Workshop

ICT Adoption in Agriculture and Agribusiness

The CIARD RING as a support tool for building integrated information systems

Ajit Maru; Valeria Pesce
Global Forum on Agricultural Research (GFAR)
Rome, Italy
e-mail: ajit.maru@fao.org

Abstract— **The need for integrated information systems in the agricultural domain is widely acknowledged.**

Relevant information on a specific topic, with a specific geographic scope, or aimed at a specific audience is rarely available from one single source. Many sources of information exist but none are comprehensive and all differ in coverage (often overlapping), semantic organization, being up-to-date, quantity and quality of the information they provide: on the part of the user who looks for information, selecting only one source is limiting and no cross-searches are possible because very few information systems share and exchange data among themselves.

In order for integrated information systems to provide information that is really relevant, comprehensive and tailored to specific audiences, information has to be gathered from as many relevant sources as possible and re-purposed for the specific needs of the prospected audiences. The real value addition of integrated information systems is the selection of quality sources and the re-purposing of contents for specific audiences and needs.

The main difficulty in building such integrated information systems is the little awareness of what information sources exist, how interoperable they are, how to tap into them and how to exploit their semantics. There is no comprehensive list or directory of agricultural information sources and technical details about these sources are often not documented and known only to the developers.

This is why the Coherence in Information for Agricultural Research for Development (CIARD) initiative launched in November 2009 a new service called the Routemap to Information Nodes and Gateways (RING), aiming at providing a map of existing information sources in agriculture. The RING (http://ring.ciard.net) is managed by one of the founding members of the CIARD initiative: the Global Forum on Agricultural Research (GFAR).

The RING is a platform where detailed information about information sources and services is collected through a submission form and made searchable through a search engine. The way the services are catalogued and categorized enhances the potential re-usability of the information they offer: the services are described in details and categorized according to criteria that are relevant to the use of the service and its interoperability. These criteria include content criteria such as thematic coverage, geographic coverage, content type, target audience; and technical criteria such as metadata sets adopted, vocabularies used, technologies used, protocols implemented.

Particular importance, especially in the context of CIARD, is given to the use of standards in the management of information: in order to build integrated information systems, sources should provide some level of technical and semantic interoperability, which can only be achieved through the use of standards. In order to provide comprehensive "authority" lists of existing information management standards, the RING harvests information from the registries available in the Agricultural Information Management Standards (AIMS) website: the registry of metadata sets and the registry of Knowledge Organization Systems (KOS). As regards other technical standards that are relevant to interoperability (protocol, notation), no comprehensive authority lists have been found, so the system provides free-tagging lists that users can extend.

Beside disseminating technical information on existing services, the RING will provide tutorials and examples of good practices on how to implement interoperability and to efficiently re-use and re-purpose information from interoperable sources.

The CIARD partners intend that the RING will become the principal global technical platform for enhancing the interoperability of information services and supporting the building of new integrated information systems.

Keywords-information systems; information services; interoperability; information management standards

I. BACKGROUND: WHY INTEGRATED INFORMATION SYSTEMS?

The definition of "integrated" systems in this paper includes both systems that integrate information from different sources and systems that integrate information of different nature (scientific, geographic, managerial, financial...) and/or available in different forms (bibliographic records, full text documents, multimedia, raw data...). And integrating in this context does not just mean "put together" or assemble, but rather "make sense of" by linking different pieces of information and contextualizing them, thus adding value to information and making it usable and applicable by specific audiences.

The need for integrated information systems in the agricultural domain is widely acknowledged. Integrated

systems are necessary because information needs in agriculture are complex and translate into complex questions that cannot find answers in mono-dimensional information systems. Farmers in particular have complex information needs that require to draw information from sources of different kinds and combine different data.

At a recent consultation on ICTs held in Hyderabad in December 2009 ("International Consultation on Agricultural Research for Development and Innovation: Addressing emerging challenges and exploiting opportunities through Information and Communication Technologies"), it was acknowledged that "The information that these farmers now need and will need in future is not only about how to grow a crop or improve productivity of a crop. It now relates to optimizing whole farm productivity and profit as also off-farm employment since many of these farmers and their households depend on off-farm activities for their livelihoods" [1]

The demand for ICTs by farmers is high and it is mostly about applications that deliver different kinds of inter-related pieces of information, contextualized and packaged in a suitable way for the prospected audience.

"In the recent regional consultations for the Global Conference on Agricultural Research for Development (GCARD 2010), farmer organizations such as of the Central Asia and South Caucuses and the Asia Pacific region made demands for application of cutting edge technologies for their development. In ICTs, they wanted applications for whole farm productivity and economic simulations, knowledge based decision support systems, the ability to access and use information for risk assessment and mitigation including that for climate change, the use of geographical information in planning and monitoring their agriculture and market related information not only of prices but of appropriate options for increasing productivity and profit and for ensuring food safety and appropriate information for consumers" [1]

In providing such information to different audiences, also the media adopted are of crucial importance. Integrated information systems should support the "re-packaging" of information for the "last mile", allowing information to be re-used and re-purposed through different media. "The Internet is no replacement for traditional information sources for farmers. Best results come from a mix of media such as radio, television, telephones, computer based information kiosks, computers, video and digital cameras and through the Internet, web and e-mail based services." [3]

The report of the workshop on ICTs held at the Science Forum 2009 [2] gives clear examples of how integrated systems can support rural people:

Delivery of various ICT-enabled services to rural people: such as market access; access to international export markets through ICT traceability systems; mobile financial services; mobile extension services

ICT can help dealing with uncertainty and complexity by integrating large amounts of data

Traditionally, the building of thematic or tailored information systems was mostly based on huge efforts of data entry, where data that were considered relevant to the specific subject area or to the specific audience were gathered, entered into the system and made searchable. Some of these efforts produced good results, creating big repositories of data, but they were not easily sustainable. In some cases data entry was distributed in order to minimize the effort at the central level, but quite often this resulted in a one-time effort and data became quickly obsolete.

As a result, many sources of agriculture-related information exist but none are comprehensive and all differ in coverage (often overlapping), semantic organization, being up-to-date, quantity and quality of the information they provide: on the part of the user who looks for information, selecting only one source is limiting and no cross-searches are possible because very few information systems share and exchange data among themselves.

Integrated information systems can overcome these difficulties: end users can reach the information they need, even when it resides 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 (see Fig. 1). 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.

.However, building such integrated information systems requires at least: a) the availability of technologies that allow systems to easily communicate or data to be interoperable; b) the availability of technologies and hardware that allows to store and quickly process large amounts of data; c) the appropriation of such technologies on the part of agricultural information managers; d) awareness of good existing information sources, assessment of their level and type of interoperability and if necessary negotiation of common protocols and vocabularies.

Regarding points a) and b) above, over the last years, new technologies and the availability of large amount of data and processing power have created favourable conditions for the development of new services and products..Last year (2009), in the workshop on ICTs held at the Science Forum 2009, the main trends identified included: "an increasing ability to collect, analyse and re-use massive, distributed collections of data. Calculations unthinkable a few years ago are now routine; completely new services and products are being developed on top of different data sets and services." [2]

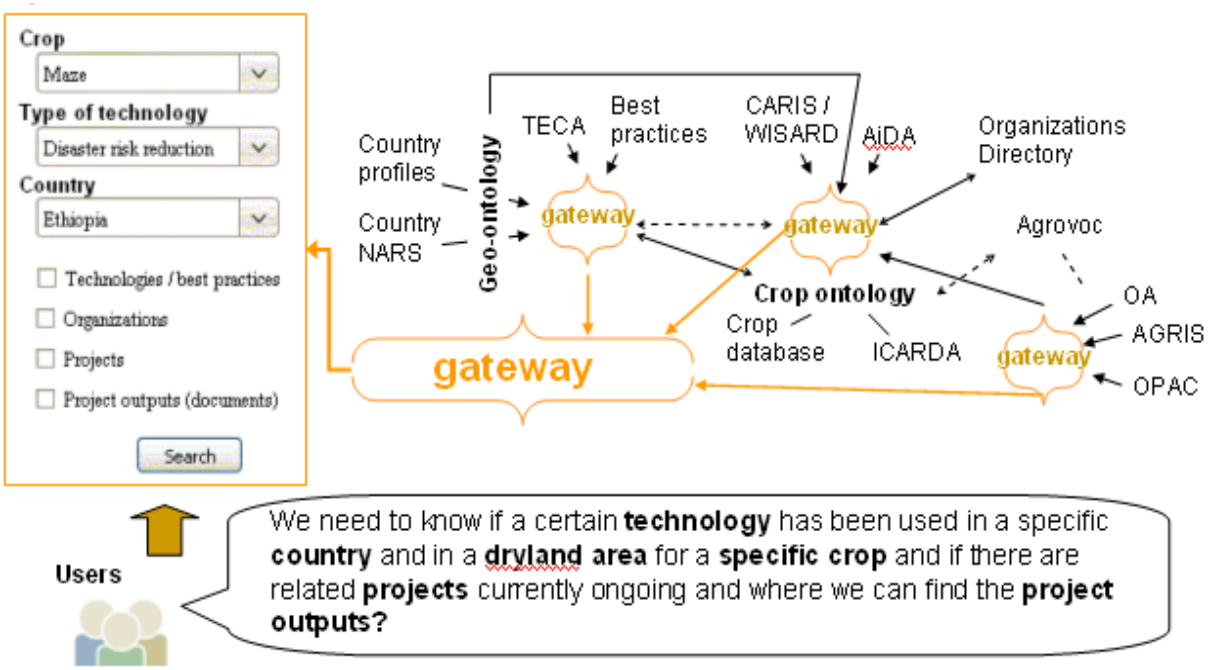

Figure 1. Example of integrated information system

Point c) has to do mostly with advocacy and capacity development, and we will not investigate on this here, because this paper focuses on technology and content related issues, although the RING, as it will appear in the next chapters, can also act as an advocacy instrument and as a training tool.

This paper explains how the CIARD RING ("Routemap to Information Nodes and Gateways") is a service that can support the implementation of new good integrated information systems by helping to overcome the difficulties arising from point d) above.

II. CURRENT SITUATION: WHY SO FEW GOOD INTEGRATED INFORMATION SYSTEMS?

Despite the availability of technologies and standards that allow to interoperate data and services over the Internet, there are very few true integrated infor4mation systems in agriculture.

The closest thing to an integrated information system in the last years have been news and events aggregators. Beside AgriFeeds, a service implemented by the CIARD partners, which harvests and re-exposes news, events and job opportunities from several registered feeds produced by agricultural information services, there is the NewsforDev service developed by CTA, where feeds are thematically categorized by the maintainers of the service, and there are several aggregators implemented by Institutions on their websites to harvest contents specifically related to the Institution's domain and scope (FARA has an aggregator of feeds from African regional, sub-regional and national sources, IAALD aggregates contents from ICM partner initiatives, EGFAR harvests contents from partner ARD Institutions and filtered news and events from AgriFeeds and so on). Indeed, at the moment RSS feeds are the easiest way to produce and consume interoperable information, although they are almost always limited to the exchange of news.

But there are also good examples of true integrated information systems that combine different data and pieces of information in order to provide a co prehensive service to their specific audiences.

A session at the latest IAALD Congress in Montpellier in April 2010 was dedicated to integrated information systems. What all systems had in common was the objective of organizing or re-purposing contents for specific audiences and needs; adding tailored value to raw data. Examples are monitoring systems, early warning systems, decision support systems.

Some of the systems presented at the session are still based on central or distributed data entry, but some (like CERISA and VELISA by INRA, VIGICULTURES by French Agricultural Institutes, LIMS by SADC, the Global Plant Health Centre by CABI) integrate data from distributed sources and one of the common problems these systems encountered was the identification and integration of sources: many presentations highlighted the importance of knowing which sources are available, where, and how they can be integrated (sources are difficult to find and "disconnected").

The main difficulties all arise from the very nature of integrated information systems: in order for integrated information systems to provide information that is really relevant, comprehensive and tailored to specific audiences, information has to be gathered from as many relevant sources as possible and re-purposed for the specific needs of the prospected audiences. The real value addition of integrated information systems is the selection of quality sources and the re-purposing of contents for specific audiences and needs. Building such systems is therefore difficult.

Without delving into political, institutional and financial conditions that may slow down the creation of integrated information systems at the decisional level, the main difficulties in implementing such systems are: a) the lack of resources and technical knowledge on the part of many information owners and the subsequent lack of really interoperable sources; b) the little awareness of what information sources exist, how interoperable they are, how to tap into them and how to exploit their semantics. There is no comprehensive list or directory of agricultural information sources and technical details about these sources are often not documented and known only to the developers.

This is why the Coherence in Information for Agricultural Research for Development (CIARD) initiative launched in November 2009 a new service called the Routemap to Information Nodes and Gateways (RING), aiming at providing a map of existing information sources in agriculture together with relevant information on how to interoperate them. The RING (http://ring.ciard.net) is managed by one of the founding members of the CIARD initiative: the Global Forum on Agricultural Research (GFAR).

III. THE RING AS A SUPPORT TOOL FOR BUILDING INTEGRATED INFORMATION SYSTEMS

The RING is a platform where detailed information about agricultural information sources and services is collected through a submission form and made searchable through a search engine. The way the services are catalogued and categorized enhances the potential re-usability of the information they offer: the services are described in details and categorized according to criteria that are relevant to the use of the service and its interoperability. These criteria include content criteria such as thematic coverage, geographic coverage, content type, target audience; and technical criteria such as metadata sets adopted, vocabularies used, technologies used, protocols implemented.

What the RING does is:

- provide 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;
- make the level and modes of interoperability of the registered information services more explicit;
- provide the examples of existing services that represent good practices on how interoperability can be implemented;
- provide instructions for building value-added integrated services that re-package information and make it accessible in different ways.

Figure 2. A geographic map of the locations where the registered services are maintained, as of August 2010

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 (from news and events services to blogs and forums); 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; structured data in any format (Excel files, microformats)..

Particular importance, especially in the context of CIARD, is given to the use of standards in the management of information: in order to build integrated information systems, sources should provide some level of technical and semantic interoperability, which can only be achieved through the use of standards. In order to provide comprehensive "authority" lists of existing information management standards, the RING harvests information from the registries available in the Agricultural Information Management Standards (AIMS) website: the registry of metadata sets and the registry of Knowledge Organization Systems (KOS). As regards other technical standards that are relevant to interoperability (protocol, notation), no comprehensive authority lists have been found, so the system provides free-tagging lists that users can extend.

Beside disseminating technical information on existing services, the RING will provide tutorials and examples of good practices on how to implement interoperability and to efficiently re-use and re-purpose information from interoperable sources. This will be done in different ways: 1) the information managers who register their services are asked to provide examples and instructions on how their services can be interoperated; 2) the RING managers and technical experts from the CIARD community will write tutorials and technical documents on the technologies involved in interoperability, from the point of view both of data providers and data consumers; 3) links will be provided to the CIARD "Pathways", sort documents that provide best practices on how to make knowledge accessible to all.

These features, essential to increase the availability of both interoperable sources and integrated systems, will make the RING also an instrument for capacity building.

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.

The RING was formally launched at the CIARD regional consultation in Lima, Peru (30 October 2009): so far, around 100 "pioneer" services have been registered and feedback is still being gathered on the registration procedure and the indexing criteria.

The RING is available at http://ring.ciard.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 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.

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 as examples for others to replicate.

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.

IV. CONCLUSIONS

Last year, in the workshop on ICTs held at the Science Forum 2009, some promising opportunities that ICTs can offer agricultural science for development were identified and they included: "Enhanced information repositories, as well as tools for information exchange and dissemination. Science results and outputs become inter-operable, open and re-usable" [2], while the issues to be addressed included: "Coherence and interoperability of data/information & quality control" and "Discovery of relevant information and putting it into use" [2] .

At the same meeting, the following points were identified among the main high-return investments:

- open access to research documents;
- develop and agree standards to exchange data and information".

The RING project places itself in the process that facilitates the exploitation of these opportunities and the achievements and implementations around such high-return investments.

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 RING will essentially contribute to the achievement of the CIARD objective ("to make agricultural research information publicly available and accessible to all ") and to the implementation of the actions identified in the Hyderabad meeting under the heading "To optimize the efficiencies in and to reduce the cost of content transformation and generation".

REFERENCES

[1] Global Forum on Agricultural Research et al.., International Consultation on Agricultural Research for Development and Innovation: Addressing emerging challenges and exploiting opportunities through Information and Communication Technologies. ICRISAT, Patancheru, Andhra Pradesh, India. 7-11 Dec. 2009.

[2] Peter Ballantyne,.ICTs transforming agricultural science, research and technology generation. Summary of the ICT Workshop at the Science Forum1. August 2009.

[3] Peter Ballantyne, Investing in Agricultural Development. Information, Knowledge and Communication the 'Fertilizers' of Future Research Harvests.