CHAPTER 6
KNOWLEDGE AND DATA

Achieving Food and Nutrition Security through Open Access Data

INDIRA YERRAMAREDDY AND SURESH CHANDRA BABU

Indira Yerramareddy is the manager for Knowledge Management and Web, Communications and Public Affairs Division, and **Suresh Chandra Babu** is a senior research fellow and head of the Capacity Strengthening Program, International Food Policy Research Institute, Washington, DC, USA.

KEY FINDINGS

- Open data can improve the performance of food systems and help achieve global food and nutrition security.
- Accessible data are critical for decision making, from the farm to the retail level of food systems.
- Open data increase both the visibility and utility of research, allowing researchers to create more knowledge products and support decision making.
- Open data allow governments to make evidence-based policy decisions and push governments toward increased accountability.
- Data quality and ease of use are essential for putting data to use, but datasets are often too large or complex to be easily handled.
- Inequality in access to knowledge is increasing. Data policies, commitments, and investments can improve access to and use of knowledge, but current commitment and action on open data are uneven.

KEY RECOMMENDATIONS

- Democratize data access and improve livelihoods by putting data tools, such as mobile-phone apps, into farmers' hands.
- Increase the efficiency of knowledge transfers to prevent loss of information and ensure uptake in the field.
- Make government "big data" public to drive high-quality analysis of food systems and better policy and decision making.
- Build open data initiatives, including to reduce inequality and address issues of data quality, use, storage, and dissemination.
- Increase data quality and ease of use through better data collection, new tools, working groups, capacity building, and improvements in big data platforms.
- Empower citizen stakeholders to demand open data through capacity building and access to data tools.

Despite global reductions in hunger, malnutrition, and poverty, food and nutrition insecurity remain a global challenge. Reducing hunger and malnutrition requires evidence-based decision making, which in turn depends upon access to knowledge and data. Support is growing for "open data"—data that can be freely used, shared, and built on by anyone, anywhere, for any purpose.[1] Widespread open access to usable knowledge and data for farmers, businesses, and governments can improve the performance of food systems. However, under increasing isolationism and nationalism, there are worrisome restrictions on the free flow of knowledge and data. Ensuring open access will be critical to supporting better food systems and strengthening food and nutrition security.

BENEFITS OF OPEN ACCESS

Data are the foundation of decision making in all subsectors of our food systems, for example, seed or fertilizers or finance. Producers depend on knowledge of inputs—including land, water, other natural resources, fertilizers, seeds, and credit—and on cropping system data that inform what crops to grow, when, where, and how. Data are equally important for other segments of the value chain, including data on quality control, markets, prices, and consumption. Open flow of knowledge and information is therefore critical from the farmer to the retailer. Likewise, knowledge is critical for government policy making and program design. Policy makers' effectiveness depends on the evidence base they can draw on. Enabling open access to knowledge and data lowers the costs associated with gathering and accessing data and drives better decision making and improved products and services.[2]

At the farm scale, field-level knowledge and information made available to farmers have the potential to improve agricultural productivity. Easily accessible information about breeding, soil management, improved irrigation, plant protection, and postharvest management can greatly enhance farmers', especially smallholder farmers', operations and improve end results.[3] As farmers face climate change—associated with variability in weather, including erratic rainfall—improved access to data and knowledge about weather patterns can help them increase their resilience. Better flow of data

This chapter benefited from research and writing assistance from Namita Paul and Sivan Yosef.

can also address the information asymmetries within markets that can put farmers' operations, especially smallholder farmers, at a disadvantage. When price data are openly available, for example, farmers can obtain more affordable inputs and sell their produce for the best farmgate price.

Open data and knowledge have been shown to increase both the visibility and utility of research.[4] Researchers can use data to create more knowledge products, conduct cross-sectoral studies, and support decision making. Government leaders benefit from having more evidence available to inform policy making and can avoid decision making based on assumptions or "gut feelings." During natural and manmade emergencies, data availability can significantly improve the timeliness and appropriateness of humanitarian responses.

In Malawi, for example, droughts resulted in a food security emergency in 2015–2016. A vulnerability assessment was conducted in May 2016 to understand the extent of the crisis and to collect information on the performance of food markets across the country. The assessment showed that 6.5 million people (later updated to 6.7 million) were affected by drought in Malawi. Food security and nutrition monitoring systems that collect, process, analyze, and openly share data can improve food and nutrition security in such situations. Currently, much of the data collected for emergency purposes are not shared through open data systems, delaying interventions and increasing the vulnerability of affected populations.[5]

Open access policies help push governments toward increased transparency and accountability. For example, India's Right to Information Act of 2005 requires that citizens have access to information regarding constitutional authorities, including the executive, legislature, and judiciary. Information obtained through this act is claimed to have revealed that shopkeepers and food grain officers had siphoned off 87 percent of wheat and 94 percent of rice from a food program meant for the poor—findings that prompted significant reforms.[6]

CHALLENGES IN ACCESSING AND USING DATA

Many governments and international organizations are working to make knowledge and data freely and easily available, but challenges remain to open access. Developing countries, in particular, are hampered by a lack of infrastructure and capacity to support data collection, processing, and management, all necessary to ensure data quality. They also face limited standardization of data collection formats needed to make data easier to compare and aggregate.[7] A cornerstone of open access must be reducing the knowledge inequality within and among societies that arises from both lack of access and lack of capacity to make use of the world's growing store of knowledge and data.

MAINTAINING DATA QUALITY

Exchange of knowledge and data between the global, national, and local levels is essential to putting it to use. However, information and data quality can be lost as knowledge flows from one level to another. For example, inaccurate or incomplete transfer of knowledge to farmers can reduce its usefulness. Likewise, aggregation of data that are not standardized can be problematic. Preserving and enhancing data quality is critical for improving production and livelihoods and for credible policy making.

For example, the National Food Balance Sheets of the Food and Agriculture Organization of the United Nations (FAO) are compiled at the country level and aggregated to provide insight into global trends. The data sheets provide essential information on a country's food system, including trends in domestic food supply, use of commodities, and changes in the types of food consumed, and reveal whether a country's food supply is adequate to meet nutritional requirements.[8] However, data collection for the Food Balance Sheets varies widely from country to country, and because there is no standardization, cross-national comparisons can be misleading. Thus, even when data are openly available, poor quality or other limitations can decrease their utility.[9]

EASE OF USE

Access to knowledge and data depends not only on availability but also ease of use. The national surveys conducted by many countries provide a clear example. Even when these extensive sets of primary socioeconomic data are open, access is severely limited by the difficulty of using the large datasets. Often datasets available to individuals and researchers are so large and complex that they cannot be handled efficiently using traditional data-processing tools. India's National Sample

Survey Organization, which since 1950 has conducted large-scale surveys to provide data needed for national income estimation, is one example. In addition to making such individual datasets readily accessible, there is a growing need to make these datasets interoperable or machine readable so that they can be combined with other datasets to produce new insights on food and nutrition security and poverty reduction. Considering the large investments made in conducting national surveys, and the potential usefulness of the data for many stakeholders, the datasets should be made easier to use to ensure they are accessible to all.

REDUCING KNOWLEDGE INEQUALITY

Increasing ease of access and use is central to reducing the inequality of knowledge between developed and developing countries. In the past few decades, a vast amount of raw data has been made available for further exploration by stakeholders working on improving food and nutrition security. However, while open access to data has increased, limited infrastructure means that data are not available to users in developing countries at the same speed as in developed countries. As the world continues to digitize, knowledge inequality is actually increasing. According to Data.gov, a US government website that aims to make government more open and accountable through open data, 53 countries in the developed and developing world have set up national open data platforms.[10] However, open data initiatives are still limited in developing countries as a result of technological, economic, political, and social barriers. In the words of the Open Data Charter, more timely access to data would help increase food security globally by "enabling better decision making, transparency and innovation."[11]

PROGRESS TOWARD OPEN DATA

Democratization of data access is the key to evidence-based decision making for stakeholders ranging from farmers to policy makers. Broad consensus exists among stakeholders—including among governments, development institutions, research organizations and universities, publishers, and nonprofit and civil society organizations—that opening data access is a critical step toward attaining the Sustainable Development Goals. Yet while open data have been widely identified as a priority, commitment and action remain uneven. Some organizations are making substantial investments in open data. The Bill & Melinda Gates Foundation, for example, makes its data freely available immediately and without restrictions. Several international organizations, including the World Bank, CGIAR centers (including the International Food Policy Research Institute [IFPRI]), and United Nations' agencies such as FAO, are moving quickly toward open data but are yet to be fully FAIR-compliant—that is, findable, accessible, interoperable, and re-usable. Other institutions have shown less commitment to open data. But even among those that have made open data a priority, efforts to translate that commitment into action have been limited (Table 1).

Some countries have been in the vanguard in making data more accessible, taking steps to connect individuals directly with the knowledge and data they require. The Katalyst program in Bangladesh aims to increase household incomes across sectors, including agriculture. As part of the program, the Soil Resource Development Institute of the Ministry of Agriculture analyzed soil-sample data in various locations and for different crops. This information was used to develop recommendations for farmers on fertilizers, in order to optimize their use of inputs and improve yields. In collaboration with Banglalink and Grameen Phone, the Katalyst program launched a mobile-based fertilizer information service in the local language, Bangla. Since the program's inception in 2009, it is claimed that farmers have experienced up to a 25 percent reduction in fertilizer costs and up to a 15 percent increase in crop yields.[12] Based on this apparent success, Katalyst is launching a similar project to provide irrigation-related information.

In Peru, access to information about weather and climate patterns is expensive and limited. Data collected by the government cover only a small portion of the country. Capitalizing on Peru's large number of mobile-phone users, the Institute for University Cooperation (Istituto per la Cooperazione Universitaria Onlus), an Italian nongovernmental organization, developed a mobile platform that permits widespread sharing of information on climate and irrigation at a low cost. Farmers have access to relevant information on their crops based on climate, meteorological, and soil data.[13]

TABLE 1 Statement of purpose, commitments, and actions of different organizations

	Statement of Purpose, Commitments, and Actions
CGIAR	The CGIAR Open Access and Open Data Policy ensures that all research outputs are made open access. Data (and any relevant data collection and analysis tools) are deposited in a suitable repository and made open access as soon as possible and in any event within 12 months of completion. CGIAR strongly believes that this commitment to open access will improve the efficiency, efficacy, and impact of its research, aid interdisciplinary research and novel computation of research literature, and allow the global public to further benefit from the wealth of CGIAR research. For example, IFPRI established policies and protocols for making data global public goods; established an IFPRI Dataverse, an open source network hosted at Harvard; and established a Data Governance Team.
Bill & Melinda Gates Foundation (BMGF)	BMGF is committed to information sharing and transparency. It believes that published research should be promptly and broadly disseminated. It adopted an open access policy that enables unrestricted access to and reuse of all peer-reviewed published research funded, in whole or in part, by the foundation, including any underlying datasets.
World Bank	The World Bank recognizes that transparency and accountability are essential to the development process and central to achieving the Bank's mission to alleviate poverty. The Bank's commitment to openness is also driven by a desire to foster public ownership, partnership, and participation in development from a wide range of stakeholders. As a knowledge institution, the Bank's first step is to share its knowledge freely and openly. For example, the primary World Bank collection of development indicators is compiled from officially recognized international sources. It presents the most current and accurate global development data available, and includes national, regional, and global estimates.
US Agency for International Development (USAID)	USAID encourages its partners, the academic and scientific communities, and the public at large to make broad use of its Development Data Library (DDL) data for innovative scientific, technological, analytical, and other applications. USAID Operating Units consult datasets available through the DDL, as those data may prove useful in supporting evidence-based decision making across all stages of the USAID Program Cycle. For example, data could be used to: inform the design and implementation of USAID projects and programs by enabling additional analysis beyond that presented in written reports; adapt projects and programs based on learning from data analysis; incorporate baseline and contextual data to plan, measure, evaluate, adapt, and improve performance of development assistance activities; and facilitate the creation of data visualizations to generate additional insight for advocacy and training purposes.
European Union (EU)	The EU Open Data Portal (EU ODP) provides access to an expanding range of data from EU institutions and other EU bodies. This allows for data to be used and reused for both commercial and noncommercial purposes.

Source: Authors' compilation.

Colombia's Ministry of Agriculture and the International Center for Tropical Agriculture worked together from 2007 to 2013 to identify what was causing reductions in rice crop yields. Using both open and private data, they developed climate-smart tools that were made available to all farmers. Giving farmers access to these knowledge resources helped them avoid extreme damage from a subsequent drought and was estimated to have saved farmers about US$3.6 million.[14]

Other countries are working to scale up information and communication technologies and are promoting the use of open-source knowledge and information flows within their borders and with other countries. India's Public Distribution System has become more cost-effective and transparent with the adoption of a computerized system—consumers receive individual identification cards that they can use at any ration shop to purchase subsidized groceries. Data collected on these transactions are recorded and can be used to analyze the amount of stock a shop owner receives and disburses and who purchases the rations, as well as to identify duplicate cards.[15] Computerized monitoring of the supply and distribution of food grains ensures a more cost-effective system and reduces the opportunities for corruption in this distribution system.[16]

NEXT STEPS

Data policies, commitments, and action to promote openness, capacity building, technologies, and other assets have the power to reduce knowledge inequality and improve access to and use of knowledge. Together these can contribute to better livelihoods and food security. A few investments can make a big difference.

PUT DATA TOOLS IN FARMERS' HANDS. Simple steps can enhance the use of data by making them available in easily downloadable formats. The number of cell phone and smartphone users has dramatically increased in developing counties; in Africa south of the Sahara, more than 60 percent of the population now has access to mobile phones. Because mobile phones are cheap, easy to use, and require only limited literacy, mobile applications can be powerful mechanisms for increasing data accessibility. Some initiatives have begun. For example, the government of India created several mobile apps—AgriMarket, Kisan Suvidha, and Crop Insurance—for the benefit of farmers. These user-friendly apps can help reduce the digital divide among and within countries.

EMPOWER STAKEHOLDERS. To overcome knowledge and data access inequality, citizens must demand open data. Data curators can promote the use of open data and help get data into the hands of stakeholders, for example, by sharing data through well-targeted tweeting, press releases, website announcements, and email distributions; catering messages to target audiences; and choosing the right Open License for data.[17] Data users can also be empowered through capacity building, including trainings and tools for analysis, such as open access software, and support for information and communications infrastructure to enhance easy access.

IMPROVE KNOWLEDGE TRANSFERS. Increasing the efficiency of knowledge transfers can prevent loss of information during the process. A study in Malawi looked at the knowledge transmission chain from researchers to agricultural extension agents to "lead" farmers to other farmers, and suggested that the greatest loss of information was in the transfer from extension agents to lead farmers, potentially due to agents' and farmers' lack of attention.[18] New options are available for knowledge transfer—Digital Green, for example, is a nonprofit organization that trains farmers through videos.[19]

MAKE GOVERNMENT DATA PUBLIC. National governments collect the bulk of data related to food systems. However, national systems are slow to make the information public and many countries place restrictions on the use of this kind of data. If made openly accessible to the public, this "big data" could drive high-quality analyses of food systems and better decision making. India's National Sample Survey, for example, has collected data for over 50 years, providing invaluable insights on the country's population and challenges related to food and nutrition security. Analysis of these data is regularly published in *Agriculture Situation in India*, a monthly journal available to the public.[20]

BUILD OPEN DATA INITIATIVES. Realizing the importance and benefits of data accessibility, several key organizations including the World Bank, FAO, the United Nations Development Programme, and CGIAR/IFPRI are promoting big data and open data movements for the benefit of food and nutrition security.[21] Initiatives such as Global Open Data for Agriculture and Nutrition (GODAN) are working toward knowledge equality—helping developing countries achieve open data parity with their developed-country counterparts—and addressing pertinent issues including data quality, interoperability, storage, and dissemination.[22] IFPRI makes primary and secondary datasets freely available and encourages their use in research and policy analysis (Box 1). The World Bank's Open Data initiative is intended to provide all users with access to Bank data.[23]

INCREASE DATA QUALITY AND EASE OF USE. A number of steps can be taken to support use of knowledge and data from both national and international datasets. These include:

- collecting more nationally representative and real-time data that enhance understanding of food systems
- introducing tools and technologies that facilitate access to data
- creating forums and working groups to discuss food system data challenges
- building capacity for data collection, processing and preparation, analysis, presentation, and showing impact results
- adapting big data platforms to better facilitate data collection, curation, storage, analysis, search, sharing, transfer, visualization, querying, and information privacy

REAPING THE BENEFITS OF OPEN DATA

Open data and knowledge are essential to reap the benefits of globalization and to mitigate its costs. But increasing data access is not sufficient to achieve food security in developing countries. The gap between availability and usability of data must be bridged, and the quality and consistency of data must be improved. While there is broad commitment to open data and some notable advances have been made, greater investment is needed to build open channels for knowledge and data at the global, national, and local levels. More freely available and usable information can both reduce the digital divide and related inequality, and contribute to poverty reduction and food and nutrition security for all.

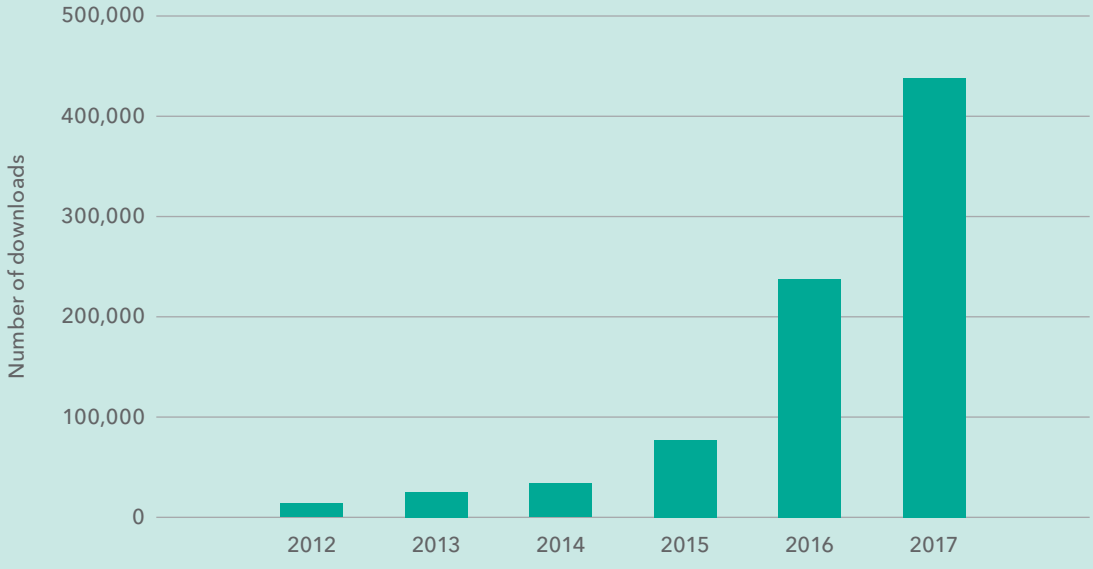
BOX 1 AN OPEN DATA INITIATIVE AT IFPRI

Open data require institutional support and strategic vision on the part of institutions collecting data from the public and using them for decision making, including the International Food Policy Research Institute (IFPRI). IFPRI's Open Data Initiative aims to amplify the potential impact of the Institute's research by disseminating it widely and in a timely fashion to researchers, policy makers, and the public. To help ensure that IFPRI's research data are easy to find, access, and use for decision making, the Institute has invested substantially in improving its data infrastructure, administration, and staff training over the past 20 years. To enhance the delivery and usability of data, IFPRI works with other CGIAR centers, government agencies, and international and academic communities in making its datasets "global public goods."

IFPRI began publishing and distributing data through CD-ROM in 1998, and two years later began requiring that its researchers publish all their datasets. These data have been shared on IFPRI's website since 2005 and, since 2008, through Harvard Dataverse, an open source network. IFPRI's data policy also establishes a timeline for making data public. In 2013, IFPRI along with other CGIAR centers endorsed CGIAR's Open Access and Open Data Policy, which requires that data be made public within one year.

To date, IFPRI has made some 370 datasets openly available, including household- and community-level surveys; country-level data; social accounting matrixes; and institution-level survey datasets. IFPRI datasets are downloaded thousands of times each year by a wide range of users working on food security and nutrition, including students, researchers, faculty, policy makers, and organizations worldwide.

IFPRI dataset downloads by year, cumulative



Source: Authors.

"A cornerstone of open access must be reducing the knowledge inequality within and among societies that arises from both lack of access and lack of capacity to make use of the world's growing store of knowledge and data."