SEMIDE

http://www.semide.org

EMWIS

http://www.emwis.org

Proposal of common metadata

Lylia ABROUK
LIRMM
EMWIS Technical Unit
l.abrouk@semide.org

EURO-MEDITERRANEAN PARTNERSHIP

Outline

➢ **Metadata : definition**

➢ **Why using Meta-data?**

➢ **Using Metadata for EMWIS**

➢ **Types of data managed by EMWIS**

➢ **Metadata associated to each type of data**

➢ **Prospective approach: a web service for each provider**

➢ **Discussion**

Metadata: definition

➢ **A standard and structured set of information describing a resource or an object: structured data about data**

➢ **Metadata machine readable information**

➢ **The described resources can be of any type and nature.**

Why using Metadata?

➢ **For information resources, metadata are used for:**
  - ➢ **Describing the resource**
  - ➢ **Facilitating search**
  - ➢ **Location**
  - ➢ **Management**

➢ **Without metadata it would be difficult to manage or search for some resources (e.g. images, video)**

➢ **Necessary for Content Management Systems:**
  - ➢ **classification, archiving …**

➢ **Allowing customisation of user interfaces:**
  - ➢ **Linked to user profiles**

➢ **Extending search criteria beyond the classical keywords.**

➢ **Providing information necessary for the administrative and technical management of the resources: date, rights.**

➢ **Facilitating exchanges and integration of data**

Metadata

➢ **Descriptive information :**

   ➢ **title, author, publication date, editor, etc.**

➢ **Information on the contents :**

   ➢ **keywords (resulting from a thesaurus or a glossary), summary**

➢ **Information on how to access :**

   ➢ **Localization, access protocol, structuring of the data ,**

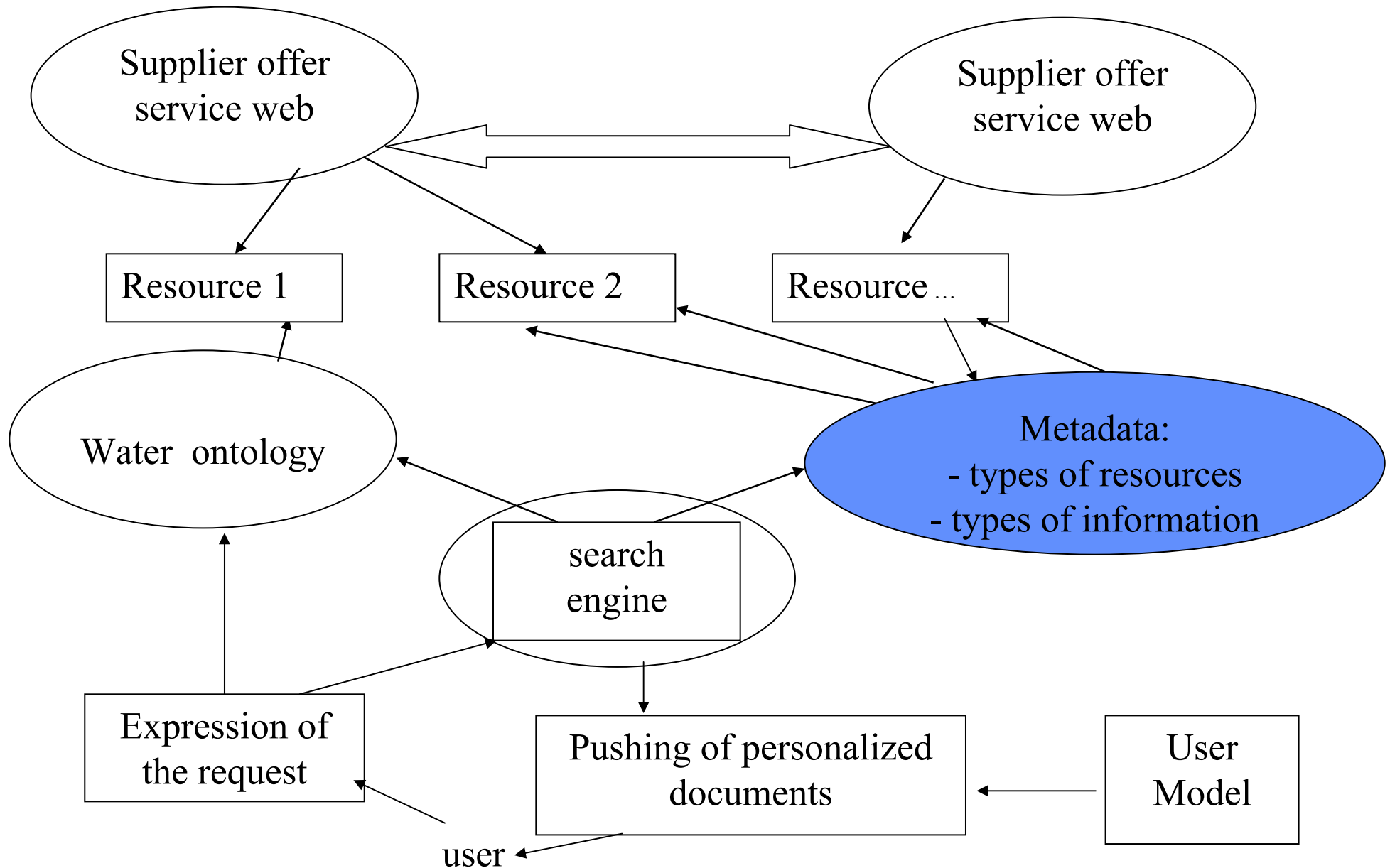➢ **Administrative information :**

   ➢ **Rights, etc.**

Standards

➤ **The Dublin Code (DC), created in 1995, is a convention shared by documentation specialists and librarians at an international level. It defines generic metadata and develops tools to implement metadata for resource.**

➤ **ARIADNE (Alliance of Remote Instructional Authoring and Distribution Networks for Europe), created in 1996, is a convention shared by economic actors both from the public and private sectors and europeen universities (plus switzerland) . This alliance is based on the FOAD (Formation Ouverte A Distance, Open Distance Learning). It extends the DC metadata.**

➤ **L'IMS (Instructional Management Systems), created in 1997, is a convention shared by universities and american busnisses. It is based on the FOAD and extends the DC metadata.**

➤ **SCORM (Sharable Content Object Reference Model), created in 1997, is a convention shared by the US defense ministary and centered on the FOAD.**
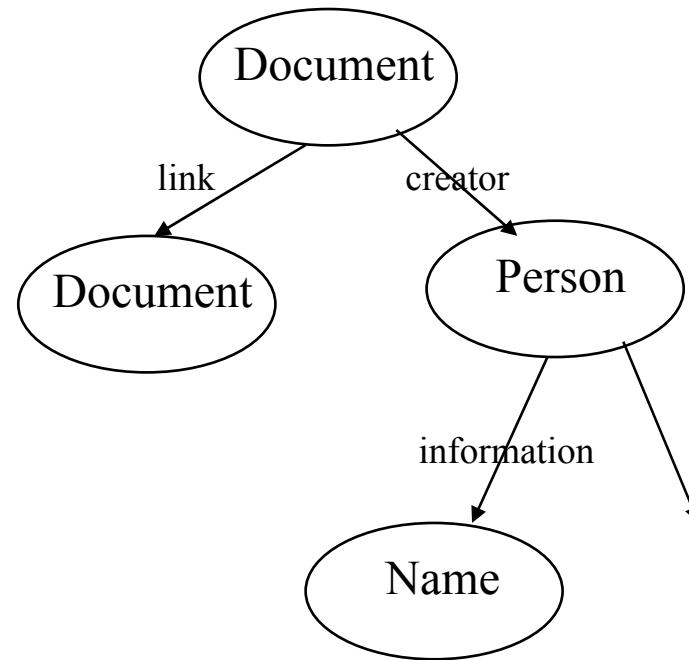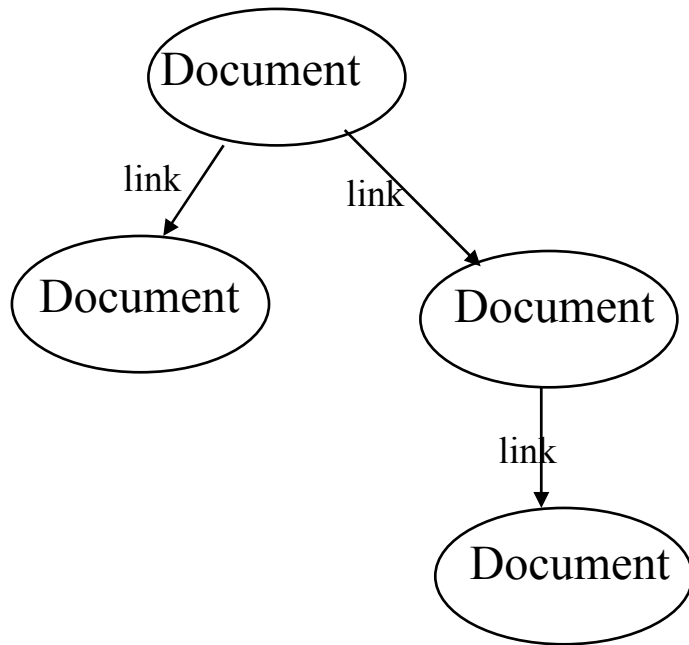
Using Metadata for EMWIS

➢ **We need to access information available on water that are coming from different remote and heterogeneous sources:**

  ➢ Identification of the information available in the Euro-Med domain
  ➢ Accessing this information
  ➢ Allow the international community to access the services developed by EMWIS

➢ **Service interoperability among the focal points:**

  ➢ Common metadata are needed to exchange and integrate content

➢ **Evolution towards a semantic web architecture**

ARCHITECTURE

Supplier offer service web

Supplier offer service web

Resource 1

Resource 2

Resource ...

Water ontology

Metadata:
- types of resources
- types of information

search engine

Expression of the request

Pushing of personalized documents

User Model

user

Metadata (EMWIS)

Actual Web =>Semantic Web

Metadata (EMWIS)

➢ **Question:**

All the projects which have a cost higher than 30000 euros

Type of resource

Meta-data

Value

Metadata (EMWIS)

HTML document (without metadata)

```
<HTML>
   <HEAD>
     <TITLE> Projects </TITLE>
   </HEAD>
   <BODY>
            <UL>
                  <LI> EMWIS project </LI >

                  <LI> cost is 35000 euros</LI >          </UL>
   </BODY>
</HTML>
```

Metadata (EMWIS)

XML document

```
<document>

    <title>  EMWIS project </ title >
    <type> projects </type>
     …….
     …….
     <funding></ funding >
    <cost>35000 </cost>


</document>
```

Types of resources

➢ **News**

➢ **Events**

➢ **Texts of law**

➢ **Multimedia**

➢ **Web sites**

➢ **Projects**

➢ **Other document full text**

Definition of metadata for each type of resource

Metadata (EMWIS)

➢ **Common to all the types of resource**

  ➢ **Title, Description, Date (creation, modification)**

➢ **Contact**

  ➢ **Who contacting to obtain additional information on this resource**

➢ **Specific to each type of resource**

  ➢ **Responding to specific needs: e.g. projects (cost, beneficiaries)**

➢ **Reference to international standards, as far as possible**

  ➢ **Visibility outside of EMWIS community (e.g. Dublin Core)**

Dublin Core elements

element	Identifier	Definition
titre	Title	The name given to the resource by the CREATOR or PUBLISHER.
créateur	Creator	The person(s) or organization(s) primarily responsible for the intellectual content of the resource; the author.
sujet et mots-clefs	Subject	The topic of the resource; also keywords, phrases or classification descriptors that describe the subject or content of the resource.
description	Description	A textual description of the content of the resource, including abstracts in the case of document-like objects; also may be a content description in the case of visual resources.
éditeur	Publisher	The entity responsible for making the resource available in its present form, such as a publisher, university department or corporate entity.
contributeur	Contributor	Person(s) or organization(s) in addition to those specified in the CREATOR element, who have made significant intellectual contributions to the resource but on a secondary basis.
date	Date	The date the resource was made available in its present form.
type	Type	The resource type, such as home page, novel, poem, working paper, technical report, essay or dictionary.
format	Format	The data representation of the resource, such as text/html,
identifiant	Identifier	A string or number used to uniquely identify the resource.
source	Source	The work, either print or electronic, from which the resource is delivered (if applicable).
langue	Language	The language(s) of the intellectual content of the resource.
relation	Relation	The relationship to other resources.
couverture	Coverage	The spatial locations and temporal duration characteristics of the resource.
droits	Rights	A link (URL or other suitable URI as appropriate) to a copyright notice, a rights-management statement or perhaps a server that would provide such information in a dynamic way.

Meta-data
1. News

- *Author name*
- Author mail
- *Rights*
- *Title*
- *keywords*
- *Description*
- *Type*
- *Format*
- *Identifier*

- File link
- *Source name*
- Source link
- *Language*
- Topic
- *References*
- *Coverage*
- *Save date*
- News date

Metadata
2. Events

- ➤ *Author name*
- ➤ **Author mail**
- ➤ *Rights*
- ➤ *Title*
- ➤ *keywords*
- ➤ *Description*
- ➤ *Type*
- ➤ *Format*
- ➤ *Identifier*
- ➤ **File link**

- ➤*Source name*
- ➤ **Source link**
- ➤ *Language*
- ➤ **Topic**
- ➤ *References*
- ➤ *Coverage*
- ➤ **Start date**
- ➤ **End date**
- ➤ *Publication date*
- ➤ **Country**

Metadata
3. Text of law

- *Title*
- *keywords*
- *Description*
- *Type*
- *Format*
- *Identifier*
- **File link**
- *Source name*
- **Source link**

- *Language*
- **Topic**
- *References*
- *Coverage*
- *Date*
- **Official journal**
- **Jounal number**
- **Origine language**

Metadata
4. Multimedia

➢ *Author name*

➢ **Author mail**

➢ *Rights*

➢ *Title*

➢ *keywords*

➢ *Description*

➢ *Identifier*

➢ **File link**

➢ *Source name*

➢ **Source link**

➢ *Language*

➢ **Topic**

➢ *References*

➢ *Coverage*

➢ *Date*

➢ *Type of file*

➢ **Type of media**

Metadata
5. Web site

- **Webmaster**
- **Mail**
- *Title*
- *keywords*
- *Description*
- *Type*
- *Identifier*
- **Link**

- *Source name*
- **Source link**
- *Language*
- **Topic**
- *References*
- *Coverage*
- **Code (xml,html)**

Metadata
6. projects

- **Id**
- **Acronym**
- ***Title***
- **Budget total**
- **Funding**
  - **Source**
  - **Programme**
  - **Type**
  - **Funding rate**

- **Programme**
- **Web site**
- **Background**
- **Objectives**
- **Results**
- **Starting date**
- **Ending date**
- **Publications**

Metadata
6. projects

➢ **Partners Nom organisme**

   ➢ **Name organisme**

   ➢ **Type organisme**

   ➢ **Description**

   ➢ **Adress**

   ➢ **Country**

   ➢ **Web site**

   ➢ **Contact**

      ➢ **Contact title**

   ➢ **Contact firstname**

   ➢ **Contact lastname**

   ➢ **Position**

   ➢ **Email**

   ➢ **Tel**

   ➢ **Fax**

➢ **Field site**

   ➢ **Name**

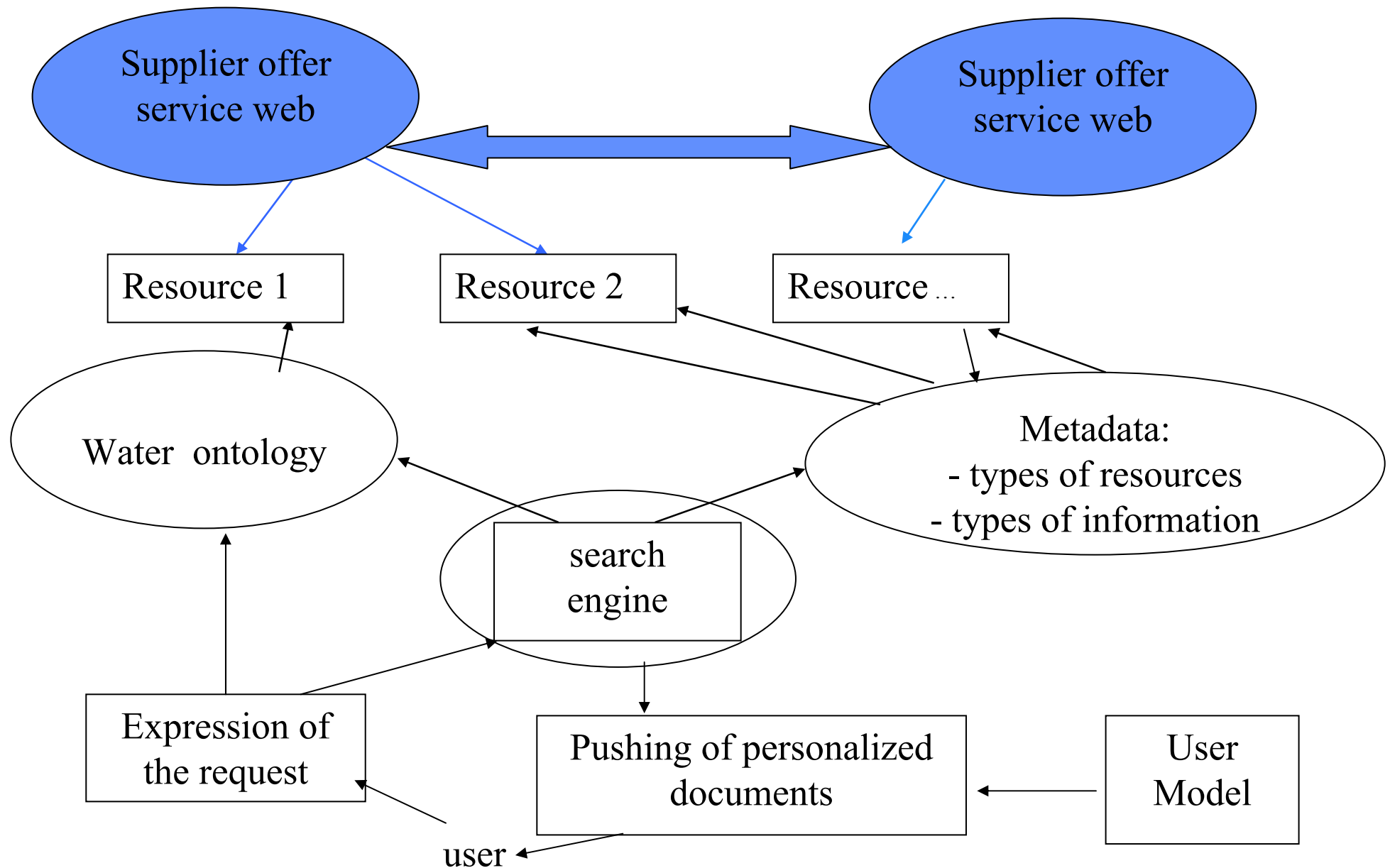   ➢ **Country**

   ➢ **river basin**

Metadata
7. Full text

- ➢ *Editor*
- ➢ *Author name*
- ➢ **Author mail**
- ➢ *Rights*
- ➢ *Title*
- ➢ *keywords*
- ➢ *Description*
- ➢ *Type*
- ➢ *Format*

- ➢ *Identifier*
- ➢ **File link**
- ➢ *Source name*
- ➢ **Source link**
- ➢ *Language*
- ➢ **Topic**
- ➢ *References*

Perspectives (Services Web)

Supplier offer service web ⟷ **Supplier offer service web**

Resource 1   Resource 2   Resource ...

Water ontology

Metadata:
- types of resources
- types of information

search engine

Expression of the request

Pushing of personalized documents

User Model

user

Discussion