FAO's experience in metadata exchange from CDS/ISIS bibliographic databases using XML format, compliant to Dublin Core standard

Stefka Kaloyanova, Food and Agriculture Organization of the United Nations,Rome, Italy
Stefka.Kaloyanova@fao.org;
Michal Okoniewski, FAO, MOkoniewski@PICR.man.ac.uk

Key Words
bibliographic metadata, CDS/ISIS, metadata exchange, DC, XML, AGRIS Application profile

Abstract

One of the problems in collaborative networks of documentation centers is the lack of interoperability between the different types of systems and applications used. The metadata is created in various places, time and using a number of tools and applications, data structures, content types, formats, and standards that may vary according to local needs. Exchange specifications that identify a minimum set of common requirements can be used for to facilitate the exchange and re-use of data.

This paper summarizes FAO's experience in the efficient management and use of multilingual bibliographic CDS/ISIS databases, applying DC metadata standards and tools for search and exchange of bibliographic metadata.

It explains in detail different approaches on how the bibliographic data from the CDS/ISIS databases can be converted and exported according to the Dublin Core (DC) and Agricultural Metadata Element Set (AgMES) based metadata set for cohesive and efficient platform-independent exchange of metadata. The second point is how different database applications can be united in a common search interface using DC elements. In addition, export features of DOS-, WINDOWS- and WEB-based CDS/ISIS application systems tools are compared. A number of examples on how the data in various input and storage formats can be reformatted without any modifications to the existing CDS/ISIS database and applications.

The experience was gathered while working with FAO's on-line library catalogue (FAOBIB) and the International Information System for the Agricultural Sciences and Technology (AGRIS). The solution adopted was to use the Dublin Core (DC) and Agricultural Metadata Element Set (AgMES) based metadata set for the platform-independent exchange of metadata between existing databases.

Introduction

The best way of ensuring and maintaining interoperability between different systems and thus to "maximize opportunities for exchange and re-use of metadata" is to use standard specifications with minimum set of common requirements .

FAO has gained a lot of experience in the efficient management and use of multilingual bibliographic databases that are using different data storage and exchange formats.

Here we explain how the bibliographic data from the local databases (mainly based on CDS/ISIS software) can be converted and exported according to the Dublin Core (DC) and Agricultural Metadata Element Set (AgMES) based metadata set for cohesive and efficient platform-independent exchange of metadata between existing databases and resource centres.

The experience was gathered while working with FAO's on-line library catalogue (FAOBIB) [1]and the International Information System for the Agricultural Sciences and Technology (AGRIS) [2]

1. Interoperability problems encountered

The main problem in collaborative networks that exchange their information is the lack of interoperability between the different type of systems and applications used. The metadata is created in different places, time and using different tools and applications, data structure and content, formats, and standards. Some of the problems that were encountered are summarized below:

1.1 Variety in the systems and input/exchange formats used

The software applications that are primarily used in the AGRIS network (AGRIN2, AGRIN3, WEBAGRIS) and FAOBIB (WEB FAOBIB) are based on UNESCO's software CDS/ISIS, and the main exchange format is ISO 2709. This format is not compliant with other non-CDS/ISIS based systems and it always requires reformatting, which is expensive and resource consuming process.

There are many different exchange formats, used within AGRIS network like: tag/text, MARC21, ISO exchange, ISO 2709, XML formats.

1.2 Different structure of the record fields (elements)

Different bibliographic database systems contain different levels of detail in their structure of elements (fields). For example in many bibliographic database management systems (for example CDS/ISIS based) the record structure is limited to two-levels of depth: fields, subfields. Therefore, there is a need to retain more metadata elements. Keeping more detailed structure at the input level is not a negative point as it gives a possibility to provide detailed information about the resource, besides allowing for validation, regrouping of elements in different ways for display, sort, search or exchange using a common format.

1.3 Differences in the representation of Record bibliographical levels

AGRIS and FAOBIB are not relational databases and they have a flat multilevel structure. Some of the existing bibliographic database systems were based on the concepts of bibliographic levels. The bibliographic levels were introduced to allow identifying, within a single reference, all bibliographic entries of which the reference itself is composed and which are required to make the description of a documentary unit complete. The main problem with this method is the inefficiency of repeating redundant information, for example, repeating information about a book when cataloguing all its chapters.

---

[1] FAOBIB contains more than 155,000 bibliographic records of FAO documents. The FAOBIB database is located at the HQ but also collects input from other field office libraries which input data directly on-line or send their records in various formats for inclusion in the central repository. FAOBIB can also be considered as a part of AGRIS network as it represents FAO's contribution to the AGRIS system.

[2] *The AGRIS database contains over 3 million bibliographic records, stored in a central repository at FAO. It is a result of the functioning of the AGRIS (International Information System for Agricultural Science and Technology) network, a decentralized cooperative system for bibliographical metadata collection, processing and dissemination with the aim to improve access to full text information resources. Participating national, international and intergovernmental centers from all over the world (totalling 240) input references of different types of literature and especially non-conventional (grey) literature on agriculture subjects. AGRIS resource centers use different formats and systems (tools) for data input, storage and submission to the central repository. Each centre sends the data in for validation, standardization, and translation of subject descriptors (according to the AGROVOC thesaurus) before it is added to the central repository for further dissemination.*

1.4 Inconsistency of existing database fields and DC (AGRIS AP) metadata elements

In local systems, there are many local fields (specific to meet the local requirements) that have no corresponding elements in the common DC compliant format and could be lost in a common repository.

1.5 Problems of different content representation of the same elements

Different systems are based on different standards of bibliographic description.The presence of the same element in all existing systems can't guarantee also unique representation of its content.

Example:

Data in alternative title with unidentified language, original title or transliterated title, coded in different ways, data that should be decoded or referred to a decoding local schema.

2. Main steps to effect the conversion of bibliographic metadata in XML format (based on DC and AGRIS AP)

The main task is how to resolve the problem with interoperability between existing metadata resources and in most of the cases without changing the existing local systems and formats.

One of the major tasks for better interoperability within the collaborative networks (like AGRIS network) for platform independent exchange and access of metadata is to establish a common metadata architecture (a common repository format) accommodating different metadata formats.

Here we present main steps identified during FAO experience with data conversion in more details.

- Creation of a common standard for metadata presentation
- Mapping of metadata fields and subfields from an existing bibliographic metadata (databases or files) to a common XML document structure and/or data architecture based on AGRIS AP
- Development of the reformatting tools and conversion procedures to export existing legacy bibliographic metadata to XML/RDF AGRIS AP compliant files
- Conversions
- Validation with agris.dtd
- Implementation of the new format in a generic model that can facilitate the development and application of common interfaces for retrieval and exchange of metadata.

2.1 Creation of a common standard for metadata presentation

The need for high quality, consistent description of the bibliographic documents led to development of AGMES AGRIS Application profile and corresponding Schema for XML/RDF representation of AGRIS and FAO bibliographic database records.

The Agricultural Metadata Element Set (AgMES), developed by FAO for the description and discovery of all agricultural information resource types, was a good starting point to address the issue of creating a data model for exchange.

AgMES is an extension of Dublin Core for the use of metadata elements/schemes/refinements that are important for the description of information resources in the domain of agriculture and in facilitating both the finding and sharing of information about those resources.

The AGRIS AP contains specifications about the metadata that should be exchanged and disseminated through the AGRIS system. An XML DTD for encoding of this metadata was created and published in February 2003 Any XML document submitted following these specifications can be considered compliant with the AGRIS database.

The document on the new exchange format, entitled AGRIS Application Profile (AP), includes detailed description of the metadata standard created to enhance the description, exchange and subsequent retrieval of agricultural Document-Like Information Objects (DLIOs) and its implementation *as a standard is* available at http://www.fao.org/docrep/008/ae909e/ae909e00.htm Another document giving more details on *Generating AGRIS AP XML from local databases* can be seen at http://www.fao.org/docrep/008/ae908e/ae908e00.htm

We usually discuss the processes used to create and manage **DC compliant metadata** from the different application systems, already based on the XML format. What about the existing metadata collected through the years with so many efforts on the part of the cataloguers?

There is no standard methodology on how to apply the XML standards to the existing bibliographic data (legacy data), which is an important issue at the data provider and service provider level.

Main question for the users is “Should existing bibliographic data be transferred to a new structure (for example AGRIS AP compliant)” or “should be kept the existing data management system and transform data on the fly when needed (for exchange within a network)”. We can’t propose a general recipe working for all cases, but we can share our experience and give some tips, advices and tools that can be of help to the user’s decision.

2.2 Mapping of metadata fields and subfields from an existing bibliographic metadata to a common XML document structure and data architecture based on AGRIS AP

Usually the existing bibliographic databases contains more detailed structure so we have to group one or more fields in an element in AGRIS AP. Keeping more detailed structure at the input level is not a negative aspect, as it gives the possibility to provide detailed information about the resource, and allows for validation and regrouping of elements in different ways for display, sort and search.

We had to define what the common and not common elements are, and where local data should be attached in a common structure (AGRIS AP). The result is compliance of the local metadata with AGRIS AP schema with hierarchical elements and qualifiers.

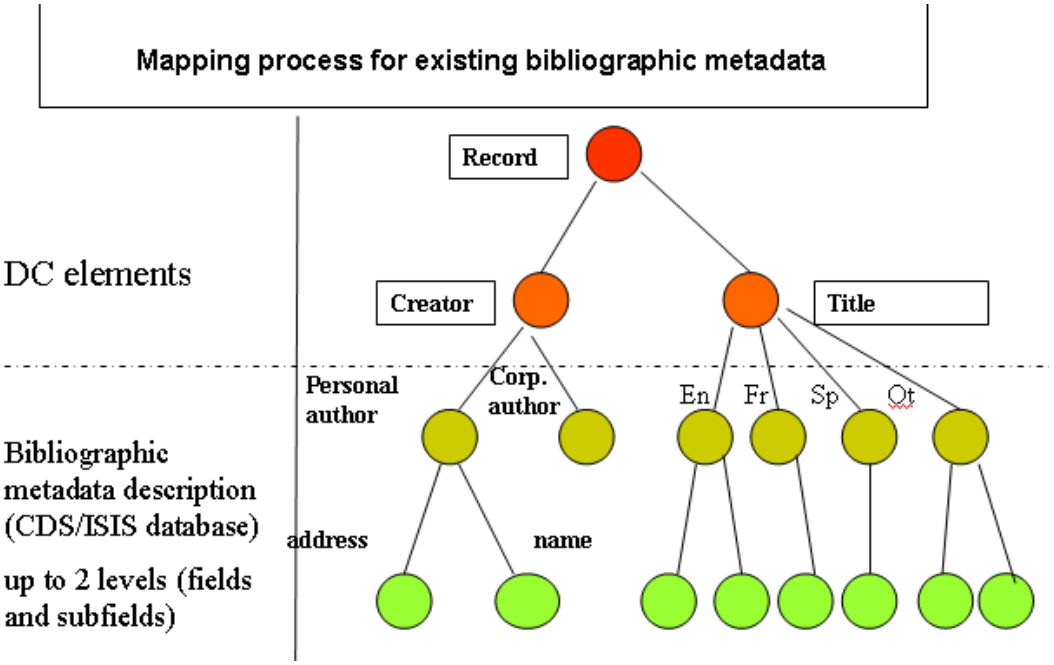


Fig. 1 Represents elements in the existing bibliographic databases and DC (AGRIS AP) elements

In order to define at what stage existing bibliographic element should be mapped and convert to become compliant to AGRIS AP depends on the application and the tools used for metadata collection and management.

The user has to be aware of:

- What format of the output file his application system ensures: XML, HTML or tag/text or ISO exchange on the output
- What is the structure of the local elements and how further can be mapped to DC and AGRIS AP DTD elements

This decision will define also and following steps of the conversion process

For most of the systems that supply XML output the only step was to map the elements from exported XML file to AGRIS AP elements and run a conversion to AGRIS AP structure (using XSLT stylesheet for encoding local database elements).

(AGS) Creator	ags:creatorPersonal	(200) Personal author
	ags:creatorCorporate	(210)Corporate Author
	ags:creatorConference	(230) Conference Name; (231) Place; (233)

		Date

Fig 2 Mapping WEBAGRIS to AGRIS AP

2.3 Development of the reformatting tools and conversion procedures to export existing legacy bibliographic metadata to XML/RDF AGRIS AP compliant files

Once the output XML/RDF data model was created and mapping of the existing local elements to AGRIS AP was done, the next step involved was the creation of **xml convertor.**
The process of preparation of XML file from existing databases includes presentation of each record in XML format according to the documentation . During the export a header and footer should be added to make a valid XML file. So the final file structure contains:

[Header]
(set of elements, refinements and schemes for describing and enforcing the structure that makes up XML format of a bibliographic record).
Example:

```
<?xml version="1.0" encoding="ISO-8859-1"?>
<!DOCTYPE agrisResources SYSTEM "agris.dtd">
<agrisResources xmlns:ags="http://purl.org/agmes/1.1/"
                xmlns:dc="http://purl.org/dc/elements/1.1/"
                xmlns:agls="http://www.naa.gov.au/recordkeeping/gov_online/agls/1.2"
                xmlns:dcterms="http://purl.org/dc/terms/">
```

Main body of the XML file with all documents formatted in XML
[Record1 in XML]
[Record2 in XML]
[Record3 in XML]

[Footer]

A. Reformatting from XML and XML AGRIS AP enabled systems.

The AGMES AGRIS AP standard is not only for the new applications generating metadata but it has to accommodate also the legacy data from the existing applications.
There are different ways of exporting and receiving well formed XML files which can later be transformed (XSLT) to another XML AGRIS AP compliant.
Almost all existing bibliographic database management systems have an extension for exporting records in XML (even if this XML is not compliant with DC or AGRIS AP structure, but has a more detailed structure and higher number of elements that could be easier to be connected and join in common elements)
Such systems are Inmagic systems (CG centers) that produce XML format files but with the structure of MARC21 elements which is not AGRIS AP compliant.
Many other systems that produce XML output can follow the same procedure of mapping and conversion : to map the elements to AGRIS AP and write a transformation XSLT for converting to AGRIS AP XML format.

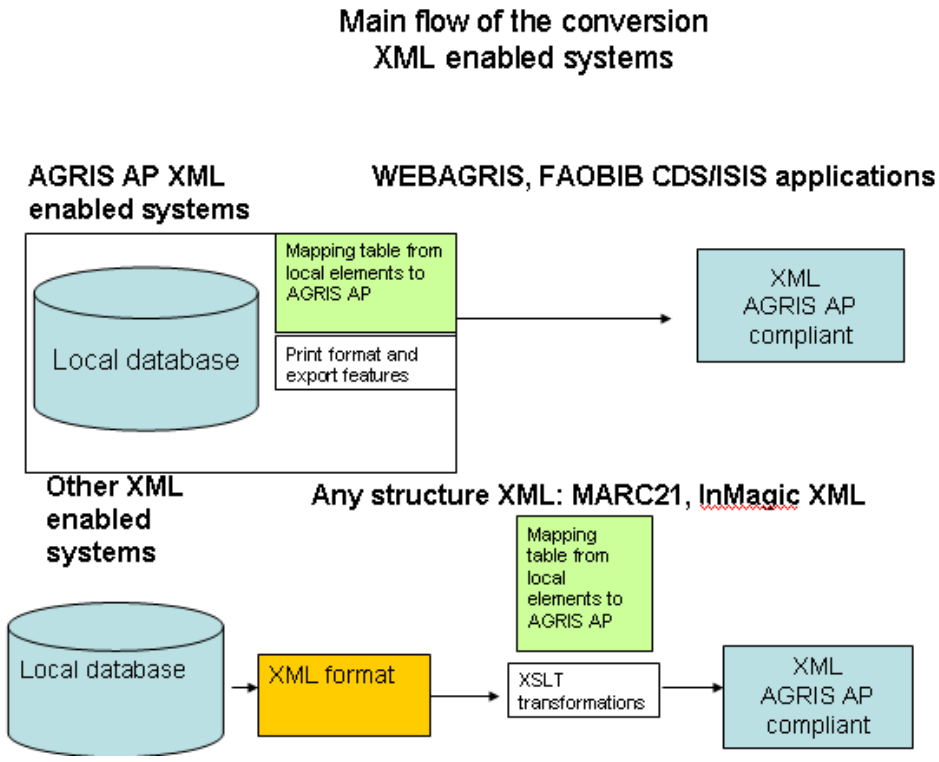


Fig 3 Flow of the conversion process

B. CDS/ISIS based applications: print/display formats, for producing XML files from existing databases.

We could always consider any systems based on DOS or WINDOW version of CDS/ISIS, XML enable as it ensures producing XML output files by applying a print formats.

CDS/ISIS software (DOS, Windows, WEB versions) supplies its users with a powerful formatting language that works as a filter for selection one or more fields, order them in appropriate way. It also gives the option of inserting text in the output. This language is used for printing or displaying records and extraction of data for index creation. The language also allows using of conditional (“”), unconditional(‘’), repeatable literals (||), before or after the fields. If commands and boolean functions allows producing of output, depending of content of data being formatted.

For example this format

'<ags:creator>'/

(|<ags:creatorPersonal>|v100,|</ags:creatorPersonal>|d100/),

|<ags:creatorCorporate>|v110,|</ags:creatorCorporate>|d110/

|<ags:creatorConference>|v130","V131,","V133,

|</ags:creatorConference>|d130/

'</ags:creator>'/

could produce the creator element from fields 100,110,130,131,133 if present:

<ags:creator>
<ags:creatorPersonal>**Garmo, T.H.**</ags:creatorPersonal>

<ags:creatorConference>**Husdyrforsoeksmoetet 1998,Aas (Norway),10-11 Feb 1998**</ags:creatorConference>
</ags:creator>

In order to ensure an XML AGRIS AP compliant output from a CDS/ISIS based local systems, we developed a print format for producing XML output of each of the exported records.

The specific header and footer parts that had to be added to the output XML file. It was partially covered by the options of DOS CDS/ISIS (in title of print worksheet) and WINDOWS CDS/ISIS export to XML option, but it didn't gave us possibility to integrate more complex header (with namespaces) as it was done (by ensuring separate print formats for generating header and footer) in the WEBAGRIS application features.

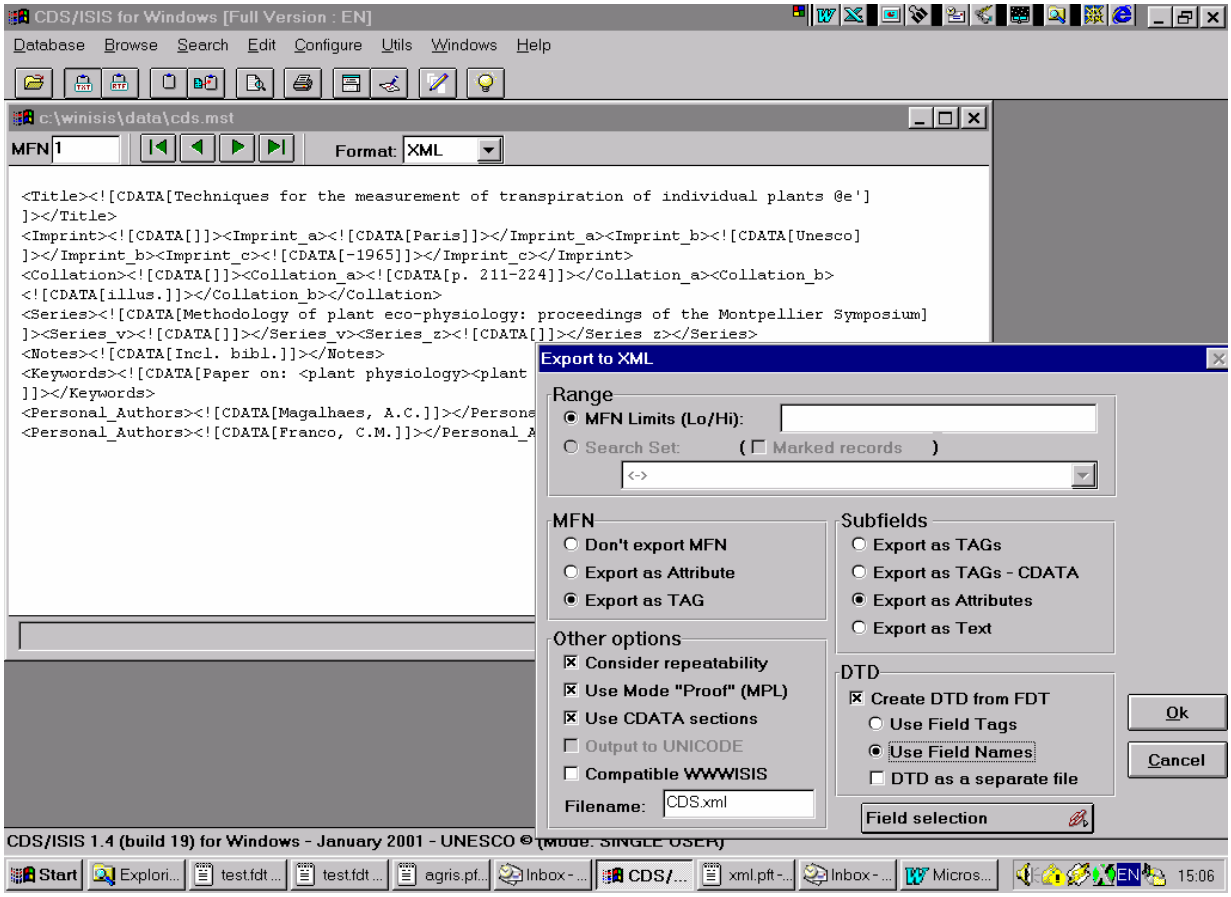


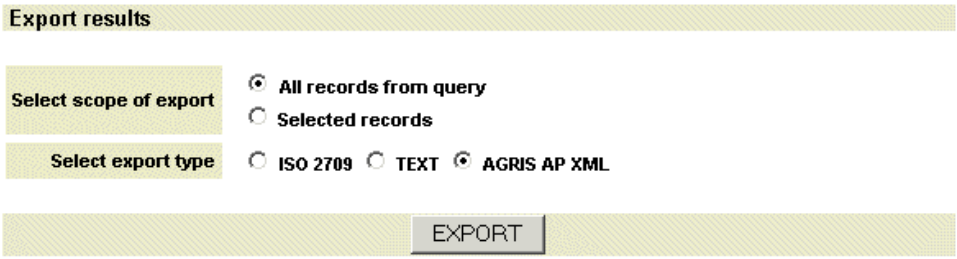
Fig .4 shows export in XML process from Windows version of CDS/ISIS

We have already developed tools integrated in the system to export with applying directly a format for output, ensuring AGRIS AP compliance so we can just go further to validate the XML file.

WEBAGRIS and FAOBIB applications are one of these new advanced multilingual WEB based interfaces for distributed data input, processing and dissemination (through the Internet or on CD-Rom), of agricultural bibliographic information. It is based on WWW-ISIS interface and working with CDS/ISIS databases. It was developed for the AGRIS resource centers.

The format for export is integrated in the system, so that in this way selecting this format for the output data, the user can download and exchange data in XML AGRIS AP format.

The following screen gives example of reformatting process from WEBAGRIS and FAOBIB applications how to export all search results or selected records in AGRIS AP XML format.



The role of these print formats is much bigger than a simple transformation of data in a repository pool as this reformatting can be done at any time when needed for example during display of search results or in the WEB services wrappers for integration in a common XML format search results from different applications and data source formats.

We develop a tool for exporting from currently used local applications based on WEBAGRIS systems to XML format according to predefined DTD structure (meta-data standard requirements). A new output from WEBAGRIS system ensures producing of XML at any local application level. This was an important achievement for all WEBAGRIS users, which can replace the existing ISO 2709 based exchange of data with exchange in XML format. This feature will make them independent for participating in cooperative systems with a common collection of metadata or offering their metadata for harvesting or multidatabase search.

One of XML/RDF implementation in FAO bibliographic databases was to convert FAOBIB records to XML (AGRIS AP compliant) as describing different type of resources, in order to finalize exportation of FAOBIB documents (more than 155000) according to a modified agris.dtd. some compromises had to be done as the original agris.dtd was too rigid for legacy FAOBIB data.

Another interesting application was moving of some local application to WEBAGRIS (AGRIS AP XML compliant) and then to use already prepared format for XML AGRIS AP. We help in moving of ICIMOD application data (28000 records) to WEBAGRIS and exporting using WEBAGRIS print format. This solution is only appropriate if WEBAGRIS could be further used as data management application and not only for conversion purposes.

C. Conversion of other formats files (XML or different than XML formats) to AGRIS AP XML files

We gained experience in converting structured legacy data coming from bibliographic repository of AGRIS pool to XML format, by bringing the legacy data (ISO exchange format) to XML AGRIS AP compliant format. Java program that uses ISIS_DLL version 7 converts raw ISO Exchange data into XML AGRIS AP compliant format (one record – one file form for Lucene search engine testing). It generates also ISIS format database for comparison and test purposes. This approach we can use to convert AGRIS legacy data (3mln records) to a common repository XML/RDF pool.

If the local system used doesn't supply any output that can be reformatted to XML may be the user wants to change his old system to a new one. Here we have already developed XML enabled systems, which already integrate the production of XML (AGRIS AP) on the output stage.

In this case mapping from the existing system (for example (AGRIN2, AGRIN3, ICIMOD CDS/ISIS based etc.) to the new one (WEBAGRIS, EIMS, FAOBIB for example) should be done and then the new systems already have been mapped and export in XML AGRIS AP was ensured .

Here we have to prepare and run a convertor at the local application or file, reformatting the output to make it compliant with the AGRIS AP XML format. The convertor can be XSLT transformation or reformatting program

Example: some AGRIS centers (Finish , Dutch) producing XML format metadata that could join AGRIS repository pool. These XML files could easily be transformed into a hierarchical result tree XML, according to the standard, using XSLT stylesheet.

This approach is recommended for non CDS/ISIS systems as it is based on the possibility of keeping flexibility for further adjustment of the record structure, according to changes done in the local database or to the metadata standard.

2.4 Some problems we had to resolve during mapping and conversion to AGRIS AP compliant elements:

During realization of our experiment there were several problems encountered and most important of which are mentioned here.

Loss of data not meeting the criteria's of the standard – many local data fields not defined in the standard structure were present as part of common fields (financial information, notes, summary, type of documents) or were missing in the final format.

Lack of some mandatory elements/information. This was one of the most difficult problems we had to resolve at the conversion. If some of the mandatory elements like Date of publication, subject, title were not present we had to take out the document as not well formed. This is a time consuming operation.

Multilingual data contained in the databases and lack of appropriate attributes identifying the language (requested by the new standard). Here we had an example with titles in different languages specified as other language title. The validation required explicit specification of the language otherwise the records could pass the validation.

Usage of unique identifier of record (as relation between the records analytic and monograph) identifying the source document and relations with the whole needs further standardization.

Consistency and quality of data (character encoding, presence of Triangle brackets<>, not Unicode compliant data from DOS WINDOWS or WEB environment) was another frequently met problem.

Applying of coded value for concepts or language identification was essential in order to facilitate implementation of ontology and semantic WEB search in different languages. In case of legacy data we met different specifications for the language. For example: (EN,FR), full text language (English, French) etc. and it was very difficult to make this compliant with the validation against namespace with the codes of languages.

the inefficiency of repeating information of the higher level in the description of the document that is part of the larger whole. We needed to separate the metadata information of the Analytical (chapter of a book) from the metadata of the larger whole Monograph (book) and have them stored as separate unit linked to each other using the Relation field.

Fig. 1 shows the systems, based on a flat record structure contained three levels of information A(analytic), M(Monograph) and S (Serial)

English title:	Shared fishery Argentine-Uruguayan common fishing zone
English title (mon):	Papers presented at the Norway - FAO Expert Consultation on the Management of S Bergen, Norway, 7-10 October 2002
Serial:	FAO Fisheries Report (FAO). 0429-9337, no. 695(Suppl.)
Authors:	Chaluleu, J.D.;
Conference:	Norway-FAO Expert Consultation on the Management of Shared Fish Stocks. Berger 2002
Corp.authors:	FAO, Rome (Italy). Fishery Policy and Planning Div.
Division:	FIP
Publ.place:	Rome (Italy)
Publisher:	FAO
Publ.date:	2003
Collation:	p. 86-104
ISBN:	92-5-104936-X
Languages:	English
IC/IY(2):	XF03
Category codes:	M11 D50
Categories:	M11-Fisheries production D50-Legislation
AGROVOC main descr. :	FISHERY MANAGEMENT; FISHERY RESOURCES; REGULATIONS; INTERNATIONAL AGREEI ASSESSMENT; FISHING AREAS
AGROVOC geogr. descr.:	ARGENTINA; URUGUAY
Publ.type:	M
Report No:	FAO-FIPP--R695(Suppl.)
Call No:	S238 8 ENG.ED
Related monograph:	406263

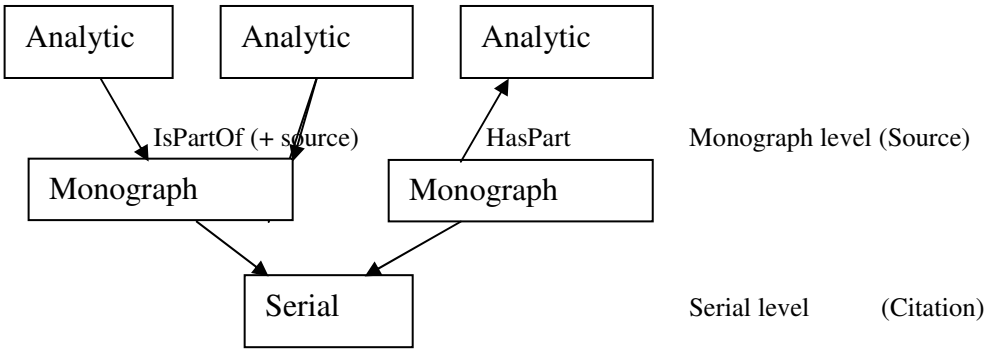


Fig 5. Presentation of sample Record from FAOBIB as flat and relational structure

The sample record (Figure 2) provides an example of FAOBIB presented as flat and relational structure. In the analytic record there were three titles: English Title, Title of the Monograph, and a Serial Title. The unique title for the catalogued unit is the lowest level (Analitic) title and other titles are presented in the relations (IsPartOf) (or Source) and Citation.

However, it is noted that the Relation element is not able to cater for the AGRIS and FAOBIB system because the current proposed value for this element is limited to a Scheme i.e. an Identifier. Where as, there is abundant metadata information of the larger whole part which is considered necessary and should be retained. This therefore led to the introduction of the elements Source and Citation.

2.5 Solutions of the problems with legacy data during validation against agris.dtd

During our experience with validation process for producing a valid XML files we had to resolve different problems mentioned in 2.4.
Here are some solutions we adapt:

- all local data not having a special element was written usually in notes field
- records with missing mandatory fields (date, availability) usually in legacy data. We had tree possibilities: to delete these records, to update them and add mandatory elements or to change in agris.dtd schema making them optional. This is a long process and decision has to be taken depending on the application which will be using these data
- for many fields we couldn't identify the language attribute (alternative title) and as the system doesn't validate it we have to create a fictive value for language or to loose the record
- Unique resource identification to be used as reference to related record. Here we met some problems in case that we prepare a subset of the database (extraction for CD-ROM production) for distribution to the centres without Internet connection.
- all coded values should be converted during the conversion process
- Unicode usage . Many input files are using different code page and so in the process of merging we have to convert them or to be aware of the character set used and reflect it in the header of the produced file

We discover that the exported in XML format files are not always readable and encoded properly.
The reason for this is that in the WEBAGRIS software the header and footer are automatically added by WebAGRIS using the hdrxml.pft and ftrxml.pft files from the upgrades.zip file.
The header format generates the following header:

```
<?xml version="1.0" encoding="ISO-8859-1"?>
<!DOCTYPE agrisResources SYSTEM "agris.dtd">
<agrisResources xmlns:ags="http://purl.org/agmes/1.1/"
                xmlns:dc="http://purl.org/dc/elements/1.1/"
                xmlns:agls="http://www.naa.gov.au/recordkeeping/gov_online/agls/1.2"
                xmlns:dcterms="http://purl.org/dc/terms/">
```

The proposed header format is ready for Latin-1 character set and it should cover all latin characters (English, French, Spanish, Italian, etc). This is reflected in the parameter: encoding="ISO-8859-1".
The format hdrxml.pft should be changed if you are not using only Latin-1 character set .
You have to discover which encoding is used for your national character set. Here are some samples for encoding values we discovered by testing different inputs

For Bulgarian the parameter should be: encoding="windows-1251" ,
Polish is using encoding="ISO-8859-2",
Ukraine is using encoding="Windows-1251"
Unfortunately in case that wrong encoding was used we can't change it after exporting has been done.

In the same HTML page or XML file only one encoding page could be specified in the header. The output from CDS/ISIS applications has to concider one encoding page (LATIN-1 or one National character set plus LATIN 1 without usage of accented characters). This is one of the main limitation of CDS/ISIS application now. The solution of creating parcer for UTF-8 decoding needs a specification of the language in all the fields of multilingual records. A good advice is to add a subfield, identifying language attribute (for example for other language title or abstract in other languages to have ^l with language specification).

XMLSpy editor was used to validate against the AGMES AP (agris.dtd was used to determine the validity of an XML document.

3. Implementation and result from the experience

This experience gives a renewed value of the existing applications based on CDS/ISIS, as it shows that CDS/ISIS database applications can comply with DC metadata standard as the matching between existing and new standards can be done at any time on data exchange (input, export, harvesting) by reformatting using print format language of CDS/ISIS to produce HTML or XML format files that can be later transformed if needed to another XML or RDF format using XSLT transformation.

The main result we are aiming at was a **centralized metadata repository** that joint (FAOBIB, AGRIS legacy, new AGRIS input) collection in a common standard format. All this metadata from AGRIS and FAOBIB databases will ensure more possibilities to re-use it applying appropriate for the user's tools: different retrieval interfaces to harvest our data or multiphase searching usually retrieving heterogeneous information resources and thus joining our metadata with other metadata resources (AGRICOLA) and Open Archives.

The experience with conversion gave us a lot of feedback how we can improve the XML schema standards for metadata description as well as our bibliographic data formats for better compliance with other systems.

4. Conclusions

FAO gained a lot of experience in creating metadata standards and applying them to the existing bibliographic metadata. Thus almost all existing bibliographic database management systems have become XML enabled by ensuring XML output on some stages of the data processing.

We are able to help the users in conversion process to AGRIS AP compliant XML format, writing appropriate convertors that use print format for CDS/ISIS applications or XSLT transformations. XSLT transformation seems to be a powerful and most appropriate tool for conversion and further comply XML with AGRIS AP.

We gained experience not only on exporting XML from CDS/ISIS databases but also on input XML data into CDS/ISIS database (which is important especially for web applications). The HTML form of data entry worksheet from our data entry system may be included and called from any system that can produce XML data output. In this way all the functionality (i.e. data entry to our database, validations) will be integrated "on the fly". This have been tested for the transfer of XML data from non-CDS/ISIS database (DOCREP) to the FAO library catalogue.

The main problem was the integration the data from all systems (DOS, WINDOWS WEB CDS/ISIS applications) using UTF-8 coding. CDS/ISIS applications being not UTF-8 compliant create problems when they assume multilingual data usage.

Links to:

1. The AGRIS Document Type Definition (DTD)
   http://www.fao.org/agris/agmes/Documents/20040421_agris_xml.dtd.txt

2. Mapping tables to AGRIS Document Type Definition (DTD) elements:Mapping WEBAGRIS fields to AGRIS AP elements
   http://www.fao.org/agris/tools/AGRIS_AP/agrisAP_agrisInputfields_R1.doc
3. WEBAGRIS application: http://www4.fao.org/agris
4. FAOBIB application: http://www4.fao.org/faobib

5. An instance of an XML AGRIS record

```
<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE agrisResources SYSTEM "http://purl.org/agmes/agrisap/dtd/">
<agrisResources  xmlns:ags="http://purl.org/agmes/1.1/"
                 xmlns:dc="http://purl.org/dc/elements/1.1/"
                 xmlns:agls="http://www.naa.gov.au/recordkeeping/gov_online/agls/1.2"
                 xmlns:dcterms="http://purl.org/dc/terms/">

<agrisResource ags:ARN="NL2004700134">
        <dc:title xml:lang="en">Effect of oxidation ditch horizontal velocity on the nitrogen removal process</dc:title>
        <ags:creator>
            <ags:creatorPersonal>Abusam, A.</ags:creatorPersonal>
            <ags:creatorPersonal>Keesman, K.J.</ags:creatorPersonal>
            <ags:creatorPersonal>Spanjers, H.</ags:creatorPersonal>
        </ags:creator>
        <dc:date>
            <dcterms:issued>2002</dcterms:issued>
        </dc:date>
        <dc:subject xml:lang="en">
            <ags:subjectClassification scheme="ags:ASC">P10</ags:subjectClassification>
            <ags:subjectThesaurus xml:lang="en" scheme="ags:CABT">WASTE WATER</ags:subjectThesaurus>
            <ags:subjectThesaurus xml:lang="en" scheme="ags:CABT">NITRATES</ags:subjectThesaurus>
            <ags:subjectThesaurus xml:lang="en" scheme="ags:CABT">REMOVAL</ags:subjectThesaurus>
            <ags:subjectThesaurus xml:lang="en" scheme="ags:CABT">OXIDATION DITCHES</ags:subjectThesaurus>
            <ags:subjectThesaurus xml:lang="en" scheme="ags:CABT">PERFORMANCE</ags:subjectThesaurus>
        </dc:subject>
        <dc:description>
            <ags:descriptionNotes>12 refs</ags:descriptionNotes>
        </dc:description>
        <dc:identifier scheme="dcterms:URI">http://www.ewaonline.de/journal/2002_06.pdf</dc:identifier>
        <dc:format>
            <dcterms:extent>p. 213</dcterms:extent>
            <dcterms:medium>internet</dcterms:medium>
        </dc:format>
        <dc:language scheme="ISO639-1">en</dc:language>
        <agls:availability>
            <ags:availabilityLocation>Library Wageningen University and Research Centre, Postbus 9100, 6703 BK
            Wageningen ub.library@wur.nl http://library.wur.nl/desktop/</ags:availabilityLocation>
            <ags:availabilityNumber>1700134</ags:availabilityNumber>
        </agls:availability>
        <ags:citation>
            <ags:citationTitle xml:lang="en">European water management online</ags:citationTitle>
            <ags:citationChronology>2002</ags:citationChronology>
        </ags:citation>
    </agrisResource>
</agrisResources>
```