



Portal U

Fachliches Feinkonzept

InGrid 1.1

Version 1.6

11.10.2005

Änderungshistorie

Alle wesentlichen Änderungen und Ergänzungen für jede Version des Dokumentes werden in der untenstehenden Tabelle kurz aufgeführt.

Version	Datum	Änderung
1.0	24.01.2005	Entwurf der KUG
1.1	17.05.2005	gein [®] 2.1 umbenannt in InGrid 1.1; Anmerkungen der AG IUTE eingearbeitet
1.2	01.06.2005	Abschnitte in Kap. 4.7 und 4.8 überarbeitet; Anmerkungen des UBA eingefügt
1.3	16.06.2005	Überarbeitung nach Diskussion in der 9. Sitzung der AG IUTE
1.4	04.07.2005	Überarbeitung im Nachgang zur 9. Sitzung der AG IUTE
1.5	05.09.2005	Einpflegen redaktioneller Änderungen
1.6	11.10.2005	Redaktioneller Änderungen nach 11. Sitzung LA UDK/GEIN

Anlagen:

- Fachliches Feinkonzept PortalU / InGrid 1.0
- DV-technisches Feinkonzept PortalU / InGrid 1.0: iBus
- DV-technisches Feinkonzept PortalU / InGrid 1.0: iPlug
- DV-technisches Feinkonzept PortalU / InGrid 1.0: Portal
- DV-technisches Feinkonzept PortalU / InGrid 1.0: WMS
- DV-technisches Feinkonzept PortalU / InGrid 1.0: Styleguide
- DV-technisches Feinkonzept PortalU / InGrid 1.0: Datenmodell
- DV-technisches Feinkonzept PortalU / InGrid 1.0: Schnittstellen
- OpenGIS[®] Catalogue Services Specification 2.0 - ISO19115/ISO19119 Application Profile for CSW 2.0
- Software Windows-UDK 5.0 (<http://www.udk-gein.de/download.shtml>)

Inhalt:

1	Einleitung	5
2	Ist-Analyse	6
2.1	<i>InGrid 1.0</i>	6
2.1.1	Recherche nach Metadaten in InGrid 1.0	6
2.1.2	Installation von InGrid 1.0	7
2.2	<i>Windows - UDK 5.0</i>	7
2.2.1	Allgemein	7
2.2.2	Der UDK-Z	8
3	Allgemeine Anforderungen an InGrid 1.1	9
3.1	<i>Metadaten-Erfassung über das Internet</i>	9
3.2	<i>Modularer Aufbau</i>	9
3.3	<i>Mehrsprachigkeit</i>	10
3.3.1	Oberfläche	10
3.3.2	Dateninhalte / Suche	10
3.4	<i>Metadaten-Kataloge</i>	10
3.5	<i>UUID (Universal Unique Identifier)</i>	11
3.6	<i>Datenmodell</i>	11
3.7	<i>Aktualität des Index</i>	11
4	Anforderungen an das Modul „Metadatenerfassung“	12
4.1	<i>Katalogverwaltung</i>	12
4.2	<i>Nutzerverwaltung</i>	12
4.3	<i>Erfassung von Objekten</i>	13
4.3.1	Allgemein	14
4.3.2	Assistent zur Unterstützung bei der Erfassung von Objekten	14
4.3.3	Hierarchien von Objekten	15
4.3.4	Einbindung von Adressen	15
4.3.5	Thesaurus-Suchbegriffe	15
4.3.6	Geografische Metadaten	16
4.3.6.1	Geografische Namen	16
4.3.6.2	Weitere geografische Metadaten	16
4.3.7	Zusätzliche Felder	17
4.4	<i>Erfassung von Adressen</i>	17
4.4.1	Allgemein	17
4.4.2	Hierarchien von Adressen	17
4.5	<i>Workflow-Kontrolle für die Erfassung</i>	18
4.6	<i>Darstellung und Funktion von Hierarchien</i>	19
4.6.1	Objekt-Hierarchie	20
4.6.2	Adress-Hierarchie	20
4.6.3	Thesaurus-Struktur	21
4.7	<i>Im- und Export – Funktionalität</i>	22
4.7.1	Import aus UDK 5.0	22
4.7.2	Im- und Export einzelner Kataloge	22
4.7.3	Im- und Export des Gesamtdatenbestandes	22
4.7.4	Teilimport und -export	23
4.7.5	Datenimport von ESRI ArcGIS Metadaten	24
4.7.6	Export einzelner Objekte über eine URL	25
4.7.7	„Harvesting“ von InGrid Metadaten	25

4.8	<i>Gesamtkatalog-Management</i>	25
4.8.1	Katalog-Statistik	25
4.8.2	Analyse-Funktionen	25
4.8.3	Pflege von Auswahllisten.....	26
4.8.4	Im- und Export von Auswahllisten und anderen Kataloginformationen.....	26
4.8.5	Absolutes Löschen von Adressen	26
4.8.6	Definition zusätzlicher Felder	26
4.8.7	Updatefunktionalität für Verwaltungsgrenzen.....	27
4.8.7.1	Hintergrund	27
4.8.7.2	Update von Verwaltungsgrenzen / Aktualisierung betroffener Objekte.....	27
4.8.7.3	Zusätzliche Hierarchiestufen von Verwaltungsgrenzen.....	27
4.8.7.4	Automatisches Update dezentraler Installationen	28

1 Einleitung

Das Umweltinformationsnetz Deutschland gein[®] (German Environmental Information Network) ist „das Portal für Umweltfragen“ in Deutschland. Es erschließt die im Internet verteilten Informationen zahlreicher öffentlicher Einrichtungen. Zurzeit sind mehr als 400.000 Internetseiten von ca. 90 Informationsanbietern sowie Schnittstellen zu sieben umweltbezogenen Datenbanken mit über 500.000 Einzelinformationen zugänglich. gein[®] wurde in den Jahren 1998 bis 2000 vom Umweltbundesamt (UBA) aufgebaut und bis Ende 2002 laufend fortentwickelt.

Der Umweltdatenkatalog (UDK) ist das Informationssystem zum Nachweis von Umweltinformationen der öffentlichen Verwaltungen in Deutschland (aktuell ca. 30.000 Einträge). Die Grundkonzeption entstand in den Jahren 1991 bis 1995 im Niedersächsischen Umweltministerium. Die Weiterentwicklung erfolgte basierend auf einer Verwaltungsvereinbarung zwischen Bund und Ländern.

Seit dem 01.01.2003 ist die Verwaltungsvereinbarung UDK/GEIN (VwV UDK/GEIN) in Kraft. Mit ihr wurden UDK und gein[®] organisatorisch zusammengeführt. Die VwV UDK/GEIN wurde bisher vom Bund und von 15 Bundesländern unterzeichnet (siehe auch <http://www.udk-gein.de>).

Technisch werden gein[®] und die Recherche-Anwendung des UDK, der WWW-UDK, derzeit zusammengefasst. Im zweiten Quartal 2006 soll die neue Plattform unter dem Namen „PortalU“ ins Netz gestellt werden. Die Software, über die PortalU realisiert wird, heißt „InGrid 1.0“ (InGrid steht für „Information Grid“).

In einem nächsten Schritt soll die Metadaten-Erfassungsanwendung, der Windows-UDK, ebenfalls als Webanwendung in InGrid integriert werden. Die zu entwickelnde Gesamtanwendung wird im Folgenden InGrid 1.1 genannt.

In Kapitel 2 wird der derzeitige Stand von InGrid 1.0 und der Metadaten-Erfassungskomponente Windows-UDK 5.0 beschrieben. In Kapitel 3 sind die allgemeinen Anforderungen an das neue System formuliert. Die fachlichen Anforderungen an das zu entwickelnde Modul zur Erfassung von Metadaten sind in Kapitel 4 beschrieben.

2 Ist-Analyse

2.1 InGrid 1.0

Das Umweltinformationsnetz Deutschland (German Environmental Information Network, gein[®]; <http://www.gein.de>) wurde von 1998 bis 2000 vom Umweltbundesamt entwickelt und erstmals zur Expo 2000 präsentiert.

gein[®] erschließt die Informationen zahlreicher öffentlicher Einrichtungen, wie Umweltbehörden, Ämter und Ministerien auf Bundes- und Landesebene. Es werden sowohl die Informationen, die sich auf den Internetseiten der Behörden befinden, als auch Daten, die in Datenbanken enthalten sind, zugänglich gemacht.

gein[®] dient damit als Informationsbroker für Umweltinformationen in Deutschland. Es ist "das (offizielle) Portal für Umweltfragen" in Deutschland.

Derzeit sind ca. 90 Anbieter von Umweltinformationen aus Behörden und anderen öffentlichen Einrichtungen aus Bund und Ländern in gein[®] zusammengeschlossen, die mehr als 400.000 Webseiten zur Verfügung stellen. Weiterhin sind sieben Datenbanken angeschlossen, über die mehr als 500.000 Datensätze zugänglich sind.

Zurzeit wird die Anwendung gein[®] zu InGrid 1.0 weiterentwickelt. Insbesondere wird die Web-Recherche-Komponente des UDK (WWW-UDK) in InGrid 1.0 integriert (siehe auch Anlage).

2.1.1 Recherche nach Metadaten in InGrid 1.0

Die Recherche nach Metadaten wird vollständig in InGrid 1.0 integriert. Als Datenmodell wird das des UDK 5.0 verwendet. Die Funktionalität von InGrid 1.0 wird sich soweit anpassen lassen, dass sich die Anwendung bei Bedarf auf die Metadaten-Recherche reduzieren lässt. Dies wird dann in etwa dem ehemaligen WWW-UDK 5.0 (<http://www.umweltdatenkalog.de>) entsprechen.

Die Anwendung InGrid 1.0 wird über die folgenden Suchfunktionalitäten verfügen, über die nach Metadaten recherchiert werden kann (siehe auch Anlagen „DV-technisches Feinkonzept PortalU / InGrid 1.0“):

- Einfache Suche
- Erweiterte Suche
- Kartenbasierte Suche

Weiterhin kann nach Adressen gesucht werden.

InGrid 1.0 wird über die Funktionalität der Detaildarstellung der Metadaten (UDK-Objekte) und der Adressen (UDK-Adressen) verfügen.

Die Funktionalitäten für Recherche und Detaildarstellung von Metadaten sollen im vollen Umfang auch für InGrid 1.1 genutzt werden.

2.1.2 Installation von InGrid 1.0







Es ist vorgesehen, die Anwendung InGrid 1.0 auf der von der KUG bei einem externen Provider angemieteten Hardware zu betreiben. Damit ergeben sich nach derzeitigem Stand folgende technische Randbedingungen:

- 3 Server der folgenden Hardware:
 - Compaq Server DL380
 - 2 Prozessoren
 - 2 GB RAM
- Software-Ausstattung und Betriebssysteme:
 - SuSE Linux SLES 9.0
 - Apache 2.0
 - Jakarta Tomcat 5.5.9
- Datenbank:
 - MySQL 4.1.10a

2.2 Windows - UDK 5.0

2.2.1 Allgemein

Der Umweltdatenkatalog (UDK) ist ein Informationssystem zum Nachweis von Umweltinformationen der öffentlichen Verwaltungen. Der UDK gibt Auskunft darüber, "wer" "wo" über "welche" Umweltinformationen verfügt. Folgende Informationskategorien werden nachgewiesen:

-  Datensammlung / Datenbank
-  Dienst / Anwendung / Informationssystem
-  Dokument / Bericht / Literatur
-  Geoinformation / Karte
-  Organisationseinheit / Fachaufgabe
-  Vorhaben / Projekt / Programm

Die Beschreibung der Umweltinformationen erfolgt über Formularfelder, die von den Informationsanbietern selbst oder mit deren fachlicher Unterstützung ausgefüllt werden. Die wichtigsten Felder bzw. Feldgruppen des UDK sind:

- Titel (Objektname)
- Informationskategorie (Objektklasse)
- Allgemeine Beschreibung
- Auskunftsadresse

- Raumbezug (administrative oder geografische Einheit)
- Zeitbezug (von ... bis)
- Fachbezug (je nach Informationskategorie)

Zur Verschlagwortung wird ein strukturiertes Wortgut verwendet (Thesaurus des Umweltbundesamtes in Dessau), das den recherchierenden Benutzer unterstützt.

Unter <http://www.udk-gein.de/allgem/udk.shtml> bzw. in den Publikationen unter <http://www.udk-gein.de/allgem/pub-udk.shtml> sind weiter gehende Beschreibungen des UDK vorhanden.

Die Erfassung der im WWW-UDK abrufbaren Metadaten erfolgt über eine Windows-Software, den Windows-UDK. Dieser liegt derzeit in der Version 5.0 vor. Die client/serverfähige Anwendung bietet weiterhin umfangreiche Rechercheinstrumente, wengleich sie primär für die Erfassung und Beschreibung von Umweltinformationen konzipiert wurde.

Eine Beschreibung des Windows-UDK ist unter der URL <http://www.udk-gein.de/software/SoftwareWindowsUdk.shtml> zu finden. Die Einzelplatzversion der Software kann unter <http://www.udk-gein.de/download.shtml> herunter geladen werden.

Die Recherchefunktionalität des Windows-UDK ist zu einem großen Teil bereits durch InGrid 1.0 abgedeckt (siehe Anlagen). Dieser Anteil soll für InGrid 1.1 weitgehend unverändert übernommen werden und um die fehlenden Recherchefunktionalitäten (siehe Kapitel 4.6) ergänzt werden.

Die Erfassungsfunktionalität des Windows-UDK 5.0 soll, soweit dies über Internet-Technologien möglich ist, in InGrid 1.1 realisiert werden.

2.2.2 Der UDK-Z

Der UDK-Z ist ein eigenständiges Modul des Windows-UDK 5.0. Er ist in die Administrationsanwendung „UDK-Administrator“ integriert und dient zur Pflege eines kompletten Kataloges. Der UDK-Administrator wird nur den Administratoren zur Verfügung gestellt, die den gesamten Katalog (Zentralkatalogadministratoren) betreuen. Die Funktionen des UDK-Z, die in InGrid 1.1 zu übernehmen sind, werden in Kapitel 4.8 beschrieben.

3 Allgemeine Anforderungen an InGrid 1.1

3.1 Metadaten-Erfassung über das Internet

Die grundsätzliche Neuerung von InGrid 1.1 gegenüber dem UDK 5.0 bzw. der Metadatenkomponente von InGrid 1.0 wird die Möglichkeit sein, Metadaten über das Internet zu erfassen. Hierzu soll eine Anwendung realisiert werden, die browserbasiert arbeitet und keine weiteren Installationen (Anwendungen, Plug-Ins, etc.) beim Nutzer erfordert.

3.2 Modularer Aufbau

Die Anwendung InGrid 1.0 wird modular entwickelt. Sie besteht aus folgenden Modulen (siehe Anlage):

1. Modul „Portal“ (Oberfläche von InGrid)
2. Modul „Informations Bus“ (Modul „iBus“)
3. Modul „iPlug“ (verschiedene Plugins wie Suchmaschine, Schnittstellen, etc.)
4. Modul „Web-Mapping-Service“ (Modul „WMS“)

Die Anwendung InGrid 1.1 ist um das folgende Modul zu erweitern:

5. Modul „Metadatenerfassung“

Die Module sind als Webservices konzipiert. Alle Wechselwirkungen zwischen den Modulen und externen Systemen können über das SOAP-Protokoll geführt werden. Zusätzlich und alternativ werden zum Teil performantere Protokolle genutzt. Das gleiche muss auch für das 5. Modul „Metadatenerfassung“ gelten.

Alle gegebenenfalls notwendigen Änderungen an den Modulen 1-4 sind im technischen Feinkonzept InGrid 1.1 zu spezifizieren.

Über Verweise (z.B. Links, Buttons oder Menüeinträge) müssen, für die recherchierenden Nutzer, die in Kapitel 4.6 beschriebenen zusätzlichen Recherche-Möglichkeiten über die Metadaten-Hierarchien in das Modul „Portal“ integriert werden.

Ansonsten muss das Modul „Metadatenerfassung“ aus der Oberfläche heraus über eine Authentifizierung der erfassenden Nutzer (siehe Kapitel 4.2) aufrufbar sein.

Die Vorgaben des DV-technischen Feinkonzepts PortalU / InGrid 1.0: Styleguide sind sicherzustellen bzw. einzuhalten.

3.3 Mehrsprachigkeit

3.3.1 Oberfläche

Ebenso wie InGrid 1.0, muss auch die Anwendung InGrid 1.1 grundsätzlich mehrsprachig konzipiert und realisiert werden. Die Oberflächen sind zunächst zweisprachig in Deutsch und Englisch zu entwickeln. Der Nutzer muss zwischen beiden Sprachen hin und her schalten können.

Die Mehrsprachigkeit muss insbesondere bei der Darstellung von Auswahllisten berücksichtigt werden. Die ISO-Listen sind zweisprachig (Deutsch/Englisch) in der Datenbank vorhanden.

Weiterhin muss die Mehrsprachigkeit auch bei der Online-Hilfefunktion berücksichtigt werden. Bei der Umschaltung der Oberflächendarstellung von deutsch auf englisch (und umgekehrt), muss entsprechend auch die Sprache der Hilfetexte umgeschaltet werden.

Die Anwendung muss so konzipiert sein, dass Oberfläche, Auswahllisten und Hilfe-seiten durch den Betreiber auf andere Sprachen erweitert werden können, ohne dass struktureller Programmieraufwand notwendig ist.

3.3.2 Dateninhalte / Suche

Die Mehrsprachigkeit der Dateninhalte muss ebenfalls berücksichtigt werden. Vorerst sollen die Sprachen Deutsch und Englisch (in Auswahllisten, etc.) berücksichtigt werden. Eine Erweiterung auf andere Sprachen durch den Betreiber ohne zusätzlichen strukturellen Programmieraufwand ist vorzusehen (siehe oben).

Für die Metadatenerfassung soll sowohl für das Feld „Sprache des Metadatensatzes“ als auch für das Feld „Sprache des Datensatzes“ der Standardeintrag auf die Spracheinstellung der Oberfläche gesetzt werden.

Die Berücksichtigung der Sprache bei der Suchfunktionalität und anderen Komponenten von InGrid bleibt, wie in der Version 1.0 realisiert, erhalten und muss bei allen Erweiterungen berücksichtigt werden.

Für die thesaurusbasierte Suche bzw. die Verschlagwortung von Datenobjekten in Deutsch und Englisch, müssen wie in InGrid 1.0 die vom Umweltbundesamt bereitgestellten Semantic Network Services (SNS) genutzt werden können. Um in Zukunft auch eine Suche und Verschlagwortung in anderen Sprachen zu ermöglichen, muss die Option zur Nutzung multilingualer Thesauri (z.B. GEMET) bzw. entsprechender Erweiterungen des SNS vorgesehen werden.

3.4 Metadaten-Kataloge

Die Anwendung muss so konzipiert sein, dass sie sowohl bei den, die Metadaten erfassenden, Partnern betrieben werden kann, als auch die Metadaten von den Partnern direkt in die Datenbank der zentralen Installation (bei der KUG) erfasst werden können (siehe auch „technische Feinkonzepte InGrid 1.0“). Daher muss die Anwen-

dung InGrid 1.1 in der Lage sein, die Metadaten-Kataloge mehrerer Partner zu verwalten.

Das Verwalten der Metadatenkataloge mehrerer Partner muss ein Konzept für Backup und Recovery umfassen. Bisher ist eine Gesamtsicherung des UDK-Datenbestandes ausreichend. Wenn aber mehrere Partner die zentrale Anwendung für die Erfassung ihrer Daten nutzen, muss z.B. auch ein selektives Zurücksetzen eines einzelnen Kataloges möglich sein.

3.5 UUID (Universal Unique Identifier)

Alle einzelnen Komponenten der Metadaten (Kataloge, Nutzer, Objekte, Adressen, etc.) erhalten einen eindeutigen Schlüssel (UUID) analog dem des UDK 5.0.

3.6 Datenmodell

Das Datenmodell wird von InGrid 1.0 übernommen.

Dabei sind die Erweiterungen des Datenmodells vorzunehmen, die sich aus den Anforderungen in Kapitel 4, wie z.B. Katalog- und Nutzerverwaltung oder Workflow, ergeben. Die Erweiterungen sind so vorzunehmen, dass möglichst keine Änderungen an den anderen Modulen notwendig werden.

Alle Änderungen des Datenmodells sind im dv-technischen Feinkonzept zu spezifizieren und entsprechend zu dokumentieren.

3.7 Aktualität des Index

Die Suchanfragen werden in InGrid 1.0 über einen aus den Metadaten erzeugten Index geführt. Es ist sicherzustellen, dass nach jeder Änderung an den Metadaten (Objekte, Adressen usw.) sowohl bei der Erfassung als auch beim Import eine Indizierung erfolgt und damit die Metadaten für den recherchierenden Nutzer sichtbar werden. Der in Kapitel 4.5 beschriebene Qualitätssicherungsmechanismus ist dabei zu berücksichtigen

4 Anforderungen an das Modul „Metadatenerfassung“

4.1 Katalogverwaltung

Unter einem Katalog wird die Sammlung von Metadaten eines Anbieters (siehe auch Kapitel 2.2.1) verstanden. Kataloge können von einem Systemadministrator eingerichtet, bearbeitet und gelöscht werden.

Die Kataloge bekommen eine räumliche Einheit zugeordnet (Deutschland für Bundeskataloge, das jeweilige Land für Kataloge, die einem Land oder einer Kommune, die in einem Land liegt, zugeordnet werden).¹ Weiterhin wird der Anbieter zugeordnet² (z.B. UBA oder BfN für die Bundeskataloge der entsprechenden Bundesämter).

4.2 Nutzerverwaltung

Da das Modul „Metadatenerfassung“ schreibenden Zugriff auf die Datenbank erhält, kommt der Nutzerverwaltung in InGrid 1.1 zentrale Bedeutung zu.

Die Nutzerverwaltung für die Metadatenerfassung wird über die Erweiterung des Rollen-Konzeptes von InGrid 1.0 (siehe technische Feinkonzepte InGrid 1.0) gesteuert.

Folgende neue Rollen werden eingeführt:

Rolle	Rechte
Katalog-Administrator	<ul style="list-style-type: none"> – Einrichtung jeweils eines Metadaten-Katalogs – Einrichtung von Nutzern und Nutzergruppen und Zuordnung der Rollen Metadaten-Administrator, Metadaten-Autor und Metadaten-Qualitätssicherung – Gesamtkatalog-Management (siehe Kapitel 4.8) – Vergabe und Änderung von Schreibrechten auf alle Metadaten-Objekte und -Adressen des Metadaten-Katalogs an die Metadaten-Administratoren und Metadaten-Autoren – Festlegung der Nutzer bzw. Nutzergruppen für die Metadaten-Qualitätssicherung der einzelnen Metadaten-Objekte und -Adressen – Anlegen und Bearbeiten von allen Metadaten-Objekten und -Adressen des Metadaten-Katalogs

¹ Grundsätzlich ist an dieser Stelle konzeptionell zu berücksichtigen, dass wie bei den Objekten zukünftig auch andere Staaten als Raumbezug eingetragen werden können müssen.

² Die ist neu gegenüber dem UDK.

Metadaten-Administrator	<ul style="list-style-type: none"> – Einrichtung von Nutzern und Nutzergruppen und Zuordnung der Rolle Metadaten-Autor – Vergabe und Änderung von Schreibrechten auf die eigenen Metadaten-Objekte und –Adressen an die Metadaten-Autoren – Anlegen und Bearbeiten von Metadaten-Objekten und -Adressen (soweit Schreibrechte vorhanden)
Metadaten-Autor	<ul style="list-style-type: none"> – Anlegen und Bearbeiten von Metadaten-Objekten und -Adressen (soweit Schreibrechte vorhanden)
Metadaten-Qualitätssicherung	Qualitätssicherung von Metadaten-Objekten (siehe Kapitel 4.5)

Es können sowohl Nutzer als auch Nutzergruppen eingerichtet werden. Jedem Nutzer wird bei der Einrichtung vom jeweiligen Administrator eine Metadaten-Adresse zugeordnet. Jede Nutzergruppe besteht aus einem oder mehreren Nutzern

Die Schreibrechte der einzelnen Nutzer bzw. Nutzergruppen werden an die Metadaten-Objekte bzw. -Adressen gebunden, wobei mehrere Nutzer Schreibrechte auf ein Metadaten-Objekt bzw. eine Adresse haben können³.

Die Rechte für das Einhängen eines Objektes (bzw. einer Adresse) in die Hierarchie wird wie folgt geregelt:

- Das Recht ein Objekt (bzw. eine Adresse) auf der obersten Ebene der Hierarchie anzulegen (also ohne Hierarchie), wird vom Katalog-Administrator vergeben. Es kann an alle, an Gruppen oder an Einzelne vergeben werden.
- Ansonsten kann ein Nutzer ein Objekt (bzw. eine Adresse) nur unterhalb eines Objektes einhängen, auf das er ebenfalls Schreibrechte hat.

Die Rolle der Metadaten-Qualitätssicherung wird ebenfalls an die einzelnen Metadaten-Objekte bzw. -Adressen gebunden.

Alle Rechte vererben sich standardmäßig in der Objekt- bzw. Adresshierarchie. Dies kann aber für einzelne Objekte explizit ausgeschaltet werden.

Die Oberfläche zur Erfassung von Metadaten wird nur den Nutzern zugänglich, die sich mit Nutzernamen und Passwort identifizieren und denen eine der oben genannten Rollen zugeordnet worden ist. Der recherchierende Nutzer bekommt diesen Teil der Gesamtanwendung InGrid 1.1 nicht angezeigt.

4.3 Erfassung von Objekten

Das Layout sowie die technische Umsetzung der Erfassungsmaske für Objekte des Moduls „Metadatenerfassung“ wird grundsätzlich im technischen Feinkonzept spezifiziert.

³ Es ist darauf zu achten, dass die Nutzerverwaltung die Pflege von Adressen durch eine zentrale Stelle erlaubt.

4.3.1 Allgemein

Das Layout der Oberfläche für die Erfassung von Objekten ist dem von InGrid 1.0 anzupassen. Dabei sind die Vorgaben des DV-technischen Feinkonzepts PortalU / InGrid 1.0: Styleguide einzuhalten. Für jede Objektklasse muss eine eigene Erfassungsmaske gestaltet werden, die den klassenspezifischen Fachbezug der Objekte berücksichtigt.

Grundsätzlich müssen alle Felder, die im Windows-UDK 5.0 im Erfassungsfenster angezeigt werden, auch auf der Erfassungsmaske des Moduls „Metadatenerfassung“ von InGrid 1.1 angezeigt werden.

Hierbei sind die Zuordnungen der Datenbankfelder zu Textfeldern sowie einzeiligen und mehrzeiligen Feldern mit Auswahllisten wie im Windows-UDK 5.0 vorzunehmen. Zahlen- und Datumsfelder dürfen nur entsprechende Einträge zulassen.

Ebenso müssen die Gruppierungen von Feldern sowohl auf der Ebene einzelner Tabellen als auch auf der Ebene der inhaltlichen Gruppierung (z.B. Raumbezug, Zeitbezug, etc.) so erhalten bleiben, wie sie in der derzeitigen Erfassungsanwendung zusammengefasst sind⁴.

Es muss gewährleistet sein, dass nur Objekte mit vollständig ausgefüllten Pflichtfeldern abgespeichert werden können⁵.

4.3.2 Assistent zur Unterstützung bei der Erfassung von Objekten

Die Metadaten-Autoren sind durch geeignete Funktionalitäten bei der Eingabe zu unterstützen, wenn Umweltinformationen, die bereits als Webseite im Internet stehen, als Metadaten erfasst werden sollen. Es ist zu prüfen, ob hierzu die bereits vom Modul „Suchmaschine“ automatisch generierten Metadaten genutzt werden können, indem diese in die Erfassungsoberfläche übertragen werden. Andernfalls ist die URL der Webseite an den SNS zu übergeben. Die daraus interaktiv ermittelten Inhalte sind anschließend an die Erfassungsoberfläche zu übergeben.

Im Einzelnen handelt es sich um folgende Informationen, die ermittelt und übergeben werden müssen:

- Titel
- Schlagworte des Umweltthesaurus
- Raumbezug
- Zeitbezug
- Angaben zum Anbieter der Internetseite und zum PortalU-Partner, dem dieser Anbieter zugeordnet ist.
- Abstract für die jeweilige Seite (Der Abstract wird aus den Metatags der Webseite übernommen. Für die Fälle, in denen kein Metatag vorhanden ist, wird

⁴ Ausgenommen hiervon sind Änderungen der Gruppierungen, die ggf. von der AG IUTE oder der KUG bis zum Auftragsbeginn festgelegt worden sind.

⁵ Technisch ist sicherzustellen, dass nach einer temporären Unterbrechung der Arbeiten ein Weiterarbeiten möglich ist.

ein Konzept erstellt, wie ein Abstract erstellt werden kann (z.B. Übernahme des ersten Absatzes der Webseite)

- Verweis auf die durchsuchte Webseite (angegebene URL)

Die konkrete technische Funktionalität der Mechanismen zur Bereitstellung der Metadaten für den Metadaten-Erfasser ist im technischen Feinkonzept zu spezifizieren. Dabei ist zu prüfen, ob es möglich ist, eine Liste mit URLs zu übergeben, die dann ohne Nutzereingaben abgearbeitet werden kann, um so Engpässe bei der SNS-Nutzung umgehen zu können. In diesem Fall muss sichergestellt werden, dass die automatisch generierten Daten über den Workflow (siehe Kapitel 4.5) ausreichend qualitätsgesichert werden.

4.3.3 Hierarchien von Objekten

Die Hierarchie der Objekte wird vom UDK 5.0 übernommen. Der Nutzer kann sein Objekt innerhalb der Hierarchie unterhalb jeden Objektes einhängen, auf das er selbst Schreibrechte besitzt.

Ebenso ist es möglich, Objekte innerhalb der Hierarchie zu verschieben, zu kopieren und zu löschen (Teilbaum).

Die Darstellung der Hierarchie (siehe auch Kapitel 4.6.1) und die Funktionalitäten (verschieben, kopieren und löschen) sind im technischen Feinkonzept zu spezifizieren. Insbesondere sind die Verweisbeziehungen zwischen den Objekten zu berücksichtigen.

4.3.4 Einbindung von Adressen

Die Adressen werden über eine getrennte Funktionalität in das System eingepflegt (siehe Kapitel 4.4). Die Zuordnung der Adressen zu den Objekten erfolgt durch Verlinkung der Objekte mit den Adressen. Dabei wird die Möglichkeit der inhaltlichen Zuordnung der Adressen (Auskunft, Anbieter, etc.) aus dem Windows-UDK 5.0 übernommen.

4.3.5 Thesaurus-Suchbegriffe

Die Thesaurus-Suchbegriffe werden über den SNS geliefert. Es sind immer mindestens drei Pflichtbegriffe pro Objekt zu speichern.

Die Metadaten-Autoren sind durch geeignete Funktionalitäten bei der Eingabe der Thesaurus-Suchbegriffe zu unterstützen. Hierzu muss es möglich sein, sowohl eine Suche nach Thesaurusbegriffen abzusetzen als auch in der Hierarchie des Thesaurus zu „blättern“. Die Suche soll nicht nur die Thesaurus-Deskriptoren sondern auch die Synonyme berücksichtigen. Das heißt, es müssen auch Thesaurus-Synonyme gefunden werden. Die zugeordneten Deskriptoren werden dann dem Erfasser zur Verschlagwortung vorgeschlagen (siehe hierzu auch den Windows-UDK oder den gein[®] Recherche-Assistenten).

Die Möglichkeiten, Thesaurusbegriffe über eine alphabetische Liste einzugeben (angelehnt an den Windows-UDK 5.0), sind zu prüfen und, wenn technisch möglich, zu realisieren.

Zusätzlich zu den Funktionalitäten des „Assistenten zur Unterstützung bei der Erfassung von Objekten“ können die bereits befüllten Textfelder der Metadaten-Objekte in geeigneter Weise ausgelesen und als Text an SNS übergeben werden. Die daraus durch SNS ermittelten Thesaurus-Deskriptoren werden dem Nutzer zur Verschlagwortung vorgeschlagen.

Die konkrete technische Funktionalität der unterschiedlichen Mechanismen zur Bereitstellung der Thesaurus-Suchbegriffe für den Metadaten-Erfasser ist im technischen Feinkonzept zu spezifizieren.

4.3.6 Geografische Metadaten

4.3.6.1 Geografische Namen

Die geografischen Metadaten müssen ebenfalls SNS-unterstützt eingegeben werden können. Hierbei werden die ursprünglich im UDK verwendeten administrativen Einheiten auf das gesamte vom SNS unterstützte geografische Namensgut erweitert⁶. Dieses basiert auf externen Referenzsystemen, die vom BKG und anderen zuständigen Stellen bereitgestellt und gepflegt werden (siehe auch Kap. 4.8.7).

Die Koordinaten der Bounding-Box werden dabei vom SNS geliefert⁷ und parallel zum jeweiligen geografischen Begriff im Datenmodell gespeichert.

Die Eingabefunktionalität ist dabei an der in Kapitel 4.3.5 beschriebenen Funktionalität für Thesaurus-Suchbegriffe zu orientieren.

4.3.6.2 Weitere geografische Metadaten

Neben den von den Referenzsystemen unterstützten geografischen Namen muss es auch weiterhin möglich sein, freie geografische Einträge in den Metadaten zu tätigen. Hierzu muss es (wie im Windows-UDK 5.0) möglich sein, sowohl textuelle Einträge als auch Koordinaten einzugeben.

Die Daten sollen, neben der textuellen Eingabeoption, auch über das Aufziehen eines Rechtecks über einer Karte einzugeben sein. Dazu ist das InGrid-Modul „WMS“ zu nutzen.

Weiterhin müssen die Bounding-Box-Koordinaten auch über ein Eingabefeld eingegeben werden können. Hierbei werden mindestens die Koordinatensysteme Gauß-Krüger, ETRS89 und Lat-Lon/WGS84 unterstützt⁸. Eine Auswahl des Koordinatensystems durch den eingebenden Benutzer muss möglich sein.

⁶ Grundsätzlich ist an dieser Stelle konzeptionell zu berücksichtigen, dass zukünftig auch andere Staaten als Raumbezug eingetragen werden können müssen.

⁷ Zurzeit werden von SNS nur Bounding-Boxen für administrative Einheiten unterstützt. Daher wird InGrid 1.1 zum Zeitpunkt der Fertigstellung auch nur für diese die Bounding-Boxen speichern können. Grundsätzlich soll die Software aber in der Lage sein, bei einer möglichen Erweiterung von SNS auch die Koordinaten von anderen geografischen Gebieten zu speichern.

⁸ Es ist zu prüfen, ob zur Umsetzung dieser Anforderung (Koordinatentransformation) ein allgemeiner Web-Service genutzt werden kann.

4.3.7 Zusätzliche Felder

Im Windows-UDK 5.0 ist es möglich, zusätzliche Felder zu nutzen (siehe Anlage „Software Windows-UDK 5.0“). Diese sind im Datenmodell verankert (siehe Anlage „DV-technisches Feinkonzept PortalU / InGrid 1.0: Datenmodell“) und werden bisher über den UDK-Z definiert. Diese Felder müssen auch zukünftig über das Gesamtkatalog-Management, also durch den Katalog-Administrator, pflegbar sein (siehe Kap. 4.8.6).

Die zusätzlichen Felder müssen in der Erfassungsmaske berücksichtigt werden. Ebenso ist die Recherche-Komponente so anzupassen, dass die Felder in der Detailansicht dargestellt werden.

4.4 Erfassung von Adressen

Das Layout sowie die technische Umsetzung der Erfassungsmaske für Adressen des Moduls „Metadatenerfassung“ wird grundsätzlich im technischen Feinkonzept spezifiziert.

4.4.1 Allgemein

Das Layout der Oberfläche für die Erfassung von Adressen ist dem von InGrid 1.0 anzupassen. Dabei sind die Vorgaben des DV-technischen Feinkonzepts PortalU / InGrid 1.0: Styleguide einzuhalten. Für jede Adressart (siehe UDK 5.0) muss eine eigene Erfassungsmaske gestaltet werden, die die spezifischen Felder der Adressart berücksichtigt.

Grundsätzlich müssen alle Felder, die im Windows-UDK 5.0 im Erfassungsfenster für Adressen angezeigt werden, auch auf der Erfassungsmaske des Moduls „Metadatenerfassung“ von InGrid 1.1 angezeigt werden.

Hierbei sind die Zuordnungen der Datenbankfelder zu Textfeldern sowie einzeiligen und mehrzeiligen Feldern mit Auswahllisten wie im Windows-UDK 5.0 vorzunehmen. Zahlen- und Datumsfelder dürfen nur entsprechende Einträge zulassen.

Ebenso müssen die Gruppierungen von Feldern sowohl auf der Ebene einzelner Tabellen als auch auf der Ebene der inhaltlichen Gruppierung (z.B. Adresse, Verschlagwortung, etc.) so erhalten bleiben, wie sie in der derzeitigen Erfassungsanwendung zusammengefasst sind.

Es muss gewährleistet sein, dass nur Adressen mit vollständig ausgefüllten Pflichtfeldern abgespeichert werden können.

Eine Adresse darf grundsätzlich nur dann gelöscht werden, wenn sie keinem Objekt mehr zugeordnet ist.

4.4.2 Hierarchien von Adressen

Die Möglichkeit Hierarchien von Adressen zu bilden, ist vom UDK 5.0 zu übernehmen. Der Nutzer kann eine Adresse innerhalb der Hierarchie unterhalb einer Adresse einhängen, auf die er selbst Schreibrechte besitzt.

Dabei ist zu berücksichtigen, dass wie im UDK 5.0 eine Adresshierarchie immer mit einer Institution beginnt. Nachgeordnet können Einheiten oder Personen sein. Einer Einheit kann wiederum eine Einheit oder eine Person nachgeordnet sein. Eine Person darf keine nachgeordneten Adressen besitzen. Außerhalb der Hierarchie stehen die „freien Adressen“, die weder übergeordnete noch nachgeordnete Adressen besitzen dürfen.

Es muss möglich sein, Adressen innerhalb der Hierarchie zu verschieben, zu kopieren und zu löschen (Teilbaum). Allerdings darf eine Adresse grundsätzlich nur dann gelöscht werden, wenn sie keinem Objekt zugeordnet ist.

Die Darstellung der Hierarchie (siehe auch Kapitel 4.6.2) und die Funktionalitäten (verschieben, kopieren und löschen) sind im technischen Feinkonzept zu spezifizieren.

4.5 Workflow-Kontrolle für die Erfassung

Im Unterschied zum UDK 5.0 soll InGrid 1.1 über die Möglichkeit einer automatischen Workflow-Kontrolle verfügen. Hierunter wird ein Vorgang verstanden, der eine kontrollierte Erfassung von Metadaten, Qualitätssicherung durch eine zentralen Stelle und anschließende Veröffentlichung vorsieht.

Um den Partnern Rechnung zu tragen, die in ihrer Erfassungskonzeption keinen bzw. einen anders definierten Workflow vorgesehen haben, wird die Workflow-Kontrolle so konzipiert, dass sie für jeden Katalog vom jeweiligen Katalog-Administrator optional eingeschaltet werden kann.

Ist die Workflow-Kontrolle für einen Katalog eingeschaltet gilt folgendes:

- Instanz für die Qualitätssicherung sind diejenigen Nutzer, die die Rolle der Qualitätssicherung für das jeweilige Metadaten-Objekt bzw. die jeweilige Metadaten-Adresse bekommen haben.
- Es wird eine Zeitspanne für den Verfall eines Objektes bzw. einer Adresse festgelegt (hierzu wird das Datenmodell um ein entsprechendes Feld erweitert).
- Für jedes Objekt (bzw. Adresse) gibt es einen zuständigen Metadaten-Autor. Im UDK 5.0 ist dies die Adresse „Geändert von“. In InGrid 1.1 muss hier ein zusätzliches Feld „Metadaten-Autor“ eingeführt werden, das bei der Altdaten-übernahme auf „Geändert von“ gesetzt wird, aber durch den Katalog-Administrator oder die Metadaten-Administratoren geändert werden kann).
- Softwaretechnisch muss sichergestellt werden, dass für jeden Nutzer mit Schreibrechten (z.B. die Rollen „Katalog-Administrator“, „Metadaten-Autor“, etc.) mindestens eine Email-Adresse eingetragen ist.
- Der Standard-Workflow für die Metadateneingabe umfasst folgende Schritte:
 1. Die von den erfassenden Nutzern neu angelegten, gelöschten (gesetztes Delete-Flag) oder geänderten Objekte und Adressen werden in der Datenbank separat gespeichert (einschließlich der Information, wo sie in der Hierarchie eingehängt werden). Sie sind dann nur für Nutzer mit entsprechenden Schreibrechten auf das Objekt bzw. die Adresse und von Nutzern mit

- der Rolle „Metadaten-Qualitätssicherung“ zugänglich. Ein öffentlicher Zugang zu diesen Daten über die Recherchefunktionalitäten ist nicht möglich.
2. Alle Nutzer mit der Rolle „Metadaten-Qualitätssicherung“ werden per Mail benachrichtigt, dass ein Objekt oder eine Adresse angelegt, gelöscht oder geändert wurde.
 3. Ergibt die Qualitätskontrolle Mängel, kann ein qualitätssichernder Nutzer (Rolle: „Metadaten-Qualitätssicherung“) entweder die Änderungen selbst vornehmen (Verwerfen des Anlegens oder Löschens, Änderungen am Objekt oder Adresse) oder einen entsprechenden Kommentar an den Metadaten-Autoren zurücksenden. In letzterem Fall muss dieser die gewünschten Änderungen durchführen und der Workflow beginnt wieder bei Punkt 1. Eine derartige Kommunikation kann beliebig oft wiederholt werden.
 4. Entspricht die Qualität des Objektes oder der Adresse den Kriterien der Qualitätssicherung, gibt die QS das Objekt frei und die beabsichtigte Aktion (Anlegen, Löschen oder Ändern) wird durchgeführt.⁹ Dabei erhält das Objekt oder die Adresse ein neues Verfallsdatum, das dem Tag der Veröffentlichung plus der Verfallszeitspanne entspricht (der Administrator kann ein Standard-Verfallsdatum festlegen, das aber für einzelne Objekte vom freigebenden Nutzer (Rolle: „Metadaten-Qualitätssicherung“) überschrieben werden kann). Das Objekt oder die Adresse ist danach auch für den recherchierenden Nutzer sichtbar.
 5. Ist das Verfallsdatum für ein Objekt, bzw. eine Adresse erreicht, bekommt der zuständige Bearbeiter automatisch eine Mail mit der Aufforderung, das Objekt (die Adresse) zu bearbeiten, sowie einen Link, um es komfortabel direkt im Schreibmodus aufrufen zu können¹⁰. Der qualitätssichernde Nutzer (Rolle: „Metadaten-Qualitätssicherung“) bekommt zur Information die Mail in Kopie.

Um Konflikte beim Bearbeiten von UDK-Objekten durch mehrere Nutzer zu vermeiden, ist sicherzustellen, dass beim Öffnen eines Objektes zum Bearbeiten dieses gegenüber anderen Nutzern zum Bearbeiten gesperrt wird (nur lesender Zugriff).

4.6 Darstellung und Funktion von Hierarchien

Im UDK 5.0 und auch zukünftig in InGrid gibt es drei Arten von Hierarchien, die in InGrid 1.1 dargestellt und von den Metadaten erfassenden Nutzern (Katalog-Administratoren, Metadaten-Administratoren und Metadaten-Autoren) bearbeitet werden müssen. Die unterschiedlichen Hierarchien werden in den folgenden Unterkapiteln diskutiert. Sie sind für die Erfassung der Metadaten zwingend erforderlich, sollen aber auch als spezielle Recherche-Funktionalitäten für Nutzer, die sich nur für Metadaten interessieren, genutzt werden können.

⁹ Werden mehrere Nutzer mit der Qualitätssicherung eines Objektes betraut, so muss organisatorisch die Verantwortung für die Freigabe geklärt werden. Technisch sind alle Nutzer mit der Rolle „Metadaten-Qualitätssicherung“ berechtigt, die Freigabe zu erklären.

¹⁰ Die Autorisierung über Nutzernamen und Passwörter muss an dieser Stelle natürlich berücksichtigt werden.

Alle Darstellungen der Hierarchie von Objekten, Adressen usw. sollen als Portlets in der Portal-Oberfläche von InGrid 1.1 (Portalsystem Jetspeed, siehe DV-technisches Feinkonzept PortalU / InGrid 1.0: Portal) realisiert werden.

4.6.1 Objekt-Hierarchie

Die Hierarchie der Objekte ist als Baumstruktur darzustellen. Dabei ist möglichst nur HTML zu verwenden. BITV-konforme Alternativen sind möglich, aber an InGrid 1.0 zu orientieren.

Das konkrete Layout und die konkrete funktionale Anwendung muss im technischen Feinkonzept spezifiziert werden.

Erfassung

Die Erfassungsmaske für Objekte muss die Möglichkeit bieten, die Objekt-Hierarchie aufzurufen, um neuangelegte Objekte in die Hierarchie einzuhängen oder aber um an den Stellen, an denen Querverweise zwischen Objekten möglich sind, diese zu erzeugen. Nach Auswahl eines Objektes muss dessen ID an die Erfassungsmaske übergeben werden.

Innerhalb der Hierarchie muss es für den erfassenden Nutzer möglich sein, Teilbäume von Objekten zu kopieren (Teilbäume und einzelne Objekte), zu verschieben und zu löschen.

Recherche

Es muss die Struktur der Objekte aller in der Datenbank enthaltenen Kataloge sowie auch die, der über das UdkDB-iPlug (siehe Anlage „DV-technisches Feinkonzept PortalU / InGrid 1.0: iPlug“) oder die XML/SOAP-Schnittstelle (siehe Anlage „DV-technisches Feinkonzept PortalU / InGrid 1.0: Schnittstellen“) angeschlossenen Kataloge angezeigt werden können¹¹. Aus der Struktur heraus muss direkt die Detaildarstellung der Metadaten aufgerufen werden können. Die Funktionalität ist dabei am WWW-UDK 5.1 bzw. der Recherche-Einstellung des Windows-UDK 5.0 zu orientieren.

4.6.2 Adress-Hierarchie

Die Hierarchie der Adressen ist als Baumstruktur darzustellen. Dabei ist möglichst nur HTML zu verwenden. BITV-konforme Alternativen sind möglich, aber an InGrid 1.0 zu orientieren.

Das konkrete Layout und die konkrete funktionale Anwendung muss im technischen Feinkonzept spezifiziert werden.

Erfassung

Die Erfassungsmaske für Adressen muss die Möglichkeit bieten, die Adress-Hierarchie aufzurufen, um neuangelegte Adressen in die Hierarchie einzuhängen. Ebenso muss die Erfassungsmaske für Objekte die Möglichkeit bieten die Adress-

¹¹ Ggf. ist hierzu die Funktionalität des UdkDB-iPlug und der XML/SOAP-Schnittstelle zu erweitern.

Hierarchie aufzurufen, um Adressen den zu bearbeitenden Objekten zuzuordnen. In beiden Fällen muss nach Auswahl einer Adresse ihre ID an die jeweilige Erfassungsmaske übergeben werden.

Innerhalb der Hierarchie muss es für den erfassenden Nutzer möglich sein, Teilbäume von Adressen zu kopieren (Teilbäume und einzelne Adressen), zu verschieben und zu löschen.

Recherche

Es muss die Struktur der Adressen aller in der Datenbank enthaltenen Kataloge sowie auch die der über das UdkDB-iPlug (siehe Anlage „DV-technisches Feinkonzept PortalU / InGrid 1.0: iPlug“) oder die XML/SOAP-Schnittstelle (siehe Anlage „DV-technisches Feinkonzept PortalU / InGrid 1.0: Schnittstellen“) angeschlossenen Kataloge angezeigt werden können¹². Aus der Struktur heraus muss direkt die Detaildarstellung der Metadaten aufgerufen werden können. Die Funktionalität ist dabei an der Recherche-Einstellung des Windows-UDK 5.0 zu orientieren.

4.6.3 Thesaurus-Struktur

Der Thesaurus muss in seiner Struktur dargestellt werden können. Es müssen sowohl die Hierarchie mit Ober- und Unterbegriffen als auch die Synonyme und verwandten Begriffe zu den einzelnen Thesaurus-Deskriptoren angezeigt werden. Neben einer hierarchischen Darstellung, muss auch eine Suche nach Thesaurusbegriffen (bzw. Synonymen) möglich sein.

Die Thesaurus-Deskriptoren sowie zugehörige Synonyme und Verknüpfungen werden über den Semantic Network Service (SNS) des Umweltbundesamtes bezogen. Eine entsprechende Schnittstelle ist im technischen Feinkonzept vorzusehen.

Für die Darstellung ist nur HTML zu verwenden. BITV-konforme Alternativen sind möglich, aber an InGrid 1.0 zu orientieren. Das konkrete Layout und die konkrete funktionale Anwendung muss im technischen Feinkonzept spezifiziert werden.

Erfassung

Die Erfassungsmasken für Objekte und Adressen müssen die Möglichkeit bieten, die Thesaurus-Struktur aufzurufen, um den zu bearbeitenden Objekten und Adressen einen oder gleichzeitig mehrere (bei Objekten mindestens drei) Thesaurusbegriffe zuzuordnen. Nach der Auswahl der Thesaurus-Deskriptoren müssen diese an die jeweilige Erfassungsmaske übergeben werden (siehe auch Kapitel 4.3.5). Die Sprache der Thesaurusbegriffe ist an der „Sprache des Metadatensatzes“ (siehe Kapitel 3.3.2) zu orientieren (Für Sprachen, die nicht von SNS unterstützt werden, ist ein Standard, wie z.B. „Englisch“, vorzusehen).

Recherche

Die oben beschriebene Struktur des Thesaurus soll auch dem recherchierenden Nutzer zur Verfügung gestellt werden. Die Funktionalität muss sich dabei am WWW-UDK 5.1 orientieren. Das heißt, die mit dem jeweils eingestellten Begriff verschlag-

¹² Ggf. ist hierzu die Funktionalität des UdkDB-iPlug und der XML/SOAP-Schnittstelle zu erweitern.

worteten Metadaten-Objekte sind als Ergebnisliste anzuzeigen. Es ist dabei zu berücksichtigen das die Objekte aller in der Datenbank enthaltenen Kataloge sowie auch die der über das UdkDB-iPlug (siehe Anlage „DV-technisches Feinkonzept PortalU / InGrid 1.0: iPlug“) oder die XML/SOAP-Schnittstelle (siehe Anlage „DV-technisches Feinkonzept PortalU / InGrid 1.0: Schnittstellen“) angeschlossenen Kataloge angezeigt werden¹³.

Die Sprache der Thesaurusbegriffe ist an der Sprache der Oberfläche (siehe Kapitel 3.3.1) zu orientieren (Für Sprachen, die nicht von SNS unterstützt werden, ist ein Standard, wie z.B. „Englisch“, vorzusehen).

4.7 Im- und Export – Funktionalität

Die verschiedenen Im- und Export-Funktionalitäten von InGrid 1.1 müssen vom Katalog-Administrator ausgeführt werden können. Die einzelnen Funktionalitäten sind in den folgenden Unterkapiteln beschrieben.

4.7.1 Import aus UDK 5.0

InGrid 1.0 unterstützt bereits den Import aus dem UDK 5.0. Dieser Import ist auch weiterhin von InGrid 1.1 zu unterstützen. Er ist, soweit notwendig, an die Datenbankänderungen anzupassen.

4.7.2 Im- und Export einzelner Kataloge

InGrid 1.1 erhält eine Im- und Exportfunktionalität, die es erlaubt, einzelne Kataloge zu importieren bzw. zu exportieren. Im Gegensatz zum UDK 5.0 ist die Nutzerverwaltung von InGrid 1.1 an den Katalog gebunden und muss beim Im- und Export berücksichtigt werden.

Als Format ist das semantische XML des Im- und Exports des Windows-UDK 5.0 zu verwenden. Die aktuelle XML-DTD muss dafür zu einem XML-Schema (in Anlehnung an das Schema der UDK-Schnittstelle von InGrid 1.0) ausgebaut und an die zusätzlichen Anforderungen von InGrid 1.1 (Datenmodellerweiterungen, katalogbezogene Nutzerverwaltung, etc.) angepasst werden.

Beim Import wird ein ganzer Katalog importiert. Sollte der Katalog bereits vorhanden sein (gleiche Katalog-ID) wird er nach einer Warnung (und Bestätigung durch den Systemadministrator) gelöscht und dann der neue Katalog importiert.

4.7.3 Im- und Export des Gesamtdatenbestandes

Über diese Funktionalität kann der Gesamtbestand an Metadaten in InGrid 1.1 importiert oder exportiert werden. XML-Schema und Funktionsweise der Anwendung entsprechen dem Katalog Im- und Export mit dem Unterschied, dass er sich auf alle Kataloge bezieht.

¹³ Ggf. ist hierzu die Funktionalität des UdkDB-iPlug und der XML/SOAP-Schnittstelle zu erweitern.

Sind beim Import bereits Kataloge vorhanden, werden diese nach einer Warnung (und Bestätigung durch den Systemadministrator) gelöscht und dann der Gesamtimport durchgeführt.

4.7.4 Teilimport und -export

Der Teilimport und -export soll einerseits zum Austausch zwischen verschiedenen Installationen des gleichen Kataloges dienen und andererseits den Import aus bzw. den Export zu externen Systemen ermöglichen.

Der Export orientiert sich grundsätzlich an dem speziellen Teilexport des Windows-UDK 5.0. Das heißt, grundsätzlich sind drei Exportarten möglich:

- Alle Datensätze, denen ein bestimmter Wert im Feld „XML-Export-Kriterium“ zugeordnet ist, werden exportiert. Alle für einen Katalog vergebenen Werte werden in der Export-Funktionalität angezeigt und können ausgewählt werden.
- Ein Teilbaum von Objekten oder Adressen kann ausgewählt werden und wird dann exportiert.
- Ein einzelnes Objekt oder eine einzelne Adresse kann ausgewählt werden und wird dann exportiert.
- Alle Objekte, die seit dem letzten Export geändert wurden, werden exportiert (inkrementeller Export)¹⁴.

Der Import orientiert sich grundsätzlich am speziellen Import des Windows-UDK 5.0. Das heißt insbesondere:

- Es können beliebig viele Objekte und Adressen importiert werden. Dabei wird das in Kapitel 4.7.2 beschriebene XML-Schema benutzt.
- Beim Import aus externen Systemen wird eine Original-ID mitgegeben, die ebenfalls im System zu speichern ist.
- Es ist möglich, Hierarchien von Objekten und Adressen zu importieren, indem auch die Informationen über die Hierarchiebeziehungen der Objekte bzw. der Adressen anhand ihrer ID's oder ihrer Original-ID's importiert werden.
- Neue Objekte werden neu angelegt. Wo keine Hierarchieangabe über den Import geliefert wird, wird interaktiv nachgefragt, an welcher Stelle der Hierarchie das Objekt bzw. die Adresse eingehängt werden soll.
- Vorhandene Objekte oder Adressen werden anhand ihrer ID bzw. ihrer Original-ID identifiziert und ersetzt. Der importierende Nutzer muss vor dem Überschreiben von Objekten oder Adressen um eine Bestätigung des Vorganges gebeten werden (optional ausschaltbar für das Update größerer Datenmengen).

¹⁴ Im technischen Feinkonzept ist hierzu ein gegenüber dem UDK 5.0 neues Konzept der Berücksichtigung von Änderungen an Objekten und Adressen zu entwickeln.

4.7.5 Datenimport von ESRI ArcGIS Metadaten

ArcGIS bietet über seinen Metadaten-Editor Funktionalitäten zur Erzeugung und Pflege von beschreibenden Informationen für die vorgehaltenen Geo-Objekte. Da die ESRI-Produktfamilie in öffentlichen Fachverwaltungen mittlerweile weit verbreitet ist und mit diesem Werkzeug Metadaten für umfangreiche Geodatenansammlungen erstellt werden, soll die Möglichkeit geschaffen werden, die Metadaten der Arc-Geo-Objekte in die Metadatenbank von InGrid 1.1 einzupflegen. Sie soll unabhängig davon funktionieren, ob die ArcGIS-Metadaten lokal als XML-Dateien, in einer Geodatabase oder zentral über ArcSDE in einer relationalen Datenbank verwaltet werden.

Dabei bezieht sich dieses Einpflegen ausschließlich auf die (UDK-) Objektklasse „Geoinformation/Karte“ und soll sowohl das Anlegen neuer als auch die Fortführung (im Sinne von Ersetzen durch eine neuere Version) bestehender Metadatenobjekte unterstützen.

Sofern für das korrekte Neu-Anlegen bzw. das Fortführen von Metadatenobjekten in der InGrid-Datenbank weitere zusätzliche Angaben erforderlich sind, wie z.B. an welcher Stelle ein neues Objekt im Strukturbaum einzuhängen ist bzw. welches Objekt fortzuführen ist, so sollen diese Informationen von der Benutzerin bzw. dem Benutzer abgefragt und verifiziert werden können.

Bei einem erneutem Import bereits in InGrid 1.1 vorhandener Metadaten aus ArcGIS werden die ESRI-spezifischen Daten des bestehenden InGrid-Objektes überschrieben. Das Objekt wird anhand seiner Original-ID identifiziert. Die für das Objekt in InGrid 1.1 bereits bestehenden, mittels InGrid-Tools eingepflegten Adressinformationen, Thesaurusbegriffe und Strukturinformationen müssen beim Import erhalten bleiben.

ArcGIS unterstützt Standards wie den CSDGM des FGDC und die ISO 19115, setzt aber auf einen sehr umfangreichen eigenen Attributsatz auf. Um die Arc-Metadaten in die Metadatenbank von InGrid 1.1 zu überführen, wird die InGrid XML/SOAP-Import-Schnittstelle (ehemals XML/SOAP Schnittstelle des UDK, vgl. Kap. 4.7.4) verwendet, über die auch UDK-Daten eingepflegt werden. Daher ist für ArcGIS ein XSLT-Stylesheet zu entwickeln, das aus den Arc-Metadaten eine XML-Datei generiert, die sich gegen das ISO19115-konforme, XML-Schema der InGrid 1.0 Importschnittstelle validieren lässt. Die Transformation soll aus ArcGIS „per Knopfdruck“ erfolgen und muss sicherstellen, dass alle ISO-Kerndatenfelder vom Bearbeiter befüllt wurden. Sollten Pflichtfelder nicht ausgefüllt sein, wird der Bearbeiter aufgefordert, die fehlenden Einträge zu ergänzen. Erst danach ist ein XML-Export möglich.

Das Mapping der ESRI-Metadaten-Attribute auf die UDK-Attribute wird im technischen Feinkonzept in enger Absprache mit der KUG realisiert und ausführlich dokumentiert. Die Frage, welche über das ISO19115-Kernmodell hinausgehenden ESRI-Attribute in InGrid abgebildet werden sollen bzw. können, wird im technischen Feinkonzept diskutiert und mit der KUG abgesprochen. Es ist weiterhin zu klären, wie die Objekt-Adress-Relationen des UDK-Datenmodells unterstützt werden können.

Importierte Metadaten, die im Sinne des ISO19115 Kerndatensatzes ausreichend dokumentiert sind, stehen nach dem Import sofort einer Indizierung, Editierung und Recherche zu Verfügung.

Eine Dokumentation zur Integration des Programmmodules zum InGrid-Metadatenexport in die Menüoberfläche von ArcGIS ist zu erstellen.

4.7.6 Export einzelner Objekte über eine URL

Der Export eines einzelnen Objekts bzw. einer einzelnen Adresse muss auch über eine URL möglich sein, die mit Hilfe der eindeutigen Objekt-ID parametrisiert wird. Das Objekt bzw. die Adresse wird dann als vollständige Objektinformation exportiert, wie bereits heute im WWW-UDK 5.0 möglich.

4.7.7 „Harvesting“ von InGrid Metadaten

Die bereits in InGrid 1.0 existierende CSW 2.0 Schnittstelle wird um die Funktionalitäten aus dem optionalen CSW Manager Interface erweitert (OpenGIS® Catalogue Services Specification 2.0 - ISO19115/ISO19119 Application Profile for CSW 2.0). Sowohl die Schnittstelle am iBus als auch das CSW-iPlug werden in die Lage versetzt, CSW2.0-Applikationsprofil-konforme harvest-operation-requests weiter zu reichen, bzw. zu beantworten. Synchrones und asynchrones Antworten ist zu ermöglichen.

4.8 Gesamtkatalog-Management

Das Gesamtkatalog-Management orientiert sich am UDK-Z. Es umfasst eine Reihe von Funktionen, die in den folgenden Unterkapiteln beschrieben werden. Diese Funktionen beziehen sich immer auf einen ganzen Katalog und können daher nur vom Katalogadministrator (Rolle: „Katalog-Administrator“) ausgeführt werden.

4.8.1 Katalog-Statistik

Es werden folgende statistische Aussagen ausgegeben (diese Funktionalität kann von allen Nutzern eines Kataloges durchgeführt werden):

- Anzahl der im Katalog erfassten Objekte und Adressen, einschließlich der Angaben wie viele Objekte den jeweiligen Objektklassen bzw. wie viele Adressen den einzelnen Adressarten zugeordnet sind.
- Anzahl der in einem Teilbaum erfassten Objekte bzw. Adressen (einschließlich Objektklassen bzw. Adressarten, siehe oben).
- Häufigkeit der verwendeten Thesaurus- und freien Suchbegriffe (siehe UDK 5.0).

Der Nutzer erhält die Möglichkeit des Ausdrucks der Statistiken.

4.8.2 Analyse-Funktionen

Folgende Analyse-Funktionen werden realisiert:

- Datenbestand-Analyse: Die Analyse-Funktionen des UDK-Z 5.0 beziehen sich auf datenmodellinterne Inkonsistenzen. Sie müssen im technischen Feinkonzept an das physikalische Datenmodell von InGrid 1.1 angepasst werden, mit dem Ziel eine Funktionalität zu schaffen, alle im Datenbestand möglichen Inkonsistenzen zu analysieren und Möglichkeiten zu deren Behebung zu bieten.

- Suchbegriffanalyse: Die freien Suchbegriffe werden daraufhin kontrolliert, ob (z.B. durch Thesaurusänderungen) in ihnen Thesaurus-Suchbegriffe enthalten sind. Diese werden in Thesaurus-Suchbegriffe überführt. Sind in den freien Suchbegriffen Synonyme enthalten, so werden alle ihnen zugeordneten Thesaurusbegriffe zusätzlich als Thesaurus-Suchbegriffe gespeichert. Die Thesaurusbegriffe werden analysiert. Sind Begriffe gespeichert die (z.B. durch Thesaurusänderungen) keine Thesaurusbegriffe mehr sind, werden diese in freie Suchbegriffe überführt.
- Dubletten-Prüfung: Es werden Prüfungen des Objektnamens nach Dubletten durchgeführt. Es wird die Möglichkeit der Änderung des Objektnamens einer gefundenen Dublette geschaffen.
- Die im Katalog vorhandenen URL's werden auf Existenz überprüft und können ggf. gelöscht oder bearbeitet werden.

4.8.3 Pflege von Auswahllisten

Alle in der Erfassungsmaske für Objekte und Adressen existierenden Auswahllisten müssen zentral vom Katalogadministrator gepflegt werden können. Dies gilt auch für die Pflege der ISO-Code-Listen.

4.8.4 Im- und Export von Auswahllisten und anderen Kataloginformationen

Die Auswahllisten und ggf. andere Informationen (z.B. Sprachinformationen), die den Inhalt eines Kataloges bestimmen, müssen über diese Funktionalität im- und exportiert werden können.

4.8.5 Absolutes Löschen von Adressen

Adressen dürfen normalerweise vom Nutzer nur dann gelöscht werden, wenn sie keinem Objekt mehr zugeordnet sind. Mit dieser Funktionalität erhält der Katalog-Administrator das Recht, eine Adresse zu löschen, die noch einigen Objekten zugeordnet ist. Dazu muss er angeben, durch welche andere Adresse diese Zuordnungen ersetzt werden sollen.

4.8.6 Definition zusätzlicher Felder

Wie im UDK-Z muss es auch in InGrid 1.1 möglich sein, zusätzliche Felder frei zu definieren, wobei es sich wie bisher um Textfelder mit einer Länge von 1 – 255 Zeichen handeln soll, die entweder mit oder ohne pflegbare Auswahlliste erstellt werden können. Die Auswahllisten müssen wie unter 4.8.3 beschrieben gepflegt werden können.

4.8.7 Updatefunktionalität für Verwaltungsgrenzen

4.8.7.1 Hintergrund

Das in den geografischen Metadaten verwendete geografische Namensgut basiert auf einem externen Referenzbestand, der vom BKG und anderen zuständigen Stellen bereitgestellt und gepflegt wird. Die Einbindung in die Anwendung InGrid soll über SNS erfolgen (vgl. Kap. 4.3.6).

Zur näheren Abstimmung der Zuständigkeiten, Schnittstellen und konzeptionellen Planungen wurde eine Arbeitsgruppe Geo-Thesaurus (AG Geo-Thesaurus) aus Vertretern des UBA, dem BKG, dem BfN und der KUG eingerichtet.

Zusätzlich zu dem geografischen Namensgut, sind in InGrid und PortalU die Verwaltungsgrenzen auf Land, Kreis und Gemeindeebene im WMS hinterlegt. Sie werden in InGrid 1.1 neben der Suche auch für die Zuordnung des Raumbezugs zu den Objekten verwandt (vgl. Kap. 4.3.6.2).

Derzeit basieren die Verwaltungsgrenzen im WinUDK auf dem VG 1000 (Verwaltungsgrenzen Gemeinden) Modell des BKG, Stand 01.01.2000. Dieser Basisdatensatz wurde vor der Übernahme in den WinUDK an verschiedenen Stellen überarbeitet, z.B. in der Region Hannover und in Hamburg (Bezirke/Stadtteile).

4.8.7.2 Update von Verwaltungsgrenzen / Aktualisierung betroffener Objekte

Da auch Verwaltungsgrenzen einer gewissen Dynamik unterliegen soll in InGrid 1.1 eine Möglichkeit geschaffen werden, den Polygondatensatz der Verwaltungseinheiten entsprechend aufzudatieren und die nötigen Anpassungen in den Metadatenbanken durchzuführen. Dazu müssen nach einem Update Objekte und Adressen identifiziert werden, die über IDs verfügen, die in der aktuellen Version der Verwaltungseinheiten nicht mehr vorhanden sind.

In enger Abstimmung mit der AG Geo-Thesaurus wird geprüft, ob ein Automatismus zur Zuordnung von älteren Versionen von Verwaltungseinheiten zu aktuellen geschaffen werden kann, der nach einem Update durchgeführt werden kann. Wird eine automatisierte Lösung gefunden, ist diese zu realisieren.

Sollte eine automatische Zuordnung nicht möglich sein, wird ein manuelles Verfahren zur Bereinigung realisiert, das z.B. wie folgt aussehen kann: Betroffene Objekte und Adressen werden aus dem aktiven Datenbestand entfernt und zwischengespeichert. Die Metadatenautoren, bzw. bei aktiviertem QS-Workflow die entsprechenden Nutzer mit der Rolle „Metadaten-Qualitätssicherung“, werden informiert und aufgefordert, den administrativen Bezug zu aktualisieren. Nach der Überarbeitung werden die Objekte wieder in den aktiven Datenbestand übernommen.

4.8.7.3 Zusätzliche Hierarchiestufen von Verwaltungsgrenzen

Der Update-Funktionalität kommt auch wegen der geplanten Einbindung von Kommunen eine große Bedeutung zu. So ist davon auszugehen, dass zukünftig Partner Funktionalitäten benötigen, um Verwaltungsstrukturen flexibel an ihre Bedürfnisse anzupassen (wie zum Beispiel im Falle der Hamburger Bezirks-/Stadtteilgrenzen bereits geschehen). Die aktuelle vierstufige Struktur (Bundesland, Regierungsbezirk,

Kreis, Gemeinde) genügt nicht den zu erwartenden Anforderungen an Flexibilität und Anpassbarkeit.

Konzeptionell muss daher eine erweiterbare Hierarchie insbesondere im Raumbezug der Objekte (sowohl nach oben, z.B. ganz Deutschland/Europa, als auch nach unten, z.B. Stadtteile) vorgesehen und implementiert werden.

4.8.7.4 Automatisches Update dezentraler Installationen

Verwaltungsgrenzen-Updates werden von der KUG in die Zentralinstallation PortalU eingepflegt¹⁵ und dann über den p2p-Mechanismus, analog zur Exportfunktion der Systemdaten im WinUDK, an die dezentral angeschlossenen Metadatenkataloge weitergereicht. Die dezentralen Installationen erkennen anschließend die Änderungen am Verwaltungsgrenzenmodell und starten den oben für die Zentralinstallation beschriebenen Mechanismus zur Anpassung der lokalen Datenbankobjekte.

Bei der technischen Konzeption der Updatefunktionalität sind die konzeptionellen Planungen der "Geo-Arbeitsgruppe" zu berücksichtigen.

Die entsprechenden manuellen oder automatischen Funktionalitäten zum Update der Verwaltungsgrenzen sind zu realisieren.

¹⁵ Es ist zu prüfen, ob ein Update der Verwaltungsgrenzen über einen Service bereitgestellt werden kann (z.B. durch das BKG), ansonsten ist an dieser Stelle eine Import-Funktionalität vorzusehen, z.B. unter Verwendung des ESRI Shapefile-Formats.