



Vom Umweltdatenkatalog zum InGrid® Data Catalog

Aktuelle Anpassungen des UDK-Datenmodells und der
Software InGrid® im Hinblick auf INSPIRE

Martin Klenke, Fred Kruse, Stefanie Uhrich, Christiane Giffei, Sybille Peters



Koordinierungsstelle
Portal U



- **INSPIRE und PortalU[®]**
- **Anpassungen am UDK-Datenmodell**
- **GDI-DE / INSPIRE Metadaten Infrastruktur**



- 15.05.2007: Richtlinie 2007/2/EG zur Schaffung einer Geodateninfrastruktur in der Europäischen Gemeinschaft (INSPIRE, *Infrastructure for Spatial Information in Europe*)

- INSPIRE ...
 - ... soll die grenzübergreifende Nutzung von Geodaten in Europa erleichtern
 - ... fordert webbasierte Online-Dienste für die Suche, die Visualisierung und Download der Daten
 - ... muss innerhalb von zwei Jahren jeweils in nationales Recht umgesetzt werden

- Die technischen Einzelheiten zur Umsetzung der Richtlinie werden über Durchführungsbestimmungen (Implementing Rules/IRs) geregelt, die derzeit erarbeitet werden (Metadata, Network Services, Data Specifications, etc.)

- INSPIRE Durchführungsbestimmungen Metadaten, Status:
 - 14.05.2008: Regelungsausschuß "INSPIRE Committee" stimmt dem Draft einstimmig und ohne Enthaltungen zu
 - Votum wird von der EC ans EP weiter geleitet, EP hat Rückäußerungsfrist von vier Wochen
 - Abschließend werden die IRs dem EU-Kabinettt vorgelegt (Veröffentlichung ~Herbst 2008 ?)
 - Nach Veröffentlichung/In-Kraft-treten sind die IRs unmittelbar geltendes Recht in den Mitgliedsstaaten
 - Ab spätestens 2010 (Annex I & II) / 2013 (Annex III) müssen nationale Metadaten zur Verfügung stehen



- UDK-Datenmodell ist bislang ISO 19115- / ISO 19119-konform
 - UDK-Klasse „Geoinformation/Karte“: ISO 19115
 - UDK-Klasse „Service/Dienstleistung/Informationssystem: ISO 19119
 - UDK wird daher von INSPIRE direkt betroffen sein
- Koordinierungsstelle arbeitete als “Legally Mandated Organisation“ im Drafting Team Metadata mit
- Zielsetzungen:
 - Zusatzaufwände für den UDK gering halten
 - ISO 19115/19119 im INSPIRE-Prozess verankern
 - Möglichst keine Erweiterung des INSPIRE-Metadatenmodells gegenüber den ISOs 19115/19119
 - Möglichst zeitnahe Umsetzung der INSPIRE-Anforderungen
- Draft Implementing Rules Metadata, Version 4 -> UDK/IDC-Umsetzungskonzept
- Andere ISO-konforme Kataloge sind wahrscheinlich in ähnlicher Weise betroffen



IR Metadata: 2.2.7 *Conformity* – „Konformität“

- Neue dreiteilige 1:n Pflichtfeldgruppe
 - Spezifikation der Konformität (konform zu welchen INSPIRE IRs, ggf. auch anderen Spezifikationen ?)
 - Referenzdatum (Publikation, letzte Überarbeitung, Erstellung)
 - Grad der Konformität (Codeliste: konform / nicht konform / nicht evaluiert)

- Altdatenübernahme:
 - Spezifikation: “INSPIRE-Richtlinie”
 - Grad der Konformität: “nicht evaluiert”
 - Zukünftig evt. Codelistenunterstützung des Freitextfeldes “Spezifikation der Konformität”



2.2.1.5 *Unique resource identifier* – „Eindeutiger Identifikator der Datenquelle“

● Neues Pflichtfeld

- Inhalt: Eindeutiger Identifikator des beschriebenen Datensatzes (!)
- Inhaltliche Ausgestaltung noch nicht klar, soll in den Durchführungsbestimmungen zu den „Data Specifications“ geklärt werden

● Altdatenübernahme:

- Neues einzeliges Pflichtfeld in der UDK/IDC-Klasse „Geoinformation/Karte“
- Befüllung: <Katalogkürzel:<UUID des Metadatensatzes>



2.2.8 *Constraint related to access and use* – „Zugangs- und Nutzungsbestimmungen“

- Zweispaltige 1:n Tabelle mit den Köpfen „Zugangsbeschränkungen“ und „Nutzungsbedingungen“
- Ein Eintrag ist verpflichtend, beide Felder müssen befüllt sein
- 2.2.8.2 Limitations on public access (Zugangsbeschränkungen)
 - bezieht sich auf die Auflistung der in Artikel 13 der INSPIRE-Richtlinie erlaubten Beschränkungen
 - Fest vorgegebene Auswahlliste:
 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
 - Default: “1. Keine”



2.2.8 Constraint related to access and use – „Zugangs- und Nutzungsbest.“ (Fortsetzung...)

- 2.2.8.1 Conditions applying to access and use (Nutzungsbedingungen)
 - Bedingungen zur Nutzung des beschriebenen Datensatzes bzw. des Services, z.B. Kosten
 - Pro Zeile der Tabelle Zugangsbeschränkungen und Nutzungsbedingungen ist ein Eintrag verpflichtend
 - Bei frei nutzbaren Daten/Services soll laut Durchführungsbestimmungen *Keine Einschränkungen* eingetragen werden
 - Sind die Bedingungen unbekannt, ist *Unbekannte Nutzungsbedingungen* einzutragen

- Altdatenübernahme:
 - UDK-Felder *Nutzungsanmerkungen* und *Kosten* in der UDK-Rubrik *Verfügbarkeit* entfallen
 - Deren aktuelle Inhalte werden wie folgt verarbeitet:
 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.



2.2.6.1 Lineage – „Fachliche Grundlage“

- Angaben zur Entstehungsgeschichte sowie zur Qualität des beschriebenen Datensatzes
- Wird zum Pflichtfeld für Geodaten
- Altdatenübernahme:
 - Bislang optionales UDK-Feld “Fachliche Grundlage” wird zum Pflichtfeld für die Klasse Geoinformation/Karte
 - Falls in existierenden Datensätzen unbefüllt automatischer Eintrag: *Unbekannt*



2.2.2.2 Spatial data service type – „Service-Klassifikation“

- Pflichtfeld für Dienste, Befüllung mit geschlossener Code-Liste:

- | | |
|----------------------------------|-------------------------------|
| 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] |

- Altdatenübernahme:

- “Servicetyp” ist UDK-Pflichtfeld für Objekte der Klasse “Service/Dienstleistung/Informationssystem”
- Wird überführt in das neue Pflichtfeld “Service-Klassifikation” (Codelistenwechsel)
- Die bisherigen Einträge werden wie folgt umgesetzt:

Bisheriger Eintrag	Neue ID (siehe oben)
CSW	1
WMS	2
WFS	3
Alle anderen Einträge (incl. leeres Feld)	6



2.2.3 Keyword, Annex B.4 – „Typ des Dienstes“

- Wird für INSPIRE *Spatial Data Services* verpflichtend
- Befüllung durch umfangreiche (~70 Einträge) feste Codeliste
- Altdatenübernahme:
 - Umsetzung als einspaltige Tabelle in der Klasse “Service/Dienstleistung/Informationssystem”
 - Erweiterung der durch INSPIRE vorgegebenen Auswahlliste um die ID=901: *Non geographic service*
 - Automatische Befüllung anhand der migrierten *Serviceklassifikation* mit folgender Tabelle:

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



2.2.6.2 *Spatial resolution* – „Erstellungsmaßstab“

- Bereits im Fachbezug *Geoinformation/Karte* vorhanden
- Zusätzliche Einführung in den Fachbezug der Klasse *Dienst/Anwendung/Informationssystem*
- Einträge optional

2.8 *Metadata point of contact* – „E-Mail-Adresse“

- E-Mail-Adresse wird für alle UDK-Adressen zum Pflichtfeld (bisher optional), die bislang verpflichtende Postadresse wird optional
- Altdatenübernahme:
 - Ist für einen Metadatensatz keine E-Mailadresse vorhanden, wird in der UDK-Adress-Hierarchie aufwärts gewandert um eine E-Mailadresse zu finden und automatisch einzutragen
 - Sollte im gesamten Baum keine gefunden werden, wird die Mailadresse des Katalogadministrators verwendet
 - Vor der Altdatenübernahme sollten die Kataloge dahingehend überarbeitet werden, dass zumindest jede Institution über eine E-Mail-Adresse verfügt



2.9 *Keyword* – „Verschlagwortung“

- UDK: Bisher drei Thesaurusschlagworte aus dem UMTHEs-Wortgut des UBA verpflichtend
- INSPIRE: Angabe eines GEMET- (General Multilingual Environmental Thesaurus) Schlagworts wird verpflichtend
- Der deutschsprachige GEMET-Teil ist komplett im UMTHEs abgebildet. Zukünftig werden Bearbeiter weiterhin mindestens drei UMTHEs-Schlagworte und, falls sich unter diesen kein GEMET-Deskriptor befindet, ein GEMET-Schlagwort erfassen müssen
- Altdatenübernahme:
 - Automatische Prüfung, ob sich unter den existierenden UMTHEs-Schlagworten ein GEMET-Deskriptor befindet
 - Falls nicht „aufwärtswandern“ der UMTHEs-Hierarchie, bis ein GEMET-Schlagwort gefunden wird
- InGrid 2.0 Metadatenerfassungs-Komponente:
 - UMTHEs/GEMET-Bereitstellung durch SNS-Webservice: <http://www.semantic-network.de>
 - Unterstützung der Verschlagwortung durch Thesaurus-Navigator/-Suche und Verschlagwortungsassistenten
 - Demo Thesaurus-Navigator/-Suche: [MDEK Demo](#)



GDI-DE Modellprojekt *Geodatenkatalog-DE*

- Aufbau einer „INSPIRE-gerechten“ Topologie existierender Katalogdienste für Deutschland
- Bündelung im Sinne einer zentralen Abfragemöglichkeit der verteilten Dienste (Prototyp eines deutschen „INSPIRE-Knoten“)
- Dienst-Provider:
 - Verschiedene fachübergreifende Geodatenkataloge der Länder
 - Geodatenkatalog des Bundes
 - Die länderübergreifenden Fachkataloge PortalU® und Adv-MIS
- Nutzer:
 - Joint Research Center EC / ISPRA
 - GIW-Kommission (Kommission für Geoinformationswirtschaft des Bundesministeriums für Wirtschaft und Technologie)



Anforderungen seitens GDI-DE und INSPIRE (I)

- GDI-DE Spezifikation für Katalogschnittstellen (lt. GDI-DE Architekturkonzept): OpenGIS® Catalogue Services Specification 2.0.2 - ISO Metadata Application Profile (1.0.0)
- Wird voraussichtlich auch der fachliche Katalogdienststandard für INSPIRE (aktueller Stand im DT *Network Services*)
- Technisch-funktionale Anforderungen: *Draft Implementing Rules for Discovery and View Services*, z.B.:
 1. Maximale Zeit zum Senden der Rückmeldung nach einer Suchanfrage (in 90% der Fälle, bei 250 gesendeten Metadatensätzen): **3 Sekunden**
 2. Minimum möglicher gleichzeitiger Anfragen an den Service: **100**

-> *Ohne Harvesting- und Indexierungsmechanismen kaum zu realisieren*



Anforderungen seitens GDI-DE und INSPIRE (II)

- Suchkriterien [Artikel 11(2) der INSPIRE-Richtlinie] - Kombinationen folgender Suchkriterien sind 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
- > *Das CSW AP ISO 1.0 kann diese Anforderungen im Grunde nicht direkt erfüllen, da z.B. die Konformität (s.o., Punkt 4) nicht abgebildet ist. Ähnliches gilt für die Nutzungsvereinbarungen*
- Die weitere Entwicklung bleibt abzuwarten: *Draft Guidelines – INSPIRE metadata implementing rules based on ISO 19115 and ISO 19119*
- Aktueller Stand: Es soll keine Erweiterung („INSPIRE-Profil“) des CSW AP ISO 1.0 geben, d.h. alle Anforderungen werden durch Mappings von ISO 19115/19119-Klassen erfüllt



Anforderungen seitens GDI-DE und INSPIRE (III)

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

- Einschätzung: *Als generelle Vorgabe für alle Katalogdienste nicht sinnvoll, da*
 - kostenintensiv
 - Datensicherheitsaspekte
 - fachliche Anforderungen erfordern keinen CSWT, zumindest an „nieder-topologen“ Diensten

- Forderung (Consolidation): *CSWT-Operationen wieder auf optional zurück stufen*

- Weitere Entwicklung bleibt abzuwarten ...



Topologie der deutschen Katalogdienste

- Grundsätzliche Konzepte:
 - fachbezogener Aufbau (z.B. AdV MIS, PortalU®)
 - administrativ (z.B. Kommune -> Land -> Bund)
 - keine mehrstufige Topologie, direkter Anschluss aller Dienste an den Bundesknoten, ggf. registry-unterstützt)
- Alle drei Formen existieren nebeneinander und müssen berücksichtigt werden
- Ohne Harvesting- und Indexierungsmechanismen wird eine INSPIRE-konforme Implementierung des deutschen INSPIRE-Knotens nicht möglich sein (Performanzanforderungen)
- Entsprechende Funktionalität wird derzeit im GDI-DE Modellprojekt Geodatenkatalog-DE vorbereitet
- Für InGrid® befinden sich Weiterentwicklungsideen im Konzeptstadium



Harvesting und Indexierung

- Harvesting I: `GetRecords` (*ElementSetName=brief*) mit einer Wildcard-Query
- Harvesting II: Herunterladen der Metadaten (`GetRecordById`, *ElementSetName=full*), verspeichern in XML-Datei
- Jede CSW-Datenquelle erhält einen eigenen Harvest-Bestand (Bessere Skalierbarkeit, Lastverteilung, Konfigurationsflexibilität)
- Indexierung (Apache Lucene): Mit Indexfeld-zu-XPath-Abbildungsregel auf der CSW-XML-Datei
- Konfigurationsmechanismus zum unaufwendigen Austauschen der Mapping-Klassen
- Abfrage der originären Inhalte der XML-Datei (*ElementSetName=original*) bleibt möglich
- Zum Anfragezeitpunkt erfolgt die Antwort standardmäßig performant direkt aus dem Index, die Quell-Datenbank muss nicht erneut angefragt werden



„Proprietäre“ InGrid-Erweiterungen

- Dublettenproblematik (Harvesting kaskadierender Kataloge) auflösen
 - Relevant für CSW-, Portal- und OpenSearch-Schnittstelle
 - Berücksichtigung Metadaten-UUID / Aktualitäts-Zeitstempel
 - Implementierung der Funktionalität am iBus und den Datenquellenadaptern (iPlugs)
- Am iBus wird ein Mechanismus implementiert, der die Metadaten der Kataloge (Partner, Provider, etc.) unabhängig von CSW-Funktionalitäten verarbeiten kann. Auf Brokerseite kann damit z.B. entschieden werden, welche an PortalU angeschlossenen Datenquellen über die CSW-Schnittstelle „nach oben“ transportiert werden
- Weitergabe aller angeschlossenen ISO-konformen Metadatenquellen via CSW-Interface ermöglichen



Zusammenfassung und Ausblick

- Die INSPIRE-Durchführungsbestimmungen gewinnen aktuell an Kontur
- Die relevanten existierenden deutschen Katalogsysteme erscheinen fachlich, durch ihre Unterstützung der ISO-Standards 19115 und 19119, zukunftsfähig aufgestellt
- Auch technisch haben sich mittlerweile die meisten Katalogbetreiber bzw. Softwarehersteller zur Unterstützung von CSW AP ISO 1.0 „committed“
- Hinsichtlich der praktischen Umsetzung der IRs, insbesondere bezüglich Dienste-Topologie und der Rolle von Harvesting und Indexierungsprozessen, bestehen demgegenüber noch erhebliche Unsicherheiten
- Zur Klärung der offenen Fragen sind Pilotprojekte mit möglichst umfangreichen praktischen Anteilen erforderlich



<http://www.portalu.de/>

<http://www.kst.portalu.de/>

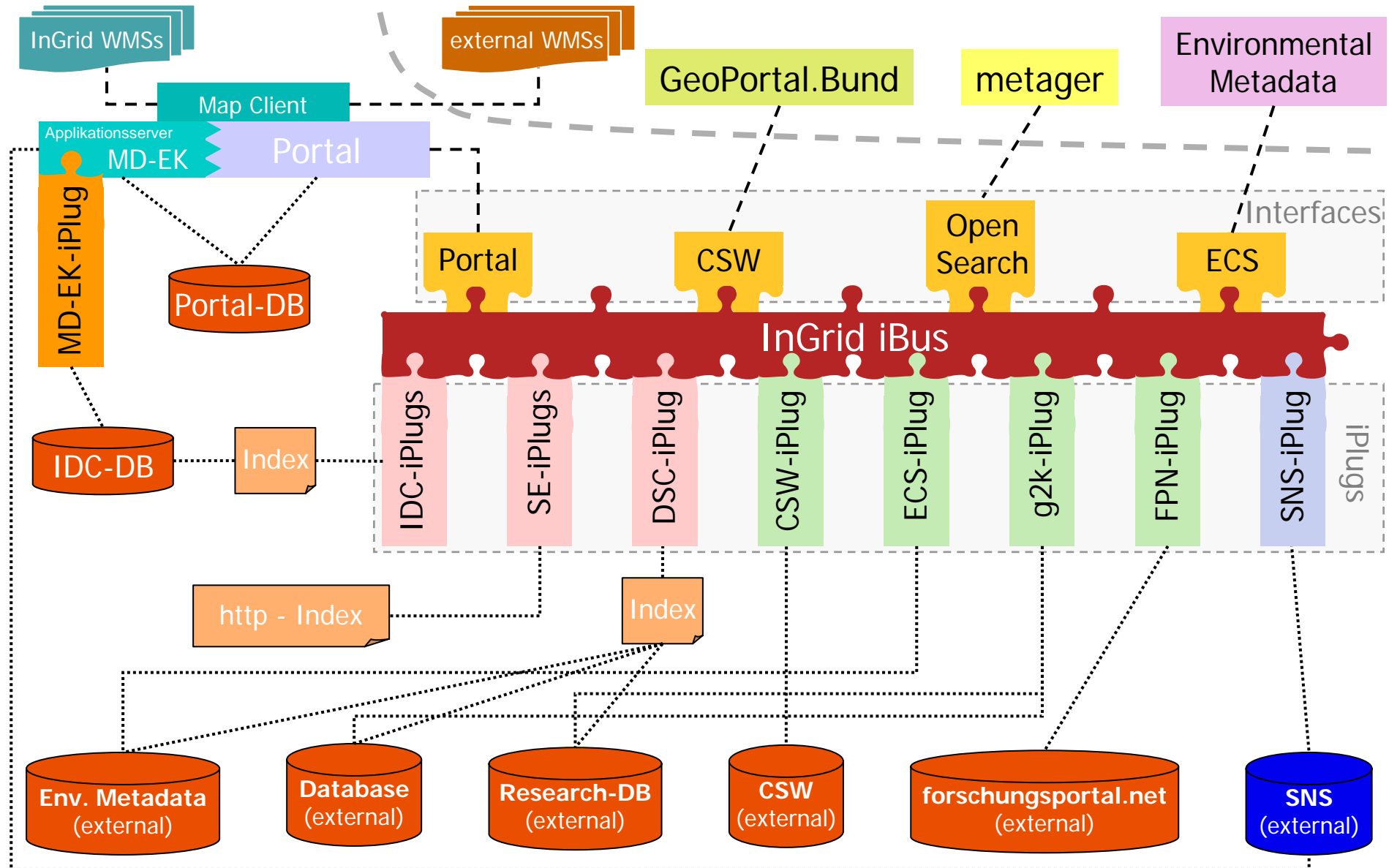
<mailto:kst@portalu.de>

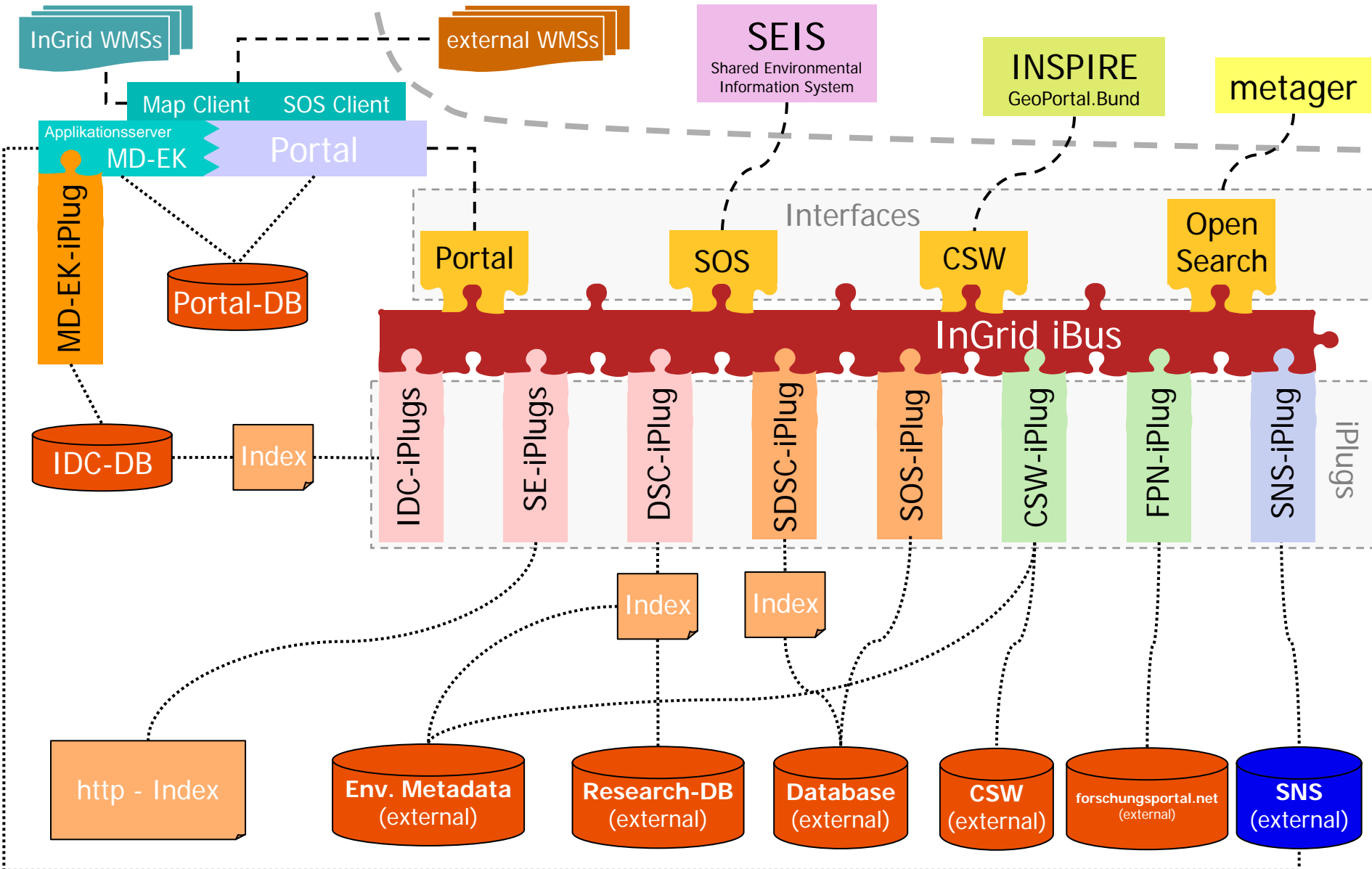
Dr. Martin Klenke

Koordinierungsstelle PortalU
im Niedersächsischen Ministerium für Umwelt und Klimaschutz

Archivstr. 2
D-30169 Hannover

+49 (0) 511 120 3407







Annex I Themen

- Koordinatenreferenzsysteme
- Geogittersysteme
- Geographische Namen
- Administrative Grenzen
- Transportwege
- Hydrographie
- Schutzgebiete

Annex II Themen

- Höhenmodelle
- Adressen
- Grundstücke
- Topographie
- Orthofotos

Annex III Themen

- Statistische Einheiten
- Gebäude
- Bodenschätzung
- Geologie
- Landnutzung
- Gesundheit
- industrielle und agrarische Produktionsstätten
- öffentliche Einrichtungen
- Bevölkerung
- Meteorologie, etc.



100. Geographic human interaction services (humanInteractionService)

- 101. Catalogue viewer
- 102. Geographic viewer
- 103. Geographic spreadsheet viewer
- 104. Service editor
- 105. Chain definition editor
- 106. Workflow enactment manager
- 107. Geographic feature editor
- 108. Geographic symbol editor
- 109. Feature generalization editor
- 110. Geographic data-structure viewer

200. Geogr. model/information management services

- 201. Feature access service
- 202. Map access service
- 203. Coverage access service
- 204. Sensor description service
- 205. Product access service
- 206. Feature type service
- 207. Catalogue service
- 208. Registry Service
- 209. Gazetteer service
- 210. Order handling service
- 211. Standing order service

300 Geographic workflow/task management services

- 301. Chain definition service
- 302. Workflow enactment service
- 303. Subscription service

400 Geographic processing services - spatial

- 401. Coordinate conversion service
- 402. Coordinate transformation service
- 403. Coverage/vector conversion service
- 404. Image coordinate conversion service
- 405. Rectification service

400 Geogr. proc. serv. – spatial (Forts.)

- 406. Orthorectification service
- 407. Sensor geometry model adjustment service
- 408. Image geometry model conversion service
- 409. Subsetting service
- 410. Sampling service
- 411. Tiling change service
- 412. Dimension measurement service
- 413. Feature manipulation service
- 414. Feature matching service
- 415. Feature generalization service
- 416. Route determination service
- 417. Positioning service
- 418. Proximity analysis service

500 Geographic processing services – thematic

- 501. Geoparameter calculation service
- 502. Thematic classification service
- 503. Feature generalization service
- 504. Subsetting service
- 505. Spatial counting service
- 506. Change detection service
- 507. Geographic information extraction services
- 508. Image processing service
- 509. Reduced resolution generation service
- 510. Image Manipulation Service
- 511. Image understanding service
- 512. Image synthesis services
- 513. Multi-band image manipulation service
- 514. Object detection service
- 515. Geoparsing service
- 516. Geocoding service

Annex B.4 Classification of spatial data services

600 Geographic processing services – temporal

- 601. Temporal reference system transformation service
- 602. Subsetting service
- 603. Sampling service
- 604. Temporal proximity analysis service

700. Geogr. processing services - metadata

- 701. Statistical calculation service
- 702. Geographic annotation service

800. Geogr. communication services

- 801. Encoding service
- 802. Transfer service
- 803. Geographic compression service
- 804. Geographic format conversion service
- 805. Messaging service
- 806. Remote file and executable management