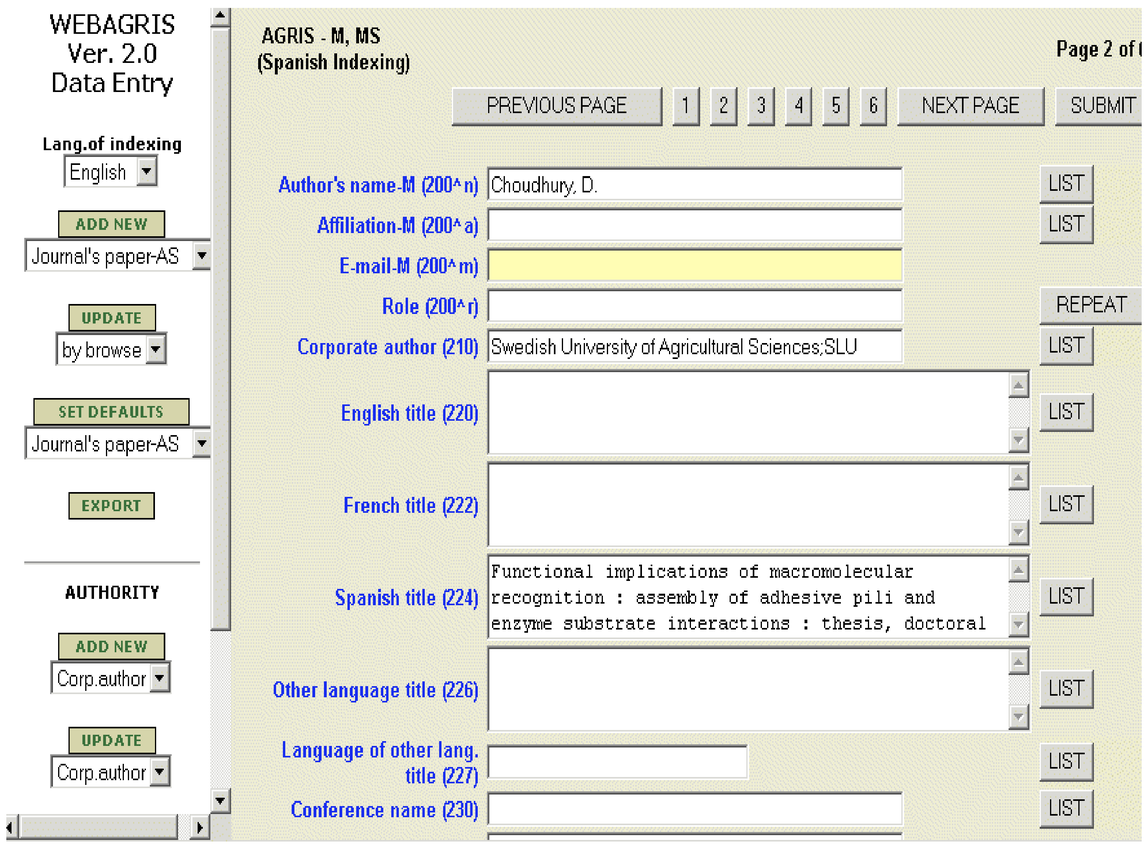
5th Workshop on Innovations in Scholarly Communication, OAI5, 18th to 20th April 2007 , Geneva (Switzerland)

Exposing metadata from Web CDS/ISIS databases to OAI framework

Background

CDS/ISIS is a broadly used system for metadata management.
Some FAO Web applications based on CDS/ISIS are:

- **WEBAGRIS** - a complete, multilingual system for distributed data input, processing and dissemination, of agricultural bibliographic information. It is based on common standards of data input and exchange formats (XML, HTML, ISO2709), as well as subject categorization schema and AGROVOC thesaurus. WEBAGRIS also allows to link to documents that are available in electronic format. It is used by almost 90 % of the AGRIS network participating centers.
- **FAOBIB** – an application for management of FAO on-line catalogue with more than 150000 records.

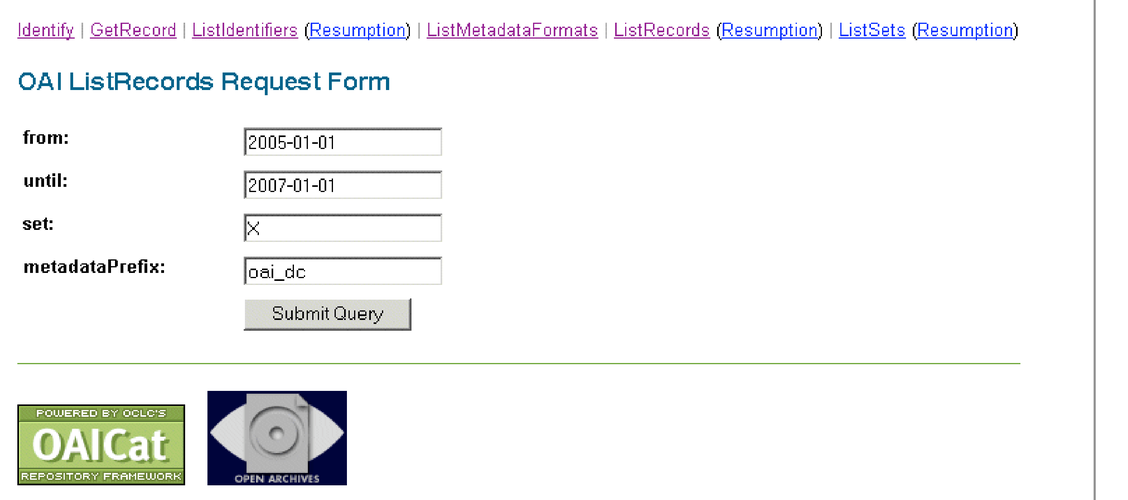


OAI framework

consists of data providers and service providers. The quality of service is proportional to the quality of the data harvested
A harvester is a client application that issues OAI-PMH requests. A harvester is operated by a service provider as a means of collecting metadata from repositories.
A repository is a network accessible server that can process the 6 OAI-PMH requests (verbs). A repository is managed by a data provider to expose metadata to harvesters.

Technical details on ISISOAI

- **ISISOAI** is an Open source Java Servlet Web application, for exposing CDS/ISIS databases to OAI-PMH v2.0 framework
- It is a customization of OAICat that provides OAI access to metadata records from a database or stored in a XML file,
- Runs under Apache's Tomcat Server
- Adapted for use with WWW/ISIS (WEBAGRIS and FAOBIB) and wxis applications. Implemented all 6 verbs



Who can use it?

ISISOAI can be easy installed at any Internet accessible CDS/ISIS data provider that uses WEBAGRIS or other CDS/ISIS Web applications in order to expose their data to OAI framework

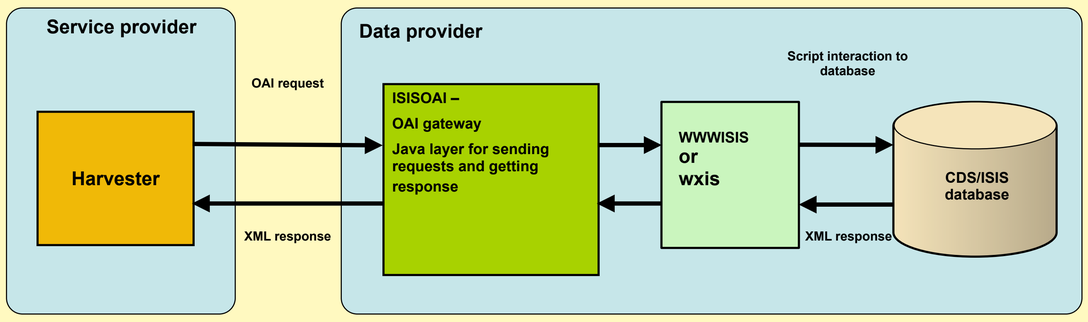
Impact: Opens CDS/ISIS community for dynamic harvesting over OAI-PMH, including AGRIS Network, FAO library etc.
Assures that the OAI functionality applies to:

- **Simple DC syntax:** The original data elements are grouped under 15 DC elements, with much larger requirements (no mandatory elements and all of them could be repeatable) and,
- **AGRIS AP syntax**: more complex than simple DC with mandatory and nested elements, exploring more complex structure of the original metadata for further integration in a value added services

Flow of the process:

Harvester request: http://www4.fao.org:8080/oaicat?verb=ListIdentifier&from=2006-01-01&until=2006-12-19&metadataPrefix=oai_dc

CDS/ISIS Script: http://www4.fao.org/cgi-bin/oai.exe?database=faobib&search_type=query&query=t=mon&FAO=FAO%2Adocuments&DS=20061219&table=mont&lang=oai&format_name=idendate



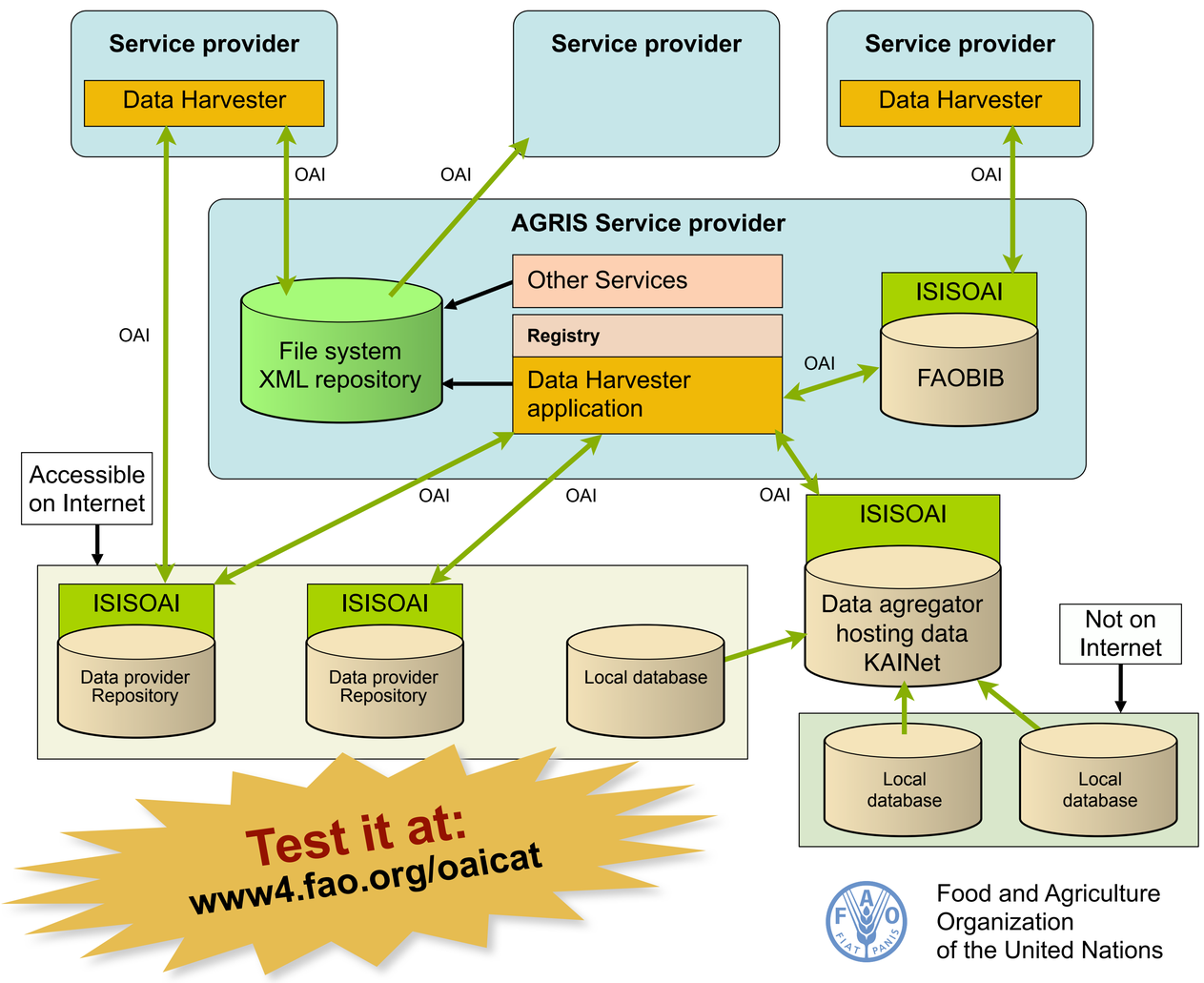
Xml response


```
<?xml version="1.0" encoding="UTF-8" ?>
- <OAI-PMH xmlns="http://www.openarchives.org/OAI/2.0/" xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xsi:schemaLocation="http://www.openarchives.org/OAI/2.0/ http://www.openarchives.org/OAI/2.0/OAI-PMH.xsd">
  <responseDate>2007-03-14T13:50:28Z</responseDate>
  <request metadataPrefix="oai_dc" until="2007-01-01" verb="ListRecords" set="X" from="2005-01-
    01">http://www.dba.it/oaicat/OAIHandler</request>
  - <ListRecords>
    - <record>
      - <header>
          <identifier>oai:isis.fao.org:XF2006427477</identifier>
          <datestamp>2006-09-08</datestamp>
          <setSpec>X</setSpec>
          <setSpec>A</setSpec>
        </header>
      - <metadata>
        - <oai_dc:dc xmlns:oai_dc="http://www.openarchives.org/OAI/2.0/oai_dc/"
            xmlns:dc="http://purl.org/dc/elements/1.1/" xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
            xsi:schemaLocation="http://www.openarchives.org/OAI/2.0/oai_dc/
            http://www.openarchives.org/OAI/2.0/oai_dc.xsd">
            <dc:title>Opportunity untapped. Water, food, agriculture and rural livelihood</dc:title>
            <dc:title>Une opportunite a saisir. L'eau pour l'alimentation, l'agriculture et le bien-etre en milieu rural</dc:title>
            <dc:title>Una oportunidad para aprovechar. Agua para alimentos, agricultura y medio de vida</dc:title>
            <dc:title>FAO reports and guidelines on water and agriculture</dc:title>
            <dc:title>Rapports et directives de la FAO concernant l'eau et l'agriculture</dc:title>
            <dc:title>Informes y directrices de la FAO sobre el agua y la agricultura</dc:title>
            <dc:creator>FAO, Rome (Italy). Land and Water Development Div.</dc:creator>
            <dc:creator>FAO, Rome (Italy). Div. de la Mise en Valeur des Terres et des Eaux</dc:creator>
            <dc:creator>FAO, Rome (Italy). Direccion de Fomento de Tierras y Aguas</dc:creator>
            <dc:creator>AGL</dc:creator>
            <dc:publisher>FAO</dc:publisher>
```


How to implement it

- Preparation of CDS/ISIS Scripts equivalent to the harvester requests
- Some modification of the local application
- Registration of the data server exposing metadata at OAI registry
- Installation of ISISOAI servlet and testing with a harvester
- Changing parameters for using it by different applications
- Schedule for automatic harvesting
- Services over harvested data

ISISOAI harvesting within the AGRIS network



Food and Agriculture Organization of the United Nations

Contact: stefka.kaloyanova@fao.org, Information Systems Analyst, Knowledge Exchange and Capacity Building Division (KCE)