

INVESTIGATION OF AGRICULTURAL DATA FOR DATA BASE CREATION LIKE AGRIS FOR FARMERS WELFARE.

¹Mr.J.N.S.S Janardhana Naidu, ²Dr.E.N.Ganesh and ³Dr.N.Shanmugasundaram

¹Swarnandhra College of Engineering & Technology,Narsapur.

²Dean school of Engineering,VISTAS, Chennai

³Associate professor Department of Electrical and Electronics Engineering,
Vels Institute of Science, Technology and Advanced Studies,(VISTAS) , Chennai

Enganesh50@gmail.com

jnss.janardhana@gmail.com

shanmugam71.se@velsuniv.ac.in

Abstract:

Agriculture Productivity can be increased by improving farming system with stakeholders by proper planning and usage of ICT Materials. Agricultural information and data are stored in distributed systems and repositories (using different pieces of software and databases) that expose a variety of research outputs with various metadata formats. Often grey literature, journal articles and technical reports are lost as technologies are not employed to ensure that such resources are online, accessible and shared widely within the agricultural stakeholders. This paper reviews AGRIS (International System for Agricultural Science and Technology) as an international collaboration and partnership in collecting and using agricultural bibliographic data to enable researchers and policy makers retrieve related agricultural and scientific information and data. AGRIS database, therefore, uses bibliographic data as an aggregator of locating related content across information systems available on the Web through taking advantage of the Semantic Web and Linked Open Data technologies. AGRIS has a potential to make the researchers to discuss and share ideas for creating a social media layer.

Keywords: AGRIS, AIMS, AGROVOC, Agriculture, Linked Data, Agricultural Information, food productions, bibliographic datasets.

1. Introduction

Agricultural information is a key ingredient in boosting agricultural productivity; hence, institutions are investing in research to provide solutions to problems be-delving agriculture. Emerging technologies are set to improve agricultural research and examples have been adapted in geographic information systems, global positioning systems, remote sensing, agricultural statistical data, and early warning systems amongst many research based systems. Initiatives are on-going to embrace open data and the implementation of data infrastructures in the management of agricultural data and information both at technical and political level. These developments are ushering in a new era that will affect farmers' day-to-day operations

and increase their ability to access agricultural knowledge and enable also researchers to access agricultural research information¹.

Availability of software and databases with agricultural data has given enough opportunity to do research for increase in yield to usage of farming system. Today's researcher is lost in the following range of content types - photos, satellite images, statistical data, journal articles, published reports, books, graphs, maps and other geographic information, genomic information, and these coming in various formats and standards. It is from these sources that farmers, researchers, policy makers, academics, governments, and multinational institutions need to delve into during the research process (to both produce and consume these research outputs). The development in the semantic technologies provides a possibility to interlink different kinds of research outputs. In this sense, AGRIS (International System for Agricultural Science and Technology) is a mashup application that allows users to query the AGRIS content, interlinking all resources to external sources of information.

This paper reviews about exploring opportunity of AGRIS data to enable researchers and thereby by finding possible ways to find solutions for many agricultural problems. A brief overview of the literature on initiatives in open data in agriculture and their possible impact on agricultural productivity and food security will be offered to contextualise the contributions of AGRIS. The paper will further explain the current status of the AGRIS repository. The AGRIS repository seeks to use bibliographic data as an aggregator of locating not only the full-text of the article, but also related content across information systems available on the Web through taking advantage of the possibilities offered by the Semantic Web. The paper proposes a plan to establish a social AGRIS layer over bibliographic records, where authors and scientists can discuss various issues emanating from the literature and of interest to agricultural scientists.

2. Brief Literature Review

Literature is awash² with authorities detailing the concept of Linked Open Data and numerous interest groups and communities in LOD have been established over the years. The AIMS team of the Food and Agriculture Organization have published articles³ on the use of semantic Web standards to improve the open unrestricted availability and accessibility of agricultural information and data. Baker & Keizer (2010:178) reviewed how the data is integrated and access to information from a variety of sources. Subirats et al., (2012) and Subirats (2014) provide a detailed explanation of how the AIMS team implemented Linked Open Data in the agricultural domain. Van Vark (2013) explained how food security can be achieved using open data in agriculture; he explained that the first step is to make agricultural data available, readable (interoperable) and thereafter develop user applications that facilitate data querying, and re-use. If this is achieved, as stated by the World Bank (2013) open data combined with agricultural knowledge can support farmers, and this was aptly phrased, 'Imagine creating the ability for farmers to use data to understand what about best crops, or

what prices after harvest, or how best to solve any challenges to yield. A number of projects and organisations in the agricultural domain already exist to expose their datasets or either to implement open data projects. For instance, Keizer et al (2014) reviewed the latest developments of opening and linking wheat agricultural data within the a framework of Research initiative. Other projects include CAB's Plant wise Knowledgebase, Cassava base, a number of open data projects within CGIAR centres, and the Global Open Data for Agriculture and Nutrition (GODAN)⁴. This paper presents how AGRIS seeks to provide to the community of agricultural researchers about the accessing agriculture data, and use of it to improve knowledge, productivity and food security. The AGRIS framework and its search engine to its present state has been well explained over the years by Jaques, et al and others.

3. Background to AGRIS

AGRIS can mean three different things – AGRIS as a network, AGRIS as web portal and AGRIS as a database. Firstly, AGRIS network refers to the collaborative network of more than 150 institutions from 65 countries. The database AGRIS has 7.7 million structured records of data of agriculture. Finally AGRIS as a Web portal (<http://agris.fao.org/>) refers to the Web application that links the AGRIS knowledge to related Web resources using Linked Open Data methodologies. The purpose of AGRIS is to show the agriculture data and provide research information in agricultural domain. The Food and Agriculture Organization of the United Nations and its technical partner Agro-Know⁵ maintains the AGRIS database. AGRIS accepts content from publishers related to agriculture, forestry, animal husbandry, aquatic sciences, fisheries, and nutrition. Some of the records received have a link to the full text and AGRIS becomes an indexer to the respective repositories. AGRIS data is indexed in Google (since 2008), and this offers greater visibility of AGRIS content.

In the formative years, AGRIS grew from data received from AGRIS centers; however, AGRIS later began to ingest data from journal publishers. Celli, et al., (2013) explains how AGRIS can ingest metadata, either by pulling them through harvesting from clients (e.g. aggregators, institutional repositories) or by pushing data from clients (national libraries or journal publishers). AGRIS's stable and structured metadata, and the use of AGROVOC⁶ concept scheme for indexing have provided good sustainability over the technological changes for some years. Currently, AGRIS tries to accept all metadata formats and the following formats/standards are still highly recommendable: PubMed NLM XML, METS XML, DOAJ XML, MODS XML, AGRIS AP, MARCXML, OAI-DC, ENDNOTE XML, and SIMPLE_DC. After the importing of AGROVOC to a Simple Knowledge Organisation System (SKOS) concept scheme published as Linked Open Data, also the AGRIS content was provided as LOD to fully exploit the potentialities of the Semantic Web. Thus, the AGRIS repository was born from a collaborative effort of AGRIS Centres, later journal publishers, research institutions, and FAO. Before content linking (see Fig 1, below), AGRIS data comprise of three RDF sources which are the AGRIS records dataset, the AGRIS serials dataset and the AGROVOC dataset (Anibaldi et al., 2013). Over 80% of

AGRIS records are journal articles which shows about Subjects, Frequency of publication, Date, Publishers with ISSN. Some of the records come with a URL linking to the full text: it was estimated that 25% of the records have links to full-text articles. Today AGRIS used to retrieve better Agriculture information. In December 2013, the new AGRIS website⁷ (AGRIS 2.0) was launched with new search functionalities. AGRIS is heavily used with an average of 300,000 visits per month from more than 190 countries.

4.AGRIS path to RDF and Open AGRIS Mashup

AGRIS is an RDF-aware system, a mashup application that allows users to query the AGRIS content, interlink all resources to external sources of information. The linkages are made possible by the use of AGROVOC, a multilingual vocabulary containing more than 40,000 concepts available in more than 21 languages. AGROVOC itself is published as LOD and is part of the LOD cloud. As stated earlier, AGRIS data consumption is comprised of centralisation of data collected from data providers and then interlinking this data with other kinds of information related to the bibliographic records in AGRIS and the AGRIS domain. These kinds of information include maps, statistics, and country profile information.

In brief, the RDF process includes the translation in to XML database into RDF and the selection of the external datasets to link AGRIS. In the transition to RDF, a unique permanent URI is assigned to each AGRIS record. Other URIs included are for AGROVOC keywords and AGRIS Journals. To assure provenance, each AGRIS records have a unique identifier (ARN) with a predefined structure and contains implicit information about the data source and year of creation. The information about AGRIS data providers is exposed as RDF, so that each AGRIS data provider has its own de-referentiable URI. Two options are available to interlink to external databases using the AGRIS application to access external sparql endpoints, where metadata have been indexed with thesauri different from AGROVOC; and AGRIS can query external Web services using scientific names extracted from AGROVOC . This is necessary to ensure that the content automatically retrieved from databases available on the Web is related to an AGRIS record.

To exemplify the benefits of interlinking AGRIS content to external datasets, one needs to play around with the platform and observe the system pulling content from different sources. The International Food Policy Research Institute (IFPRI) case exemplifies this process: when the user types a search on the AGRIS portal, the system shows some results and the user can select one of them; if the selected record was indexed with at least one AGROVOC keyword referring to a country name, the systems can query the IFPRI sparql endpoint to retrieve some information about the country and display it on the screen (See Fig 2).

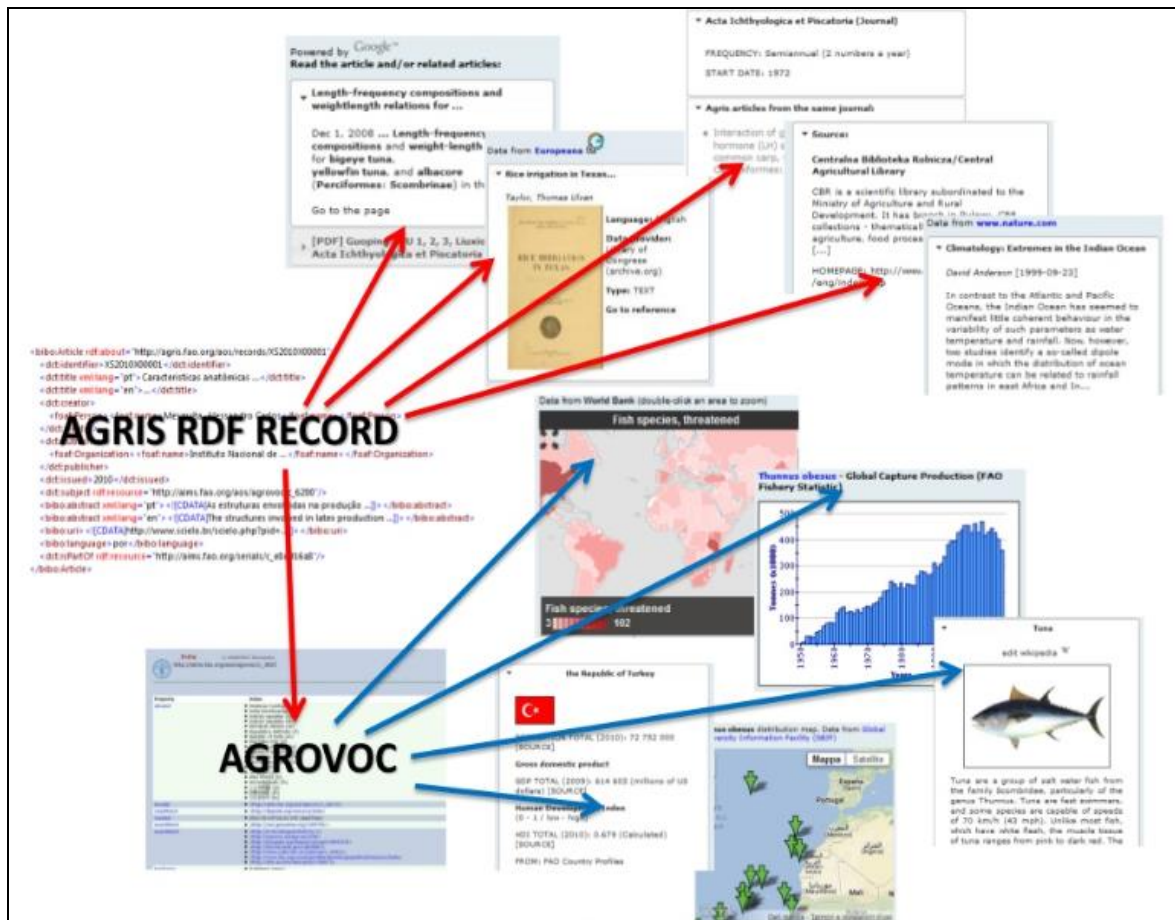


Fig 1: Content interlinking in AGRIS

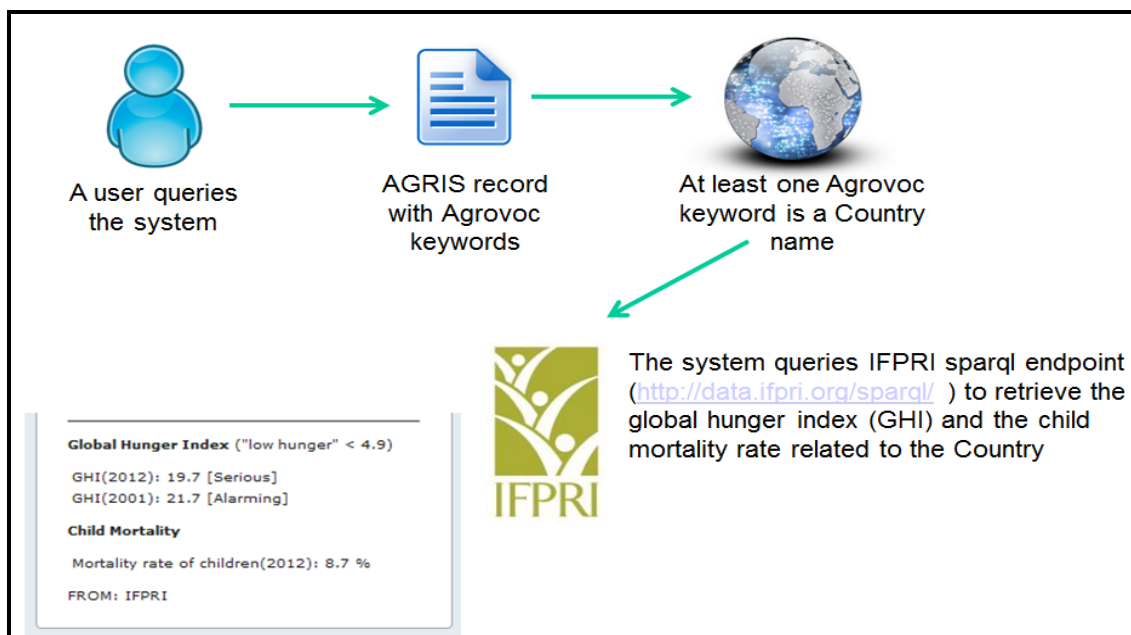


Fig 2: The IFPRI case

5.AGRIS user needs survey

In order to ensure AGRIS' usability and to evaluate new features of AGRIS 2.0 a survey of users was carried out. One of the new feature of AGRIS 2.0 is the ability to filter records by collapsing and adding AGROVOC keywords. The purpose of the AGRIS survey was to get user's feedback on the platform so as to further inform and adapt existing information resources according to user requirements. The 30 question survey⁸ was divided into the following parts, general appearance, search function, search results listing, search results display and general observations. The target audience was AGRIS users who include information management specialists, librarians and software developers, to researchers and students. The survey was distributed in English and Spanish. The overall result was a positive appreciation of AGRIS 2.0 features and the interface was considered to be clear and intuitive.

The recommendation included,

- improving the location of the advanced search function,
- reducing the number of boxes in the homepage,
- to add additional information like year and type of publication as well as information about the availability of full text was requested,
- to provide the original link of the source,
- to review the relevance of the information displayed from external resources

Work is underway to address these issues and improve the AGRIS 2.0 and assimilate the proposed and other new features. Many changes are foreseen, as the customization of the views. AGRIS users come from diverse backgrounds and disciplines, therefore AGRIS will provide users with an opportunity to register and choose their screen display. Some users may want to access the old fashion bibliographical view to make specific boolean queries to the system, other users may want to select the external sources of information to be displayed on the screen, others may want to change colours. By the end of 2015, AGRIS 3.0 will be published, with many new features and the possibility of social navigating the information in agriculture.

6.Conclusion

The foregoing case has shown that using semantic technologies have enabled the aggregation of various research outputs into a single interface. AGRIS has shown the prowess of Linked Open Data and related semantics in the provision of agricultural information and data. The visits on AGRIS 2.0 and the collaboration in growing its content and technologies testify the AGRIS key role in providing essential information for agricultural productivity. However, even if agricultural information and data is available and used at research and farming system level, there remains a risk of falling back to a situation where there is no information available at all. There is a need to develop applications that can consume these data and make it available at farming level or user level; such applications could enable querying sub-disciplinary subjects as crops, weather, planting and pests data and information. For example, in Sub-Saharan Africa and Asia mobile agriculture is revolutionising and impacting greatly

on farming information systems, which 'extensionists' and researchers could fully benefit from.

References

1. Anibaldi, Stefano and Jaques, Yves and Celli, Fabrizio and Stellato, Armando and Keizer Johannes. 2013. Migrating bibliographic datasets to the Semantic Web: The AGRIS case. Semantic Web. In *Semantic Web Journal*. <http://www.semantic-web-journal.net/system/files/swj463.pdf>
2. Baker, Thomas and Keizer Johannes. 2010. *Linked Data for Fighting Global Hunger: Experiences in setting standards for Agricultural Information Management*. <http://www.fao.org/docrep/014/am324e/am324e.pdf> [2014, May 6].
3. Celli, Fabrizio and Keizer Johannes. 2014. *AGRIS: an RDF-aware system in the agricultural domain*. A presentation in the LOD2014 Conference, Rome, Italy 20-21 February 2014.
4. Celli, Fabrizio and Jaques, Yves and Anibaldi, Stefano and Keizer, Johannes. 2013. Pushing, Pulling, Harvesting, Linking - Rethinking Bibliographic Workflows for the Semantic Web. In EFITA, Torino (Italy), 23-27 June 2013.
5. Jaques, Yves and Anibaldi, Stefano and Celli, Fabrizio and Subirats-Coll, Imma and Stellato, Armando and Keizer, Johannes. 2012. *Proof and Trust in the OpenAGRIS Implementation., 2012*. In International Conference on Dublin Core and Metadata Applications 2012, Kuching, Sarawak, Malaysia, 3-7 September 2012.
6. Kaboré, Esther Dzalé Yeumo, Madalli, Devika and Keizer, Johannes. 2014. Opening and Linking Agricultural Research Data. *D-Lib Magazine*. Vol. 20 (1&2): <http://www.dlib.org/dlib/january14/kabore/01kabore.html> [2014, March 30].
7. Subirats-Coll, Imma and Malapela, Thembani and Dister, Sarah and Zeng, Marcia and Goovaerts, Marc and Pesce, Valeria and Jaques, Yves and Anibaldi, Stefano and Keizer, Johannes. 2012. *Reorienting open repositories to the challenges of the Semantic Web: Experiences from FAO's contribution to the resource processing and discovery cycle in repositories in the agricultural domain*. 2012. In 6th Metadata and Semantics Research Conference (MTSR 2012), Cádiz, Spain, 28-30 November 2012.
8. WorldBank. 2013. *Open data+ agriculture can transform farmers response to crisis*. <http://www.worldbank.org/en/news/feature/2013/04/26/open-data-can-transform-farmers-response-to-crisis> [2014, March 13].