



Modul 6: Voraussetzungen einer GDI
Vertiefende Dokumente | Stand: 24.01.2012

Metadaten

Warum muss man Metadaten erfassen und veröffentlichen und welche Vorschriften gibt es hierzu? Was ist notwendig, damit die Metadaten als eines der wichtigsten Themen einer GDI funktionieren? Dieses Dokument beantwortet diese Fragen.

Was sind Metadaten und was bieten sie?

In einem Versandkatalog informiert man sich über ein bestimmtes Produkt mit Hilfe beschreibender Informationen. Auf diese Weise können Produkte anhand ihrer technischen Merkmale miteinander verglichen werden. Bei den hierbei zur Entscheidungsfindung beitragenden beschreibenden Daten handelt es sich um Metadaten. Metadaten können aber nicht nur Produkte, sondern auch Daten, welche letztlich auch Produkte sind, beschreiben. In der Geoinformation beschreiben Metadaten Geodaten oder Geodatendienste, nicht zu verwechseln mit Sachdaten, die den thematischen Inhalt eines raumbezogenen Objekts wiedergeben. Bei Naturschutzgebieten wären dies zum Beispiel Name, Datum des Inkrafttretens, Schutzzweck und ähnliche Angaben. Weitere Metadaten über Geodaten sind zum Beispiel Koordinatenreferenzsysteme, Datum der inhaltlichen Aktualität, Maßstabsgrundlagen oder Ansprechpartner.

Das Beispiel der Metadaten in einem Produktkatalog zeigt, dass es sinnvoll ist, Metadaten zu kategorisieren. Um Produkte miteinander vergleichen zu können, sollte eine möglichst einheitliche Form vorgegeben sein. Ein Prosatext erschwert das Herausarbeiten der wesentlichen Unterschiede. Erst durch eine Vereinheitlichung der Beschreibung bringen sie dem Anwender, der wissen will, ob die Daten anderer Anbieter für den geplanten Einsatzzweck geeignet sind, einen Mehrwert.

Bevor der Anwender über die Eignung entscheiden kann, muss er die Metainformationen zunächst finden. In Metadateninformationssystemen (MIS) hat er die Möglichkeit, gezielt nach Daten und Diensten mit bestimmten Inhalten und Eigenschaften zu suchen. Im Falle von WMS-Diensten findet der Anwender in den Metadaten die URL des Dienstes und kann die räumlichen Daten somit direkt im eigenen Geoinformationssystem verwenden oder in einer entsprechenden Viewing-Komponente visualisieren.

Metadaten dienen der längerfristigen Wertsicherung von Geodaten. Innerhalb einer Institution sind Metadaten unentbehrlich bei der Dokumentation zu archivierender Datenbestände und zur Nachvollziehbarkeit von Analyseergebnissen.

Der Nutzen von Metadaten liegt zusammengefasst in folgenden Punkten:

- Durch die Bereitstellung von Metadaten sind Suchfunktionen realisierbar, bei denen auf der Grundlage von Anwenderkriterien die entsprechenden Datenbestände identifiziert werden.
- Metadaten helfen dem Nutzer, über die Eignung der recherchierten Daten und Dienste zu entscheiden.
- Metadaten sind Voraussetzung für einen weitreichenden Datenaustausch.
- Innerhalb einer Institution dienen Metadaten zur internen Organisation und Verwaltung von Datenbeständen.

Metadaten in einer Geodateninfrastruktur

In einer Geodateninfrastruktur (GDI) sind Metadaten der Dreh- und Angelpunkt zur gegenseitigen Datenbereitstellung. Der Datenanbieter beschreibt seine Geodaten und Geodatendienste durch Metadaten und veröffentlicht die Metadaten in einem Portal. Das Portal stellt dem Anwender eine Oberfläche zur Suche in den Metadaten zur Verfügung und/oder es stellt Schnittstellen (OGC Catalogue Service for the Web) bereit, die es wiederum anderen Portalen ermöglichen, in dem Portal die Metadaten zu durchsuchen. Auf diese Weise können Metadaten dezentral vorgehalten und über standardisierte Suchanfragen von verschiedenen Metadateninformationssystemen aufgefunden werden.

Sollen die Metadaten in einer dezentralen Organisation über Dienste bereit gestellt und über eine Recherche-Anwendung, einem Metadateninformationssystem (beispielsweise dem Geodatenkatalog Hessen), zur Nutzung angeboten werden, so muss eine „Broker-Software“ implementiert werden, mit der die verschiedenen Dienste registriert werden. Bei einer Anfrage führt der Broker dann eine verteilte Suche über die angebotenen Dienste durch. Der Broker kann seinerseits als Suchdienst fungieren. Auf diese Weise können verschiedene Broker miteinander kommunizieren, so dass eine kaskadierende Suche eingerichtet werden kann („Harvesting“ von Metadateneinträgen). So kann beispielsweise ein europäisches MIS die Kataloge der Mitgliedsstaaten abfragen. Diese wiederum fragen die Kataloge der nächsten Ebene – in Deutschland die Bundesländer – ab und diese wiederum die regionale Ebene.

Erfassung von Metadaten

Um in Metadaten suchen zu können, bedarf es eines einheitlichen Grundmusters, in dem die Metadaten vorliegen. Die Erfasser der Metadaten müssen sich auf eine einheitliche Struktur und Semantik verständigen. Zu diesem Zweck wurden in den verschiedenen Ländern Metadaten-Standards entwickelt. Zu nennen sind hier zum Beispiel der internationale Standard ISO 19115 sowie die Metadaten-Verordnung (Durchführungsbestimmung) der europäischen INSPIRE-Richtlinie (zum Teil basierend auf ISO 19115).

In Deutschland haben die ISO-Standards 19115 für Metadaten über Geodaten und 19119 für Metadaten über Geodatendienste in den letzten Jahren viel Beachtung gefunden. Nachteil des ISO-Standards 19115 ist die große Zahl seiner Elemente und, damit verbunden, dessen Unübersichtlichkeit. Es besteht die Notwendigkeit, sich auf eine Teilmenge der Elemente zu beschränken, also ein Profil zu entwickeln. Für die Entwicklung eines Profils sehen sich in Deutschland die Bundesländer zuständig. Die Länder Baden-Württemberg und Berlin/Brandenburg beispielsweise machen die Einhaltung dieses Profils für ihre Landesbehörden zur Pflicht. Als die Metadaten-Verordnung der INSPIRE-Richtlinie verabschiedet wurde, gab es diese Profile bereits. Sie wurden 2010 an die INSPIRE-Anforderungen angepasst.

Leider gibt es nicht den einen Standard – die eine Sprache – über den sich alle verständigen könnten. So werden scheinbar gleichbedeutende Elemente von den Standards unterschiedlich interpretiert. Aber nicht nur zwischen verschiedenen Metadatenstandards, sondern auch innerhalb eines Standards gibt es Interpretationsspielräume. In vielen Fällen wird die Bedeutung eines Elementes nicht exakt definiert. Dies ist allerdings auch verständlich, angesichts der Heterogenität der Daten und Anwendungen, die beschrieben werden müssen. Nachteil ist: Es führt zwangsläufig zu unterschiedlichen Auslegungen bei den Anwendern. Ein Leser versteht möglicherweise unter einem Revisionsdatum etwas anderes als das, was der Erfasser aussagen wollte. Es ist daher sinnvoll, einen Erfassungsleitfaden anzubieten, der die einzelnen Elemente näher erläutert und auch dem nichtspezialisierten Anwender erleichtert, Metadaten aussagekräftig und korrekt zu erfassen. Auf Bundesebene gibt es dazu seit 2008 ein Dokument der GDI-DE, das eine deutsche Übersetzung der ISO 19115 sowie kurze Erläuterungen der Elemente aus Sicht verschiedener Institutionen bietet.

Zusätzlich zu den allgemeinen Standards gibt es fachbezogene oder auf einen bestimmten Aspekt der Geodaten spezialisierte Metadatenprofile. So hat die Arbeitsgemeinschaft der Vermessungsverwaltungen der Länder der Bundesrepublik Deutschland (AdV) einen AAA-Metadatenkatalog veröffentlicht. Die Staatlichen Geologischen Dienste (SGD) haben ein Profil für die komplexen geologischen Daten entwickelt und der Deutsche Dachverband für Geoinformation e.V. (DDGI) veröffentlichte ein Modell speziell für die Beschreibung der Qualität von Geodaten.

Organisatorische Fragen

Das Ziel, Metadaten bereit zu stellen, scheitert oft an organisatorischen Umsetzungshürden. Die Erfassung der Metadaten muss innerhalb einer Organisation abgestimmt werden. Sind die Erfasser auf die Bedeutung der Elemente geeicht? Wer pflegt welche Metadaten, beziehungsweise welchen Teil der Metadaten? Wer ist verantwortlich für die Pflege und Bereitstellung? Werden die Metadaten separat von den Geodaten in einer Metadatenbank verwaltet oder sind sie direkt an die Geodaten gekoppelt? Wer sorgt im ersten Fall auf welche Weise für die Konsistenz zwischen Geodaten und Metadaten? Werden die Daten zentral oder dezentral erfasst? Wer übernimmt die Qualitätssicherung? Es muss geklärt werden, wer die technische Infrastruktur für ein Metainformationssystem trägt. Im Rahmen einer GDI ist damit die finanzielle Lastenverteilung verbunden. Hinzu kommt, dass die Vorgaben (Standards, Modelle, Profile) komplex und technisch orientiert sind. Dies wirft die Frage nach dem Einarbeitungsaufwand und der Qualifikation auf, beziehungsweise eröffnet Interpretationsspielräume, die zu Missverständnissen führen.

Ein weiterer organisatorischer Aspekt ist das Problem, dass während der Bearbeitung der Geodaten oft nicht an deren Dokumentation gedacht wird. Erst am Ende der Projektbearbeitung werden die Metadaten – wenn überhaupt – erfasst. Zu diesem Zeitpunkt sind aber Motivation und Gedächtnis nicht mehr optimal ausgeprägt. Eine Institution sollte sich also überlegen, wie sie eine zeitnahe Dokumentation erreichen kann.

Im Rahmen eines speziellen GDI-Themas, zum Beispiel dem Bebauungsplaninformationssystem der GDI-Süd Hessen, ist eine Einigung über die Struktur der Inhalte und die Inhalte selber erforderlich: Beispielsweise welche Schlagwörter werden benutzt und wie werden sie geschrieben? Zuvor sollte geklärt sein, für welchen Zweck und welche Zielgruppe die Metadaten bereit gestellt werden sollen. Gibt es über INSPIRE-konforme Metadaten hinaus interne Metadaten, die nicht für die Öffentlichkeit gedacht sind, oder spezifische Metadaten für Experten? Organisatorisch ist auch zu klären, ob die Geodaten in aggregierter Form beschrieben werden sollen. Bei einer Serie topographischer Karten erscheint es sinnvoller, die Serie und nicht jedes einzelne Kartenblatt zu beschreiben. Werden Metadaten zu einem Dienst, also beispielsweise einem WMS, veröffentlicht, ist zu hinterfragen: Inwiefern spiegelt die Beschreibung des WMS die Geodaten wider? Sollen zum Beispiel mit den WMS-Metadaten auch Nutzungsbedingungen über die Geodaten mitgeteilt werden?

INSPIRE

Zur Umsetzung der INSPIRE-Richtlinie ist Ende 2008 die Metadaten-Durchführungsbestimmung (VERORDNUNG (EG) Nr. 1205/2008 DER KOMMISSION vom 3. Dezember 2008 zur Durchführung der Richtlinie 2007/2/EG des Europäischen Parlaments und des Rates hinsichtlich Metadaten) in Kraft getreten. Ergänzt wird die Verordnung durch eine technische Anleitung (Technical Guidelines). Welche Geodathemen beschrieben werden müssen, regelt die INSPIRE-Richtlinie in ihren Anhängen I bis III.

Im Verordnungstext werden die Erfordernisse für die Erstellung und Pflege von Metadaten über Geodatenätze, Geodatenatzreihen und Geodatendienste festgelegt. Die Durchführungsbestimmungen (Implementing Rules) sind rechtlich verbindlich und geben vor, was getan werden muss, d.h. welche Metadaten erfasst werden müssen. Die meisten Metadatenelemente sind verpflichtend, einige optional und viele bedingt verpflichtend. Zum

Beispiel muss mindestens ein Datumselement ausgefüllt werden oder es muss die Sprache angegeben werden, wenn ein Geodatensatz Texte enthält.

Die Technical Guidelines sind rechtlich nicht verbindlich. Dadurch können sie in kürzeren Abständen auf einen technisch neueren Stand gebracht werden. Sie beschreiben, wie die Metadaten erfasst werden sollen. In den Technical Guidelines findet der Metadatenerfasser die für ihn wichtigsten Informationen. Sie legen vor allem dar, wie die in der Durchführungsbestimmung definierten Metadatenelemente auf Basis der ISO-Standards umgesetzt werden. Für jedes Metadatenelement wird anhand eines Beispiels ein ausführliches Mapping auf die Elemente des ISO 19115 bzw. ISO 19119 Encodings durchgeführt. Es werden Instruktionen zur Implementierung gegeben und ein Beispiel für eine konkrete XML-Umsetzung geliefert. Die angeführten XML-Pfade beruhen wiederum auf dem Standard ISO 19139 für Geodatensätze und -reihen bzw. auf dem OGC CSW 2.0.2 Application Profile für Geodatendienste.

Der Überblick über die Metadatenelemente wird dadurch erschwert, dass in den jeweiligen Datenspezifikationen der INSPIRE-Themen zusätzliche Metadatenelemente gefordert werden. Je nach Thema muss man also verschiedene Metadaten zusätzlich liefern. Für die Themen Verkehrsnetze und Hydrographie des Annex I ist dies beispielsweise die Topologische Konsistenz.

Zur besseren Übersicht wurden die INSPIRE-Metadatenelemente in zehn Gruppen eingeteilt. Diese sind Identification, Classification of spatial data and services, Keyword, Geographic Location, Temporal Reference, Quality and Validity, Conformity, Constraints related to access and use, Responsible Organisation und Metadata on Metadata.

Unter Identification (Identifizierung) versammeln sich zentrale Elemente wie Titel, Kurzfassung und ID. In Classification (Klassifizierung) wählt man einen oder mehrere Begriffe aus der ISO-Themenkategorie aus und in Keyword (Schlüsselwort) mindestens ein Schlüsselwort aus dem europäischen Umwelt-Thesaurus (GEMET), das sich auf das in den Anhängen I, II und III der Richtlinie 2007/2/EG aufgeführte einschlägige Geodathema bezieht. In Geographic Location (Geografischer Standort) folgt die Angabe der Koordinaten des umgebenden Rechtecks und in Temporal Reference (Zeitbezug) können verschiedene Datumsangaben gemacht werden. Eine sehr wichtige Gruppe ist Quality and Validity (Qualität und Gültigkeit), wo Aussagen zur Herkunft und Entstehung der Daten sowie geometrische Genauigkeitsangaben zu finden sind. Unter Conformity (Übereinstimmung) wird in Zukunft zu erkennen sein, ob ein Datenbestand konform ist zu einem ihn betreffenden INSPIRE-Datenmodell. Im Abschnitt Constraints lassen sich Nutzungsbedingungen und Zugangseinschränkungen konkret formulieren. Im Kapitel Responsible Organisation (Zuständige Stelle) sind die Kontaktinformationen des Datenbereitstellers verzeichnet und schließlich gibt es auch noch (Meta-)Informationen über die Metadaten, welche im Abschnitt Metadata on Metadata erfasst werden können.

Suche in Metadaten

Wie schon erwähnt, müssen Metadaten mit Hilfe von Katalogdiensten gesucht werden können. Auch hierfür sind wieder Spezifikationen und Standards erforderlich. Dementsprechend definiert INSPIRE einen Suchdienst (Discovery Service). Der spezifizierte Dienst unterstützt die Veröffentlichung von und die Suche nach Metadaten sowie die Verlinkung von Suchdiensten untereinander (Kaskadierung). Die Einhaltung des OGC CSW-Standards (Catalogue Service for the Web) wird dabei vorgeschrieben. Der Suchdienst stellt die Verbindung zwischen Recherche-Anwendung und einer Metadatenbank her. Die Kommunikation erfolgt durch standardisierte CSW-Operationen.

Damit der Dienst weiß, nach welchen Elementen er suchen kann, bedarf es wiederum einer Liste von Metadatenelementen, einem Profil. Für Geodatensätze und Geodatensatzreihen liegt dem Profil das ISO-19115-Modell und für Dienste das ISO-19119-Modell in Verbindung mit dem OGC CSW 2.0.2 Application Profile zugrunde.

Der Endanwender, der sich in einem Katalog auf die Suche begibt, bleibt vor dieser Komplexität bewahrt. In der Regel bietet ihm ein Metadateninformationssystem eine „Einfache Suche“ an, bei der er ein Schlagwort eingeben kann, sowie eine „Erweiterte Suche“, bei der er die Recherche thematisch und räumlich einschränken kann. Der Anwender hat außerdem die Option, eine verteilte Suche zu starten. Dadurch werden andere Suchdienste eingebunden, d.h. es werden weitere Metadatenbanken abgefragt. Auf diese Weise kann jede Organisation ihren eigenen Metadatenserver betreiben und bei einem übergeordneten Metadatenkatalog registrieren, wodurch eine wesentlich größere Anwenderschaft erreicht werden kann.