GIGAS
www.thegigasforum.eu

Catalog and Metadata sessions

Frédéric Houbie (Erdas), Marie-Lise Vautier (IGN)

Parallel Working Session 1/5

GIGAS
www.thegigasforum.eu

Scope

- Review the issues and recommendations identified in the analysis of the approaches of INSPIRE, GEOSS, GMES, etc. Identify which of these issues and recommendations are generally agreed and which require further cross-initiative discussions.
- Jointly work on a plan how progress can be made. Depending on the level of agreement about an issue this may mean continuing cross-initiative discussions or taking concrete actions to address the issue. Together, we will try to identify the right fora for further discussions, the right instruments for actions and the right actors to involve.

Issues and Recommendations

Issues concern:

- Session 1
  - Service metadata
  - Discovery metadata mapping
  - Metadata discovery profile
  - Access to other types of resources
- Session 5
  - ISO 19119 XML Schemas : versions, repository…
  - Unique identification of the resources
  - Synchronization of catalogues content

© 2010 GIGAS Consortium http://www.thegigasforum.eu

Service metadata (1)

- INSPIRE has defined the ISO19119 metadata elements to be used to describe their web services.
- The OGC OWS Common specification defines the Capabilities XML structure to describe an OGC Web Service.
- OGC XML Capabilities are not sufficient to build a complete ISO19119 record, therefore:
    - OGC XML Capabilities cannot be used to populate a Catalog compliant with INSPIRE
    - As OGC CS-W ISO AP uses the ISO19119 schemas, OGC XML Capabilities cannot be used to populate a Catalog compliant with OGC CS-W ISO AP

© 2008 GIGAS Consortium http://www.thegigasforum.eu

GIGAS
www.thegigasforum.eu

Service metadata (2)

1. It is recommended to OGC to add a in the OGC Capabilities a link to a ISO 19119 metadata instance encoded using ISO 19139 so that it can be published into an OGC CS-W ISO AP Catalog
   This recommendation has already been pushed at OGC (OWS Common 1.2 SWG) and taken into account by the group.
   INSPIRE must specify if this link is to be taken into account and how.
   - GetRecordById functionnality is not mandatory in the INSPIRE regulation.
   - Avoid duplication of access to metadata (I.e through discovery service and simple HTTP/FTP download)
   - Link to metadata is supposed to take place through the Discovery Services.
   - Question of SOAP encoding which would not be possible with this new element (URL).
   - GET/POST method not specified in the regulation but member state practices show use of POST for GetRecords.

© 2008 GIGAS Consortium http://www.thegigasforum.eu

Service metadata (3)

- Discussion on WSDL/Capabilities/19119
  - WSDL and capabilities are perceived as being more machine oriented, 19119 as human oriented
- Need for a generic mechanism for linking to an OGC Web Service from the Capabilities ? (not simply a URL)

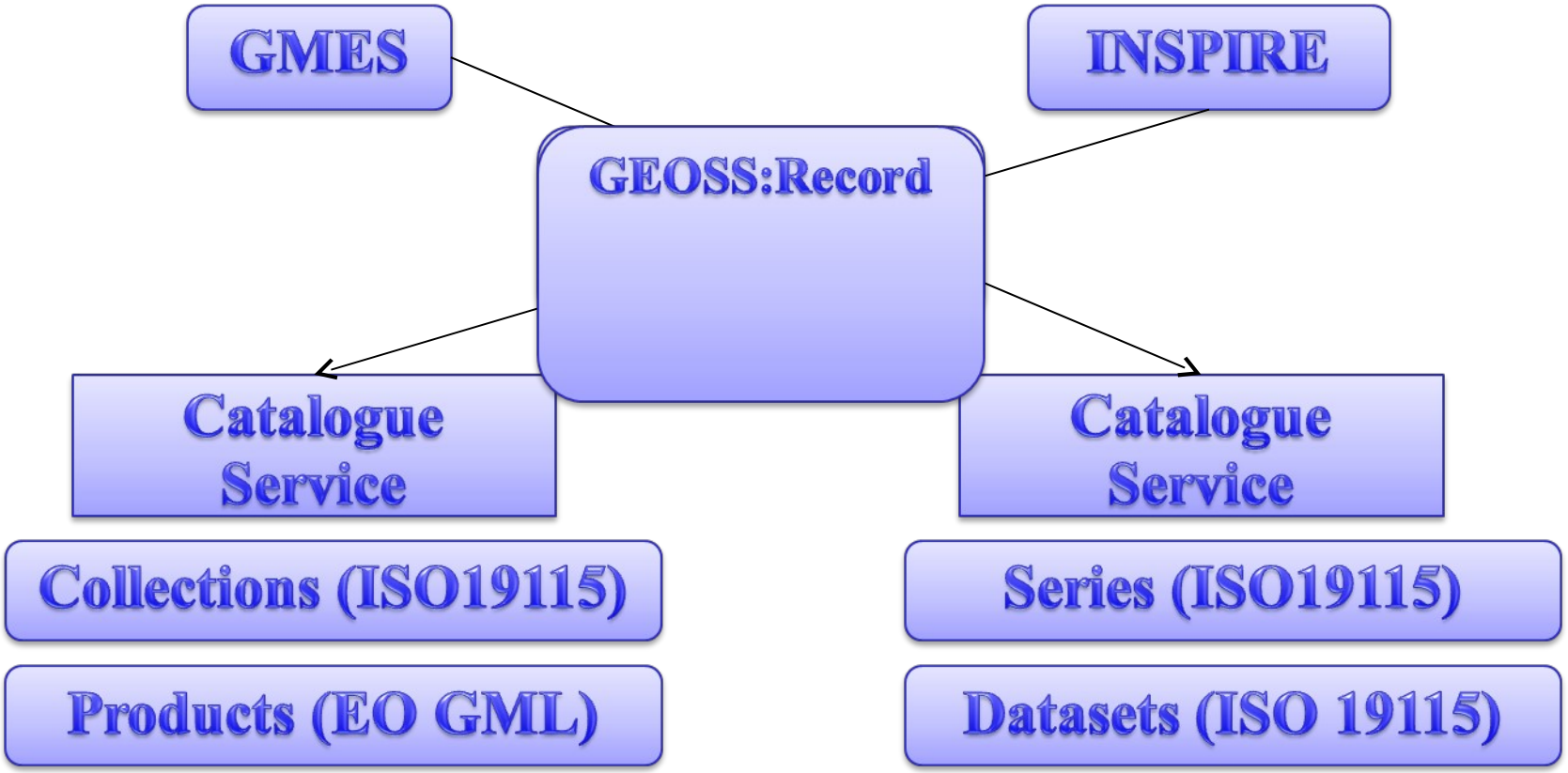
2. It is recommended to INSPIRE to define a standard mapping between OGC service description (Capabilities) and INSPIRE metadata and publish it as an OGC Best Practice.
   - To be discussed in the INSPIRE task force

Discovery metadata elements mapping (1)

Currently, the different initiatives/projects have defined a set of discovery metadata elements which go beyond the CS-W Metadata Record (based on DC). These elements are not the same and thus cannot be used to do a global discovery of datasets or services across the catalogue services of the initiatives.

© 2008 GIGAS Consortium http://www.thegigasforum.eu

GIGAS
www.thegigasforum.eu



Find the color images from SPOT 5 on Mediterranean Coast during last summer with a cloud coverage less than 15% and a resolution of 2m

Find the ortho images on Mediterranean Coast during last summwer described by beach, sand, vegetation keywords

GIGAS
www.thegigasforum.eu

Discovery metadata elements mapping (2)

It is recommended to INSPIRE, GEOSS and GMES technical teams to elaborate a set of core queryable metadata elements be used in common by the three initiatives.

- Recommendation on the standardisation body ?
    - INSPIRE Discovery elements are fixed in the Directive and cannot be changed
- Need to identify the intersection
    - Solve 1 to n mappings issues
- Who will do that ?? GEOSS has an experience on this, having defined the GEOSS Record based on existing implementations (Engineering Report AIP2)
- To be considered in Catalogue Service 3.0 SWG;
- INSPIRE has identified some limitations in Dublin Core; this report will be submitted to the Catalogue Service 3.0 SWG
- All the annexes' themes may lead to the definition of additional metadata elements, but at the evaluation/use levels not at the discovery level

© 2008 GIGAS Consortium http://www.thegigasforum.eu

GIGAS
www.thegigasforum.eu

Discovery metadata elements mapping (3)

- Rephrased recommendation :
- It is recommended to the GIGAS consortium
  - to prepare a cross walk of the INSPIRE, GEOSS and GMES common metadata elements,
  - to identify issues of semantic mismatch, cardinality issues, etc. and resolve questions raised in the 3 communities
  - to prepare guidelines for mapping community metadata into the core queryables
  - to prepare CR to CSW 3.0 for the addition of new queryables to the baseline

Metadata Discovery Profiles (1)

- The catalogues of the different initiatives are built on the Catalogue Service for the Web standard of the OGC (CS-W), but the usage of the base catalogue profiles differs strongly.
  - GEOSS supports all the available standard service interfaces, i.e.: CS-W ISO AP v1.0.0, CS-W ebRIM AP Basic Package, CS-W ebRIM AP EO EP, CS-W ebRIM AP CIM EP
  - INSPIRE requests the use of the elements defined in the Discovery implementing rules, and suggests to use the **CS-W ISO profile** (CS-W ISO AP INSPIRE Profile).
  - GMES is supporting profiles based on the **CS-W ebRIM Profile**, i.e. the Basic Package, the EOProduct Ext. Package for the discovery of Earth Observation metadata (encoded acc. to the GML App. Schema for EO Products) and the CIM Ext. Package for the discovery of collection metadata (encoded acc. to ISO 19139).

Metadata Discovery Profiles (2) - Recommendation and issues

- Issue #1: the CSW ISO Profile and the CIM EP based on the CSW ebRIM Profile are not interoperable solutions for the discovery of ISO 19115 and ISO 19119 metadata, except by using the very limited CSW Record format.
- Recommendation : At the minimum, a bridge between the ISO Application Profile and the CIM extension package should be defined.
  - Work has been done on this issue in the HMA project, cf. http://wiki.services.eoportal.org/tiki-download_wiki_attachment.php?attId=425&download=y
- Issue #2: In the future, should the initiatives use the same Application Profile ?
  - What if they don't have the same metadata requirements ?
- Issue #3: In the future, should the initiatives profile the base CS-W specification for each type of resource (e.g. feature catalogues, ISO 19115-2 metadata, etc.) or should they instead define extension packages of the ebRIM Application Profile?
  - The ebRIM Application Profile defines a generic metamodel that can be instantiated for the discovery of different types of resources.
- Use of ISO 19115 for these resources ? ISO 19110 is referenced from ISO 19115. THE CSW ISO AP would need to be extended for these types of resources (with additional queryables)

GIGAS
www.thegigasforum.eu

Access to other types of resources (1)

1. It is recommended to
   • the INSPIRE, GEOSS and GMES technical teams
   - Question : who is targeted by this recommendation? INSPIRE: Task Forces ?
   that they consider together the requirements to set up and provide standardized and possibly shared access to other types of geographic resources. Resources to be considered include:
   • Feature catalogues;
   • Property types;
   • Sensors;
   • Application schemas;
   • Geodetic codes and parameters
   • Portrayal information - work underway in ISO/TC 211 (on modelling not registers)
   • Feature concept dictionaries, thesauri, glossaries and other registers of terms
   • Data quality measures – Discussion about measurement accuracy; requirement missing in INSPIRE ? - work underway in ISO/TC 211 (on modelling not registers)
   • Ontologies – work underway in ISO/TC 211 (on modelling not registers)
   • Code lists
   • Object identifier namespaces
   • Units
   • Data specifications
   • Encoding schemas

Access to other types of resources (2)

2. It is also recommended to consider the use of the OASIS RegRep standard and/or the OGC RegRep profile currently being developed for the registry services used to manage these resources.

GIGAS
www.thegigasforum.eu

Next steps

- Which issues/recommendations require further discussion?
  - Can we record the key issue(s)?
  - Identify the right fora for contuining a cross-initiative discussion
- If there is agreement, which are the right steps in each initiative / organisation to address the issue
  - Identify concrete actions including responsibilities and milestones

GIGAS
www.thegigasforum.eu

Catalog and Metadata sessions

Frédéric Houbie (Erdas), Marie-Lise Vautier (IGN)

Parallel Working Session 5

Issues and Recommendations

Issues concern:

- Session 5
  - XML implementation of ISO 19119
  - Unique identification of the resources
  - Synchronization of catalogues content

© 2010 GIGAS Consortium http://www.thegigasforum.eu

GIGAS
www.thegigasforum.eu

XML implementation of ISO 19119 – the problem

- Schema in OGC repository is bound to *interim* versions of 19115/19139 (GCO, GMD) and 19136 (GML)
  - Can't easily combine in applications that also use the official versions of GML etc
  - Market confusion
  - Reluctance of INSPIRE members and IOC to implement
- Other schema repositories
  - http://standards.iso.org/ittf/PubliclyAvailableStandards/ISO_19139_Schemas/
  - http://eden.ign.fr/xsd/csw_profile/20071126/20071126

GIGAS
www.thegigasforum.eu

Technical solution

- Schema is easily fixed
  - <import> final versions of gco, gmd, gml
  - New targetNamespace (because not backward compatible)
    - http://www.opengis.net/srv/2.0 (?)
- Very minor changes to specification document
  - Annex D.1 – delete one sentence, change one sentence (namespace)
  - Fix examples
- Designate one repository as ‘normative’

GIGAS
www.thegigasforum.eu

Implementation/deployment consequences

- New deployments
  - Compatible with other GML applications
  - Removes confusion + impediment to wider deployment
  - INSPIRE IOC will have better target so will move forward faster
- Existing deployments
  - Current schema in OGC repo. stays put, so existing deployments will not break
  - … but, must upgrade in order to link with new deployments
    - Doesn't have to be immediate
    - Unfortunate price of early adoption
- Can anything be done to make it easier on the early adopters?

Governance question

- Schema repository – should ISO/TC 211 have one?

GIGAS
www.thegigasforum.eu

Unique identification of resources

The three initiatives should consider how resources can be uniquely identified in each initiative's catalogues. If a service is made available in INSPIRE, GEOSS and/or GMES, can it be identified as the same (particularly if the access URLs are different for each initiative but the exposed content and functionality is the same).
Where metadata or content profiles are different but the underlying database(s) is/are the same, some identifier may be useful to identify the underlying resource, with different identifiers for the different service profiles.

- Problem under discussion in INSPIRE
- 2 aspects
    - Need for unique/persistent identifiers
    - Need to identify where to put this unique identifier
- Several identification mechanisms in ISO 19115; Clarification could take place in the 19115 revision;
- Best practice on both aspects needed; INSPIRE has defined some mechanism in the Conceptual Framework document;
- Study made by Erdas on identification practices in IGN, INSPIRE and Germany (provided by con terra) : all had different practices
- UUID vs URI vs legacy identifiers (etc.) : should there be best practices as well ? Difficult for legacy resources
    - CSW ISO AP has defined some rules

© 2008 GIGAS Consortium http://www.thegigasforum.eu

GIGAS
www.thegigasforum.eu

Unique identification of resources (2)

- fileIdentifier is mandatory in ISO AP, one identificationInfo only (one dataset, service, application… only)
- Semantic web: practices are emerging
- Additional layer : Identity resolution service; Delegate the identification issue to another component
- This is not an issue specific to the geo community
- GEOSS : redundant resources, no solutions yet
- OGC NA dealing with multiple identifiers and syntax
- Reformulated recommendation
  - The three initiatives should consider how resources can be uniquely identified and share experience/practices on this topic

GIGAS
www.thegigasforum.eu

Synchronization of catalogues content

This is a further ellaboration of previous issue. Where a resource is being served into more than one initiative, there is a danger that the catalogues representing the resource may be desynchronised in the different initiatives. This may come about where metadata harvesting is employed on different update schedules, or where cached and live queries are made to the resource or the initiative catalogues.

The three initiatives are recommended to work together to ensure that any resource is represented consistently.

INSPIRE experience: under test, preliminary results could be shared with the other initiatives
GEOSS also has experience/practice on this topic (work on patterns); OGC NA too.

© 2008 GIGAS Consortium http://www.thegigasforum.eu

GIGAS
www.thegigasforum.eu

Next steps

- Which issues/recommendations require further discussion?
  - Can we record the key issue(s)?
  - Identify the right fora for contuining a cross-initiative discussion
- If there is agreement, which are the right steps in each initiative / organisation to address the issue
  - Identify concrete actions including responsibilities and milestones