


PortalU German Environmental Information Portal

## Shaping PortalU®/InGrid® to meet the INSPIRE requirements

*First experiences and some considerations on how to efficiently organize the German spatial data infrastructure as INSPIRE node*

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



1/15 ENVIRONFO 2008 / Martin Klenke

PortalU German Environmental Information Portal

## outline

- introduction  
PortalU®, InGrid®, Coordination Center PortalU, INSPIRE ...
- metadata model adjustments needed to emerge from ISO 19115/19119-compliant to "INSPIRE-proof" metadata
- the German geospatial data infrastructure (GDI-DE) with respect to INSPIRE, how to organize efficiently ?

2/15 ENVIRONFO 2008 / Martin Klenke

PortalU German Environmental Information Portal

## introduction (1)

### PortalU / InGrid / InGridCatalog

- PortalU is Germany's **environmental information portal**: <http://www.portalu.de>
- 'one stop' for environmental data and metadata held **by the German federal government & states**
- search engine accesses web pages, databases ("hidden web"), **environmental data catalogs**
- currently > 200 information providers, ~ 2.5 million web pages indexed
- **InGrid**, „information grid“, **the software behind PortalU**
- **InGridCatalog** (IGC): the InGrid Metadata component, successor of the "Umweltdatenkatalog" (UDK)
- IGC is **ISO 19115/19119 compliant. Near future: INSPIRE conform**

### Coordination Center PortalU

- funded by an administrative agreement between the German federal government and the federal states
- operation and further development of PortalU/InGrid
- support the metadata retrieval and maintenance within the german environmental authorities
- as LMO ("Legally Mandated Organisation") member of the INSPIRE drafting team metadata

3/15 ENVIRONFO 2008 / Martin Klenke

PortalU German Environmental Information Portal

## introduction (2)

### INSPIRE

- INSPIRE: *Infrastructure for Spatial Information in Europe*
- May 15th, 2007: the INSPIRE Directive (2007/2/EG) entered into force
- INSPIRE ...
  - ... demands web-based online-services for discovery, visualisation and download of spatial data
  - ... will ease pan-european use of geodata
  - ... has to be adopted to national rights within two years (!)
- the technical details are covered in so called "implementing rules" which are currently under development (metadata, network services, data specifications, etc.)
- INSPIRE Metadata Implementing Rules, status:
  - May 14th, 2008: IRs metadata approved by the INSPIRE Committee, adoption procedure started
  - final publication expected soon
  - after publication the IRs are directly applicable law in the member states
  - 2010 (Annex I & II) / 2013 (Annex III) national metadata has to be provided for INSPIRE

4/15 ENVIRONFO 2008 / Martin Klenke

Portal U German Environmental Information Portal

## INSPIRE / ISO / OGC

### INSPIRE: metadata and metadata exchange

- INSPIRE directive (16): *...implementing rules should be based, where possible, on international standards...*
- our main intentions to participate in the DT Metadata and the INSPIRE process in general:
  - keep INSPIRE as close as possible to ISO 19115/19119 in terms of metadata content: No excessive costs for existing, ISO-compliant catalogs like the IGC
  - tie OGC CSW metadata exchange specifications (DT Network Services) to INSPIRE: Again, no excessive costs for german catalog systems which commit themselves to OGC CSW 2.0.2 AP ISO 1.0 through the GDI-DE architecture concept
- both have been **nearly** met:
  - IR Metadata (4.3.3): *...The standardisation activity in this area has reached level of maturity. A metadata standard dedicated to Geographic Information is available with the publication of ISO 19115... INSPIRE IR Metadata are very close to ISO 19115/19119*
  - Draft IR Discovery and View Services (7.4.1): *...In order to provide search support for all search metadata elements as defined by the INSPIRE Metadata IRs, the OGC CSW/ISO 19115/19119 Application Profile (CSW ISO AP) shall be used as the reference specification for the INSPIRE Discovery Service...*

ENVROINFO 2008 / Martin Kienle

Portal U German Environmental Information Portal

## ISO vs. INSPIRE

### From ISO-compliant to INSPIRE-proof metadata – discrepancies between ISO 19115/19119 and INSPIRE metadata

- IR Metadata, 2.2.7: *conformity*
- IR Metadata, 2.2.1.5: *unique resource identifier*
- IR Metadata, 2.2.8: *constraint related to access and use*
- IR Metadata, 2.2.6.1: *lineage*
- IR Metadata, 2.2.2.2: *spatial data service type*
- IR Metadata, 2.2.3: *keyword, Annex B.4 – ‚Service type‘*
- IR Metadata, 2.2.6.2: *spatial resolution*
- IR Metadata, 2.8: *metadata point of contact – ‚E-Mail-Address‘*
- IR Metadata, 2.9: *keyword*

ENVROINFO 2008 / Martin Kienle

Portal U German Environmental Information Portal

## INSPIRE extensions

### IR Metadata: 2.2.7 conformity

- new, INSPIRE-related three-column 1:n mandatory field group
  - specification: INSPIRE Implementing Rules the dataset applies to
  - specifications reference date (e.g. generation, modification, publication)
  - degree of conformity: conformant / not conformant / not evaluated (Codelist B.5 Degree of conformity)
- refers to the requirement of the Directive in Articles 5-2 (a) and 11-2 (d) that metadata shall include information on the degree of conformity with the Implementing Rules provided in Art. 7-1, which refer to the interoperability and where practicable harmonization of spatial data sets and services.

ENVROINFO 2008 / Martin Kienle

Portal U German Environmental Information Portal

## INSPIRE extensions

### 2.2.1.5 unique resource identifier

- a value uniquely identifying the resource (!)
- uncertain how this will be realised, clarification by DT Data Specifications ?
- IGC first step solution (automatic): `<cat-id><UUID of the metadata-object>`
- example (SAPOS® - Reference Station 0655 Rotenburg):
 

```
idc-ni:015C0BE7-76F6-11D3-AE10-00104B57C66D
```

### 2.2.10.1 metadata point of contact – contact e-mail

- e-mail address becomes mandatory (ISO 19115 Cl\_Address + electronicMailAddress [0..\*])

ENVROINFO 2008 / Martin Kienle

Portal U German Environmental Information Portal

## INSPIRE extensions

### 2.2.8 constraint related to access and use

- two-column 1:n field group with columns
  - conditions applying to access and use of spatial datasets, services and fees (where applicable) – value domain: free text
  - limitations on public access (fixed codelist specified in INSPIRE directive Article 13)
- regarding (1) *no conditions apply and conditions unknown* are possible
- regarding (2) *no limitations* is a possible option

### 2.2.6.1 lineage

- process history and/or overall quality of the spatial dataset
- becomes mandatory for geodata (ISO 19115 LI\_Lineage + statement [0..1])

10/15 ENVRI/ENFO 2008 / Martin Kleinke

Portal U German Environmental Information Portal

## INSPIRE extensions

### 2.2.2.2 spatial data service type

- becomes mandatory for spatial data services
- value domain: fixed codelist
  - Discovery Service (CSW)
  - View Service (WMS)
  - Download Service (e.g. WFS)
  - Transformation Service (WCTS)
  - Invoke Spatial Data Service
  - Other Service

### 2.2.3 keyword, annex B.4 – ‘classification of spatial data services’

- becomes mandatory for spatial data services
- value domain: huge codelist (~70 entries) based on the geographic services taxonomy of EN ISO 19119, will be further developed with input from stakeholder communities

10/15 ENVRI/ENFO 2008 / Martin Kleinke

Portal U German Environmental Information Portal

## INSPIRE extensions

### 2.9 keyword

- at least one GEMET- (GEneral Multilingual Environmental Thesaurus) keyword relating to the INSPIRE Annex Themes becomes mandatory (<http://www.eionet.europa.eu/gemet/about>)
- GEMET has been developed by the European Topic Centre on Catalogue of Data Sources (ETC/CDS) and the European Environment Agency (EEA) since 1996
- consortium has been led by the Ministry of the Environment of Lower Saxony, members from Germany, Austria, Italy and Sweden
- EEA currently develops a list of GEMET keywords for the INSPIRE Annex Themes
- PortalU/IGC uses the UMTHES-thesaurus of the German Federal Environment Agency (Umweltbundesamt) for keyword generation
- for every IGC/UDK metadata object, three UMTHES keywords are mandatory
- GEMET is a subset of UMTHES
- IGC INSPIRE catalog migration: Automatic GEMET/INSPIRE keyword generation based on the existing UMTHES keywords will be possible by means of the Semantic Network Webservice SNS (<http://www.semantic-network.de>)

10/15 ENVRI/ENFO 2008 / Martin Kleinke

Portal U German Environmental Information Portal

## GDI-DE/INSPIRE - Infrastructure

### GDI-DE pilot project *Geodatenkatalog-DE*

- setup of a topology of existing German catalog services in line with INSPIRE
- grouping of the services to supply a central German metadata node for INSPIRE
- project service provider
  - several multidisciplinary geodata catalogs from German federal states
  - geodata catalog of the federal government (geoMIS.Bund)
  - the cross-state catalogs PortalU<sup>®</sup> (environment) and Adv-MIS (surveying)
- project user
  - Joint Research Center, European Commission, Ispra (<http://ec.europa.eu/dgs/jrc/>)
  - GIW-Commission (Federal Ministry of Economics and Technology, <http://www.geobusiness.org/>)

10/15 ENVRI/ENFO 2008 / Martin Kleinke

Portal U German Environmental Information Portal GDI-DE/INSPIRE - Requirements

## GDI-DE & INSPIRE requirements

- GDI-DE specification for catalog services: OpenGIS® Catalogue Services Specification 2.0.2 - ISO Metadata Application Profile (1.0.0)
- probably as well INSPIRE metadata exchange standard (state of discussion in DT Network Services):
  - ... *In order to provide search support for all search metadata elements as defined by the INSPIRE Metadata IRs, the OGC CSW ISO 19115/19119 Application Profile (CSW ISO AP) shall be used as the reference specification for the INSPIRE Discovery Service....* (Draft Implementing Rules for Discovery and View Services, 7.4.1)
- technical/functional requirements are formulated in the *Draft Implementing Rules for Discovery and View Services, 7.4.3, e.g.:*
  - time for sending initial response to service request in normal situation shall be **3 seconds**. This time shall allow to send **250 Metadata records**
  - the minimum number of **simultaneous service requests** served according to the performance requirement shall be **100** for the Discovery Service

-> hard to realize without caching and indexing mechanisms !

15/15 ENVRIINFO 2008 / Martin Klenke

Portal U German Environmental Information Portal Geodatenkatalog-DE

## topology of (INSPIRE) catalog services

- INSPIRE (softly) demands a central metadata node for each member state -> geoPortal.Bund
- basic topology concepts
  - topically-based composition (PortalU, AdV MIS)
  - administrative hierarchy (local authority -> ... -> federal state -> ... -> federal government)
  - no cascading topology, direct connection of every single catalog to the national node, possibly supported by a registry service
- Status Quo: All three concepts already co-exist and have to be considered
- Geodatenkatalog-DE, first findings:
  - caching/harvesting/indexing mechanisms are necessary, at least for the central german INSPIRE metadata node, in order to meet INSPIRE demands
  - functionality is currently under development within the pilot project
- similar functionality is needed for information broker systems like PortalU/InGrid

15/15 ENVRIINFO 2008 / Martin Klenke

Portal U German Environmental Information Portal Summary

## summary & conclusion

- INSPIRE implementing rules currently taking more and more shape
- the relevant existing german geo- and environmental data catalogs are based on ISO 19115/19119, they seem to be sustainably and INSPIRE-ready
- technically, most german catalog owners and software suppliers committed themselves to provide/support OGC CSW AP ISO 1.0 interfaces
- when it comes to the practical implementation of the IRs, particularly with respect to service topologies and performance demands, a lot of questions are not answered yet
- this underlines that practical pilot projects and early prototyping is necessary to develop feasible strategies for national and european metadata flows
- PortalU/InGrid metadata components will be INSPIRE-proof end of 2008

15/15 ENVRIINFO 2008 / Martin Klenke

Portal U German Environmental Information Portal Contact

*Thank you for your attention !*

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

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

[kst@portalu.de](mailto:kst@portalu.de)

[martin.klenke@portalu.de](mailto:martin.klenke@portalu.de)

Dr. Martin Klenke  
 Coordination Center PortalU  
 Ministry of Environment and Climate Protection  
 Archivstr. 2  
 D-30169 Hannover  
 +49 (0) 511 120 3407

15/15 ENVRIINFO 2008 / Martin Klenke

Portal U German Environmental Information Portal

## INSPIRE Annex-Themes

### Annex I Themes

- Coordinate reference systems
- Geographical grid systems
- Geographical names
- Administrative units
- Addresses
- Cadastral parcels
- Transport networks
- Hydrography
- Protected sites

### Annex II Themes

- Elevation
- Land cover
- Orthoimagery
- Geology

### Annex III Themes

- Statistical units
- Buildings
- Soil
- Land use
- Human health and safety
- Utility and governmental services
- Environmental monitoring facilities
- Production and industrial facilities
- Agricultural and aquaculture facilities
- Population distribution – demography
- Area management/restriction/regulation zones and reporting units
- Natural risk zones
- Atmospheric conditions
- Meteorological geographical features
- Oceanographic geographical features
- Sea regions
- Bio-geographical regions
- Habitats and biotopes
- Species distribution
- Energy resources
- Mineral resources

ENVROINFO 2008 / Martin Kienle

Portal U German Environmental Information Portal

## 2.2.8.2 Limitations on public access - Codelist

- the confidentiality of the proceedings of public authorities, where such confidentiality is provided for by law,
- International relations, public security or national defence,
- the course of justice, the ability of any person to receive a fair trial or the ability of a public authority to conduct an enquiry of a criminal or disciplinary nature,
- the confidentiality of commercial or industrial information, where such confidentiality is provided for by national or Community law to protect a legitimate economic interest, including the public interest in maintaining statistical confidentiality and tax secrecy,
- intellectual property rights,
- the confidentiality of personal data and/or files relating to a natural person where that person has not consented to the disclosure of the information to the public, where such confidentiality is provided for by national or Community law,
- the interests or protection of any person who supplied the information requested on a voluntary basis without being under, or capable of being put under, a legal obligation to do so, unless that person has consented to the release of the information concerned,
- the protection of the environment to which such information relates, such as the location of rare species.

ENVROINFO 2008 / Martin Kienle

Portal U German Environmental Information Portal

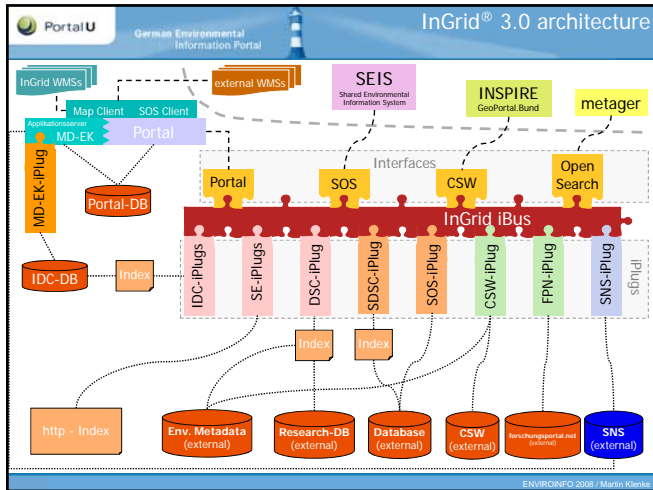
## Annex B.4 Classification of spatial data services

<p><b>100 Geographic human interaction services (humanInteractionService)</b></p> <p>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</p> <p><b>200 Geogr. model/information management services</b></p> <p>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</p> <p><b>300 Geographic workflow/task management services</b></p> <p>301. Chain definition service 302. Workflow enactment service 303. Subscription service</p> <p><b>400 Geographic processing services - spatial</b></p> <p>401. Coordinate conversion service 402. Coordinate transformation service 403. Coverage/vector conversion service 404. Image coordinate conversion service 405. Rectification service</p>	<p><b>400 Geogr. proc. serv. – spatial (ForTs.)</b></p> <p>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</p> <p><b>500 Geographic processing services – thematic</b></p> <p>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</p>	<p><b>600 Geographic processing services – temporal</b></p> <p>601. Temporal reference system transformation service 602. Subsetting service 603. Sampling service 604. Temporal proximity analysis service</p> <p><b>700 Geogr. processing services - metadata</b></p> <p>701. Statistical calculation service 702. Geographic annotation service</p> <p><b>800 Geogr. communication services</b></p> <p>801. Encoding service 802. Transfer service 803. Geographic compression service 804. Geographic format conversion service 805. Messaging service 806. Remote file and executable management</p>
--	---	---

ENVROINFO 2008 / Martin Kienle

Portal U German Environmental Information Portal

ENVROINFO 2008 / Martin Kienle



**Portal U** German Environmental Information Portal

### GDI-DE/INSPIRE - Requirements

#### GDI-DE & INSPIRE requirements (2)

- required search criteria (INSPIRE Directive Art.11.2), combinations of:
  - keywords
  - classification of spatial data and services
  - the quality and validity of spatial data sets
  - degree of conformity with the implementing rules provided for in Article 7(1)
  - geographical location
  - conditions applying to the access and use of spatial data sets and services
  - the public authorities responsible for the establishment, management, maintenance and distribution of spatial data sets and services
- not fully covered by CSW AP ISO 1.0 yet (how to handle *conformity, conditions for access und use* ?)
- JRC provided an implementation guideline: *Draft Guidelines – INSPIRE metadata implementing rules based on ISO 19115 and ISO 19119*
- goal: No extended INSPIRE-ISO-profile, all demands should be met by mappings from the standard ISO-classes

ENVROINFO 2008 / Martin Kienle

**Portal U** German Environmental Information Portal

### InGrid-CSW-extensions

#### caching and indexing

- caching I: **GetRecords** (*ElementSetName=brief*) with Wildcard-Query
- caching II: metadata download (**GetRecordById**, *ElementSetName=full*), storage as XML file
- every connected CSW data source gets its own cache (better scalability, load distribution, configuration flexibility)
- indexing (Apache Lucene) with an indexfield-to-XPath mapping on the XML file
- at query time, a performant response can be generated directly from the index, no (slow) contact to the original data source is necessary

ENVROINFO 2008 / Martin Kienle