7th Framework Programme
ENV.2010.4.1.2-2
Integrating New Data Visualisation Approaches of Earth Systems into GEOSS Development

Project Nr: 265178

QUAlity aware VIsualisation for the Global Earth Observation system of systems

Deliverable D4.2
Smart searchable interface component report

Version 1.0

Due date of deliverable: 31/10/2013
Actual submission date: 16/12/2013

Document control page				
Title	D4.2 Smart searchable interface component report			
Creator	Lorenzo Bigagli (CNR)			
Editor	Lorenzo Bigagli (CNR) and Roberto Roncella (CNR)			
Description	This document describes the GeoViQua smart searchable interface component for improved quality-enabled discovery of geospatial resources in the context of GEOSS			
Publisher	GeoViQua Consortium			
Contributors	GeoViQua Partners			
Type	Text			
Format	MS-Word			
Language	EN-GB			
Creation date	25/10/2013			
Version number	1.0			
Version date	16/12/2013			
Last modified by	Lorenzo Bigagli			
Rights	Copyright © 2013, GeoViQua Consortium			
Dissemination level		CO (confidential, only for members of the consortium)		
	X	PU (public)		
		PP (restricted to other programme participants)		
		RE (restricted to a group specified by the consortium)		
Nature	X	R (report)		
		P (prototype)		
		D (demonstrator)		
		O (other)		
Review status		Final Draft	*Where applicable:*	
		WP leader accepted		Accepted by the PTB
		PMB quality controlled		Accepted by the PTB as public document
	X	Coordinator accepted		
Action requested		to be revised by all GeoViQua partners		
		for approval of the WP leader		
		for approval of the PMB		
		for approval of the Project Coordinator		
		for approval of the PTB		
Requested deadline	31/10/2013			

Revision history			
Version	Date	Modified by	Comments
1.0-draft	25/10/2013	CNR_LB	Document created
	26/11/2013	CNR_LB, CNR_RR	Table of content, initial draft
	29/11/2013	CNR_RR	First draft
	02/12/2013	CNR_RR	Revision and consolidation
	03/12/2013	CNR_RR	Final draft
1.0	16/12/2013	CNR_LB	Final revision

Institution	Contributors
CNR_LB	CNR: Lorenzo Bigagli
CNR_RR	CNR: Roberto Roncella

Table of Contents

List of Figures

List of Tables

Executive Summary

This document describes the GeoViQua smart searchable interface component for improved quality-enabled discovery of geospatial resources in the context of GEOSS.

Deliverable 4.1 Graphical search interface report focused on the client graphical user interface, while this one introduces the GeoViQua smart searchable interface architecture and the main related components focusing on the GeoViQua server side components related to data search: the GeoViQua Broker and the GeoViQua Capacity.

1. Introduction

1.1 Scope

The present document is the Deliverable 4.2 "Smart searchable interface component report" of the GeoViQua Project. [GVQ-DOW 2010]

It describes the searchable server side components and the interface components for quality-enabled discovery of geospatial resources in the context of GEOSS. The objective of this work was to extend the main functionalities of a catalogue service component to support quality-related aspects and constraints. To establish the requirements, indications and guidance have been gathered from the users and the project partners by a number of means, including project deliverables, formal and informal meetings and other communications.

This document is mainly targeted at a technical audience and is concerned with ICT aspects only; all the other concerns (e.g. quality of data and scientific methods, legal aspects) are out of scope of this deliverable and can be found in other previously released documents.

This document is structured in five main sections. This **Section 1** provides an introduction to the project Work Package 4 (WP4) and its background, and establishes the terminology and notation used throughout the rest of the document. **Section 2** describes the smart searchable interface component architecture. **Section 3** describes the GeoViQua Broker and its components. **Section 4** describes the GeoViQua Capacity component and its services. **Section 5** summarizes the conclusions and the foreseen future improvements.

1.1.1 Relation to GeoViQua tasks

This deliverable is directly related to Task 4.1 "Quality-aware catalogue service" of WP4 "Enhanced geo-search components".
It relates to: WP3 "Data quality elicitation mechanisms", particularly to Task 3.5 "User feedback and quality assessment of data sets in GEOSS", which defines the GeoViQua data model; and to WP6 "Delivery of solutions to end users", which integrates everything.

Task 4.1 description
This task aims to develop a catalogue service component (i.e. a service provider) that is compliant with the GCI and supporting queries based on data quality constraints. Therefore, the main features of the catalogue service are the following: a) adoption of international standard interfaces for discovery; b) support of discovery through both "harvesting" and "query distribution" approaches – as required by the GCI Recommendations; c) multi-disciplinarity support to address applications in all the nine GEOSS SBAs; d) quality-awareness to support discovery clauses based on quality metadata characterizing the addressed resources using the data model (encoding) developed in Task 6.1.
Specific RTD actions in order to implement functionalities realizing the introduced features:

- Adoption of standard interfaces for discovery: in order to enable the catalogue service to work in the planned GCI, and more generally in the context of the most relevant System of Systems initiatives for geo-information sharing (e.g. INSPIRE, GMES, SEIS), the component will be based on the most relevant international standards (e.g. OGC CSW 2.02 which is the standard adopted by GEOSS Clearinghouse Catalogue and Community Catalogues). Welladopted Application Profiles and Extension Packages will be implemented and supported, such as: CSW/ISO profile (recognized by INSPIRE); CSW-ebRIM/EO (recognized by GMES). Lightweight technologies for discovery interoperability (e.g.

OpenSearch) will be also implemented in order to open the catalogue services to the world of Web 2.0 applications (e.g. through the publication of RSS or ATOM feeds) and support Semantic Web search functionality; [CNR, 52N]

- Harvesting and distributed search functionalities: the GCI Recommendations recognize that "many data holders have shown a clear preference for registering their catalogue(s) of data holdings, as opposed to the actual datasets themselves. [...] However, to access the resources held in these catalogues requires the implementation of either a distributed search, or they are harvested on a recurrent basis". Which approach is the best depends on the characteristics of the registered catalogues (e.g. rate of update, catalogue size, etc.). Thus a flexible and extensible catalogue service component should support both the approaches. This requires specific RTD actions. In the "distributed search" approach, a query directed to the central catalogue must be replicated towards every single registered catalogue. Since they may expose different interfaces, a "brokering & mediation" functional module is required in order to asynchronously "distribute" and "translate" queries. While, in the "harvest" approach the query is performed on metadata local copies; the registered catalogues are periodically queried in order to collect copies of the metadata of all the registered datasets; the metadata collection strategy must be flexible and configurable in order to optimize performances and provide up-to-date results. Besides synchronization strategies must be carefully implemented; [CNR]
- Multidisciplinarity support: In order to enable applications on all the nine GEOSS SBAs, it is necessary to support queries on multi-disciplinary catalogues. This implies to address different metadata models that need to be harmonized. The catalogue service component must implement mediation functionalities in order to provide a single view of different metadata models. Besides, the catalogue component must accept the SBAs and SBA categories as valid query clauses; [CNR].
- Quality-awareness: The GCI Recommendations recognize the importance of quality information: "The GCI should enable users to discover the declared quality of resources registered in the GEOSS, enabling them to query and filter the GCI information content quality, based on the metadata provided." The proposed component will profile/extend the discovery interface to support the specification of quality constraints. This task will investigate the design and implementation of quality-aware search capabilities through the use of both metadata and semantic relationships (i.e. RDF encoded). Metadata can be used when quality information is available and modeled employing a suitable quality metadata model (this will be based on the work in Task 6.1). In some cases quality can be better express through semantics relationships (e.g. "Layer Y is the uncertainty field for Layer X"). In this case the experience from Semantic Web applications will be considered, e.g. implementing inferences on RDF representations of the semantic relationships. The quality information will be used for searching or to sort the results by quality. [CNR, UREAD]
- Statistics of the dataset collection: Some statistics for different metadata parameters of the dataset collections (e.g. number of datasets per scale, time, SBAs, quality indicators etc) could be requested and retrieved from the catalogue. [CREAF]

This task includes the design and implementation of the catalogue service component, extending the GI-cat middleware technology. GI-cat is an implementation of distributed catalogue services, conceived as an extensible framework, and it has been tested and used in contexts like: GEOSS IP3 (Interoperability Process Pilot Project), GEOSS AIP-2 (Architecture Implementation Pilot), FP7 EUROGEOSS project, ESA HMA project, FP7 SeaDataNet, FP7 GENESI-DR, etc.

1.2 GeoViQua enhanced geo-search components architecture

Figure 1 shows the simplified component diagram of the overall architecture of GeoViQua enhanced geo-search components.

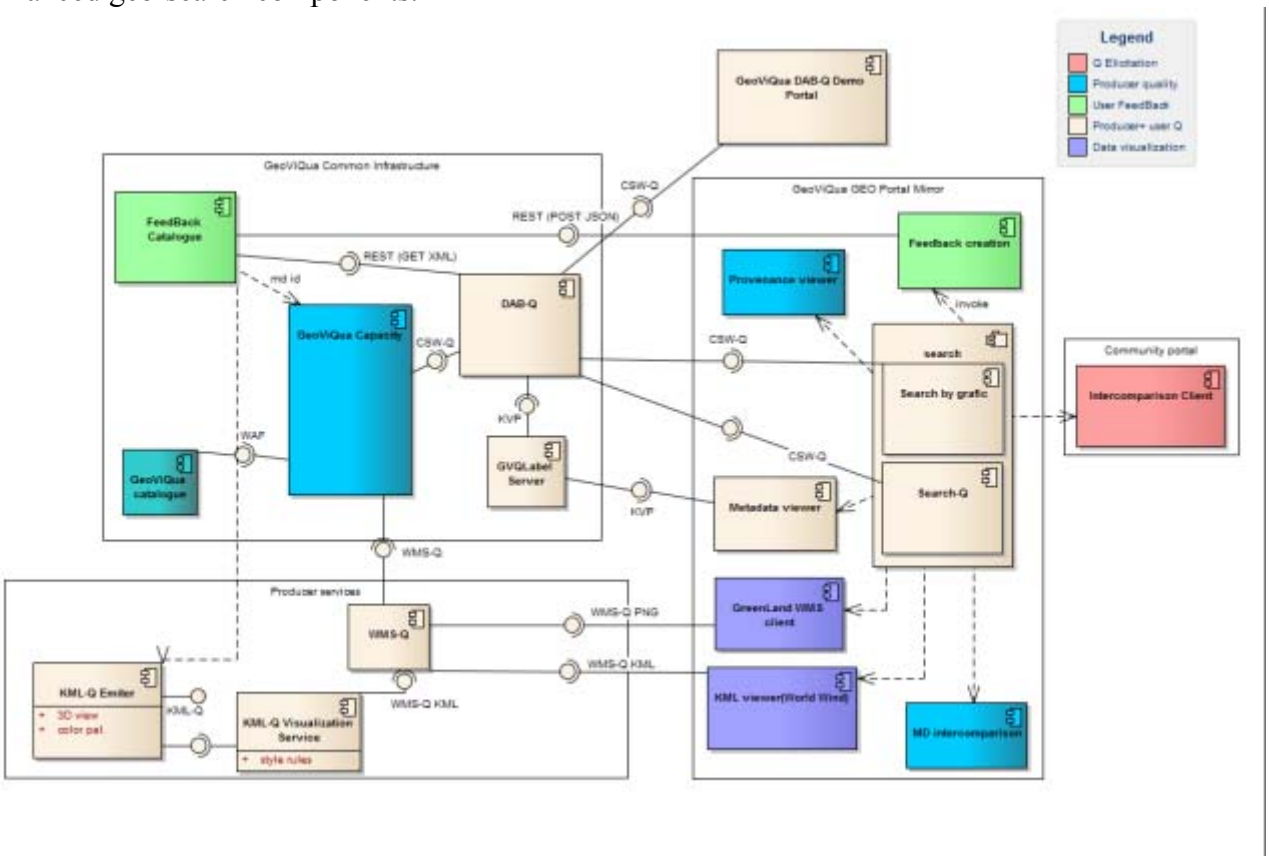

Figure 1 – Simplified Components Diagram

In this deliverable we focus on the server side Smart searchable interface components, which are contained in the GeoViQua Common Infrastructure (top left box in the picture).

1.3 Conventions

1.3.1 Acronyms and abbreviations

CSW – Catalogue Service for the Web
DAB – Discovery and Access Broker
FP7 – Seventh EU Framework Programme for Research and Technological Development
GCI – GEOSS Common Infrastructure
GEO – Group on Earth Observation
GEOSS – Global Earth Observation System of Systems
GVQ – GeoViQua
ICT – Information and Communication Technology
ISO – International Organization for Standardization
OGC – Open Geospatial Consortium
PQM – Producer Quality Model
UML – Unified Modeling Language
UQM – User Quality Model
WAF – Web Accessible Folder
WMS – Web Map Service

WP – Work Package

1.3.2 Notation

Several diagrams appearing in this document are presented using the UML notation. UML class diagram notation is briefly described below.

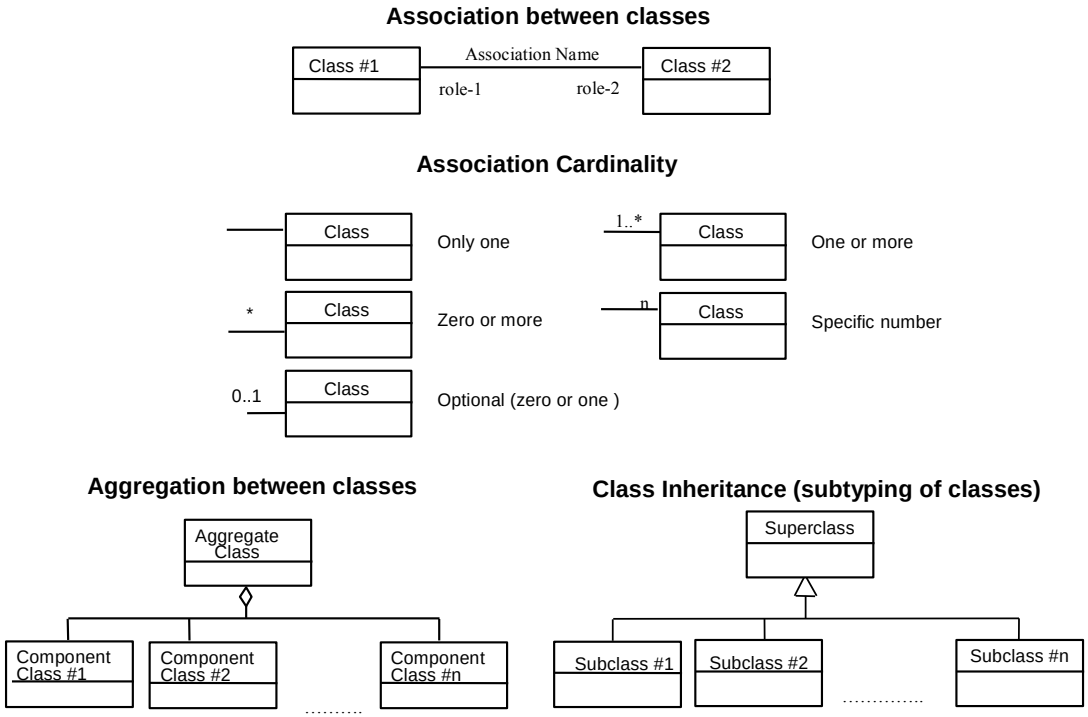

Figure 2 – UML class diagram notation

In this document, the following stereotypes of UML classes may be used:

a) <<Interface>> is a definition of a set of operations that is supported by objects having this interface. An Interface class cannot contain any attributes.
b) <<DataType>> is a descriptor of a set of values that lack identity (independent existence and the possibility of side effects). A DataType is a class with no operations whose primary purpose is to hold the information.
c) <<CodeList>> is a flexible enumeration that uses string values for expressing a list of potential values.
d) <<GUI>> is a Graphical User Interface, providing graphical controls for human interaction with underlying software instances.

In this document, the following standard data types may be used:

a) CharacterString – A sequence of characters.
b) Integer – An integer number.
c) Double – A double precision floating point number.
d) Float – A single precision floating point number.

2. Smart searchable interface component architecture

The GeoViQua enhanced geo-search components architecture illustrated in Figure 1 is a general representation of the components identified in the GeoViQua project. Figure 3 shows a more detailed view of the components that will be described in this deliverable: the Smart searchable interface components, which are contained in the GeoViQua Common Infrastructure.

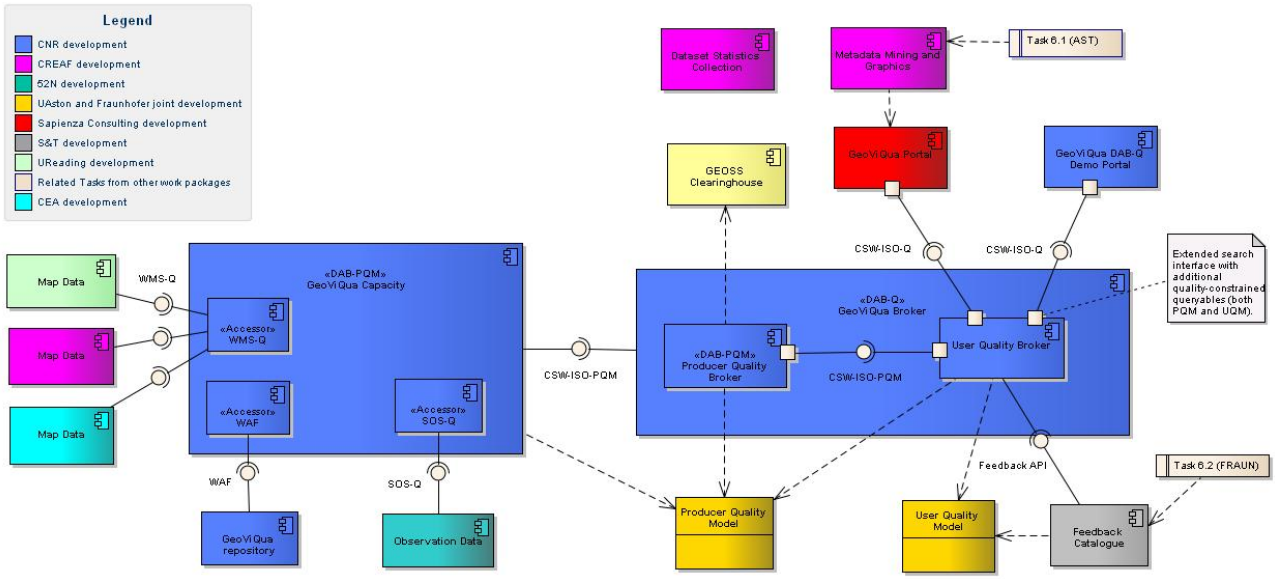

Figure 3 – Smart searchable interface components architecture

There are two main components in the diagram:

- The GeoViQua Broker (DAB-Q)
- The GeoViQua Capacity (DAB-PQM)

The next sections describe the above components and their internal sub-parts.

3. GeoViQua Broker

The GeoViQua Broker is an instance of the Quality-enabled Discovery and Access Broker (DAB-Q) service, developed in the context of GeoViQua project. The DAB-Q is an extension of the GEO DAB (formerly known as EuroGEOSS Broker) and enables smart search and discovery functionalities using parameters relevant to GEO products (date, scale/detail, bounding box, SBA, etc.) extended to include quality information.
Quality information can be stated both by data producers and by data users, resulting in two conceptually distinct data models, the Producer Quality Model (PQM) and the User Quality Model (UQM) (also known as User Feedback model). These models together form the GeoViQua Quality Model, now in its final version 4.0 (http://schemas.geoviqua.org).
The GeoViQua Broker homepage is available at the following reference endpoint:

http://geoviqua.essi-lab.eu/dabq-demo/

The GeoViQua Broker implements several internal components: the CSW/ISO-Q interface, the UQM and the PQM.

3.1 CSW/ISO-Q

CSW/ISO-Q is an extension of the standard OGC Catalog Service for the Web 2.0.2 ISO Application Profile (CSW ISO AP) to support quality-constrained queries. Following the standard CSW behavior, the CSW/ISO-Q interface requires explicit quality statements in order to match query constraints. In other words, datasets without quality statements will never be returned as a result of a quality-constrained query.
The Capabilities document of the CSW/ISO-Q interface is available at the following endpoint:

http://geoviqua.essi-lab.eu/dabq-demo/services/cswisoq?service=CSW&version=2.0.2&request=GetCapabilities

Table 1 summarizes the ISO standards quality indicators and its definition. For further information about the quality indicators considered by GeoViQua, see http://qualityml.geoviqua.org/.

Table 1 – Classification and definition of quality elements by ISO 19115

Quality Class	Quality indicator	Definition
Positional accuracy	Absolute or external positional accuracy	Closeness of reported coordinate values to values accepted as or being true
	Relative or internal accuracy	Closeness of the relative positions of features in a dataset to their respective relative positions accepted as or being true
	Gridded data positional accuracy	Closeness of gridded data position values to values accepted as or being true
Completeness	Commission	Excess data present in a dataset

Logical consistency	Omission	Data absent from a dataset
	Conceptual consistency	Adherence to rules of the conceptual schema
	Domain consistency	Adherence of values to the value domains
	Topological consistency	Correctness of the explicitly encoded topological characteristics of a dataset
	Format consistency	Degree to which data is stored in accordance with the physical structure of the dataset
Temporal accuracy	Accuracy of a time measurement	Correctness of the temporal references of an item (reporting of error in time measurement)
	Temporal consistency	Correctness of ordered events or sequences
	Temporal validity	Validity of data with respect to time
Thematic accuracy	Quantitative attribute accuracy	Accuracy of quantitative attributes
	Non-quantitative attribute correctness	Correctness of non-quantitative attributes
	Thematic classification correctness	Comparison of the classes assigned to features or their attributes to a universe of discourse (*e.g.,* ground truth or reference dataset)

Quality-constrained discovery is enabled by query constraints based on the above quality indicators. Such quality constraints are expressed against queryable properties (also referred to as "queryables"), which are selected concrete expressions of quality parameters/indicators. In this context, "selected" means that they are a subset of all the possible ones; "concrete" means that they are implementable in an efficient way. We distinguish between two categories of queryables: PQM and UQM queryables. The queryables, chosen with the GeoViQua partners, represent a significant set of the quality measures captured by the GeoViQua quality model.
Queryables are structured as follows (some parts may be omitted for some quality indicators):

- Name
- Related quality parameter

- Relevant quality model elements
- Default measurement description
- Type
- Units of measure
- Comment

The next paragraphs detail the PQM and UQM queryables and the related implementation in the Broker.

3.2 PQM Broker

The Producer Quality Model introduces indicators to record qualitative and quantitative quality information, and to identify resources (i.e. datasets) in order to relate metadata in hierarchical or other ways. The model extends ISO 19115, 19115-2 and 19157, adding means to report publications, discovered issues, reference datasets used for quality evaluation, traceability and statistical summaries of quantified uncertainty.

Table 2 defines the core PQM queryables. Only the following structural parts are shown: name, quality indicator, default measurement description and type.

Table 2 – Producer Quality Model queryables

Queryable Name	Quality parameter indicator	Default measurement description	Type
QualityInfoCount	Completeness	-	Integer
QualityInfoReportCount	Completeness	-	Integer
TCC.ResultValue	Thematic accuracy	Misclassification rate	Double
CO.ResultValue	Completeness	Number of missing items	Double

DC.ResultValue	Logical consistency	Value domain conformance rate	Double
PA.ResultValue	Positional accuracy	Root mean square error	Double
PublicationCount	Completeness	-	Integer
QAA.ResultValue	Thematic accuracy	Uncertainty at 68.3% significance level	Double
CC.ResultValue	Logical consistency	Number of items noncompliant to the rules of the conceptual schema	Double
MeasureDescription	-	-	String
LI.ProcessCount	-	Number of process steps cited	Integer
LI.SourceCount	-	Number of sources cited	Integer

LI.Process	-	Records that uses a process with the following description or identifier	String
LI.Source	-	Records that uses a source with the following title or identifier	String

The PQM Broker implements all the above queryables for quality-enabled query support. Clients can query the PQM Broker through the CSW/ISO-Q interface in order to retrieve quality and quantitative quality information.

Some simple PQM queries examples are reported below.

Example 1: Give me the first record having quality info count greater than or equal to 1.

```
<csw:GetRecords xmlns:csw="http://www.opengis.net/cat/csw/2.0.2"
        xmlns:dc="http://purl.org/dc/elements/1.1/" xmlns:dct="http://purl.org/dc/terms/"
        xmlns:ows="http://www.opengis.net/ows" xmlns:ogc="http://www.opengis.net/ogc"
        xmlns:gml="http://www.opengis.net/gml"
xmlns:xlink="http://www.w3.org/1999/xlink"
        xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" resultType="results"
        outputSchema="http://www.geoviqua.org/QualityInformationModel/4.0"
startPosition="1"
        maxRecords="1" service="CSW" version="2.0.2"
        xsi:schemaLocation="http://www.opengis.net/cat/csw/2.0.2
http://schemas.opengis.net/csw/2.0.2/CSW-publication.xsd">
        <csw:Query xmlns:gvq="http://www.geoviqua.org/QualityInformationModel/4.0"
                typeNames="gvq:GVQ_Metadata">
                <csw:ElementSetName
typeNames="gvq:GVQ_Metadata">full</csw:ElementSetName>
                <csw:Constraint version="1.1.0">
                        <ogc:Filter>
                                <ogc:PropertyIsGreaterThanOrEqualTo matchCase="true">

    <ogc:PropertyName>apiso:QualityInfoCount</ogc:PropertyName>
                                        <ogc:Literal>1</ogc:Literal>
                                </ogc:PropertyIsGreaterThanOrEqualTo>
                        </ogc:Filter>
                </csw:Constraint>
        </csw:Query>
</csw:GetRecords>
```

Example 2: Give me a maximum of five records starting from the first record, having value domain greater than 3 with a measure description equals to "Value domain conformance rate".

```
<csw:GetRecords xmlns:csw="http://www.opengis.net/cat/csw/2.0.2"
        xmlns:dc="http://purl.org/dc/elements/1.1/" xmlns:dct="http://purl.org/dc/terms/"
        xmlns:ows="http://www.opengis.net/ows" xmlns:ogc="http://www.opengis.net/ogc"
        xmlns:gml="http://www.opengis.net/gml"
xmlns:xlink="http://www.w3.org/1999/xlink"
        xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" resultType="results"
        outputSchema="http://www.geoviqua.org/QualityInformationModel/4.0"
startPosition="1"
        maxRecords="5" service="CSW" version="2.0.2"
        xsi:schemaLocation="http://www.opengis.net/cat/csw/2.0.2
http://schemas.opengis.net/csw/2.0.2/CSW-publication.xsd">
        <csw:Query xmlns:gvq="http://www.geoviqua.org/QualityInformationModel/4.0"
                typeNames="gvq:GVQ_Metadata">
                <csw:ElementSetName
typeNames="gvq:GVQ_Metadata">full</csw:ElementSetName>
                <csw:Constraint version="1.1.0">
                        <ogc:Filter>
                                <ogc:And>
                                        <ogc:PropertyIsGreaterThan>
                                                <ogc:PropertyName>apiso:DC.ResultValue
</ogc:PropertyName>
                                                <ogc:Literal>3</ogc:Literal>
                                        </ogc:PropertyIsGreaterThan>
                                        <ogc:PropertyIsEqualTo matchCase="true">
                                <ogc:PropertyName>apiso:MeasureDescription</ogc:PropertyName>
                                                <ogc:Literal>Value domain conformance
rate</ogc:Literal>
                                        </ogc:PropertyIsGreaterThanOrEqualTo>
                                </ogc:And>
                        </ogc:Filter>
                </csw:Constraint>
        </csw:GetRecords>
```

3.3 UQM Broker

The UQM describes the structure and attributes of comments, citations, discovered issues, ratings and reports of usage. It reuses some ISO quality and metadata elements, along with elements of the PQM, but is less strictly bound to existing ISO schemas.

Table 3 lists the core UQM queryables. Only the following structural parts are shown: name, quality parameter indicator, type and comment. In particular, we omit the default measure description element (shown for PQM queryables) as it is empty for all UQM queryables.

Table 3 – User Quality Model queryables

Queryable Name	Quality parameter indicator	Type	Comment
Rating	User quality	Integer	-
FieldPresence	Completeness	CodeList (userComment, usage, rating, citation, qualityOverride, domainURN, externalFeedback, tags)	Presence of feedback fields, e.g. user comment, usage, rating, citation, quality override, domain, reply_to, external feedback, tags. E.g. records with gvq:userComment OR with a gvq:domainURN.
SearchTerm	-	String	Internally in the feedback server an "OR" search is done
AverageRating	User quality	Double	Average of all rating score
ReportAspect	-	CodeList (usage, problem, alternative, fitness for purpose)	Category of the report. This can be composed with other queryables, e.g.: average of all rating score from usage by hidrologists.

FeedbackItemCount	Completeness	Integer	Count of feedback items
FeedbackDomain	-	URN	Domain(s) in which the feedback is deemed relevant
UserDomain	-	URN	User domain(s)
UserRole	-	CodeList	Role of the user when submitting the item

The UQM Broker implements only a subset of the above UQM queryable: AverageRating, Rating, Report Aspect, SearchTerm, UserDomain.

The UQM Broker associates feedback information to query results. Clients can query the UQM Broker through the CSW/ISO-Q interface for searching based on user-quality aspects of the resources/datasets.

Some simple UQM queries examples are reported below.

Example 1: Give me the first record having average rating greater than or equal to 2.

```
<csw:GetRecords xmlns:csw="http://www.opengis.net/cat/csw/2.0.2"
        xmlns:dc="http://purl.org/dc/elements/1.1/" xmlns:dct="http://purl.org/dc/terms/"
        xmlns:ows="http://www.opengis.net/ows" xmlns:ogc="http://www.opengis.net/ogc"
        xmlns:gml="http://www.opengis.net/gml" xmlns:xlink="http://www.w3.org/1999/xlink"
        xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" resultType="results"
        outputSchema="http://www.geoviqua.org/QualityInformationModel/4.0" startPosition="1"
        maxRecords="1" service="CSW" version="2.0.2"
        xsi:schemaLocation="http://www.opengis.net/cat/csw/2.0.2
http://schemas.opengis.net/csw/2.0.2/CSW-publication.xsd">
        <csw:Query xmlns:gvq="http://www.geoviqua.org/QualityInformationModel/4.0"
            typeNames="gvq:GVQ_Metadata">
            <csw:ElementSetName typeNames="gvq:GVQ_Metadata">full</csw:ElementSetName>
            <csw:Constraint version="1.1.0">
                <ogc:Filter>
```

```
                <ogc:PropertyIsGreaterThanOrEqualTo matchCase="true">
                    <ogc:PropertyName>apiso:AverageRating</ogc:PropertyName>
                    <ogc:Literal>2</ogc:Literal>
                </ogc:PropertyIsGreaterThanOrEqualTo>

            </ogc:Filter>
        </csw:Constraint>
    </csw:Query>
</csw:GetRecords>
```

Example 2: Give me the first record having average rating greater than or equal to 1 with problems reported by research end-users.

```
<csw:GetRecords xmlns:csw="http://www.opengis.net/cat/csw/2.0.2"
    xmlns:dc="http://purl.org/dc/elements/1.1/" xmlns:dct="http://purl.org/dc/terms/"
    xmlns:ows="http://www.opengis.net/ows" xmlns:ogc="http://www.opengis.net/ogc"
    xmlns:gml="http://www.opengis.net/gml" xmlns:xlink="http://www.w3.org/1999/xlink"
    xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" resultType="results"
    outputSchema="http://www.geoviqua.org/QualityInformationModel/4.0" startPosition="1"
    maxRecords="1" service="CSW" version="2.0.2"
    xsi:schemaLocation="http://www.opengis.net/cat/csw/2.0.2
http://schemas.opengis.net/csw/2.0.2/CSW-publication.xsd">
    <csw:Query xmlns:gvq="http://www.geoviqua.org/QualityInformationModel/4.0"
        typeNames="gvq:GVQ_Metadata">
        <csw:ElementSetName typeNames="gvq:GVQ_Metadata">full</csw:ElementSetName>
        <csw:Constraint version="1.1.0">
            <ogc:Filter>
                <ogc:And>
                    <ogc:PropertyIsGreaterThanOrEqualTo matchCase="true">
                        <ogc:PropertyName>apiso:AverageRating</ogc:PropertyName>
                        <ogc:Literal>1</ogc:Literal>
                    </ogc:PropertyIsGreaterThanOrEqualTo>
                    <ogc:PropertyIsEqualTo matchCase="true">
                        <ogc:PropertyName>apiso:ReportAspect</ogc:PropertyName>
                        <ogc:Literal>Problem</ogc:Literal>
                    </ogc:PropertyIsEqualTo>
                </ogc:And>
            </ogc:Filter>
        </csw:Constraint>
    </csw:Query>
</csw:GetRecords>
```

It is also possible make queries with PQM and UQM queryables together, as in the following example, where we search the first record having quality info count greater than or equal to 1 and average rating greater than or equal to 2:

```
<csw:GetRecords xmlns:csw="http://www.opengis.net/cat/csw/2.0.2"
    xmlns:dc="http://purl.org/dc/elements/1.1/" xmlns:dct="http://purl.org/dc/terms/"
    xmlns:ows="http://www.opengis.net/ows" xmlns:ogc="http://www.opengis.net/ogc"
```

```
    xmlns:gml="http://www.opengis.net/gml" xmlns:xlink="http://www.w3.org/1999/xlink"
    xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" resultType="results"
    outputSchema="http://www.geoviqua.org/QualityInformationModel/4.0" startPosition="1"
    maxRecords="1" service="CSW" version="2.0.2"
    xsi:schemaLocation="http://www.opengis.net/cat/csw/2.0.2
http://schemas.opengis.net/csw/2.0.2/CSW-publication.xsd">
    <csw:Query xmlns:gvq="http://www.geoviqua.org/QualityInformationModel/4.0"
      typeNames="gvq:GVQ_Metadata">
      <csw:ElementSetName typeNames="gvq:GVQ_Metadata">full</csw:ElementSetName>
      <csw:Constraint version="1.1.0">
        <ogc:Filter>
          <ogc:And>
            <ogc:PropertyIsGreaterThanOrEqualTo matchCase="true">
              <ogc:PropertyName>apiso:QualityInfoCount</ogc:PropertyName>
              <ogc:Literal>1</ogc:Literal>
            </ogc:PropertyIsGreaterThanOrEqualTo>
            <ogc:PropertyIsGreaterThanOrEqualTo
            matchCase="true">
              <ogc:PropertyName>apiso:AverageRating</ogc:PropertyName>
              <ogc:Literal>2</ogc:Literal>
            </ogc:PropertyIsGreaterThanOrEqualTo>
          </ogc:And>
        </ogc:Filter>
      </csw:Constraint>
    </csw:Query>
  </csw:GetRecords>
```

3.4 DAB Response

The DAB-Q interface can be interrogated to return the ISO19115 or a GeoViQua quality model. To select between them, the right outputSchema parameter has to be provided.
To get an ISO 19115 document you have to specify:

outputSchema="http://www.isotc211.org/2005/gmd"

To get a GeoViQua quality model you have to specify:

outputSchema=http://www.geoviqua.org/QualityInformationModel/4.0

This is a schema fragment of the request for a dataset named “test_data_1”
http://geoviqua.essi-lab.eu/dabq-demo/services/cswisoq?service=CSW&request=GetRecordById&id=mtri2an1ib&outputschema=http://www.geoviqua.org/QualityInformationModel/4.0&elementSetName=full

The response is structured in 2 parts: The Producer Quality Model and the User Feedback Model.
Header part:

```
<csw:GetRecordByIdResponse>
  <gvq:GVQ_Metadata>
    <gmd:fileIdentifier xmlns:gmd="http://www.isotc211.org/2005/gmd">
      <gco:CharacterString>mtri2an1ib</gco:CharacterString>
    </gmd:fileIdentifier>
```

....

Producer Quality Model part

```
<gvq:dataQualityInfo>
  <gvq:GVQ_DataQuality>
    <gmd19157:scope xlink:href="#datasetScope"/>
    <gvq:discoveredIssue>
      <gvq:GVQ_DiscoveredIssue>
        <gvq:target xlink:href="#mtri2an1ib"/>
        <gvq:knownProblem>
          <gco:CharacterString>Some methodological constraints can be
defined</gco:CharacterString>
        </gvq:knownProblem>
        <gvq:workAround>
          <gco:CharacterString>First of all be careful with extrapolated areas (checking
them in the provided binary mask). Moreover, be careful with areas with high uncertainty (checking
them in the provided residuals map).</gco:CharacterString>
        </gvq:workAround>
      </gvq:GVQ_DiscoveredIssue>
    </gvq:discoveredIssue>
  </gvq:GVQ_DataQuality>
</gvq:dataQualityInfo>
```

User Feedback Model part

```
<gvq:userFeedback>
  <gvq:GVQ_FeedbackCollection>
    <gvq:itemUnderReview>
      <updated19115:MD_Identifier>
        <gmd:code
                                    xmlns:gmd="http://www.isotc211.org/2005/gmd">
          <gco:CharacterString>mtri2an1ib
                                    </gco:CharacterString>
        </gmd:code>
...
      </updated19115:MD_Identifier>
    </gvq:itemUnderReview>
    <gvq:item>
      <gvq:subject>Digital Climatic Atlas</gvq:subject>
      <gvq:userRole>
        <gvq:GVQ_UserRoleCode
codeList="http://schemas.geoviqua.org/GVQ/4.0/resources/Codelist/gvqCodelists.xml#GVQ_User
RoleCode"
codeListValue="NonResearchEndUser">NonResearchEndUser</gvq:GVQ_UserRoleCode>
      </gvq:userRole>
      <gvq:dateStamp>
        <gco:DateTime>2013-12-05T12:05:52.466000+00:00
                                    </gco:DateTime>
      </gvq:dateStamp>
      <gvq:userComment>
```

```
                <gvq:comment>Second test comment on 'Mean Annual Temperature'
                </gvq:comment>
                <gvq:mime-type>plain/text</gvq:mime-type>
            </gvq:userComment>
            <gvq:rating>
                <gvq:score>4</gvq:score>
                <gvq:justification> </gvq:justification>
            </gvq:rating>
        </gvq:item>
    </gvq:GVQ_FeedbackCollection>
</gvq:userFeedback>
</gvq:GVQ_Metadata>
</csw:GetRecordByIdResponse>
```

4. GeoViQua Capacity

The GeoViQua Capacity is a component of the smart searchable interface used to connect several services: WMS-Q, WAF and other community catalogues (e.g. the GEOSS Clearinghouse).
Specific components, termed *accessors*, implement mediation services for interfacing heterogeneous service providers that expose multiple standard specifications. These mediating components map the heterogeneous provider metadata models into a uniform data model which implements ISO 19115 based on ISO 19139 schemas.
The GeoViQua Capacity is mainly composed by two internal components: WMS-Q accessor and WAF accessor.

4.1 WMS-Q accessor

WMS-Q is a convention for encoding quality information in a Web Map Service. WMS-Q is primarily concerned with conveying quality information at the level of variables representing fields or categorical values and also individual pixels (or features for the vector case). After the first publication as a OGC engineering report, GeoViQua has reviewed and perfected it and released the WMS v.2 internally and will be delivered as a new document later. This accessor is basedin WMS-Q v2.

In standard WMS, the essential unit of information is the Layer. Layers that have a *Name* element can be displayable (i.e. can be requested in a GetMap operation) whereas Layers that don't have a *Name* are not displayable. Layers can be organized hierarchically (i.e. can be arbitrarily nested) and any Layer can be displayable, even if that Layer contains child Layers.

The WMS specification allows metadata to be attached to Layers within a WMS, using the *MetadataURL* element of the WMS capabilities document. This URL points to an online resource that resolves to a separate document that provides metadata about the Layer. However the limitation of WMS 1.3 is that the *Type* attribute of the *MetadataURL* element only supports ISO 19115 and FGDC-STD-001-1998.
The idea of WMS-Q derives from the need to extend the standard WMS expanding the range of values permitted by the *Type* attribute of the *MetadataURL* element. In particular these types should also include the ISO 19157 and the GeoViQua quality models.

The WMS-Q v2 components developed are described in Deliverable 5.4 Final versions of quality-aware visualisation components.

4.1.1 Mapping description

The purpose of the WMS-Q accessor is to map the WMS-Q Layers into ISO datasets according to specific requirements described in the WMS-Q version 2.0 [GVQ-WMS-Q].

The current WMS accessor of the DAB maps all the Layers of the Capabilities document. The bottom-level Layers are mapped into the ISO concept of dataset and all the Layers above in the hierarchy are mapped to the concept of dataset collection.

A variable in a WMS-Q shall always map to an ISO concept of a dataset, so there will be no dataset collection.

In WMS-Q, not all Layers will necessarily have quality information attached. It is allowed to mix quality-enabled and regular Layers in the same WMS service instance.

In the WMS-Q accessor only the variable Layers will be mapped. To identify a variable Layer, there are two possible cases to consider: variable Layer with pixel level quality components (that may also have quality at the dataset level) and variable Layer without pixel level components (only dataset level quality).

A variable Layer with pixel level components has the following properties:

- *Keyword* element "http://qualityml.geoviqua.org/1.0/qualityCollection";
- *metadataURL* element pointing to ISO or GeoViQua extension quality documents;
- children representing the components;
- NO *Name* element available.

A variable Layer without pixel level components has the following properties:

- NO *Keyword* "http://qualityml.geoviqua.org/1.0/qualityCollection";
- *metadataURL* element pointing to ISO or GeoViQua extension quality documents;
- NO children;
- *Name* element available;
- NO parent being a variable Layer.

The following code extracted from the CREAF-Miramon WMS-Q Capabilities document represents a variable Layer with pixel level components:

```
<Layer>
        <Title>Annual Air Temperature of Iberian Peninsula</Title>
        <Abstract>Annual climatic digital maps of air temperature of Iberian Peninsula. Maps
          have been generated by means of statistical techniques, Geographic Information
          Systems and spatial interpolation.</Abstract>
        <KeywordList>
          <Keyword
            vocabulary="http://www.isotc211.org/2005/resources/codeList.xml#MD_KeywordTypeCode/theme"
            >Temperature</Keyword>
          <Keyword
            vocabulary="http://www.isotc211.org/2005/resources/codeList.xml#MD_KeywordTypeCode/theme"
            >Spatial interpolation</Keyword>
          <Keyword
            vocabulary="http://www.isotc211.org/2005/resources/codeList.xml#MD_KeywordTypeCode/place"
            >Iberian Peninsula</Keyword>
          <Keyword
            vocabulary="http://www.uncertml.org/distributions/MultivariateNormalDistribution"
            >multivariate regression</Keyword>
          <Keyword vocabulary="http://www.uncertml.org/statistics/Mean">annual
            mean</Keyword>
          <Keyword
            vocabulary="http://qualityml.geoviqua.org/1.0/measures/ExtrapolatedAreas"
            >uncertainty areas</Keyword>
          <Keyword vocabulary="http://qualityml.geoviqua.org/1.0/"
            >qualityCollection</Keyword>
        </KeywordList>
        <Layer queryable="1" opaque="1">
          <Name>AnnualTemperature</Name>
          <Title>Annual Air Temperature of Iberian Peninsula</Title>
          <Abstract>Annual climatic digital maps of temperature. This layer is an
```

aggregator for all the components of the temperature value. Maps have been generated by means of statistical techniques, Geographic Information Systems and spatial interpolation.</Abstract>
<KeywordList>
<Keyword vocabulary="http://www.uncertml.org/statistics/mean">mean annual</Keyword>
<Keyword vocabulary="http://qualityml.geoviqua.org/1.0/">qualityCollection</Keyword>
</KeywordList>
<EX_GeographicBoundingBox>
<westBoundLongitude>-9.340148</westBoundLongitude>
<eastBoundLongitude>3.858061</eastBoundLongitude>
<southBoundLatitude>35.697178</southBoundLatitude>
<northBoundLatitude>43.731057</northBoundLatitude>
</EX_GeographicBoundingBox>
<BoundingBox CRS="EPSG:23030" minx="-74000.00" miny="3969000.00" maxx="1052400.00" maxy="4865000.00" resx="100.00" resy="100.00"/>
<MetadataURL type="ISO19115:2003">
<Format>application/xml</Format>
<OnlineResource xmlns:xlink="http://www.w3.org/1999/xlink" xlink:type="simple" xlink:href="http://www.ogc.uab.cat/cgi-bin/GeoViQUA/WMSQ/MiraMon.cgi?SERVICE=CSW&REQUEST=GetRecordById&OUTPUTSCHEMA=http://www.isotc211.org/2005/gmd&ELEMENTSETNAME=full&id=AnnualTemperature:EPSG:23030&LANGUAGE=eng&OUTPUTFORMAT=application/xml" />
</MetadataURL>
</Layer>
<Layer queryable="1" opaque="1">
<Name>MeanAnnualTemperature</Name>
<Title>Mean Annual Air Temperature</Title>
<Abstract>Annual climatic digital maps of mean temperature. Maps have been generated by means of statistical techniques, Geographic Information Systems and spatial interpolation.</Abstract>
<KeywordList>
<Keyword vocabulary="http://www.uncertml.org/statistics/mean">mean annual</Keyword>
</KeywordList>
<EX_GeographicBoundingBox>
<westBoundLongitude>-9.340148</westBoundLongitude>
<eastBoundLongitude>3.858061</eastBoundLongitude>
<southBoundLatitude>35.697178</southBoundLatitude>
<northBoundLatitude>43.731057</northBoundLatitude>
</EX_GeographicBoundingBox>
<BoundingBox CRS="EPSG:23030" minx="-74000.00" miny="3969000.00" maxx="1052400.00" maxy="4865000.00" resx="100.00" resy="100.00"/>
<MetadataURL type="ISO19115:2003">
<Format>application/xml</Format>
<OnlineResource xmlns:xlink="http://www.w3.org/1999/xlink" xlink:type="simple" xlink:href="http://www.ogc.uab.cat/cgi-bin/GeoViQUA/WMSQ/MiraMon.cgi?SERVICE=CSW&REQUEST=GetRecordById&OUTPUTSCHEMA=http://www.isotc211.org/2005/gmd&ELEMENTSETNAME=full&id=MeanAnnualTemperature:EPSG:23030&LANGUAGE=eng&OUTPUTFORMAT=application/xml" />
</MetadataURL>
</Layer>
<Layer opaque="1">
<Name>ExtrapolationAreas</Name>

```
    <Title>Extrapolation areas</Title>
    <Abstract>Extrapolation areas of the regression model, these are uncertain areas
      and thus indicate the place where extrapolation of the regression model is
      applied(i.e. prediction of values without adjustment points within the range
      or area).</Abstract>
    <KeywordList>
      <Keyword
        vocabulary="http://qualityml.geoviqua.org/1.0/measures/ExtrapolatedAreas"
        >Extrapolated Areas</Keyword>
    </KeywordList>
    <EX_GeographicBoundingBox>
      <westBoundLongitude>-9.340148</westBoundLongitude>
      <eastBoundLongitude>3.858061</eastBoundLongitude>
      <southBoundLatitude>35.697178</southBoundLatitude>
      <northBoundLatitude>43.731057</northBoundLatitude>
    </EX_GeographicBoundingBox>
    <BoundingBox CRS="EPSG:23030" minx="-74000.00" miny="3969000.00"
      maxx="1052400.00" maxy="4865000.00" resx="100.00" resy="100.00"/>
  </Layer>
  <Layer opaque="1">
    <Name>ResidualSpatialRegression</Name>
    <Title>Residual spatial regression</Title>
    <Abstract>Detection of cells that have different weather from the predicted
      averages conditions. Integrates both the uncertainty caused by the
      unsatisfactory modeling as well as the natural uncertanty of areas with
      particular climatic conditions.</Abstract>
    <KeywordList>
      <Keyword vocabulary="http://www.uncertml.org/statistics/ResidualAccuracy"
        >Residual regression</Keyword>
    </KeywordList>
    <EX_GeographicBoundingBox>
      <westBoundLongitude>-9.340148</westBoundLongitude>
      <eastBoundLongitude>3.858061</eastBoundLongitude>
      <southBoundLatitude>35.697178</southBoundLatitude>
      <northBoundLatitude>43.731057</northBoundLatitude>
    </EX_GeographicBoundingBox>
    <BoundingBox CRS="EPSG:23030" minx="-74000.00" miny="3969000.00"
      maxx="1052400.00" maxy="4865000.00" resx="100.00" resy="100.00"/>
  </Layer>
  <Layer opaque="1">
    <Name>StabilityTestAreas</Name>
    <Title>Map of test stability areas (discrete standard deviation)</Title>
    <Abstract>Map of areas obtained from a stability test in which 12 maps have been
      generated using different sets of tuning (60%) and test (40%) stations, and
      it has also been obtained a discretized standard deviation where the value 1
      represents areas with less deviation and value 2 represents areas of greater
      deviation.</Abstract>
    <KeywordList>
      <Keyword vocabulary="http://www.uncertml.org/statistics/deviation">stability
        test</Keyword>
    </KeywordList>
    <EX_GeographicBoundingBox>
      <westBoundLongitude>-9.340148</westBoundLongitude>
      <eastBoundLongitude>3.858061</eastBoundLongitude>
      <southBoundLatitude>35.697178</southBoundLatitude>
      <northBoundLatitude>43.731057</northBoundLatitude>
    </EX_GeographicBoundingBox>
    <BoundingBox CRS="EPSG:23030" minx="-74000.00" miny="3969000.00"
      maxx="1052400.00" maxy="4865000.00" resx="100.00" resy="100.00"/>
```

```
    </Layer>
</Layer>
```

If there is not *metadataURL* element associated with main Layer, we consider the *metadataURL* element of the first child Layer. The ISO 19115 dataset representing the Layer is built from WMS accessor, taking into account the *OnlineResource* element located inside the *metadataURL* element. In particular the created dataset is a copy of the online resource metadata document, in which only element compliant with ISO 19115 are mapped.

The variable Layer without pixel level components is mapped in the same way.

At the moment of writing this document only CREAF-Miramon WMS-Q instance is fully compliant with the WMS-Q v2 conventions and the WMS-Q accessor already federates it. It is possible to discover the correspondent datasets by means of the "WMS-Q" keyword. UREAD and CEA WMS-Q instances are not yet fully WMS-Q v2 compliant, so they are federated as standard WMS.

4.2 WAF accessor

A Web Accessible Folder (WAF) is an HTTP or FTP accessible directory of files, typically metadata files in XML format in which all files and their time-stamps are visible to a web browser or client. It is a simple way to share metadata documents or data by exposing them in a web folder.

The GeoViQua WAF currently used by the DAB-Q is published at:

http://essi-lab.eu/gvq/waf4/.

It is accessible for reading and writing files via FTP.

The WAF accessor simply reads the XML files in the folder and tries to map them into the ISO 19115 standard. In the context of GeoViQua the WAF contains metadata based on the PQM, which will be mapped into a uniform data model implementing ISO 19115, based on official ISO 19139 schemas.

Currently there are only two XML files: GeometricCorrections_GCP and PureISO19115 because the other generated examples where moved to WMS-Q Metadata URL elements. Additional metadata records based on the GeoViQua PQM (i.e. GVQ_Metadata element) can be published by simply copying them to the GeoViQua WAF.

5. Conclusions

This deliverable is directly related with Deliverable D4.1 "Graphical search interface report". In D4.1 the focus was the Quality-aware catalogue client (client side), while in this deliverable we concentrate on the description of the Quality-aware catalogue service (server side).
We have designed and developed a smart searchable interface for improved quality-enabled discovery of geospatial resource in the context of GEOSS.
In particular we presented the GeoViQua Broker and the GeoViQua Capacity components for supporting data users and data providers in quality enabled filtering and querying.

5.1 Future work

By the end of the project, the smart searchable interface will be improved with the integration of the GeoViQua label for supporting additional (optional) query constraints. For details on GeoViQua label see [GVQ-D6.2] and visit http://www.geolabel.info/.

The integration of the GVQ label as queryable requires the insertion of the *graphicOverview* element in the datasets, as in the following example:

```
…
<gmd:graphicOverview>
  <gmd:MD_BrowseGraphic>
    <gmd:fileName>
      <gmx:FileName src="ADD_URL _HERE">GeoViQua.GeoLabel</gmx:FileName>
    </gmd:fileName>
    <gmd:fileDescription>
      <gco:CharacterString>A GEOLabel with drill-down capabilities for this
dataset</gco:CharacterString>
    </gmd:fileDescription>
    <gmd:fileType>
      <gmx:MimeFileType type="image/svg+xml">Scalar Vector
Graphics</gmx:MimeFileType>
    </gmd:fileType>
  </gmd:MD_BrowseGraphic>
</gmd:graphicOverview>
…
```

The *src* attribute of *FileName* element should contain the GVQ service URL with metadata or feedback as query URL fields (e.g: http://www.geolabel.net/api/v1/geolabel?metadata=http://www.creaf.uab.es/temp/AgriculturalDrought19139.xml).

6. References

[GVQ-DOW 2010]
GeoViQua Consortium, "GeoViQua - Annex I – 'Description of Work'", 2010.

[GVQ-D4.1]
GeoViQua Consortium, Deliverable D4.1 "Graphical search interface report".

[GVQ-D6.1]
GeoViQua Consortium, Deliverable D6.1 "Best practice document for quality encodings".

[GVQ-D6.2]
GeoViQua Consortium, Deliverable D6.2 "GEO Label description final document".

[OGC, 2012]
Open Geospatial Consortium, "OGC® OWS 9 Data Quality and Web Mapping Engineering Report, OGC 12-160r1, http://www.opengis.net/def/doc-type/per/ows9-data_quality_web_mapping.

[GVQ-WMS-Q]
GeoViQua Consortium, "GeoViQua -WMS-Q 2.0 document with a list of requirements, 2013.