

Functional characteristics of a metadata catalogue tool

Agenda

- Metadata search & view facilities
 - Metadata discovery, Data view and download
 - Catalogue Service
- Metadata editing
 - multilingual content
 - terminology of reference (ie. thesaurus)
- Interoperability access
 - a web access point
 - import/export XML formats (ISO & DC)
 - metadata exchange, harvesting, ...
- on going improvements

Search and view

Metadata search & view facilities

Standard catalogue interface (search/results)

User: Prunayre FX | Logout

Mediterranean Water Information Datasource Catalogue

*search here ...

What?

Title

Abstract

Keyword

Template no

Search accuracy

Where?

When?

[Hide advanced options](#)

Home | Administration | Contact us | Links | About | Help | English

Find Interactive Maps, GIS datasets, Satellite Imagery and Related Applications

Most popular

- [Zones vulnérables 2003](#)
- [Zones vulnérables 2007](#)
- [BDCarthage / Zones hydrographiques](#)
- [Masse d'eau de transition](#)
- [Zones sensibles - Arrêté 2005](#)

Recent Changes [GeoRSS](#)

- [Template for Vector data in ISO19115](#)
- [Template for Vector data in ISO19115](#)
- [Multilingual template](#)
- [Balkans population density 2007-10 - \(UNEP\)](#)
- [ESA Tiger website \(draft for demo\)](#)
- [OSS website \(draft for demo\)](#)
- [Statistiques diverses sur l'alimentation en eau potable au Maroc](#)
- [Secteur de l'eau en chiffres - Maroc](#)
- [Suivi des progrès dans le domaine de l'eau au Maroc](#)
- [Rapport National 2004 sur les ressources en eau au Maroc](#)

Refine Search

Category

- [dataset \(66\)](#)
- [geo-dataset \(vector/raster\) \(5\)](#)
- [info. system international \(28\)](#)
- [info. system national/local \(15\)](#)
- [maps \(50\)](#)
- [otherresources \(9\)](#)
- [z-interactive resources \(1\)](#)

keyword

- [ad \(3\)](#)
- [ad, gw \(6\)](#)
- [adm \(1\)](#)
- [administrative boundary \(24\)](#)
- [africa \(2\)](#)
- [agence de l'eau \(1\)](#)
- [agriculture \(1\)](#)
- [agriculture \(13\)](#)
- [ain \(15\)](#)
- [aisne \(9\)](#)
- [albania \(7\)](#)
- [algeria \(2\)](#)
- [allier \(8\)](#)
- [alpes-de-haute-provence \(15\)](#)
- [alpes-maritimes \(19\)](#)
- [andorra \(1\)](#)
- [annual \(2\)](#)
- [ardeche \(15\)](#)
- [ardennes \(8\)](#)
- [ariege \(13\)](#)

Metadata search & view facilities

... with a catalogue summary for preferred records and latest changes

User: Prunayre FX | Logout

Mediterranean Water Information Datasource Catalogue

What?

Title

Abstract

Keyword

Template no

Search accuracy

Where?

When?

Hide advanced options

Find Interactive Maps, GIS datasets, Satellite Imagery and Related Applications

Most popular

- [Zones vulnérables 2003](#)
- [Zones vulnérables 2007](#)
- [BDCarthage / Zones hydrographiques](#)
- [Masse d'eau de transition](#)
- [Zones sensibles - Arrêté 2005](#)

Recent Changes GeoRSS

- [Template for vector data in ISO19115](#)
- [Template for Vector data in ISO19115](#)
- [Multilingual template](#)
- [Balkans population density 2007-10 - \(UNEP\)](#)
- [ESA Tiger website \(draft for demo\)](#)
- [OSS website \(draft for demo\)](#)
- [Statistiques diverses sur l'alimentation en eau potable au Maroc](#)
- [Secteur de l'eau en chiffres - Maroc](#)
- [Suivi des progrès dans le domaine de l'eau au Maroc](#)
- [Rapport National 2004 sur les ressources en eau au Maroc](#)

RSS feeds

Home | Administration | Contact us | Links | About | Help | English

Refine Search

Category

- [dataset \(66\)](#)
- [geo-dataset \(vector/raster\) \(5\)](#)
- [info. system international \(28\)](#)
- [info. system national/local \(15\)](#)
- [maps \(50\)](#)
- [otherresources \(9\)](#)
- [z-interactive resources \(1\)](#)

keyword

- [ad \(3\)](#)
- [ad, gw \(6\)](#)
- [adm \(1\)](#)
- [administrative boundary \(24\)](#)
- [africa \(2\)](#)
- [agence de l'eau \(1\)](#)
- [agriculture \(1\)](#)
- [agriculture \(13\)](#)
- [ain \(15\)](#)
- [aisne \(9\)](#)
- [albania \(7\)](#)
- [algeria \(2\)](#)
- [allier \(8\)](#)
- [alpes-de-haute-provence \(15\)](#)
- [alpes-maritimes \(19\)](#)
- [andorra \(1\)](#)
- [annual \(2\)](#)
- [ardeche \(15\)](#)
- [ardennes \(8\)](#)
- [ariege \(13\)](#)

Metadata search & view facilities

... with catalogue summary for quick access & global overview

User: Prunayre FX | Logout

Mediterranean Water Information Datasource Catalogue

What?

Title

Abstract

Keyword

Template no

Search accuracy

Where?

When?

Search

Hide advanced options

Home | Administration | Contact us | Links | About | Help | English

Find Interactive Maps, GIS datasets, Satellite Imagery and Related Applications

Most popular

- [Zones vulnérables 2003](#)
- [Zones vulnérables 2007](#)
- [BDCarthage / Zones hydrographiques](#)
- [Masse d'eau de transition](#)
- [Zones sensibles - Arrêté 2005](#)

Recent Changes GeoRSS

- [Template for Vector data in ISO19115](#)
- [Template for Vector data in ISO19115](#)
- [Multilingual template](#)
- [Balkans population density 2007-10 - \(UNEP\)](#)
- [ESA Tiger website \(draft for demo\)](#)
- [OSS website \(draft for demo\)](#)
- [Statistiques diverses sur l'alimentation en eau potable au Maroc](#)
- [Secteur de l'eau en chiffres - Maroc](#)
- [Suivi des progrès dans le domaine de l'eau au Maroc](#)
- [Rapport National 2004 sur les ressources en eau au Maroc](#)

Refine Search

Category

- [dataset \(66\)](#)
- [geo-dataset \(vector/raster\) \(5\)](#)
- [info. system international \(28\)](#)
- [info. system national/local \(15\)](#)
- [maps \(50\)](#)
- [otherresources \(9\)](#)
- [z-interactive resources \(1\)](#)

keyword

- [ad \(3\)](#)
- [ad, gw \(6\)](#)
- [adm \(1\)](#)
- [administrative boundary \(24\)](#)
- [africa \(2\)](#)
- [agence de l'eau \(1\)](#)
- [agriculture \(1\)](#)
- [agriculture \(13\)](#)
- [ain \(15\)](#)
- [aisne \(9\)](#)
- [albania \(7\)](#)
- [algeria \(2\)](#)
- [allier \(8\)](#)
- [alpes-de-haute-provence \(15\)](#)
- [alpes-maritimes \(19\)](#)
- [andorra \(1\)](#)
- [annual \(2\)](#)
- [ardeche \(15\)](#)
- [ardennes \(8\)](#)
- [ariege \(13\)](#)

Metadata search & view facilities

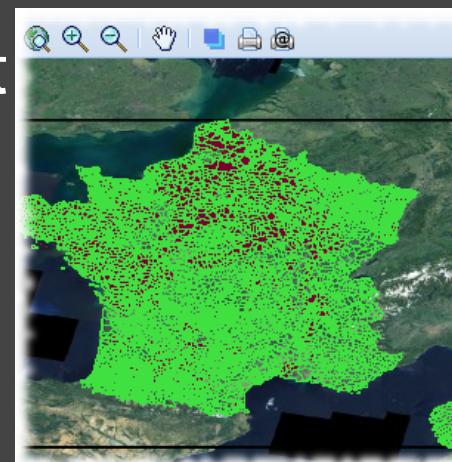
Narrow your search: Quick sum-up of your catalogue and quick access/filter

The screenshot shows a search interface with the following elements:

- Search Bar:** A text input field containing "for demo from EEA".
- Sort By:** A dropdown menu set to "Relevance".
- View:** A dropdown menu showing options like "List View", "Map View", and "Thumbnail View".
- Hits per page:** A dropdown menu set to "10".
- Results Area:** A list of entities:
 - elevation model
 - 18 are available: 30 9 km², including a
 - a, Austria, Belarus,
 - atia, Cyprus, Czech
 - ria, France, Georgia,
 - ity, Latvia,
 - re-the Former
 - of, Montenegro,
 - a, Russian
 - eden, Switzerland,
- Refine Search Sidebar:** Titled "Refine Search" with a "Clear" button.
 - Refinements:** A section with a "Clear" link.
 - Category:**
 - geo-dataset (vector/raster) [remove](#)
 - keyword:**
 - environmental data [remove](#)
 - Category:**
 - dataset (2)
 - info_system international (2)
 - keyword:**
 - administrative boundary (1)
 - albania (2)
 - armenia (2)
 - austria (2)
 - belarus (2)
 - bulgaria (2)
 - bosnia_and_bosnianerzegovina (2)
 - bulgaria (2)
 - croatia (2)
 - cyprus (2)
 - czech_republic (2)
 - georgia (2)
 - estonia (2)
 - latvia (2)
 - montenegro (2)
 - russia (2)
 - switzerland (2)
 - switzerland (1)

Metadata search & view facilities

- Identify sources (partner logo)
- Links to remote resources (when available):
 - Website
 - Data
 - Map your results / map on demand:
 - Bounding box highlighter
 - WMS display
 - GoogleEarth output



The screenshot shows a user interface for viewing metadata records. On the left is a map of France with green dots representing hydrographic zones. To the right are two detailed record cards.

BDCarthage / Zones hydrographiques

Abstract: Définition Sandre : "La circulaire n°91-50 du 12 1991 relative à la codification hydrographique et au repérage des milieux aquatiques superficiels en France métropolitaine définit la zone hydrographique comme suit : "L'ensemble du territoire français est divisé en zones élémentaires appelées zones hydrographiques. Leurs limites s'..."

Keywords: COURS D'EAU, ZONE HYDROGRAPHIQUE, BASSIN HYDROGRAPHIQUE, BASSIN VERSANT, EAU DE SURFACE, REFERENTIELS, BDCARTHAGE, FRANCE

(modified: 2008-08-13)

Zones vulnérables 2007

Abstract: Une zone vulnérable est une partie du territoire où l'environnement est exposé à une pollution des eaux par le rejet direct ou indirect de nitrates et de phosphates.

Editing

Metadata editing

- The prototype make metadata editing ...
 - easier
 - Start from templates or duplicate existing records
 - Suggest terms (eg. most usual scales, main keywords)
 - Inline & complete validation report
 - Custom widgets (eg. period, date datatype)
 - Batch processing tools for administrator (eg. keyword mapping or formatter)
 - advanced
 - multilingual editing
 - access constraint management

Multilingual content

- Multilingual content is different from multilingual interface
 - GeoNetwork user interface is available in **Arabic, Chinese, English, French, Holland, Russian, Spanish**
 - Lack of tools to do multilingual editing.
- Multilingual content needs to address the following items:
 - Editing
 - View
 - Indexing
- Method define in ISO Standard but NOT so much detailed in OGC spec. & INSPIRE guidelines

Multilingual content / editing

- Define :
 - main metadata language (INSPIRE recommandation)
 - define other possible languages
- ISO standards allow translation of every text fields
- Translate only the main one for discovery purpose (ie. title, abstract)
- Add rules to advice editor when editing :
 - "Metadata language is mandatory" (INSPIRE)

The screenshot shows a web-based form for 'Identification info' with several input fields and a language dropdown menu.

- Title***: Nationale Bedeutung: historischer Verlauf
- Date***: 2008-12-02T10:02:24
- Date type***: revision
- Abstract***: Nationale Bedeutung: Historische Verkehrswege sind von nationaler Bedeutung, wenn sie eine herausragende historische Bedeutung oder eine außerordentliche traditionelle Wegsubstanz aufweisen.
Historischer Verlauf: historischer Verlauf bedeutet, dass der historische Verkehrsveg in seinem Verlauf (als Linie) noch zu erkennen ist, dieser Verlauf aber keine oder nur wenige Elemente in traditioneller Form mehr
- Language Selection**: A dropdown menu shows 'German' selected from a list that also includes French, English, and Italian.

Multilingual content / ISO mechanism

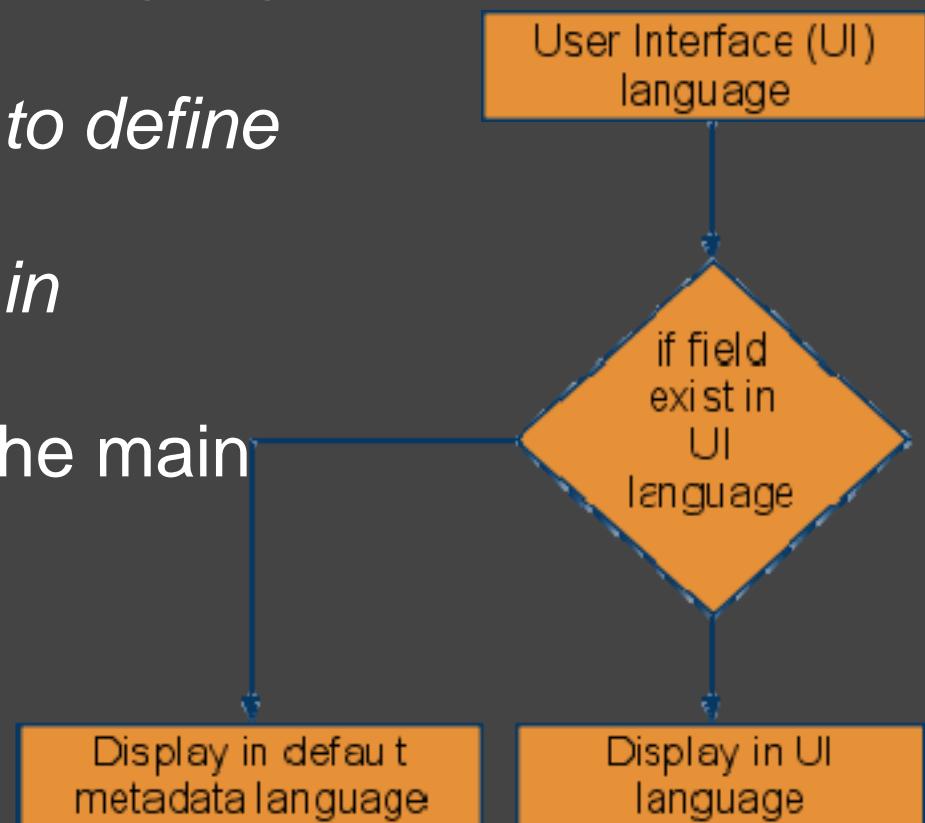
- How to store multilingual content in metadata records ?

```
<gmd:descriptiveKeywords
  xlink:href="http://localhost:8082/geonetwork/srv/eng/xml.keyword.get?thesaurus=external.theme.gemet&id=http://www.eionet.europa.eu/gemet/concept/12226&locales=FR,IT,SP,AR,EN" xlink:title="water salinity" xlink:role="embed">
<gmd:MD_Keywords>
  <gmd:keyword>
    <gmd:PT_FreeText>
      <gmd:textGroup>
        <gmd:LocalisedCharacterString locale="#SP">salinidad de agua</gmd:LocalisedCharacterString>
      </gmd:textGroup>
      <gmd:textGroup>
        <gmd:LocalisedCharacterString locale="#FR">salinité de l'eau</gmd:LocalisedCharacterString>
      </gmd:textGroup>
      <gmd:textGroup>
        <gmd:LocalisedCharacterString locale="#EN">water salinity</gmd:LocalisedCharacterString>
      </gmd:textGroup>
    </gmd:PT_FreeText>
  </gmd:keyword>
</gmd:MD_Keywords>
```

Multilingual content / viewing

Principles:

- "*Search information in all known languages in the catalogue*"
- "*Display information in end-user language when possible*"
- "*Use user interface information to define preferred language*"
- "*Use translation in metadata or in thesaurus*"
 - title, abstract, keywords are the main ones



Multilingual content / indexing

- One index by languages
 - Language specific analyzer
 - stopwords list,
 - per field analyzer (tokenizer),
 - per languages analyzer (eg. French, German)
 - Allows mixing european and non european language
- Search in all languages and boost current user interface languages

Multilingual content / issues

- INSPIRE recommandation : define metadata main language
- OGC does not handle multilingual contents and languages specific issues:
 - No languages parameters (all specs.)
 - GetCapabilities documents only in one language (ows:*)
 - No languages specific behaviour in CSW spec.

=>GeoNetwork implemented multilingual metadata engine following ISO rules.

Terminology

- Controlled vocabulary define in MEDWIP project for:
 - theme
 - countries
 - river basins & aquifers
- GEMET is used and has been integrated in GeoNetwork
- Questions :
 - More information needed on GEMET lifecycle
 - GEMET concepts could be removed ?
 - new concept could be proposed to GEMET ?



Metadata editing / validation

Validation report levels :

- 1.ISO schema
- 2.ISO conditional rules
- 3.INSPIRE metadata rules
- 4.Profil specific rules

Complete validation report

The screenshot shows a software interface for metadata validation. At the top, there's a toolbar with icons for 'New', 'Open', 'Save', 'Close', 'Print', and 'Help'. Below the toolbar, a preview area displays a small map thumbnail with the text 'No preview available' over it. To the right of the preview is a 'Validation report' window. The report window has several sections:

- Validation report**: A title bar with 'View errors only' checked.
- Compliance to metadata standard (XML Schema)**: A section with a green checkmark and the text '[INSPIRE] - Title is mandatory'.
- Compliance to metadata recommendations (Schematron)**: A section with a green checkmark and the text '[INSPIRE] - Abstract is mandatory'.
- INSPIRE implementing rules**: A section with a yellow warning icon and the text '[INSPIRE] - Resource type is mandatory'. It also lists 'INSPIRE - Resource type is mandatory'.
- Implementing instructions:** A section describing MD_ScopeCode values and dataset types (dataset, series, services).
- Implementation notes:** A section about hierarchyLevel and its relation to ISO 19115.
- Conformance results:** A section listing various conformance items with status indicators (green checkmarks, yellow warning icons) and associated error codes.
- Implementation instructions:** A section about the first two degrees of conformity defined in Part D 5 of the INSPIRE Implementing rules.
- Example:** A section providing an example of a title statement.
- Additional compliance:** A section listing three more INSPIRE implementing rules.

Metadata editing / validation

Inline validation report :

The screenshot shows a validation report with two main sections:

- Identification info**: A red exclamation mark icon indicates an error: **TOPICCATEGORY IS MANDATORY IF MD_METADATA.HIERARCHYLEVEL EQUAL 'DATASET' OR 'SERIES' OR DOESN'T EXIST.**
- INSPIRE - ISO TOPIC CATEGORY IS MANDATORY**: A detailed message explaining that topic categories are mandatory and derived from ISO 19115/ISO 19119 metadata elements.
- INSPIRE - KEYWORD IS MANDATORY**: A section for keyword validation.

Below these sections, there are dropdown menus for date type and edition information, along with input fields for edition date and citation identifier.

Date type	creation
Edition	1
Edition date	Date (gco:Date)
Citation identifier	Identifier (gmd:RS_Identifier)

Metadata access constraints

- Hide specific elements to restrict access to some sections
- Access constraint levels : public/intranet/internet

The screenshot shows a user interface for managing metadata. At the top left is a checkbox labeled "Enable element visibility editing". On the right, there is a "No preview available" message with a small globe icon.

On the left, a sidebar lists various metadata fields with visibility status indicators:

- File identifier: Visible to all (green)
- Character set: Intranet only (yellow)
- Parent identifier: Intranet only (yellow)
- Hierarchy level: Visible to all (green)
- Hierarchy level name: Intranet only (yellow)
- Date stamp: Intranet only (yellow)
- Metadata standard name: Visible to all (green)
- Metadata standard version: Intranet only (yellow)
- Contact: Visible to all (green)
- Responsible party: Intranet only (yellow)
- Organisation name: Intranet only (yellow)
- Position name: Intranet only (yellow)
- Contact Information: Intranet only (yellow)
- Contact: Intranet only (yellow)
- Phone: Intranet only (yellow)
- Telephone: Intranet only (yellow)
- Voice: Intranet only (yellow)

On the right, several input fields are displayed:

- File identifier: 537128b6-9f7e-45be-8fb0-27b3d8999f46
- Character set: German (dropdown menu)
- Encoding: utf8
- Dataset: dataset (dropdown menu)
- Field session: fieldSession (dropdown menu)
- Date stamp: 2009-04-20T11:26:15
- Version: GM03 2+
- Preview: Agroscope FAL Reckenholz, Eidgenössische Forschungsanstalt für German (dropdown menu)

A legend titled "Element visibility" is located in the top right corner, defining the color coding for visibility levels:

- Visible to all (Green circle)
- Intranet only (Yellow circle)
- Identified user only (Purple circle)

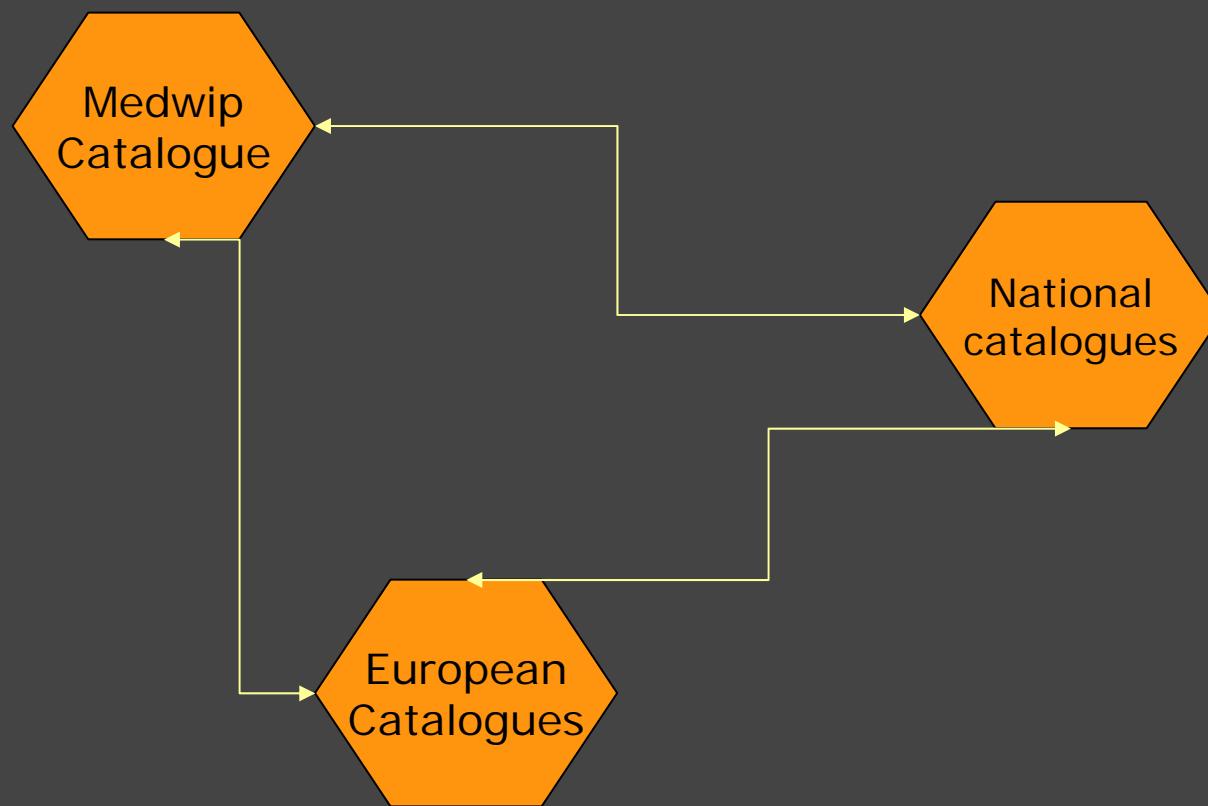
"Interoperability"

Interoperability

- News feeds:
 - RSS and GeoRSS
 - Google KML
- Export / Import: metadata in XML format
- Catalogue exchange:
 - CSW2.0.2 server and client
 - Harvesting:
 - Catalogue (CSW, OAI, ArcSDE, ...)
 - OGC services: create metadata from OGC services like WMS

Interoperability

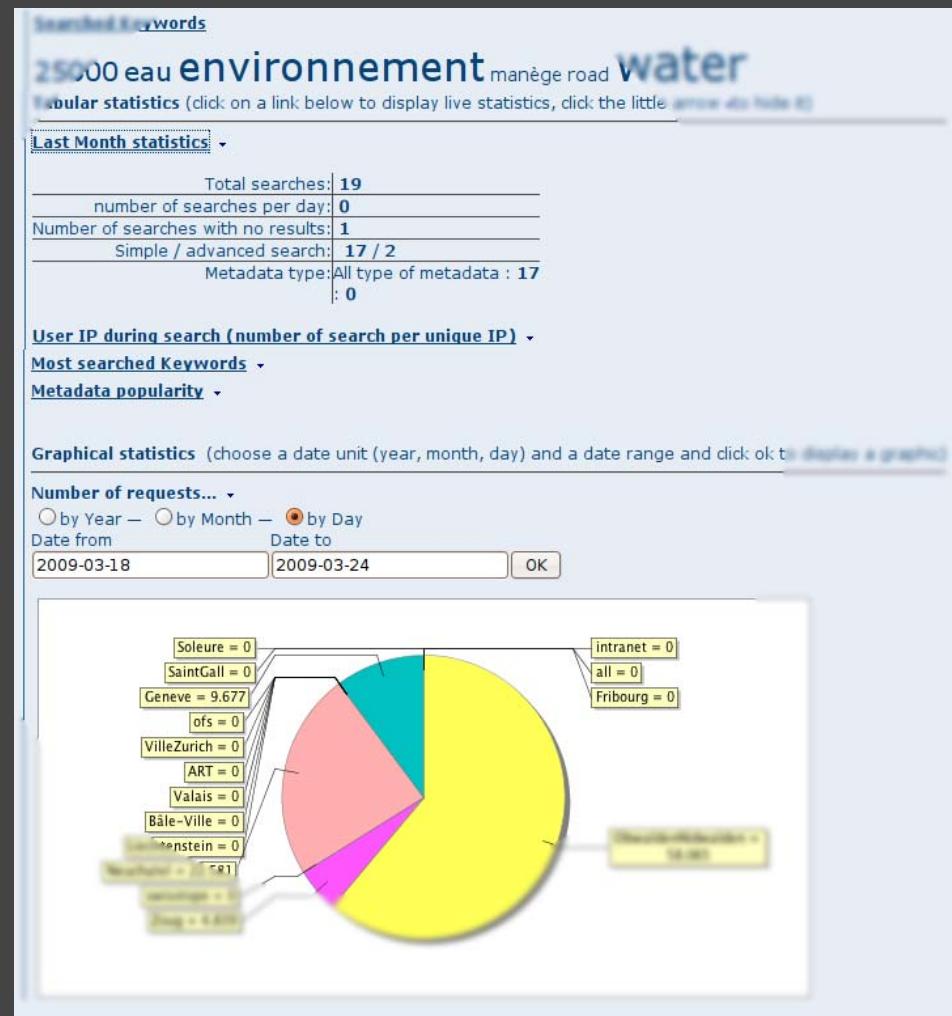
Communication between catalogues will mainly use OGC CSW protocol



ongoing improvements

Catalogue usage & statistics

- Reporting & statistics:
 - tagcloud
 - number of search
 - main search
 - metadata popularity
 - ...



Generate metadata & related resources monitoring

- Metadata describes resources: datasets, services, ...
- Metadata editing could be improved by analyzing remote resources and generating automatically part of the metadata
- Relations could be monitored:
 - parent/child relations (eg. basic geodata, aggregate, series)
 - dataset and service metadata
 - Link with resources
- ... in order to report information (eg. Catalogue footprint/coverage – spatial, temporal and themes - , dead links, services availability) about catalogue content.

Thanks for your attention