Paper for the IAALD XIIIth World Congress, 26-29 April 2010, Montpellier

The CIARD RING, an infrastructure for interoperability of agricultural research information services

Valeria PESCE 1, Ajit MARU 2, Johannes KEIZER 3

1. Global Forum on Agricultural Research (GFAR), Italy, valeria.pesce@fao.org
2. Global Forum on Agricultural Research (GFAR)
3. Food and Agriculture Organization of the United Nations (FAO)

Abstract

Creating integrated information services in agriculture giving access and adding value to information residing in distributed sources remains a major challenge.

In distributed architectures, value added services by definition interface several information sources / services. Therefore value added services cannot be built without an awareness of what others have done: which sources are available, how to tap into them, how to exploit their semantics.

The Coherence in Information for Agricultural Research for Development (CIARD) Routemap to Information Nodes and Gateways (RING) is a portal offering an interlinked registry of existing information services in agriculture.

The CIARD RING covers both information services and sources: in nowadays information architectures, the distinction between the two is very fluid. In the RING, the definition of "service" includes any form of providing information from one server instance (website, mail server, web services, XML archive...) to many clients (browsers, email clients, news readers, harvesters...)

The services registered in the RING are described in details and categorized according to criteria that are relevant to the use of the service and its interoperability. The RING categorizes and interlinks the featured services according to criteria such as: standards adopted, vocabulary used, technology used, protocols implemented, level of interoperability etc. In addition, it features detailed instructions on how the registered services can be "interoperated".

The vision is that the RING will become the common global technical platform for the community of agricultural information professionals for accessing, sharing and exchanging information through web services.

This paper describes how the RING provides an infrastructure for enhancing interoperability of information sources and thus paves the way towards better accessibility of information through value-added and better targeted services.

1. Background and rationale

While information needs for agricultural research, innovation and development become more and more complex, agriculture-related information is widely scattered among a huge quantity of sources that in most cases do not share and exchange data among themselves. On the whole, agricultural information on the web is not organized for easy access. The consequence is that those who need it do not find it or at least do not find all the relevant pieces of information that are potentially available to meet their specific needs.

First of all, users cannot be aware of all the available sources of information, as there is no comprehensive list or directory of existing information services in agricultural research for development (ARD). Secondly, even out of the sources of which they are aware, they have to search several, which are not and cannot be comprehensive and therefore provide partial answers to complex questions. Lastly, they cannot always get the information in a format or through a channel that is accessible to them, be it due to language barriers, literacy barriers, technological or infrastructural barriers..
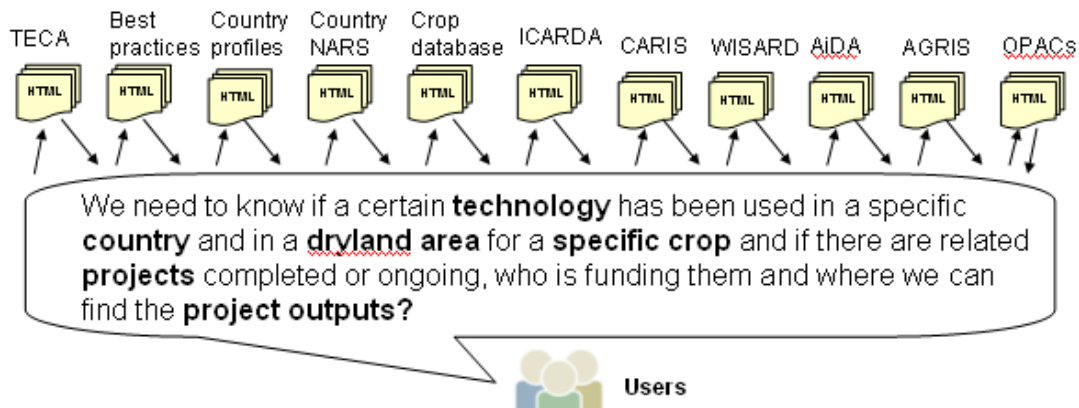
Fig. 1. Information needs and information services in recent years

The problem is not the quantity of existing information sources or that they are different but sometimes overlapping in coverage or that none of them is comprehensive: on the contrary, the fact that information sources are managed at the "lowest" level possible (from a geographic or an administrative point of view) in a distributed way allows for better and more sustainable maintenance. However, from the point of view of the consumers, this makes valuable information out of their reach, keeping it hidden and therefore only virtually available.

How can users reach the information they need when it is residing in widely scattered distributed sources? Only through highly user-targeted services that enable their specific audience to search, collate and integrate information from various sources acting as gateways to them. Such value-added services re-package the collated information and make it available through different browsing and search options, different formats, different channels according to the target users.

Creating such value-added information services in agriculture remains a major challenge, as they should by definition interface several information sources that use different forms of semantic organization, different languages and different levels of aggregation, and then integrate, semantically enhance and re-package the original information. Therefore value added services cannot be built without at least an awareness of what others have done: which sources are available, how to tap into them, how to exploit their semantics.
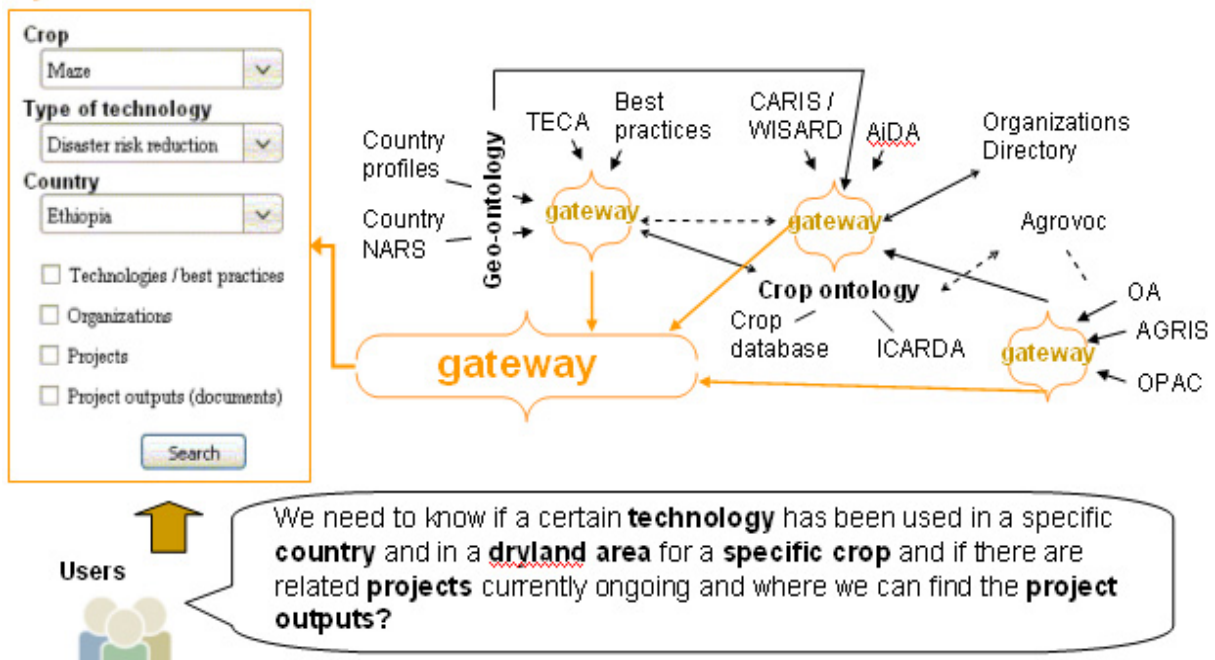

Fig. 2. Value-added services acting like gateways

In order for this to happen, information sources need to become more easily "discoverable" and to allow other services to semantically query, re-use and re-package their information; in other words they must become truly interoperable and allow for easy automatic retrieval of information, while work on mapping between vocabularies or advanced natural language processing must be done to improve the semantic accessibility of information.

The community of agricultural information specialists has been well aware of these issues for quite some years now, and the issue of information sharing and interoperability has been at the core of programmes carried on by international organizations in the years 2000-2008 (like the Global.RAIS and ICM4ARD programmes in GFAR, the AGRIS initiative and related Consultations on Agricultural Information Management in FAO, the "Fertile Ground" Study from IICD/DFID and the ICT-KM programme in the CGIAR) and, after the "International Information Systems for Agricultural Science and Technology" (IISAST) global consultations in 2005 and 2007, several major international organizations have joined forces in 2008 under the Coherence in Information for Agricultural Research for Development (CIARD) initiative

Within CIARD, the Global Forum on Agricultural Research (GFAR), building on the former idea of the GLOBAL.RAIS "Web Ring" that aimed at improving integration of national and regional agricultural information systems, proposed and implemented the "CIARD RING": the Routemap to Information Nodes and Gateways, a registry of information services in agriculture that is foreseen to act as a routemap towards better interoperability of information services.

This paper describes how the RING provides an infrastructure for enhancing interoperability of information sources and thus paves the way towards better accessibility of information through value-added and better targeted services.

2. The RING as a registry and as a routemap

The CIARD Routemap to Information Nodes and Gateways (R.I.N.G) is a web-based registry of existing information sources in ARD that supports CIARD in monitoring and improving the accessibility of information by:

- providing a map of where available sources of information (on a certain subject domain, in a certain format, for a certain audience) can be found and instructions on how they can be effectively searched;
- providing the examples of existing services that represent good practices on how interoperability can be implemented;
- making the level and modes of interoperability of information services more explicit;
- providing all the necessary information for building value-added integrated services that re-package information and make it accessible in different ways.

The RING is first of all a registry where detailed information about information sources is collected through a submission form and made searchable through a search engine.

It is important to define what an "information source" is and why in the context of the RING the terms information source and information service are used indifferently: in nowadays information architectures, the distinction between the two is very fluid. This is why the concept of "nodes and gateways" in the RING covers both information sources and information services, that is, both "static" files available in some structured format (like XML or RDF, but also data text files like .csv) and interactive services like search engines and web services. The reason is that both ways of making information available can be made interoperable and can contribute to improve the accessibility of information on the whole.

In a broader sense, the definition of "service" in this context includes any form of providing information from one server instance (website, mail server, web services, XML archive...) to many clients (browsers, email clients, news readers, parsers, harvesters...).

Examples are:

- services that provide RSS feeds
- services that provide XML exports of information based on agreed metadata sets (e.g. the AGRIS data providers)
- Open Archive Initiative (OAI) data providers
- OAI harvesters
- services that offer web services for accessing and re-using their information
- RDF stores
- SPARQL engines

The services registered in the RING are described in details and categorized according to criteria that are relevant to the use of the service and its interoperability. The RING categorizes and interlinks the featured services according to criteria such as: standards adopted, vocabulary used, technology used, protocols implemented, level of interoperability etc. In addition, the RING features detailed instructions on how the registered services can be "interoperated".

Fig. 3. The RING advanced search function shows the major indexing criteria

This is why the RING goes beyond being a registry and acts in fact as a "routemap":

- Thanks to the way services are indexed and made searchable, users are easily guided to find the type of information service they are looking for.
- Thanks to the way services are indexed and described, users can learn how to "exploit" them for building value-added integrated services.
- Thanks to the type of information gathered, a map of existing services can be drawn, highlighting the relationships between them and the flow of information among them

The potential impact of the RING is not so much in the collected information itself as in what can be built out of it. Providing structured information on the metadata sets, the formats, the protocols and the vocabularies used in each registered source will facilitate the building of applications like:
- services that offer a common browsing or searching interface to different sources;
- aggregating and harvesting services;
- integrated services providing relations between entities (organizations, projects, experts, documents) through semantic-web technologies;
- services that re-package information and make it available through different channels (text messaging, radio etc.);
- services that interface the different knowledge organization systems (KOS) used by different sources;
- applications providing value-added services like digests, bibliographies, best practices, surveys etc.

While the final objective is that of benefitting the consumers of agricultural information by increasing the quantity and improving the quality of value-added targeted services, the immediate users of the RING are mainly information professionals and web developers, who can on the one hand provide the relevant technical information about their services at registration time and on the other exploit the technical information shared by others in the RING in order to tap into the existing sources and feed their own systems.

Also the managers of information services may find the RING useful in different ways: they can identify useful information sources for the services they manage, they can find information services into which they can feed their contents for better dissemination and they can publicize the services they manage by just registering them.

Finally, once the RING is populated with a significant number of information services, also consumers of agricultural information may find it useful as a "bookmark" list to relevant information services in agriculture.

3. The present stage and the way forward
The RING is currently still in its launching phase. It has only been presented at two regional venues: at the CIARD regional consultations in Lima, Peru (30 October 2009) and in Hyderabad, India (10 December 2009). A few "pioneer" services have been registered and feedback is still being gathered on the registration procedure and the indexing criteria.
The RING is available at www.ciardring.net.

The first phase, just started, consists in building the registry.
In this phase, the RING is gathering information on which information sources / services are currently available and how to tap into them.
The metadata that are so collected allow to:
- categorize and interlink the registered services according to specific criteria: standards adopted, vocabulary used, technology used, protocols implemented, level of interoperability, target audience etc.;
- link the services to each other through provider / consumer relationships;

- feature detailed instructions on how the featured services can be "interoperated".

Fig. 4. A geographic map of the physical location of the registered services

The services featured in the RING are submitted directly by their managers and technical staff, which ensures ownership and reliability of the data. Any person who is responsible for an information service can register it. Each record describing a service must link to the record of an organization / institution registered in the system: these records can be created on the fly while registering the service or can be just referenced if they exist. A mandatory element is the email address of the institution: in order to ensure the correctness of attribution of the services to their owners, the organizations responsible for the service will be alerted upon submission and periodical checks will be run by the RING administrators.

Already in this first phase information professionals and web developers can exploit information in the RING to build advanced services that tap into the registered services.

The second phase will start when the technical information collected about the services is detailed and structured enough and when the number of registered services that have a good level of interoperability is significant enough: at this stage, some advanced services can be built semi-automatically directly on the RING website.
Examples are:
- an Open Archive Initiative (OAI) harvester harvesting all the registered OAI providers;
- an RDF viewer / navigator of the registered RDF stores;
- sample thematic RSS aggregators that harvest from the registered RSS feeds;
- sample consumers of web services.

Additionally, sample programming code on how to implement similar services will be provided on the RING website.

4. Conclusions

The effectiveness of the RING in achieving its objectives will be proportional to the number and quality of the services registered in the system. A high number of interoperable services correctly indexed in the system will allow to provide a comprehensive registry and a real infrastructure that can

be leveraged for building new services; the examples of advanced services featured in the system will provide guidance and incentives to the building of new ones; and the collected data will make it possible to generate interesting views on the flows of information between the services.

The core of the system and the real infrastructure for interoperability will be the registered web services: with the broad meaning of the term "web service" introduced by the advent of the so-called "RESTful" web services, these now include RSS feeds, RDF stores, SPARQL engines and OAI providers, and these are presently the best technologies for interoperability.

The vision is that the RING will become the common global technical platform for the community of agricultural information professionals for accessing, sharing and exchanging information through web services.

The RING is a CIARD project led by GFAR and its functionalities and the information therein registered are free and publicly available, thus constituting Global Public Goods that can be leveraged by any organization, person or information service.