

Publikationen der Deutschen
Gesellschaft für Photogrammetrie,
Fernerkundung und Geoinformation e.V.



Band 21

2012



Vorträge

**32. Wissenschaftlich-Technische
Jahrestagung der DGPF**

14. – 17. März 2012
in Potsdam

Erblicke – Perspektiven für die Geowissenschaften

ISSN 0942-2870

Publikationen der
Deutschen Gesellschaft für Photogrammetrie, Fernerkundung
und Geoinformation (DGPF) e.V.
Band 21 487 S., Potsdam 2012
Hrsg.: Eckhardt Seyfert

© Deutsche Gesellschaft für Photogrammetrie, Fernerkundung
und Geoinformation (DGPF) e.V.
Oldenburg 2008

Zu beziehen durch:

Geschäftsstelle der DGPF
c/o EFTAS GmbH
Oststraße 2-18
D-48145 Münster
Tel.: (0251) 133 070, Fax: (0251) 133 0733, E-Mail: klaus.komp@eftas.com

Bearbeitung:

Landesvermessung und Geobasisinformation Brandenburg
Heinrich-Mann-Allee 103, 14473 Potsdam
Tel.: (0331) 8844 -229, Fax: (0331) 8844-126

Semantische und geometrische Aspekte der Interoperabilität

INGRID CHRIST¹ & ROLF LESSING²

Zusammenfassung: Die Aktualisierung bestehender Landbedeckungs- und Landnutzungsdaten basiert im Projekt DeCOVER auf einem eigenen Objektartenkatalog, der den Anforderungen der Fernerkundung entspricht. Die beiden Datenmodelle sowie die darauf basierenden Datensätze müssen interoperabel sein, um eine Datenvergleichbarkeit zu erhalten und eine Datentransformation ausführen zu können. Zur Erreichung dieses Ziels kommen verschiedene Interoperabilitäts-Methoden zum Einsatz, die alle ‚semantisches Wissen‘ nutzen. Landbedeckungsdaten bieten semantisches Wissen über Objektarten-Beschreibungen, geometrische Bedingungen, topologische Beziehungen und zeitliche Veränderungsmöglichkeiten. Mit Hilfe dieser Eigenschaften und der darauf abgebildeten Regeln wird eine (semi-)automatische Überführung von der Ausgangs-Datenbasis in eine Datenbasis eines Zielkatalogs ermöglicht. Neben der Unterstützung der mittels Fernerkundungsmethoden detektierten Aktualisierungen kann dieses Wissen auch zur Qualitätsprüfung oder Qualitätsbewertung eingesetzt werden.

1 Einleitung

Mit dem Verbundvorhaben DeCOVER werden Methoden zur Aktualisierung und Erweiterung bestehender Landbedeckungsinformationen entwickelt. Die Change Detection-Methode zur Aktualisierung der Daten findet auf Satellitensensoren wie RapidEye und TerraSAR-X statt. Als Grundlage dient ein eigens in DeCOVER spezifizierter Objektartenkatalog mit 35 Objektklassen, die ihren Schwerpunkt in der Landbedeckung haben und nachgeordnet auch Landnutzungsinformationen liefern. Somit beziehen sich die gesamten Klassifikationsmethoden und Algorithmen auf ein DeCOVER Datenmodell. Auf der anderen Seite befindet sich der Nutzer, der an den Veränderungen der bestehenden Landbedeckungsdaten interessiert ist. Der Nutzer möchte sich nicht mit weiteren Katalogspezifikationen und spezifischen Klassen beschäftigen, sondern er möchte ganz konkret wissen, welche Objekte seiner Daten sich aktuell geändert haben und welche aktuelle Landbedeckung diese Objekte aufweisen.

Abb. 1 zeigt die Schnittstellen der Interoperabilitäts-Methoden zu dem DeCOVER-Klassifikationsprozess auf. In der Mitte befinden sich die Fernerkundungsmethoden, die eine Ausweisung von Veränderungsobjekten sowie die Neuklassifikation zum Ziel haben. Dieser Klassifikationsprozess findet auf Basis der DeCOVER Nomenklatur statt. Davor und danach erfolgt die Transformation zwischen Nutzerdaten und DeCOVER-Daten auf Basis von Interoperabilitätsmethoden. Im dargestellten Beispiel wurde ATKIS BasisDLM verwendet.

¹ Ingrid Christ, DELPHI IMM GmbH, Friedrich-Ebert-Straße 8, 14467 Potsdam; E-Mail: ingrid.christ@delphi-imm.de

² Rolf Lessing, DELPHI IMM GmbH, Friedrich-Ebert-Straße 8, 14467 Potsdam; E-Mail: rolf.lessing@delphi-imm.de

Die Schaffung der Interoperabilität von Geodaten verschiedener Herkunft ist essentiell, um eine durchgängige DeCOVER Prozesskette zu erlangen und damit den Anforderungen der Aktualisierung von Landbedeckungsinformationen gerecht zu werden. Dies bedeutet, dass die unterschiedlichen Landbedeckungsinformationen interpretierbar und vergleichbar sein müssen. Ontologiebasierte Methoden aus dem Bereich der Wissensrepräsentanz kommen für die Beschreibung der Objektartenkataloge zum Tragen. Daneben spielen aber auch geometrische sowie räumliche Eigenschaften und Beziehungen eine Rolle, da es sich nicht um ‚einfache‘ Daten, sondern um Geodaten handelt. Geometrische Angaben von Katalogspezifikationen, wie z.B. die Angabe der Mindestkartierfläche (MKF), müssen beachtet werden. Gleichzeitig bieten sich auch weitere Möglichkeiten durch die Nutzung dieser geodaten-spezifischen Eigenschaften, um Rückschlüsse auf die semantischen Beziehungen abzuleiten oder logische und topologische Inkonsistenzen aufzuzeigen. Im vorliegenden Beitrag wird erläutert, welche Interoperabilitäts-Methodenkette angewandt wird, um eine Überführung der Geodaten herzustellen und damit eine Aussage zu Veränderungen bestehender Landnutzungsdaten zu ermöglichen.

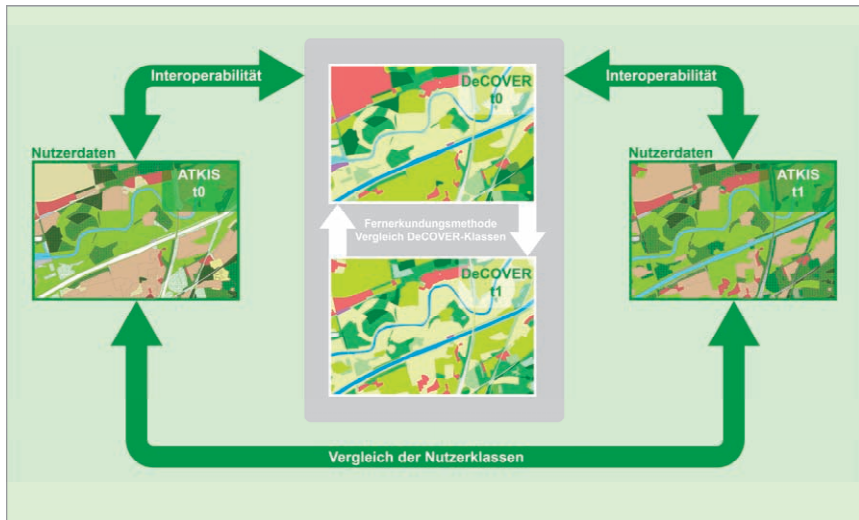


Abb. 1: Interoperabilität im Rahmen des DeCOVER Prozesses

2 Interoperabilitätsmethoden

Der semantische Vergleich von Landbedeckungsdaten unterschiedlicher Datenquellen und Objektartenkatalogen birgt bereits auf der Schemaebene eine Vielzahl an Konfliktmöglichkeiten [FRAUNHOFER ISST]. Die flache Datenstruktur von CORINE LandCover ist beispielsweise anders modelliert als das Datenmodell von DeCOVER, bei der die Angaben zur Landbedeckung und Landnutzung getrennt vorgehalten werden. Die Konfliktmöglichkeiten treten jedoch insbesondere auf der Datenebene auf. Für einfache Modelltransformationen oder Klassen-Mappings gibt es bereits Produkte auf dem Markt. Werkzeuge für semantische Datenharmonisierungen sind jedoch nicht bekannt.

Für die Datentransformation in DeCOVER hat sich herausgestellt, dass zwischen Ansätzen auf der Klassenbasis und Ansätzen auf der Instanzebene unterschieden werden kann. Daraus hat sich eine prozessorientierte Methodik in drei Schritten entwickelt (siehe Abb. 2).

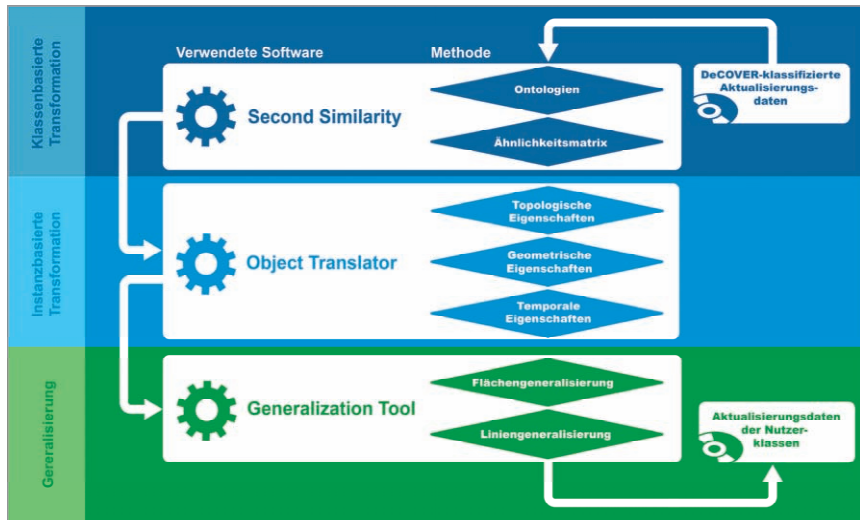


Abb. 2: Methodenkette zur Schaffung einer semantischen und geometrischen Interoperabilität

Klassenbasierte Transformation

Im ersten Schritt findet die semantische Transformation auf Katalogbasis statt. Für den Vergleich von Landbedeckungsklassen werden Eigenschaften extrahiert und in einem Wissensmodell in Beziehung gesetzt. Das semantische Wissen aus der Katalogbeschreibung wird in OWL (WebOntologyLanguage) formalisiert. Dieses abgestimmte Basisvokabular dient zur Beschreibung der Klassen und ermöglicht damit eine einheitliche Dateninterpretation. Mit Hilfe

eines Reasoners findet ein Matching von identischen Klassenpaaren oder Subklassen statt. Bei den vorliegenden Landbedeckungsklassen konnte jedoch festgestellt werden, dass die meisten Klassen keinen eindeutigen Partner besitzen, da es oft nur eine Schnittmenge an übereinstimmenden Eigenschaften gibt. Die gängigen semantischen Klassen-Relationen von ‚1:1‘, ‚n:1‘ und ‚1:n‘ kommen damit nicht zum Tragen. Die Entwicklung eines Algorithmus zur Berechnung von Ähnlichkeiten trug durch den Vergleich der Ähnlichkeitswerte erheblich zum Lösen dieses Problems bei. Der Ähnlichkeitswert von Grünland zu Ackerflächen ist beispielsweise höher als von Grünland zu bebauten Flächen. Baut man auf diese semantischen Werte Mapping-Regeln auf, dann lassen sich die Transformationen definieren.

Die Standard-Mapping-Regel sieht beispielsweise im Klartext so aus:

„Bei der Überführung einer Klasse eines Katalogs in einen anderen Katalog wird die Klasse mit der höchsten Ähnlichkeit zugeordnet, wobei nur Ähnlichkeiten ab einem Wert von 60% Berücksichtigung finden. Gibt es kein Klassen-Paar, das diesen Schwellenwert erreicht, dann kann für diese Klasse keine Transformation hergestellt werden.“

Schwellenwerte, Ausnahmen und auch Prioritäten können bei den Regeln berücksichtigt werden. Prioritäten sind ein wichtiges Kennzeichen für die Charakteristik eines Objektartenkataloges. So ist ATKIS eher im Bereich der Landnutzung einzuordnen, während dessen CORINE LandCover Katalog schon alleine durch die fernerkundliche Datengrundlage ein Landbedeckungskatalog darstellt.

2.1 Instanzbasierte Transformation

Bei der klassenbasierten Transformation in Schritt 1 ist von einem allgemeineren zu einem detaillierteren Katalog, wie dies bei DeCOVER zu ATKIS der Fall ist, die Anzahl der möglichen Zielpartner (1:n-Beziehungen) noch sehr hoch. Diese Relationen sind auf Katalogebene durchaus berechtigt. Die DeCOVER Klasse BVS-Straßen- und Eisenbahnen beinhaltet beispielsweise die ATKIS-Klassen Straßenverkehr, Bahnverkehr und Plätze und damit kann eine BVs-Neufläche auch potentiell alle drei ATKIS-Klassen betreffen. Die Herausforderung, hierfür weitere Spezifizierungen zu erhalten wird in dem zweiten Schritt, der instanzbasierten Transformation, angenommen.

Die Methode der instanzbasierten Transformation nutzt das semantische Wissen jeder individuellen Objektinstanz. Geometrische Eigenschaften der Objekte werden mit den Wertebereichen aller Objekte gleicher Ausgangsklasse verglichen, um zu prüfen, ob potenzielle Zielklassen ausgeschlossen werden können. Der Wertebereich der Flächengröße, Kompaktheit sowie das Längen- und Breitenverhältnis, um nur ein paar geometrische Eigenschaften zu nennen, werden automatisch für alle Klassen aus den Objekten ausgelesen und zur Klassentrennung verwendet. Das gleiche gilt für die topologischen Eigenschaften, die ebenfalls die 1:n-Beziehungen weiter spezifizieren und damit die Überführungsklassen qualitativ verbessern können. So handelt es sich bei einer schmalen Gehölzfläche zwischen Straßenflächen mit hoher Wahrscheinlichkeit um ‚Straßenbegleitgrün‘ und nicht nur um ‚Gehölz‘ im allgemeineren Sinne. Analog dazu sieht es auch bei schmalen Gehölzflächen entlang von Gewässern aus. Hierbei handelt es sich sehr wahrscheinlich um Ufergehölz.

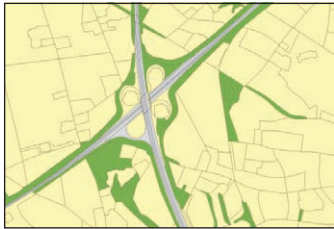


Abb. 3: DeCOVER Strauchvegetation



Abb. 4: hellgrün: ATKIS Gehölzflächen; dunkelgrün: ATKIS Verkehrsbegleitgrün



Abb. 5: DeCOVER Gewässerläufe



Abb. 6: hellblau: ATKIS Strom, Fluss, Bach; dunkelblau: ATKIS Kanäle

2.2 Geometrische Interoperabilität (Geometrische Anpassung – objektbasierte Manipulation)

Nach der Durchführung der ersten beiden Schritte sind die Daten nun semantisch vergleichbar, d.h. dass die in DeCOVER aktualisierten Daten nun nach der Transformation der semantischen Zielspezifikation entsprechen, wie beispielsweise der Spezifikation von CORINE Land Cover. Die geometrische Spezifikation wurde jedoch noch nicht ausreichend berücksichtigt. Bei der Aktualisierung in DeCOVER werden Objekte ab 0,5 ha berücksichtigt, während der europaweite Datensatz von CORINE Land Cover erst ab 25 ha Objekte ausweist. Beim Betrachten der aktualisierten CORINE Daten im Arbeitsmaßstab von 1:100.000 fällt der ‚Flickenteppich‘ recht deutlich auf, der beim letzten Schritt, der Generalisierung, modifiziert wird.

Generalisierung ist ein weites Feld und betrifft die erwähnten Flächenaggregationen. Weiterhin muss bei der Generalisierung auch eine maßstabsabhängige Linienanpassung berücksichtigt werden, um tatsächlich von geometrisch interoperablen Geoinformationen zu sprechen. Für das Projekt DeCOVER wurde die Flächenaggregation unter Angabe der Pflichtparameter der MKF sowie der Bodenauflösung der verwendeten Fernerkundungsdaten in Form eines Entscheidungsbaumes automatisiert in die Prozesskette eingebunden. Nacheinander werden Objekte anhand der Flächengröße und Nachbarpolygone bestätigt oder auf Basis der Ähnlichkeiten zugunsten der Nachbarobjekte aufgelöst, so dass auch hier wieder das semantische Wissen der Objektarten zum Tragen kommt.



Abb. 7: Aktualisierte CLC-Daten



Abb. 8: Flächenaggregierte CLC-Daten

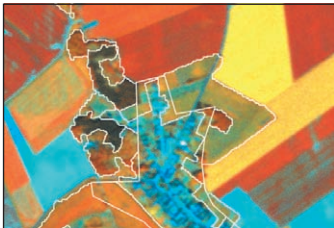


Abb. 9: Aktualisierte CLC-Daten

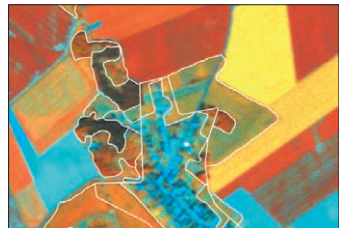


Abb. 10: Liniengeneralisierte, aktualisierte CLC-Daten
Smooth-PEAK Algorithmus mit 50m-Toleranz, Maßstab 1:5.000

Semantik und Geometrie bedingen sich häufig. Besonders gut ist dies bei den Mischklassen zu erkennen. So besteht ein Mischwald in CLC aus Laub- und Nadelgehölzen mit jeweils mindestens 25% Flächenanteil. Wird ein Objekt geometrisch aufgelöst, dann kann es bei der Gesamtbetrachtung zu einer anderen Flächenzusammensetzung kommen, die semantisch nicht mehr dem Mischwald entspricht.

3 Semantik zur Qualitätsverbesserung

In dem vorigen Kapitel wurde aufgezeigt, wie Semantik für die Verwendung der gesamten Überführungskette verwendet werden kann und letztlich zu einem interoperablen Datensatz führen kann. Der Einsatz des semantischen Wissens und den semantischen Regeln kann darüber hinaus noch wesentlich vielseitiger eingesetzt werden. Klassen-Inkonsistenzen auf Basis von logischen, topologischen, temporalen und geometrischen Eigenschaften können aufgezeigt werden und damit zur Qualitätsverbesserung des Datensatzes beitragen.

Dem Faktum ‚Raumbezug‘ ist es zu verdanken, dass diese Informationen konkret eingesetzt werden können, um kombinierte und teilweise auch komplizierte Abfragen zu machen. Es gibt zahlreiche Beispiele, von denen drei vorgestellt werden.

- Ein einfacher Fall stellt die Veränderung einer bebauten Fläche dar. Hierbei kann es möglicherweise zu einer veränderten bebauten Nutzung, wie lockere Wohnbebauung zur

Industriefläche oder ehemalige Industrieflächen zu Sport- und Freizeitflächen kommen. Als sehr unwahrscheinlich ist die Veränderung zu Agrarland einzustufen.

- Bei neu entstandenen Wasserflächen handelt es sich meist um Überflutungsflächen aufgrund von Hochwasser oder Flutungsflächen von ehemaligen Tagebauseen. Somit kann die Aussage getroffen werden, dass „neue Wasserflächen in der Nachbarschaft von Wasser oder Feuchtfächen oder auf ehemaligen Abbaufächen“ vorliegen können. Bei allen anderen zeitlichen und topologischen Kriterien ist eine Entstehung als sehr unwahrscheinlich anzusehen.
- Aus ökologischer Betrachtung ist die Veränderung von langjährig genutztem, artenarmen Dauergrünland hin zu naturnahen, artenreicheren Wiesen innerhalb von wenigen Jahren kaum möglich. Selbst das Einbringen von samenhaltigem Material wird als wenig aussichtsreich eingestuft. Hierbei gilt: Je länger die Zeitreihen vorliegen, desto besser können Rückschlüsse auf die heutige Landbedeckungssituation vorgenommen werden.

Die Nutzung des semantischen Wissens kann zur Qualitätsprüfung oder gezielten Qualitätsverbesserung eingesetzt werden. Es sind jedoch auch Grenzen gesetzt, die berücksichtigt werden sollten. Bei fehlerhaften Ausgangsdaten bleibt die reale, unwahrscheinliche Veränderungsrichtung zwar bestehen; die aktuell klassifizierte Klasse könnte jedoch richtig sein und die Fehlerquelle in diesem Fall in der früheren Landnutzungskartierung liegen.

4 Zusammenfassung und Ausblick

Semantik ist mehr als die Klassenhierarchie von Geodaten. Mit der der Wissensbeschreibung von Objektarten, der Beachtung der (maßstabsbezogenen) geometrischen Bedingungen und vor allen Dingen der Nutzung der Nachbarschaftsbeziehungen bis hin zum Einsatz der Veränderungsbeziehungen können sehr viele Bewertungen und Überführungsaussagen getroffen werden. In Projekt DeCOVER werden all diese Eigenschaften genutzt, Regeln aufgestellt und damit eine Prozesskette zusammengestellt, die (semi-) automatisch einen Ausgangsdatensatz in einen Zieldatensatz eines anderen Kataloges überführt. Die Schaffung von interoperablen Daten wird damit erreicht. Synergieeffekte können genutzt und Kosten für die mehrfache Datenaufnahme eingespart werden, indem ein aktualisierter Nutzerdatensatz rein technisch in weitere Nutzerkataloge überführt werden kann.

Der Prozess funktioniert, wenn gleich es auch Grenzen der Machbarkeit gibt. Dies betrifft hauptsächlich die Berechnung von ontologiebasierten Ähnlichkeiten, da die Zuordnung von semantisch korrespondierenden Objektbeschreibungen in verschiedenen Datensätzen das Expertenwissen und die Anwendersicht, die je nach Person und Einsatzgebiet anders eingeschätzt wird, bedarf. Somit stellt die automatisiert-erzeugte Mapping-Tabelle die Grundlage für weitere Abstimmungen dar, dessen Erfolg dennoch nicht zu verschmälern ist.

Der Einsatz des semantischen Wissens (als Oberbegriff aller inhaltlichen, geometrischen, topologischen und zeitlichen Eigenschaften und Prozessen) bietet zahlreiche Möglichkeiten, die bisher auch im Projekt DeCOVER nur ‚angeritzt‘ wurden. Ein zukünftiges Ziel sollte die Veränderung des bisher klassenbezogenen Regelwerkes unter Zur Hilfenahme des

wissensbasierten Ansatzes sein. Dann bezögen sich die Veränderungsrichtungen nicht mehr auf Objektarten, sondern auf Eigenschaften, wie z.B. der Versiegelung. Die Veränderung einer Fläche mit einer höheren in eine geringere Versiegelungsdichte wird dann als unwahrscheinlich angesehen, unabhängig, ob es sich dabei um Wohnbebauung oder Freizeitanlagen handelt.

5 Acknowledgements

Die Arbeit ist Teil des Projektes DeCOVER und wird vom Bundesministerium für Bildung und Forschung (BMBF) unter der FKZ 50EE0909 gefördert.

6 Literaturverzeichnis

- ADV 2008. ATKIS-Objektartenkatalog. <http://www.atkis.de> (2008/04/30).
- BOBZIEN, M., 2005: Methodische Aspekte der Generalisierung von Geodaten, BONN.
- FRAUNHOFER ISSI, 2009: Semantische Interoperabilität. Band 2: Semantische Konflikte. White Paper. Berlin
- KIELER, B., 2007: A Geometry-Driven Approach for the Semantic Integration of Geodata Sets. IN: PROC. OF 23RD. International Cartographic Conference, 4.-10. August 2007, Moscow, Russia. (http://www.ikg.uni-hannover.de/fileadmin/ikg/staff/publications/Konferenzbeitraege_abstract_review/Kieler_icc2007.pdf)
- LOIDOLD, M., 2005: Entwicklung einer automatisierten Konflikterkennung für partizipative GIS auf Basis dynamischer Ontologien, Salzburg. ([HTTP://WWW.IKG.UNI-HANNOVER.DE/FILEADMIN/IKG/STAFF/PUBLICATIONS/KONFERENZBEITRAEGE_ABSTRACT_REVIEW/KIELER_ICC2007.PDF](http://www.ikg.uni-hannover.de/fileadmin/ikg/staff/publications/Konferenzbeitraege_abstract_review/Kieler_icc2007.pdf))
- LUTZ, M., CHRIST, I., SCHUBERT, C., KLIEN, E. & HÜBNER, S., 2008: Overcoming Semantic Heterogeneity in Spatial Data Infrastructures, *in* Computers and Geoscience. Special Issue on Geoscience Knowledge Representation in Cyberinfrastructure, in press. Available online 25 January 2008, <http://www.elsevier.com/locate/cageo>.
- SCHUBERT, C., WOLPERT, I. & I. CHRIST, 2008: Comparison of Different Land Use Object Classes by Means of Semantic Measurements. Brady. S.R., Sinha, A.K. and Gundersen, L.C., editors, 2008, Geoinformatics 2008 – Data to Knowledge, Proceedings: U.S. Geological Survey Scientific Investigations Report 2008-5172, p. 66, ISBN 978-1-4-4113-2279-0.
- STAUB, P., 2009: Semantische Interoperabilität – Voraussetzung für Geodaten -Infrastrukturen. Géomatique Suisse, **6** (2009), S. 304-307.