



**How can we improve
agriculture, food
and nutrition
with open data?**

What is open data?

Data anyone can access, use and share. It must:

- be accessible, which usually means published on the web
- be available in a machine-readable format
- have a licence that permits anyone to access, use and share it

Closed data

Data that only data owners or people within an organisation can access, for reasons like **privacy**, **commercial sensitivity** and **security**.

Shared data

Data that is shared with specific people and organisations for a specific purpose: to provide services, connect information and contribute to research.

Find out more at theodi.org/what-is-open-data

Acknowledgements

Authors: Liz Carolan (The ODI), Fiona Smith (The ODI), Vassilis Protonotarios (Agro-Know), Ben Schaap (Wageningen UR), Ellen Broad (The ODI), Jack Hardinges (The ODI), William Gerry (The ODI).

With thanks to: Jeni Tennison (The ODI), Tim Davies (Practical Participation), Sander Janssen (Wageningen University), Martin Parr (CABI), Ana Brandusescu (CABI), Johannes Keizer (FAO), Fabrizio Celli (FAO), Nikos Manouselis (Agro-Know), Daniel Jimenez (CIAT), Shaun Hobbs (CABI), Christopher Brewster (Aston University), Gerbert Roerink (Wageningen UR), Medha Devare (CGIAR).

Editing: Anna Scott (ODI)

Production: Phil Lang (ODI)

Design: Adrian Philpott



Table of contents

1. Executive summary	4
2. Introduction	6
3. How open data is solving problems in agriculture and nutrition? 14 use cases	9
<p>Enabling more efficient and effective decision making</p> <ul style="list-style-type: none"> • Protecting crops from pest outbreaks with vegetation maps: GroenMonitor • Helping farmers forecast with weather apps and SMS: AWhere • Boosting crop yields with a best practice knowledge bank: Plantwise • Saving \$3.1m in drought damage with a climate-smart tool: CIAT Colombia • Managing the California drought with data visualisations: California Department of Water Resources 	9
<p>Fostering innovation to benefit everyone</p> <ul style="list-style-type: none"> • Saving crops and cash with weather simulation and smart insurance: Climate Corporation • Improving crop varieties with open data on breeding trials: AgTrials • Bringing agricultural research to the masses: FAO AGRIS portal • Making agri-food data more discoverable: the CIARD RING 	14
<p>Driving organisational and sector change through transparency</p> <ul style="list-style-type: none"> • Tracking water, pesticide, water and fuel use with an open, collaborative platform: Syngenta • Exposing misspent farm subsidies in Mexico: FUNDAR • Empowering consumers to make smart food choices: US national nutrient database • Helping consumers understand the risks of the food they eat: US Food alerts • Highlighting restaurant inspection scores and improving food safety: LIVES 	19
4. How ready are agriculture and nutrition for widespread open data innovation?	24
5. Realising the full potential of open data: next steps	26
About GODAN and the ODI	29
Glossary: Key data concepts	30
Appendix 1: Useful references and tools	31

1. Executive summary

In today's evolving global landscape we face complex challenges. Populations are growing, climates are changing and markets are often volatile.

As the world's population grows to around 9 billion by 2050,¹ global demand for food, feed and fibre is predicted to nearly double. The number of people at risk of hunger is likely to increase from 881 million in 2005 to more than one billion.²

Alongside these challenges come huge opportunities: a global data infrastructure is emerging and, with innovative business models and international political will, we believe now is the time to invest in open data-driven solutions in agriculture and nutrition.

In this discussion paper we highlight **three specific ways open data can help solve practical problems** in the agriculture and nutrition sectors:

- 1. Enabling more efficient and effective decision making**
- 2. Fostering innovation that everyone can benefit from**
- 3. Driving organisational and sector change through transparency**

We present a series of **14 use cases** showing how **open data can be useful in different stages of agriculture, food production and consumption**. From managing scarce water resources during the California drought or helping farmers in Africa estimate the outbreak of animal diseases, to helping consumers avoid harmful allergens in their food – open data is becoming a valuable tool for policy-makers, industry, small-scale farmers and consumers alike.

These are some examples of the range of stakeholders involved in agriculture and nutrition. We hope they **stimulate discussion about the potential uses, needs and challenges** of people, organisations and governments interested in exploring open data within this field.

There are still obstacles to realising open data's full potential in agriculture and nutrition.³ In a recent stakeholder survey we conducted for the Global Open Data for Agriculture and

1 Estimates and UN Population Division (2012), World Population Prospects: The 2012 Revision, available at: <http://www.unfpa.org/world-population-trends#sthash.MYReYwCX.dpuf>, accessed 18/05/15

2 IFPRI (2014), Food Security in a World of Natural Resource Scarcity: The role of agricultural technologies: <http://www.ifpri.org/publication/food-security-world-natural-resource-scarcity>, accessed 27/04/15

3 See Overseas Development Institute (2015), The data revolution: Finding the missing millions: http://www.developmentprogress.org/sites/developmentprogress.org/files/case-study-report/data_revolution_-_finding_the_missing_millions_-_final_20_april.pdf, accessed 15/05/15

Nutrition (GODAN) initiative,⁴ we found that although the amount of data openly available is constantly increasing, there are still **challenges related to data management, licensing, interoperability and exploitation**. There is a need to evolve policies, practices and ethics around closed, shared, and open data.

The best outcomes are realised when open data is used to enable, enhance and unlock the wider potential through focused collaboration: these use cases were developed as solutions to specific problems, such as drought, pest infection or food safety concerns.

We propose establishing projects and initiatives that start with **real-world problems**, and engage with problem owners to identify how data can provide actionable information that helps to address needs. To achieve this, we set out five steps for pursuing **solution-focused open data initiatives for agriculture and nutrition**:

1. **Engage with the growing open data community**, including key problem owners and experts at GODAN, to identify the challenges that open data can help solve.
2. **Build open data strategies and projects** with a focus on finding solutions to these agriculture and nutrition problems.
3. **Develop the infrastructure, assets and capacities for open data** in relevant organisations and networks.
4. **Use open data and support users of relevant data.**
5. **Learn through ongoing evaluation, reflection and sharing** to ensure we can all continue to improve our practice. GODAN provides a forum for shared learning, through creating case studies, mapping partner activity, and bringing partners together. Advocate for the data you use from others, and that you produce, to be provided as part of the commons of open data in order to help stimulate network learning.

Do you have an idea for a solution-focused open data initiative for agriculture and nutrition? Submit it at bit.ly/godan-ideas.

⁴ Results forthcoming in 2015.

2. Introduction

Agriculture and food are major contributors to the global economy, underpinning livelihoods and economic growth in the developed and developing world alike. Overall, the agriculture and food industries account for 6% of GDP in the EU, comprising 15 million businesses and 46 million jobs.⁵ Meanwhile, agriculture accounts for 65% of Africa's workforce and 32% of the continent's GDP.⁶ In some of Africa's poorest countries, including Chad and Sierra Leone, it accounts for more than 50% of GDP.⁷

Yet the global food system is struggling under the combined pressures of a growing population, climate uncertainty and volatile market forces. As the world's population grows to around 9 billion by 2050,⁸ global demand for food, feed and fibre is predicted to nearly double, with the number of people at risk of hunger increasing from 881 million in 2005 to more than a billion.⁹ Fundamental concerns around inequality of access to nutritious food also persist. The Food and Agriculture Organisation of the United Nations (UN FAO) estimates that one in nine people in the world suffered from chronic undernourishment between 2012-2014.¹⁰

Alongside these shifts, new and traditional demand for agricultural produce will put growing pressure on scarce resources. While agriculture competes with sprawling urban settlements for land and water, it will also be required to serve on other major fronts: adapting to (and helping to mitigate) climate change, preserving natural habitats, protecting endangered species and boosting biodiversity. Climate change is expected to make it more difficult to continue to grow crops and raise livestock in the ways and places we have been.¹¹ The increased unpredictability of weather patterns will also make it more difficult to make sound decisions.

5 European Commission (2013), The common agricultural policy and agriculture in Europe: Fact sheet http://europa.eu/rapid/press-release_MEMO-13-631_en.htm, accessed 27/04/15

6 World Bank (2013), World Bank Agriculture and Africa: Fact sheet, <http://web.worldbank.org/WBSITE/EXTERNAL/COUNTRIES/AFRICAEXT/0,,contentMDK:21935583~pagePK:146736~piPK:146830~theSitePK:258644,00.html>, accessed 18/05/15

7 The World Bank (2015), <http://data.worldbank.org/indicator/NV.AGR.TOTL.ZS>, accessed 05/05/15

8 Estimates and UN Population Division (2012), as above.

9 IFPRI (2014), as above.

10 FAO (2014), State of food insecurity in the world 2014, <http://www.fao.org/publications/sofi/2014/en/>, accessed 27/04/15

11 USEPA (2013), Climate impacts on agriculture and food supply, <http://www.epa.gov/climatechange/impacts-adaptation/agriculture.html>, accessed 27/04/15

Poor access to market information and credit has made it hard for smallholder farmers to respond to extremely volatile food prices.¹² Farmers struggle to plan their economic activities and make decisions regarding when and how much to plant. Providing farmers with more accurate, accessible, timely market information – from large agriculture groups to the individual smallholders – will help to ensure food commodity markets function well in future.

The ‘Sustainable Development Goals’ (SDGs) – a framework for the next global development agenda, led by the United Nations – have highlighted “achieving food security and improved nutrition and promoting sustainable agriculture” as a global priority for the next 15 years.¹³

As the World Bank has stated, increasing smallholder farmers’ productivity and access to markets can have “a profound impact on the livelihoods and general prosperity of literally millions of the world’s poor.”¹⁴

This is no simple task. It will mean sustainably increasing agricultural productivity, while shaping more efficient and equitable markets. Progress will be driven largely by providing better access to accurate, timely information for individual smallholder farmers, businesses and policy-makers alike. **Open data can and should be part of the solution.**

Fortunately, information and social structures exist in different parts of the world that can support the uptake and implementation of open data. Open data builds upon and extends a rich tradition of knowledge and information sharing in the sector. Since the 1990s, ‘agricultural knowledge systems’ have been used to promote knowledge exchanges among farmers (farmer field schools, for example), and with industry, public and private research and education institutions.¹⁵ Data from cultivar trials, as we discuss below in the AgTrials use case, has been shared in other formats even before the internet era.

12 See for example: Committee on World Food Security (CFS), Price Volatility and Food Security. A report by the High Level Panel of Experts on Food Security and Nutrition, [http://www.fao.org/fileadmin/user_upload/hlpe/hlpe_documents/HLPE-price-volatility-and-food-security-report- July-2011.pdf](http://www.fao.org/fileadmin/user_upload/hlpe/hlpe_documents/HLPE-price-volatility-and-food-security-report-July-2011.pdf), accessed 08/05/15

13 UN-DESA/DSD (2014), Open Working Group proposal for Sustainable Development Goals, <https://sustainabledevelopment.un.org/?page=view&nr=164&type=230&menu=2059>, accessed 08/05/15

14 World Bank (2013), as above.

15 Röling, N. (1990), The agricultural research-technology transfer interface: a knowledge systems perspective, in Kaimowitz, David (ed) (1990), Making the link: Agricultural research and technology transfer in developing countries: 1-42. http://pdf.usaid.gov/pdf_docs/pnabe269.pdf#page=15, accessed 15/05/15

In this paper we outline how open data is boosting innovation in agricultural and nutritional business and service models. We explain how **open data promotes transparency across the sector to accelerate progress, identify areas for improvement and help create new insights**. Finally, we suggest how this important progress can continue to scale.

We do this through presenting a series of use cases from around the world. These demonstrate some of the ways open data can help solve problems for a range of stakeholders, from smallholder farmers to agribusinesses. These use cases were identified through a process of rapid desk research and key informant interviews with a sample of GODAN members. It therefore does not represent the voices of all 125 GODAN members. We hope that this paper will lead to more conversations among the GODAN network and wider global development community about the potential for open data in agriculture and nutrition.

3. How is open data solving problems in agriculture and nutrition?

Because open data is data that anyone can access, use and share, it is shaping solutions to problems that would otherwise be expensive, time intensive or impossible to solve using closed data sources. Through speeding up innovation, open data fosters collaboration between governments, businesses, NGOs and individuals to make new discoveries and help sustainably feed a growing population.

We have identified three key ways in which open data can play a critical role in addressing challenges in agriculture and nutrition:

1. **Enabling more efficient and effective decision making**
2. **Fostering innovation that everyone can benefit from**
3. **Driving organisational and sector change through transparency**

1. **Enabling more efficient and effective decision making**

The World Wide Web has fundamentally transformed the way we access and use information. We have reached a point where information can be gained from a variety of online sources as and when we need it. As more information is made available online, this helps us make faster, more informed decisions.

Open data works in the same way as the web, except it enables computers to pull data from various sources and process it for us, and does not rely on humans to interpret and integrate information contained in web pages. Open data underpins new products and services by presenting information from a wide range of sources that helps everyone from policy-makers to smallholders find gaps in markets or fine-tune their products or services.

Open data

- **lowers direct costs associated with accessing data** – i.e. licensing fees and subscriptions that can put off new providers and users
- **lowers indirect costs of data access** – the time and resources it takes to locate, request and negotiate access to data
- **drives better decision making and improved products and services** based on lower costs and the wider availability of data

One key way in which open data is transforming the agricultural sector is by underpinning a variety of **tools for farmers** themselves. A range of applications and services have been built using open data to help keep farmers informed of various issues that affect their work: from pest prevalence and weather predictions to crop growth and market prices.

Use cases

Protecting crops from pest outbreaks with vegetation maps: GroenMonitor

Farm productivity is often hit by crop damage caused by pests. Mice and other pests are difficult to detect on large farms through manual inspection alone. The **GroenMonitor** (GreenMonitor) is a tool that shows a current vegetation map of the Netherlands, based on satellite images and maps made publicly available through the European Space Agency (ESA).¹⁶ This makes pest outbreaks easy to identify and mitigate relatively quickly. In 2014, the **GroenMonitor** helped to identify 12,000 hectares (29,652 acres) of fields affected by mice.¹⁷ The tool is now being exploited for various other applications, including plant phenology, crop identification and yield, identification of agricultural activities (e.g. mowing, ploughing and harvesting), nature and water management.

Helping farmers forecast with weather apps and SMS: AWhere

It is often difficult for farmers to access essential information that affects their cultivation practices – such as temperature highs and lows, humidity and precipitation, especially in areas with low internet access. A number of data providers now provide individual farmers with much needed weather information, including commercial companies like **AWhere**. Via its global database, Weather Terrain, AWhere combines weather observations, forecast models and historical records from all over the world and downscales them to field level, helping farmers better forecast and plan activities.

Many farmers, especially in the developing world, use mobile phones (rather than computers) as their main communication tool. With this in mind, one group in Ghana worked with AWhere to develop an app on top of Weather Terrain's open API to make

¹⁶ GroenMonitor website:
<http://www.groenmonitor.nl>, accessed 12/05/15

¹⁷ See http://www.groenmonitor.nl/sites/default/files/Groenmonitor_muizenschade_9jan2015.pdf, accessed 7/05/15

their rich data accessible through mobile phones. The weather information was transformed into a simple SMS messaging service, which uses basic keywords (e.g. partly sunny, partly cloudy, windy) and images. Farmers can now access weather information at a low cost to help them make decisions regarding cultivation.¹⁸

In other cases, openly licensed data may be combined with other private (restricted) data sources to gain new insights. This can provide a more accurate picture of issues affecting the sector as a whole, where information is otherwise dispersed or inaccessible.

Use cases

Boosting crop yields with a best practice knowledge bank: Plantwise

Plant pests and disease are currently responsible for about 40% of global crop production losses.¹⁹ **Plantwise** helps smallholder farmers in developing countries deal with plant health issues. It aims to increase food security and improve rural livelihoods by reducing crop losses from pests and diseases. It does so by combining global and local open access data from sources such as CABI's databases, research publications and governmental data. It makes the data available and easy to search for via an online platform. Reports of disease from plant clinic operations on the ground are also used to supplement the knowledge bank and notify local partners of pest issues.

In two years the Plantwise knowledge bank has become a vital tool to support plant clinic operations in 33 countries. Over 600,000 farmers from 198 countries have visited the knowledge bank including over 9,000 factsheets to access critical agricultural data on crop pest prevalence and best practices to help manage and prevent potential crop loss from pests and diseases.²⁰

18 O'Reilly Radar Podcast (2015), <http://radar.oreilly.com/2015/03/big-datas-impact-on-global-agriculture.html>, aired 13 March 2015, accessed 7/05/15

19 CABI (2010), CABI in review 2010, <http://www.cabi.org/Uploads/CABI/about-us/4.8.1-annual-reviews-and-corporate-accounts/annual-review-2010.pdf>, accessed 27/04/15

20 In recognition of this the Plantwise programme was announced as the winner of the Open Data Award for Social Impact in 2014: <https://www.plantwise.org/news-and-media/odi-award-for-plantwise>

Saving \$3.6m in drought damage with a climate-smart tool: CIAT Colombia

A recent collaboration between public (open) and private data sources is helping farmers take precautions to avoid drought damage in Colombia. Between 2007-2013 an association of farmers (National Federation of Rice Growers; Fedearroz), an international research centre (Centro Internacional de Agricultura Tropical; CIAT) and Colombia's Ministry of Agriculture joined forces to identify the issues behind yearly reductions in rice crop yields, one of the most important food crops for the country.²¹

Making use of both open and private data (the latter obtained from companies through special agreements), CIAT analysed large datasets from annual rice surveys, harvesting records, field experiments and weather data and identified the complex and region-specific issues behind the decreasing rice crop yields. This led to the development of a climate-smart agriculture decision-making tool for the Colombian rice growers, which is openly available to anyone.²²

The impact on the agricultural sector as well as the Colombian economy was significant. Actions informed by this data helped farmers avoid extreme damage from the drought saving an estimated \$3.6m of potential economic losses.²³

Open data can also provide a **tool for policy-makers, researchers and organisations** to plan and distribute resources around critical information. Making basic geospatial data available for anyone to use and share, for example, has been particularly valuable when responding to natural disasters such as the 2010 Haiti earthquake. By offering more detailed and accurate insights into local needs, and the extent of damage, open data can help to improve coordination in disaster response, as demonstrated by the government response to the California drought, explained in the following use case.

21 Overseas Development Institute (2015), The data revolution Finding the missing millions, p.37, <http://www.odi.org/sites/odi.org.uk/files/odi-assets/publications-opinion-files/9604.pdf>, accessed 12/05/2015

22 Clima y Sector Agropecuario Colombiano, <http://www.aclimatecolombia.org>, accessed 12/05/2015

23 UN Global Pulse (2014), Big Data Climate Change Winners Announced, <http://www.unglobalpulse.org/big-data-climate-challenge-winners-announced>, accessed 6/05/15

Managing the California drought with data visualisations: California Department of Water Resources

California is experiencing one of its most severe droughts on record. Water shortages pose a critical threat to the agriculture sector, which accounts for 80% of water consumption across the state. Economic estimates in 2014 predicted direct costs of \$1.5bn to the agricultural sector, and the loss of 17,000 jobs on top of reduced food production.²⁴ To ensure a safe and sustainable water supply, the **California Department of Water Resources** announced a water rationing plan, reducing water allocation to farmlands and cutting consumer water usage by 25%.

Open data is being used to inform how the state allocates its scarce water resources under these conditions.²⁵ The drought has been visualised by the US Geological Survey (USGS) with publicly accessible open data collected by the USDA's network of research facilities, on long-term physical, chemical, and biological data on agricultural sustainability, climate change, and natural resource conservation at the watershed or landscape scale.²⁶ This allows researchers and policy-makers to monitor conditions and plan water management.²⁷ Models based on the data, which is constantly updated, estimate actual water volume levels, water consumption and other factors, allowing for timely forecasting and decision making regarding how much water to use in agriculture.

The US Department of Agriculture (USDA) is also making research data related to the California drought open.²⁸ Dr Catherine Woteki, Under Secretary for the United States Department of Agriculture's (USDA) Research, Education and Economics division, hopes this will "stimulate the use of open data by both the private and public sectors to give decision support for farmers on water use, crop choice [and so on]."²⁹

24 Howitt, Richard et al (2014), Economic analysis of the 2014 drought for Californian agriculture, https://watershed.ucdavis.edu/files/biblio/Economic_Impact_of_the_2014_California_Water_Drought_1.pdf, accessed 7/05/15

25 The Open Governance Council (2015), Online Conference "Open Data from Space to Solve Drought in California" <http://www.meetup.com/The-Open-Governance-Council/events/221756087/?eventId=221756087&action=detail>, accessed 30/04/15

26 Data providers include U.S. Drought Monitor, U.S. Bureau of Reclamation, U.S. Army Corps of Engineers, CA Data Exchange, U.S. Natural Resources Conservation Service SNOTEL, U.S. Geological Survey GAGESII and NWIS. See http://cida.usgs.gov/ca_drought

27 USGS - CIDA (2015), Datasets and docs for CIDA's data visualization challenge, <https://github.com/USGS-CIDA/CIDA-Viz>, accessed 30/04/15

28 <https://twitter.com/godanSec/status/590603899310108674>

29 <https://twitter.com/godanSec/status/590613377745080321>

2. Fostering innovation to benefit everyone

As a raw material for creating tools, services, insights and applications, open data makes it inexpensive and easy to create new innovations. When data is open for all to experiment with, there is no need to invest large amounts in repeating already completed trials. When data is openly licensed, it also allows for novel combinations with other data to gain new insights.

Open data provides SMEs, startups and other organisations with a level playing-field, exposing gaps in markets and helping them compete against established market players to deliver new products and services. It also benefits established companies, who learn from and react to innovation in their sector – they might invest in these new products and services being delivered, acquire new talent and adjust their own business practices. In the agricultural sector, large open datasets have stimulated business creation and provided farmers with advisory services that boost their productivity.

Entities and people sometimes might release data as open data, without the tools or capacity to make use of it themselves, especially in the context of smallholder farmers. It is important to explore ways to overcome this issue:

“Open data is an opportunity that we simply cannot miss for creating new businesses and jobs. Creating new products or enhancing existing ones using open data requires a wide, diverse and often domain-specific skill-set that can boost our knowledge economy. In the agri-food sector, I would like to see new data-powered innovation ecosystems to emerge, grow and flourish in the years to come. And I would like to see new business models that will bring value, and possibly revenue, to the public and private stakeholders that release open data.”

Nikos Manouselis, Agro-Know

Use case

Saving crops and cash with weather simulation and smart insurance: Climate Corporation

In the past, farmers struggled with predictive climate models that failed to take into consideration local conditions, leading to inefficient risk calculations. Climate Corporation is an open data business that offers more accurate insurance³⁰ and a commercial advisory service to help farmers manage and adapt to climate change.³¹

They do this through analysing huge volumes of data points from open and other data sources, to simulate weather events and assess risk to the yield of specific crops. The company utilises open data from sources including the National Oceanic and Atmospheric Administration (NOAA), the Next Generation Radar (a network of 159 Doppler radar stations operated by the National Weather Service), as well as maps of terrain and soil types from the U.S. Geological Survey.³²

Farmers can use the detailed weather forecasting data to enhance their cultivation practices and activities like spraying, fertilising and seeding. For example, using moisture and precipitation maps provided by the company, farmers can tell if a specific part of their fields is too wet to be ploughed. On an industry level, the impact of the company's open data-powered service could be significant. In 2013 Climate Corporation's customers farmed more than 10 million acres.³³

30 See for example the Climate Insurance offering
http://www.denodo.com/en/system/files/document-attachments/case_study_the_climate_corporation.pdf

31 Climate Corporation (2015), Climate Insurance,
<http://www.climate.com/products/mp-federal-crop-insurance>, accessed 14/05/15

32 Open Data Now (2013), Updated: How Open Data Can Fight Climate Change,
<http://www.opendatanow.com/2013/06/new-how-open-data-can-fight-climate-change>, accessed 30/04/15

33 The New Yorker (2013), Climate by Numbers,
<http://www.newyorker.com/magazine/2013/11/11/climate-by-numbers>, accessed 12/05/15

Open data helps governments, established companies and other organisations to outsource and/or crowdsource their research and development. By releasing the data for anyone to use they can benefit from a large group of people – outside the original data holders – who are solving problems with the data. For example, the online atlas RTBMaps aggregates individual maps provided by researchers and scientists through ‘scientific crowdsourcing’ to produce a unique map showing variables affecting crops, such as susceptibility to pests and diseases or vulnerability to failed harvests.³⁴ Its 32 map layers are made available online for anyone to download and use for advanced analysis.

In many cases agricultural scientists focus on answering very specific questions. They may not consider other potential uses for their raw data and other research data; but even when they do, they need a centralised platform or repository to share their results. In this context, opening data up can help promote better data management and workflows.

Meanwhile, accessing research findings in the form of closed publications can be prohibitively expensive.³⁵ But thanks to the open access mandates of an increasing number of major research funders,³⁶ more of such findings (including data) will be made open and free to anyone. In particular, the opening of scientific research and global datasets for application to local conditions by smallholder farmers has the potential to transform the structure of agricultural research and innovation systems.³⁷

“Open data makes all research accessible to target users and helps in the efforts of open access initiatives to strengthen agricultural research for development, support decision makers, improve livelihoods, and support food security.”

**Ahlam Musa, Agricultural Research Corporation (ARC) Central Library,
Sudan AGRIS Resource Center**

34 CIAT blog (2014), Open data’s contribution to better research for development: The example of RTBMaps, <http://www.ciatnews.cgiar.org/2014/10/20/open-data-rtbmaps>, accessed 15/05/15

35 Center for Global Development Blog (2015), The Price of Knowledge in a World of Fee-Based Scientific Journals, <http://www.cgdev.org/blog/price-knowledge-world-fee-based-scientific-journals>, accessed 17/05/15

36 University of Glasgow (2015), Research Funders: Open Access Policies, <http://www.gla.ac.uk/services/openaccess/managingyourresearchpublications/funders>, accessed 14/5/15

37 Maru, A. et al (2014) ICTs Improving Family Farming in Agrifuture Days 2014, Villach, Austria.

Use cases

Improving crop varieties with open data on breeding trials: AgTrials

Cultivar testing is an important means of improving crop varieties. A wide range of trials are taking place on sites all over the world, addressing issues such as drought tolerance, heat stress, and soil management. However, almost all of the data generated has been inaccessible to other researchers – filed away on laboratory hard drives, or sometimes lost completely due to bad data management.

By compiling data from agronomic and plant breeding trials and making it open, the Global Agricultural Trial Repository (**AgTrials**)³⁸ hosted by a CGIAR Research Programme on Climate Change, Agriculture and Food Security (CCAFS) offers a rich knowledge base to inform ongoing, collaborative research, while eliminating unnecessary and costly duplication of efforts.

Scientists used 250 open AgTrials datasets to build crop models specific to the West Africa region. The models are used to project the local impacts of climate change, and define breeding programmes for adaptation.³⁹

Bringing agricultural research to the masses: FAO AGRIS portal

AGRIS⁴⁰ is an international network of research institutions and information nodes making agricultural research information globally available. It collects and disseminates bibliographic information on diverse food and agricultural publications, from over 150 data providers in 65 different countries.

AGRIS uses bibliographic data as an aggregator for locating related content online and organises it via an open data repository (of over 8 million records). An application combines records with other open data repositories and links to other quality sources of data such as the World Bank, Nature, and the Chinese Germplasm Database.

38 AgTrials website: www.agtrials.org, accessed 03/05/15

39 CGIAR CCAFS (2015), AgTrials helps repurpose data for adaptation research, <http://ccafs.cgiar.org/AgTrials>, accessed 15/05/15

40 FAO AGRIS Portal: <http://agris.fao.org>, accessed 03/05/15

The AGRIS portal has received more than 7.5 million visitors from 204 countries. For many undergraduate and graduate students – AGRIS is the most important access point to scientific and technical information.

Making agri-food data more discoverable: the CIARD RING

Despite the fact that there is already a wealth of information related to open data (like datasets, portals and standards), the retrieval of relevant information is still an important issue that needs to be addressed. In this context, the **CIARD Routemap to Information Nodes and Gateways (RING)**⁴¹ acts as a global registry of web-based information services and datasets for agricultural research for development (ARD).

The registry allows data providers to register and categorise their services. It also promotes and leverages standards by ensuring all datasets are accompanied by metadata about which standards are adopted (such as vocabularies, dimensions and protocols). This promotes data reuse and discovery, and allows for greater automation. It currently features over 1000 information services of which about a third are agri-food datasets. More than 500 registered data providers have published about 30 million records through the registered datasets. The RING is used by AGRIS, AgriFeeds,⁴² the agINFRA harvester API⁴³ and Drupal modules, the International Livestock Research Institute (ILRI)⁴⁴ and other global initiatives in agri-food research.

41 CIARD RING: <http://ring.ciard.net>, accessed 18/05/15

42 AgriFeeds RSS service: <http://www.agrifedds.org>, accessed 15/05/15

43 agINFRA (2015), agINFRA REST API docs, <http://aginfra.eu/api/>, accessed 15/05/15

44 International Livestock Research Institute (ILRI) website, <http://www.ilri.org>, accessed 12/05/15

3. Driving organisational and sector change through transparency

Open data helps shape best practice within a sector, market or organisation. Transparency around targets, subsidy distribution and pricing, for example, creates incentives which affect the behaviours of producers, regulators and consumers.

Where an organisation has targets, open data about its work makes progress transparent and accessible across the sector chain. When an outcome or target is measured, and those measures made openly available, organisations and individuals adjust their behaviour so that they perform better against those measures. As a basic example, to reduce the number of single-use carrier bags provided by supermarkets, the UK Government requires sellers to keep records and report the numbers of carrier bags that are used.⁴⁵ Here, transparency motivates sellers to reduce their use.

By requiring companies, government departments and other organisations to publish key datasets – performance data, spend data or supply-chain data, for example – governments, regulators and companies can **monitor, analyse and respond to trends** in that sector. More importantly, publishing this data across a sector can ultimately transform how products and services are delivered.

Many resources can be overused or misspent in agriculture systems, resulting in wasted money and environmental damage.

The unnecessary use of agrochemicals (such as fertilisers and pesticides) can damage the environment and raise production costs for farmers. In the United States alone, \$10bn are annually spent on pesticides. While pesticide use saves around \$40bn on crops that otherwise would be lost to pest destruction, this does not take into account any of the negative effects that result from the use of these pesticides. The issues related to the use of pesticides lead to a social cost of more than \$12bn in addition to the \$10bn spent for these pesticides.⁴⁶ If standards and best practices are openly available, these outputs can be monitored by farmers and their unnecessary costs can be prevented.

45 House of Commons (2014), Plastic bags - Environmental Audit Committee Contents, <http://www.publications.parliament.uk/pa/cm201314/cmselect/cmenvaud/861/86104.htm>, accessed 03/05/2015

46 Steven Gilbert (2014), Economics of Pesticide Usage, <http://www.toxipedia.org/display/toxipedia/Economics+of+Pesticide+Usage>, accessed 12/05/15

Farm subsidies generally represent a huge cost to governments; for example the European Union spends approximately €59bn on subsidies each year. However, subsidies do not always reach those who need them most. By publishing data about the distribution of agricultural payments, such as subsidies, government agencies can adjust targeting so it reaches those who really need it, thereby helping to strengthen the sector as a whole.⁴⁷

Use cases

Tracking water, pesticide, water and fuel use with an open, collaborative platform: Syngenta

In 2013, Syngenta announced its Good Growth Plan: six commitments to improve crop productivity, protect soil and biodiversity, train smallholders and ensure labour standards, with targets to be achieved by 2020. The initiative aims to allow farmers to increase crop yields in a sustainable way, through monitoring activities like nutrient and pesticide application and water and fuel usage.

A data management system has been established to track these agricultural outputs and inputs using farm and public data, which is collected, validated and analysed by independent companies. Baseline data for 2014 on all commitments are available in a machine-readable format (as CSV files) and published using different Creative Commons licenses⁴⁸ (such as NonCommercial-NoDerivatives and Attribution-ShareAlike, an open data license).⁴⁹ Through these actions, Syngenta in collaboration with ODI, aims to build an open, collaborative platform to co-create solutions that minimise the use of resources required to feed a growing population and preserve habitats for biodiversity.

“Making this data public will allow people to make their own assessments of the progress of our Good Growth Plan. It is also blurring the traditional roles of business, government and NGOs by highlighting our collective responsibility to address acute global challenges. Above all, the data will be of value to farmers, enabling them to increase productivity sustainably and to enhance their livelihoods.”

Mike Mack, Syngenta

⁴⁷ See for example <http://farmsubsidy.openspending.org>

⁴⁸ Creative Commons (2015), About the Licenses, <http://creativecommons.org/licenses>, accessed 12/05/15

⁴⁹ Syngenta (2014), The Good Growth Plan, <http://www.syngenta.com/global/corporate/en/goodgrowthplan/home/Pages/homepage.aspx>, accessed 10/05/15

Exposing misspent farm subsidies in Mexico: FUNDAR

PROCAMPO is the largest federal farm subsidy programme in Mexico supporting the poorest farmers. There have been concerns since 2007 that its subsidies were not received by those meeting the requirements, who were in dire need of support.

To better understand the situation, a Mexican NGO named **FUNDAR Center of Analysis and Research**⁵⁰ called for information related to the distribution of subsidies from the Mexican Ministry of Agriculture. After initial requests resulted in incomplete data in non-machine readable formats, the agency in charge finally published the data. Analysis showed that 57% of the benefits were distributed among the wealthiest 10% of recipients,⁵¹ confirming initial fears.

An important outcome of the data release was the development of a database (Subsidios al Campo en México)⁵² by FUNDAR and other NGOs, which publishes ongoing information about the farm subsidies to ensure more transparency over the process.⁵³ A series of resignations followed the revelations and the government imposed limits on the eligibility of subsidies.⁵⁴

Open data also **enables consumers to demand better standards**, through promoting transparency around elements of sustainable food production and safety. This has a dual effect of influencing consumer behaviour, while creating incentives for food producers to improve their performance.

In the food sector, requiring the publication of certain information – such as ingredient lists, greenhouse gas emissions, allergen information and the origin of produce – serves three purposes:

50 Fundar (<http://fundar.org.mx>) acts as an independent, interdisciplinary and plural organization devoted to research issues related to democracy and citizen participation.

51 Pirker, K. (2010), Subsidios al campo en México: Contraloría social, nuevas tecnologías e incidencia en políticas públicas, UNAM, 1 September 2010.

52 Subsidios al Campo en Mexico website: <http://subsidiosalcampo.org.m>

53 53 Cissé, H., Madhava Menon, N.R.; Cordonier Segger, M-C., Nmehielle, V. (2014), The World Bank Legal Review, Volume 5: Fostering Development through Opportunity, Inclusion, and Equity. Washington, DC: World Bank. <https://openknowledge.worldbank.org/handle/10986/16240>, accessed 03/05/15

54 See above.

1. It helps regulators **identify areas where food standards are not being complied with** or where targets are not being met
2. It **informs consumers** so they can make better choices
3. It **drives sector change** in response to this increased consumer choice

An initiative recently developed to meet these aims in the UK is **FoodTrade Menu** by Food Trade,⁵⁵ which combines data provided by the food business, government open data, and data from food producers to provide restaurants with tailored food menus flagging these allergens in their food to their customers.⁵⁶

Use cases

Empowering consumers to make smart food choices: US national nutrient database

Consumers have clearly indicated that they want to be more well-informed on the quality and ingredients of the food that they are consuming. Although basic information already exists on food packaging, more detailed information on food nutrients could allow people to make better decisions regarding food selection based on their individual needs (e.g. following the advice of a dietitian).

The **USDA National Nutrient Database for Standard Reference (SR25)**⁵⁷ is the major source of food composition data in the United States, and provides data sources for most public and private sector databases. SR25 contains nutrient data for more than 8,500 food items and about 150 food components, such as vitamins, minerals, amino acids, and fatty acids. The use of this data is not limited to commercial applications (e.g. smartphones apps). It provides the basis for new services like ChooseMyPlate.gov, an initiative launched by US First Lady Michelle Obama and USDA Secretary Tom Vilsack to provide “practical information to individuals, health professionals, nutrition educators, and the food industry to help consumers build healthier diets with resources and tools for dietary assessment, nutrition education, and other user-friendly nutrition information.”⁵⁸

55 Nesta (2015), Food Open Data Challenge, <http://www.nesta.org.uk/project/the-open-data-challenge-series/food-open-data-challenge>, accessed 10/05/15

56 Nesta (2015), Being a food open data “prosumer” - FoodTrade Menu winner of Food Open Data Challenge, <http://www.nesta.org.uk/blog/being-food-open-data-prosumer-foodtrade-menu-winner-food-open-data-challenge>, accessed 10/05/15

57 USDA (2015), National Nutrient Database for Standard Reference Release 27: <http://ndb.nal.usda.gov/ndb/search/list>, accessed 15/05/15

58 USDA (2015), Choose My Plate: <http://www.choosemyplate.gov/about.html>, accessed 18/05/15

Helping consumers understand risks of the food they eat: EU Food alerts

Food safety is another important issue affecting consumers. The European RASFF (**Rapid Alert System for Food and Feed**) portal⁵⁹ provides access to a database of publicly available information about recently transmitted food safety alerts and notifications.

Consumers can access data on food safety issues such as the presence of allergens, pathogens, toxins or other harmful substances in food products,⁶⁰ and share preventative information.⁶¹ Following the March 2011 Fukushima nuclear power plant incident, the RASFF was used to monitor fishery and other marine products caught in the Pacific region for the presence of radioactive substances which would be harmful to consumers.⁶²

Highlighting restaurant inspection scores and improving food safety: LIVES

Open data is also being used to help consumers choose where to dine, while incentivising improvements in food safety. **Local Inspector Value-Entry Specification (LIVES)** aims to “normalise restaurant inspection scores across jurisdictions, allowing consumers to get a sense for restaurant food safety compliance across municipalities and within their home town.”⁶³ LIVES was launched in 2013 as a project between San Francisco, Socrata, Code for America, and Yelp, and is providing the standard for publishing open data on restaurant inspections⁶⁴. By allowing citizens to make better use of inspection results, LIVES facilitates food transparency and decision making on approved restaurants. When the City of Los Angeles began to require that restaurants displayed hygiene grade cards on their entrances, studies found it was associated with a 13% decrease in hospitalisations due to foodborne illness.⁶⁵

59 RASFF (2015), The RASFF portal: <https://webgate.ec.europa.eu/rasff-window/portal>, accessed 16/05/15

60 RASFF (2015), Consumers' Portal: <https://webgate.ec.europa.eu/rasff-window/consumers>, accessed 16/05/15

61 RASFF - The Rapid Alert System for Food and Feed (2009), 30 years of keeping consumers safe: http://ec.europa.eu/food/safety/rasff/docs/rasff30_booklet_en.pdf, accessed 18/05/15

62 RASFF - The Rapid Alert System for Food and Feed (2011), Annual Report http://ec.europa.eu/food/safety/rasff/docs/rasff_annual_report_2011_en.pdf, accessed 15/05/15

63 Socrata blog (2015), Could the LIVES standard reduce food poisonings?, <http://www.socrata.com/blog/lives-helps-prevent-food-poisoning-in-restaurants>, accessed 18/05/15

64 As above.

65 Code for America (2013), Foodies and Open Data Enthusiasts Rejoice, <http://www.codeforamerica.org/blog/2013/01/17/foodies-and-open-data-enthusiasts-rejoice>, accessed 18/05/15

4. How ready are agriculture and nutrition for widespread open data innovation?

“I look forward to a time when stakeholders will be able to have easy, unbridled access to data on agricultural practice, market realities and attendant issues, thus making agriculture the sector of choice.”

Richard-Mark Mbaram, Agro-Nigeria

As we have shown, there are a number of emerging examples of open data making a difference in the agriculture and food sectors. However, there is still a lot of untapped potential. There are a number of reasons to be optimistic that **agriculture and nutrition are well-positioned to use more open data.**

First, a **rich data infrastructure is emerging.** Data collection tools (such as sensors, satellites and storage) cover multiple parts of the global food chain – from production through to markets and consumption. Sophisticated analytical tools, algorithms and increasingly accessible and affordable technology put information in the hands of people who can make decisions. This includes smallholder farmers with access to mobile phones and consumers having access to nutritional information before the food reaches their table.

Second, this is an area where **innovation is needed by public and private sectors working together to develop new solutions.** In most geographic regions of the world fewer people will be living in rural areas and even fewer will be farmers. They will need new technologies to grow more from less land, with fewer hands.⁶⁶

In this environment, there is **wide scope for locally contextualised solutions** to find a market. For example, there is increasing demand from smallholder farmers to purchase tailored market information services built on open data to inform their business decisions. iShamba⁶⁷ is one of a growing number of mobile phone services providing farmers in the developing world with agricultural advice on various topics such as how to increase maize yield. It focuses on localised weather forecasts and market prices; both of them aggregated through open data sources.

Lastly, there is a **growing sense of political will to address the sustainability of agriculture and nutrition** through open data and open access within the international policy, research and development and private sector communities. At the 2012 G8 summit, world leaders committed

⁶⁶ Expert Meeting on “How to Feed the World in 2050” (2009), Concept Note & Programme, <http://www.oecd.org/tad/agricultural-policies/43109054.pdf>, accessed 18/05/15

⁶⁷ Shamba Shape-Up (2015), iShamba: <http://www.shambashapeup.com/ishamba>, accessed 18/05/15

to “share relevant agricultural data available from G8 countries with African partners”⁶⁸ and to convene the April 2013 G8 International Conference on Open Data for Agriculture⁶⁹ which led to the establishment of GODAN.⁷⁰

Meanwhile **major funding bodies of agri-food and nutrition research are making open access mandatory**, requiring research outcomes and research data produced through their funding to be made publicly available. These include the European Commission,⁷¹ the US Department of Agriculture (USDA),⁷² the CGIAR Fund,⁷³ the UK Department for International Development (DFID),⁷⁴ the Bill and Melinda Gates Foundation⁷⁵ and the Ford Foundation.⁷⁶

Individual businesses are also recognising potential efficiency and productivity gains and voluntarily embarking on their own open data initiatives. In 2015, Syngenta, one of the world’s largest agribusinesses, published agricultural sustainability data as part of its Good Growth Plan.⁷⁷

Nonetheless, there are still **challenges to realising open data’s full potential**.⁷⁸ A recent stakeholder survey for GODAN indicated that although the amount of data openly available is constantly increasing, there are still issues related to data management, licensing, interoperability and exploitation, as well as capacity to implement open data initiatives. More than this, there remains vast, untapped potential for innovation in the sector using open data. There are practical steps everyone can take – as individuals and as communities – to realise this.

68 G-8 International Conference on Open Data for Agriculture, <https://sites.google.com/site/g8opendataconference/home>, accessed 18/05/15

69 See above, <https://sites.google.com/site/g8opendataconference/home>, accessed 18/05/15

70 Global Open Data for Agriculture and Nutrition (GODAN)(2015) Statement of Purpose <http://www.godan.info/about/statement-of-purpose>, accessed 18/05/15

71 European Commission, Research & Innovation (2015) Open Science (Open Access) Policy: <http://ec.europa.eu/research/swafs/index.cfm?pg=policy&lib=science>, accessed 18/05/15

72 The White House, Office of Science & Technology Policy (2013) Expanding Public Access to the Results of Federally Funded Research: <http://www.whitehouse.gov/blog/2013/02/22/expanding-public-access-results-federally-funded-research>, accessed 18/05/05

73 CGIAR (2015), CGIAR is committed to Open Access and Open Data, <http://www.cgiar.org/resources/open>, accessed 18/05/2015

74 Department for International Development (2014) Open Data Strategy http://data.gov.uk/sites/default/files/DFID%20Open%20Data%20Strategy_10.pdf, accessed 18/05/15

75 Bill & Melinda Gates Foundation (2014), Open Access Policy, <http://www.gatesfoundation.org/how-we-work/general-information/open-access-policy>, accessed 18/05/15

76 Ford Foundation (2015), Ford Foundation Expands Creative Commons Licensing for all Grant-funded projects, <http://www.fordfoundation.org/newsroom/news-from-ford/934>, accessed 15/05/15

77 See <http://www.syngenta.com/global/corporate/en/goodgrowthplan>, accessed 12/05/15

78 See the Overseas Development Studies report for examples of the challenges http://www.developmentprogress.org/sites/developmentprogress.org/files/case-study-report/data_revolution_-_finding_the_missing_millions_-_final_20_april.pdf

5. Realising the full potential of open data: next steps

“Simply making data available is only one half of the open data equation; the other is interoperability – making our data machine readable so it can be mixed with others’ datasets to produce new information and insights.”

Krysta Harden, U.S. Department of Agriculture

“Our challenge is the opening of access to this data not just through coding but through awareness and use.”

Chris Addison, CTA

Open data is helping, and can continue to help, the agriculture and nutrition sector meet the challenge of sustainably feeding the world in the context of population growth, climate change and volatile markets. Open data is still a relatively new concept, and there remains a vast, untapped potential for using it to innovate. Many open data initiatives have been established to help unlock this potential, with many taking the publication of data and development of data standards and capacity as their starting points.

However, as the use cases in this report show, the best outcomes are realised when an open data approach is used to enable, enhance and unlock the wider potential through focused collaboration. These use cases were mostly developed as solutions to specific problems, such as drought, pest infection or food safety concerns. When it comes to realising the full potential for open data, it is important to **start with real world problems**, and to engage with problem owners, and not simply focus on the data.

Open data strategies should be developed with both the common good, and specific goals in mind: building projects that bring together a range of actors to work on problems, using and driving the improved publication of open data. Effective strategies will enable communities to work through some of the technical and social steps to realising the potential of open data at a manageable scale. These challenges are best addressed at the level of a particular problem, where standards can be identified or developed, and data released as part of solving a problem. This is especially true when advocates can point to a clear theory of change.

To help achieve this, we have set out **five steps for pursuing solution-focused open data initiatives for agriculture and nutrition**:

1. Engage with the growing open data community, including key problem owners and experts at GODAN, to **identify the challenges** that open data can help solve, and commit to work together on open data projects that address them.

2. Build open data strategies and projects with a focus on **finding solutions** to agriculture and nutrition problems. At an organisational level, identify the datasets most likely to be used by others to address pressing issues. At a collaborative level, build projects with partners in order to demonstrate impact. An open data approach enables you to build diverse partnerships, including data owners, problem owners, innovators, intermediaries, open data experts and crucially, end-users, from smallholder farmers to consumers, who stand to benefit from change.

3. Develop the infrastructure, assets and capacity for open data in your organisation and networks. At the organisational level, publish your own data, prioritising datasets identified through engagement, and build capacity of your team to publish, use open standards and apply open licenses. Collaborating on projects with partners will highlight gaps in datasets, standards and capacity. It will also present opportunities to take action to fill these gaps, by advocating for data from others to be published and standards developed and implemented, for example. The GODAN secretariat is mandated to advocate at the global level and can support this work with partners.

4. Use open data and support users of relevant data. Advocate for the data you use from others to be provided as part of the commons of open data. Increased use contributes to increased data quality – as feedback loops lead to data being made more relevant, or links being made between datasets that enhance it. Think also about how you can support use of your own data, and listen to feedback from users. Supporting use may involve thinking about the creation of intermediary tools, or investing in support for capacity building of users at different stages of the open data value chain. Consider where you have a role to play building the capacity for effective use of open data.

5. Learn through ongoing evaluation, reflection and sharing to ensure we can all continue to improve our practice. GODAN provides a forum for shared learning, through creating case studies, mapping partner activity, and bringing partners together.

This paper has outlined the growing use of open data to meet agriculture and nutrition challenges, and the potential that remains. This potential will be realised more quickly if we focus the starting point of our efforts away from the data and instead towards the problems faced in the sector. There are practical steps that can help us realise the potential of open data, as individuals as well as part of collaborative action – as part of the GODAN community.

About GODAN

The Global Open Data for Agriculture and Nutrition (GODAN) initiative supports global efforts to make agricultural and nutritionally relevant data available, accessible, and usable for unrestricted use worldwide. The initiative focuses on building high-level policy and public and private institutional support for open data. The initiative encourages collaboration and cooperation among existing agriculture and open data activities, without duplication, and brings together all stakeholders to solve long-standing global problems.⁷⁹ It aims to achieve that by facilitating access to agri-food & nutrition data that may unlock the agricultural productivity; and by increasing agricultural productivity and promoting good nutrition the global issues of hunger and poverty will be addressed.

The GODAN initiative is a growing network of 125 organisations that advocate for global open data in agriculture and nutrition. GODAN encourages:

- **new and existing open data initiatives** with a core focus on agriculture and nutrition
- agreement on and **release of a common set of agricultural and nutrition data**
- **collaborative efforts on future agriculture and nutrition open data programmes**, good practices, and lessons learned that enable the use of open data, particularly by and for the rural and urban poor

Any organisation or individual is invited to participate in GODAN and can take part in its activities and events. More information can be found at www.godan.info.

About the ODI

The Open Data Institute (ODI) is an independent, non-profit and non-partisan company based in London, UK. The ODI convenes world-class experts from industry, government and academia to collaborate, incubate, nurture and explore new ideas to promote innovation with open data. It was founded by Sir Tim Berners-Lee and Professor Sir Nigel Shadbolt, and offers training, membership, research and strategic advice for organisations around the world looking to explore the possibilities of open data.

⁷⁹ See <http://www.godan.info/about/statement-of-purpose>

Glossary: Key data concepts

- **Open data** - data anyone can access, use and share. It must: be accessible, which usually means published on the web; be available in a machine-readable format; and have a licence that permits anyone to access, use and share it.
- **Closed data** - data that only the data owners or people within an organisation can access. It might be personal or commercially sensitive information, or information for national security.
- **Shared data** - data that is shared with specific people or organisations for a specific purpose: to provide services, connect information and contribute to research. Data about consumer shopping habits, or electoral registries, are two examples of data that is shared with specific groups of people. Publishing information about how data is shared and with whom as open data can be an important transparency mechanism.
- **Personal data** - personally identifiable information. Some personal data can be open data: for example, MP salaries, information about company directors and personal insolvency records. Open data can help people understand how their personal data is being accessed and used.
- **Big data** - a vendor-driven term used to describe large quantities of rapidly changing data being collected from various sources that is used for (statistical) analysis. Big data can contain open data, and can be released as open data if licensed for anyone to access, use and share.
- **Linked data** - a way of publishing data so that it can be easily linked with other data and automatically read by computers.
- **Open government** - requiring the publication of open government data enables greater civic participation in government affairs through transparency, accountability and providing a platform for engagement.
- **Open access** - within the research community, outputs such as raw research data can be released as open data as part of an organisation's 'open access' policy.
- **Open source** - within the software community, "open source" denotes software made available under a licence that enables anyone to access, modify and use its underlying code.

Appendix 1: Useful references and tools

Links available at bit.ly/godan-resources

Tools for publishing your data

- ODI Certificates
- ODI Pathway
- ODI Guide:
 - Engaging with re-users
 - Publisher's guide to open data licensing
 - How to make a business case for open data
 - How to prioritise open data to drive global development
- Software tools for collecting, cleaning, scraping and publishing data
 - Dataverse data repository software
 - CKAN open-source data portal platform
 - DKAN Drupal-based open data platform
 - Zenodo open digital repository for publications & data
 - Fraunhofer Open Data Platform (ODP)
 - Socrata Open Data Portal
 - Open Data Kit
 - Biomart data software and services portfolio
 - CSV Lint
 - OpenRefine
 - DM2E Tools
 - Metadata Interoperability Services (MINT)
 - OpenDataSoft Open Data Platform opendatasoft.com
 - DataPress Open Data Platform
 - Scraperwiki PDF Tables
 - Import.io
 - Zizo
 - Sibdo Desktop 5.0
 - Datactics FlowDesigner DQM
 - DigIn
 - ArcGIS Online esri
 - Redlink
 - PublishMyData Swirrl
 - Epimorphics
 - Elda

- EUDAT Research Collaborative Data Infrastructure (CDI)
- re3data.org global registry of research data repositories
- DataDryad digital data repository
- agINFRA open access recommendations for the agri-food research community
- the CIARD RING for registering open data sets and make them accessible

Tools for becoming an open data user and creating open data projects

- ODI Guide: 5 ways to be a better open data re-user
- School of Data
- Open Development Toolkit
- Open Knowledge Global Network
- Open Knowledge Working Groups
- Open Knowledge Network Projects
-

Examples of organisations with open access and data policies:

- The European Commission
- The US Department of Agriculture (USDA)
- CGIAR
- The World Bank
- The World Health Organisation
- The UK Department for International Development (DFID)
- The Bill and Melinda Gates Foundation
- The Ford Foundation
- CABI
- Overseas Development Institute (ODI)
- Association of European Research Libraries (LIBER)

