AgriVIVO for enabling global networking for agriculture
Concept Note

Jon Corson-Rikert (Cornell University)
John Fereira (Cornell University)
Valeria Pesce (Global Forum on Agricultural Research)
Johannes Keizer (Food and Agriculture Organization of the United Nations)
Ajit Maru (Global Forum on Agricultural Research)

Part 1. Goals, background and rationale

What we want to achieve

The primary goal of AgriVIVO is to facilitate better networking of individual researchers and the organizations they belong to for better collaborations and less duplication of efforts.

Networking happens through many channels now – through institutional directories, Web search by person or organization name, discovery via print or online publications, professional channels such as LinkedIn, and direct interaction at meetings or conferences. Face-to-face meetings are of course optimal for establishing a connection but are often not practical due to distance, time, and financial constraints.

Although incentives for improved collaboration are increasing, as detailed below, we rely heavily on serendipity to discover and build our networks. The sheer amount of information being posted on the Web and its lack of systematic organization risks blinding us to most useful opportunities for connection – one recent study indicates that new links posted on social networking sites receive half their total lifetime clicks in just a few hours[1].

This project will demonstrate the potential for AgriVIVO to leverage existing sources of information about agricultural researchers by discipline, organization, or topic –

- by integrating data using a consistent ontology framework
- by bridging gaps between selected established local, regional, and global membership directories, databases, and websites
- by supporting bi-directional pathways within AgriVIVO and between AgriVIVO and its initial data sources

We don't need more information – we need more context. Most of the linkages we seek exist implicitly within the information sources we use every day, but we rarely leverage the common data elements and connections to make them visible and persistent (figure 1).

[1] http://www.readwriteweb.com/enterprise/2011/09/measuring-the-lifespan-of-shar.php

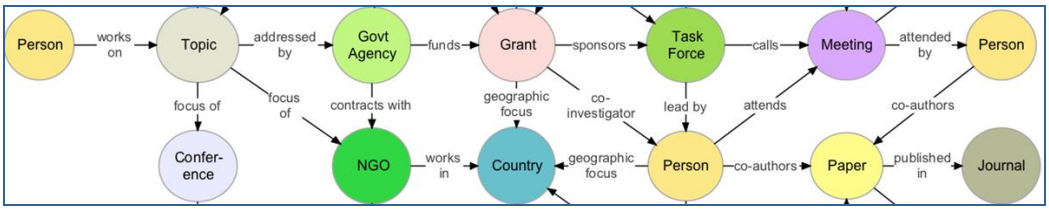


Fig. 1. Beyond serendipity: Building a network by representing the nature of relationships and using them to connect people, organizations, topics, activities, and outcomes

The AgriVIVO pilot will demonstrate a strategy for bridging across disjoint communities and directories without requiring centralized control or updating. Capturing information about researchers and their activities close to the source retains local review and will continue to meet many local needs. What AgriVIVO will provide is the ability to search several databases and directories and to share people profiles, affiliations, competencies, and publications across communities.

Background

The need for more collaboration between agricultural researchers, practitioners, information managers across Institutional and geographical boundaries has been highlighted in many conferences, publications and projects in recent years. The more complex the issues faced by agricultural development become, including reactions to natural disaster and pests, the more timely the response of research and agricultural public information has to be. And as agricultural research becomes more knowledge-intensive, the ability to quickly collaboration and share results with other relevant actors becomes more essential.

The GCARD RoadMap[2] states (Section 6.4) that "Regional capacity development partnership is needed to: i) generate economies of scale in collaborative AR4D, ii) foster inter-country and inter-regional cooperation, learning and exchange of experiences and develop national capacities and iii) promote more effective regional and sub-regional collaborative research and networking to make better use of available resources and enhance capacity development in the smaller and weaker national systems."[3] More specifically, "The convening role of Regional and Global Fora and their functions in the sharing of knowledge and innovation between regions is crucial in facilitating capacity strengthening and networking of skills where required to support national development processes and hasten development through inter-regional learning."[4]. And the need for more effective collaboration is clearly stated in the Conclusions of the GCARD RoadMap: "As the Roadmap challenges institutional divides and transcends the conventional boundaries of disciplines, effective collaboration must be forged among diverse stakeholders to provide technological, socio-economic and policy solutions."

---

[2] The main outcome of the CGARD 2010 Conference held in Montpellier, France, in March 2010, outlining a process of reform of agricultural research for development.

[3] GFAR and CGIAR, 2011. The GCARD Road Map. Transforming Agricultural Research for Development (AR4D) Systems for Global Impact. Available on line at (http://www.fao.org/docs/eims/upload/294891/GCARD%20Road%20Map.pdf

[4] Ibidem.

Also a recent report by the CGIAR recognizes the key importance of collaboration and partnerships: "As advanced science breaks down old barriers to crop improvement, new information and communications technologies are widening the scope for collaboration in research on crops and natural resources by creating new spaces for innovation that are partly virtual. In fact, the whole concept of technological innovation has evolved in recent years into a more collaborative enterprise involving more diverse actors. This has spurred researchers to assume new roles in development and enter into broader partnerships that involve not just the public sector but private companies and civil society as well, including local producer associations."[5]

To ensure that collaboration is as effective as it can be, it is essential to identify the right collaborators across Institutional and geographical boundaries and to make sure that no potential collaboration is overlooked or ignored due to lack of connections or partial coverage of our networks. Besides, in emergencies, it is essential to be able to quickly identify the right people or the key Institution or the relevant project that can make the difference in providing the necessary knowledge.

Rationale

Better networking is needed for:

- Fostering collaboration and synergy through greater awareness
- Reducing duplication of research
- Determining strategic trends based on strengths and weaknesses of the network
- Identifying missing expertise
- Improving responsiveness to calls for proposals
- Facilitating team formation
- Providing a marketing tool for research

Today it is not easy for a researcher, a research manager or a practitioner to identify / discover:

- *his/her potential best collaborators all over the world for a project*
- *a person with an answer to his/her question*
- *an organization running a project on a specific area of research*
- *an organization funding projects in a specific area of research*
- *all the publications written by a potential collaborator*
- *numbers or geographic distribution of available competencies or ongoing projects*

Part 2. Why VIVO

What is VIVO

2003 - 2009: VIVO at Cornell University

VIVO started in the library as a collaboration website for the Genomics Initiative at Cornell and grew into

[5] CGIAR, 2011. The CGIAR at 40 and Beyond. Available on line at http://www.cgiar.org/pdf/cgiar@40_final_LOWRES.pdf

a research-focused discovery tool that enables collaboration among scientists across all disciplines at Cornell University. VIVO supports browsing or searching information on research & expertise across Cornell: people, departments, courses, grants, and publications.

2009 - 2011: VIVO across the US

A $12.2 million, two-year grant from the National Institutes of Health's National Center for Research Resources enables the creation of a network of VIVO institutions, initially at seven sites[6]. VIVO supports interoperability through linked open data,[7] and other prominent research networking systems adopt the VIVO ontology to facilitate searching and linking across authoritative, institutionally-managed data[8]. VIVO becomes an open source community with contributions from many new adopters.

2010: VIVO at USDA

The U.S. Department of Agriculture is the first federal organization to commit to using VIVO, a web application designed to enable better national networking between scientists from different disciplines and locations[9], to link its 45,000 research scientists with each other, researchers in other federal agencies, and academia.

A global AgriVIVO?

- "*Addressing the critically important agricultural issues facing the world today requires an interdisciplinary approach between scientists across the United States and* ***around the world***" said Agriculture Secretary Tom Vilsack.[10]

- "*VIVO will be an excellent way to* ***make research more effective*** *and help researchers forge important new collaborations that can lead to the kind of ground breaking results that we need to help solve the problems we face today.*"[11]

AgriVIVO

From VIVO to AgriVIVO

- **VIVO at Cornell**: greater *interaction*, with the goal of catalyzing networks of **campus-wide** scholarship, research and educational activities*[12]
- **VIVO across the U.S.**: greater *interaction*, with the goal of catalyzing networks of **national** scholarship, research and educational activities in health science
- **VIVO at USDA**: better **national** networking among scientists in agriculture

---

[6] http://vivoweb.org
[7] http://linkeddata.org
[8] as demonstrated at http://vivosearch.org
[9] http://www.usda.gov/wps/portal/usda/usdahome?contentidonly=true&contentid=2010/10/0507.xml
[10] http://www.usda.gov/wps/portal/usda/usdahome?contentidonly=true&contentid=2010/10/0507.xml
[11] http://www.usda.gov/wps/portal/usda/usdahome?contentidonly=true&contentid=2010/10/0507.xml
[12] http://www.dlib.org/dlib/july07/devare/07devare.html

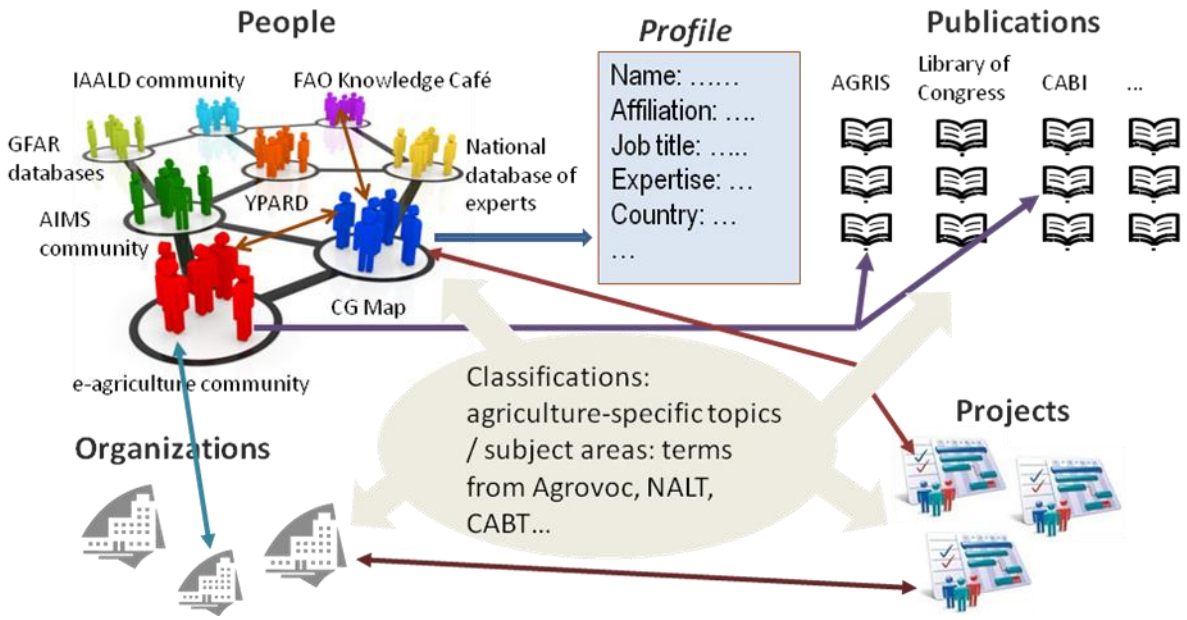
- → **AgriVIVO**: greater *interaction*, with the goal of catalyzing networks of **world**wide agricultural research, educational and policy activities

What is AgriVIVO.

AgriVIVO will be a global cross-institutional version of VIVO to help researchers, policy makers, practitioners, extensionists, information managers, students in agriculture discover common interests and make connections. AgriVIVO will enable global networking for agriculture.
The goal is to foster alliances, making agricultural research and innovation move faster.

AgriVIVO **will not replace** any existing community or database, it will work as **a common registry to interlink the data** managed in the existing communities and databases. Communities and databases will share data **through** AgriVIVO.

AgriVIVO will only store **pointers** to and **relations** between the Agricultural Research Management data managed anywhere:



How VIVO is different

VIVO vs. Google

- VIVO only searches **relevant** communities / directories
- Information in VIVO is automatically gathered but can be **curated** by the community members:
  - Editing one's profile
  - Claiming publications, associating / dis-associating oneself with/from projects

VIVO vs. Linkedin
(or other similar social channels)

- VIVO searches across communities / directories, Linkedin only uses its internal database
- People profiles in VIVO are shared across communities
- In VIVO, subject areas, research topics and categorization criteria in general are customized for the community that is using it, typically by selecting from a common standard such as Agrovoc
- Data in VIVO can be easily re-used by other applications to build customized search engines

VIVO vs. EuroCRIS
(or similar research information systems and career databases)

- EuroCRIS focuses on European research data
- EuroCRIS is based on a GRID architecture, more complex and less "open": data cannot be automatically imported from other communities / databases; VIVO can import data from other systems and can expose data for other applications easily because it uses standard semantic technologies
- EuroCRIS CVs are available in the EuroCRIS database; people profiles in VIVO are shared across communities
- In VIVO, subject areas, research topics and categorization criteria in general are customized for the community that is using it
- Data in VIVO can be easily re-used by other applications to build customized search engines

N. B. *VIVO and EuroCRIS agreed in November 2011 to map their data models to allow data exchange and common searches*

VIVO vs. WISARD / CARIS / InfosysPlus
(or similar agricultural research management information systems)

Many databases of institutions / projects / experts in agriculture already exist, but **they are managed in silos**

- Each uses its own format / structure / classifications
- Each stores data in its own database with limited or no import / export functionalities → no data exchange and no common search possible

Part 3. AgriVIVO applications and scenarios

1. Global search engine

AgriVIVO will initially integrate data from several large biobliographic and eAgriculture databases, including information on people, organizations, publications, events, projects, and expertise. A search index over this integrated VIVO and other designated sources (e.g., USDA VIVO or VIVO at land grant

universities in the U.S.) patterned on http://vivosearch.org will demonstrate the potential of a global search.

2. Specialized search engines in any website

VIVO's search functionalities can be integrated in other websites through remote calls. In this way, specialized and targeted search engines can give access to and offer highly customized "views" of the data coming from AgriVIVO.
(AgriVIVO data can be retrieved remotely, through pre-defined queries designed to support directories or researcher profiles or more open-ended, user-driven search returning a ranked list of results faceted by type).

3. Measuring impact

AgriVIVO data can be semantically aggregated and be visualized as maps, charts and different kinds of statistics. An example of such usage is the "research impact" portal at Cornell: http://impact.cals.cornell.edu/.

4. Many communities, one profile

AgriVIVO will demonstrate the reuse and enrichment of profile data from existing agricultural sites that manage people profiles as a model for maintaining one profile that can provide consistent information across multiple websites.
In order to allow for the management of profiles through existing websites – besides the direct editing in the VIVO environment - the VIVO development team is exploring ways of propagating editing changes, not only from source systems to VIVO, but from VIVO back to the original source system when new or supplemental information becomes available.

5. Interactive data curation

AgriVIVO can also be used as a community platform for interactive data curation. Users can add/remove "relations" in which they are part of the relation (person A "is author of" publication B, person A "participates in" project C): in this way they can "claim / disclaim" publications or associate / remove themselves with / from a project.
Current research activities in the VIVO project include the management of overlapping or duplicate data from multiple sources through "sameAs" assertions and the ability to assert that an author is not the author of a publication when ambiguous author names may risk repeat assertions of incorrect data.
These VIVO functionalities may in the future allow for the creation of "authority data" for publications' authors, something that is missing and very difficult to achieve without a semantic and data-curation infrastructure.

More information

- VIVO at Cornell: http://vivo.cornell.edu/
- VIVO network project: http://vivoweb.org/
- On VIVO: http:/www.dlib.org/dlib/july07/devare/07devare.html

- VIVO going national: http://www.news.cornell.edu/stories/Oct09/VIVOweb.ws.html
- VIVO at USDA: http://www.usda.gov/wps/portal/usda/usdahome?contentidonly=true&contentid=2010/10/0507.xml
- GCARD RoadMap: http://www.fao.org/docs/eims/upload/294891/GCARD%20Road%20Map.pdf