

NEWSLETTER

e-geo.ch
Geoinformation



Inhalt

April 2009



Status quo: Überlagerung von WMS und WFS zur interoperablen Visualisierung von Relief, Kantons-grenzen und SwissNames

Schwerpunkt:

Semantische Interoperabilität

- 2 **Editorial: GDI und semantische Interoperabilität**
Rolf Buser, swisstopo

- 4 **Die Bedeutung der Semantik für Interoperabilität und Datentransfer**
Alessandro Carosio, ETH Zürich

- 8 **Fachliche Hürden für semantische Interoperabilität**
Peter Jordan, Steuerungsorgan e-geo.ch

- 9 **Gesucht: Innovative Projekte im Bereich Geoinformation**

- 10 **Semantische Modelltransformation im Kontext von INSPIRE**
Univ.-Prof. Dr.-Ing. Matthäus Schilcher, Dr.-Ing. Andreas Donaubauer, Dipl.-Geogr. Astrid Fichtinger, Dipl.-Inf. Tatjana Kutzner, TU München

- 14 **Datenaustausch, Begriffe, Normen**
Hans Rudolf Gnägi, ETH Zürich

- 16 **Semantische Interoperabilität als Voraussetzung für die Realisierung von INSPIRE**
Christine Giger, INSPIRE-Kontaktstelle Schweiz, wissenschaftliche Koordinatorin HUMBOLDT Projekt

- 20 **Voraussetzungen und Hindernisse zur interoperablen Nutzung verteilter Geodaten in GDI**
Peter Staub, swisstopo

- 22 **Konzeptionelle Beschreibungssprache für Modellabbildungen**
Andreas Morf, FKL & Partner AG, ETH Zürich



Rolf Buser,
swisstopo

Liebe Leserinnen, liebe Leser

Betreiber von Geodateninfrastrukturen (GDI) kennen die Problematik von heterogenen Geodaten, welche in mühsamen und fehleranfälligen Prozessen integriert werden müssen. Gut dokumentierte oder gar harmonisierte und qualitätsgeprüfte Geodaten sind noch nicht die Regel. Je grösser das Einsatzgebiet dieser Geodaten ist, von kommunal, kantonal, regional, national bis international, desto schwieriger wird eine vollständige Normalisierung der Datenstrukturen und der Semantik. Auch im gleichen Fachthema werden wir es immer mit unterschiedlichen Datenstrukturen und -modellen zu tun haben. Wir haben aber die Vision der Geodateninfrastrukturen vor Augen: Jeder möchte via Internet schnell, einfach und möglichst kostenlos auf vernetzbare und standardisierte Geodienste zugreifen können und künftige Geodateninfrastrukturen sollen untereinander vernetzt sein.

Die Grundsteine für eine gute Interoperabilität müssen zuerst auf der organisatorischen Ebene gelegt werden. Interoperabilität heisst auch «zusammenarbeiten können». In so genannten Fachinformationsgemeinschaften «inter operieren» die Beteiligten untereinander und müssen sich über die Bedeutung der Begriffe einigen um möglichst präzise die Struktur der Daten beschreiben zu können. Es ist entscheidend, dass man untereinander mit einer gemeinsamen Sprache kommuniziert und sich immer wieder vergewissert, ob man das gleiche meint, als Ziel immer ein Geodatenmodell vor Augen.

Im Geoinformationsgesetz (GeoIG) werden für die Themenbereiche der Geobasisdaten nach Bundesrecht minimale Geodatenmodelle gefordert. Wird uns der Ansatz der semantischen Interoperabilität den grossen Aufwand, diese minimalen Geodatenmodelle zu erstellen, erleichtern? Ich denke nicht. Umgekehrt: Geodatenmodelle, d.h. möglichst präzise Datenbeschreibungen erleichtern den Weg zu mehr Interoperabilität.

Gibt es aber konkrete Anforderungen an minimale Geodatenmodelle, wie sie das GeoIG verlangt? Minimale Geodatenmodelle müssen in einer produkt- und herstellerneutralen Sprache beschrieben werden. In einem Objektkatalog werden die Objekte dann mit ihren Eigenschaften möglichst präzise inklusive Erfassungsrichtlinien dokumentiert. Das Modell soll den fachlichen Anforderungen genügen und möglichst einfach sein, d.h. nicht alles Wünschbare sondern nur das Nötigste wird modelliert. Weiter soll es über längere Zeit möglichst stabil bleiben. Minimal

bedeutet aber nicht minimalistisch. Es ist eine Herausforderung an die Fachinformationsgemeinschaften, den grössten gemeinsamen Nenner zu suchen und zu erarbeiten. Oft ist die Fachgesetzgebung heute noch zu ungenau um eine hinreichend genaue Vorgabe für Geodatenmodelle zu sein. Anpassungen der Fachgesetze werden deshalb nötig sein.

Mit den minimalen Geodatenmodellen im Sinne des GeoIG wird eine verbindliche Abmachung im Sinne eines Vertrages geschaffen. Die Betroffenen haben sich per Gesetz daran zu halten.

In der Schweiz werden für die Modellierung von Geobasisdaten nach Bundesrecht UML und INTERLIS als Modellierungssprachen eingesetzt. Das UML-Diagramm (grafisch), die INTERLIS-Modelldatei (textlich) und der Objektkatalog (beschreibend) bilden zusammen eine Einheit, um das Modell zu dokumentieren. Das Schöne an dieser schweizerischen Vorgabe ist, dass damit die Geodatenmodelle genau beschrieben werden können und damit die Voraussetzungen für Interoperabilität verbessert werden.

Auch mit einer guten Beschreibungssprache kann man schlechte Modelle machen. Es braucht gute Modellierer, welche einen breiten und reichhaltigen Erfahrungsschatz in unterschiedliche Fachthemen und mit unterschiedlichen GIS haben. Das nötige Fachwissen bringen die Fachexperten in die Geodatenmodellierung mit ein.

Bei den Geobasisdaten nach Bundesrecht haben wir nun den Auftrag aber auch die Chance, minimale Geodatenmodelle zu erstellen, welche auf nationaler Ebene harmonisiert, präzise beschrieben und vollständig dokumentiert sind. Bei der Erstellung neuer Modelle sollten aber wo vorhanden, bestehende Modelle (aus Bundessicht v.a. auch im internationalen Umfeld) konsultiert und wo sinnvoll übernommen werden.

Im vorliegenden Newsletter werden Sie auch Beispiele finden, wo trotz Vorgabe von detaillierten Geodatenmodellen das Ziel der technischen Interoperabilität nicht erreicht wurde. Der Bund hat ähnliche Erfahrungen gemacht. Man hat sich im Bereich der Metadaten auf das Datenmodell von ISO (ISO TC211 19115) gestützt. Obwohl unsere Nachbarländer ihre Metadaten auf den gleichen Grundlagen und nach dem gleichen UML-Modell von ISO aufgebaut haben, ist heute die technische Interoperabilität nicht gegeben. Wir sind zwar ISO-kompatibel, aber nicht interoperabel! Die bestehenden internationalen Geonormen genügen offenbar den Ansprüchen der Interoperabilität nicht. Es bestehen zu viele Freiheitsgrade und UML alleine kann den Anspruch einer präzisen Beschreibungssprache nicht erfüllen. In der Schweiz schliessen wir diese Lücke mit der Kombination von UML mit INTERLIS.

Gut dokumentierte oder gar harmonisierte und qualitätsgeprüfte Geodaten sind noch nicht die Regel.

Wir sind zwar ISO-kompatibel, aber nicht interoperabel!



Alessandro Carosio,
ETH Zürich

Interoperabilität und Datentransfer: Eine zentrale Funktion jedes GIS

Das Entstehen und der erfolgreiche Einsatz von Geoinformationssystemen sind mit einer wirksamen Lösung des Problems der Interoperabilität und des Transfers der Geodaten untrennbar gekoppelt.

Wir verfügen heute über eine immer grösser werdende Menge geografischer Daten in digitaler Form, und immer mehr Personen arbeiten mit Daten dieser Art. Das Risiko, dass Doppelspurigkeiten und Widersprüche entstehen, ist gross. Oft existieren die benötigten Daten bereits. Sie werden aber wieder erfasst, weil eine geeignete Dokumentation fehlt oder sie eine inkompatible Struktur aufweisen. Die Geoinformationstechnologie wird aber nur erfolgreich sein, wenn sie die Anforderungen möglichst vieler User erfüllt. Wir brauchen technische Lösungen, um die gemeinsame Verwendung der Geodaten zu ermöglichen.

Die wirtschaftlich entwickelten Länder realisieren zur Zeit nationale Geodateninfrastrukturen (NGDI), in welchen Private, Ämter, Lieferanten und Anwender integriert werden, damit sie gemeinsam Technologien und existierende Daten nutzen können. Voraussetzung dafür ist die Realisierung von Koordinationszentren, Standard, technischen Lösungen und Richtlinien, die für den Informationsaustausch und für die interoperable Nutzung der Geodaten zwischen Systemen unterschiedlicher Hersteller erforderlich sind.

Die Lösungsansätze für diese Aufgabe sind besonders aktuell, da man zurzeit sowohl auf nationaler Ebene als auch in internationalen Institutionen (Europa, Welt) an der Entwicklung von technischen Normen arbeitet, die die Realisierung von interoperablen Systemen ermöglichen sollen.

Erforderliche Funktionalitäten

Die Vielfalt der Bedürfnisse ist gross. Ausgetauscht oder angefragt werden z.B.

- Grafische Darstellungen (digitale Karten und Pläne)
- Beschreibungen der GIS-Inhalte (Metadaten)
- Resultate von Abfragen (Tabellen, Karten usw.)
- Strukturierte Datenbankinhalte (z.B. Tabellen, Attribute) ohne Änderung der Datenstrukturen
- Daten von einer Datenbank in eine andere Datenbank mit unterschiedlichen Datenstrukturen
- Vollständige Objekte in einer objektorientierten Umgebung (inkl. Operationen)

Interoperable GIS müssen daher über unterschiedliche Funktionalitäten verfügen:

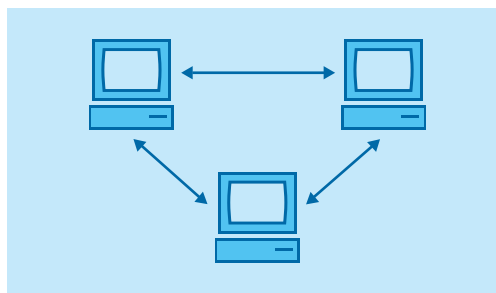


Abb. 1: Die Kommunikation zwischen Geoinformationssystemen ist eine Notwendigkeit

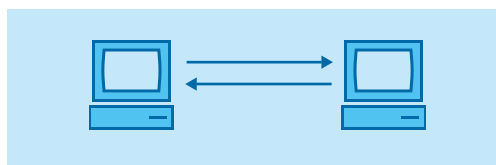


Abb. 2: Datenaustausch und Interoperabilität

- Datenaustausch zwischen gleichen Systemen
- Datenaustausch mit Standardformaten
- Modellbasierte Datentransfermethoden
- Interoperabilität (nach OGC = Open Geospatial Consortium)
- Semantische Transformationen für den Datentransfer
- Semantische Interoperabilität

Proprietäre Transferformate

Alle GIS verfügen über die Möglichkeit, die gespeicherten Daten in Standarddateien (in der Regel sequentielle Dateien) auszugeben und aus den gleichen Dateien wieder zu übernehmen. Die verwendeten Formate sind ausschliesslich geeignet für eine bestimmte GIS-Software (proprietäres Format) und erfordern beim Sender und beim Empfänger die identische Datenstruktur für die transferierten Daten.

Standard-Formate

Wenn sowohl die geometrischen als auch die thematischen Inhalte feststehen (dies ist vor allem in stark hierarchischen Organisationen der Fall: NATO, Staatsbetriebe usw.) kann man ein passendes Format definieren, das die Information aufnehmen kann, und es ist möglich, in jedem System die entsprechende Schnittstelle für das Lesen und Schreiben der Transferdateien in das definierte Format einzubauen. So wurden zum Beispiel die Austauschformate DIGEST (Digital Geographic Information Exchange Standard) und das Format NTF (National Transfer Format) definiert. Bei solchen Anwendungen hat man mit Geoinformationssystemen zu tun, in welchen die zu verwaltenden Informationen für alle Systeme einheitlich festgelegt sind.

Modellbasierte Transferverfahren

Geoinformationen der heutigen Zeit bieten den Anwendern die Möglichkeit, die thematischen Inhalte frei zu definieren. Das System bildet dann aufgrund der Datenstrukturbeschreibung (Eingabe des DB-Schemas) die erforderlichen Entitätsklassen. Somit passt sich die Datenverwaltung den Bedürfnissen an.

Wir brauchen technische Lösungen, um die gemeinsame Verwendung der Geodaten zu ermöglichen.

Abb. 3: Modellbasierte Transferverfahren



Abb. 4: Komponente des modellbasierten Transferverfahren

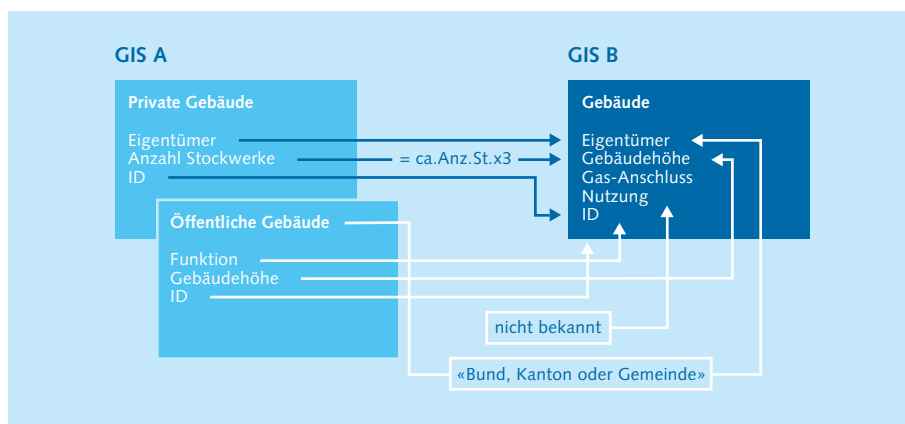
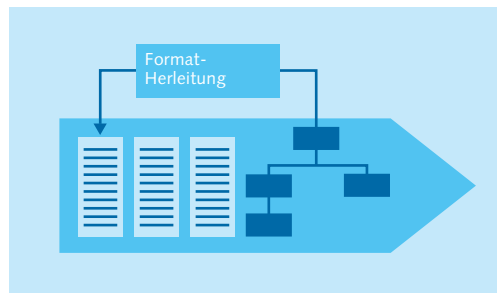


Abb. 5: Ein Beispiel von semantischer Transformation

übernehmen und damit die erforderliche Datenstruktur im eigenen System implementieren. Bei vielen GIS kann diese Operation automatisch ausgeführt werden, wenn die Datenbeschreibung in INTERLIS vorliegt.

Wenn hingegen im Empfängersystem eine Datenstruktur bereits besteht und sie nicht identisch mit der erhaltenen ist, muss eine semantische Transformation ausgeführt werden. Diese Operation bedeutet nichts anderes, als die Elemente des erhaltenen Datenmodells in die Elemente des bestehenden Modells zu transformieren. Die Elemente, die im empfangenden System nicht vorhanden sind, gehen verloren; die Elemente, die in den übermittelten Daten nicht vorhanden sind, können selbstverständlich nicht übernommen werden.

Das konstruierte Beispiel lässt sofort erkennen, dass die Abbildung in der Regel nicht automatisch entstehen kann. Um eine semantische Transformation ausführen zu können, wird man sich zuerst erkundigen oder vorhandene Beschreibungen lesen müssen, um aufgrund der eigenen Interpretationen und Zielsetzungen die Datenfelder in Beziehung zu bringen. Allenfalls sind Transformationsberechnungen erforderlich (Anzahl Stockwerke > Gebäudehöhe). Die Abbildung kann auch Annahmen beinhalten oder keine Information ergeben. Normalerweise sind diese Abbildungen nicht vollständig und nicht invertierbar. Sie müssen explizit in beide Richtungen definiert werden.

Man kann diese Idee auch brauchen, um den Datentransfer flexibel zu gestalten: Die Struktur der Daten, die man übertragen möchte, wird beschrieben, und daraus kann ein Format für diese Daten hergeleitet werden. Dafür benötigt man zwei Komponenten: erstens eine standardisierte Datenbeschreibungssprache, um die Struktur der Daten, die man transferieren möchte, eindeutig und konsistent zu beschreiben, und zweitens ein genormtes Verfahren, um aus der Datenstruktur ein Format herzuleiten.

In der Schweiz wurde die Sprache INTERLIS definiert, welche die Beschreibung der Thematik in einem GIS nach einem relationalen Modell und der Geometrie aufgrund von festgelegten Geometrieelementen ermöglicht. Die Sprache verfügte von Anfang an über einen Compiler für die automatische Herleitung der Transferformate. Zurzeit kann weltweit keine andere Lösung als operationell betrachtet werden.

Die neue INTERLIS-Version (INTERLIS 2) ist mit den ISO-Normen (Objektorientierung, inkrementelle Nachführung, usw.) voll kompatibel. Das Transferformat wurde im Informatik-Standard XML und im XML-Profil GML definiert.

Wenn der Empfänger noch keine Datenstruktur definiert hat, kann er die Datenbeschreibung

Interoperabilität

Die Datentransfermethoden haben als Ziel, die Daten von einem System zu einem anderen zu übertragen. Es sind aber auch Kommunikationsmöglichkeiten denkbar, die ohne Verschiebung der Originaldaten auskommen.

Diese Lösung ist besonders interessant, wenn nur einfache Auswertungen der Information benötigt werden (z.B. eine grafische Darstellung von einem Ausschnitt oder die Auflistung von bestimmten Objekten). In diesen Fällen ist es einfacher, die Auswertungsbefehle und -ergebnisse zu standardisieren und zu transferieren als die Grunddaten selbst.

Die Interoperabilität nach OGC bedeutet die Parallelnutzung verschiedener GIS, indem die Befehle (Anfragen) und die daraus entstehenden Ergebnisse (Antworten) ausgetauscht werden. Die GIS-Industrie (Softwarehersteller) hat das Open Geospatial Consortium (OGC) gegründet, um diese Technik möglichst weit zu entwickeln.

In den OGC Interoperabilitäts-Standards ist eine semantische Transformation nicht vorgesehen, sodass die Anfragen immer in der Datenstruktur (Namen der Entitätsklassen, Namen der Attribute usw.) des angefragten Systems formuliert werden müssen. Die Antworten erhält man ebenfalls in

der Datenstruktur des angefragten Systems. Man spricht von einer syntaktischen Interoperabilität.

Die semantische Transformation: Voraussetzungen und Grenzen

Die obigen Beschreibungen zeigen, dass Datenaustausch in der Regel eine semantische Transformation (Anpassung der Datenstruktur) erfordert, um ausgeführt werden. Bei der OGC-Interoperabilität ist diese Transformation nicht vorgesehen: Die Abfragen zu den anderen Systemen und die entsprechenden Antworten finden in der Datenstruktur des angefragten Systems statt. Der Empfänger muss selber das Erhaltene interpretieren und transformieren.

In der Praxis möchte man aber die wertvollen Informationen, die an verschiedenen Orten erfasst und verwaltet werden, gemeinsam nutzen und unabhängig von der Herkunftsdatenstruktur verwenden, sodass eine semantische Transformation als unentbehrlich betrachtet werden muss.

Ontologien

In Systemen, die unter gemeinsamen Regeln organisiert sind, wird es in spezifischen Bereichen möglich sein, Ontologien zu entwickeln, welche die Semantik beschreiben. In diesen Fällen wird vielleicht die Herstellung automatischer Transformationsmodule (Agenten) möglich werden, die nur die Beschreibung der Datenstruktur benötigen. Die Ontologien sind Gegenstand der heutigen Forschung. Man wird allerdings nur die Semantik in klar definierten und abgegrenzten Sektoren eindeutig beschreiben können. Eine Ontologie entspricht einer Standardisierung der Begriffe, die man für die Modellbildung und für die Bezeichnung der Elemente der Datenstruktur im GIS verwendet.

Heutige Bedürfnisse, die Antworten der Forschung

Der Bedarf nach Interoperabilität nimmt laufend zu. Weltweit möchte man NGDI realisieren. Dies wird wirtschaftlich nur tragbar sein, nur wenn die vorhandenen Datenbestände möglichst automatisch ohne aufwändige Operationen (z.B. Neuerfassung, mehrfache Verwaltung usw.) zugänglich gemacht werden können. Dies erfordert eine leistungsfähige Lösung des Problems der semantischen Transformationen sowohl bei eigentlichen Datentransferoperationen als auch beim Einsatz der OGC-Interoperabilität. Das grosse Vorhaben der EU (Programm INSPIRE), mit welchem die europäischen Geodaten über die Umwelt gemeinsam genutzt werden sollen, zeigt die Aktualität dieser Problematik. Die Schweiz ist im Rahmen der e-geo.ch Aktivitäten ebenfalls täglich mit Fragen der semantischen Interoperabilität konfrontiert. In der letzten Zeit wurden daher viele Projekte in Angriff genommen, um technische Lösungen zu entwickeln und zu erproben. Die vorhandene Publikation bietet einen Einblick in diese Thematik.

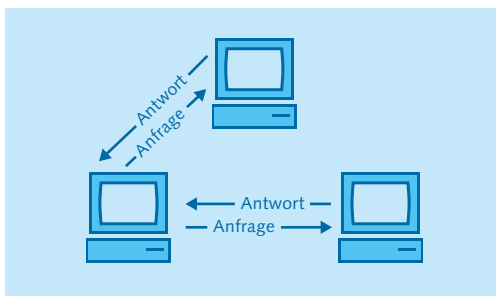


Abb. 6: Interoperabilität

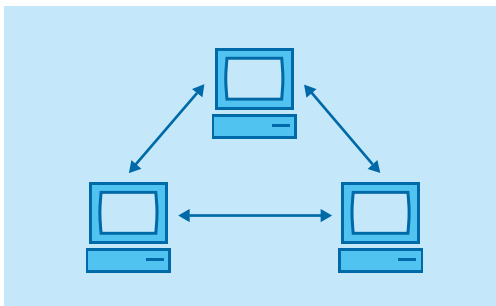


Abb. 7: Gewünscht wird Interoperabilität zwischen beliebigen Systemen mit beliebigen Datenstrukturen

Bis jetzt bestand die einzige Möglichkeit, semantische Transformationen auszuführen, auf Format-Ebene. Die Inhalte eines GIS wurden in eine sequentielle Datei mit dem proprietären Format des Systems geschrieben und mit einem speziell entwickelten Programm in eine andere proprietäre Datei konvertiert. Um diese Arbeit zu unterstützen, haben Software-Hersteller Module entwickelt, die leichter die Zuordnung der Tabellenfelder (Schema Matching) erlauben (eine der erste Lösungen war INTERLIS-Studio von Leica, heute hat FME von Safe Software grosse Verbreitung). Im Rahmen der Forschungsprojekte der GIS-Gruppe der ETH Zürich wurden in der letzten Jahren mehreren Projekte (Doktorarbeiten, internationale Normung, Kooperationen mit Industrie und Verwaltung) abgewickelt über diese aktuelle Thematik. Ein Teil davon wird in dieser Publikation von den Autoren beschrieben.

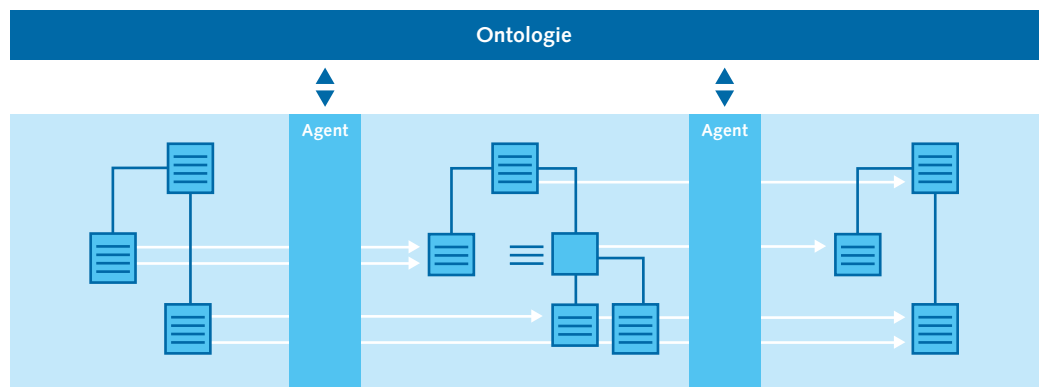
Die wichtigsten Entwicklungen der letzten Jahre sind die Erkenntnisse, die eine Bearbeitung der semantischen Transformationen auf konzeptueller Ebene anstatt auf Formatebene, wie es bisher der Fall war, ermöglichen.

Um dies zu erreichen, benötigt man eine Sprache (d.h. ein Darstellungsverfahren), um die Abbildungen zwischen zwei bereits definierten Datenstrukturen zu beschreiben. Eine solche Sprache kann als Erweiterung bestehender Modellierungssprachen (INTERLIS, UML) realisiert werden.

Eine solche Sprache wurde im Rahmen der Doktorarbeiten von A. Morf und P. Staub entwickelt. Es handelt sich um UMLT als Ergänzung von UML als grafische Methode für Abbildungsdarstellungen sowie eine textuelle Form von UMLT als Erweiterung von INTERLIS 2.

Der Bedarf nach Interoperabilität nimmt laufend zu. Weltweit möchte man NGDI realisieren.

Abb. 8: Semantische Transformationen mit Hilfe von Ontologien



Die nächsten Schritte der Forschung sind im Bereich der semantischen Interoperabilität und der semantischen Transformationen zu erwarten.

Eine weitere Komponente dieser neuen Entwicklungen sind die Software-Module (Compiler), welche die beschriebenen Modellabbildungen in Operationen umwandeln. Damit ist ein wesentlicher Schritt zur Automatisierung der Transformationen gemacht worden. Im Bereich der modellbasierten Methoden verfügen wir mit dem INTERLIS-Compiler über ein Instrument, um aus konzeptionellen Modellbeschreibungen Transferformate herzuleiten. In Kombination mit der Beschreibung von Modellabbildungen kann man jetzt die semantische Transformation auf konzeptioneller Ebene definieren und automatisch ausführen lassen.

Im Rahmen einer internationalen Zusammenarbeit zwischen dem Bundesamt für Landestopographie (Schweiz), dem Bundesamt für Kartografie und Geodäsie (Deutschland), der ETH Zürich und der Technischen Universität München wurde ebenfalls eine Erweiterung der OGC-Interoperabilität entwickelt und mit einer Prototypsoftware getestet. Es handelt sich um das *Model Driven Web Feature Service (mdwfs)*, mit welchem geografische Elemente eines GIS direkt mit den eigenen Objektbezeichnungen abgefragt werden können, um Antworten in den Bezeichnungen der eigenen Datenstruktur zu erhalten. Voraussetzung dafür ist eine vorgängige Beschreibung der Modellabbildungen auf konzeptioneller Ebene mit UMLT oder mit dem textuellen UMLT (INTERLIS-ähnlich).

Schlussfolgerung

Die Kommunikation zwischen Geoinformationssystemen ist eine vielseitige Aufgabe, für welche eine Vielfalt von Lösungsansätzen zur Verfügung steht. Die Probleme, die Anforderungen und die technischen Grenzen sind erst in neuester Zeit in ihrer ganzen Breite wahrgenommen worden. Die heute eingesetzten Datentransferverfahren und die Interoperabilität werden sich in den nächsten Jahren stark entwickeln. Mehrere Methoden werden sich als Standard durchsetzen. Die nächsten Schritte der Forschung sind im Bereich der semantischen Interoperabilität und der semantischen Transformationen zu erwarten.

Diese Arbeiten sind die Voraussetzung, damit Geoinformationssysteme ihre wertvollen Daten ohne Verzögerung und Hindernisse der ganzen Gesellschaft zur Verfügung stellen und die geplanten nationalen und internationalen Geodateninfrastrukturen tatsächlich in nützlicher Frist realisiert werden können.

Literatur:

A. Carosio, 2005: Interoperabilität für die breite Nutzung von Geoinformation, Weiterbildungstagung 17. und 18. März 2005, Zürich, ETH Höggerberg



Peter Jordan,
Steuerungsorgan e-geo.ch,
Mitglied Geschäftsleitung
Böhringer AG

Die rasche Zunahme digital erfasster Geoinformationen weckt den Wunsch, diese direkt für gebietsübergreifende Analysen einzusetzen und sie einem koordinierten überregionalen Zugriff zugänglich zu machen. Eine direkte Verwendung scheitert aber oft daran, dass die Geodaten in den einzelnen Regionen oder Verwaltungseinheiten in verschiedenen Systemen mit abweichenden Techniken und mit unterschiedlichen Zielsetzungen erhoben wurden.

In den vergangenen Jahren wurden grosse Anstrengungen unternommen, die durchgängige Nutzung von Geodaten über System- und Technikgrenzen hinweg zu vereinfachen, auch wenn sie auf unterschiedlichen Standards basieren. Man spricht dann von «semantischer Interoperabilität», da die Standards vorab durch eine semantische Transformation aufeinander abgebildet werden müssen. Der Fokus dieser Anstrengungen lag bislang vor allem im technischen Bereich. Eine Analyse verschiedener Datenharmonisierungsvorhaben im Rahmen des Projektes e-geo.ch 06-09 zeigte jedoch klar auf, dass der technischen meist eine begriffliche Harmonisierung auf fachlicher Ebene vorausgehen muss, bevor Geodaten auf einen gleichen Standard gebracht, oder ein Schlüssel für deren Austausch und Zusammenführung definiert werden können. Oder, pointiert ausgedrückt, die Fach-Semantik muss «interoperabel» sein, bevor Fach-Geodaten sinnvoll ausgetauscht oder zusammengeführt werden können.

Es geht also nicht nur darum, Datenmodelle aneinander anzugleichen, Datenstrukturen aufeinander abzubilden oder allenfalls Masseneinheiten (1 barrel = 159 l) zu übersetzen. Vielmehr muss vorab sichergestellt werden, dass unter gleichen oder ähnlich lautenden Begriffen auch das Gleiche verstanden wird, bzw. dass unterschiedlich benannte Objekte als identisch erkannt werden. Dieser Prozess muss unter Einbezug der jeweiligen Fachleute erfolgen.

Bei einem Überblick über den planerischen Naturschutz in Europa ist beispielsweise zu beachten, dass ein Nationalpark in Deutschland (IUCN-Kategorie II oder V) nicht mit dem Schweizer Nationalpark (IUCN-Kategorie I) gleichgesetzt werden darf. Bei der Erstellung einer Strassenkarte kann eine Schweizer Nationalstrasse nur ausnahmsweise oder unter gewissen Aspekten (Betrieb, Finanzierung) mit einer französischen «route national» verglichen werden.

Innerhalb der Schweiz können die aufgrund kantonaler oder kommunaler Erlasse erhobenen Geodaten meist nicht unreflektiert untereinander

Bevor die semantische Interoperabilität gewährleistet werden kann, muss die Semantik interoperabel sein.



Abb. 1: Prinzip der **interoperablen Semantik**

verglichen werden. Eine Wohnzone W2 im Kanton Luzern mag zwar vom Prinzip her einer Zone W2 im Kanton Zürich entsprechen. Aber bereits innerhalb des Kantons Zürich können sich entscheidende Auflagen in W2 von Gemeinde zu Gemeinde signifikant unterscheiden. Und wie weit ist die Zone W2B der Gemeinde Henggart mit der Stadtzürcher Zone W2bII vergleichbar? Gibt es – neben den zwei Stockwerken – überhaupt Eigenschaften, welche allen W2(+)-Zonen der Schweiz gemeinsam sind? Andererseits gibt es auch Begriffe, die zwar unterschiedlich tönen, in der Sache aber durchaus etwas Vergleichbares bezeichnen: so z.B. die «Überbauungsordnung» des Kantons Bern und der «Gestaltungsplan» des Kantons Solothurn.

Bemerkenswert ist, dass ähnliche Begriffsunschärfen auch in Gebieten vorhanden sind, welche durch die Bundesgesetzgebung klar geregelt werden. So beschränkt das Raumplanungsgesetz seit jeher die mögliche Unterteilung der Nutzungspläne abschliessend auf Bau-, Landwirtschafts- und Schutzzonen. Also sollten – mindestens aus der Sicht der GIS-Spezialisten – alle kommunalen Zonen eindeutig einer dieser Klassen zugeordnet werden können. In der Praxis tauchen aber durchaus auch Zwitter auf, wie z.B. «Bauernhofzonen», welche fallweise der Bau- oder der Landwirtschaftzone zugerechnet werden. Und im Bereich des Grundwasserschutzes hat sich in den Jahren vor der Inkraftsetzung der aktuellen Bundesverordnung eine Vielzahl von Unterteilungen der Schutzzone S eingebürgert. So wurde die Engere Schutzzone S2 oft in ein und demselben Kanton in eine Zone S2 und S2A oder in eine Zone S2A und S2B unterteilt. Entsprechend gilt $S2A = S2B$ und $S2A \neq S2A$.

Dieser komplexe und oft auch diffuse Sachverhalt kann in einem Datenmodell schlicht nicht abgebildet werden. Bei der Interoperabilität hilft – bei spezifischen Fragestellungen – allenfalls die Fuzzylogik. Allerdings wären IT-Spezialisten zu deren Formulierung zwingend auf die Beratung durch Fachexperten angewiesen, welche sich wiederum mit der Logik der Datenmodellierung auseinandersetzen müssen. In diesem Fall wäre

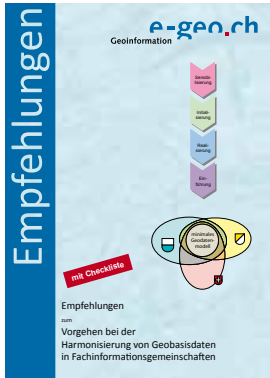


Abb. 2: **Empfehlungen Vorgehen Harmonisierung Geodaten**
(s.a. Talon auf der Umschlagrückseite)

es wohl sinnvoller und zweckmässiger, wenn sich die Fachexperten vorgängig zur Datenmodellierung mit der Begrifflichkeit in ihrem Fachgebiet auseinandersetzen und diese wo nötig revidieren und harmonisieren. Insbesondere in der mehrsprachigen, unterschiedliche Verwaltungstraditionen pflegenden Schweiz ist aber auch immer eine interoperable Fach-Semantik gefragt, welche fallspezifisch die Zusammenführung unterschiedlicher Daten erlaubt. Entsprechend können die Schnittstellen zu Gewerbe und Hochschulen sowie zum Ausland sichergestellt werden.

Das Projekt e-geo.ch 06-09 kam zum Schluss, dass eine nachhaltige Datenmodellierung und -interoperabilität nur durch den Einbezug der jeweiligen Fachinformationsgemeinschaften gewährleistet werden kann, welche neben den IT-Experten auch die Fachspezialisten und Nutzer umfasst. Wie man dabei zweckmässig vorgeht, wird in den entsprechenden Empfehlungen¹ skizziert.

¹ zu beziehen unter: www.e-geo.ch > Publikationen > Empfehlungen

Es wäre wohl sinnvoller und zweckmässiger, wenn sich die Fachexperten vorgängig zur Datenmodellierung mit der Begrifflichkeit in ihrem Fachgebiet auseinandersetzen und diese wo nötig revidieren und harmonisieren.

INNOVATIONSPREIS E-GEO.CH

Gesucht: Innovative Projekte im Bereich Geoinformation

**2009
doppelte
Preissumme:
Fr. 10 000.-**

Im November 2009 wird zum zweiten Mal der e-geo.ch-Innovationspreis verliehen. Es werden sowohl Studentenarbeiten als auch Projekte und Entwicklungen von Organisationen oder Firmen ausgezeichnet, die sich den Themen der zukünftigen NGDI (Nationale Geodaten-Infrastruktur der Schweiz) widmen.

Der Preis wird in drei Kategorien verliehen:

1) Nachwuchspreis:

Abschlussarbeiten von Studierenden einer Universität, Fachhochschule, Geomatiktechnikerlehrgang oder Mittelschule (Bachelor-, Master-, Diplom-, Maturaarbeiten oder Dissertationen).

2) Innovationspreis:

Projekte oder Entwicklungen von Organisationen oder Firmen, die einen besonders innovativen Beitrag zur NGDI leisten.

3) Ehrenpreis:

Personen, die sich in besonderer Weise um die NGDI verdient gemacht haben.

Anmeldung:

Ab dem 1. Juni 2009 können Sie sich auf www.e-geo.ch für die Teilnahme am Innovationspreis anmelden. Die Arbeiten müssen bis am 31. Juli 2009 eingereicht werden.

Teilnahmebedingungen:

Die detaillierten Teilnahmebedingungen sowie Angaben und Richtlinien zur Einreichung der Bewerbung finden Sie auf www.e-geo.ch.

Trägerschaft:

Die Verleihung des e-geo.ch-Innovationspreises wird von e-geo.ch organisiert und durchgeführt. Die Verantwortung für das Programm «e-geo.ch» trägt ein Steuerungsorgan, in dem 15 Vertreter von Bund, Kantonen, Schweizer Städten und Gemeinden sowie der Schweizerischen Organisation für Geo-Information (SOGI) vereint sind.

Infos: www.e-geo.ch



Univ.-Prof. Dr.-Ing. Matthäus Schilcher
 Dr.-Ing. Andreas Donaubaue, Akademischer Oberrat a.Z.
 Dipl.-Geogr. Astrid Fichtinger, Wissenschaftliche Mitarbeiterin
 Dipl.-Inf. Tatjana Kutzner, Wissenschaftliche Mitarbeiterin
 TU München

Das Projekt «Modellbasierter Ansatz für den Web-Zugriff auf verteilte Geodaten am Beispiel grenzübergreifender GIS-Anwendungen»

Vorwort

Aus einer langjährigen engen Zusammenarbeit zwischen der ETH Zürich und der TU München auf dem Gebiet der Geoinformatik ist das hier beschriebene gemeinsame Forschungsprojekt entstanden, welches von den zuständigen Behörden in der Schweiz (Bundesamt für Landestopographie, swisstopo) und Deutschland (Bundesamt für Kartographie und Geodäsie, BKG) unterstützt wird.

Die aktuelle Situation des weltweiten Geoinformationsmarktes lässt sich folgendermassen charakterisieren: Die Technik von Geoinformationssystemen (GIS) ist ausgereift. Die GIS-Technologie ist in vielen Bereichen von Wirtschaft und Verwaltung zu einem unentbehrlichen Standardwerkzeug für raumbezogene Aufgaben geworden. Bei behördlichen und privaten Stellen stehen flächendeckend Geodaten aus vorhandenen GIS für interne und externe Nutzer zur Verfügung. Dies gilt insbesondere für die Vermessungsverwaltungen und Umweltbehörden, die mit ihrem breiten Spektrum an amtlichen Geobasisdaten bzw. Umweltdaten jeweils zu den führenden Datenanbietern eines Landes zählen.

Trotz der unbestritten grossen Fortschritte in den letzten Jahren und zusätzlichen Innovationen wie dem Internet/Intranet (für Web-GIS) oder GNSS für den mobilen Einsatz, sowie den Erfolgen, die bei der Standardisierung und Normung für den Datenaustausch erzielt worden sind, gibt es Defizite, die einer noch besseren Ausschöpfung des GIS-Potenzials im Wege stehen. Ein Hindernis bildet die Verteiltheit und Heterogenität von vorhandenen Geodatenbeständen, die für lokale, regionale, nationale und länderübergreifende Aufgaben nicht ohne grössere Aufwände und spezielle Fachkenntnisse nutzbar sind. Die Überwindung dieser Probleme in Geodateninfrastrukturen (GDI) stellt gegenwärtig eine der grossen internationalen Herausforderungen für die Geoinformatik dar.

Die politisch Verantwortlichen in vielen Ländern haben diese Problematik erkannt und fördern seit einiger Zeit intensiv die Realisierung von nationalen und länderübergreifenden Geodateninfrastrukturen (GDI). Von hervorgehobener Bedeutung ist die von der EU beschlossene Richtlinie INSPIRE, die Einfluss auf die nationalen GDI der Mitgliedsstaaten hat und die in den nächsten Jahren in die Praxis umgesetzt werden soll.

Einführung und Problemstellung

Die INSPIRE Richtlinie der EU zum Aufbau einer europaweiten Geodateninfrastruktur verlangt von den Mitgliedstaaten, Geodaten über interoperable Geodienste bereitzustellen. Hierfür werden in den INSPIRE Durchführungsbestimmungen standardisierte Schnittstellen erarbeitet, die einen einheitlichen Zugriff auf die Geodienste aller Mitgliedstaaten ermöglichen und die Daten über einheitliche Datentransferformate bereitstellen sollen. Damit für diese Daten auch auf semantischer Ebene, also auf Ebene der Datenmodelle Interoperabilität gegeben ist, werden im Rahmen der INSPIRE Durchführungsbestimmungen zudem harmonisierte, europaweit einheitliche Datenmodelle (INSPIRE Data Specifications) für bestimmte Themen erstellt, welche in drei Anhängen zur INSPIRE Richtlinie (INSPIRE Annex I–III) aufgeführt sind, wie z.B. Verwaltungsgrenzen oder Gewässernetz. Ende 2008 wurden für die Themen aus Annex I Entwürfe der Data Specifications veröffentlicht. Aus Sicht der Anbieter derartiger Daten (z.B. Vermessungsverwaltungen) bedeutet dies, dass eine Transformation der Daten aus ihren eigenen Datenmodellen in die von der EU vorgegebenen Modelle durchzuführen ist. In Deutschland wird somit beispielsweise eine Transformation des ATKIS Basis-DLM als Teil des AFIS-ALKIS-ATKIS Referenzmodells (AAA) nach INSPIRE notwendig. Die INSPIRE Richtlinie sieht für diese Transformationen sogenannte Modelltransformationssdienste vor. Für derartige Dienste gibt es jedoch noch keine gültigen Standards [Müller 2008] und es herrscht Forschungsbedarf in diesem Bereich (vgl. z.B. [Donaubaue et al. 2007], [Staub et al. 2008], [Lehto 2007], [Balley 2007], [ORCHESTRA 2008]).

Grundlagen der semantischen Modelltransformation
 Interoperabilität bezeichnet die Fähigkeit verteilter Systeme zur Zusammenarbeit. Aus technischer Sicht charakterisieren zwei Hauptaspekte die Interoperabilität:

Syntaktische Interoperabilität (Dateninteroperabilität): Die Fähigkeit mehrerer Systeme oder Prozesse, Daten über bestimmte Datentransferformate und Schnittstellen auszutauschen. Die Struktur der Datenformate bzw. Schnittstellen (Syntax) ist dabei zwischen den beteiligten Systemen bekannt und nutzbar.

Semantische Interoperabilität (Modellinteroperabilität): Hier ist neben der Syntax auch die Bedeutung der ausgetauschten Daten (Semantik) zwischen den beteiligten Systemen bekannt und nutzbar.

Trotz der unbestritten grossen Fortschritte in den letzten Jahren und zusätzlichen Innovationen gibt es Defizite, die einer noch besseren Ausschöpfung des GIS-Potenzials im Wege stehen.

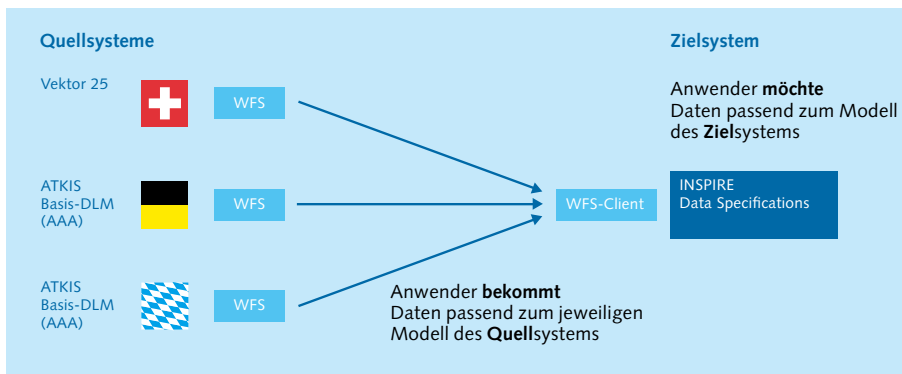


Abb. 1: Datenbezug mittels des OGC WFS

Die zentrale Idee des modellbasierten Ansatzes ist die Beschreibung von Geodatenmodellen mittels einer konzeptionellen Schemasprache.

Um semantische Interoperabilität zu erreichen, ist es unabdingbar, konzeptionelle Modellabbildungen zwischen den zu transformierenden konzeptionellen Datenmodellen zu definieren, um damit semantische Modelltransformationen auszuführen. Mittels eines «Übersetzers» lassen sich dann die Abbildungsregeln interpretieren, anschließend kann die Transformation der Datenobjekte automatisch ausgeführt werden.

Die zentrale Idee des modellbasierten Ansatzes ist die Beschreibung von Geodatenmodellen mittels einer konzeptionellen Schemasprache (engl. Conceptual Schema Language, CSL). Die Verwendung einer CSL erlaubt es, Datenstrukturen auf konzeptioneller Ebene, also unabhängig von Systemen ODER spezifischen Transferformaten wie XML oder GML zu beschreiben. Ein mit einer CSL beschriebenes Datenmodell kann als Träger semantischer Information im Sinne der semantischen Interoperabilität bezeichnet werden. Aus einem konzeptionellen Datenmodell können unter Zuhilfenahme von sog. Kodierungsregeln (engl. Encoding Rules) und geeigneter Compiler praktisch beliebige Transferformate automatisch abgeleitet werden. Im vorliegenden Forschungsprojekt kommt für die Datenmodellierung die textuelle CSL «INTERLIS» zum Einsatz. INTERLIS (so wie im Projekt zur Anwendung kommend) basiert auf einem UML 2-Profil. Mit einem Compiler können XML-Schemata im INTERLIS-Format oder auch GML-Applikationsschemata für beliebige INTERLIS-Modelle erzeugt werden.

Bestehende OGC Web Services wie der Web Feature Service (WFS) weisen Einschränkungen bezüglich der semantischen Interoperabilität auf: Sie sind nur auf syntaktische Interoperabilität ausgelegt (siehe oben) und bieten keine Handhabe für die semantische Interoperabilität. Abbildung 1 verdeutlicht dies: Ein Anwender möchte über einen WFS Daten passend zum Modell seines Zielsystems abrufen. Da jedoch der WFS keine semantische Interoperabilität unterstützt, bekommt der Anwender die Daten stets passend zum Modell des gewählten Quellsystems geliefert und nie zum Modell des Zielsystems. Mit einer CSL beschriebene Modelle sind in einem WFS – wenn überhaupt vorhanden – für Zielsysteme nicht sichtbar und semantische Transformationen werden nicht unterstützt. OGC Web Services sind

somit per se nicht in der Lage, neben der Dateninformation (GML) mit zugehöriger Beschreibung des Transferformats (GML Anwendungsschema) auch eine Modellinformation zu behandeln [Donaubauer et al. 2005].

Lösungsansatz zur web-basierten Modelltransformation

Die Spezifikation eines Web Service mit der Fähigkeit zur Modelltransformation lässt sich aufteilen in die Spezifikation einer Sprache zur Formulierung der Transformationsregeln (siehe hierzu den Beitrag von A. Morf) und in die Spezifikation eines Protokolls für die Kommunikation mit dem Service, das nachfolgend erläutert wird:

Beim Entwurf der Schnittstelle für einen Web Service mit der Fähigkeit zur Semantischen Modelltransformation (im Projekt «mdwfs» genannt, für model driven Web Feature Service) wurde darauf geachtet, dass sich der Service bestmöglich in eine OGC Web Services Infrastruktur integrieren lässt. Um dies zu erreichen, wurde als Grundlage für den mdwfs die OGC Spezifikation des Web Feature Service [OGC 2005] herangezogen und entsprechend den Anforderungen an den mdwfs erweitert.

Zum einen wurde für den mdwfs die WFS Operation DescribeFeatureType dahingehend erweitert, dass der mdwfs konzeptionelle Datenmodelle im Format XMI, einer XML-Serialisierung für UML-Modelle, bereitstellen kann. Zum anderen wurden für den mdwfs zwei neue Operationen definiert:

GetCapabilities: Diese Operation ermöglicht es einem mdwfs-Client, dienstbezogene Metadaten abzurufen. In Abwandlung der WFS Spezifikation des OGC enthält die Antwort auf eine GetCapabilities-Anfrage keine «feature type list» sondern eine Liste mit den Namen aller konzeptionellen Modelle, auf die die jeweilige mdwfs-Instanz einen Zugriff erlaubt.

DoTransform: Durch Aufruf dieser Operation wird die semantische Transformation ausgeführt. Jeder erfolgreiche Aufruf der Operation resultiert in einer neuen Instanz eines herkömmlichen WFS, über den die Daten im transformierten Datenmodell (Zielschema) abrufbar sind.

Die Antwort des mdwfs auf eine DoTransform-Anfrage besteht aus einer Statusmeldung sowie aus einer GetCapabilities-Request-URL als eindeutigem Identifikator für die neu angelegte WFS-Instanz. Ein herkömmlicher WFS-Client kann diesen URL nutzen, um auf Funktionalität und Daten der WFS-Instanz zuzugreifen.

Anwendungsbeispiel

Als Demonstration für das Konzept wurde ein Prototyp entwickelt, welcher eine Transformation von Verwaltungsgrenzen aus Deutschland und der Schweiz in ein nach dem Entwurf des EuroSpec

Feature Catalog «Administrative Boundaries» erstelltes Modell vornehmen kann. Dabei wurden aus dem deutschen Modell Administrative Gebietseinheiten (Bundesland, Regierungsbezirk, Kreis, Kommune) und aus dem Schweizer Modell Gemeindegrenzen transformiert. Als Testdaten standen auf deutscher Seite ATKIS-Daten im Masstab 1:25 000 zur Verfügung, die teilweise vom Bundesamt für Kartographie und Geodäsie in Kooperation mit der Firma AED-SICAD in das neue AAA-Modell [AdV 2006] migriert wurden. Aus der Schweiz wurde der Datensatz GG25 verwendet [Swisstopo 2006].

Für die Demonstration wurde folgendes Anwendungsszenario festgelegt: Ein Kunde möchte Verwaltungsgrenzen aus topografischen Informationssystemen aus Deutschland und der Schweiz über das Internet beziehen. Die Daten sollen bezogen auf das Zielmodell (Eurospec Feature Catalog) ausgeliefert werden. Stellt der Kunde nun GetCapabilities- sowie GetFeature-Anfragen, so erhält er daraufhin, unter Nutzung des derzeitigen Stands der Technik, also OGC WFS ohne mdwfs, die Daten bezogen auf die Quellmodelle. Es ist somit nur eine syntaktische Interoperabilität vorhanden. Nach Import und Darstellung der Daten in einem GIS, das hier als WFS Client genutzt wird, ergibt sich das in Abbildung 2 dargestellte Resultat. Anhand der Layer- und der Spaltennamen der Attributtabelle ist zu erkennen, dass die Daten zum Thema Verwaltungsgrenzen aus Deutschland und der Schweiz deutlich unterschiedlich modelliert sind. Grenzübergreifende Abfragen und Analysen müssen deshalb immer getrennt für beide Datenmodelle definiert werden, was die Arbeit mit diesen Daten aufwändig macht.

Abbildung 3 hingegen zeigt das Ergebnis, das durch die Nutzung des mdwfs erreicht werden kann. Anhand der Layer- und der Spaltennamen der Attributtabelle wird deutlich, dass die Daten nun bezogen auf ein harmonisiertes Datenmodell (Eurospec Feature Catalog) abgerufen wurden. Die Daten aus Deutschland und der Schweiz weisen nach der semantischen Transformation ein identisches Datenmodell auf, es ist somit nun neben der syntaktischen Interoperabilität auch eine semantische Interoperabilität vorhanden. Grenzübergreifende Abfragen und Analysen müssen nur noch bezogen auf das harmonisierte Datenmodell (Zielmodell) formuliert werden, was vor allem bei komplexeren Analysen den Arbeitsaufwand beim Anwender deutlich reduziert.

Potenzial des Lösungsansatzes im Hinblick auf INSPIRE

Der hier beschriebene Lösungsansatz für eine Web-basierte Modelltransformation zeichnet sich durch folgende Merkmale aus:

- Web-Schnittstelle für modellbasierten Datentransfer (Web-Zugriff auf Daten und CSL-Beschreibung des konzeptionellen Datenmodells)

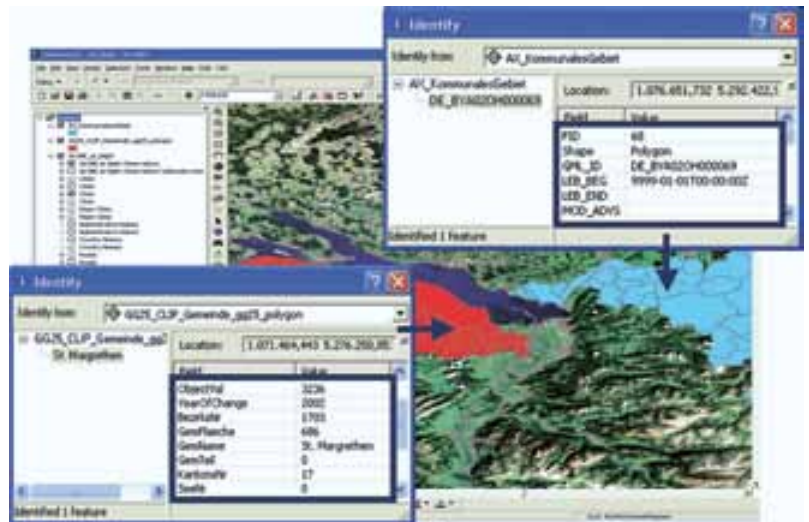


Abb. 2: WFS-GetFeature-Anfrage basierend auf syntaktischer Interoperabilität

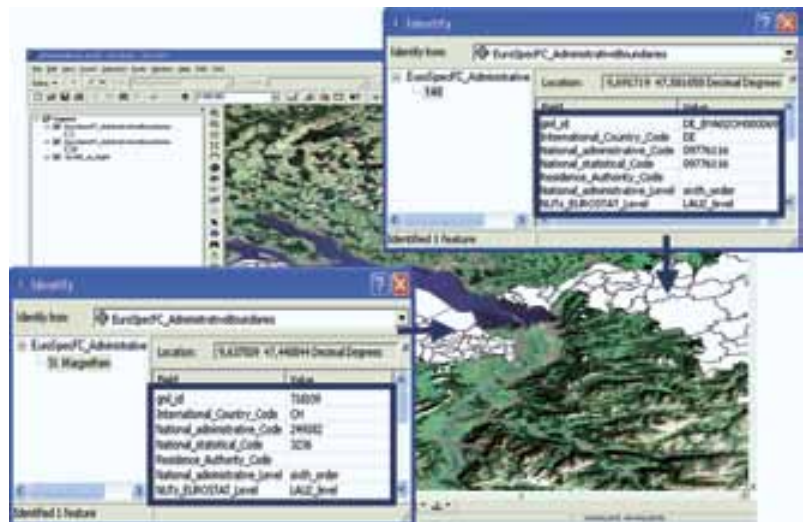


Abb. 3: WFS-GetFeature-Anfrage basierend auf semantischer Interoperabilität

- Definition von Transformationsregeln auf konzeptioneller Ebene mittels der im Projekt entwickelten Sprache UMLT
- Web-Schnittstelle für das Durchführen einer semantischen Transformation (Operation DoTransform)

Aufgrund dieser Merkmale ergeben sich folgende Potenziale:

- *Modellbasierter Datentransfer:*
Der im Projekt konzipierte und prototypisch umgesetzte mdwfs ermöglicht den modellbasierten Datentransfer auf der Ebene von Web Services. Zusätzlich zu den Geodaten kann eine Beschreibung des konzeptionellen Datenmodells abgerufen werden.
- *Semantische Interoperabilität:*
Durch den Betrieb eines einzigen Web Service können Daten bezogen auf beliebige, benutzerdefinierte Datenmodelle abgegeben werden, sofern sinnvolle Transformationsregeln zwischen dem Quellmodell und einem benutzerdefinierten Zielmodell erstellt werden können. Die Web-Schnittstelle für das Durch-

Kontakt

Technische Universität München,
Fachgebiet Geoinformationssysteme,
Arcisstraße 21, D-80333 München

Telefon +49 (0) 89 289 22578,
E-Mail schilcher@bv.tum.de,
andreas.donaubauer@bv.tum.de,
astrid.fichtinger@bv.tum.de,
tatjana.kutzner@bv.tum.de

Eine sinnvolle Nutzung des Lösungsansatzes setzt voraus, dass die beteiligten Datenmodelle das gleiche Metamodell der CSL verwenden.

führen einer semantischen Transformation ermöglicht es jedem Datenbezieher, mit dem Wissen zur Erstellung von Transformationsregeln Daten bezogen auf sein Zielmodell abzurufen, ohne dass hierfür Aufwand beim Datenbereitsteller entsteht.

- **Syntaktische Interoperabilität:** Die Schnittstelle für den Abruf von Geodaten entspricht der OGC-WFS-Spezifikation ohne projektspezifische Erweiterungen (wird durch die Designentscheidung erreicht, jede semantische Transformation in einer neuen WFS-Instanz resultieren zu lassen). Die Herstellerunabhängigkeit und die Interoperabilität mit bestehenden kommerziellen oder Open Source Komponenten z.B. in einer GDI auf Basis OGC Web Services sind damit voll gewährleistet.

Das Potenzial des Lösungsansatzes für die Praxis lässt sich z.B. anhand von Anwendungsszenarien im Kontext der INSPIRE Richtlinie darstellen:

- Im Rahmen der INSPIRE Durchführungsbestimmungen werden Datenspezifikationen für die Themen gemäss INSPIRE Anhang 1 bis 3 entwickelt. Diese Datenspezifikationen werden mit einer konzeptionellen Beschreibungssprache (CSL) beschrieben.
- Teile der in INSPIRE Anhang 1 bis 3 genannten Themen werden von den topografischen Informationssystemen z.B. von Deutschland und der Schweiz abgedeckt.
- Die Modelle dieser Informationssysteme sind ebenfalls mittels einer CSL beschrieben. Organisationen, die gefordert sind, Daten gemäss der INSPIRE Richtlinie bereitzustellen, sind mittels des mdWFS hierzu sowohl semantisch (durch semantische Transformation) als auch syntaktisch (OGC WFS Schnittstelle) in der Lage, ohne ihren Quelldatenbestand zu ändern. Auf Änderungen in den Quell- oder Zielmodellen kann flexibel durch Ändern der Transformationsregeln reagiert werden.

Die im gewählten Lösungsansatz vorgesehene Definition der Transformationsregeln auf Ebene der konzeptionellen Modelle hat sich im Zusammenhang mit den zu Testzwecken verfügbaren Modellen als machbar, wegen der Komplexität der Modelle jedoch als nicht trivial erwiesen.

Die Voraussetzungen für eine sinnvolle Nutzung des Lösungsansatzes sind in der Praxis nur teilweise erfüllt. Folgende Probleme konnten identifiziert werden:

- Eine sinnvolle Nutzung des Lösungsansatzes setzt voraus, dass die beteiligten Datenmodelle das gleiche Metamodell der CSL verwenden. Bei den verfügbaren Quellmodellen aus Deutschland und der Schweiz war dies nicht der Fall.
- Eine sinnvolle Nutzung des Lösungsansatzes setzt voraus, dass für das Encoding der Daten, d.h. für die Herleitung von GML-Anwendungs-

schemata aus den konzeptionellen Modellen, einheitliche oder zumindest standardisierte Regeln verwendet werden. Bei den verfügbaren Testdaten war dies nicht der Fall.

Danksagung

Die in diesem Beitrag vorgestellten Ergebnisse entstanden im Rahmen des laufenden Forschungsprojekts «Modellbasierter Ansatz für den Web-Zugriff auf verteilte Geodaten am Beispiel grenzübergreifender GIS-Anwendungen (mdWFS)», das vom Bundesamt für Kartographie und Geodäsie (BKG) sowie vom Bundesamt für Landestopographie (swisstopo) gefördert wird. Das Projekt wird von Mitarbeitern des Fachgebiets Geoinformationssysteme (Prof. Schilcher) der Technischen Universität München und der Gruppe GIS und Fehlertheorie (Prof. Carosio) der Eidgenössischen Technischen Hochschule Zürich in einer gemeinsamen Forschungsgruppe bearbeitet.

Literatur:

- AdV** (Hrsg.) (2006): Dokumentation zur Modellierung der Geoinformationen des amtlichen Vermessungswesens (GeoInfoDok). Hauptdokument, Version 5.1.
- Balley, S.** (2007): Aide à la restructuration de données géographiques sur le Web. Dissertation an der Université de Marne-la-Vallée, Frankreich.
- Donaubauer, A.; Straub, F.; Schilcher, M.** (2007): mdWFS: A Concept of Web-enabling Semantic Transformation. Proceedings of the 10th AGILE Conference on Geographic Information Science, Aalborg.
- Donaubauer, A.; Schilcher, M.; Huber, A.** (2005): OGC-Lösungen: Möglichkeiten und Grenzen. In: Carosio, A. (Hsg.): Tagungsband zum Seminar «Interoperabilität für die breite Nutzung von Geoinformation, 17. und 18. März 2005, Zürich, ETH Hönggerberg.
- Lehto, L.** (2007): Real-time content transformations in a web service-based delivery architecture for geographic information. Dissertation am Finnischen Geodätischen Institut.
- Müller, M.** (2008): INSPIRE – Bedeutung für Dienste und Portale. In: Runder Tisch GIS e.V. (Hrsg.): Tagungsband zum 13. Münchner Fortbildungseminar Geoinformationssysteme, München, 2008.
- Object Management Group** (Hrsg.) (2003): MDA Guide Version 1.0.1. OMG specification omg/2003-06-01.
- Object Management Group** (Hrsg.) (2005): MOF 2.0 Query/Views/Transformations Specification. OMG specification ptc/05-11-01.
- Open Geospatial Consortium** (Hrsg.) (2005): Web Feature Service Implementation Specification: 1.1.0. OpenGIS implementation specification OGC 04-094.
- ORCHESTRA Consortium** (Hrsg.) (2008): Implementation Specification of the Translating Feature Access Service.
- Schweizerische Normenvereinigung** (Hrsg.) (2006): INTERLIS Referenzhandbuch, Version 2.3. Schweizer Norm SN 612031.
- Staub, P.** (2007): A Model-Driven Web Feature Service for Enhanced Semantic Interoperability. OSGeo Journal 1(3): 38–43.
- Staub, P.; Gnägi, H.R.; Morf, A.** (2008): Semantic Interoperability through the Definition of Conceptual Model Transformations. Transactions in GIS 12(2): 193–207.
- swisstopo** (Hrsg.) (2006): GG25. Die digitalen administrativen Grenzen der Schweiz.



Hans Rudolf Gnägi,
ETH Zürich

Interoperabilität = ?

«Inter» bedeutet «zwischen, untereinander, zusammen», «Operabilität» bezeichnet etwa «Bearbeitbarkeit», unter «Interoperabilität» kann man also verstehen «gemeinsame Bearbeitbarkeit, Zusammenarbeit». Wann spricht man praktisch von Interoperabilität? Wenn Systeme Daten gemeinsam bearbeiten können, heissen sie interoperabel. Das ist vor allem interessant bei Daten, welche mühsam zu erfassen und nachzuführen, daher teuer sind, wie das insbesondere bei Geodaten der Fall ist.

Zwei grundsätzlich verschiedene Typen von Interoperabilität

Wie im Beitrag von A. Carosio ausgeführt ist gibt es zwei grundsätzlich verschiedene Typen von Interoperabilität: Einerseits kann gemeinsame Nutzung von Daten stattfinden, indem sie *von* einem anderen System übernommen werden durch Transfer und Integration. Andererseits kann gemeinsame Nutzung *auf* einem anderen System stattfinden durch Übergabe eines Befehls und Übernahme des Resultats vom anderen System ohne Transfer und Integration von Basisdaten. Das Versenden von Befehle und Empfangen von Resultate über Web-Dienste ist natürlich sehr bequem, hat aber schon seine Grenzen. Etwa muss die gewünschte Operation auf dem angefragten System überhaupt implementiert sein, die Anfrage sollte nicht zu komplex sein und das Resultat nicht zu umfangreich. Da bekanntlich gilt, «l'appétit vient en mangeant», möchte man gerne auch umfangreicheren Datenaustausch mit Web Diensten besorgen.

Verschiedene Lösungen für den Datenaustausch

Die verschiedenen Lösungen, welche auch im Beitrag von A. Carosio ausführlich zur Sprache kommen, seien hier kurz zusammengefasst:

Proprietären Formate. Damit können gleiche Systeme Daten austauschen. Wollen das Systeme verschiedenen Typs machen, müsste bilateral umformatiert werden, was aufwändig wird. Bei n Systemen wären $n(n-1)$ Umformatierungsprogramme nötig.

Formatbasierte TransfERNorm, d.h. ein bestimmtes Format wird ausgewählt. Das reduziert die Anzahl Prozessoren radikal von $n(n-1)$ auf $2n$. Nachteil: Hinter dem ausgewählten Format steckt *eine* bestimmte Datenstruktur, ändert diese, dann muss das Format geändert werden, als die Norm.

Modellbasierte TransfERNorm: Diese ist ideal, weil sie Formate liefert, die sich der Datenstruktur anpassen. Die Methode lautet: Man beschreibe die Datenstruktur exakt durch ein systemunabhängiges konzeptionelles Datenmodell und leite daraus

Ein grosser und vermutlich der wesentliche Teil der Semantik (Bedeutung) eines Begriffs kann erfasst werden durch präzise Beschreibung der Struktur von Daten, die ihn charakterisieren.

nach festen Regeln die Formatbeschreibung des Standardformats her.

Kleiner (ontologischer) Exkurs zur Semantik

Der Transfer gleicher Datenstrukturen ist ein erster einfacher Schritt zur Interoperabilität. Der Austausch gleichem Inhalt zwischen verschiedener Datenstrukturen hingegen eine echte Herausforderung. Denken wir etwa an Umweltmonitoring in benachbarten Ländern. Da sind sicher gleiche oder mindestens vergleichbare Inhalte in sehr verschieden organisierten Datenbanken vorhanden. Oder auch Katastrophendaten sollten rasch über Kantons- und Landesgrenzen ausgetauscht werden können, sind aber sicher auch lokal total verschieden strukturiert.

Es gibt das Semantik-Problem: Gleiche Inhalte kommen sehr verschieden strukturiert als Daten vor (Beispiel: Leitungen in Figur 1), oder es gibt auch Inhalte, die verschieden sind, aber Daten haben, die gleich heissen (Beispiel Wildstatistik von Jägern und vom WWF, Gebäude für Architekten und BfS).

Und es gibt die Semantik-Frage: Wie kann man feststellen, ob in zwei unterschiedlichen Datenstrukturen dieselben Inhalte sind? Ob sie dasselbe bedeuten, d.h. ob sie denselben Realitätsausschnitt beschreiben, d.h. ob sie dieselbe Semantik haben? Auf die Semantik-Frage gibt es folgende Antwort-Versuche:

Es gibt den Begriff «Baum» im Deutschen wie im Polnischen: Der Einsatz eines Wörterbuches erlaubt den Transfer des Begriffs aus der einen Sprache in die andere. Ein «native speaker» wird aber feststellen, dass das polnische Wort Baum eher dem deutschen Begriff Baumart entspricht.

Es gab den gleichen Begriff «Punkt» in 26 verschiedenen Kantonen beim Start der RAV: Der Vergleich der proprietären Formate und die Suche nach einem gemeinsamen Format (erster Versuch, so die AVS zu definieren) endete nach ca. 4-stündiger heftiger Auseinandersetzung mit roten Köpfen. Die Lösung war: Exakte Beschreibung der hinter den Punkt-Formaten steckenden Datenstruktur und Harmonisierung derselben, danach automatische Herleitung des harmonisierten Formats aus der harmonisierten Struktur. Damit war auch die modellbasierte Methode INTERLIS erfunden.

Anwendung dieser Methode auf den Begriff «Baum», d.h. möglichst präzise die Struktur von Daten beschreiben, die einen Baum charakterisieren, ergibt z.B. folgendes objektorientierte Datenmodell: Damit ist klar, dass ein Objekt der Klasse Baum nicht nur die Art eines Baumes beschreibt, sondern den Baum als Ganzes viel allgemeiner.

Es gibt das Semantik-Problem: Gleiche Inhalte kommen sehr verschieden strukturiert als Daten vor.

Und es gibt die Semantik-Frage: Wie kann man feststellen, ob in zwei unterschiedlichen Datenstrukturen dieselben Inhalte sind?

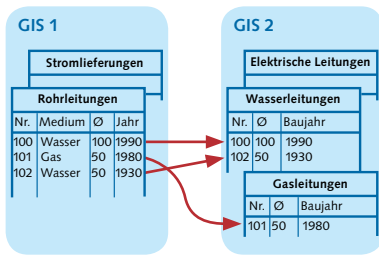


Abb. 1: Zwei verschiedene Datenstrukturen für Leitungen

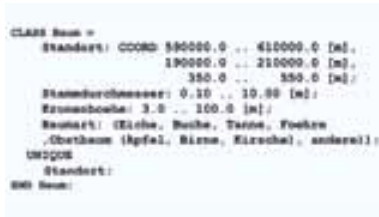


Abb. 2: Konzeptionelles Datenmodell für die Klasse Baum

Fazit: Ein grosser und vermutlich der wesentliche Teil der Semantik (Bedeutung) eines Begriffs kann erfasst werden durch präzise Beschreibung der Struktur von Daten, die ihn charakterisieren. Man ist versucht zu sagen: So bekommt man interoperable Semantik.

Herausforderung: Semantische Interoperabilität

Zurück zur Herausforderung: Austausch entsprechender Inhalte trotz je verschiedener Datenstruktur. Man spricht dann von semantischer Interoperabilität oder – eigentlich treffender – von Semantik erhaltender Interoperabilität. Das braucht sicher Datenumbau. Der kann immer geschehen durch bilaterales Umformatieren von je proprietären Formaten. Dies kann aber in aufwändige Programmierung ausarten, wenn geometrische Aspekte umgebaut werden müssen, etwa Spaghettilinien in berandete Flächen. Wirtschaftlicher ist sicher, solche und ähnliche umfangreiche Basisfunktionalitäten in geeigneten Werkzeugen zusammenzufassen.

Wie müsste ein solches Werkzeug aussehen? Das Fazit zu Semantik-Problem und Semantik-Frage legt ein modellbasiertes Vorgehen nahe: Man beschreibe die Datenstruktur des Startsystems und des Zielsystems exakt mit je einem Datenmodell, man definiere den Strukturumbau auf konzeptioneller Ebene, zwischen den Modellen, und lasse auf Grund dieses konzeptionellen Strukturumbaus den Datenumbau auf Format-ebene automatisch ausführen. Da kommt auch die modellbasierte Transfernorm ins Spiel, denn damit kann sowohl dem Startmodell wie dem Zielmodell ein Standard-Transferformat zugeordnet werden. Für den Strukturumbau auf konzeptioneller Ebene müssen Umbaufunktionen als Schlüsselwörter/grafische Icons zur Verfügung stehen, die entsprechenden Programme für den Formatumbau sind dann nur einmal zu programmieren.

- Verschiedene Realisierungen dieses Konzepts Semantik erhaltender Interoperabilität:
- Info Grips Conversion System (skriptgesteuert, formatnahe, viele Funktionen)
 - INTERLIS Studio (grafikgesteuert, weniger Funktionen)
 - UMLT (grafisch)/ILIT (INTERLIS ähnlich) (sehr wenig Funktionen, Prototypen Oracle-basiert bzw. FME-basiert, siehe Artikel von A. Morf und P. Staub, mit Integration in Web-Dienste durch mdwfs, siehe Artikel von M. Schilcher et al.)

Was braucht es für Normen?

- Eine konzeptionelle Datenbeschreibungssprache (Conceptual Schema Language CSL) ist normativ festzulegen, grafisch übersichtlich und textuell präzise, praxisnah, vollständig.
- Format-Herleitungs-Regeln (Encoding) müssen eindeutig festgelegt sein, d.h. der Zusammenhang zwischen der konzeptionellen und system- bzw. format-unabhängigen Modellstruktur (PIM = platform independent model) und der physischen und system- bzw. format-spezifischen Struktur der Instanzen (PSM = platform specific model). Ist das nicht der Fall, wird es prohibitiv aufwändig, wesentliche Tools zu implementieren.
- Eine konzeptionelle Umbausprache mit grafischer und textueller Ausprägung ist konsistent zur oben definierten CSL eindeutig festzulegen.
- Umbaufunktionen für die semantische Transformation sind als Programmschnittstellen zu definieren.

Was gibt es für Normen? Wie ist der Stand?

Wünschbar wäre automatisches Umstrukturieren (semantisches Web). Hauptschwierigkeit ist dabei das automatische Erkennen gleicher Semantik in verschiedenen Strukturen und die Definition des passenden Strukturumbaus.

Abgesehen davon, dass der Mensch für die Definition von Semantik erhaltendem Strukturumbau noch lange nicht entbehrlich sein wird (wenn überhaupt), braucht es Tool-Unterstützung, um nötige Erfahrung sammeln zu können. Die Praxistauglichkeit eines Umbautools ist davon abhängig, ob es eine überzeugende grafische Benutzeroberfläche hat, die auch den präzisen Text zu definieren erlaubt, analog zum UML/INTERLIS Editor, mit dem auch die Details des INTERLIS 2 Modells festgelegt werden können.

Normen für Semantik erhaltende Interoperabilität		
	ISO-OGC-CEN	SNV-eCH
CSL	Nur grafisch ISO 19103: UML	Grafisch: UML Textuell: INTERLIS 2 SN612031, eCH0031
Encoding-Regeln	Nicht eindeutig, modellspezifisch ISO 19136: GML ISO 19139: Metadata Encoding	Eindeutig, modellunabhängig, unterschiedliche Formate: SN 612030/31, Kapitel 3: ITF, ILI2 XML, GML
Umbausprache	Kein Kandidat	Kandidaten: UMLT, ILIT
Umbaufunktionen	Keine Kandidaten	Kandidaten: Namen, Einheiten ändern, Koordinatentransformationen, Datum- und Zeit-Transformationen, implizite/explicite Beziehungen, Klassen umstrukturieren.

Semantische Interoperabilität als Voraussetzung für die Realisierung von INSPIRE



Christine Giger,
INSPIRE-Kontaktstelle Schweiz,
wissenschaftliche Koordinatorin
HUMBOLDT Projekt

Der juristische Hintergrund

INSPIRE ist der Name einer Richtlinie der EU zum Aufbau einer Europäischen Geodateninfrastruktur. Ziel dieser Infrastruktur ist die Unterstützung der Europäischen Umweltpolitik. Die Richtlinie hat für EU-Mitgliedsstaaten Rechtsverbindlichkeit und enthält explizite Umsetzungsfristen.

In der Schweiz besteht keine direkte Verbindlichkeit. Rechtsgrundlage für die Schweiz ist das neue Geoinformationsgesetz (GeoIG) und die Geoinformationsverordnung (GeoIV), die in wesentlichen Grundsätzen allerdings INSPIRE-konform sind. Die Regelungen betreffend Vorgaben an Geodaten, Metadaten und Dienste sind analog in GeoIG/GeoIV und der INSPIRE-Richtlinie. Nicht enthalten sind die Einschränkungen auf Umweltdaten wie bei INSPIRE und es entsteht natürlich durch das Inkrafttreten von GeoIG/GeoIV auch keine Berichtspflicht gegenüber der Europäischen Kommission wie dies für EU-Länder der Fall ist. Allerdings gibt es zahlreiche Bereiche und internationale Vereinbarungen im Zusammenhang mit dem Umweltschutz und der Europäischen Umweltagentur aber auch im Bereich der Statistik und Vermessung, in denen die Anwendung der INSPIRE-Richtlinie auch für Schweizer Institutionen zum Datenaustausch mit internationalen Institutionen erforderlich und sinnvoll ist.

Gründe für INSPIRE

Die Gründe für die Entwicklung der INSPIRE-Richtlinie sind sogar im Gesetzestext explizit und ausführlich festgehalten. Es wird dort gesagt, dass Umweltpolitik in Europa sich auf den geeigneten Umwelt- und Geodaten aus den verschiedenen Europäischen Regionen und Verwaltungsebenen abstützen müsse, dass aber auch viele Probleme für die Nutzung dieser Daten bestünden. Explizit heisst es im Absatz (3) der Begründung: *«Die Probleme bei der Verfügbarkeit, Qualität, Organisation, Zugänglichkeit und gemeinsamen Nutzung von Geodaten betreffen in gleicher Weise zahlreiche Bereiche der Politik und Information und nahezu alle Verwaltungsebenen. Ihre Lösung erfordert Massnahmen für den Austausch, die gemeinsame Nutzung, die Zugänglichkeit und die Verwendung von interoperablen Geodaten und Geodatendiensten über die verschiedenen Verwaltungsebenen und Sektoren hinweg. Deshalb sollte in der Gemeinschaft eine Geodateninfrastruktur geschaffen werden.»*

In den weiteren Begründungen wird dann vor allem explizit auf (verschiedene) «Formate und Strukturen» der Geodaten eingegangen, d.h. die syntaktische Interoperabilität angesprochen. Die

semantische Interoperabilität wird nicht so explizit in einem Absatz verankert, allerdings an vielen Stellen (in etwa 15 der 34 Begründungsabsätze der Richtlinie) implizit erwähnt, indem etwa auf die Probleme der Kombinierbarkeit von Daten aus unterschiedlichen Quellen hingewiesen wird.

Konzeptionelle Modelle auf EU-Ebene

Der in der INSPIRE-Richtlinie verankerte Lösungsansatz sieht die Definition von Datenspezifikationen einschliesslich der zugehörigen konzeptionellen Modelle auf europäischer Ebene vor. Dazu werden in drei Anhängen 34 Themen gelistet, für die Datenspezifikationen in Durchführungsbestimmungen festgelegt werden müssen. Zur Zeit liegen erste Entwürfe für die 9 Themen des Anhang I vor. Dabei handelt es sich vor allem um Geodaten, die als Referenzinformationen für verschiedene Anwendungen benötigt werden, wie zum Beispiel Koordinatenreferenzsysteme, Verkehrs- und Gewässernetze. Die EU-Mitgliedsstaaten müssen später entweder ihre eigenen Datenbestände entsprechend diesen Datenspezifikationen aufbauen oder aber mindestens die Daten über Geodienste gemäss den vorgegebenen Datenmodellen bereitstellen. Die Anpassung der internen Strukturen ist nur für ganz neu erfasste Datenbestände zwingend. Für bereits vorhandene Daten genügt die Bereitstellung über Dienste.

Das Vorgehen, vorhandene Datenbestände über Geodienste gemäss einem EU-Datenmodell anzubieten, soll einerseits zentralen Stellen (z.B. der Europäischen Umweltagentur) ermöglichen, Daten aus verschiedenen Quellen ohne Harmonisierungsaufwand direkt zu integrieren und zu nutzen. Andererseits sollte das Vorgehen den EU-Ländern eine aufwendige Umstrukturierung ihrer internen Datenbestände ersparen.

Voraussetzungen und Datenharmonisierung

Allerdings ist es für die Umsetzung dieses Ansatzes notwendig, dass lokale Datenbestände mit lokalen Quellmodellen in Datenbestände gemäss der vorgegebenen Zielmodelle überführt, d.h. transformiert werden (können). Dieser Vorgang wird als Harmonisierungsprozess bezeichnet und muss technisch in den Webdiensten ausgeführt werden, die dafür sorgen, dass die Geodaten auf EU-Ebene (oder für andere Anwendungen) gemäss INSPIRE-Datenmodellen bereitstehen.

Die notwendigen Harmonisierungsprozesse enthalten dabei viele unterschiedliche Aspekte:

- Schematransformationen, d.h. die Abbildungen von konzeptionellen Quelldatenmodellen auf Zieldatenmodelle
- Geometrische Transformationen, z.B. wegen Edge Matching Problemen
- Koordinationentransformationen
- Sprachtransformationen
- Formatttransformationen, etc.

Einige dieser Aspekte lassen sich mehr oder weniger leicht durch automatisierte Prozesse abdecken, ein Grossteil erfordert jedoch fundamentales Expertenwissen und Entscheidungskompetenz durch menschliche Benutzer.



Die Untersuchungen im Rahmen von HUMBOLDT zeigten sehr klar auf, dass viele der bestehenden Probleme bei der Umsetzung von INSPIRE eher organisatorischer als technischer Art sind.

Einige dieser Aspekte lassen sich mehr oder weniger leicht durch automatisierte Prozesse abdecken, ein Grossteil erfordert jedoch fundamentales Expertenwissen und Entscheidungskompetenz durch menschliche Benutzer.

Förderung der Forschung und Entwicklung

Auf Seiten der Europäischen Kommission ist man sich durchaus bewusst, dass die weltweit existierenden Ansätze zur technischen Unterstützung der Datenharmonisierung noch nicht so trivial anwendbar sind, dass den EU-Ländern keine oder nur minimale Aufwände (wie in der INSPIRE-Richtlinie gefordert) durch die Harmonisierung entstehen. Aus diesem Grund wurden in den letzten Jahren zahlreiche Projekte gefördert, die sich entweder mit der konkreten Harmonisierung ganz bestimmter Datenbestände und -Themen befassen oder aber das Ziel haben, technische Werkzeuge zur Datenharmonisierung zu entwickeln.

Das EU-Projekt HUMBOLDT fällt in die letztgenannte Kategorie von Projekten. Es handelt sich um ein Projekt, das im Rahmen der EU-Umweltinitiative GMES (Global Monitoring for Environment and Security) gefördert wird. Es startete im Oktober 2006, hat eine Laufzeit von 4 Jahren und ein Gesamtprojektbudget von ca. 13,5 Millionen Euro. 28 Projektpartner aus 14 Ländern (darunter die Schweiz) entwickeln gemeinsam Konzepte und Werkzeuge zur Umsetzung von INSPIRE, mit Fokus auf den notwendigen Harmonisierungsprozessen. Das Projekt wird vom Fraunhofer-Institut für Graphische Datenverarbeitung (Darmstadt, Deutschland) koordiniert. Die Autorin dieses Artikels ist wissenschaftliche Koordinatorin (Scientific Manager) des Projektes im Auftrag des Partners ETH Zürich (Prof. Hurni).

Wichtige Ergebnisse

Im ersten Projektjahr von HUMBOLDT wurden in einer aufwendigen und umfangreichen Analyse der Stand der Technik im Bezug auf Harmonisierungsmethoden und Werkzeuge untersucht. Ebenso wurde eine zweistufige, intensive Benutzeranalyse durchgeführt um festzustellen, welche konkrete Harmonisierungsprobleme auftreten und wie gross deren Bedeutung für die Anwender ist.

Anwenderprobleme

Die Untersuchungen im Rahmen von HUMBOLDT zeigten sehr klar auf, dass viele der bestehenden Probleme bei der Umsetzung von INSPIRE eher organisatorischer als technischer Art sind. So wirken sich z.B. fehlende Prozesse und Vorgaben zur Erfassung und Verwaltung von Metadaten negativ auf die Nutzbarkeit der Geodaten aus. Es ist jedoch weitgehend unbestritten, dass es zwischenzeitlich ausreichende technische Voraussetzungen und Werkzeuge gibt, diese Probleme zu lösen.

Demgegenüber sind allerdings die tatsächlich noch vorhandenen technischen Probleme so gravierend, dass eine Nutzung von relevanten Geodaten in vielen Fällen sehr schlecht oder sogar gar nicht möglich ist. Die vier am häufigsten genannten Probleme im Bezug auf Harmonisierung waren (in dieser Reihenfolge):

- Probleme mit verschiedenen Datenformaten
- Probleme mit verschiedenen Datenmodellen
- Probleme mit fehlenden/inkonsistenten/nicht mehr aktuellen Metadaten
- Probleme mit unterschiedlichen Bedeutungen, d.h. mit der Semantik von Objekten

Es muss dazu bemerkt werden, dass das Know-how auf Seiten der Anwender im Bezug auf Datenharmonisierung sehr unterschiedlich war. Teilweise wurden mit hoher technischer Kompetenz sehr konkrete Effekte dokumentiert, teilweise wurden aber auch Probleme berichtet, von denen den Anwendern nicht klar war, dass es sich um Harmonisierungsprobleme handelte, bzw. welches technische Problem sich hinter den nicht-nutzbaren Daten verbarg. Aufgrund der im Rahmen von HUMBOLDT geführten Interviews vermuten wir daher, dass auch für den ersten Punkt, der die Datenformate betrifft, eher semantische als syntaktische Probleme eine Rolle spielen. Insbesondere, da es auf dem Markt bereits zahlreiche und weit verbreitete Werkzeuge zur Formatkonvertierung gibt.

Technische Ansätze

Im Rahmen von HUMBOLDT wurden insgesamt 176 Software-Werkzeuge, die (im weitesten Sinne) Geodatenharmonisierung unterstützen, untersucht. Ausserdem wurden ca. 60 nationale und internationale Projekte und dokumentierte Methoden zur Datenharmonisierung analysiert. Dabei wurde festgestellt, dass es eine ganze Reihe von immer wieder vorkommenden Schritten gibt, die im Rahmen einer Geodatenharmonisierung typischerweise ausgeführt werden (müssen). Diese Schritte lassen sich wie folgt in Hauptaufgaben oder Phasen gruppieren:

- Spezifikation der Harmonisierungsziele und des Harmonisierungsansatzes (dazu gehört auch die Definition eines konzeptionellen Datenmodells für das Ziel)
- Analyse der Quell- und Zieldaten sowie Definition der notwendigen Harmonisierungsschritte
- Ausführung der Prozessschritte für die Harmonisierung
- Sammeln von Feedback zum Harmonisierungsergebnis und allfällige Wiederholung einzelner Prozessschritte

Alle untersuchten Methoden und Ansätze zeigten, dass Geodatenharmonisierung grundsätzlich das Wissen und die Interaktion menschlicher Wesen erfordert. Vollautomatische Harmonisierungsprozesse wurden nicht dokumentiert und deren

Existenz sogar von einer Vielzahl von Autoren explizit in Zweifel gezogen. Nur nachdem ein erstes Mal ein vollständiger und erfolgreicher Harmonisierungsprozess durchgeführt wurde, lassen sich einzelne Prozessschritte oder in einfachen Fällen auch der ganze Prozess automatisieren.

Die weitaus grösste Anzahl der auf dem Markt befindlichen Software-Werkzeuge betreffen die syntaktische, d.h. unter anderem die strukturelle Interoperabilität der Geodaten, z.B. Formatkonverter oder Werkzeuge, die auf der Ebene der logischen Datenmodelle operieren. Wenige kommerzielle oder Open-Source Werkzeuge unterstützen semantische Interoperabilität, insbesondere mit dem modellbasierten Ansatz (MDA), wie er auch in der Informatik heute im Mainstream genutzt und von der OMG (Object Management Group (www.omg.org)) propagiert wird. Dagegen konnte festgestellt werden, dass ein Grossteil der untersuchten Forschungs- und Entwicklungsprojekte zur Datenharmonisierung sehr wohl auf dem MDA basieren. Der Grundsatz, konzeptionelle Modelle bzw. Ontologien möglichst formal und vollständig zu beschreiben, um anschliessend Werkzeuge zur Modell- bzw. Ontologie-Transformation für die Datenharmonisierung anwenden zu können, ist sogar in allen uns bekannten Ansätzen verfolgt worden.

Modellierungsprinzipien und Datenharmonisierung

Im zweiten Projektjahr von HUMBOLDT wurden im Rahmen von zwei Anwendungsszenarien (Katastrophenmanagement in Hochwasser-Situationen und Management von Schutzgebieten in grenzübergreifenden Regionen) konkrete Harmonisierungsprozesse untersucht und man versuchte, diese mit vorhandenen Werkzeugen zu lösen. Dies geschah vorrangig mit dem Ziel herauszufinden, wie vorhandene Werkzeuge erweitert werden müssten bzw. welche neuen Werkzeuge benötigt würden. Dabei ergaben sich jedoch auch sehr interessante Rückschlüsse auf die ursprünglich verwendeten Modellierungsprinzipien und verwendeten Werkzeuge. So wurden zum Beispiel von der ETHZ, der TU München und der Firma Intergraph Daten im Grenzgebiet um den Bodensee untersucht. Dabei trat der interessante Effekt auf, dass sich topografische Strassendaten aus Baden Württemberg und Bayern massiv bzgl. ihrer Struktur sowie der Anzahl, der Namen und Ausprägungen ihrer Attribute unterscheiden, obwohl dasselbe, gut dokumentierte konzeptionelle Modell (ATKIS) zu Grunde liegt und die Daten auch im selben Format (Shape File) von den beiden deutschen Bundesländern geliefert wurden. Das hier sich manifestierende Problem bestand in der unvollständigen Anwendung des MDA: die Kodierung, d.h. die Abbildung des konzeptionellen Modells auf ein konkretes Format wurde nie zentral festgelegt und in den beiden Bundesländern wurde dies einfach nach Interpretation verschie-

Alle untersuchten Methoden und Ansätze zeigten, dass Geodatenharmonisierung grundsätzlich das Wissen und die Interaktion menschlicher Wesen erfordert.

dener Sachbearbeiter vorgenommen. Keine Schema-Transformation und keine Formatkonvertierung sind selbstverständlich in der Lage dies zu lösen. Einzig mühsame Interpretation und Handarbeit und unterschiedliche Transformationen für die beiden Bundesländer in ein anderes (wohldefiniertes) Datenmodell lösen das Problem.

Noch schwieriger als das gerade beschriebene Problem der unvollständigen Nutzung des MDA, gestaltet sich jedoch ein anderer Effekt. Im Rahmen der Harmonisierungstests in HUMBOLDT trat deutlich zu Tage, dass es für die konzeptionelle Modellierung von Geodaten (im Rahmen der Anwendung des MDA) grundsätzlich zwei verschiedene Prinzipien mit sehr unterschiedlichen Konsequenzen für die Harmonisierung gibt:

- *Bereitstellung einer maximalen Anzahl vordefinierter Elemente.* Die Modellierung erfolgt durch Einschränkungen oder Verbot bestimmter Elemente sowie Erweiterung vorhandener Elemente.
- *Bereitstellung einer minimalen Anzahl vordefinierter Elemente.* Die Modellierung erfolgt ausschliesslich durch Erweiterung vorhandener Elemente.

Das erste Prinzip ist in den Normen von ISO/TC211 verankert, auf denen zur Zeit auch alle INSPIRE-Vorgaben aufbauen. Das zweite Prinzip wird beispielsweise in INTERLIS verfolgt und wurde auch als Grundlage für die weiteren technischen Entwicklungen im Projekt HUMBOLDT gewählt.

Beide Prinzipien haben verschiedene Vor- und Nachteile für Anwender und für technische Werkzeuge zur Unterstützung der Harmonisierung, die in der nebenstehenden Tabelle sehr plakativ und grob zusammengefasst sind.

Studiert man diese Tabelle etwas, merkt man schnell, dass beide Modellierungs-Prinzipien ihre Berechtigung haben und je nach Zielsetzung das eine oder das andere Prinzip vorzuziehen ist. Grundsätzlich besteht die Möglichkeit, Modelle, die nach Prinzip 2 erstellt wurden, in Modelle gemäss Prinzip 1 umzuwandeln und umgekehrt. Man sollte sich jedoch auch bewusst machen, dass in beiden Fällen verschiedene Dinge (Teilaspekte der Realität) abgebildet werden und dass eine solche Transformation immer mit einem grundsätzlichen Informationsverlust sowie aber auch mit einem Informationsgewinn verbunden ist. Letzter kann selbstverständlich nicht automatisch erzeugt werden. Er erfordert menschliches Expertenwissen, das in der Definition der Transformation eingebracht werden muss.

Prinzip	Maximale Anzahl Elemente/Modellierung durch Einschränkung	Minimale Anzahl Elemente/Modellierung durch Erweiterung
Vorteile für Anwender	<ul style="list-style-type: none"> – Fast keine Notwendigkeit (auch nicht in Spezialfällen) eigene Konzepte erfinden zu müssen 	<ul style="list-style-type: none"> – Die minimalen Elemente können so gewählt werden, dass sie per Definition eine grosse Anzahl der Einschränkungen (z.B. INSPIRE requirements and recommendations) erfüllen. Anwender müssen daher die Einschränkungen nicht mehr explizit kennen, um sie einzuhalten – Da Software-Werkzeuge nur die Basiselemente unterstützen müssen, sind diese immer implementiert und der Harmonisierungsprozess (im einfachsten Fall ein Datentransfer) kann immer durchgeführt werden
Nachteile für Anwender	<ul style="list-style-type: none"> – Einschränkungen (z.B. OCL constraints aber auch zahlreiche INSPIRE requirements and recommendations) müssen vor dem Modellieren gelesen, verstanden und beachtet werden – Da so viele Elemente prinzipiell vorkommen können, sind niemals alle in allen Werkzeugen implementiert, dadurch wird allenfalls die Weiterverarbeitung in folgenden Prozessschritten verunmöglicht – Es entstehen sehr grosse, komplexe Modelle (z.B. in UML), von denen anschliessend Teile explizit per Text oder Beschreibungssprache (z.B. OCL) wieder ausgeschlossen werden 	<ul style="list-style-type: none"> – Elemente, Strukturen und Konzepte, die noch nicht vorhanden sind, müssen selbst entworfen und auf die vorhandenen Basiselemente abgebildet werden – Selbst hinzugefügte, neue Strukturen und Konzepte können allenfalls von den Software-Werkzeugen nicht ideal verarbeitet werden
Vorteile für Implementierung	<ul style="list-style-type: none"> – (Fast) alle Elemente und Strukturen, die vorkommen können, sind von vornherein bekannt und können implementiert und daher in Werkzeugen zur Weiterverarbeitung unterstützt werden 	<ul style="list-style-type: none"> – Klare, stabile Vorgaben für notwendige Datenstrukturen durch wenige zu unterstützende Elemente – Software-Werkzeuge funktionieren immer – Harmonisierungsprozess muss nie wegen inkompatiblen Schnittstellen abgebrochen werden – Automatische Überprüfung bzgl. Einschränkungen (Qualitäts-Checks) möglich
Nachteile für Implementierung	<ul style="list-style-type: none"> – Automatische Überprüfung, ob Einschränkungen eingehalten wurden, nicht möglich – Die Anzahl der möglicherweise vorkommenden Elemente und Strukturen ist so gross, dass es ineffizient wäre, alle zu implementieren 	<ul style="list-style-type: none"> – Funktionalität für neue Elemente eines bestimmten Anwenders muss explizit hinzugefügt werden

In den konkreten Tests im Rahmen von HUMBOLDT verursachten die unterschiedlichen Modellierungsprinzipien insbesondere folgende Probleme:

Wenn man sich unsicher war, ob gelieferte Daten tatsächlich dem dokumentierten Datenmodell entsprachen, war nur eine mühsame manuelle Überprüfung möglich.

Wenn ein eigentlich sehr einfaches Ziel-Datenmodell (z.B. EuroRoadS, www.euroroads.org) nach Prinzip 1 modelliert vorlag, musste man sich in der Analysephase mit einem relativ komplexen UML-Modell auseinandersetzen, bei dem einzelne Attribute oder Klassen explizit durch Einschränkungen ausgeschlossen waren, was im reinen UML-Modell nicht erkennbar war.

Lösungsansätze in HUMBOLDT

Wie bereits angedeutet, wird in HUMBOLDT klar der MDA zur Entwicklung von Werkzeugen für die Unterstützung der Datenharmonisierungsprozesse genutzt. Zum gegenwärtigen Zeitpunkt wurde auch bereits ein Modelleditor zur Beschreibungen von konzeptionellen Geodatenmodellen entwickelt, der das Prinzip 2 (minimale Anzahl vorgegebener Elemente) zur Modellierung umsetzt. Der Editor erlaubt nur solche Modelle zu erstellen, die auch den Vorgaben von INSPIRE zur Datenmodellierung entsprechen. Damit wollte man einerseits den Anwendern in den 8 Anwendungsszenarien in HUMBOLDT ermöglichen, korrekt und INSPIRE-konform zu modellieren, ohne sich mit den ca. 700 (!) Seiten Dokumentation sowie zusätzlichen mit allen Normen von ISO/TC211 auseinandersetzen zu müssen.

Ausserdem ist ein klares Ziel von HUMBOLDT, allgemeine und in vielen Anwendungen notwendige Harmonisierungsschritte zu unterstützen. Eine Beschränkung auf wenige Basiselemente widerspricht dem in keiner Weise und reduziert aber gleichzeitig deutlich den Aufwand und die Komplexität der Implementierungen von Werkzeugen für die Harmonisierungsschritte. Zur Zeit wird in HUMBOLDT intensiv an der Entwicklung eines Alignment Editors gearbeitet, der die Datenmodelle aus dem Modelleditor lesen und als Vorgaben für Quell- und Zieldatenmodelle nutzen wird. Der Alignment Editor wird dann Werkzeuge anbieten, um möglichst einfach die notwendigen Transformationen zwischen Quell- und Zielmodell definieren zu können. Für diese Arbeiten wurde auch im Vorfeld intensiv mit den Mitarbeiter(inne)n der ETHZ aus der Gruppe von Prof. Carosio sowie weiteren Schweizer Expert(inn)en aus dem INTERLIS-Umfeld aber auch anderen internationalen Wissenschaftler(inne)n diskutiert, die sich bereits seit vielen Jahren mit dieser Problematik befassen. Insbesondere fliessen Ergebnisse des Projektes «Modellbasierter Ansatz für den Web-Zugriff auf verteilte Geodaten» (Deutsch-Schweizerisches Kooperationsprojekt) in die Arbeiten von HUMBOLDT ein.

Weitere Informationen zu INSPIRE, zur schweizerischen INSPIRE-Kontaktstelle und zu HUMBOLDT finden sich im Internet:

- <http://inspire.jrc.ec.europa.eu>
- www.swisstopo.admin.ch > INSPIRE
- <http://esdi-humboldt.eu>

Voraussetzungen und Hindernisse zur interoperablen Nutzung verteilter Geodaten in GDI



Peter Staub,
swisstopo

Interoperabilität: Status quo und Begriff, Anforderungen aus GDI

Ohne Zweifel ist die Realisierung interoperabler Lösungen im Rahmen von GDI für die gemeinsame Nutzung verteilter Geodaten und Systeme essentiell. Die Anwendung von standardbasierten Diensten zum Informations- und Datenaustausch ermöglicht Interoperabilität auf Formatebene durch die einheitliche Verwendung eines Transferformates. Dies lässt dem Nutzer jedoch keine Möglichkeit, die Struktur der bezogenen Daten im Detail einzusehen oder gar seinen Bedürfnissen anzupassen. Mit anderen Worten: Interoperabilität ist nur zum Teil realisiert. Ein Transferformat übermittelt im Allgemeinen nicht die komplette Modellinformation. Der Zugang zu solchen Informationen und darüber hinaus auch die Möglichkeit, Datenstrukturen gemäss Nutzerbedürfnissen anzupassen, stellen in bestehenden Ansätzen für die Realisierung umfassender Interoperabilität grosse Hürden dar.

«Interoperabilität» ist ein facettenreicher Begriff. Grundsätzlich kann grob zwischen organisatorischer und technischer Interoperabilität unterschieden werden (Abb. 1). Erstere wird durch Normen und Standards sichergestellt, wohingegen die technische Ebene durch verschiedene Methoden und Werkzeuge realisiert wird. Der Datentransfer und die Services (Dienste) sorgen für die so genannte syntaktische Interoperabilität. Profile als praktikable Anwendungen von Normen und Standards sowie die konzeptionelle Datenmodellierung stellen für die so genannte semantische Interoperabilität ebenso eine wesentliche Voraussetzung dar wie die semantische Transformation durch Modellstruktur-Umbau.

GDI-Entwicklungen stellen ganz spezifische Anforderungen an die Interoperabilität. So fordert beispielsweise die INSPIRE-Richtlinie in Art. 3, Abs. 7: «Im Sinne dieser Richtlinie bezeichnet [...] «Interoperabilität» im Falle von Geodatensätzen ihre mögliche Kombination und im Falle von Diensten ihre mögliche Interaktion [...]».

Im Zusammenhang mit den Durchführungsbestimmungen wird in Art. 7 der Richtlinie explizit die Harmonisierung der Datensätze verlangt. Offensichtlich ist es wirtschaftlich und politisch nicht denkbar, den Datenherren vorzuschreiben, ihre bestehenden Modelle zugunsten eines harmonisierten Modells aufzugeben. Der einzige praktikable Weg führt über Modelltransformationen, so dass aus bestehenden Anbietermodellen Geodatensätze in ein harmonisiertes Datenmodell einer GDI integriert werden

können. Der allgemeine Ansatz setzt einen Transformationsmechanismus voraus, der es einem Nutzer erlaubt, beliebige Geodatensätze seinem eigenen Modell entsprechend umzustrukturieren und zu beziehen.

Erfahrungen mit semantischer Interoperabilität

Im Rahmen eines Forschungsprojektes in Zusammenarbeit der ETH Zürich mit der TU München ist ein Lösungsansatz für die semantische Interoperabilität im Rahmen der Datenintegration für GDI entwickelt und getestet worden (vgl. Artikel Schilcher et. al.). Dieser Ansatz kombiniert die Methoden der syntaktischen Interoperabilität im Sinne von OGC Webdiensten mit der semantischen Transformation. Damit wird es möglich, im Rahmen von GDI heterogene, verteilte Geodaten integriert zu nutzen und in entsprechenden Anwendungen einen Mehrwert zu erzielen.

Es wurden verschiedene Datenmodelle aus der Schweiz (GG25, VECTOR25), aus Deutschland (AFIS-ATKIS-ALKIS) sowie auf europäischer Ebene (EuroSpecFC, INSPIRE Data Specification) untersucht. Dabei stand die Integration der Ländermodelle in die EU-Modelle im Vordergrund. Die transformierten Datenmodelle beschreiben jeweils dieselbe Thematik; in einer Pilotanwendung sind dies administrative Gebiets-einheiten; weitere Anwendungen behandeln umweltrelevante Themen wie Schutzgebiete oder das Gewässernetz.

Bei der Definition der Abbildungsregeln von Quell- auf Zielmodelle wird deutlich, dass Modelltransformationen auf konzeptioneller Ebene ziemlich anspruchsvoll sind. Semantische Unterschiede können bisweilen ohne zusätzliche Informationen nicht oder nur ungenügend erkannt und in Übereinstimmung gebracht werden. Die unterschiedliche Herangehensweise

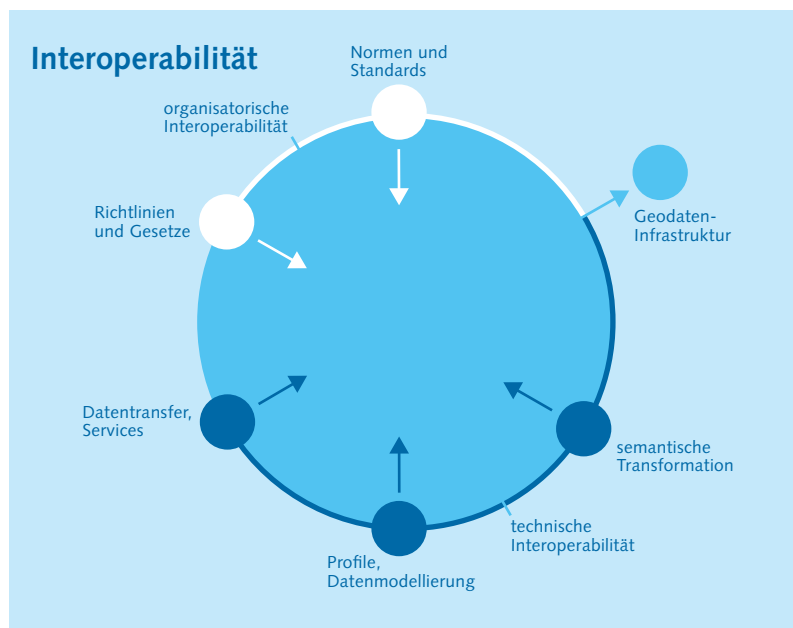


Abb. 1: Charakteristika des Interoperabilitäts-Begriffs

Interoperabilität ist nur zum Teil realisiert. Ein Transferformat übermittelt im Allgemeinen nicht die komplette Modellinformation.

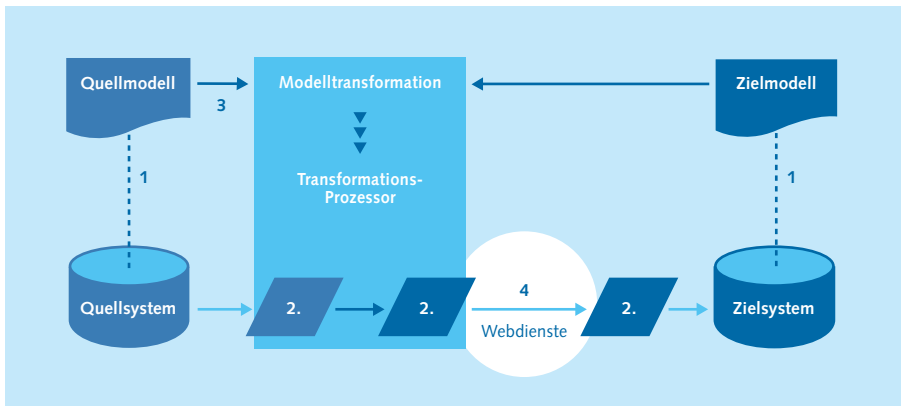


Abb. 2: Voraussetzungen für umfassende Interoperabilität

Die Erfahrung mit der erwähnten Pilotanwendung zeigt, dass das Potenzial der semantischen Transformation unabhängig von Diensten genutzt werden kann.

der Fachgemeinschaften und Länder an ein Datenmodell resultieren in verschiedenen «Modellierungsphilosophien», die eine Transformation grundsätzlich erschweren.

Potenzial und Grenzen, wirtschaftlicher Einsatz

Die Erfahrung mit der erwähnten Pilotanwendung zeigt, dass das Potenzial der semantischen Transformation unabhängig von Diensten genutzt werden kann. Dies hat insbesondere für den modellbasierten Datentransfer ausserhalb von GDI eine Bedeutung. In erster Linie stellt die Modelltransformation im Zusammenhang mit der geforderten semantischen Interoperabilität im Rahmen von GDI ein wesentlicher Fortschritt dar.

Wie oben gezeigt, ist die Machbarkeit von Modelltransformationen durch die unterschiedliche Konzeption der Modelle limitiert. Ein Modell «zur gleichen Thematik» bedeutet nicht, dass die komplette Modellinformation transformiert werden kann. Dies ist in der unterschiedlichen fachlichen Sicht auf das zu modellierende Thema begründet.

Wenn verschiedene Sprachen zur Datenmodellierung verwendet werden, ist die Definition von Abbildungsregeln teilweise unklar. Weiter erschweren verschiedene Transferformate die syntaktische Interoperabilität. Bei Formatkonversionen ist im Allgemeinen davon auszugehen, dass Informationen verloren gehen.

Schliesslich sind ganz praktische Grenzen erkennbar, denn es ist davon auszugehen, dass ein bestimmtes Quellmodell nur einen Teil eines Zielmodells abdeckt; mit weiteren Quellmodellen ist die gewünschte Information herzuleiten.

Durch folgende Voraussetzungen können oben genannte Grenzen zumindest weitgehend überwunden werden und umfassende Interoperabilität wird realisierbar (Abb. 2):

- Klare und eindeutige Spezifikation von Modellierungssprachen sowie deren einheitliche Anwendung. Je nach Umfang und Brauchbarkeit dieser Spezifikationen sind Anwendungsprofile zu definieren (1).
- Einheitliche und eindeutige Definition von Kodierungsregeln für zu transferierende Datenobjekte (2).
- Spezifikation einer konzeptionellen Sprache zur Definition von Modelltransformationen und Implementierung eines Transformationsprozessors mit entsprechender Funktionalität (3).
- Anwendung von etablierten Standards zum Informations- und Datenaustausch über (Web-)Dienste; Erweiterung dieser Dienste, um auch Modellinformationen transferieren zu können. (4)

Literatur:

P. Staub: Über das Potenzial und die Grenzen der semantischen Interoperabilität von Geodaten – Ein operationelles Verfahren zur Nutzung verteilter Systeme in Geodaten-Infrastrukturen. Diss. ETH Nr.18201, ETH Zürich, 2009



Andreas Morf
FKL & Partner AG, ETH Zürich

Motivation

In vielen Fachinformationsgemeinschaften werden mit der Einführung des Geoinformationsgesetzes weitere Datenmodelle für die bereit-zustellenden Geodaten entstehen. Diese neue Ausgangslage muss als Ansporn für eine verbesserte Interoperabilität von Geodaten zum Aufbau von Geodateninfrastrukturen genutzt werden.

Das Subsidiaritätsprinzip ist hier aber wohl einer der wesentlichen Gründe, dass auch bei guter Koordination und Kommunikation der Fachexperten Variationen in den Modellen und Datenstrukturen weiterhin vorhanden sein werden. Oder etwas provokativ ausgedrückt: Mit dem kleinsten gemeinsamen Nenner wird wohl ein Teil der Ansprüche befriedigt – die Bedürfnisse an zusätzlicher Information wird aber sicher zu Erweiterungen und damit zu einer gehörigen Modellvielfalt auch im selben Fachgebiet führen. Ganz zu schweigen von den sich im Aufbau befindlichen Infrastrukturen in Grenzregionen und auf europäischer Ebene.

Es wird also auf absehbare Zeit nicht zu vermeiden sein, dass zum Beispiel eine Gebietskörperschaft A denselben Aspekt anders modelliert wie das die Gebietskörperschaft B tut.

Bei verteilter Nutzung beziehungsweise der Integration von Geodatenbeständen wurden bislang (mit Erfolg) sogenannte ETL-Tools (Extract, Transform, Load) verwendet. Beispiele für derartige Software sind FME (Safe Software), ICS (Infogrips Conversion System) oder spezifisch erstellte Programme und Schnittstellen.

Jede dieser Lösungen wird mit einer proprietären Oberfläche oder Sprache bedient und erfordert den Aufbau spezifischen Know-hows – und allzu oft muss daher jemand, der mit einer Transformationsaufgabe konfrontiert ist mehr oder weniger bei Null beginnen.

Wenn man Geodaten schon systemunabhängig auf konzeptioneller Ebene mit UML und INTERLIS modelliert sollte auch die Beschreibung der Transformationen von Geodaten auf diesem Niveau möglich sein!

Grundlagen

Das Thema «Modelltransformation» ist im Modellierungsumfeld bereits seit längerem ein viel diskutiertes Thema. Auch die OMG (Object Management Group), welche die Modellierungssprache UML bereitstellt und pflegt, hat bereits einen Standard zu diesem Thema verabschiedet: MOF-QVT (Query, Views, Transformations). Ziel dieses Standards ist jedoch nicht so sehr die konzeptionelle Beschreibung von Transforma-

tionen zwischen Datenmodellen, sondern die Umsetzung der Modelle in Richtung Implementierung von Software-Systemen.

Kommt noch dazu, dass die Komplexität dieses Standards gross und die Verständlichkeit der damit formulierten Transformationsmodelle für den Nicht-Informatiker kaum gegeben ist.

Anforderungen

Aufgrund bisheriger Erfahrungen mit der Modellierung und Transformation von Geodaten scheinen die folgenden Komponenten für eine Transformationssprache unabdingbar:

- Grafischer Formalismus, welcher die Darstellung der Transformationsmodelle in intuitiver Form erlaubt.
- Textuelle Notation analog zu den INTERLIS «ili»-Dateien, welche insbesondere für die rechnergestützte Verarbeitung verwendet werden kann.
- Funktionalität zur Behandlung von strukturierten Geodatentypen (Inspektion, Konstruktion und Analyse von Geometrie-Attributen).
- Anbindung von benutzerdefinierten Abbildungsfunktionen zur Realisierung spezieller Aufgaben und Bedürfnisse (Koordinatentransformationen, GIS-Analysefunktionen, Datenprüfprozesse etc.)
- Eine prototypische Implementierung muss zudem die Umsetzbarkeit anhand vorhandener Modelle und Daten aufzeigen.

UMLT: konzeptuelle Transformationssprache für Geodaten

Zur Modellierung von Softwaresystemen und -architekturen bietet die Modellierungssprache UML neben den in der Geoinformatik häufig angewendeten Klassendiagrammen eine Auswahl an weiteren Diagrammtypen. Das Aktivitätsdiagramm mit welchem die dynamischen Aspekte (Verhalten) des modellierten Systems beschrieben werden können, eignet sich insbesondere auf den zweiten Blick erstaunlich gut zur Beschreibung von Transformationsprozessen: mit der Version 2 von UML wurde dieser Diagrammtyp stark den sogenannten Petri-Netzen angeglichen, welche in der Informatik seit geraumer Zeit zur Beschreibung von nebenläufigen Systemen und Transformationsprozessen verwendet werden.

Mit einem Aktivitätsdiagramm kann eine Modelltransformation als gerichtetes Netz von einzelnen Aktionen, welche die einzelnen Schritte der Transformation darstellen, aufgebaut werden. Die Kanten des Netzes werden durch den Objektfluss und einen fakultativen Kontrollfluss realisiert, welcher den Ablauf des gesamten Transformationsprozesses ermöglicht. Das folgende Beispiel (Abb. 1) in UML veranschaulicht dies anhand der Bildung eines Flächennetzes aus Begrenzungslinien und Gebietsreferenzpunkten (Centroide).

Die Erarbeitung von Modell-Abbildungen führt automatisch zu Überlegungen und Erfahrungenwerten betreffend Datenmodellierung.

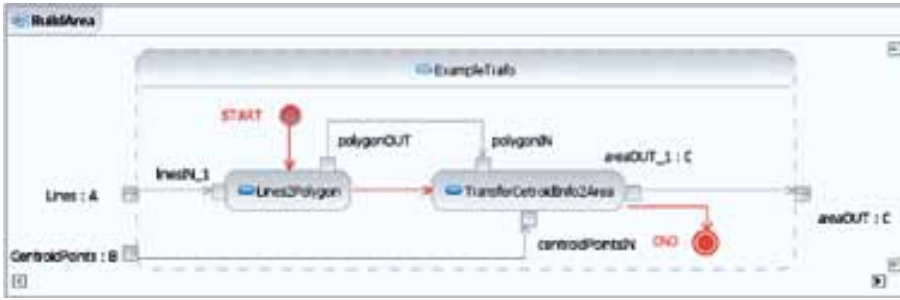


Abb. 1: ●●●Legende●●●

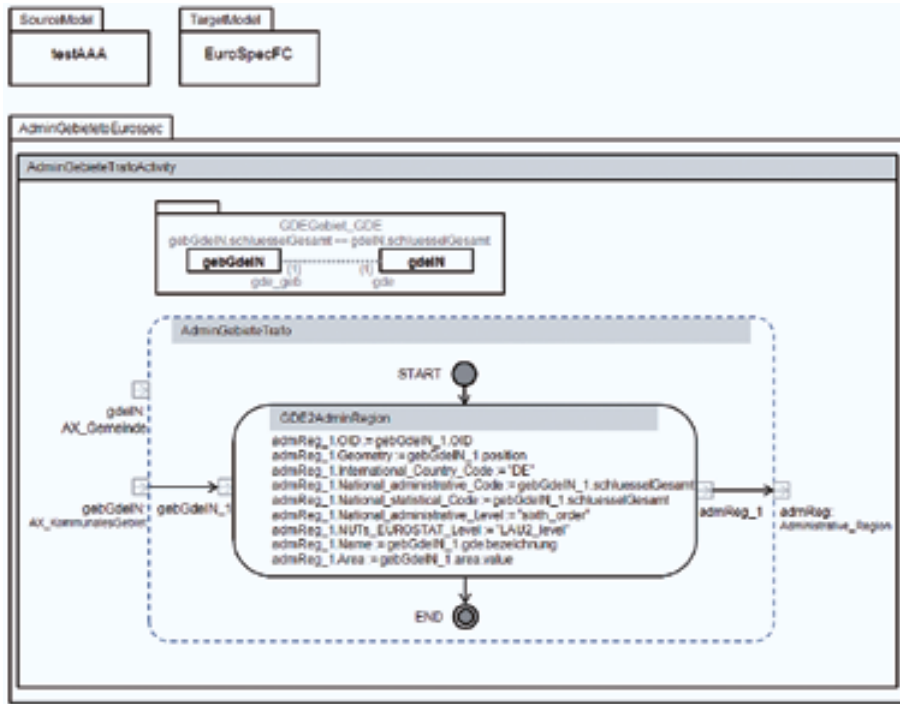


Abb. 2: Beispiel für eine Transformation in UMLT. Das Diagramm beschreibt die Abbildung der Klasse «Gemeindegebiete» des deutschen AAA-Modells (AFIS, ALKIS, ATKIS: amtli. Vermessung und Topografie) auf das Äquivalent «Administrative_Region» des EuroSpec FeatureCatalog.

Die Elementarbausteine einer solchen strukturierten Aktivität sind die sogenannten UML-«Actions» (graue, abgerundete Rechtecke). Darin findet die Abbildung eines oder mehrerer Geodaten-Objekte in eine neue Struktur statt. Die abzubildenden Daten- und Beziehungsobjekte werden über den Objektfluss (graue Pfeile) den Actions über die sogenannten Input-Pins zugeführt bzw. ab den Output-Pins entnommen. Der Kontrollfluss (rot) steuert die Reihenfolge des Ablaufs. Er ist fakultativ da die Logik des Ablaufs meistens bereits durch den Objektfluss hinreichend spezifiziert ist. Im vorangehenden Beispiel dient er nur der Veranschaulichung des Konzeptes).

Ein wesentlicher Bestandteil fehlt nun aber noch: Die Spezifikation der Abbildungsfunktion in den einzelnen Aktionen zur Überführung der ankommenden Objekte in die gewünschte Zielstruktur gemäss Typ des Output-Pins.

Da die UML hier keine Möglichkeiten bietet, wird eine Erweiterung des UML Metamodells vorgenommen, welche die Spezifikation dieser Abbildungsfunktionen erlaubt. Die Notation wird innerhalb der Aktionen in der Form von Zuweisungen vorgenommen (vgl. Abb. 2).

Zusätzliche Erweiterungen der UML sind für die Spezifikation der Transformation unabdingbar:

- Bildungsregeln für sogenannte virtuelle oder abgeleitete Beziehungen (zum Beispiel aus vorhandenen Schlüsselwerten oder der Berechnung von räumlichen Beziehungen).
- Bindungsart von Beziehungen/Rollen (vergleichbar mit dem Join-Typ in relationalen Datenbanken).
- Spezifikation von Selektionskriterien mit logischen Ausdrücken
- Steuerung der Handhabung von Objekt-identifikatoren

Bezug zu INTERLIS 2

Transformationsmodelle, welche mit UMLT erstellt werden referenzieren die Klassen und Pakete (TOPICS) der UML-Geodatenmodelle. Die textuelle Notation von UMLT orientiert sich stark an der INTERLIS 2 Syntax und verwendet wo möglich auch deren Konstrukte.

Die meisten Konzepte und Regeln von INTERLIS 2 haben sich auch in diesem Kontext bewährt – der Bereich der algebraischen und logischen Ausdrücke sollte jedoch im Sinne der Vollständigkeit überarbeitet werden. Ebenso wäre eine Angleichung der Notation von Ausdrücken in INTERLIS und UML wünschenswert.

Implementierung und Ausblick

Die prototypische Implementierung von UMLT im Rahmen des ETH-TUM Forschungsprojektes hat gezeigt, dass die automatische Umsetzung der Transformationsregeln mit einem kommerziellen räumlichen Datenbanksystem möglich ist. In einer weiteren Projektphase wird zusätzlich die Umsetzung mit dem ETL-Instrument FME erarbeitet.

Neben den theoretischen und technischen Erkenntnissen scheint folgender Punkt von Bedeutung: Die Erarbeitung von Modell-Abbildungen führt automatisch zu Überlegungen und Erfahrungswerten betreffend Datenmodellierung. Ein willkommener Nebeneffekt, wenn damit die Qualität anschliessend erstellter Geodatenmodelle zunimmt!

Auskünfte:

e-geo.ch
c/o Bundesamt für Landestopografie
Seftigenstrasse 264, Postfach
CH-3084 Wabern

Telefon +41 31 963 21 11
Fax +41 31 963 24 59
e-mail info@e-geo.ch
www.e-geo.ch

Herausgeber: e-geo.ch

Redaktion: René Sonney, e-geo.ch
Konzept: MKR Consulting AG, Bern
Gestaltung: Atelier Ursula Heilig SGD, Gümliigen
Druck: swisstopo
Auflage: 2450 Exemplare (1800 deutsch, 650 franz.)
Bilder: KOGIS, e-geo.ch, Autoren

Antworttalon

Sie können die Antwortkarte faxen (031 963 24 59) oder Ihre Bestellung mailen an info@e-geo.ch.
Immer aktuell informiert über das Programm e-geo.ch:

Bitte senden Sie uns regelmässig den **Newsletter e-geo.ch**:

Anzahl Exemplare deutsch Anzahl Exemplare französisch
 per Post per E-Mail

Bitte senden Sie uns die Broschüre «**Das Umsetzungskonzept zur Strategie für Geoinformation beim Bund**»:

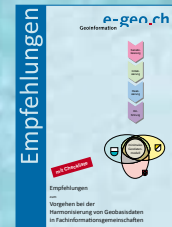
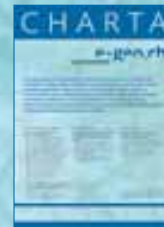
Anzahl Exemplare deutsch Anzahl Exemplare französisch

Bitte senden Sie uns die **Charta e-geo.ch**:

Anzahl Exemplare deutsch Anzahl Exemplare französisch

Bitte senden Sie uns die Broschüre «**Empfehlungen zum Vorgehen bei der Harmonisierung von Geobasisdaten**»:

Anzahl Exemplare deutsch Anzahl Exemplare französisch



Organisation/Firma

Name, Vorname

Adresse

E-Mail

Blank white form fields for entering contact information.