformation kann entweder horizontal (zwischen verschiedenen Schemata auf derselben Abstraktionsebene) oder vertikal (zwischen verschiedenen Schemata auf unterschiedlichen Abstraktionsebenen) erfolgen.
- **Grad der Automatisierung:** Schematransformation kann bzgl. des Mappings zwischen Entitäten in unterschiedlichen Schemata mit unterschiedlichen Automatisierungsgraden ausgeführt werden. In der Regel wird das Mapping „von Hand“ ausgeführt, es gibt aber auch Ansätze zum automatischen Schema Mapping mit Hilfe von Ontologien, auf die hier aber nicht näher eingegangen wird.

Schematransformationen werden in internationalen Initiativen und Projekten im Bereich Geodateninfrastrukturen

(z.B. INSPIRE, RISE, EuroSpec, etc.) als wichtiges Interoperabilitätsthema angesehen. Es wurden jedoch bisher noch keine detaillierten Spezifikationen bzgl. Typ und Ebene der Transformationen oder über Tools zur Ausführung dieser festgelegt. In einer Reihe von internationalen Projekten wie z.B. OGC GOS-TP (OGC 2003a; OGC 2003b), GiMoDig (Sarjakoski und Lehto 2004) und SDI GER (Orlova und Bejar 2005) wurden horizontale Transformationen zwischen verschiedenen Schemata (z.B. Transformation von Straßendaten aus nationalen Schemata in ein gemeinsames Schema) unter Verwendung von Web Services erprobt und Prototypen sog. „translating WFS“ entwickelt.

Die Transformationen in den oben genannten Projekten wurden alle auf der logischen Ebene ausgeführt, beispielsweise mit Hilfe von Xquery oder XSLT. Transformationen zwischen unterschiedlichen konzeptionellen Schemata und die Definition einer Sprache zur Abbildung solcher Transformationen werden als zukünftige Forschungsfelder identifiziert. Im Projekt mdWFS (siehe Kapitel 3.3) werden diese Themen aufgegriffen mit dem Ziel, Transformationen auf konzeptioneller Ebene zu ermöglichen.

3.2 Das Projekt mdWFS

Mit dem Ziel, die Nutzung von verteilten, heterogenen Geodaten zu vereinfachen, kooperieren die Technische Universität München (TUM) und die Eidgenössische Technische Hochschule Zürich (ETHZ) in dem Forschungsprojekt „Modellbasierter Ansatz für den Web-Zugriff auf verteilte Geodaten am Beispiel grenzübergreifender GIS-Anwendungen“ (mdWFS), das vom Bundesamt für Kartographie und Geodäsie und swisstopo finanziert wird. In der ersten Projektphase (2006) werden die Möglichkeiten und Grenzen der beiden oben genannten Ansätze „Direkte Kopplung zwischen Quell- und Zielsystemen mit OGC Web Services“ und „modellbasierter Datentransfer“ evaluiert. In einer Machbarkeitsstudie wird die Möglichkeit der Kombination der beiden Ansätze beurteilt und analysiert, welche Voraussetzungen für eine solche Kombination erfüllt sein müssen. In Phase zwei (2007) wird ein Prototyp eines modellbasierten Web Feature Services (mdWFS) entwickelt und implementiert, der die ►

Defizite der bestehenden OGC Web Services in Bezug auf semantische Interoperabilität überwinden soll. Dazu werden die Konzepte des modellbasierten Datentransfers sowie der Schematransformation auf konzeptioneller Ebene für einen WFS adaptiert. Dies beinhaltet auch die Implementierung eines Formalismus für die konzeptionelle Beschreibung von Schematransformationen (vgl. Gnägi et al. 2006). Folgendes Szenario kann zur Illustration des mdWFS Workflows herangezogen werden: Eine EU Behörde fordert Daten (z.B. über Verwaltungsgrenzen) bei einer deutschen Vermessungsverwaltung an, die einen mdWFS betreibt. Die konzeptionelle Beschreibung des deutschen Datenmodells wird als Antwort auf die DescribeFeatureType Anfrage geliefert (z.B. im Format XMI). Der Nutzer kann dann Transformationsregeln zwischen dem deutschen und dem EU Schema formulieren (oder bereits vorformulierte Transformationsregeln für diese Transformation laden). Die Transformation wird im mdWFS unter Verwendung der Transformationsregeln und der Beschreibung des Zielschemas ausgeführt. So können die Daten aus dem deutschen Quellsystem dem Nutzer in der EU Behörde bereits bezogen auf das gewünschte EU Schema geliefert werden.

4. Schlussfolgerungen

Eine Reihe von Ansätzen für die grenzüberschreitende, web-basierte Nutzung von Geodaten existieren bereits. Das Projekt mdWFS zielt auf die Überwindung einiger wichtiger Defizite dieser bisherigen Ansätze ab. Die Harmonisierung von Datenmodellen wird mittels Web Services unter Verzicht auf eine Datenintegration in ein zentrales System möglich. Gegenüber bestehenden Ansätzen zur Schematransformation (Transformationen auf der logischen oder Formatebene, vgl. Kapitel 3.2) sind durch den Ansatz im mdWFS Projekt (Transformation auf konzeptioneller Ebene) u.a. folgende Vorteile zu erwarten:

- Die Möglichkeit der Formulierung von Transformationsregeln auf der konzeptionellen Ebene, z.B. durch die graphische Zuordnung von Klassen des Quellschemas zu Klassen des Zielschemas (möglicherweise unterstützt durch eine graphische Benutzerober-

Literatur:

Donaubauer, A. 2004: Interoperable Nutzung verteilter Geodatenbanken mittels standardisierter Geo Web Services. Dissertation, Technische Universität München, Institut für Geodäsie, GIS und Landmanagement, München.

Donaubauer, A. 2005: A Multi-Vendor Spatial Data Infrastructure for Local Governments based on OGC Web Services, FIG Working Week 2005 and GSDI-8, 16.-21. April, 2005, Kairo.

Gläsel, H. et al. 2006: Der Bodensee-Geodatenpool, Baustein für eine internationale Geodateninfrastruktur. In: Mitteilungen des DVW-Bayern e.V., Bd. 2, 2006, S. 265-290.

Gnägi, H.R.; Morf, A.; Staub, P. 2006: Semantic Interoperability through the Definition of Conceptual Model Transformations. 9th AGILE Conference on Geographic Information Science, 20-22 April 2006, Visegrád.

Gnägi, H.R.; Plabst, S. 2004: Modellbasierter Datenaustausch und Geoweb services im Internet. Zwei Wege zur Geodateninfrastruktur. 9. Münchner Fortbildungsseminar Geoinformationssysteme, 10.-12. März 2004, München.

Huber, A. 2005: Aufbau eines grenzübergreifenden GIS in der Bodenseeregion. Möglichkeiten und Grenzen der Ansätze zur kombinierten Nutzung verteilter, vorhandener Geodaten aufgezeigt am Beispiel der grenzübergreifenden Planung. Diplomarbeit, Technische Universität München, Institut für Geodäsie, Gis und Landmanagement, München.

Jaenicke, K. 2003: Set-up of a cross-border spatial data infrastructure for an interope-

table use of data through the internet. ESRI User Conference 2003, 07.-11. Juli 2003, San Diego.

KOM 2004 = Kommission der Europäischen Gemeinschaften (Hrsg.) 2004: Vorschlag für eine Richtlinie des Europäischen Parlaments und des Rates zur Schaffung einer Raumdateninfrastruktur in der Gemeinschaft (INSPIRE). KOM(2004) 516 endgültig, 2004/0175 (COD), SEC(2004) 980

OGC 2003a = Open Geospatial Consortium Inc. (Hrsg.) 2003a: GOS-TP Translating WFS (WFS-X) - Galdos's Experience.

OGC 2003b = Open Geospatial Consortium Inc. (Hrsg.) 2003b: WFS-X Schema Translation: The GOS-TP Experience.

Orlova, N; Bejar, R. 2005: SDIGER project. INSPIRE Ad hoc meeting, 8. November 2005, Bonn.

Riecken, J.; Bernard, L.; Portele, C.; Remke, A. 2003: North-Rhine Westphalia: Building a Regional SDI in a Cross-Border Environment / Ad-Hoc Integration of SDIs: Lessons learnt. 9th EC-GI & GIS Workshop, 25-27 June 2003, A Coruña.

Sarjakoski, T.; Lehto, L. 2004: The GiMoDig Schema translation based on XSLT. EuroSDR Workshop, 16 April 2004, Paris.

Steudle, G. 2006: Der Bodensee-Geodatenpool – mehr als "grenzenlos". In: Vermessung und Geoinformation, Bd. 1/2, 2006, S. 63-77.

Straub, F. 2005: Modellbasierter Ansatz für den Web-Zugriff auf verteilte, objektorientierte Geodaten. Master's Thesis, Technische Universität München, Fakultät für Informatik, München.

fläche) kann den Zugang zu Schematransformationen für den Nutzer vereinfachen. Der Nutzer braucht zwar Expertenwissen bzgl. der konzeptionellen Datenmodelle, nicht aber über die Formulierung von Transformationsregeln in einer dafür vorgesehenen maschinenlesbaren Sprache.

- Der modellbasierte, plattformunabhängige Ansatz macht es möglich, die Daten bezogen auf das Zielschema des Nutzers zu liefern, statt nur im gewünschten Transferformat. Die Wartbarkeit von Transformationsregeln wird dadurch ebenfalls verbessert. Wenn entweder im Quellschema oder im Zielschema Veränderungen auftre-

ten, können die Transformationsregeln auf der konzeptionellen Ebene modifiziert und die entsprechenden neuen Transferformate automatisch abgeleitet werden. ■

ANSPRECHPARTNER

Dr.-Ing. Andreas Donaubauer
Technische Universität München
Fachgebiet Geoinformationssysteme
andreas.donaubauer@bv.tum.de

Dipl.-Ing. Peter Staub
Eidgenössische Technische Hochschule
Zürich
Institut für Geodäsie und Photogrammetrie
staub@geod.baug.ethz.ch