
A Framework for Data and Information sharing for Agricultural Research for Development: A perspective for its development

Ajit Maru
Global Forum on Agricultural Research Secretariat

In developing a framework for data and information sharing for agricultural research for development (ARD), the first question we need to ask is to why we need a framework, and prior to it, why do we need to share data and information for agricultural research for development?

The answer lies in the fact that current and emerging challenges in agricultural development are global in their scope. Every community and country needs increased agricultural productivity from their farms, and just returns for their farmers and producers. They need to reduce hunger and malnutrition, alleviate extreme poverty and use natural resources such as land and water and energy more efficiently and sustainably. Climate change affects their agriculture now and will affect more in the future and they need to adapt their agriculture rapidly to it as also contribute, as responsible world citizens, to mitigating and reducing climate change and its effects. Desertification, trans-boundary diseases, loss of agrobiodiversity and equitable participating in global agricultural markets are also universal problems. And all need research and innovation to be solved.

Data and information sharing for agricultural research for development enable, among many benefits:

- New information and knowledge to be generated
- Localize globally available information and knowledge and enable it to be used more effectively
- Increase efficiency and effectiveness of research and its outputs and innovation in time, cost, quality and human effort
- Reduce reinvention and repetition of research efforts
- Allow greater inclusiveness and participation in research and innovation
- Bring cross-disciplinary and specialized skills to agricultural research

- Reduce “market failure” or inability to use research outputs effectively and/or efficiently
- Create new research directions and avenues
- Bring greater equity in using agricultural knowledge across and among communities

The second question that needs to be asked is what do we do to improve data and information sharing in ARD? The answer, in part, lies in the potential of new information and communications technologies (ICTs). These ICTs have enabled new avenues to generating, processing, sharing and exchanging data and information. The processing power of computing devices has increased dramatically, doubling almost every two years while their size and energy consumption has reduced enabling more mobile use. Similarly, connectivity of these devices has increased not only in speed but in spread and is becoming ubiquitous facilitating further mobility in the use of these devices even in remote, rural areas. “Cloud” computing is enabling huge amounts of data to be stored, accessed and used remotely. “Crowd” computing enables communities to collectively generate, process and use information, many a times through mobile devices, to solve a community’s problems. The Web 2.0 technologies allow an interactive web and in the near future a “semantic” web, which provides information as a human mind understands “information” and further down in the near future, a “symbiotic” web, where the web symbiotically or through mutual benefits provides information to all who are dependent on it rather than only to those who can afford it.

To move faster in harnessing the potential of new ICTs in ARD, and it is important here, that information (and knowledge) is available and accessible to those who are dependent on it and that they can use information effectively, we need to start looking at the information related to agriculture available in the public domain and making it more accessible. And why the focus on the public domain? Because, a majority of information generators, assemblers, amalgamators and disseminators of agricultural information **for development** are from the public sector and funded through public funds. They are obliged to put their information in the public domain transparently.

To make the information more easily **accessible** as also **interoperatable** so as to make its effective use in different situations possible and easier, we need to work at the level of **information objects or data**. And therefore we need to create **open data repositories** and couple them to concepts of **“linked open data”** or data objects that can be **hyperlinked** across common attributes. Of course, to develop these data repositories that can be located and used globally we will need **standards, norms and regulations** to **govern** effective, efficient, transparent

and equitable flows and uses. We will need **standard descriptors**, mutually agreed **taxonomies** and **vocabularies** based on **common ontologies** to describe and organize the storage and access to the data. To reduce costs and time in assembling data to meet the information needs for meeting the most critical challenges to agricultural development, we may need to define **Core Data Sets** that Institutions and National Systems of agricultural research and Innovation universally may contribute and work in the spirit of global shared responsibility, collaboration and partnership. And for this we will need **policies** and **strategies** at Institute, Organization and Country level. We will need **governance** structures, **legislation** and **International treaties** to share data and information openly as also protect interests of all parties in information management and use. We may need to create **Trust organizations** that will manage the data and metadata that is shared openly at national and global levels. We will need to look at **investments**, financial and in human skills that will be needed to manage the entire information infrastructure in global ARD. And of course for all these ideas to take off, we will need **advocacy** and **promotion** and new **capacities** to advocate and manage open data repositories sustainably and all the issues related to them.

The third question that we should ask is, how do we do all that is needed to be done for data and information sharing for ARD? In my opinion, we need to start, through inclusive dialogue, by creating a framework for data and information sharing for ARD for all involved to follow and contribute to systematically. A lot is already being done and we need not all do the same individually and independently. We simply do not have the resources for it. For example, the concept of open repositories is now well developed and there is constant development of the concept of "linked open data" with basic standards such as XML, RSS, RDF etc already in place. The taxonomies and vocabularies, such as AGROVOC, are being rapidly evolving. We need to look at Core Data Sets though some such as for weather and climate, plant germplasm and spatial data already exist and some for example in agronomy are in the making. The issue is who defines them? We need to look at Trust Organizations. The CIARD.RING led by GFAR is one of them. We need to look at implications of open data access in ARD at organizational and national levels and this can be done through GFAR and FAO. We need to advocate for increased investment in information infrastructure in global ARD and this can be done at GCARD 2012. And much more already done, being done and need to be done can be mapped once we have a framework.

Contact:

Ajit Maru

Senior Knowledge Officer
Global Forum on Agricultural Research (GFAR) Secretariat
OEKD, B 648,OEKD, FAO,
Viale delle Terme di Caracalla, 00153, Rome, Italy
Email: Ajit.Maru@fao.org