GeoViQua
Components and Achievements

Tuesday, January 14th 2014

The research leading to these results has received funding from the European Union Seventh Framework Programme (FP7/2010-2013) under grant agreement no. 265178.
This is a GeoViQua initiative

GeoViQua provides a set of scientifically developed software components and services that facilitate the creation, search and visualization of quality information on EO data integrated and validated in the GEOSS Common Infrastructure.
The team

Universities
University of Reading
Aston University

Research centers

Small or Medium Enterprises
s&i
dependable solutions
52north
extploring horizons

International org.
OGC
Open Geospatial Consortium
GeoViQua (265178)

QUALity aware

VIssualisation for the

Global Earth Observation system of systems

FP7 Collaborative Project Environment
10 partners with CREA coordinating

Duration: 36 months
1 February 2011 – 31 January 2014

Budget: 4.024.256,42 €
EC Contribution: 3.266.803,98 €

2 reporting periods:
1. Months 1-18
2. Months 19-36
Elements that are not in the normal portal

- **Documentation phase**
  - No **quality** focused metadata **model**. No pixel level quality.

- **Discovery phase**:
  - **Query**
    - It is not possible to use a quality indicator as a way to **refine** the search
    - No **user feedback** consideration
  - **Response**
    - Quality is not a criteria for **ranking** results
    - Quality is not emphasize in the results
    - Metadata can not be easily **compared**
    - No visual representation of the metadata **completeness**

- **Evaluation phase**
  - No pixel level quality **visualization** strategies

- **Access phase**
  - No data access with **pixel level quality**

- **Interaction phase**
  - No immediate **user feedback**
Embedding quality in all phases of the geospatial data exploitation

GeoViQua Producer and User Quality Model, UncertML & QualityML

DAB-Q: Query with quality information

WMS-Q extension to relate data with spatialized quality layers

GEOLabel Graphical picture that helps to evaluate data

WCS-Q: GMLCov extension for spatialized quality layers (in project)

User feedback system has been introduced

Document Discover Evaluate Access Interact
Assessing fitness-for purpose

- peer and expert review
- better traceability and provenance information
- information on citations and usage of a dataset
- warnings about problems identified with a dataset and potential workarounds,
- ‘soft knowledge’ from data producers (e.g. recommendations for use which are not easily encoded using the existing standards)

Chrisman, 1988: asked for ‘user experience’ to be included in data quality specs
Epstein et al 1998: disclaimers only cover ‘reasonably foreseeable use’
Comber et al., 2006: identify mismatches between producer/user ontologies
Devillers et al., 2007: tools needed to help users understand quality information.
Boin & Hunter, 2007: users want a simple summary that they can interrogate for more detail as appropriate.
Lush et al., 2011: Interviews identified that these needs are still not being met.
**Producer quality model**
- ISO 19115 / 19157
- Extensions to the GCI:
- Citations
- Reference datasets
- Traceability
- Discovered Issues
- Pixel and dataset level – UncertML
- INSPIRE-recommended identifiers and codelists

**User Quality model**
- User feedback
- More flexible
- New model - not present in the GCI
- Ratings
- Examples of use in named application domains
- ‘Soft knowledge’ from producers
European Commission Speakers Corner

Identification information::MD_Usage
+ specificUsage :CharacterString
+ usageDateTime :DateTime [0..1]
+ userDeterminedLimitations :CharacterString [0..1]
+ userContactInfo :CL_ResponsibleParty [1..*]

GVQ_Usage
+ referenceDoc :GVQ_Publication [0..*]
+ response :CharacterString [0..1]

For publications using this dataset (e.g., deriving new data)

Identification information::MD_Dataldentityfication
+ spatialRepresentationType :MD_SpatialRepresentationTypeCode [0..1]
+ spatialResolution :MD_Resolution [0..1]
+ language :CharacterString [1..*]
+ characterSet :MD_CharacterSetCode [0..1] = "utf8"
+ topicCategory :MD_TopicCategoryCode [1..*]
+ environmentDescription :CharacterString [0..1]
+ extent :EX_Extent [0..*]
+ supplementalInformation :CharacterString [0..1]

GVQ_Dataldentityfication
+ referenceDoc :GVQ_Publication [1..*]

For publications describing this dataset (e.g., assessing or quantifying its characteristics). This member is already included in the new 19115-1 MD_Leantification named “additionalDocumentation”

Data quality information::LI_Lineage
+ statement :CharacterString [0..1]

GVQ_Lineage
+ referenceDoc :GVQ_Publication [1..*]

For publications about the methodology used to get this dataset (e.g., algorithms, models...)

Documents describing this dataset

Title: A new land-cover map of Africa for the year 2000, DOI: 10.1111/j.1365-2699.2004.01073.x
Category: journalArticle

Citations
Reference datasets
Populating the producer model

- Geonetwork plugin to create, edit and publish GeoViQua-compliant metadata
User feedback model

- ** GVQ_FeedbackTarget**: 0..1
  - parent : GVQ_FeedbackTarget
  - resourceRef : MD_Identifier
- ** GVQ_FeedbackGroup**: 1..*
  - timestamp : CI_Date
  - user : GVQ_UserInformation
  - roles : GVQ_UserRoleCodeEnum [1..*]
- ** GVQ_UserInformation**: + ratingValue : int
  - user : CI_RespParty [0..1]
  - applicationDomain : string [0..*] [ordered]
  - expertiseLevel : int
- ** GVQ_FeedbackItem**: + primaryFocus
  - secondaryFoci
  - supplementaryFoci
- ** GVQ_FeedbackFocusType**: +primaryFocus
- ** GVQ_ExternalFeedback**: + items 1..*
  - resourceURL : String
  - mime : String
- ** GVQ_QualityOverride**: + alternativeDataQualityEstimate : DQ_DataQuality
- ** GVQ_Rating**: + ratingValue : int
- ** GVQ_UserComment**: - comment : String
  - mime-type : String = text/plain
- ** GVQ_UsageReport**: + use-as : GVQ_ReportAspectCode [0..*]
  + citation : CI_Citation [0..1]
  + usageDescription : string
  «XSDelement»
  + alternativeDatasets : MD_Identifier [0..-1]
User feedback

—Mandatory information on user role (application domain, expertise level etc).

—Subject, application domain and keyword tags

—Other optional information such as rating, comments, a quality override (superseding producer quality information), a report of usage or a citation.

—Focus -spatial, temporal or other subsets to which feedback pertains
Populating the user model
DAB-Q capabilities

• Feedback Catalog integration
  – Discover of user feedback items to the Feedback Catalog
  – Upload of user feedback items to the Feedback Catalog
  – Merging of discovered metadata records with related user feedback items

• Support to quality-related constraints
  – Producer Quality Model constraints (e.g.: misclassification rate, count of data quality info elements, etc...)
  – User Quality Model constraints (e.g.: rating score, user domain, etc...)

• WMS-Q integration

• Insertion of GEO label link to all discovered records
CSW-Q extension

**Producer Quality model queryables**
- Count of data quality info elements
- Count of report elements
- Misclassification rate
- Number of missing items
- Domain conformance rate
- Positional accuracy value
- Count of publication elements
- Uncertainty level value
- Number of items noncompliant to the rules of the conceptual schema
- Measurement method
- Number of process steps cited
- Number of sources cited
- Records that uses a process with the following description or identifier
- Records that uses a source with the following title or identifier

**User Quality model queryables**
- Rating score
- Presence/absence of feedback fields (e.g. user comment, usage, rating, citation,...)
  - Any text
- Average of all rating score
  - Category of the report
- Count of feedback items
  - Domain in which the feedback is deemed relevant
  - User domain
  - Role of the user when submitting the item
Components diagram
GEOSS Portal Integration
**Metadata comparison. GeoViQua Project**

<table>
<thead>
<tr>
<th>Identification</th>
<th>Dataset 1</th>
<th>Dataset 2</th>
<th>Dataset 3</th>
<th>Dataset 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Title</td>
<td>Digital topographic Karte 1: 10 000 - Vorläufige Ausgabe 4159-NW Lette</td>
<td>Digital topographic Karte 1: 10 000 - Vorläufige Ausgabe 4159-NW Lette</td>
<td>Horizontal Positional Accuracy</td>
<td>Horizontal Positional Accuracy</td>
</tr>
<tr>
<td>Abstract</td>
<td>Die verlässigen Aussagen bestehen aus georeferenzierten Rasterdaten der gesamten topographischen Karte 1: 10 000 (TK10). Zudem sind die DTK aus den Basis-DLM abgeleitet. Die Daten sind in Ebenen (Geschoss) im wechselseitigen Deckelverhältnis (entsprechend) oder kombiniert aus diesen als Summenlagerung verfügbar. Sie liegen in 5 Inhaberbeben vor: Situation Schwarzwald, Waldflächen, Gewässerschutz, Gewässerflächen, Relief</td>
<td>Summenlagerung verfügbar. Sie liegen in 5 Inhaberbeben vor: Situation Schwarzwald, Waldflächen, Gewässerschutz, Gewässerflächen, Relief</td>
<td>Resolution as reported</td>
<td>Resolution as reported</td>
</tr>
</tbody>
</table>

**DQ Element**

<table>
<thead>
<tr>
<th>Scope</th>
<th>Dataset 1</th>
<th>Dataset 2</th>
<th>Dataset 3</th>
<th>Dataset 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Measure Name: Horizontal</td>
<td>Horizontal</td>
<td>Horizontal</td>
<td>Horizontal Positional Accuracy</td>
<td>Horizontal Positional Accuracy</td>
</tr>
<tr>
<td>Measure Description:</td>
<td></td>
<td></td>
<td>Resolution as reported</td>
<td>Resolution as reported</td>
</tr>
<tr>
<td>Measure Identification:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Evaluation Method Type:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Positional Accuracy</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Absolute External Positional Accuracy:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Metadata comparison**

![Metadata comparison diagram](image-url)
Semantic groupings of WMS Layers

• We need a method to convey that individual Layers are related semantically
  – E.g. one Layer represents the variance of another Layer

• We use Layer nesting + coupled with keywords from the UncertML vocabulary

• See fragment of Capabilities document (right, simplified)
  – In this example uncertainties are normally distributed

• Also could be applied to other kinds of semantic groupings
  – E.g. components of a velocity field

```xml
<Layer>
  <!-- Non-displayable container -->
  <Title>Sea Surface Temperature</Title>
  <KeywordList>
    <Keyword vocabulary="http://uncertml.org/distributions">normal</Keyword>
  </KeywordList>
</Layer>

<Layer>
  <Name>sst</Name>
  <Title>Sea Surface Temperature Mean</Title>
  <KeywordList>
    <Keyword><normal</Keyword>
  </KeywordList>
</Layer>
```
Styling of Layers

• There are many different ways of representing uncertainties visually:
  – Contours, textures, shading, transparency, bivariate colour maps...
• Different methods suit different datasets and users
  – WMS provides two methods:
    – Named Styles – simple but inflexible
    – Styled Layer Descriptors and Symbology – more flexible but still rather basic for raster data
• ncWMS provides some simple extensions to WMS
• None of these meet the use cases for visualization of uncertainty
• Hence we have developed a new XML language for specifying styles for raster data
  – Named styles can map to WMS definitions for backward compatibility

Contours, textures, shading
Published by OGC as Public Eng. Report

- **OGC® OWS 9 Data Quality and Web Mapping Engineering Report** 12-160r1 Jon Blower, Xiaoyu Yang, Joan Masó and Simon Thum, 2013-06-18

- This Engineering Report specifies conventions for conveying information about data quality through the OGC Web Map Service Standard (known hereafter as the “WMS-Q conventions”), OGC Web Map Tile Service Standard (known hereafter as the “WMTS-Q conventions”), OGC KML (known hereafter as the “KML-Q conventions”) and OGC Augmented Reality Markup Language.

- [https://portal.opengeospatial.org/files/?artifact_id=52884](https://portal.opengeospatial.org/files/?artifact_id=52884)
MiraMon WMS-Q client and server

- [http://www.ogc.uab.cat/geoviqua/wmsq](http://www.ogc.uab.cat/geoviqua/wmsq)
- Three different layers with several quality indicators

Ebro delta rice fields flooding maps

Catalan Landsat land use classification

Iberian annual air temperature

- Service GetCapabilities
  - [http://www.ogc.uab.cat/cgi-bin/GeoViQUA/WMSQ/MiraMon.cgi?REQUEST=GetCapabilities&VERSION=1.3.0&SERVICE=WMS](http://www.ogc.uab.cat/cgi-bin/GeoViQUA/WMSQ/MiraMon.cgi?REQUEST=GetCapabilities&VERSION=1.3.0&SERVICE=WMS)
Greenland: Visualisation of DATA quality / uncertainty

http://giv-uw.uni-muenster.de/vis/v2/
GEO label

- a quality **indicator** for GEOSS geospatial data and datasets
- assist in **searching** - provide users with visual clues of dataset quality / relevance.
- provide **accreditation**, **provenance**, **monitoring**
- increase **visibility** of EO data
- emphasize **open access** and **easy availability**
**GEO label meaning**

- **Producer** of the dataset, e.g., organisation or individual who produced the dataset, their contact information, etc.

- **Lineage/provenance** information, e.g., processing applied to data and number of process steps.

- **Informal comments** about the dataset quality as provided by the dataset **producer**, e.g., any identified problems, suggested use, etc.

- **Dataset’s compliance** with international standards, e.g., compliance with ISO 19115, Dublin Core, etc.

- **Quality measures** of the dataset, e.g., uncertainty measures recorded in UncertML, errors, accuracy information, etc.

- **Feedback, comments and ratings** provided by the users of the dataset, e.g., general comments on dataset quality, identified problems, suggested use for the dataset, etc.

- Domain **experts’ comments** on dataset quality, e.g., results of formal quality checks, expert suggestions on the dataset applications, etc.

- **Citations** where the dataset was used and cited, e.g., formal reports on dataset quality checks, journal articles, etc.
GEO Label : Metadata+feedback summary with drill-down

Citations for this dataset

Title: Objective air temperature mapping for the Iberian Peninsula using spatial interpolation and GIS
DOI: 10.1002/joc.1462

Category: journalArticle

Title: Monthly precipitation mapping of the Iberian Peninsula using spatial interpolation tools implemented in a Geographic Information System
DOI: 10.1007/s00704-006-0264-2
http://www.springerlink.com/content/05603p257731lx5j5/

Category: journalArticle
GEO Label-based Dataset Intercomparison Tool
GeoViQua Scenarios booklet

Agriculture flooding practices control

The EU agri-environmental measures (Regulation CEE 1257/1999) consist of a subsidy that farmers receive to support agricultural practices that favour the conservation of affected ecosystems. In Catalonia, the local government monitors the farmers’ practices by means of a fair decision support system based on remote sensing imagery and flooding practices detection, replacing time-consuming and expensive field surveys.

Global Carbon Atlas

Through the Carbon Atlas website, carbon balances information is provided to derive both technical reports and dissemination material for the media, general public and, most importantly, inputs for national policy makers who ultimately agree on the emission market rules. The Global Carbon Atlas is able to return datasets in several visualization styles (e.g., legends, layers styling, etc) through GetMap WMS requests. Flux models can be visually compared and analyzed, detecting trends, anomalies and calculating statistics related to the quality and performance of the models.
URLs for the public components code and API’s

- **Schemas**
  - schemas.geoviqua.org

- **Catalogues** DAB-Q, CSW, GI-CAT (CNR)
  - http://geoviqua.essi-lab.eu/dabq-demo/

- **GEO label services, API, github** (ASTON+52N)
  - http://www.geolabel.net/api.html
  - http://twiki.geoviqua.org/twiki/bin/view/GEO_SIF/SifGeoLabel
  - https://github.com/lushv/geolabel-service

- **WMS-Q, WMS, ncWMS, contact** MiraMon (UREAD + CEA)
  - http://ncwms.geoviqua.org/godiva2.html

- **KML-Q, KML, github** Contact Fraunhofer (FRAUN)
  - https://github.com/igd-geo/pcolor

- **GECAaaS WPS, Contact** S&T

- **Q emitter tool (github) API, (ASTON+S&T)**
  - https://github.com/GeoViQua/computeqi-web
  - https://github.com/GeoViQua/emulatorization-api

- **GeoNetwork**
  - https://github.com/GeoViQua/geoviqua-geonetwork-plugin
  - For Geonetwork 2.8 use the 2.8.x-dev branch
www.geoviqua.org
Come to our booth!!
GeoViQua scenarios

- Agriculture policy control (UAB)
- Global Carbon (CEA)
Agriculture flooding practices control
Scenario

• The EU agri-environmental measures (Regulation CEE 1257/1999) consist of a subsidy that farmers receive to support agricultural practices that favour the conservation of affected ecosystems. In Catalonia, the local government monitors the farmers’ practices by means of a fair decision support system based on remote sensing imagery and flooding practices detection, replacing time-consuming and expensive field surveys.
...Farmers’ flooding practices are extracted by combining Landsat image classification procedures with rigorous statistical quality assessments. Both the Landsat imagery and sample of the classified maps are available through the GEOSS Portal.

GEOSS PORTAL
http://geoportal.geoviqua.org

This enhanced GEOSS Portal includes:
• discovery and access of quality indicators and traceability embedded in metadata
• a GEO Label and referencing scientific reports
• smart quality visualization
  • at pixel level (quality enriched Web Map services [WMS-Q] visualized in Greenland)
  • at feature level (Quality enriched [KML-Q] visualized in Google Earth).
...A sudden failure of the Landsat TM/ETM+ sensor occurs during a monitoring critical period. Time is critical to find a suitable alternative dataset. A discovered issue is found.

**ADDITIONS IN THE METADATA QUALITY MODEL**

The Producer Quality Model extends ISO 19115-1, 19115-2 and 19157 standards to include information such as discovered issues, new and better formalized quality indicators (reference datasets for quality evaluation, traceability, and statistical summaries of uncertainty), citations to publications, pixel level quality and more.
...The government technicians can search for these alternative datasets through the GEOSS Portal thanks to the quality-aware component (DAB-Q)

QUERY BY QUALITY INDICATOR

The Quality-enabled Discovery and Access Broker (DAB-Q) is an extension of the GEO-DAB that implements the Producer Quality Model (and also the User Quality Model) as a quality extensions of the OGC Catalogue Service for the Web (CSW-ISO-Q) interface.
4...The metadata of the possible alternative datasets can be compared to facilitate the choice of the right product. The most fit-for-purpose dataset is identified through metadata comparison.

METADATA COMPARISON

The metadata comparison tool presents datasets in columns, with attributes and metadata parameters aligned in rows.
...At the end of the campaign, the local government releases KML-Q with a summary of the season.

### KML-Q AND WMS-Q

http://www.ogc.uab.es/geoviqua/wmsq

Quality can be described at dataset level (overall indicator in the metadata) or at pixel level and feature level (a quality value for each pixel or feature). GeoViQua has incorporated into the ncWMS and in the KML-Q innovative symbolization.
Global Carbon Atlas Scenario

- Through the Carbon Atlas website, carbon balances information is provided to derive both technical reports and dissemination material for the media, general public and, most importantly, inputs for national policy makers who ultimately agree on the emission market rules. The Global Carbon Atlas is able to return datasets in several visualization styles (e.g., legends, layers styling, etc) through GetMap WMS requests. Flux models can be visually compared and analyzed, detecting trends, anomalies and calculating statistics related to the quality and performance of the models. http://www.globalcarbonatlas.org
When a scientist detects an anomaly he can send his **feedback** with a detailed report and contribute to improve the accuracy of the models on the basis of local studies.

**USER FEEDBACK**

https://geoviqua.stcorp.nl/submit_feedback.html

Offers a framework for storing and querying user comments, ratings, expert reviews, metadata overrides and other feedback elements. The user **feedback** component allows introduction of new feedback items, while the DAB-Q is able to search in the user feedback database and integrate the response in its search results.
...User feedback and expert reviews can be accessed by means of a GEO label. In the carbon atlas, each dataset is associated a GEO label.

GEO LABEL

http://www.geolabel.info

The GEO label is a dynamic graphic representation which visually summarizes the availability of quality information for the dataset it represents, comprising 8 informational facets: producer profile, producer comments, lineage information, standards compliance, quality information, user feedback, expert review and citations Information.
In the Carbon Atlas, a scientist wants to compare two type of carbon flux models. They have different geographic projections and a different averaging period. **GECA Toolset** harmonizes them.

**DATA COMPARISON TOOL: GECAaaS**

The **GECA Toolset** is a toolkit for ingesting, processing and inter-comparing satellite data against collocated measurements, which can be either in-situ data or other satellite data.
The scientist gets a new quality indicator that convinces him about the fit for purpose of the data. From now on, other users will also get the quality emitter result.

QUALITY EMITTER

The WPS utilized by the GECA Toolset can also be queried directly to generate a ZIP file that contains the intercomparison report and collocated values. The quality emitter extracts a QI and adds it to the Producer Quality Model metadata. The quality emitter also provides a validation for quality indicators.

![Reliability diagram](image-url)