

Vom Umweltdatenkatalog zum *InGrid*[®] *Data Catalog* - Aktuelle Anpassungen des UDK-Datenmodells und der Software *InGrid*[®] im Hinblick auf INSPIRE

Martin Klenke, Fred Kruse, Stefanie Urich, Christiane Giffei, Sybille Peters,
Koordinierungsstelle PortalU im Niedersächsischen Ministerium für Umwelt und
Klimaschutz, martin.klenke@portal.de

Zusammenfassung

Im Zuge der Entwicklung und Bereitstellung einer web-basierten Metadaterfassungs- und Pflegekomponente im Rahmen des Projekts *InGrid*[®] 2.0 [Klenke et al., 2007] als Ersatz für den in der deutschen Umweltverwaltung verbreiteten *Windows-UDK* (Umweltdatenkatalog) werden verschiedene Anpassungen am ISO 19115- und ISO 19119-konformen UDK-Datenmodell erforderlich, teils um die Abfrageperformanz externer Schnittstelle zu erhöhen, teils um neuen Funktionalitäten und inhaltlichen Anforderungen Rechnung zu tragen. Von besonderer Relevanz sind inhaltliche Erweiterungen, die derzeit auf Grundlage der *INSPIRE*¹ *Draft Implementing Rules for Metadata (Version 4)* durchgeführt werden um sicher zu stellen, dass über PortalU^{®2}/*InGrid*[®] die Inhalte der angeschlossenen Umweltdatenkataloge INSPIRE-konform abgegeben werden können. Im ersten Teil unseres Artikels beschreiben wir die diesbezüglich durchgeführten Anpassungen am UDK-Datenmodell.

Mit der zunehmenden Konkretisierung der verschiedenen INSPIRE-Durchführungsbestimmungen stellt sich im nationalen Umfeld die Frage, mit welchem Architekturkonzept GDI-DE³ (Geodateninfrastruktur Deutschland) den Anforderungen der INSPIRE-Richtlinie begegnet. Verschiedene Ansätze werden aktuell im Rahmen des GDI-DE-Modellprojekts *Geodatenkatalog-DE* diskutiert. Im Modellprojekt soll u.a.

¹ <http://www.ec-gis.org/inspire/>

² <http://www.portal.de/>

³ <http://www.gdi-de.org/>

prototypisch ein zentraler, deutscher INSPIRE-Metadatenknoten entwickelt werden. Im zweiten Teil unseres Artikels geben wir einen kurzen Überblick über den aktuellen Stand der Diskussion im Modellprojekt sowie über geplante Erweiterungen der PortalU[®]-Software *InGrid*[®], die im Rahmen von INSPIRE relevant erscheinen.

Abstract

Currently, a web-based application for the collection and maintenance of environmental metadata is under development within the scope of the project *InGrid*[®] 2.0. This new application will supersede the *Windows-UDK* desktop application, actually being widely distributed within the German environmental authorities. Along with the front-end development several data model adjustments are necessary in order to improve the performance of external interfaces as well as to calibrate to new demands with respect to content. Regarding the latter, of most importance are requirements emerging within the scope of the INSPIRE¹ directive. For the metadata community, of particular significance are the INSPIRE Implementing Rules for Metadata which are supposed to be published officially in the second half of 2008. In the first part of our paper we describe modifications of the ISO 19115 / ISO 19119 conformal UDK metadata model, initiated by the *INSPIRE Draft Implementing Rules for Metadata (Version 4)*.

While the INSPIRE implementing rules are getting more concrete, discussions arise within the German environmental and geospatial data and metadata communities on how to establish an efficient national metadata infrastructure to serve INSPIRE. At present, different approaches are discussed in the context of the GDI-DE² (Spatial Data Infrastructure Germany) pilot project *Geodatenkatalog-DE*. One aim of the pilot project is to establish the prototype of a central German metadata-node for INSPIRE. In the second part of our article we give a brief overview on the state of discussion in the pilot project as well as on some desired *InGrid*[®]-functionalities, probably being relevant for the INSPIRE metadata flow in the near future.

1 Einführung: INSPIRE und PortalU[®]

Am 15. Mai 2007 trat die Richtlinie 2007/2/EG zur Schaffung einer Geodateninfrastruktur in der Europäischen Gemeinschaft (INSPIRE, *Infrastructure for Spatial Information in Europe*) in Kraft. Ziel der Richtlinie, die innerhalb von zwei Jahren von den Mitgliedstaaten der EU jeweils in nationales Recht umgesetzt werden muss, ist es, die grenzübergreifende Nutzung von Daten in Europa zu erleichtern. INSPIRE fordert webbasierte Online-Dienste für die Suche, die Visualisierung und den Download der Daten. Die technischen Einzelheiten zur Umsetzung der Richtlinie werden über Durchführungsbestimmungen geregelt, die zurzeit erarbeitet werden.

Als erste Durchführungsbestimmung soll diejenige zum Metadaten-Modell von INSPIRE veröffentlicht werden. Der Regelungsausschuß „INSPIRE Committee“ hat am 14. Mai 2008 den Durchführungsbestimmungen zu Metadaten ohne Gegenstimmen und Enthaltungen zugestimmt. Das Votum des Regelungsausschusses wird von der EU-Kommission nun dem Europäischen Parlament (EP) zugeleitet. Das EP hat eine Rückäußerungsfrist von 4 Wochen. Im Anschluss werden die Durchführungsbestimmungen dem „EU-Kabinett“ (EU-Kommissare) zur Zustimmung vorgelegt. Danach können die Durchführungsbestimmungen veröffentlicht werden und treten in Kraft. Die EU-Kommission rechnet damit, dass die Metadaten-Durchführungsbestimmungen zum Ende des Sommers veröffentlicht werden. Sie sind dann unmittelbar geltendes Recht in den Mitgliedsstaaten.

Die inhaltliche Ausarbeitung, an der die Koordinierungsstelle PortalU als Mitglied des INSPIRE Drafting Teams Metadata direkt beteiligt war, ist nach der Zustimmung des Regelungsausschusses in einem Stadium, dass für die Veröffentlichung keine Änderungen gegenüber dem derzeitigen Entwurf (Draft Implementing Rules for Metadata, Version 4) mehr erwarten läßt. Für die PortalU[®]-Software InGrid[®], deren Metadatenkatalog (UDK) bislang ISO 19115 / ISO 19119 konform gestaltet ist, wurden die für INSPIRE erforderlichen Datenmodellerweiterungen daher aus dem Draft / Version 4 abgeleitet.

2 Erforderliche Anpassungen am UDK-Datenmodell

In den folgenden Unterkapiteln werden die notwendigen Anpassungen und Erweiterungen am Metadatenmodell von InGrid[®] (UDK-Datenmodell) feldweise

beschrieben. Dort, wo es sich um echte Erweiterungen handelt, ist zu erwarten, dass die beschriebenen Anpassungen auch für andere, nicht auf InGrid[®]-Software basierende, ISO-19115/19119-konforme Kataloge Relevanz besitzen.

Der Umstieg auf das neue Datenmodell erfordert für die bestehenden Umweltdatenkataloge eine Altdatenübernahme. Wo möglich, soll die Befüllung der neuen Felder initial automatisiert erfolgen. Die Überführungsregeln werden, soweit vorhanden, ebenfalls dargestellt. Die überführten Umweltdatenkataloge werden im Folgenden auch als *InGrid[®]-Datenkataloge* (IDCs) bezeichnet.

2.1 Konformität (Draft: 2.2.7 Conformity)

Die Feldgruppe *Konformität* ist eine neue Pflichtfeldgruppe, in der beschrieben wird, in wie weit und zu welcher Spezifikation die in dem Metadatenobjekt beschriebenen Daten konform sind. Die Feldgruppe besteht aus einer 1-zu-N-Zuordnung von Feldpaaren.

Bei der *Spezifikation der Konformität* soll angegeben werden, zu welcher Durchführungsbestimmung der INSPIRE-Richtlinie die beschriebenen Daten konform sind. Hierbei soll die Möglichkeit bestehen mehr als einen Eintrag vorzunehmen. Es können grundsätzlich also auch Nicht-INSPIRE-Spezifikationen aufgenommen werden. Pro Spezifikationen sollen der Name der Spezifikation und ein entsprechendes Referenzdatum (Datum der Publikation, Datum der letzten Überarbeitung oder Erstellungsdatum) aufgeführt werden. Das Feld *Grad der Konformität* wird über eine Auswahlliste mit folgenden Einträgen zu füllen sein:

1. konform: Die Datenquelle ist vollständig konform zur zitierten Spezifikation
2. nicht konform: Die Datenquelle ist nicht konform zur zitierten Spezifikation
3. nicht evaluiert: Die Konformität der Datenquelle wurde noch nicht evaluiert

Standardvorgabe bei der Altdatenübernahme für *Grad der Konformität* ist *nicht evaluiert*, Standardvorgabe für *Spezifikation der Konformität* ist die *INSPIRE-Richtlinie*. Zukünftig, wenn die verschiedenen Durchführungsbestimmungen (insbesondere *Data Specifications*) verabschiedet sind, kann das Freitextfeld *Spezifikation der Konformität* ggf. durch eine Auswahlliste ersetzt bzw. unterstützt werden.

2.2 Eindeutiger Identifikator der Datenquelle (Draft: 2.2.1.5 *Unique resource identifier*)

Das Feld *Identifikator der Datenquelle* ist ein neues Pflichtfeld, das für jeden in den Objekten beschriebenen Datensatz einen eindeutigen Identifikator enthalten soll. Das Feld ist nur für die UDK-Objektklasse *Geoinformation/Karte* vorgeschrieben.

Bisher ist unklar, wie dieser Identifikator aussehen wird, dieser Punkt soll über die Durchführungsbestimmungen zu den *Data Specifications* geklärt werden.

Der Fachbezug der UDK-Objektklasse *Geoinformation/Karte* wird um ein einzeiliges Textfeld erweitert, dessen Befüllung verpflichtend ist.

Im Rahmen der Altdatenübernahme wird das Feld *Identifikator der Datenquelle* für jedes Objekt der Objektklasse *Geoinformation/Karte* automatisch mit der Kombination aus dem Katalognamen gefolgt von einem Doppelpunkt und dem Identifikator des Metadaten-Objektes (automatisch für jedes UDK-Objekt generierte UUID) gefüllt.

2.3 Zugangs- und Nutzungsbeschränkungen (Draft: 2.2.8. *Constraint related to access and use*)

Die INSPIRE-Anforderungen werden mit einer zweispaltigen Tabelle erfüllt, die sich aus den Feldern *Zugangsbeschränkungen* und *Nutzungsbedingungen* zusammensetzt. Mindestens ein Eintrag ist verpflichtend, wobei beide Felder ausgefüllt sein müssen.

2.3.1 Zugangsbeschränkungen (Draft: 2.2.8.2 *Limitations on public access*)

Das Feld *Zugangsbeschränkungen* bezieht sich auf die Auflistung der in Artikel 13 der INSPIRE-Richtlinie erlaubten Beschränkungen. Pro Zeile der Tabelle *Zugangsbeschränkungen* und *Nutzungsbedingungen* ist ein Eintrag verpflichtend.

Das Feld soll über die folgende, fest vorgegebene, Auswahlliste befüllt werden:

1. Keine
2. Aufgrund der Vertraulichkeit der Verfahren von Behörden
3. Aufgrund internationaler Beziehungen, der öffentlichen Sicherheit oder der Landesverteidigung
4. Aufgrund laufender Gerichtsverfahren

5. Aufgrund der Vertraulichkeit von Geschäfts- oder Betriebsinformationen
6. Aufgrund der Rechte des geistigen Eigentums
7. Aufgrund der Vertraulichkeit personenbezogener Daten
8. Aufgrund des Schutzes einer Person
9. Aufgrund des Schutzes von Umweltbereichen

Die Standard Vorgabe für den ersten Eintrag ist *keine*.

2.3.2 Nutzungsbedingungen (Draft: 2.2.8.1 Conditions applying to access and use)

Das Feld *Nutzungsbedingungen* soll die Bedingungen zur Nutzung des beschriebenen Datensatzes bzw. des Services enthalten. Es sollen beispielsweise die Kosten für die Nutzung der Daten angegeben werden. Pro Zeile der Tabelle *Zugangsbeschränkungen und Nutzungsbedingungen* ist ein Eintrag verpflichtend.

Bei frei nutzbaren Daten bzw. Services soll laut Durchführungsbestimmungen *Keine Einschränkungen* eingetragen werden. Sind die Bedingungen unbekannt, ist *Unbekannte Nutzungsbedingungen* einzutragen. Die UDK-Felder *Nutzungsanmerkungen* und *Kosten* in der UDK-Rubrik *Verfügbarkeit* entfallen.

2.3.3 Altdatenübernahme

Bei der Altdatenübernahme wird geprüft, ob eines der existierenden, derzeit optionalen UDK-Felder *Nutzungsanmerkungen* bzw. *Kosten* in der UDK-Rubrik *Verfügbarkeit* einen Eintrag enthält. Es werden folgende Fälle unterschieden:

1. Kein Eintrag: Das Feld *Zugangsbeschränkungen* erhält einen Eintrag *Keine*, das Feld *Nutzungsbedingungen* erhält einen Eintrag *Keine Einschränkungen*
2. Existierender Eintrag: Das Feld *Zugangsbeschränkungen* erhält einen Eintrag *Aufgrund der Rechte des geistigen Eigentums*. Das Feld *Nutzungsbedingungen* wird mit dem Inhalt der jetzigen beiden Felder gefüllt. Sind in beiden Feldern Einträge enthalten, werden Sie durch eine Absatzmarke getrennt zusammengeführt.

2.4 Fachliche Grundlage (Draft: 2.2.6.1 Lineage)

In den Durchführungsbestimmungen zu INSPIRE wird das Feld *Fachliche Grundlage* zum Pflichtfeld für Geodaten (Draft: *spatial datasets*) erklärt, also für diejenigen Objekte, die über die UDK-Objektklasse "Geoinformation/Karte" beschrieben werden. Es soll u.a. Angaben zur Entstehungsgeschichte sowie zur Qualität des Datensatzes enthalten.

Das Feld *Fachliche Grundlage* ist derzeit im UDK-Fachbezug *Geoinformation/Karte* ein optionales Feld. Es wird zukünftig als Pflichtfeld gekennzeichnet und zu den weiteren Pflichtfeldern an den Anfang des Fachbezugs verschoben. Beim Speichern wird, wie bei allen Pflichtfeldern, die Befüllung überprüft. Ist das Feld in einem existierenden Katalog nicht ausgefüllt, wird es im Rahmen der Altdatenübernahme mit dem Eintrag *Unbekannt* befüllt.

2.5 Service-Klassifikation (Draft: 2.2.2.2 *Spatial data service type*)

Das aktuelle UDK-Feld *Servicetyp* im Fachbezug der UDK-Objektklasse *Dienst / Anwendung / Informationssystem* wird in das Feld *Service-Klassifikation* überführt. Es wird wie bisher ein über eine Auswahlliste zu füllendes Pflichtfeld ausschließlich für die Objektklasse *Dienst / Anwendung / Informationssystem*. Die bisherige Auswahlliste wird durch die folgende geschlossene Auswahlliste (Draft: *B.3 Spatial data service type*) ersetzt:

- | | |
|----------------------------------|-------------------------------|
| 1. Suchdienste (CSW) | [Discovery Service] |
| 2. Visualisierungsdienste (WMS) | [View Service] |
| 3. Zugriffsdienste (WFS) | [Download Service] |
| 4. Transformationsdienste (WCTS) | [Transformation Service] |
| 5. Verkettete Geodatendienste | [Invoke Spatial Data Service] |
| 6. Andere Dienste | [Other Service] |

Im Rahmen der Altdatenübernahme wird der existierende Service-Eintrag anhand von Tabelle 1 umgesetzt.

Bisheriger Eintrag	Neue ID
CSW	1
WMS	2
WFS	3

Alle anderen Einträge (einschl. leeres Feld)	6
--	---

Tabelle 1: Service-Klassifikation

2.6 Typ des Dienstes (Draft: 2.2.3 Keyword, Annex B.4)

Dieses Feld wird mit INSPIRE für *spatial data services* verpflichtend. Umgesetzt wird es als einspaltige Tabelle im Fachbezug der UDK-Objektklasse *Dienst / Anwendung / Informationssystem*. Die von INSPIRE fest vorgegebene Auswahlliste (Draft: Annex B.4) wird für InGrid®-Datenkataloge um die ID=901: *Non Geographic Service* erweitert.

Bei der Altdatenübernahme wird gemäß Tabelle 2 die erste Zeile der neuen Tabelle mit einem Wert vorbelegt, der sich aus dem existierenden UDK-Feld *Service-Klassifikation* (siehe 2.5) ergibt.

ID <i>Service-Klassifikation</i>	ID <i>Typ des Dienstes</i>
1	207
2	202
3	201
6	901

Tabelle 2: Typ des Dienstes

2.7 Erstellungsmaßstab (Draft: 2.2.6.2 Spatial resolution)

Die bereits im Fachbezug *Geoinformation/Karte* vorhandene Feldgruppe *Erstellungsmaßstab* wird zusätzlich im Fachbezug der Objektklasse *Dienst / Anwendung / Informationssystem* eingeführt. Einträge in diese Feldgruppe sind optional.

2.8 Email-Adresse (Draft: 2.2.10.1 Metadata point of contact)

Für alle Adressen gilt, dass die Email-Adresse, abweichend vom aktuellen UDK-Datenmodell, ein Pflichtfeld wird. Im InGrid®-Datenkatalog wird die Angabe mindestens einer Email-Adresse daher verpflichtend. Die zugehörige Feldgruppe wird zu den anderen Pflichtfeldern an den Anfang der Erfassungsmaske gerückt und optisch als Pflichtfeld gekennzeichnet. Beim Speichern einer Adresse wird die Angabe mindestens einer Email-Adresse abgeprüft.

Die Angabe der Postadresse, die derzeit im UDK verpflichtend ist, wird durchgängig zu einem optionalen Feld, d.h. es verbleibt lediglich die Email-Adresse als Pflichtfeld in der Erfassungsmaske.

2.8.1 Altdatenübernahme

Ist für einen Adresseintrag keine Email-Adresse angegeben, wird in der UDK-Adress-Hierarchie schrittweise aufwärts bis zur Institution nach einer Email-Adresse gesucht. Die erste bei diesem Vorgang gefundene Email-Adresse wird dem Adresseintrag zugeordnet.

Ist der gesamten Institution keine Email-Adresse zugeordnet, wird die Email-Adresse des Katalogadministrators allen Adressen der Institution zugeordnet. Hierzu wird die Eingabe der Email-Adresse des Katalogadministrators beim Start der Altdatenübernahme vorgesehen. Im Normalfall sollen die Kataloge vor der Altdatenübernahme dahingehend überarbeitet werden, dass zumindest jeder Institution eine Email-Adresse zugeordnet ist.

2.9 Verschlagwortung (Draft: 2.4 Keyword)

Im UDK ist bislang die Vergabe von drei Thesaurusschlagworten pro Metadatenobjekt obligatorisch. Genutzt wird das Wortgut des *UMTHES*⁴, das zurzeit noch im Windows-UDK-Erfassungswerkzeug integriert ist und ab InGrid[®] 2.0 über den *Semantic-Network-Service*⁵ des Umweltbundesamts angeschlossen wird.

Für INSPIRE wird die Vergabe eines *GEMET*⁶-Schlagwortes (General Multilingual Environmental Thesaurus) verpflichtend. Da der Wortgutumfang des GEMET erheblich geringer ist als der des UMTHES und darüber hinaus der deutsche GEMET-Anteil komplett im UMTHES abgebildet ist, ergeben sich für InGrid[®] 2.0 nur marginale Änderungen. Auch zukünftig werden bei der UDK/IDC Metadatenerfassung drei UMTHES-Schlagworte verpflichtend sein. Über einen Prüfmechanismus wird abgeprüft, ob sich unter den drei gewählten UMTHES-Deskriptoren ein GEMET-Schlagwort befindet. Ist dies nicht der Fall, werden die Bearbeiter vor dem endgültigen Verspeichern des Metadatenobjekts bzw. vor Übersendung an die Qualitätssicherung zur Eingabe eines GEMET-Schlagworts

⁴ <http://www.umweltbundesamt.de/uba-info/dokufabib/thes.htm>

⁵ <http://www.semantic-network.de>

⁶ <http://www.eionet.europa.eu/gemet>

verpflichtet. Dabei werden Sie von verschiedenen Thesaurus-Assistenten unterstützt: Dem Thesaurus-Navigator, der ein blättern in der Thesaurushierarchie (Abb. 1) sowie eine Suche im UMTHEs-Wortgut (Abb. 2) ermöglicht sowie einem Verschlagwortungsassistenten, der Inhalte von Textfeldern an den SNS übergibt und vom Service Schlagwort-Vorschläge zurück erhält.

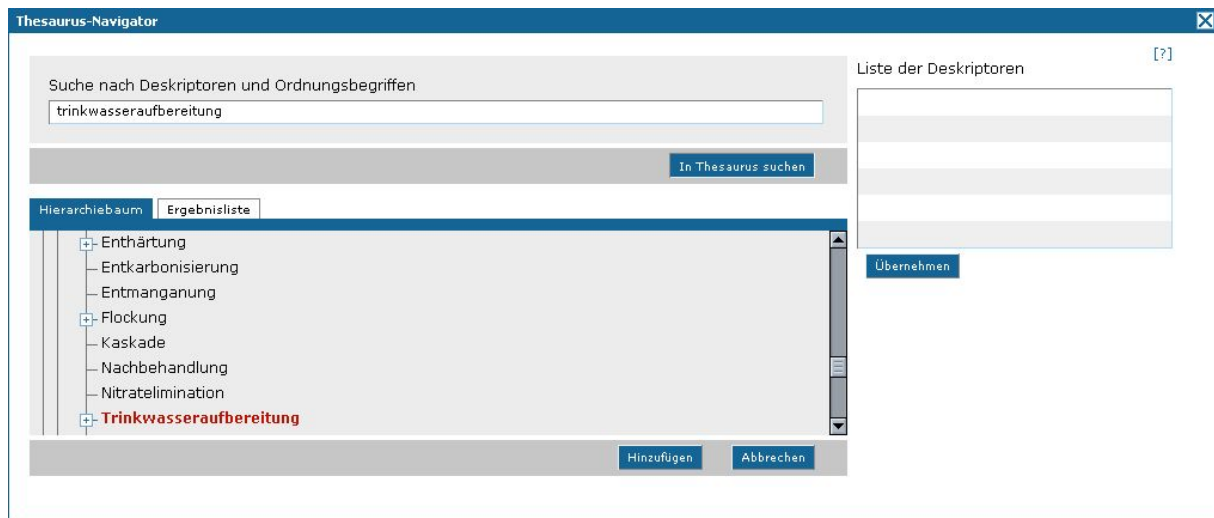


Abb. 1: InGrid® 2.0 Thesaurus-Navigator / Hierarchiebaum

Für die Altdatenübernahme ist für jedes Objekt eine automatische Überprüfung der existierenden Schlagworte vorgesehen. Wird auf diesem Wege zunächst kein GEMET-Schlagwort gefunden, wird in der UMTHEs-Hierarchie von den eingetragenen Schlagworten aus so weit „nach oben“ gewandert, bis ein entsprechender GEMET-Deskriptor gefunden wird. Dieser wird als zusätzliches Schlagwort zum bestehenden Metadatenobjekt hinzu gefügt.

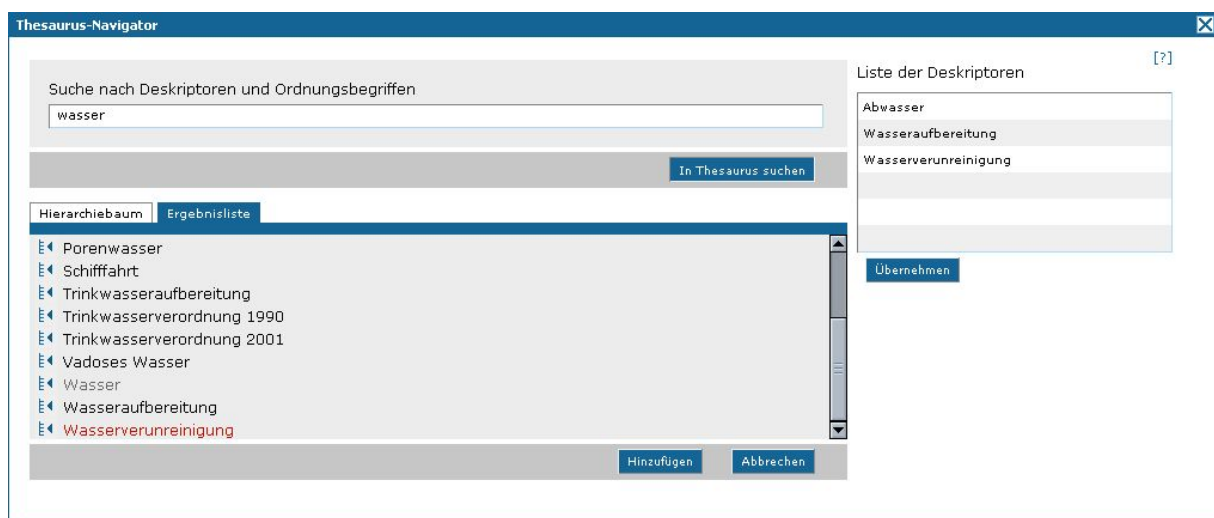


Abb. 2: InGrid® 2.0 Thesaurus-Navigator / Suche im Thesaurus

3 GDI-DE / INSPIRE Metadaten Infrastruktur

Im folgenden werden einige Überlegungen zur Frage, wie eine deutsche Metadateninfrastruktur für INSPIRE realisiert werden könnte, vorgestellt. Die Darstellung bewegt sich dabei entlang des Diskussionsstands im GDI-DE Modellprojekt Geodatenkatalog-DE, in dem die Koordinierungsstelle PortalU aktiv mitarbeitet. Aus den GDI-DE- und INSPIRE-Anforderungen werden, als Diskussionsgrundlage, verschiedene Weiterentwicklungsoptionen für die Software InGrid[®] abgeleitet.

3.1 GDI-DE Modellprojekt Geodatenkatalog-DE

Ziel des Modellprojekts ist der Aufbau einer Topologie existierender Katalogdienste und deren Bündelung im Sinne einer zentralen Abfragemöglichkeit der verteilten Dienste. Hierzu ist es erforderlich, typische Metadateninformationssysteme auf höheren Verwaltungsebenen (z.B. Bundes-MIS, Landes-MIS, länderübergreifende Fach-MIS) im Projekt zu verknüpfen. Konkret sollen eine Auswahl an fachübergreifenden Geodatenkatalogen der Länder, der Geodatenkatalog des Bundes sowie die für INSPIRE absehbar relevanten und länderübergreifenden Fachkataloge *PortalU*[®] und *Adv-MIS*⁷ zusammengeführt werden.

Nutzeranforderungen werden im Projekt u.a. über die Beteiligung des JRC⁸ und der GIW-Kommission⁹ berücksichtigt.

Die technische Umsetzung wird den ersten Prototyp eines zentralen, deutschen INSPIRE-Metadatenknotens bilden. Die angestrebte Lösung soll einen performanten Zugriff und eine hohe Verfügbarkeit sowie Redundanzfreiheit gewährleisten.

3.2 Anforderungen seitens GDI-DE und INSPIRE

Im GDI-DE-Architekturkonzept¹⁰ ist als Spezifikation für Katalogdienstschnittstellen das *OpenGIS[®] Catalogue Services Specification 2.0.2 - ISO Metadata Application Profile (1.0.0)*¹¹ festgeschrieben. Dieses Applikationsprofil wird, unter Berücksichtigung des aktuellen Diskussionsstands im INSPIRE Drafting Team Network Services, auch der fachliche Katalogdienst-Standard für INSPIRE werden.

⁷ <http://www.geodatenzentrum.de/isoinfo/>

⁸ <http://ec.europa.eu/dgs/jrc/>

⁹ <http://www.geobusiness.org/Geobusiness/Navigation/gjwk.html>

¹⁰ http://www.gdi-de.org/de/download/GDI_ArchitekturKonzept_V1.pdf

Die technisch/funktionalen Anforderungen an INSPIRE-konforme Katalogdienste werden im Drafting Team Network Services diskutiert und in den *Draft Implementing Rules for Discovery and View Services* dokumentiert. Wenngleich die zu erfüllenden Performanzwerte noch nicht endgültig fest stehen, zeichnet sich ab, dass Harvesting- und Indexierungsvorgänge erforderlich sein werden um einen INSPIRE-konformen deutschen Metadatenknoten bereit zu stellen. So werden in den *Draft Implementing Rules for Discovery and View Services (IR1) Version 2.0* u.a. folgende Anforderungen gestellt:

1. Maximale Zeit zum Senden der Rückmeldung nach einer Suchanfrage (in 90% der Fälle, bei 250 gesendeten Metadatensätze): 3 Sekunden
2. Minimum möglicher gleichzeitiger Anfragen an den Service: 100

Die Suchkriterien für INSPIRE-Suchdienste werden in Artikel 11(2) der INSPIRE-Richtlinie festgelegt. Danach sind zumindest Kombinationen von folgenden Suchkriterien zu gewährleisten:

1. Schlüsselwörter
2. Klassifizierung von Geodaten und Geodatendiensten
3. Qualität und Gültigkeit der Geodatensätze
4. Grad der Übereinstimmung mit den in Artikel 7(1) vorgesehenen Durchführungsbestimmungen
5. geografischer Standort
6. Bedingungen für den Zugang zu und die Nutzung von Geodatensätzen und -diensten
7. Behörden, die für die Erzeugung, Verwaltung, Erhaltung und Verbreitung von Geodatensätzen und -diensten zuständig sind

Für existierende Kataloge stellen sich damit Probleme analog zu den bereits weiter oben im Kontext der Metadatendurchführungsbestimmungen diskutierten. Zum Beispiel wird das Feld *Konformität* als neues Pflichtfeld eingeführt, um das in Punkt

¹¹ http://portal.opengeospatial.org/files/?artifact_id=21460

(4) definierte Suchkriterium zu erfüllen. Im Rahmen von Suchdiensten kommt hinzu, dass sichergestellt werden muss, dass dieses Feld durchsuchbar ist.

Ähnliches gilt auch für die *Nutzungsvereinbarungen* (6), die z.B. im InGrid[®]-Metadatenmodell derzeit als optionales Feld enthalten sind. Zum Einen wird dieses Feld zum Pflichtfeld erklärt, zum Anderen muss es auch über die Schnittstelle durchsuchbar gemacht werden.

Bezüglich der Funktionen der INSPIRE-Suchdienste fordern die *Draft Implementing Rules for Discovery and View Services* gegenüber CSW AP ISO 1.0 erweiterte Funktionalitäten, indem die CSWT-Operationen *Manager.Harvest* und *Manager.Transaction* von optional auf verpflichtend umgesetzt werden.

Unter einem Harvesting-Mechanismus wird die Fähigkeit eines Suchdienstes verstanden, einen anderen Suchdienst abfragen zu können, wobei die Daten des abgefragten Suchdienstes in dem des Anfragenden gespeichert werden. Mit der Software-Version InGrid[®] 2.0 wird PortalU[®] diese Fähigkeit besitzen.

Unter einem Transaktionsmechanismus wird die Funktionalität eines Suchdienstes verstanden, über die CSW- (bzw. die CSWT-) Schnittstelle in den Datenbestand hineinschreiben zu können und zwar mit den Funktionen *Einfügen*, *Ändern* und *Löschen*. Diese Funktionalitäten, die von außen auf einem Katalog durchgeführt werden können, wurden für PortalU[®] als zu risikoreich und damit als explizit nicht gewünscht erachtet.

Aus Sicht der Koordinierungsstelle PortalU können beide Funktionalitäten für einzelne Suchdienste interessant sein. Eine verpflichtende Vorgabe für alle INSPIRE-Suchdienste erscheint aber sowohl aus Kostengründen als auch im Licht der fachlichen Anforderungen nicht sinnvoll und kann insbesondere für den Transaktionsdienst mit Risiken für die Datensicherheit verbunden sein. Die Koordinierungsstelle hat daher in der Kommentierungsphase zu den Draft Implementing Rules for Discovery and View Services gefordert, die Funktionalitäten Harvesting und Transaktion auf optional zurückzustufen.

3.3 Topologie der Katalogdienste

Grundsätzlich kann die Strukturierung von Diensten einem fachbezogenen Aufbau folgen, vergleichbar dem Ansatz von AdV MIS und PortalU[®], oder auch streng

administrativ, basierend auf zentralen Kommunen- und Landeskatalogen bis hin zum Bundesknoten erfolgen. Eine weitere Möglichkeit besteht im Verzicht auf eine mehrstufige Topologie, d.h. unterhalb des zentralen Knotens werden alle Kataloge unabhängig von ihrer Herkunft direkt angeschlossen. Faktisch existieren derzeit alle genannten Topologiearten parallel und in Mischformen und sollen im Modellprojekt berücksichtigt werden. In einem ersten Arbeitsschritt müssen die vorhandenen Topologien charakterisiert und in eine Registry überführt werden. Der Registry-Gedanke wird derzeit in einem weiteren GDI-DE-Modellprojekt konkretisiert.

Insbesondere die hohen Performanzanforderungen der Draft Implementing Rules for Discovery and View Services (IR1) Version 2.0 (s.o.) stecken den Rahmen für geeignete Implementierungstechnologien ab. Ohne Harvesting- und Indexierungsmechanismen wird eine INSPIRE-konforme Katalogimplementierung nicht möglich sein. Die Vorteile solcher Verfahren liegen in der Erhöhung der Antwortgeschwindigkeit sowie in der Möglichkeit gemischte, datenquellenübergreifende und nach Relevanz im Hinblick auf den Suchbegriff sortierte, Ergebnislisten zu erstellen. Allerdings entstehen durch das Harvesting in zentrale Repositorien auch Probleme wie Dublettenpotential, redundante Datenhaltung, Daten-Aktualität, rechtliche Fragestellungen und das Erfordernis, große Datenmengen regelmäßig über das Netz zu transportieren. Diese sollen im Modellprojekt adressiert und praktisch analysiert werden.

3.4 Erweiterungen der Software InGrid[®]

Um den steigenden Anforderungen an Brokersysteme wie PortalU[®] gerecht zu werden, sollen insbesondere die InGrid[®]-CSW-Komponenten erweiterte Funktionalitäten erhalten. Die Anpassungen befinden sich aktuell im Abstimmungsstadium und sollen im Folgenden vorab zur Diskussion gestellt werden.

3.4.1 Harvesting und Indexierung

In einem ersten Schritt ist ein Mechanismus vorzusehen, der über Harvesting-Abfragen in der Lage ist, den gesamten Dateninhalt einzelner Kataloge zu ernten und anschließend zu indexieren. Als Indexierer ist *Apache Lucene*¹² zu verwenden, um die Kompatibilität des Rankings mit anderen Schnittstellenadaptern (*iPlugs*) von

¹² <http://lucene.apache.org/>

InGrid[®] zu gewährleisten. Aus Gründen der Skalierbarkeit, Lastverteilung und Flexibilität erhält jede angeschlossene Datenquelle einen eigenen Index, dessen Harvestingrythmus individuell konfigurierbar ist. Der Index wird am jeweiligen iPlug vorgehalten und konfiguriert. Die Zusammenführung der unterschiedlichen Indizes erfolgt zum Suchanfragezeitpunkt am InGrid[®]-iBus [Klenke et al., 2006].

Das Harvesting soll über die Abfrage `GetRecords (ElementSetName=brief)` mit einer Wildcard-Query und anschließendem Herunterladen der Metadaten (`GetRecordById, ElementSetName=full`) realisiert werden. Die Inhalte werden am jeweiligen iPlug in Form einer XML-Datei temporär gespeichert.

Der Indexprozess wird nach dem Harvesting automatisch angestoßen. Er indiziert die CSW-XML Datei(en). Die Indizierung erfolgt nach einer Indexfeld-zu-XPath-Abbildungsregel. Indexfelder werden dabei einem bestimmten XPath zugeordnet. Damit eine einfache Integration in die InGrid[®] 1.x Systemlandschaft möglich ist werden als Feldnamen die Bezeichnungen verwendet, die im iPlug auf Basis des InGrid[®]-Datenmodells vorgegeben sind. Es wird ein Konfigurationsmechanismus implementiert, der das Austauschen der Mapping-Klassen vorsieht. Dadurch wird es möglich, unaufwendig verschiedene Mappings im iPlug zu unterstützen, z.B. AP ISO 0.9.3. und AP ISO 1.0.

Es ist sicherzustellen, dass trotz der Abbildung der Indexinhalte auf die iPlug-Feldnamen eine Abfrage der originären Inhalte der geharvesteten XML-Daten möglich ist (`ElementSetName=original`).

Standardmäßig erfolgt die Beantwortung von Anfragen performant aus dem Index, alternativ ist ein Mechanismus vorzusehen, der es ermöglicht, für die Detailliergeberdarstellung (`GetRecordById`) auch auf die Originaldatenquelle zurück zu greifen.

Es ist, insbesondere bei mehrstufigen Abfragekaskaden, möglich, dass aus unterschiedlichen CSW-Quellen die selben Datensätze geliefert werden. Ein Auftreten von Duplikaten in der resultierenden Antwort im Portal bzw. an CSW- und OpenSearch¹³-Schnittstelle ist daher zu berücksichtigen. Jeder CSW-Datensatz hat eine eindeutige ID, die zur Filterung von Dubletten verwendet werden kann. Sollten

¹³ <http://www.opensearch.org/>

über die ID-Dubletten gefunden werden, ist zu prüfen, welcher Datensatz gemäß Zeitstempel der aktuellere ist. Dieser ist weiter zu geben. Der Harvest-Prozeß sowie der iBus sind um einen entsprechenden Mechanismus zu erweitern.

3.4.2 Erweiterung des CSW-Profiles um ein PortalU[®]-Resultset

Das AP ISO 1.0 wird um ein schmales *PortalU[®]-Resultset* erweitert. Ziel ist es, Informationssysteme zu unterstützen, die nur ein eingeschränktes, nicht CSW/ISO-konformes Datenmodell besitzen, jedoch über eine eigene Web-Darstellung der Datensätze verfügen. Das Resultset ist optional, das heißt, ISO- oder INSPIRE-konforme CSW müssen es nicht unterstützen. Tabelle 3 zeigt Inhalte des Resultsets.

CSW-Name	Pflicht (M) / Optional (O)	Beschreibung
Title	M	Titel des Datensatzes
Abstract	M	Beschreibung des Datensatzes
OnlineResource	M	URL zur Darstellung des Datensatzes im Original-System
Subject	O	Schlagworte
GeographicDescriptionCode	O	Name des geographischen Objekts
BoundingBox	O	geografische Bounding Box (WestBoundLongitude, NorthBoundLatitude, EastBoundLongitude, SouthBoundLatitude)
TemporalExtent	O	Zeitbezug der beschriebenen Daten (TempExtent_begin, TempExtent_end)

Tabelle 3: PortalU[®]-Resultset

3.4.3 Server für PortalU[®]-Resultset

Es wird eine einfacher Server erstellt, der es einem Anwender erlaubt, seine Datenbank auf das PortalU[®]-Resultset abzubilden.

Hierzu wird eine Administrationsoberfläche entwickelt, die dem Administrator die Struktur der Datenbank grafisch anzeigt und über die er das Mapping auf die Felder des PortalU[®]-Resultsets festlegen kann.

Eine Anfrage wird dann über die gemappten Felder der Datenbank geführt und das PortalU[®]-Resultsets zurückgegeben.

Es müssen nur Anfragen beantwortet werden, die auch von InGrid[®] unterstützt werden. Die Unterstützung weitergehender Anfragen, die ggf. auch durch das CSW-Profil definiert werden, ist nicht gefordert.

4 Zusammenfassung und Ausblick

Wenngleich die INSPIRE-Durchführungsbestimmungen aktuell an Kontur gewinnen bestehen noch große Unsicherheiten hinsichtlich deren praktischer Umsetzung. Grundsätzlich erscheinen die relevanten existierenden deutschen Katalogsysteme fachlich, durch ihre Unterstützung der ISO-Standards 19115 und 19119, zukunftsfähig aufgestellt. Auch auf technischer Ebene werden momentan bei vielen Betreibern, nicht zuletzt auf Grundlage des aktuellen GDI-DE-Architekturkonzepts, die Weichen in Richtung einer AP ISO 1.0-konformen CSW-Schnittstelle gestellt.

Der Ausgestaltung der Topologie deutscher Katalogdienste im Hinblick auf die ab 2010 verpflichtende Bereitstellung INSPIRE-relevanter Informationen ist demgegenüber noch völlig offen und birgt verschiedenste Schwierigkeiten, die nicht primär konzeptionell/theoretisch sondern insbesondere anhand konkreter technischer Umsetzungen beleuchtet und gelöst werden können. Insofern kommt dem GDI-DE-Modellprojekt Geodatenkatalog-DE eine maßgebliche Bedeutung für die Ausgestaltung der deutschen INSPIRE-Architektur zu.

5 Literaturverzeichnis

[Klenke et al., 2007]

Klenke, Martin; Vögele, Thomas; Kruse, Fred; Lehmann, Hanno: PortalU[®] & InGrid[®] - Werkzeuge zur Erstellung, Recherche und Verteilung von Metadaten. AGIT 2007 - Symposium und Fachmesse für Angewandte Geoinformatik an der Universität Salzburg, 2007. <http://kurzlink.de/kle-2007-2>

[Klenke et al., 2006]

Klenke, Martin; Kruse, Fred; Lehmann, Hanno; Riegel, Thomas; Vögele, Thomas: InGrid[®] 1.0 – The Nuts and Bolts of PortalU[®]. EnviroInfo 2006, Graz, 2006. <http://kurzlink.de/kle-2006-1>