Bundesamt für Landestopografie swisstopo, Koordination, Geoinformation und Services (KOGIS)

SEMANTISCHE INTEROPERABILITÄT – DER SCHLÜSSEL ZUR INTEGRIERTEN NUTZUNG HETEROGENER GEODATEN

Peter Staub

Zusammenfassung: Die gemeinsame Nutzung verteilter, heterogener Geodaten wird im Rahmen von Geodaten-Infrastrukturen immer wichtiger. Methoden der Interoperabilität bieten hierfür eine Lösung. Etabliert ist vor allem syntaktische Interoperabilität mittels OGC Webservices. Um semantische Interoperabilität zu erreichen, müssen neben den Daten auch Modellinformationen zugänglich gemacht werden und es braucht einen Mechanismus, um semantische Transformationen zu definieren und auszuführen. Ein praktischer Lösungsansatz dafür wird beschrieben.
Wesentlich ist die Erkenntnis, dass die konzeptionelle Datenmodellierung und die konzeptionelle Definition von Abbildungsregeln für Modelltransformationen essentiell sind. Durch die einheitliche Definition von konzeptionellen Schemasprachen wird zudem eine übergeordnete Ebene der Interoperabilität erreicht. Der Einsatz der semantischen Interoperabilität wird allerdings durch konzeptionelle, technische und praktische Grenzen eingeschränkt.

Schlüsselwörter: Interoperabilität, Semantik, Datenmodellierung, Modelltransformation, GDI

// SEMANTIC INTEROPERABILITY – THE KEY TO INTEGRATED USE OF HETEROGENEOUS GEODATA

//Abstract: The need for an integrated use of distributed, heterogeneous geodata is growing, particularly in the context of spatial data infrastructures. Interoperability methods provide a solution thereto whereas syntactic interoperability using OGC web services is established most. To reach the level of semantic interoperability, model information must be made accessible on top of data exchange. Furthermore, a means to both define and execute semantic transformations is needed. A pragmatic solution approach for that problem is described.
It is important to understand that conceptual data modelling and the conceptual definition of mapping rules for model transformations are essential. Through the coherent specification of conceptual schema languages, a higher level of interoperability is achieved. However, the application of semantic interoperability is constrained due to conceptual, technical, and practical limitations.

Keywords: Interoperability, semantics, data modelling, model transformation, SDI

Anschrift des Autors

Dr. sc. ETH Peter Staub
Bundesamt für Landestopografie swisstopo
Koordination, Geoinformation und Services (KOGIS)
Seftigenstrasse 264
CH-3084 Wabern
E: peter.staub@swisstopo.ch

1. EINFÜHRUNG

Interoperable Lösungen zur integrierten Nutzung verteilter heterogener Geodaten entsprechen einem großen Bedürfnis – insbesondere im Zusammenhang mit der Realisierung von Geodaten-Infrastrukturen (GDI) (vgl. Abschnitt 2.3). Grundlegend für umfassende Interoperabilität sind aber zunächst nicht technische Lösungen, sondern das gemeinsame Verständnis der Semantik von Geodaten. In diesem Zusammenhang erhält die semantische Datenmodellierung eine sehr große Bedeutung. Mit der so genannten modellbasierten Methode werden Datenstrukturen unabhängig von Systemen und Formaten konzeptionell beschrieben und mit geeigneten Mechanismen automatisiert in ein logisches Schema oder ein Transferdatenformat überführt. Die Konzepte für die Beschreibung der Semantik und der Strukturierung sowie die Kodierungsregeln sind dabei eindeutig spezifiziert (Eidgenössische Vermessungsdirektion 1999).

Ausgehend von der modellbasierten Methode zur Datenmodellierung können mit semantischen Modelltransformationen unterschiedliche Datenstrukturen durch die konzeptionelle Definition von Abbildungsregeln aufeinander abgebildet werden. Im Umfeld des Open Geospatial Consortium (OGC) spielt hingegen der Format-basierte Informations- und Datenaustausch über Geowebdienste die Hauptrolle. Beide Methoden weisen spezifische Defizite auf. Die semantische Transformation bietet keine Handhabe für webbasierten Informationsaustausch; die OGC-konforme Interoperabilität bietet keinen Zugang zu Modellinformationen, womit semantische Interoperabilität ausgeschlossen ist.

Die Kombination von standardbasiertem Datenaustausch und semantischen Modelltransformationen ermöglicht ein hohes Maß an Interoperabilität. Die semantische Interoperabilität besitzt ein großes Potenzial wie in Kapitel 4 gezeigt wird, unterliegt aber auch konzeptionellen, technischen und praktischen Grenzen, die qualitativ bestimmt werden können (vgl. Abschnitt 5.3).

Nach einer umfassenden Betrachtung des Interoperabilitätsbegriffs in Kapitel 2 werden Schwerpunkte der konzeptionellen Datenmodellierung erläutert (Kapitel 3). Im Zusammenhang mit GDI erhält die (semantische) Interoperabilität eine große Bedeutung für die im Rahmen eines Forschungsprojektes in Zusammenarbeit von ETH Zürich und TU München ein praktischer Lösungsansatz entwickelt wurde (Kapitel 4). Das Forschungsprojekt wurde gemeinsam vom Bundesamt für Landestopografie swisstopo (Wabern, Schweiz) und vom Bundesamt für Geodäsie und Kartographie BKG (Frankfurt/M., Deutschland) finanziert.

Der vorliegende Artikel basiert auf der Dissertation des Autors an der ETH Zürich (Staub 2009).

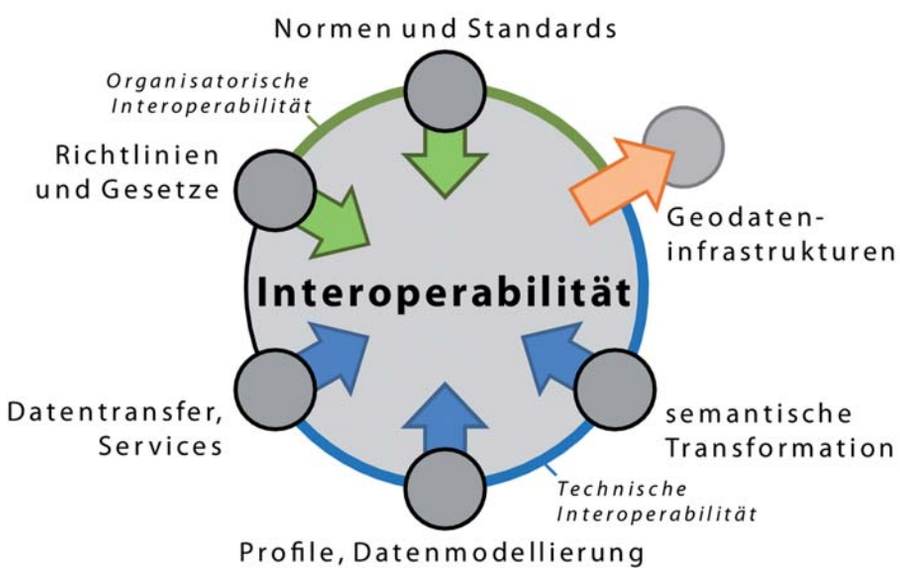

Abb. 1 – Charakterisierung der Interoperabilität

2. INTEROPERABILITÄT – EINE BEGRIFFSKLÄRUNG

2.1. EBENEN DER INTEROPERABILITÄT

Im Zuge der Entwicklung von GDI wird die Problematik der Nutzung heterogener Daten akut. Durch die große Verbreitung von Geowebdiensten gemäß Standards des OGC entsteht mithin der Eindruck, dass der Begriff *„Interoperabilität"* mit *„OGC Webservices"* gleichzusetzen sei. Für ein umfassendes Verständnis greift dieser Schluss jedoch zu kurz. Interoperabilität ist kein für sich allein stehendes Konzept sondern vielmehr das Zusammenspiel verschiedener Faktoren. Eine anerkannte Definition lautet: *„Interoperabilität ist die Fähigkeit zur Kommunikation, zur Ausführung von Programmen und zum Austausch von Daten zwischen funktionalen Einheiten in einer Art und Weise, die von Anwendern wenige oder gar keine Kenntnisse über die Besonderheiten dieser Einheiten erfordert."* (Bartelme 2005).

Konkret sind es verschiedene technische und organisatorische Aspekte, die den Begriff *„Interoperabilität"* charakterisieren. Insbesondere hängen umfassende Interoperabilität und funktionierende GDI sehr eng zusammen. In folgender Abbildung 1 ist das Zusammenspiel organisatorischer Aspekte (Richtlinien/Gesetze; Normen/Standards) (Buser 2005) und technischer Aspekte (Datentransfer/Services; Profile/Datenmodellierung; semantische Transformation) schematisch dargestellt.

Normen und Standards im Bereich der Geoinformation sind detaillierte technische Dokumente, die beispielsweise Format-Kodierungsregeln oder Schnittstellen zu Diensten festlegen (OGC 2008; OGC 2010). Wesentlich ist aber auch der organisatorische Charakter von Normen und Standards. Die Festlegung entsprechender Spezifikationen und deren einheitliche Anwendung leisten einen wichtigen Beitrag zur Realisierung (organisatorischer) Interoperabilität.

Aus technischer Sicht sind drei unterschiedliche Ebenen der Interoperabilität zu identifizieren. Auf der untersten Ebene steht die *„syntaktische Interoperabilität"* oder *„Interoperabilität der Systeme"*. Diese Ebene wird durch Standard-Transferformate und OGC Webservices realisiert. Auf der nächsten Ebene ist die *„semantische Interoperabilität"* oder *„Interoperabilität der Konzepte"* verortet (Müller, Portele 2005). Mittels konzeptioneller Datenmodelle sowie geeigneten Konzepten und Mechanismen, um semantische Modelltransformationen

zu definieren und auszuführen, wird diese Ebene erreicht (vgl. Kapitel 2.4). Auf der obersten Ebene kann durch die Definition von konzeptionellen Schemasprachen (CSL) mittels Metamodellen so genannte *„Meta-Interoperabilität"* erreicht werden. Dies aber nur, wenn entsprechende Konzepte auf der zweiten Ebene auch einheitlich angewendet werden. Einerseits kann dies durch die Spezifikation sinnvoller Anwendungsprofile geschehen oder auch durch die direkte, einheitliche Verwendung einer CSL innerhalb einer Informationsgemeinschaft (vgl. Abschnitt 3.2).

2.2. DIE SEMANTISCHE MODELLTRANSFORMATION

Um semantische Interoperabilität realisieren zu können, ist die strukturelle Modifikation von Datenmodellen unabdingbar. Datenmodelle können dann auf konzeptioneller Ebene transformiert werden, wenn es Informationen und Strukturelemente gibt, die vom Quellmodell in das Zielmodell abgebildet werden können. Die semantische Modelltransformation bietet hierfür einen Mechanismus (vgl. Abbildung 2) durch den die Struktur der Quelldaten in jene des Zielmodells umgebaut werden. Eine semantische Transformation ist eine Semantik-erhaltende Abbildung zweier Datenmodelle aufeinander. Dabei sind entsprechend Abbildung 2 vier Hauptschritte zu durchlaufen:

1 konzeptionelle Datenmodellierung im Quell- und Zielsystem + Herleitung Standard-Transferformat automatisch gemäss modellbasiertem Ansatz
2 Definition Modelltransformation auf konzeptioneller Ebene (*„schema mapping"*)
3 Ausführung der semantischen Transformation auf den Transferdaten
4 Transfer der transformierten Daten zum Zielsystem und Nutzung, beispielsweise über OGC Webservices.

Während die syntaktische Interoperabilität vornehmlich für den Informationsaustausch eingesetzt wird, dient die semantische Interoperabilität dem modellbasierten Datenaustausch mit semantischer Modelltransformation (Staub et al. 2008). Neben technischen Herausforderungen bestehen bei Modelltransformationen aber zwei grundsätzliche Probleme: Erstens kann aufgrund fachlicher/semantischer Unterschiede nicht die komplette Modellsemantik eines Quellmodells auf ein Zielmodell abgebildet werden (vgl. Abschnitt 3.3). Zweitens sind Abbildungsregeln für die Objekttransformation in vielen Fällen nicht umkehrbar. So können etwa Aufzählungen aggregiert, jedoch nicht deaggregiert werden (Beispiel: *„Eiche"*, *„Buche"*, *„Ahorn"* →„Baum"/„Baum" → ?).

2.3. GEODATEN-INFRASTRUKTUREN

Zahlreiche aktuelle Entwicklungen im Geobereich zielen auf die Realisierung von GDI ab. Diese können regional, national oder auch international aufgebaut werden. Die Komponenten einer GDI sind nicht nur technischer Natur, vielmehr müssen große administrative, politische und organisatorische Hürden überwunden werden. So setzt sich eine GDI allgemein aus den Komponenten Netzwerk, Regelwerke, Normen/Standards zusammen, die es einem Benutzer erlauben, verfügbare Daten zu finden, beziehen und nutzen zu können (Bernard et al. 2005). Die zentrale Herausforderung bei der Realisierung von GDI besteht darin, die organisatorischen und technischen Voraussetzungen zu schaffen, damit die Nutzung verteilter Systeme ermöglicht werden kann.

Sowohl für Nutzer als auch für Betreiber respektive Datenanbieter entstehen durch die GDI zahlreiche Vorteile, beispielsweise kosteneffizientere Datenproduktion, Vermeidung von Mehrfacherhebungen, Zeit- und Kostenersparnis beim Datenzugriff, effizientere Datennutzung, Angebot veredelter Daten zur Entscheidungsunterstützung, Marktexpandierung (Williamson et al. 2003).

Beispielhaft werden hier die wesentlichen Aspekte bei der Umsetzung der nationalen GDI (NGDI) der Schweiz dargestellt. Die NGDI der Schweiz verfolgt zwei Hauptziele. Erstens eine optimale Unterstützung der E-Government-Strategie des Bundes, zweitens die Erzielung volkswirtschaftlichen Nutzens sowohl innerhalb der Verwaltung, für Privatwirtschaft und Wissenschaft als auch für den Bürger (Buser 2005; Staub 2010). Durch die Schaffung der Geoinformationsgesetzgebung in der Schweiz ist eine wesentliche organisatorische Grundlage für die Realisierung der NGDI gegeben (Schweizerische Eidgenossenschaft 2007).

Auf technischer Ebene werden die Ziele der NGDI durch die Integration von verteilten, heterogenen Geodaten erreicht. Im *„Katalog der Geobasisdaten nach Bundesrecht"* ist definiert, welche Datensätze in die NGDI zu integrieren sind. Für diese Datensätze müssen konzeptionelle Datenmodelle entwickelt werden und je nach Zugangsberechtigungsstufe sind Darstellungsdienste und Downloaddienste anzubieten. Darunter werden weitgehend OGC WMS und OGC WFS verstanden. Geoportale und Geodienste ermöglichen so die breite Nutzung der Geoinformation in der NGDI. Die Bedeutung der Interoperabilität wird dadurch unterstrichen.

2.4. BEDEUTUNG DER INTEROPERABILITÄT FÜR GDI

Um der Komplexität von GDI und den vielfältigen Anforderungen gerecht zu werden, beziehungsweise GDI zielführend aufzubauen und zu betreiben, sind geeignete Standards zu schaffen und einzuhalten. Diese betreffen aber nicht nur die Organisation von GDI, sondern auch der Umgang mit Geodaten und ihre Nutzung. Somit stehen organisatorische und syntaktische Interoperabilität im Zentrum der aktuellen Datennutzung in GDI. Für die Datenintegration sind hingegen Modellgrenzen zu überwinden. Ein Konzept dazu wurde in Abschnitt 2.2 erläutert; ein Lösungsansatz wird in Kapitel 4 skizziert. Für GDI ist Interoperabilität also von zentraler Bedeutung (Bernard et al. 2005; EU 2007). Das Augenmerk liegt bei aktuellen Anstrengungen

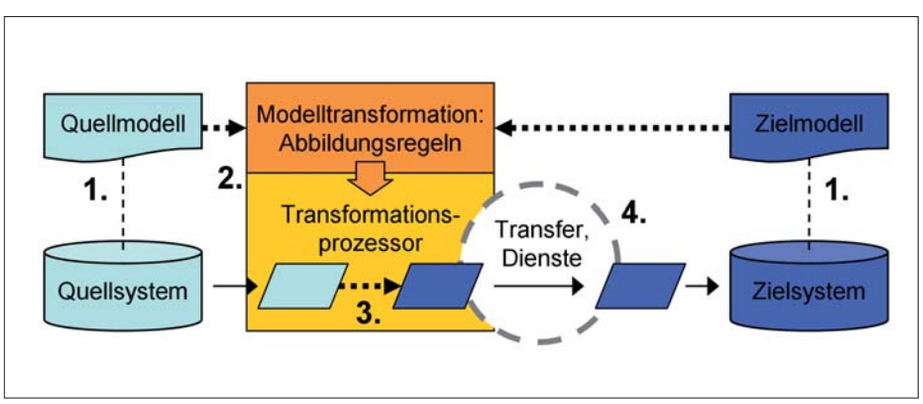

Abb. 2 – die semantische Modelltransformation

vor allem bei der Interoperabilität durch Services. Semantische Interoperabilität mittels Modelltransformationen kann den Nutzen von GDI erhöhen und die Datenintegration vereinfachen.

3. GEODATENMODELLIERUNG

3.1. MODELLIERUNGSPHILOSOPHIE, MODELLBASIERTER ANSATZ

Ausgehend von Kapitel 2, wo an verschiedenen Stellen Datenmodelle erwähnt werden, lohnt es sich, einige grundsätzliche Gedanken zum Vorgehend bei der Modellierung von Geodaten festzuhalten. Die Datenmodellierung wir als wesentlicher Baustein am Gebäude der Interoperabilität (vgl. Abbildung 1) erachtet. Die Datenmodellierung bedingt eine vertiefte Auseinandersetzung mit dem Fachthema, beschreibt die Daten im Sinne einer (nachhaltigen) Dokumentation, verschafft Klarheit und ein gemeinsames Verständnis des Fachthemas, erlaubt die systemunabhängige Mehrfachnutzung von Geodaten und vereinfacht den Datenaustausch.

Bei der Datenmodellierung ist so weit als möglich zwischen Beschreibung der Semantik und der Strukturierung zu trennen. Die Semantik ist durch Fachexperten vollständig und eindeutig zu beschreiben – als Prosatext, durchaus mit bildhaften Darstellungen ergänzbar. Dieser Text ist innerhalb der Fachinformationsgemeinschaft abzustimmen. Der Beizug eines Modellierungsexperten ist wünschbar, um späteren (technischen) Problemen vorzubeugen. An diesem Punkt wird die *„räumliche Modellierung"* abgeschlossen; technische Umsetzungen sind noch nicht relevant. Im zweiten Schritt kann ein Objektartenkatalog folgen, der bereits eine erste Strukturierung beinhaltet. Dieser Katalog steht zwischen Semantik und Strukturierung. Schließlich wird die Semantikbeschreibung durch den Modellierungsexperten mit einer CSL in ein konzeptionelles Datenmodell umgesetzt, formalisiert (z.B. Unified Modeling Language, UML). Analog zum Leitsatz aus Architektur und Design *„Form follows function"* kann man einen Leitsatz für die Datenmodellierung formulieren: *Structure follows semantics – Die Struktur folgt aus der Semantik*.

Die Implementierung des Modells auf ein GIS und die Formatherleitung für den Datenaustausch ist mit entsprechenden Schnittstellen automatisiert vorzunehmen und ist nicht mehr Teil der Datenmodellierung. (Staub 2010).

3.2. PROFILE

Wenn technische Spezifikationen zu abstrakt, zu allgemein oder zu umfangreich für einen Anwendungsbereich sind, bietet sich die Definition von Anwendungsprofilen an, um spezifische Anforderungen zu erfüllen. Ein Profil kann eine Spezifikation einschränken oder aber auch erweitern. In jedem Fall sind die zwingenden Elemente der Spezifikation im Profil zu übernehmen. Ein großer Vorteil der Definition und Verwendung von Profilen liegt einerseits im optimalen Zuschnitt auf die adressierte Informationsgemeinschaft und andererseits in der Gewährleistung der Meta-Interoperabilität durch die einheitliche Anwendung eines Profils. Ein Beispiel für die Notwendigkeit von Profilen ist die Anwendung von UML als Formalismus für die Datenmodellierung. UML definiert u.a. Grund-Datentypen, die sich aber nicht direkt zur Modellierung von Geodaten eignen (OMG 2009). Also müssen in einem UML-Anwendungsprofil geeignete Geo-Datentypen definiert werden die auf diese Weise die Spezifikation erweitern resp. spezialisieren. Zahlreiche UML-Konzepte aber, sind für die Geodatenmodellierung nicht relevant und werden im Anwendungsprofil weggelassen. Eine Möglichkeit zeigt die International Organisation for Standardization (ISO) auf, indem mit den Normen ISO 19103 Geographic Information – Conceptual Schema Language (ISO/TC211 2005) und ISO 19107 Geographic Information – Spatial Schema (ISO/TC211 2003) ein Anwendungsprofil von UML für die Modellierung von Geodaten definiert wird. Allerdings weisen diese Normen stellenweise zu große Freiheitsgrade auf, so dass für konkrete, praktikable Modelle ggf. *„strengere"* Profile benötigt werden. Ein Beispiel dafür ist ISO/TS 19137 Geographic Information – Core profile of the spatial schema (ISO/TC211 2006).

3.3. FACHLICHE SICHT

Wie man in Abschnitt 3.1 sieht, steht und fällt eine sinnvolle Geodatenmodellierung mit der Beschreibung der im Modell abzubildenden Semantik. Unschwer stellt man fest, dass jede Informationsgemeinschaft, die ein bestimmtes Datenmodell zu entwickeln gedenkt, eine ganz spezifische fachliche Sicht auf das Thema hat und somit auch nicht in der Lage ist, die Semantik des zu modellierenden Sachverhalts vollständig objektiv und allgemein zu beschreiben. So wird wahrscheinlich ein Forstingenieur das Thema *„Wald"* fachlich anders betrachten als ein Raumplaner dies tut und folglich auch anders modellieren.

Wenn man davon ausgeht, dass ein bestimmter Realweltausschnitt (*„Wald"*) eine in sich abgeschlossene, *„vollständige"* Semantik innehat und eine CSL oder ein Profil zur Anwendung vorgegeben ist, müsste die fachliche Sicht auf diese Semantik in verschiedenen Fällen identisch sein, wenn nur eine Modellversion entstehen sollte. Offensichtlich ist das unmöglich! Daraus werden drei Konsequenzen ersichtlich:

1. Die fachliche Sicht hat wesentlichen Einfluss auf die Modellgestaltung;
2. Konventionen bezüglich der einheitlichen Anwendung CSL sind unabdingbar;
3. Mechanismen zur semantischen Integration heterogener Datenmodelle müssen verfügbar und praktikabel sein.

4. LÖSUNGSANSATZ

4.1. ANSATZ FÜR DEN WEB-ZUGRIFF AUF VERTEILTE GEODATEN

Etablierte Lösungsansätze im Bereich der Geodaten-Interoperabilität decken den Bereich der syntaktischen Interoperabilität mittels OGC Webservices ab; semantische Interoperabilität ist weitgehend akademisch behandelt und praktikable Lösungen fehlen bislang. Die Nutzung verteilter Geodaten – vor allem auch im Rahmen von GDI – verlangt jedoch die eigentliche Integration von Datensätzen, also den physischen Austausch mit anschließender Integration in ein Empfängersystem, was ohne strukturellen Datenumbau nicht realisierbar ist. Es braucht also eine Lösung, die semantische Modelltransformationen ermöglicht und bestehende Service-Spezifikationen für den Datenaustausch nutzt. Im Rahmen von INSPIRE wurde ein Lösungsansatz für die grenzübergreifende Nutzung räumlicher Daten entwickelt indem erstens mittels eines so genannten *„modellbasierten WFS"* der Modelldaten-Austausch bewerkstelligt wird und zweitens mittels eines konzeptionellen Formalismus zur Datentransformation Interoperabilität auf hohem Niveau realisiert wird (vgl. Abschnitt 4.2). Die Ergebnisse dieser Forschungszusammenarbeit ETH Zürich—TU München wurden in (Donaubauer et al. 2006; Donaubauer et al. 2008) und (Staub 2009) publiziert.

4.2. MODELLBASIERTER WFS MIT SEMANTISCHER TRANSFORMATION

Am Beispiel der Nutzung von verteilten Geodaten im Rahmen von INSPIRE wird die Problematik unzureichender Interoperabilität verdeutlicht (vgl. Abbildung 3): Ein Nutzer möchte für eine grenzübergreifende Anwendung topografische Daten aus der Schweiz (Modell VECTOR25) und aus Deutschland (Modell AAA) gemeinsam nutzen. Dank OGC WFS können entsprechende Datensätze angefragt und als GML-Instanzdokumente bezogen werden. Die Nutzbarkeit bleibt aber stark eingeschränkt, da aufgrund unterschiedlicher Modelle die Daten im Zielmodell nicht integriert werden können. Die Modellinformation der Quellsysteme bleibt vor dem Nutzer verborgen und kann natürlich nicht angepasst werden. Es

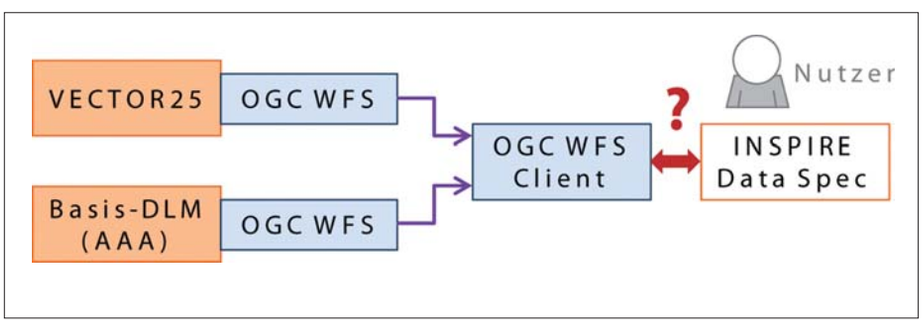

Abb. 3 – Syntaktische Interoperabilität mit OGC WFS

besteht eine semantische Barriere.

Um heterogene Daten sinnvoll nutzen zu können, muss dem Nutzer erstens die Modellinformation der Quellsysteme zugänglich gemacht werden und zweitens müssen die Quell-Datenmodelle in das Zielmodell transformiert werden, damit die Transferdaten in das Zielsystem integriert werden können. Die eigentliche Datennutzung passiert nach wie vor über Standard-OGC WFS.

Die in Abschnitt 4.1 präsentierte Lösung wurde in drei Stufen realisiert. Die erste Stufe umfasst die Erweiterung des OGC WFS zum *„model-driven WFS"* (mdWFS), der es erlaubt, entfernte Datenmodellinformationen über einen Webservice abzufragen (Staub 2007; Staub et al. 2008). Der zweite Schritt ist die Definition der Modelltransformation (Abbildungsregeln) auf konzeptioneller Ebene. Dazu wurde ein Formalismus entwickelt, der auf UML 2-Aktivitäten basiert (Donaubauer et al. 2008; Staub et al. 2008). Der dritte Schritt umfasst eine weitere Erweiterung der Service-Schnittstelle des mdWFS zum *„mdWFS mit semantischer Transformation"* (mdWFS-ST). Dieser hat die Fähigkeit, Abbildungsregeln zu transferieren, semantische Modelltransformationen auf den Quelldaten auszuführen und die transformierten Daten in einer generischen OGC WFS-Instanz dem Benutzer zur Verfügung zu stellen.

In Abbildung 4 ist ein typischer Nutzungsvorgang dargestellt. Über den mdWFS-Client bezieht der Nutzer die Quellmodelle und definiert auf konzeptioneller Ebene die semantische Modelltransformation. Die entsprechenden Abbildungsregeln werden dem Dienst übermittelt, der die Transformation auf den Quelldaten ausführt (1). Nach erfolgter semantischer Transformation werden generisch OGC WFS-Instanzen gemäß Zielmodell konfiguriert (2). Der Nutzer kann die Quelldaten entsprechend seinem Datenmodell nutzen und in sein System integrieren (3).

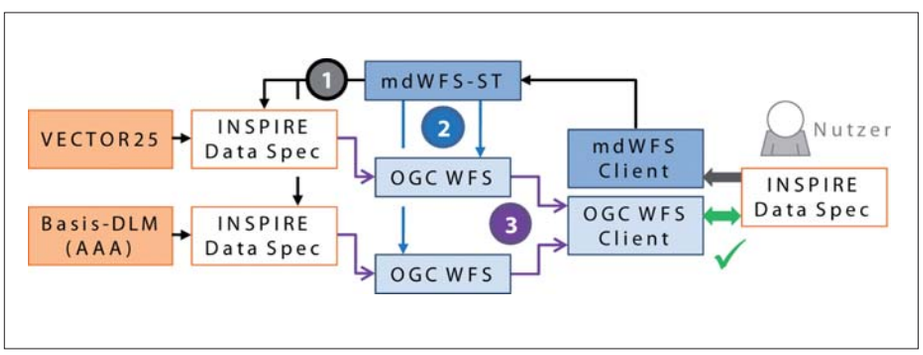

Abb. 4 – Semantische Interoperabilität durch Modelltransformationen mit mdWFS und automatischer OGC WFS-Konfiguration

5. POTENZIAL UND GRENZEN DER SEMANTISCHEN INTEROPERABILITÄT

5.1. ERKENNTNISSE AUS DER FORSCHUNGSARBEIT

Statt mit gegebenen Datenformaten einen (unvollständigen) Strukturumbau auf Formatebene zu programmieren, sollen semantische Modelltransformationen definiert und ausgeführt werden. Dafür lassen sich drei Argumente aufführen:

1. Die konzeptionelle Definition von Modelltransformationen ist in das Konzept des modellbasierten Ansatzes eingebettet. Die semantische Transformation ist grundsätzlich Format-unabhängig.
2. Die Herleitung eines Transferformates aus einem XML-Schema ist nicht ein-eindeutig. Aus einem XML-Anwendungsschema eines Modells können verschiedene Interpretationen vom Zielsystem erzeugt werden (Shi 2004).
3. Im Allgemeinen kann nicht davon ausgegangen werden, dass in einem Transferformat die komplette Modellsemantik repräsentiert werden kann.

Modelltransformationen sind also flexibler, eindeutiger und vollständiger als Strukturumbauten auf Formatebene.

Die Entwicklung des in Kapitel 4 dargestellten Lösungsansatzes zeigt, dass eine sinnvolle konzeptionelle Modellierung unabdingbar ist für praktikable semantische Transformationen. Dazu ist aber ein Formalismus nötig, der kohärent spezifiziert ist. Der entwickelte Ansatz setzt dabei voraus, dass Quell- und Zielmodell mit dem gleichen Formalismus modelliert wurden – womit Meta-Interoperabilität gegeben ist. Für Formalismen wie etwa UML sind spezifische Anwendungsprofile zu definieren, damit geeignete Datentypen zur Verfügung stehen (vgl. Abschnitt 3.2).

Ein Vorteil des präsentierten Ansatzes ist die einfache Nutzung von transformierten Datensätzen über OGC WFS. Die Definition von Modelltransformationen muss wohl

durch Experten vorgenommen werden, die Ausführung kann aber gekapselt werden oder es müssen transformierte Datensätze persistent vorgehalten werden. Somit sind auch *„Transformations-Laien"* in der Lage, das Potenzial der semantischen Interoperabilität zu nutzen.

5.2. EMPFEHLUNGEN FÜR DIE REALISIERUNG SEMANTISCHER INTEROPERABILITÄT

Falls bestimmte Konzepte, Formalismen und Werkzeuge zur Verfügung stehen und auf der organisatorischen Ebene entsprechende Vorkehrungen getroffen werden, ist semantische Interoperabilität realisierbar (vgl. Kapitel 2). Das Potenzial lässt sich am ehesten anhand von Voraussetzungen definieren, die für semantische Interoperabilität erfüllt sein müssen:

- Schaffung und Anwendung von CSL-Spezifikationen die auf anerkannten Normen/Standards basieren, also typischerweise auf UML. Ggf. sind existierende Spezifikationen mittels Anwendungsprofilen zu präzisieren;
- Einheitliche Anwendung der CSL/-Profile zur Gewährleistung der Meta-Interoperabilität;
- Eindeutige Kodierungsregeln für Transferformate, damit jedes Datenmodell in das gleiche Transferformat überführt werden kann;
- Schaffung eines konzeptionellen Formalismus zur Definition von Modelltransformationen bzw. Abbildungsregeln;
- Implementierung eines Transformationsprozessors zur Ausführung von semantischen Transformationen mit geeigneten Transformations-Funktionalitäten;
- Verwendung standardisierter Schnittstellen zum Datenaustausch, beispielsweise OGC WFS und Erweiterung der Schnittstellen um die Fähigkeit, Modellinformationen zu transferieren und semantische Transformationen auszulösen.

5.3. GRENZEN DER SEMANTISCHEN INTEROPERABILITÄT

Bei Modelltransformationen kann jeweils nur der Teil im Zielmodell hergeleitet werden, dessen Semantik in beiden Modellen durch entsprechende Information und/oder Interpretation abbildbar ist. Dieser abbildbare Anteil von Modellinformation wird durch verschiedene Faktoren beeinflusst, wie Sprache, Thema, Präzision/Umfang der Beschreibung des Modellinhalts. Meta-Interoperabilität kann als grundlegende Voraussetzung betrachtet werden. Andernfalls können sehr schnell keine Aussagen mehr gemacht werden, ob eine Modellinformation transformierbar ist, denn es kommen nicht die gleichen Modellierungskonzepte zur Anwendung. Die fachliche Sicht bestimmt auf semantischer Ebene den Umfang transformierbarer Modellinhalte (vgl. Abschnitt 3.3). Es entstehen Datenmodelle, die zum Teil eine gemeinsame Semantik abdecken. So ist schließlich ein Zieldatensatz aus einem Quelldatensatz nur teilweise herleitbar, wie in Abbildung 5 dargestellt.

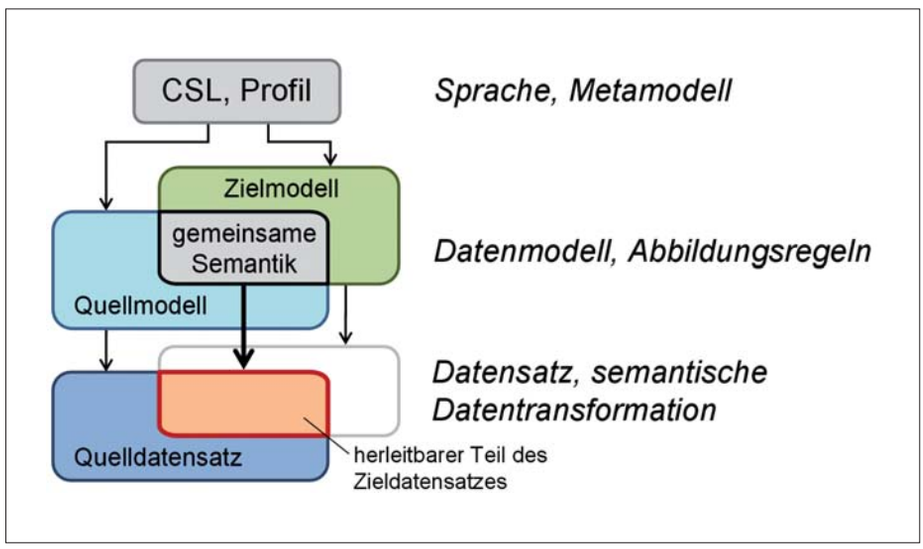

Abb. 4 – Semantische Interoperabilität durch Modelltransformationen mit mdWFS und automatischer OGC WFS-Konfiguration

Aus den vorangehenden Ausführungen lassen sich die Grenzen der semantischen Interoperabilität ableiten. Diese Grenzen werden in konzeptionelle, technische und praktisch relevante Aspekte gegliedert.

1 Konzeptionelle Grenzen: Grundsätzlich schränken unzureichende Meta-Interoperabilität sowie unpräzise Beschreibung oder Definition der Modellsemantik semantische Interoperabilität stark ein. Ein weiterer Aspekt ist die Nicht-Bijektivität von semantischen Transformationen im Allgemeinen (vgl. Abschnitt 2.2). Eine Modelltransformation bedeutet praktisch zwangsläufig ein bestimmtes Maß an Informationsverlust.
2 Technische Grenzen: Dieser Aspekt betrifft in erster Linie die Datenkodierung. Im Allgemeinen kommen für unterschiedliche Formalismen verschiedene Kodierungsregeln zum Einsatz. Sind diese Kodierungsregeln nicht eindeutig und werden sie darüber hinaus nicht einheitlich angewendet, wird insbesondere die (teil-) automatisierte Datentransformation erschwert. Weiter sind für Modelltransformationen praktisch beliebig umfangreiche und komplexe Transformationsfunktionen denkbar. Diese zu implementieren stellt einen immensen Aufwand dar. Die Nutzung existierender Produkte kann dafür eine Lösung darstellen, erfordert aber die Implementierung einer geeigneten Schnittstelle (vgl. Kapitel 6).
3 Praktische Grenzen: Ein einzelnes Quellmodell ist im Allgemeinen kaum in der Lage, ein Zielmodell vollständig abzudecken. Der Beizug weiterer, fachlich verwandter Datenmodelle kann den herleitbaren Anteil der Zielmodell-Informationen erhöhen. Ganz wesentlich ist die fachliche Barriere zwischen unterschiedlichen Anwendungs-Welten. Die fachliche Sicht auf einen Realweltausschnitt ist individuell und lässt sich kaum standardisieren. Fachvokabulare und die umfangreiche, exakte Beschreibung der Modellsemantik helfen, Modelle besser zu verstehen und Modelltransformationen gezielter zu definieren.

Weitere Ausführungen dazu finden sich in (Staub 2009).

6. FAZIT, AUSBLICK

Anforderungen aus GDI auf nationaler und internationaler Ebene verlangen nach Datenintegration und geeigneten Mitteln zur interoperablen Nutzung heterogener Daten. Die Kombination von semantischen Modelltransformationen und Standard-OGC WFS-Schnittstellen realisiert ein hohes Maß an semantischer Interoperabilität. Dafür müssen allerdings bestimmte organisatorische, konzeptionelle, technische und praktische Voraussetzungen erfüllt werden.

Das Bewusstsein für die Bedeutung der Datenmodellierung – mit einheitlichen CSL/Profilen – und der Datenintegration im Rahmen von GDI muss weiter gestärkt werden. Semantische Interoperabilität lässt sich auf technischer Ebene weitgehend realisieren, der Umgang damit hat ein großes Potenzial. Die Grenzen können durch Maßnahmen auf konzeptioneller, technischer und praktischer Ebene nur zum Teil überwunden werden.

Zusammen mit einem Industriepartner wird der beschriebene Lösungsansatz weiterentwickelt. Dem entsprechend werden das Potenzial und der Nutzen anerkannt (Safe Software 2008). Für den Einsatz in der Praxis ist es unabdingbar, den konzeptionellen Ansatz der semantischen Transformation mit etablierten Werkzeugen, die umfangreiche Transformationsfunktionalität anbieten, zu kombinieren. ◀

Literatur

Bartelme, N. (2005): Geoinformatik-Modelle, Strukturen, Funktionen. Springer, S. 363.

Bernard, L.; Fitzke, J.; Wagner, R.M. (Hrsg.) (2005): Geodateninfrastruktur-Grundlagen und Anwendungen. Wichmann.

Buser, R. (2005): Nationale Geodaten-Infrastruktur (NGDI): Organisatorische Aspekte der Interoperabilität beim Bund. In: Carosio, A. (Hrsg.): Interoperabilität für die breite Nutzung von Geoinformation. IGP-Bericht Nr. 298d, ETH Zürich, Institut für Geodäsie und Photogrammetrie.

Eidgenössische Vermessungsdirektion (1999): INTERLIS, Ein Datenaustausch-Mechanismus für Land-Informationssysteme.

Europäisches Parlament und Europäischer Rat (2007): Richtlinie 2007/2/EG des Europäischen Parlaments und des Rates vom 14. März 2007 zur Schaffung einer Geodateninfrastruktur in der Europäischen Gemeinschaft (INSPIRE). EU-Richtlinie, Amtsblatt der Europäischen Union, Nr. L 108/1.

Donaubauer, A.; Fichtinger, A.; Schilcher, M.; Straub, F.; Carosio, A.; Gnägi, H. R.; Morf, A.; Staub, P. (2006): Grenzübergreifende Web-GIS-Lösungen. GIS Zeitschrift für Geoinformatik, 9/2006, S. 29-34.

Donaubauer, A.; Staub, P.; Straub, F.; Fichtinger, A. (2008): Web-basierte Modelltransformation - eine Lösung für INSPIRE?. GIS Zeitschrift für Geoinformatik, 2/2008, S. 26-33.

ISO/TC211 (2003): ISO 19107 Geographic Information - Spatial Schema. International Standard 19107.

ISO/TC211 (2005): ISO 19103 Geographic Information - Conceptual Schema Language. International Standard 19103.

ISO/TC211 (2006): ISO 19137 Geographic Information - Core profile of the spatial schema. International Standard 19137.

Müller, M.; Portele, C. (2005): GDI-Architekturmodelle. In: Bernard et al. (Hrsg.): Geodateninfrastruktur - Grundlagen und Anwendungen. Wichmann, S. 83-92.

OMG (2009): Unified Modeling Language Version 2.2 - Infrastructure/Superstructure. Object Management Group OMG Specification formal/2009-02-04/ formal/2009-02-02.

OGC (2008): OGC Reference Model Version 2.0. Open Geospatial Consortium OGC 08-062r4.

OGC (2010): OGC Web Services Common Specification Version 2.0.0. Open Geospatial Consortium OGC 06-121r9.

Safe Software (2008): FME Addresses INSPIRE's Critical Challenges: Semantic Translations an Data Sharing. Safe Software Press Room. (online: http://www.safe.com/aboutus/news/2008/115/, Zugriff 04/2010).

Schweizerische Eidgenossenschaft (2007): Bundesgesetz über die Geoinformation (Geoinformationsgesetz, GeoIG). Gesetz der Schweizerischen Eidgenossenschaft, SR 510.62.

Shi, W. (2004): Zum modellbasierten Austausch von Geodaten auf Basis XML. Dissertation, Universität der Bundeswehr München.

Staub, P. (2007): A Model-Driven Web Feature Service für Enhanced Semantic Interoperability. OSGeo Journal, 3/2007, S. 38-43.

Staub, P.; Gnägi, H. R.; Morf, A. (2008): Semantic Interoperability through the Definition of Conceptual Model Transformations. Transactions in GIS, 12(2), S. 193-207.

Staub, P. (2009): Über das Potenzial und die Grenzen der semantischen Interoperabilität von Geodaten. Dissertation, ETH Zürich Nr. 18201.

Staub, P. (2010): Datenmodelle zur Datenintegration für die NGDI der Schweiz. In: Tagungsband des 15. Münchner Fortbildungsseminars Geoinformationssysteme. Runder Tisch GIS e.V.

Williamson, I.; Abbas, R.; Feeney M.-E. (Hrsg.) (2003): Developing Spatial Data Infrastructures: From Concept to Reality. Taylor & Francis.