



AGRIS: an RDF-aware system in the agricultural domain

Fabrizio Celli, Johannes Keizer, Maria
Folch, Armando Stellato

Fabrizio Celli, Food and Agriculture Organization of the United Nations
LOD 2014 – Roma, 20th / 21st Feb 2014

Outline

- What is AGRIS?
- The RDF-ization process
- Interlinking

WHAT IS AGRIS?

One or more definitions?

- A **network**: AGRIS is collaborative network of more than 150 institutions from 65 countries
- A **database**: AGRIS is a multilingual bibliographic database for agricultural science
- A **Web portal**: AGRIS (<http://agris.fao.org/>) is a Web application that links the AGRIS knowledge to related Web resources using the Linked Open Data methodology
 - Purpose: providing as much information as possible about a topic within the agricultural domain

The setting

- The AGRIS database is a collection of more than 7.7 million bibliographic references in the agricultural domain
- They are enhanced by the AGROVOC thesaurus, which is extensively used by cataloguers to enrich data indexing in agricultural information systems
- AGROVOC is the FAO multilingual vocabulary containing more than 40 000 concepts in up to 21 languages. It's part of the LOD cloud!

The setting (2)

- AGRIS is an RDF-aware system, a mashup application that allows users to query the AGRIS content, interlinking all resources to external sources of information
- 7.7 million bibliographic references result in 190 million triples (~24 triples/reference)
- 7.7 million bibliographic references become 7.7 million mashup pages!

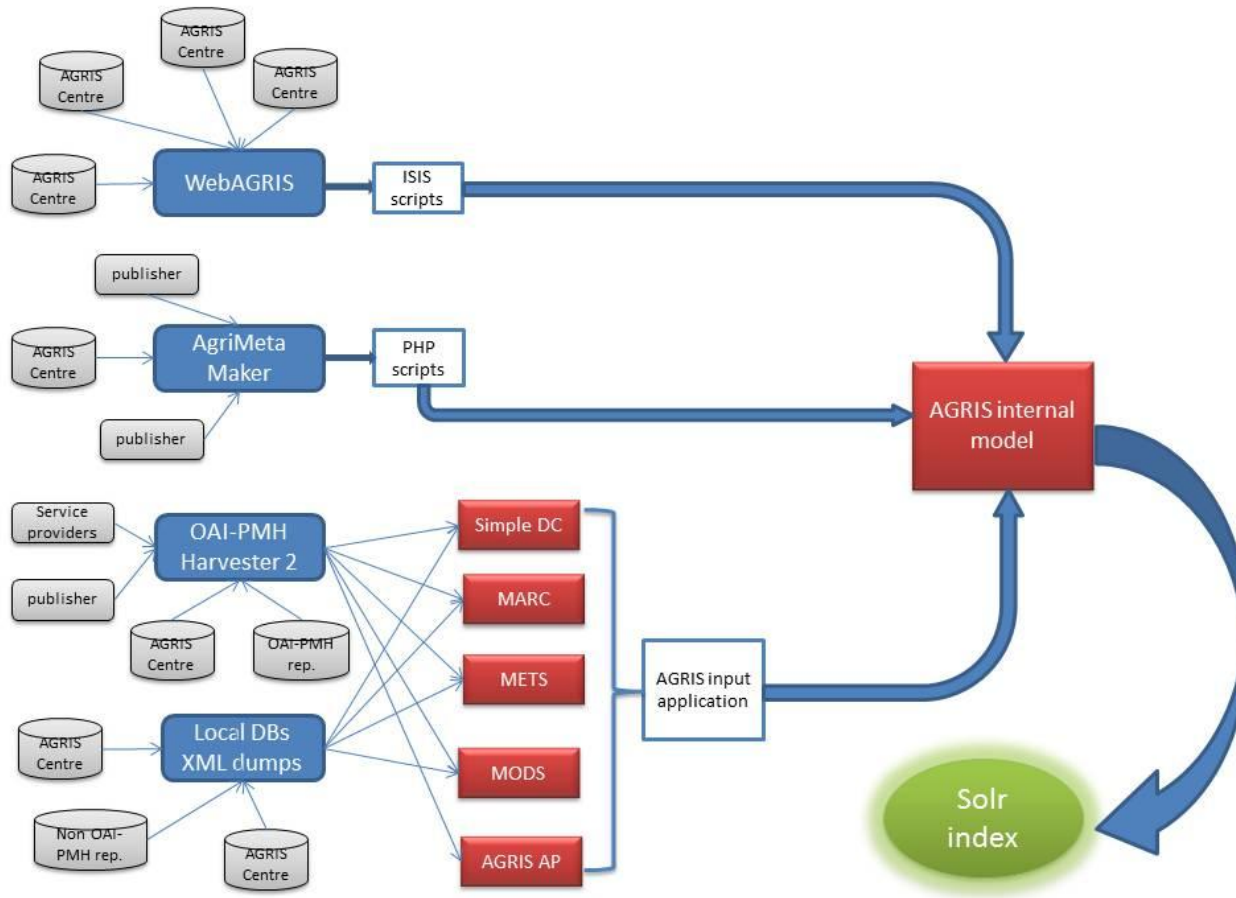
Some statistics

- 7.7 million bibliographic references
- 190 million triples
- ~ 300.000 visits/month
- World wide used (accessed from more than 200 countries)

How data come to AGRIS

- **Centralization:** bibliographic references in the AGRIS domain (agriculture, forestry, animal husbandry, aquatic sciences and fisheries, and human nutrition)
- **Interlinking:** other kinds of information related to the AGRIS domain (statistics, maps, country profiles, etc.)

Accept any input format!



THE RDF-IZATION PROCESS

The RDF-ization process

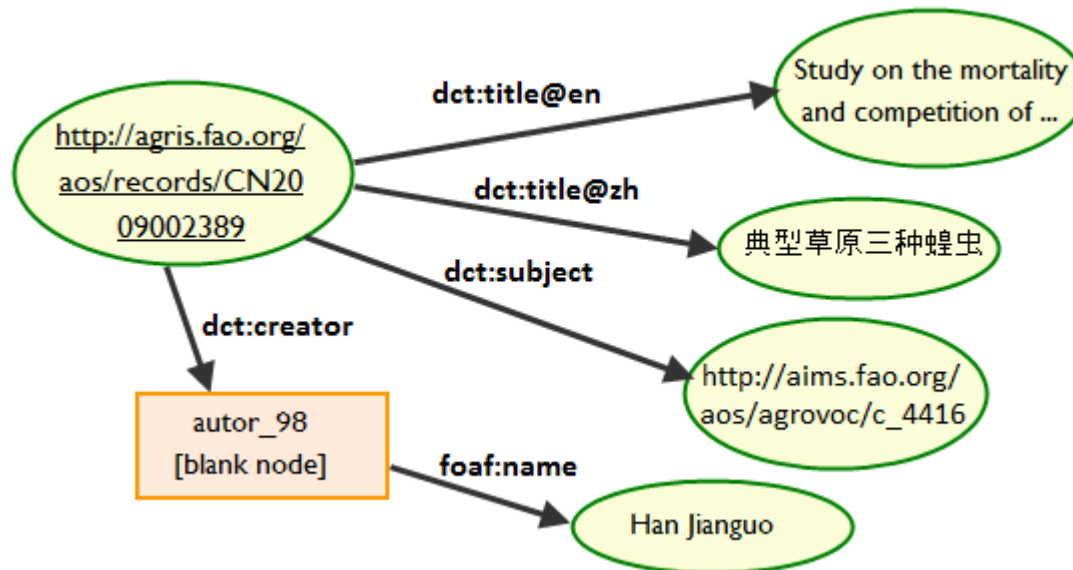
- Translation of the AGRIS AP XML database to RDF
 - Selection of existing vocabularies to model triples (when desired properties are already available)
 - Creation of new properties when needed
 - Data cleaning and normalization
 - Index all records with the Agrovoc thesaurus, to help interlinking to other datasets
 - Run the conversion and publish RDF data!
- Selection of external datasets we want to interlink to AGRIS

An AGRIS AP XML record

```
<ags:resource ags:ARN="CN2009002389">
  <dc:title xml:lang="en">Study on the mortality and competition of ...</dc:title>
  <dc:title xml:lang="Zh">
    <dcterms:alternative>典型草原三种蝗虫种群死亡率和竞争的研究</dcterms:alternative>
  </dc:title>
  <dc:creator>
    <ags:creatorPersonal>Lu Hui</ags:creatorPersonal>
    <ags:creatorPersonal>Han Jianguo</ags:creatorPersonal>
  </dc:creator>
  <dc:date><dcterms:dateIssued>Sep. 2008</dcterms:dateIssued></dc:date>
  <dc:subject>
    <ags:subjectThesaurus scheme="ags:AGROVOC" xml:lang="en">ACRIDIDAE</ags:subjectThesaurus>
  </dc:subject>
  <dc:description>
    <dcterms:abstract xml:lang="Zh">通过田间...</dcterms:abstract>
    <dcterms:abstract xml:lang="En">The three ...</dcterms:abstract>
  </dc:description>
  <dc:language scheme="ags:ISO639-1">Zh</dc:language>
  <ags:citation>
    <ags:citationTitle xml:lang="Zh">Acta Agrestia Sinica</ags:citationTitle>
    <ags:citationIdentifier scheme="ags:ISSN">1007-0435</ags:citationIdentifier>
  </ags:citation>
</ags:resource>
```

Translation to RDF

- Assign a unique, permanent URI to the AGRIS record:
<http://agris.fao.org/aos/records/CN2009002389>
- The URI should be dereferenciable, i.e. used in a browser, it should return the description of the resource
- Other URIs: Agrovoc keywords, AGRIS journals



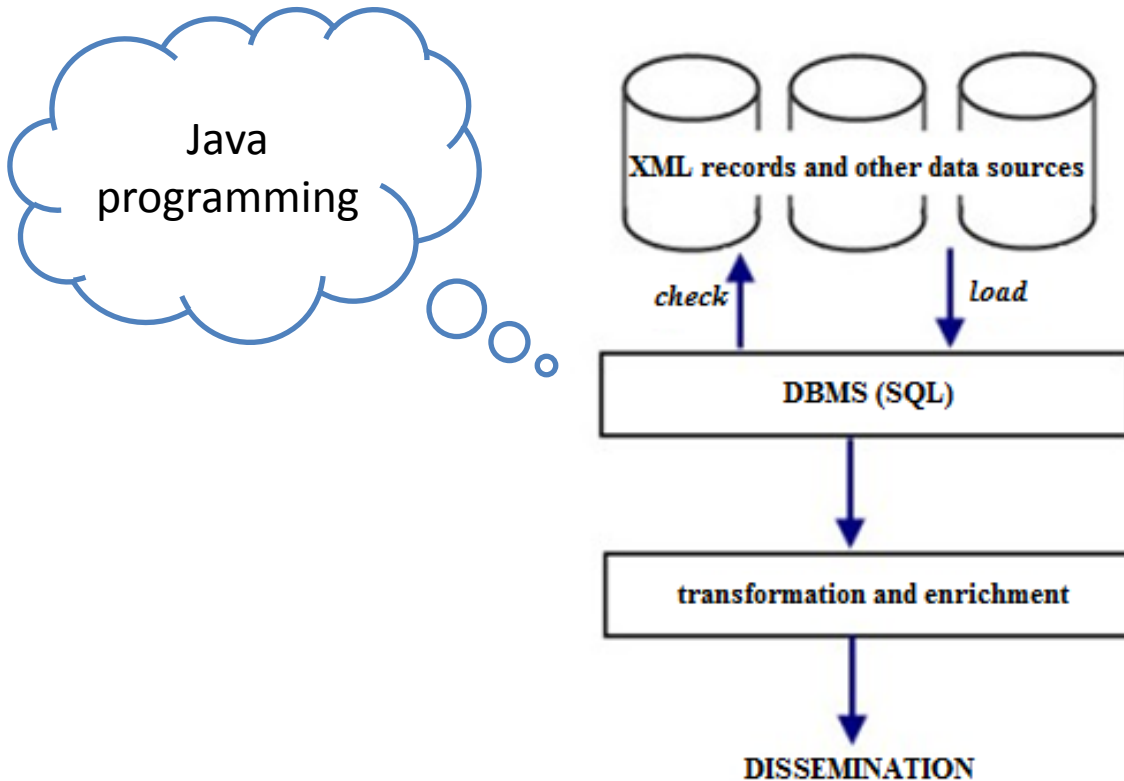
AGRIS RDF

bibo:Article
bibo:abstract
bibo:doi
bibo:isbn
bibo:presentedAt ->
 bibo:Conference -> dct:title
bibo:uri
dct:alternative
dct:creator ->
 foaf:organization -> foaf:name
dct:creator ->
 foaf:Person -> foaf:name
dct:dateSubmitted
dct:description
dct:extent
dct:identifier
dct:language
dct:isPartOf
dct:issued
dct:publisher ->
 foaf:Organization -> foaf:name
dct:source
dct:subject
dct:title
dct:type
dct:rights

RDF/XML snapshot

```
<bibo:Article rdf:about="http://agris.fao.org/aos/records/XS2010X00001">
  < dct:identifier>XS2010X00001 </dct:identifier>
  < dct:title xml:lang="pt"> Características anatómicas ... </dct:title>
  < dct:title xml:lang="en"> ... </dct:title>
  < dct:creator>
    < foaf:Person> < foaf:name> Mesquita, Alessandro Carlos </foaf:name> </foaf:Person>
  </dct:creator>
  < dct:publisher>
    < foaf:Organization> < foaf:name> Instituto Nacional de ... </foaf:name> </foaf:Organization>
  </dct:publisher>
  < dct:issued> 2010 </dct:issued>
  < dct:subject rdf:resource="http://aims.fao.org/aos/agrovoc/c_6200"/>
  < bibo:abstract xml:lang="pt"> <![CDATA[As estruturas envolvidas na produção ...]]> </bibo:abstract>
  < bibo:abstract xml:lang="en"> <![CDATA[The structures involved in latex production ...]]> </bibo:abstract>
  < bibo:uri> <![CDATA[http://www.scielo.br/scielo.php?pid=...]]> </bibo:uri>
  < bibo:language> por </bibo:language>
  < dct:isPartOf rdf:resource="http://aims.fao.org/serials/c_e8d916a8"/>
</bibo:Article>
```

The process



«OpenAGRIS: using bibliographical data for linking into the agricultural knowledge web»; Fabrizio Celli, Stefano Anibaldi, Maria Folch, Yves Jaques, Johannes Keizer.

Provenance

- AGRIS records come with an identifier (ARN) which has a predefined structure and contains implicit information about the data source together with the bibliographic record's year of creation
- “IT 2008 0 00091” refers to the 91st record created in 2008 by the AGRIS data provider in Italy (IT 0)
- Data providers information are stored in the CIARD RING (<http://ring.ciard.net/>) and published as RDF

INTERLINKING

Interlinking

- Centralization: bibliographic references in the AGRIS domain (agriculture, forestry, animal husbandry, aquatic sciences and fisheries, and human nutrition)
- **Interlinking**: other kinds of information related to the AGRIS domain (statistics, maps, country profiles, etc.)

What good is AGROVOC?

- AGROVOC is the backbone, the magic that allows the interlinking to external datasets
- Two ways to implement the interlinking:
 - Using AGROVOC formal alignments to other thesauri (skos:exactMatch, skos:closeMatch)
 - Querying external WebServices with scientific names, extracted from AGROVOC (no RDF, simply Java programming)



Acrididae

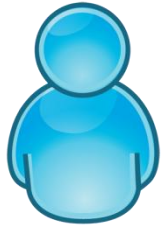
at AGROVOC Thesaurus

http://aims.fao.org/aos/agrovoc/c_4416

Property	Value
altLabel	<ul style="list-style-type: none">■ Criquet (fr)■ آکریدیده (fa)
broader	<ul style="list-style-type: none">■ <http://aims.fao.org/aos/agrovoc/c_5433>
closeMatch	<ul style="list-style-type: none">■ <http://dbpedia.org/resource/Acrididae>
exactMatch	<ul style="list-style-type: none">■ <http://lod.nal.usda.gov/nalt/3016>
<i>date of creation</i>	<ul style="list-style-type: none">■ 1981-01-09
<i>date of last update</i>	<ul style="list-style-type: none">■ 2007-04-20
hasMember	<ul style="list-style-type: none">■ <http://aims.fao.org/aos/agrovoc/c_31924>
hasTaxonomicLevel	<ul style="list-style-type: none">■ <http://aims.fao.org/aos/agrovoc/c_330937>
isThemeOf	<ul style="list-style-type: none">■ <http://aims.fao.org/aos/agrovoc/c_49974>
narrower	<ul style="list-style-type: none">■ <http://aims.fao.org/aos/agrovoc/c_1210>■ <http://aims.fao.org/aos/agrovoc/c_2349>■ <http://aims.fao.org/aos/agrovoc/c_26587>■ <http://aims.fao.org/aos/agrovoc/c_31925>■ <http://aims.fao.org/aos/agrovoc/c_31926>■ <http://aims.fao.org/aos/agrovoc/c_32308>■ <http://aims.fao.org/aos/agrovoc/c_32309>■ <http://aims.fao.org/aos/agrovoc/c_33397>■ <http://aims.fao.org/aos/agrovoc/c_33742>
prefLabel	<ul style="list-style-type: none">■ ACRIDIDAE (de)■ Acrididae (lo)■ Acrididae (hu)■ Acrididae (sk)■ Acrididae (cs)■ Acrididae (th)■ Acrididae (es)■ Acrididae (en)■ Acrididae (fr)

The IFPRI case

(International Food Policy Research Institute)



A user queries the system



AGRIS record with
Agrovoc keywords



At least one Agrovoc
keyword is a Country name



The system queries IFPRI sparql endpoint (<http://data.ifpri.org/sparql/>) to retrieve the global hunger index (GHI) and the child mortality rate related to the Country

Global Hunger Index ("low hunger" < 4.9)

GHI(2012): 19.7 [Serious]

GHI(2001): 21.7 [Alarming]

Child Mortality

Mortality rate of children(2012): 8.7 %

FROM: IFPRI

Powered by Google™
Read the article and/or related articles:

▼ Length-frequency compositions and weightlength relations for ...
Dec 1, 2008 ... Length-frequency compositions and weight-length for bigeye tuna, yellowfin tuna, and albacore (Perciformes: Scombrinae) in th
Go to the page
[PDF] Guopina ... U 1, 2, 3, Liuxia ... Acta Ichthyologica et Piscatoria

Data from Europeana

▼ Rice irrigation in Texas...
Taylor, Thomas Ulvan
Language: English
Data provider: Library of Congress (archive.org)
Type: TEXT
Go to reference

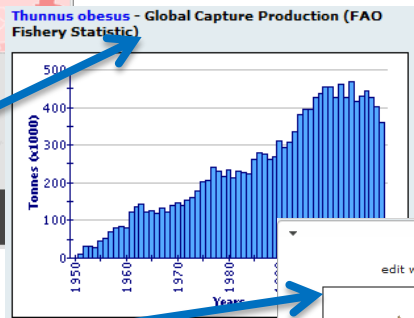
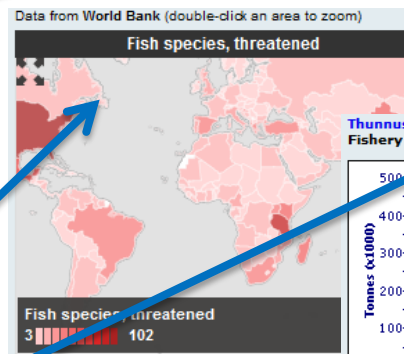
▼ Acta Ichthyologica et Piscatoria (Journal)
FREQUENCY: Semiannual (2 numbers a year)
START DATE: 1972

▼ Agris articles from the same journal:
Interaction of growth hormone (LH) and common carp, *Cyprinus carpio* L.
Source:
Centralna Biblioteka Rolnicza/Central Agricultural Library
CBR is a scientific library subordinated to the Ministry of Agriculture and Rural Development. It has broad collections - thematic: agriculture, food processing
Homepage: <http://www.cbr.gov.pl/eng/index.php>

Data from www.nature.com
▼ Climatology: Extremes in the Indian Ocean
David Anderson [1999-09-23]
In contrast to the Atlantic and Pacific Oceans, the Indian Ocean has seemed to manifest little coherent behaviour in the variability of such parameters as water temperature and rainfall. Now, however, two studies identify a so-called dipole mode in which the distribution of ocean temperature can be related to rainfall patterns in east Africa and In...

```
<bibo:Article rdf:about="http://agris.fao.org/aos/records/XS2010X00001"
<dct:identifier>XS2010X00001</dct:identifier>
<dct:title xml:lang="pt">Características anatómicas ...</dct:title>
<dct:title xml:lang="en">...</dct:title>
<dct:creator>
  <foaf:Person rdf:resource="http://aims.fao.org/aos/agrovoc/c_62000" />
</dct:creator>
<dct:publisher>
  <foaf:Organization rdf:resource="http://aims.fao.org/aos/agrovoc/c_62000" />
</dct:publisher>
<dct:issued>2010</dct:issued>
<dct:subject rdf:resource="http://aims.fao.org/aos/agrovoc/c_62000" />
<bibo:abstract xml:lang="pt"><![CDATA[As estruturas envolvidas na produção ...]]></bibo:abstract>
<bibo:abstract xml:lang="en"><![CDATA[The structures involved in latex production ...]]></bibo:abstract>
<bibo:uri><![CDATA[http://www.scielo.br/scielo.php?pid=...]]></bibo:uri>
<bibo:language>por</bibo:language>
<dct:isPartOf rdf:resource="http://aims.fao.org/serials/c_e80016a8" />
</bibo:Article>
```

AGRIS RDF RECORD



Tuna
edit wikipedia W

Tuna are a group of salt water fish from the family Scombridae, particularly of the genus Thunnus. Tuna are fast swimmers, and some species are capable of speeds of 70 km/h (43 mph). Unlike most fish, which have white flesh, the muscle tissue of tuna ranges from pink to dark red. The

India
http://aims.fao.org/aos/agrovoc/c_3823

AGROVOC

Property: alt_label
Value: India (English), India (Indonesian), República de la India (Spanish), República de l'Índia (Catalan), República de l'Índia (Basque), India (Portuguese), India (Urdu), India (Hindi), India (Tamil), India (Telugu), India (Kannada), India (Malayalam), India (Sinhala), India (Burmese), India (Thai), India (Vietnamese), India (Khmer), India (Javanese), India (Sundanese), India (Balinese), India (Malay), India (Tagalog), India (Filipino), India (Vietnamese), India (Khmer), India (Javanese), India (Sundanese), India (Balinese), India (Malay), India (Tagalog), India (Filipino)

the Republic of Turkey

POPULATION TOTAL (2010): 72 752 000 [SOURCE]
Gross domestic product
GDP TOTAL (2009): 614 603 (millions of US dollars) [SOURCE]
Human Development Index (0 - 1 / low - high)
HDI TOTAL (2010): 0.679 (Calculated) [SOURCE]
FROM: FAO Country Profiles



Bibliography

- «OpenAGRIS: using bibliographical data for linking into the agricultural knowledge web». Fabrizio Celli, Stefano Anibaldi, Maria Folch, Yves Jaques, Johannes Keizer. AOS 2011
- «Proof and Trust in the OpenAGRIS Implementation». Yves Jaques, Stefano Anibaldi, Fabrizio Celli, Imma Subirats, Armando Stellato, Johannes Keizer. DC-2012
- «Pushing, Pulling, Harvesting, Linking - Rethinking Bibliographic Workflows for the Semantic Web». Fabrizio Celli, Yves Jaques, Stefano Anibaldi, Johannes Keizer. EFITA-2013
- «Migrating bibliographic datasets to the Semantic Web: The AGRIS case». Stefano Anibaldi, Yves Jaques, Fabrizio Celli, Armando Stellato, Johannes Keizer. Semantic Web journal

Thank you !