
Metadaten: ISO – konformes Profil als Schritt in die Praxis in Österreich

Manfred MITTLBOECK*, Paul SCHREILECHNER**

* iSPACE, Research Studios Austria

** BIOGIS Consulting GmbH

Zusammenfassung

Metadaten sind als strukturierte Dokumentation von Datensätzen zu verstehen, die es Erstellern, Verwaltern und Nutzern ermöglicht, den Inhalt und damit den anwendungsbezogenen Nutzwert der beschriebenen Datensätze zu verstehen. Die Suche in Katalogen nach diesen Datensätzen wird wesentlich erleichtert, wenn ein einheitliches Datenmodell für Metadaten eingesetzt wird.

Für Metadaten wird daher dringend angestrebt, ein einheitliches Profil auf Basis der (für breite Anwendung zu umfassenden) ISO Norm 19115 für Österreich umzusetzen. Diese Standardisierung-innerhalb-des-Standards soll den Datenzugang und übergreifende Kompatibilität wie auch den transparenten Zugriff über Applikationsbereiche hinweg ermöglichen. Operatives Ziel ist somit die Entwicklung eines ISO-konformen Metadatenprofils für Österreich. In weiterer Folge steht die Entwicklung einer verteilten Katalog-Infrastruktur für die integrative Suche über Metadatenbestände hinweg im Mittelpunkt, dies unter Berücksichtigung der Metadaten-Implementierungsspezifikation ISO 19139.

1 Einleitung

Auch qualitativ hochwertige Geodaten nutzen im Kontext von GI-Infrastrukturen wenig, wenn sie nicht gut dokumentiert und nach variablen Suchkriterien auffindbar sind. Unter Metadaten - "Daten über Daten" - versteht man strukturierte Deskriptoren (z.B. Information über Herkunft, Inhalt, Struktur, Gültigkeit, Genauigkeit, Zugriffsmöglichkeit etc.), mit deren Hilfe eine geographische Ressource wie Geodaten, Online-Dienste oder Karten beschrieben wird. Ziel ist zunächst das gezielte Finden einer solchen Ressource, in weiterer Folge die Klärung der Nutzbarkeit für einen bestimmten Anwendungskontext.

In welcher Form sollen nun Geo-Metadaten definiert, gespeichert bzw. transportiert werden, und wie sollen bestehende Metadaten und Datenkataloge in die Geoinformations-Infrastruktur in Österreich eingebunden werden? Mit diesen Fragen beschäftigt sich eine Arbeitsgemeinschaft der Fa. BIOGIS, des Z_GIS an der Universität Salzburg und das Forschungsstudio iSPACE mit Unterstützung durch AGEO als nationalem Dachverband.

2 ISO-konformes Metadaten-Profil

Für Metadaten liegen definierte Standards vor, die allerdings aufgrund der großen Anzahl möglicher Deskriptoren eine praktikable Eingrenzung auf sogenannte „Profile“ für definierte „information communities“ nahe legen. Ziel der vorliegenden Arbeit ist es, ein einheitliches Profil auf Basis der ISO-Norm 19115 für Österreich um- und durchzusetzen. Diese Standardisierung-innerhalb-des-Standards soll die Durchsetzung umfassender Geodaten-Dokumentation in der Anwendungspraxis ermöglichen und in der Folge als Baustein einer nationalen GI-Infrastruktur den transparenten Zugriff über Kataloge auf Geodaten quer über Applikationsbereiche hinweg verbessern. In weiterer Folge wird die Entwicklung einer interoperablen Katalog-Infrastruktur für die integrierte Sicht verteilter Metadatenbestände unter Berücksichtigung der Metadaten Implementierungsspezifikation ISO 19139 angestrebt.

Mit den Normen ISO 19115:2003 und der in Vorbereitung befindlichen Spezifikation ISO 19115-2 für Bilddaten wurden (auf der Basis von im OpenGIS Konsortium erarbeiteten Grundlagen) vom ISO TC211 Modelle zur Abbildung von Metadaten für geographische Datenbestände (Raster und Vektor) entwickelt. Ziel der Normen ist die Definition einer Struktur für die Beschreibung von digitalen geographischen Daten. Darin definiert wird eine gemeinsame Metadaten-Terminologie, ein konzeptuelles Modell mit konkreten Metadatenelementen, sowie die Möglichkeit von Erweiterungen.

Dieser Normierungsansatz geht davon aus möglichst alle denkbaren Aspekte komplett abzubilden, was zu einem sehr umfangreichen Rahmenstandard von über 300 Merkmalen geführt hat, wobei die meisten Deskriptoren als optional definiert sind.

Sogenannte „Profile“ erlauben es nun, Teilmengen aus dem umfassenden Modell für eine spezifische Anwendungs-Domäne bzw. „community“ zu erstellen, sowie das Modell mit zusätzlichen Metadatenelementen zu erweitern bzw. zu ändern. Die Norm schreibt genau vor, wie derartige Änderungen umzusetzen sind.

Um eine minimale obligatorische Beschreibung von Geodaten garantieren zu können, wurde das sogenannte ISO Kernmodell festgelegt. Um diesen verpflichtenden Kern an Deskriptoren herum gilt es nun ca. 50 – 80 Deskriptoren zu identifizieren, die als Basis eines nationalen Profils vereinbart werden können.

Um von dieser Sicht ausgehend eine möglichst einheitliche Implementations - Richtlinie in Form eines ISO-Metadaten-Profiles für Österreich vorzuschlagen, wurde in einer Expertenrunde Übereinstimmung angestrebt. Im Mittelpunkt steht der erforderliche Schritt zur 'nächsten Generation' der Länder-Datenkataloge (*Biogis, 2002*) zunächst die Definition eines "community profile" für diesen Nutzerkreis, das jedoch in weiterer Folge generell für Basis- wie auch unterschiedliche Anwendungs-spezifische Geodaten im öffentlichen wie auch privaten Sektor anwendbar ist. Das Ergebnis wird mit anderen Profil-Entwicklungs-Initiativen im internationalen Bereich verglichen und in der Folge frei verfügbar über einschlägige Homepages angeboten werden..

3 Umsetzungsstrategie für ISO-konforme Metadaten

Ausgehend von den Inhalten des bereits jetzt zur Anwendung kommenden „Geo-Katalog“ der Fa. BIOGIS (*BIOGIS 2002*) wurde in Abstimmung mit Domänen-Experten versucht jene Metadatenelemente in der ISO 19115:2003 Norm zu identifizieren, die eine informative Identifikation eines Datensatzes unter Berücksichtigung möglichst effizienter Metadatenerfassung und Nachführung ermöglichen.

3.1 Geokatalog - Metadatenverwaltung für GIS

Geodaten sind die Grundlage jedes geographischen Informationssystems, die Mengen gespeicherter Geodaten steigen stetig an. Man steht daher vor zwei Herausforderungen: einerseits Daten zugänglich zu machen, und andererseits die Werterhaltung von Daten sicherzustellen. Ohne Information über Ausgangsdaten und modifizierende Transformationen können Datensätze trotz sorgfältiger Erhebung schnell zu „Datenfriedhöfen“ werden. Die Firma BIOGIS setzte mit dem Produkt „Geokatalog“ eine erste umfassende Metadatenbank mit folgenden Zielsetzungen um:

- Orientierung an internationalen Standards für Metadaten: Identifikation und thematische Klassifizierung von Datensätzen; Dokumentation des räumlichen Bezugssystems; Datendefinition durch Angaben über Datenstruktur und -format bzw. Attributinformationen; Qualitätsmerkmale des Datensatzes; Angaben über Verfügbarkeit und Bezugsquellen; administrative Metadaten.
- Minimierung der Eingabefelder, um den personellen Aufwand der Metadatenerfassung möglichst gering zu halten.
- Ermöglichung des Zugriffs auf Metadaten sowohl Behörden-intern wie auch für externe Partner über Intranet und Internet.
- Wo möglich und sinnvoll automatische Ableitung von Metadaten aus vorliegenden Datenbeständen durch Koppelung von Datenbank und GIS.

Bei der Orientierung an internationalen Standards war von Beginn an erkennbar, dass ein Kompromiss zwischen „vollständiger“ Metadatenerfassung und einer möglichst praktikablen Metadatenerfassung nötig ist, also eine Teilmenge als „Profil“ für eine bestimmte Anwendergruppe als beste Lösung anzustreben ist.

3.2 Internationale Normen ISO 19115:2003 und 19139

ISO TC 211 entwickelt mit der Standards-Serie 19100 geographische Informationsschemata, die deren Integration in die allgemeine IT-Infrastruktur sicherstellen sollen. Die generell zunehmende Bedeutung geographischer Information und die verstärkte Nutzung neuer elektronischer Verteilungs- und Zugangsmöglichkeiten erforderte die Entwicklung eines Standards zur Beschreibung von Geodaten.

Die elektronisch gestützte Suche in diesen Metadaten wird wesentlich erleichtert, wenn für diese ein einheitlich anerkanntes Datenmodell benützt wird. Über verschiedene Entwicklungsphasen hinweg kam es schließlich 2003 zu Verabschiedung der Norm ISO

19115:2003, die ein allgemein gültiges Schema zur Beschreibung geographischer Datensätze und Dienste vorgibt.

Durch die abstrakte Form der Spezifikation von ISO 19115:2003 gibt es unterschiedliche Alternativen zur praktischen Umsetzung, je nach Interpretation einzelner Komponenten. Mit der Entwicklung der „(Draft) Application Specification ISO 19139“ wird nun der Versuch gestartet, eine eindeutige UML Implementation und ein XML Schema zur Umsetzung des Standards ISO 19115 festzulegen. Eine führende Rolle in der Entwicklung einer solchen allgemein gültigen Umsetzung hat die Digital Geographic Information Working Group (DGIWG)¹.

3.2.1 Profile und Erweiterungen

Die Norm ISO 19115:2003 wurde für einen möglichst universellen Einsatz entwickelt. Sie besteht aus über 300 Einzelementen, wobei die meisten optional verwendet werden können. Mit den Werkzeugen der Profile-Definition und von Erweiterungen bietet die Norm die Möglichkeit den Gesamtumfang der Norm domänenspezifisch anzupassen (zu erweitern oder eben auch einzuschränken).

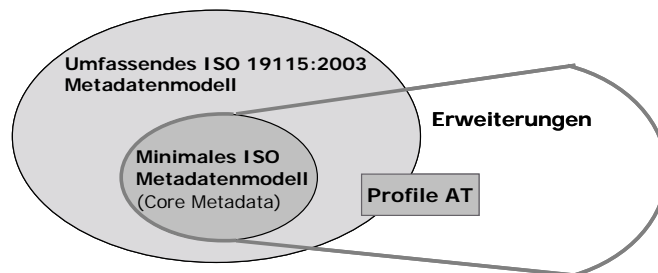


Abb. 1: ISO 19115:2003 Metadatenmodell

Profile:

Die Mehrzahl der definierten Metadatenelemente können optional verwendet werden. Organisationen, Interessensgruppen, Nationalstaaten können sogenannte „community profiles“ entwickeln die es ihnen ermöglichen, eine Teilmenge von Meta-Deskriptoren auszuwählen, die den Anforderungen am besten entspricht. In jedem Profil muss zumindest das minimale Metadatenmodell („Core Metadata“) abgebildet werden. Außerdem kann das Profil nach Bedarf erweitert werden.

Erweiterungen:

Reicht das generische Metadatenmodell nicht aus, gibt der ISO Standard 19115:2003 den Anwendern ein Regelwerk zur Erweiterung der vorgegebenen Strukturen mit. Mögliche Erweiterungen bestehen z.B. in der Berücksichtigung von rechtlichen Rahmenbedingungen bzw. in der Einbeziehung der Preisinformation im Bereich der Datenabgabe.

¹ DGIWG: metadata.dgiwg.org

3.3 Umsetzung des Metadaten-Profiles

In Anlehnung an die im ISO Standard 19115:2003 definierten Metadaten-Klassen, wurden von den Autoren folgende prioritären Kategorien identifiziert:

- Metadatensatz (MD_Metadata)
- Identifikation (MD_Identification)
- Inhalt (MD_ContentInformation)
- Datenqualität (DQ_DataQuality)
- Räumliche Ausprägung (MD_SpatialRepresentation)
- Räumliches Bezugssystem (MD_ReferenceSystem)
- Verteilung ((MD_Distribution)
- Einschränkungen (MD_Constraints)

Im Folgenden wurde die Zuordnung der bestehenden Inhalte des BIOGIS - Geokatalog mit den Anwendern dieses Kataloges in einer Expertenrunde diskutiert und die Überführung der Inhalte in die entsprechenden Kategorien der Norm ISO 19115:2003 umgesetzt.

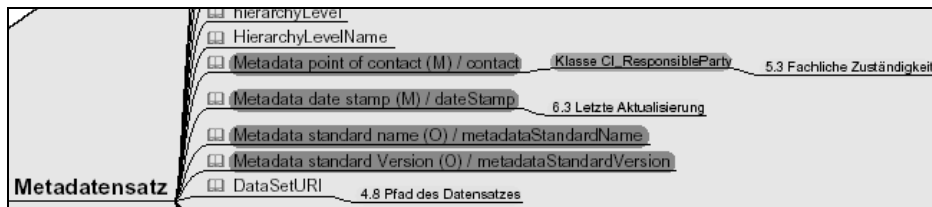


Abb. 2: Überführung der aktuellen Deskriptoren des bestehenden „Geokatalog“ in entsprechende ISO 19115:2003 Kategorien.

Als nächster Schritt erfolgte die Ergänzung des Profils mit jenen Metadatenelementen, die im minimalen Metadatenmodell definiert sind. Als Beispiel hierfür sind z.B. die in der ISO-Norm festgelegten Themengebiete (topic categories) für die einzelnen Datensätze zu nennen.

Zur Prüfung der Praxistauglichkeit wurden verschiedene Benutzergruppen in einer Testphase eingeladen, das aus diesen Anforderungen heraus entstandene Standard-konforme Profil an konkreten Geodatensätzen auszutesten. Vor allem im Bereich der Definition von Metadatenelementen für Bilddaten (z.B. Bodenauflösung des Bildsensors) konnte mit der bestehenden Norm ISO 19115:2003 nicht das Auslangen gefunden werden. Es wurde daher vereinbart, den Draft Standard ISO 19115-2 (Erweiterung für Image and Gridded Data) in die Profildefinition mit einzubeziehen.

3.4 ISO 19139 XML Implementierung

Um ambivalente Interpretationen dieses an die Norm ISO 19115:2003 angepassten Profils zu vermeiden und die zukünftige Interoperabilität zu gewährleisten befinden sich eine eindeutige Spezifikation in UML und ein an die Norm ISO 19139 (Implementation Specification) – angepasstes Schema in Ausarbeitung. Das RSA iSPACE hat das entwickelte Geo - Metadatenprofil für Österreich an dieses Schema angepasst.

ISO 19115 Metadatenprofil für Österreich / XML-Implementation 19139 v0.8

profile.AT

- MD profile.AT
 - Metadatenansatz
 - Allgemein
 - Hilfe
 - Identifikation
 - Inhalt
 - Räumliche Information
 - Räumliche Ausprägung
 - Datenqualität
 - Einschränkungen
 - Nachführung
 - Verteilung

Metadata entity Set Information (MD_Metadata)

Allgemeine Information:
Metadatenpaket mit wichtigen Schlüsselinformationen mit Beziehung zu den brigen Paketen.

	ISO 19115 value
*Eindeutige Identifikation: <input type="text" value="F565BB68-705F-493B-BAB0-61DD976E3262"/>	1, mdFileID
*Metadatensprache: <input type="text" value="deutsch"/>	2, mdLang
*CharacterSet: <input type="text" value="utf8"/>	3, mdChHr
*Abstammungsinformation: <input type="text" value="F565BB68-705F-493B-BAB0-61DD976E3262"/>	5, mdParentID
*HierarchieLevel: <input type="text" value="dataset"/>	6, mdHrLv
*HierarchieLevelName: <input type="text" value="Orthophotokatalog 2003 Tile-Beschreibung"/>	7, mdHrLvName
*Fachliche Zuständigkeit:	8, mdContact
Name: <input type="text" value="Huber, Andreas, Dr."/>	8.1,375, rpIndName
Organisation: <input type="text" value="Universität Salzburg"/>	8.2,376, rpOrgName
Position: <input type="text" value="Datenbearbeiter"/>	8.3,377, rpPosName
Kontakt:	8.4, 378, rpCntInfo (CI_Contact, 387, B 3.2.3)

Abb. 3: Web - Interface zur Eingabe eines Geo-Metadatenansatzes mit dem für Österreich angepassten Metadatenprofil mit ISO 19139 konformer XML - Implementation

4 Literatur

- ISO (Hrsg.), 2003, Geographic information – Metadata – Implementation Specification – Working Draft 3, [=International Standard, Reference Number ISO 19139.3] – Genf
- ISO (Hrsg.), 2003, Geographic information – Metadata, [=International Standard, Reference Number ISO 19115:2003] – Genf
- ISO (Hrsg.), 2004, Geographic information – Metadata – Part 2: Extensions for imagery and gridded data –Working Draft 2, [=International Standard, Reference Number ISO 19115-2.2] – Genf
- KOGIS (Hrsg.), 2001, Entwicklung eines ISO/DIS kompatiblen Metadatenmodells für die Schweiz. Version 1.2. – Wabern.
- LOCHTER F., et al., 2003, Metadaten für Geodaten im Rahmen der Initiative zum Aufbau der Geodateninfrastruktur in Brandenburg – GIB. – Potsdam.
- REMKE A., BERNARD, L., 2001, GDI NRW Geodateninfrastruktur Nordrhein-Westfalen. Testpad 1 Spezifikation 1.0. – Münster.
- BIOGIS, 2002, Bedienungsanleitung Biogis Geokatalog Metadateneditor für ESRI ArcCatalog – Version 1.1. - Salzburg