CTA
Working Paper
15/01

Open Data and Smallholder Food and Nutritional Security

Andre Jellema,
Wouter Meijninger
and Chris Addison

CTA and Alterra are members of GODAN (the Global Open Data for Agriculture and Nutrition initiative)

CTA Working Paper 15/01 | February 2015

Open Data and Smallholder Food and Nutritional Security

Andre Jellema,[1] Wouter Meijninger[1] and Chris Addison[2]

1. Alterra, P.O. Box 47, 6700 AA Wageningen, The Netherlands

2. Technical Centre for Agricultural and Rural Cooperation (CTA), PO Box 380, 6700 AJ Wageningen, The Netherlands

CTA and Alterra are members of GODAN (the Global Open Data for Agriculture and Nutrition initiative)

About CTA

The Technical Centre for Agricultural and Rural Cooperation (CTA) is a joint international institution of the African, Caribbean and Pacific (ACP) Group of States and the European Union (EU). Its mission is to advance food and nutritional security, increase prosperity and encourage sound natural resource management in ACP countries. It provides access to information and knowledge, facilitates policy dialogue and strengthens the capacity of agricultural and rural development institutions and communities.

CTA operates under the framework of the Cotonou Agreement and is funded by the EU.

For more information on CTA visit, www.cta.int

About Alterra

Alterra is the research institute for our green living environment. We offer a combination of practical and scientific research in a multitude of disciplines related to the green world around us and the sustainable use of our living environment: Flora and fauna, soil, water, the environment, geo-information and remote sensing, landscape and spatial planning, man and society. These are just a few of the numerous aspects of our green environment that Alterra focuses on.

Alterra is part of the Wageningen University and Research Centre (Wageningen UR). In research and education we closely co-operate with the school of Environmental Sciences from Wageningen University. With this partner we contribute to a high quality and sustainable green living environment. The exchange of expertise and capacity and the match between fundamental and practical research in various projects give us a scientific advantage.

This work is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License. This license applies only to the text portion of this publication.

Table of contents

Table of contents v

Key data sets vii

Government vii

Space and meteorological agencies vii

International organisations and partnerships vii

NGOs vii

Business sector vii

Executive summary ix

Introduction 1

The GODAN initiative invites all stakeholders to open up their agricultural and nutritionally relevant data to enforce global food and nutrition security 1

In modern agriculture data are becoming more and more an important resource for food production, facilitation of the value chain and governance 1

CTA wants to know the impact of the open data on the smallholder food and nutrition security 2

Open data and their impact on developing countries 3

Open data are data that can be freely used, reused and redistributed by anyone – subject only, at most, to the requirement to attribute and share alike 3

Open data fuel the pyramid of wisdom enabling better decision-making 3

Intermediaries between suppliers and consumers are key in the functioning of open data community 4

Although the potential value of open data generally is estimated to be high, the actual measured impact of open data in (developing) countries is low 5

The potential impact of open data on the smallholder ecosystem 7

The availability of open data can catalyse the functioning of the smallholder ecosystem by providing each of the actors with relevant information about the ecosystem, its actors and its functioning 7

Mobile operators and ICT service providers connecting smallholder farmers play a key role in achieving impact of open data 8

The potential impact of open data on the smallholder food and nutrition security 9

Open data and farmers' rights 10

Overview of the different sources of open data for food and nutrition security and their impact on the smallholder ecosystems 10

Open government data 10

Which data are being made available? 11

Impact on the smallholder ecosystem 13

Impact on governance 13

Impact on the value chain and sustaining services 14

Impact on agricultural production 14

Space and meteorological agencies 14
Which data are being made available? 15
Impact on the smallholder ecosystem 16
International organisations and partnerships 18
Which data are being made available? 18
Impact on the smallholder ecosystem 18
Science 20
Agricultural data 21
Impact on the smallholder ecosystem 23
NGOs 23
Which data are being made available? 24
Impact on the smallholder ecosystem 25
Business sector 25
Which data are being made available? 26
Which data are being made available? 26
Impact on smallholder ecosystem 29
Synthesis and outlook of the impact of the open development on smallholder food and nutrition security .. 30
The current impact of open data on smallholder food and nutrition security is low 30
Potentially there is a large impact of open data on smallholder food and nutrition security 31
Other options to further improve the uptake and availability of open data for smallholder food and nutrition security 32
There is a trade-off between the aggregation level of data, the amount of information it contains and farmers' rights 32
References 34

Key data sets

Government

Space and meteorological agencies

- United States Geological Survey (USGS) Landsat archive (http://earthexplorer.usgs.gov/; http://glovis.usgs.gov/)
- NASA MODIS archives (http://modis.gsfc.nasa.gov/data/)
- ESA Copernicus Space Component Data (http://copernicusdata.esa.int/web/gsc/home)
- Google Earth Engine (https://earthengine.google.org)

International organisations and partnerships

- World Bank (http://data.worldbank.org/)
- African Development Bank Group – Open Data for Africa (www.opendataforafrica.org)
- FAO (http://faostat3.fao.org/faostat-gateway/go/to/home/E)
- UNEP (http://geodata.grid.unep.ch/)
- UN (http://data.un.org)
- WTO (http://www.wto.org/english/res_e/statis_e/looking4_e.htm#summary)
  Science:
- The Harvard Dataverse Network http://thedata.harvard.edu/dvn/)
- ICSU World Data System (https://www.icsu-wds.org/)
- CGIAR Consortium Data Management System (http://www.cgiar.org/resources/open/data-management-system/)

NGOs

- Open development data (http://www.openaidsearch.org/)
- CABI's Plantwise project (http://www.plantwise.org/)

Business sector

- Google Maps (https://www.google.nl/maps?source=tldso)
- Google Earth (https://www.google.com/earth/explore/products/)
- Google Trends (http://www.google.com/trends/)

Executive summary

This report was commissioned Technical Centre for Agricultural and Rural Cooperation (CTA) as a member of the Global Open Data for Agriculture and Nutrition (GODAN) initiative. It aims to provide a better understanding of the actual impact of the open data movement on the food and nutrition security of smallholders and highlight the areas of potential unfilled opportunity. This study was carried out by Alterra as a rapid desk-based study to identify possible relevant sources of open data and determine the current and potential impact of these sources by assessing the content of the sources and possible applications. The sources identified were governments, meteorological and space agencies, international, science-based and non-governmental organisations and businesses. The impact was determined by looking at the applicability of the data to facilitate the smallholder ecosystem, specifically looking at the potential to enhance smallholder production, the value chain, support services and governance.

The study found that, despite the potential value of open data to smallholder farmers in developing countries being high, there are few readily available examples of direct impact on food and nutrition security of smallholders. However, there is a clear indirect benefit of open data usage for smallholders, including the contribution to better governance

Most impact in other domains is from open space and meteorological data. However, even in this area open data access is still in its infancy and there are remaining issues, such as reliability of data at smallholder resolution (i.e. the spatial resolution of the satellite data with respect to the size of most agricultural fields) and lack of connectivity to provide services to the smallholder ecosystem.

Most mature open data sources are international organisations such as the World Bank and FAO. These data sources are very much oriented towards global governance and have limited impact in a local smallholder context.

Potentially, businesses, such as mobile service providers, are a promising source of information. With the rapidly increasing use of mobile phone technology there are potentially large information sources about the smallholder ecosystem within these mobile networks. Corporate data-sharing is indicated as the best way to obtain fine-grained information about the smallholder ecosystem and to provide better production advice, facilitate the value chain, provide better services and governance. However issues around privacy, farmers' rights and business confidentially remain to be solved.

Other potential areas to develop open data sets relevant to enhancing the food and nutrition security of smallholders are:

1) Developing dedicated scientific open data sets targeted at specific stakeholders needs in the smallholder ecosystem.
2) Opening NGO project monitoring and evaluation data and encouraging NGOs to collect specific data during project implementation.

	Impact on Production	Impact on Value Chain	Impact on Support Services	Impact on Governance	Pot. Impact on Production	Pot. Impact on Value Chain	Pot. Impact on Support Services	Pot. Impact on Governance
Goverment Data	0	0	0	2	0	1	2	3
Meteo and Space Data	1	0	2	2	2	0	3	3
Data from Int.Organisations	0	0	0	2	0	0	0	3
Science Data	0	0	1	2	3	3	3	3
NGO Data	0	0	0	2	3	3	3	3
Business Data	0	0	0	0	2	4	4	4

Figure 1: The impact and potential impact of open data from different sources on smallholder production, the value chain, support services and governance rated between 0 and 4 (0: no/little impact, 4: high impact).

Introduction

The GODAN initiative invites all stakeholders to open up their agricultural and nutritionally relevant data to enforce global food and nutrition security

At the G8 Conference in London in October 2013, the Global Open Data in Agriculture and Nutrition initiative (GODAN) was launched (*The Guardian*, 2013). The objective of GODAN is to build high-level policy and institutional support for open data relevant to agriculture and nutrition across the public and private sector to enforce global food and nutrition security (GODAN, 2014). Open data are: "data that can be freely used, reused and redistributed by anyone – subject only, at most, to the requirement to attribute and share alike" (The Open Foundation, 2012). Open sharing of data is considered important because data collected for a specific task may have value to other people or organisations in different contexts and/or for different reasons. The value of reusing the data may be even larger than the original value of the data, considering the original purpose it has been collected for. The potential global value is estimated US$3 trillion a year (McKinsey, 2014). By making data available, the alternative values can be harvested, although it will be impossible to predict precisely how, where, and by whom this value will be created in the future. Open data can contribute to (Opendatahandbook, 2014):

- participation and self-empowerment;
- improved or new products and services;
- new knowledge from combined data sources and patterns in large data volumes;
- improved effectiveness of government services and impact measurement;
- transparency and democratic control on public processes.

In line with other global movements for open data and open access, GODAN advocates for:

- open data and open-access policies by default, in both public and private sectors, while respecting and working to balance openness with legitimate concerns in relation to privacy, security, community rights and commercial interests;
- the release and re-usability of data in support of innovation and economic growth, improved service delivery and effective governance, and improved environmental and social outcomes.

Open access to agricultural and nutritionally relevant data is vital for innovation in agriculture and value chain development driven by farmers, farmer organisations, researchers, extension experts, policy-makers, governments and other private sector and civil society stakeholders (GODAN, 2014).

In modern agriculture data are becoming more and more an important resource for food production, facilitation of the value chain and governance

A strong example of data use is precision agriculture. Precision agriculture, also known as precision farming, satellite farming or site-specific crop management (SSCM) uses GPS (global positioning system), soil testing, yield monitors, remote sensing and variable-rate technologies, information technology and geographic information systems (GIS) and the like to observe, measure and respond to spatial variations in crops (within one field or between different fields). Precision data together with computer-based decision support systems help optimise production (yield), conserve resources (e.g. water and nutrients) and reduce costs

(Venkatalakshmi and Devi, 2014). Examples of (satellite-based) crop monitoring services are Cropio, FarmSat, FieldLook and ClimatePro (Precision Agriculture, 2014).

In dairy farming, automatic milking machines are collecting data down to an individual level; each cow can be tracked and examined, and the farmer will be alerted when there are unusual changes in the animal that might indicate illness or injury. Farmers can use these data to analyse the effect of various animal feeds on milk yield (Automatic milking, 2014).

Data are also collected at all stages of food value chains, from the farm to the consumer. Each partner is challenged to be more efficient, more sustainable and more effective. Information is added to the produce along the chain and producers and customers are increasingly looking beyond one chain-partner back or ahead. Allergy information, fair production and trade, footprints and many other quality characteristics are not only relevant to consumers but in the end, all chain partners need the information (Lundqvist *et al.,* 2012).

Governments are implementing all kinds of e-governance data services, including facilitating agricultural accountability, obtaining subsidies and participatory governance. In the Netherlands, examples can be found in the 'National Single Window for Trade and Transport' to avoid repeated data entry for business to with different government and in the national agricultural statistical survey used for monitoring, policy-making, research and agricultural subsidies (Wassenaar, 2000; RVO, 2014).

CTA wants to know the impact of the open data on the smallholder food and nutrition security

The Technical Centre for Agricultural and Rural Cooperation (CTA) is a joint international institution of the African, Caribbean and Pacific (ACP) Group of States and the European Union (EU). The mission of CTA is to advance food and nutritional security, increase prosperity and encourage sound natural resource management in the ACP countries by strengthening the agricultural policy processes; improving smallholder agricultural value chains and enhancing information, communication and knowledge management capacities for rural development in ACP countries. As a member of the GODAN initiative, CTA wants to know what the actual impact is of the open data movement on the food and nutrition security of smallholders and which opportunities remain unfulfilled. In this report we attempt to answer this question based on a quick desk study and describe the results. We do this in both a narrative form and by ranking each data source and impact field. This rank is subjectively based on our current best knowledge and not based on an analytical methodology. The rank varies from 0 to 4, where 0 means no impact on the food and nutrition security of the smallholders found, 1 means initial impact, 2 means substantial impact, 3 means a large impact and 4 means a very large impact. This ranking is not a final judgement, but a means of communicating a quick scan overview. It needs to be discussed within the wider open data community for further confirmation.

- In the first section of the report, we provide a brief explanation of open data, the open data community and the current impact of open data in developing countries.
- In the second section, we describe the smallholder ecosystem and how the actors within the smallholder ecosystem may benefit from open data development.
- In the third section, we provide an overview of the status of the different sources of open data relevant to food and nutrition security and their (potential) impact, including the ranking.
- In the last section, we provide the conclusions and outlooks for the further development of open data to enhance food and nutrition security of smallholder communities.

Open data and their impact on developing countries

Open data are data that can be freely used, reused and redistributed by anyone – subject only, at most, to the requirement to attribute and share alike

The full Open Definition (Opendefinition.org) gives details about what this means. To summarise the most important are:

- **availability and access:** The data must be available as a whole and at no more than a reasonable reproduction cost, preferably by downloading over the internet. The data must also be available in a convenient and modifiable form;
- **reuse and redistribution:** The data must be provided under terms that permit reuse and redistribution including the intermixing with other data sets;
- **universal participation:** Everyone must be able to use, reuse and redistribute - there should be no discrimination against fields of endeavour or against persons or groups. For example, 'non-commercial' restrictions that would prevent 'commercial' use, or restrictions of use for certain purposes (e.g. only in education), are not allowed (Open Knowledge Foundation, 2012).

Data can be 'open' at different levels. Tim Berners-Lee, founder of the World Wide Web, proposed a five-star model of 'openness' (Wikipedia Open Data, 2014):

* Data are online available in any format.

** Data are online available in a structured file format which is appropriate for automatic reuse (a table in Excel format rather than a JPG picture).

*** Data are online available in an open file format (CSV rather than Excel).

**** All of the above-mentioned and data formats are used like Resource Description Framework (RDF) and SPARQL, which allow others to specifically point at data objects.

***** All of the above-mentioned and links are made to other related data sets providing more contexts about the data set.

The spirit of open data development is of the same nature as other openness developments focusing on different aspects, such as open access (CIARD, 2014a; Peters, 2014) focusing on the access of scientific and other information and knowledge, and open source focusing in the open access to software codes (Open Source, 2014).

Open data fuel the pyramid of wisdom enabling better decision-making

Data and access to data is not directly useful for most actors in society. The data needs to be contextualised and combined with other data in order to produce relevant, comprehensive information and new knowledge ultimately leading to wisdom. This framework for adding value to data by combining data and adding

information is captured in the knowledge pyramid in Figure 2 of Lokers and Janssen (2014). The basic concern at the bottom of the pyramid is to have enough data available from different sources in order to be able explore and combine and understand the world better. Governments, international organisations and others who make their data available in open format are contributing to the pool of data, fuelling the basic layer of the pyramid of knowledge. However intermediaries, scientists, data analysts, modellers and IT experts are needed to take the intermediary steps from data to information and from information to knowledge. Ideally, the data at the bottom of the pyramid is annotated and linked to other data (linked open data). These links provide information about the data and their quality. Linked open data help experts to find the right data and to make the right interpretation of the data. The next step is to bring the data into a domain (e.g. food security, climate change, biodiversity loss) and to analyse the data from the domain perspective in combination with other data sources.

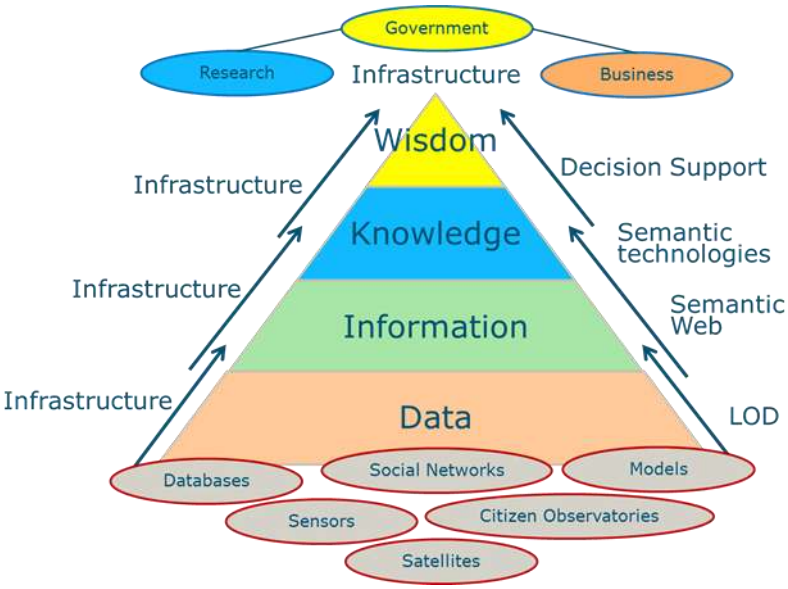

Figure 2. The pyramid of knowledge, demonstrating what is needed to get from a pool of data to knowledge and ultimately to wisdom for evidence-based decision-making.

Source: Lokers and Janssen (2014)

Intermediaries between suppliers and consumers are key in the functioning of open data community

Deloitte (2012) describes the open data community, as simultaneously simple and complex. On the one hand, the chain between the suppliers of open data and those who demand their services is short; on the other hand, almost every entity in the open data community can link to every other entity (Figure 3). The same organisations can be found in different roles, suppliers, users or intermediary in the open data chain. The intermediaries play a key role in the open data community, enabling the wider use of the open data sources.

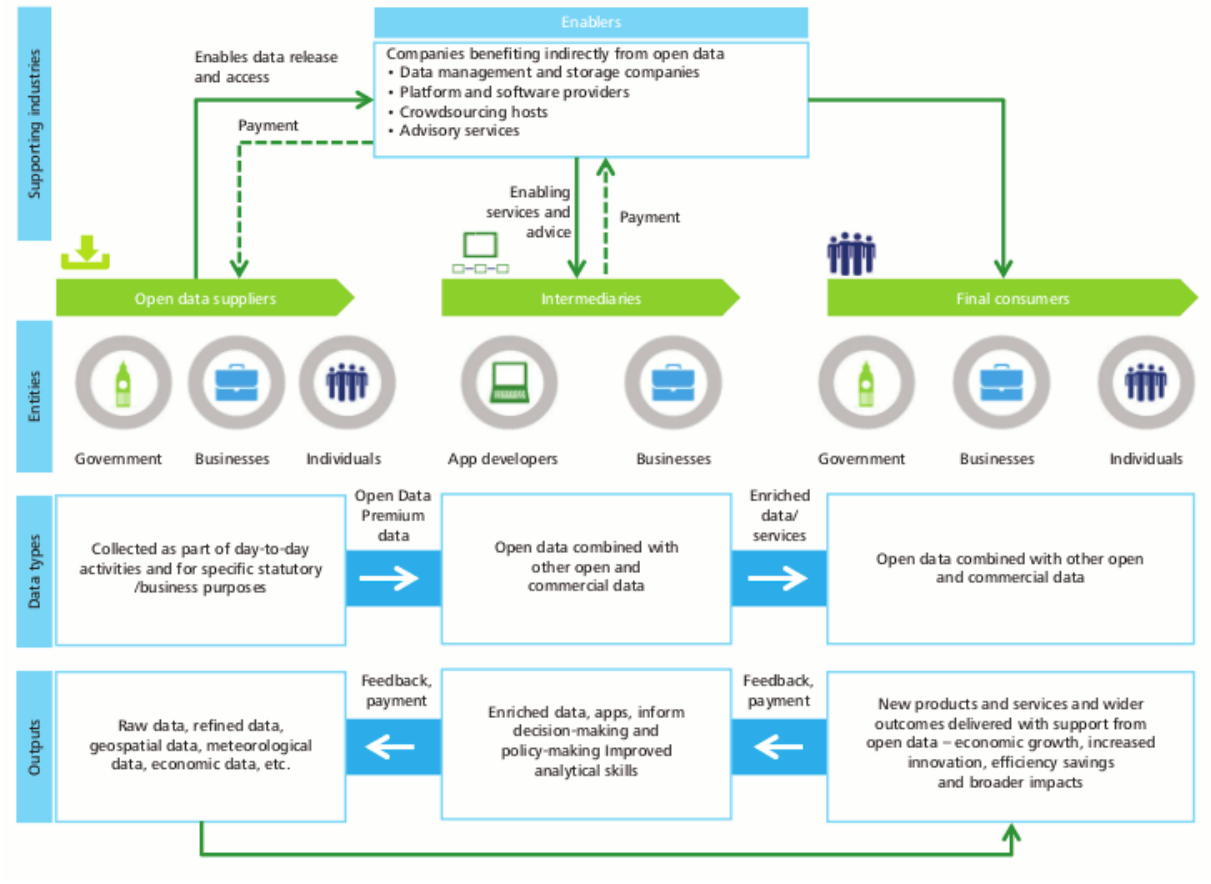

Figure 3. The open data community is simultaneously simple and complex. The chain between suppliers of open data and demanders of services is short, but almost every entity in the open data community can link to every other entity.

Source: Deloitte (2012)

Deloitte distinguish five archetypes of actors:

- Suppliers: organisations that publish their data via an open interface to allow others to use and reuse it
- Aggregators: organisations that collect aggregate open data and sometimes, other proprietary data, typically on a particular theme, find correlations, identify efficiencies or visualise complex relationships
- Developers: organisations and software entrepreneurs that design, build and sell web-based, tablet or smartphone applications for individual consumption
- Enrichers: organisations (typically larger, established businesses) that use open data to enhance their existing products and services through better insights
- Enablers: organisations that facilitate the supply or use of open data, such as the competition initiatives.

Although the potential value of open data generally is estimated to be high, the actual measured impact of open data in (developing) countries is low

In literature, the estimated value of open data for society varies from large to enormous:

- Jamaica benefitted by US$21 million in 2013 (CTA/AgriHack, 2015a)

- Ireland could benefit by EUR 126.4 million in the for geospatial information sector alone (Lee *et al.*, 2014)
- A report produced by accountancy firm Deloitte (2012) estimates the economic value of the data held by the public sector in the UK and released for use and reuse to be around £5 billion per year (UN, 2014)
- The European Commission (EC) estimates the aggregated direct and indirect economic impact from applications based on open data across the EU27 economy to be €140 billion annually (EC, 2011)
- A report from McKinsey Global Institute (McKinsey, 2013) puts the global value of better and more open data at US$3 trillion per year with most of this benefit accruing to the USA and Europe.

Whether this value actually materialises depends on the functioning of the open data community in a country. In the Open Data Barometer, Davies (2013) assesses a country's ability to secure and sustain the benefits of open data. Focusing on government data, these components are:

- **the government's capacity and commitment to open data**, addressing the political will and organisational ability of governments to both make open data available, and to secure benefits from open data, such as increased operational efficiency;
- **citizen and civil society freedoms and engagement with the open data agenda**, including the presence of strong Right to Information and Data Protection regimes, which are important for empowering citizens to hold government to account, and protecting citizens from potential abuses of open data (Davies*,* 2013);
- **resources available to entrepreneurs and businesses** to support economic reuse of open data and to catalyse intermediary actions, including internet penetration, the availability of training for businesses and government support for open-data-led innovation.

Figure 4 shows the deviation of different regions in the world from the global average in open data readiness. The African continent has the lowest open data readiness (Davies*,* 2013). This is caused by limited internet penetration and a scarcity of entrepreneurs and civic technologists who often act as key intermediaries between open data and wider use of that data. To achieve impact, a substantial focus on capacity building and sustainability of intermediaries is required, as well as an exploration of different approaches to making data accessible that do not rely on internet penetration, such as through print media, community radio and mobile phones (Davies*,* 2013). In the Barometer, the Caribbean and Pacific regions are included in larger regions, Americas and Asia, respectively, and therefore no specific conclusions can be drawn for these regions. Recently CTA has carried out an open-data readiness assessment for open data in the Caribbean (CTA/AgriHack, 2015b). This research indicates a rising Caribbean tech ecosystem and interest in open data. However, few of the developers interviewed were using open data and developing agricultural apps was seen as challenging. Lack of domain knowledge was given as main reason. Also, the need for capacity building and sustainability of intermediaries and data availability was mentioned.

Overall, there is a large gap in terms of access and uptake of ICT between the 'advanced economies' and the rest of the world as also demonstrated in the UN report *A World that Counts* (UN, 2014). As a result, open data development will currently mainly impact 'advanced economies'.

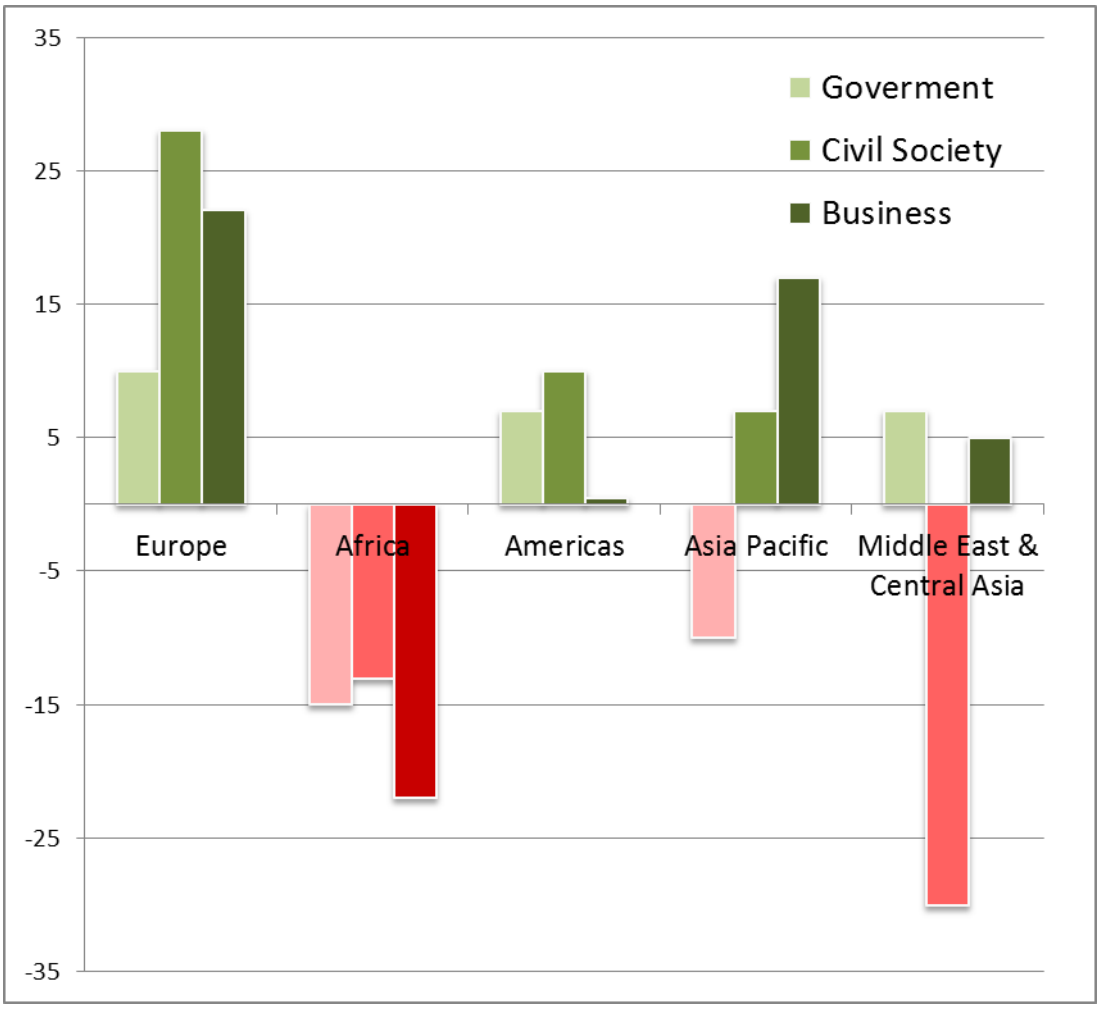

Figure 4. The difference in open data readiness in the regions of the world compared to the global average. The index for open data readiness ranks from 0 to 100. The global average scores for government, civil society and business are 50, 50 and 40.

Source: Davies (2014)

The potential impact of open data on the smallholder ecosystem

The availability of open data can catalyse the functioning of the smallholder ecosystem by providing each of the actors with relevant information about the ecosystem, its actors and its functioning

The smallholder ecosystem consists of many value chain actors including smallholders, cooperatives, input providers, traders, processors, exporters and wholesalers and global businesses (Figure 5). Around the value chain there are different service providers: financial services (credit, insurance), logistic services (transport, storage, grading, certification) and the extension services (farm management advice, business advice) facilitating the value chain. There are 'governance' actors such as the local, regional and national governments,

donors, NGOs and researchers. Each of these actors is interested in information about the production, the functioning of the value chain and the availability of services and governance. The more information that is available about the ecosystem and its functioning, the better the different actors can fulfil their role, ultimately strengthening the food and nutrition security of the smallholder farmers. Not all actors are interested in the same information or look at the same information in the same way. A smallholder wants to have the information needed for decision-making in the context of his or her farm: What crops should I grow? How do I grow these crops? Where do I store my harvest? Which inputs do I use and where can I get them? Where do I sell my crops and at which price? A government may be more interested in the general picture of an area so it can adapt its policies accordingly. Ease of access to information in the ecosystem will ensure smooth functioning of the smallholder ecosystem. Sharing open data makes the information available in a transparent and efficient way.

Figure 5. The smallholder ecosystem and its actors, including the upcoming mAgri services.

Source: Freely adapted from Fonzi and Chau (2012)

Mobile operators and ICT service providers connecting smallholder farmers play a key role in achieving impact of open data

Mobile operators and ICT service providers are the most recent emergent actors in the farmer ecosystem. Sharing of data or information exchange between smallholders or between smallholders and other actors is traditionally very difficult in rural areas of developing countries because of large distances, bad roads and sparse, weak landlines for telecommunication. However, this picture has been changing in the last 10 years. Mobile operators are penetrating the rural areas with their networks and the prices of mobile devices are falling, resulting in more smallholder farmers who are connected to the outside world (World Bank, 2011). The mobile operators and ICT service operators are developing and hosting agricultural advice services on these mobile platforms, providing information based on text messages, structured menus, voice messages etc. In some cases, these services are developed as part of the rural marketing strategy of mobile providers. The GSM association has current identified 122 deployed services worldwide (GSMA, 2014). Mobile agricultural advice services vary from: services enhancing productivity on the farm; services facilitating farmers' access to microcredit and insurance; services helping cooperatives organise their stock and trade, services allowing inclusion of farmers and cooperatives in agribusiness supply chains or providing access to the global market (World Bank, 2011). Smallholders are getting connected to the global economy and agricultural knowledge

base via the mobile network. On top of this, development information is not only flowing from the global community to the smallholders, but the farmers are also providing information about themselves and their environment through these mAgri services. Opening up these data streams will provide the opportunity to better understand the smallholder world and therefore the world at large, enabling better progress in terms of development and governance and achievement of the Millennium Development Goals (UN, 2014).

The potential impact of open data on the smallholder food and nutrition security

The potential impact of open data on the food and nutrition security of smallholder farmers and the smallholder ecosystem is manifold. Table 1 provides an overview.

Table 1: Overview of potential impact of open data on smallholder food and nutritional security.

	General	Smallholder perspective
Impact on governance	• Improved effectiveness of government services and impact measurement • Better targeted development programmes • Transparency and democratic control • Better contextualised science	• Participation and self-empowerment
Impact on services	• Improved or new products and services • More clients	• Better access to logistic, extension, financial, input, trade services
Impact on the value chain	• Improved traceability and quality standards for buyers • More efficient value chain	• Better access to the (global) markets • Better price negotiations • Better functioning cooperatives
Impact on production	• More stable supply • Continuous market	• Higher yields • Less perishing yields • Higher availability of inputs • Better pest control

Currently the main sources of open data are:

- governments
- government agencies such as space and meteorological agencies
- science
- NGOs
- business

In the next chapter, each of these sources will be explored and the current and potential impact for the smallholder food and nutrition security will be discussed. Relevant applications or potential applications are presented. Each of the data sources will be ranked for its current and potential impact on smallholder food and nutrition security. The rank is subjectively based on our current best knowledge and not based on an analytical methodology. The rank varies from 0 to 4, where 0 means no impact on the food and nutrition security of

smallholders, 1 means initial impact, 2 means substantial impact, 3 means large impact and 4 means very large impact. This ranking is not a final judgement, but a means of communicating a quick scan overview. It needs to be discussed with the wider open data community for further confirmation.

Open data and farmers' rights

Although open data has a large potential for positive impact on smallholder food and nutrition security, this does not mean that all data should be automatically open. A number of issues must be considered (Maru, 2014):

- Smallholders should benefit from the data they provide. Open data about the smallholder ecosystem should be made accessible to the smallholders in a timely, fair and equitable manner; they should be affordable, relevant, useful and trustworthy for famers to effectively use them. To realise this, smallholders should be included in processes related to the decision on which data and information they want to generate, share and exchange, according to their needs and preferences.
- Open data about smallholders may create or increase the inequality between smallholders and other actors in the smallholder ecosystem. Smallholders may lack the capacity or the technical means to benefit from the information provided. Therefore, open data development may also imply capacity building, technical enablement and the implementation of legislation.
- Open data should not violate the privacy of smallholders. Data may contain elements that are sensitive from a business, political, social, religious or traditional perspective and should not be spread automatically or only in such a way that these issues are dealt with.

Overview of the different sources of open data for food and nutrition security and their impact on the smallholder ecosystems

Open government data

The number of countries with open data programmes has grown rapidly over the last few years. As at mid-2014, there are at least 50 national governments running open data portals and initiating OGD initiatives (Davies, 2014b). One of the organisations catalysing this development is the Open Government Partnership (OGP, 2014) providing an international platform for domestic reformers committed to making their governments more open, accountable, and responsive to citizens. Although the OGP declaration does not explicitly mention open data, many governments made commitments to open data development as a result (Davies, 2014b). As can be seen in Figure 6, participation in the OGP is not evenly distributed across the globe. The number of APC countries participating in the OGP is limited when compared to the Americas, Europe and Australia. African countries who are participating are: Ghana, Kenya, Liberia, Malawi, Sierra Leone and Tanzania. In the Caribbean, the Dominican Republic and Trinidad and Tobago are participating. In the Pacific no partner countries are participating (OGP, 2014).

Of the 16 member-countries of the Caribbean Community (CARICOM), seven had enacted freedom-of-information laws, four had drafted such legislation and two had guaranteed freedom of information as a constitutional right in 2011 (CTA/AgriHack, 2015).

Figure 6. Participating countries in the Open Government Partnership.

Source: OGP (2014)

Which data are being made available?

A description of different data sources generally provided by national governments as identified by the Open Data Barometer study (Davies, 2013) is presented in Table 2.

Table 2: A description of different data sources generally provided by national governments.

Innovation cluster	Social policy cluster	Accountability cluster
Data commonly used in open data applications by entrepreneurs or with significant value to business.	Data useful in planning, delivering and critiquing social policies and with the potential to support greater inclusion and empowerment.	Data central to holding governments and corporations to account.
• Map data • Public transport timetables • Crime statistics • International trade data	• Health sector performance • Primary or secondary education Performance data • National environment statistics • Detailed census data • Land ownership data	• Legislation • National election results • Detailed government budget • Detailed government spend • Company register

Source: Davies (2013)

According to Davies (2013) categories of data managed by statistical authorities are most often accessible online (Figure 7), but are often only released in aggregated forms with unclear or restrictive licences. National budgets are available more often than the spending data and when available, spending data are often published in very aggregated forms. Land and company registration data are least likely to be openly available, reflecting both the absence of coherent land and company registry data sets in a number of countries and a low priority placed by many OGD initiatives on making these data sets available.

In developing countries, much government information is still managed on paper at local offices and is not digitised. Data sets are seldom clearly open licensed and there is poor understanding of what open licences entail. There is a frequent mismatch between open data supply and demand in developing countries; politically

sensitive data sets are among the least likely to be published; key data sets such as company registers, digital maps and land registration databases are not held in digital format (Davies, 2014a).

Many data sets are of low quality, which hinders their usage and limits their value. Data may be aggregated, outdated data sets or poorly structured data. Also, the navigation through data sets and limited information about the data sets may hinder uptake and usability (Mutuku and Mahihu, 2014).

Less than 7% of the data sets surveyed in the Open Data Barometer study were published both in bulk machine-readable forms and under open licences. This makes it unnecessarily difficult for users to access, process and work with government data, and potential entrepreneurs face significant legal uncertainty over their rights to build businesses on top of government data sets. (Davies, 2013). In the second edition of the Barometer (Opendatabarometer, 2015), the general outlook on development of government open data has changed very limited. The total amount of government open data has grown only 3%. In the survey of 2014, 31 countries have at least one open data set, but only 50% of the data sets surveyed among the 11 top-ranked countries qualified as fully open.

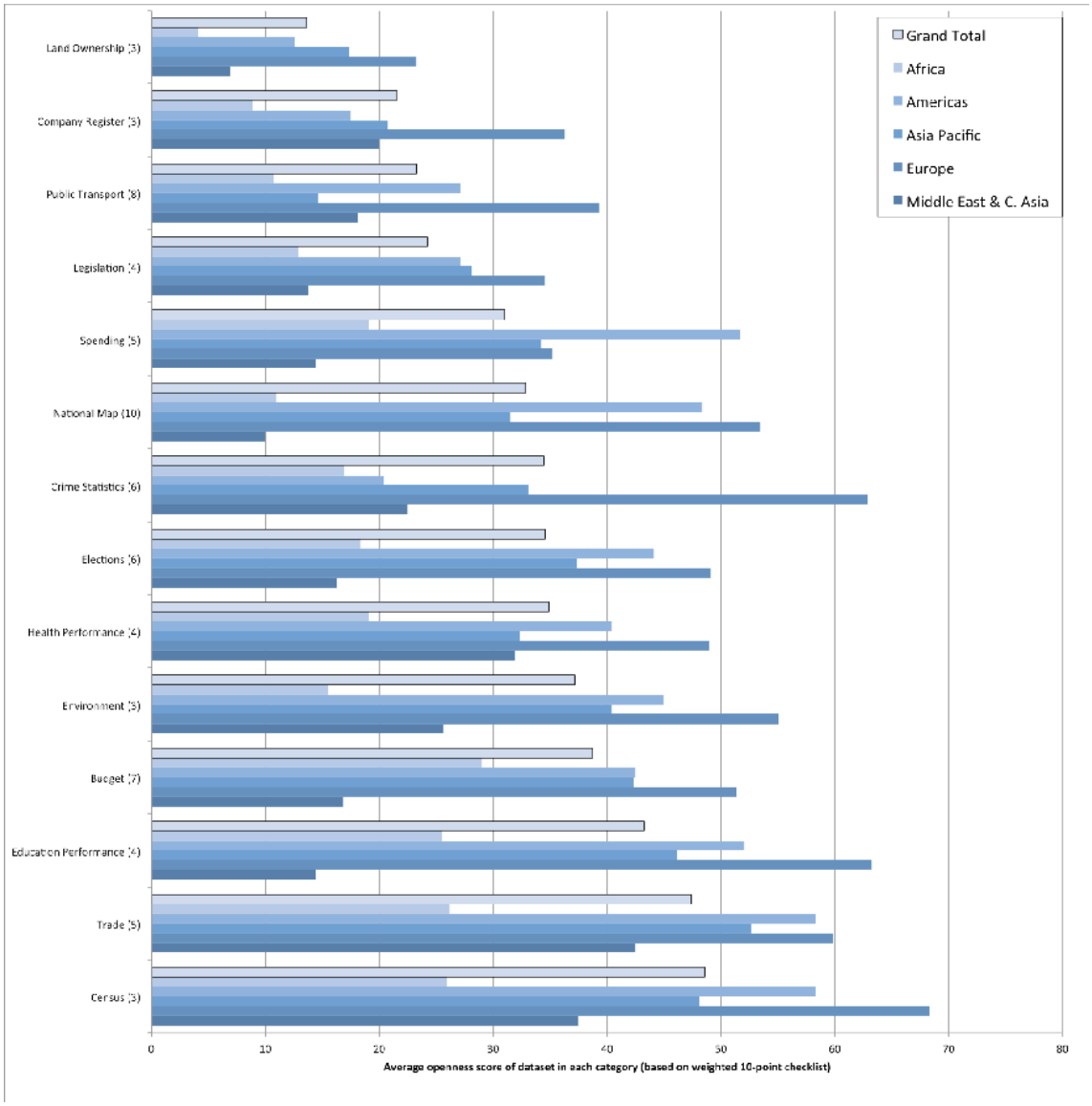

Figure 7. Average openness score of the data sets.

Source: Davies (2013)

Impact on the smallholder ecosystem

Evidence on the impact of open government data is almost universally lacking. Few open government data programmes have yet been evaluated and the majority of discussions of impacts are based on anecdote (Davies, 2013). The Open Data Barometer study asked about six kinds of open government data impact (government efficiency, transparency and accountability, environmental sustainability, inclusion of marginalised groups, economic growth, and supporting entrepreneurs). In countries with some form of open government data policy no examples of impact could be found in 45% of the impact questions and on average evidence of impact was scored at just 1.7 out of 10 (Figure 8, Davies, 2013).

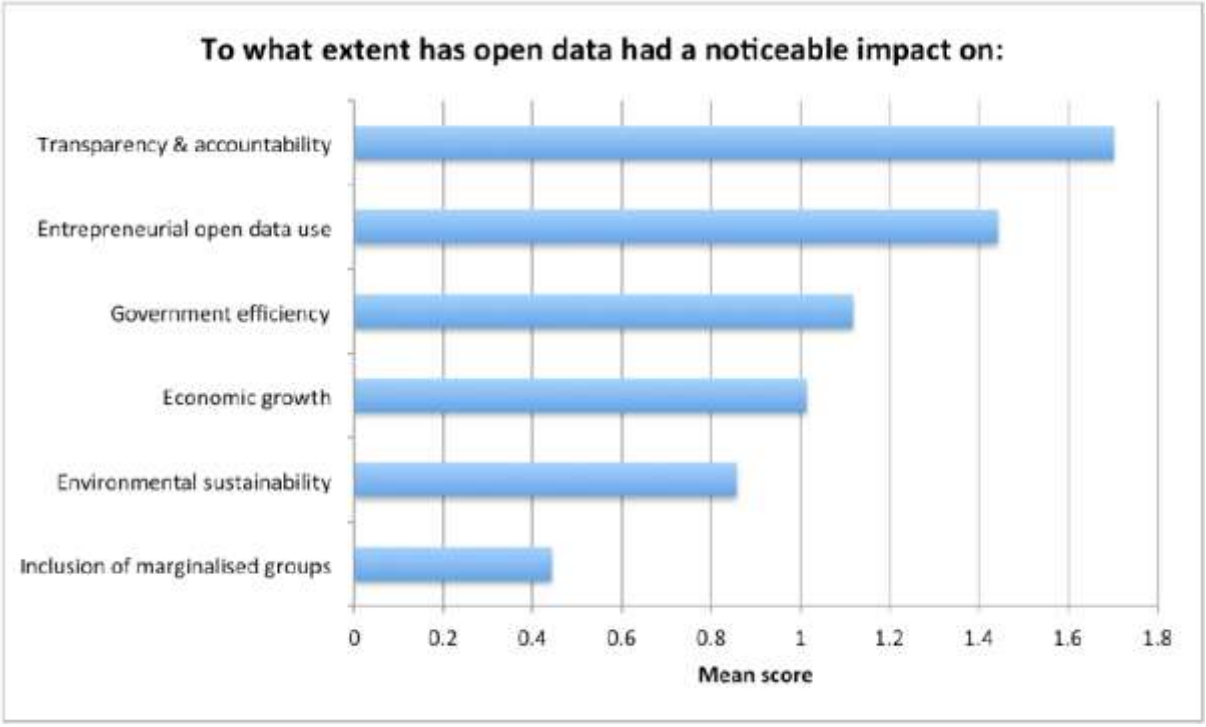

Figure 8. Average impact score across all countries based on an expert survey.

Source: Davies (2013)

Impact on governance

In general, the smallholder will benefit, like all citizens, from a transparent government. Data about government budgets, government spending and the performance public institutes such as schools and health clinics will contribute to this transparency.

An example can be found in the UN report on the data revolution for development:

> *In Mexico, a budget research and advocacy group called Fundar developed an online database of government farm subsidies. One of the problems brought to light was the way in which billions of dollars of the funds were distributed. Though many farm subsidy programs claim to target the neediest farmers, the database revealed that a small group of wealthy farmers had captured the vast majority of subsidy funds over time (the top 10% of recipients had received over 50% of the funds). The studies contributed to the government decision to review and change the distribution of the subsidies. — (UN, 2014)*

The impact of open government data on better governance starts emerging, but considering all improvements that still can be made we ranked the current impact 1, with a potential impact of 3 (on a scale of 0 to 4).

Impact on the value chain and sustaining services

Potentially, the value chain and sustaining services can also benefit from government data as presented above (Davies, 2013). In particular, the following data would be relevant at a regional or local scale. This is also one of the recommendations in the second edition of the Open Data Barometer (Opendatabarometer, 2015):

- ownership and legal status of companies; input suppliers, traders, financial advisers;
- land ownership;
- (regional) trade statistics and prices;
- infrastructure, transport.

The impact of open government data on better sustaining service and the value chain is nil, but relevant information could be made available; therefore, we rank the current impact as 0 with a potential impact of 2 (on a scale 0 to 4). We expect the impact to be less than on the governance itself.

Impact on agricultural production

The data as presented by Davies (2013) will have limited impact on agricultural production. The open data portal of the US Government gives some examples of data that can impact agricultural production.

- The plant Hardiness Zone Map. This is the standard by which gardeners and growers can determine which plants are most likely to thrive at a particular location (USAID, 2014).
- SoilWeb: An online soil survey browser, providing access to soil survey data (CSRL, 2014).

The impact of open government data related to governance on actual production in developing countries is ranked as 0. A government can decide to generate specific data sets like the American examples above, but to us this could also be seen as a scientific output hosted by the government. There is a boundary issue here. Therefore, we rank the current impact as 0, with a potential impact of 1. Governments may subsidise specific data sets dedicated to production.

Space and meteorological agencies

In 1991, NASA adopted the Earth Science Data Policy to promote full and open sharing of all data with the research and applications communities, private industry, academia and the general public. NASA was the first agency in the United States and the first space agency in the world to provide full and open access in a timely manner and at no cost. Gradually, other US agencies and international space agencies have adopted similar open-access policies and practices.

- 2008: The USGS decided to change their data policy of the Landsat programme, meaning that all Landsat data (since 1972) is freely available to any user. As a result, the distribution of Landsat satellite images increased dramatically (~25,000 in 2001 to more than 2.5 million in 2010) and they are used in a wide range of disciplines, including agriculture and nutrition (e.g. to monitor crop water use, crop growth and crop yield) (Wulder *et al.,* 2012).
- 2010: This enabled Google, in partnership with NASA, USGS and others to launch Google Earth Engine (Google, 2014a). Google Engine is a cloud-computing platform for processing satellite imagery and other observation data and facilitates the usage of satellite imagery by non-expert scientists. Currently, the platform contains over 40 years of Landsat data, taken from the USGS Landsat archive and MODIS data from NASA. Not only does it provide easier access to a large warehouse of 'ready-to-use' satellite imagery

(compared to the Landsat and MODIS archives of the USGS and NASA) but it also provides unprecedented computational power for the individual scientist. The latter is a major relief in big-data research and processing. Researchers can log on (after registration), access all the data and run their own algorithms.

- 2013: The ESA (2013) announced the European Delegated Act on Copernicus on data and information policy (together with EUMETSAT). This act provides free, full and open access to users of environmental data from the Copernicus programme, including data from the Sentinel satellites (the first of a total of six satellites was launched in April 2014).

Open-access policy also takes place, though slowly, in the national meteorological and hydrological services worldwide. In 1995, the WMO (World Meteorological Organisation) committed itself to broadening and enhancing the free and unrestricted exchange of meteorological and related data and products (Resolution 40). The WMO has issued many requests to Member States to provide their data to international data centres so that the data may be freely available for research and operational use. However, in practice, there are still many obstacles. For example, in Europe, the databases are primarily a national matter. And there is still a lack of data in international repositories and for some of these, data restrictions are imposed by the data providers, which may limit accessibility (Klein Tank *et al.,* 2010). Nevertheless, a number of weather services follow an open data policy.

- National Weather Service (NWS, part of the National Oceanic and Atmospheric Administration [NOAA]) in the United States. As the NWS is a government agency, most of its products are in the public domain and are available free of charge, ranging from satellite observations and station data to radio soundings and oceanic buoy data.
- Norwegian Meteorological Institute (MET Norway): Official data and products are regarded as public-sector information and are freely available to the public for use, distribution and processing.
- The Royal Netherlands Meteorological Institute (KNMI): Since 1999 the KNMI gradually released their weather station observation data via the internet.

Which data are being made available?

Initially most space agencies provided only raw satellite data, which require expert knowledge for processing and interpretation. Nowadays, many suppliers realise that in order to improve the applicability of their data, so-called higher-order-level products (or end products) need to be provided. For example, with the launch of MODIS in 1999, a wide range of higher-order-level products were developed, such as vegetation indices (NDVI), leaf area index (LAI), land surface temperature (LST), anomalies and fires. These products are all archived and available (most of them at no charge), and have been developed by MODIS – science teams in four discipline groups: atmosphere, calibration, land and ocean. This has been further elaborated by Google with the Earth Engine and other space agencies also provide 'ready-to-use' products.

The range of products is diverse and extensive (based on operational meteorological satellites (EUMETSAT and NOAA) and 'research type' satellites (NASA and ESA)): soil moisture data and anomalies (SMOS, ASCAT, SMAP); rainfall data (TRMM, FEWS and GPCP), 10-daily global vegetation index data (SPOT); LAI; downward radiation (LandSAF); digital elevation maps (SRTM); flood maps; land use/cover maps; lake level data; and many more.

The openness of space agencies has triggered free access of other data sources. For example, a number of global data sets on surface soil moisture have become available in the last decade. These products are based on different satellite sensors. Ground validation is required in order to demonstrate their applicability and further improve such products. This resulted in the International Soil Moisture Network initiated by GEWEX and ESA (ISMN, 2014) to establish and maintain a global *in situ* soil moisture database (available after registration), which is essential for validating and improving global satellite observations and land surface models.

More and more meteorological services release most of their 'standard' weather station observation data (e.g. air temperature, relative humidity, wind speed and rainfall). This is not the case for weather forecasts, with the exception of the NWS in the United States and MET in Norway. The spatial resolution for most of the weather forecasts is limited to national and regional levels.

The weather services in the United States (NOAA) and in Europe (European Organisation for the Exploitation of Meteorological Satellites (EUMETSAT)) have a number of operational satellites. NOAA also follows an open data policy. The role of EUMETSAT is different (but changing, e.g. the Copernicus project), although their data and services are provided free to all African countries (EUMETSAT, 2014).

Impact on the smallholder ecosystem

Impact on governance

In general, the smallholder will benefit, like all citizens, from a government that performs well and can act swiftly to sudden events. The open science data policy has led to a number of global food security applications, most of them specifically intended for local governments, NGOs and decision-makers, which in turn also benefits smallholder farmers:

- Crop explorer (http://www.pecad.fas.usda.gov/cropexplorer/Default.aspx): The global Crop Explorer provided by USDA's Foreign Agricultural Service enables global food supply monitoring, and allows users to explore by crop or region. The explorer combines weather data and coarse satellite observations (e.g. MODIS) with soil moisture and crop models to monitor agro-meteorological variables and crop conditions throughout the world.
- FEWS NET (http:fews.net): Famine Early Warning System Network created by USAID. The goal of FEWS NET is to lower the incidence of drought- or flood-induced famine by providing to decision-makers, timely and accurate information regarding potential food-insecure conditions. With early warning, appropriate decisions regarding interventions can be made. The agricultural and vegetation conditions are monitored in Africa, based on information, satellite (MODIS - NDVI) and weather data (rainfall estimates) provided by the USGS, NASA and NOAA. Beside satellite information, FEWS contains information from other sources e.g. commodity prices.
- SERVIR Global (https://www.servirglobal.net/): This is a regional visualisation and monitoring system (a joint effort of NASA, USAID, World Bank and CCAD) specifically intended for decision-makers. The system provides earth observation (EO) and predictive models (to monitor and forecast ecological changes and respond to natural disasters (e.g. droughts, floods, frost, and fire events)) based on data from orbiting satellites.
- Global Forest Watch (GFW, http://www.globalforestwatch.org/): The launch of the Google Earth Engine has led to an improved (higher spatial resolution) forest-mapping tool. GFW is an interesting example how open satellite data can be used to empower the tribal of people in the forest. GFW is free and follows an open data approach in putting decision-relevant information in the hands of governments, companies, NGOs and the public.

There is considerable experience in the application of space and meteorological data to the governance of food and nutrition security. Data from this effort are now more and more 'openly' available. The impact of open space and meteorological data on better governance is clearly there, but will develop and improve further over time. Therefore, we ranked the current impact as 2, with a potential impact of 3.

Impact on the value chain and sustaining services

Potentially, many of the above-mentioned services facilitate the **value chain and sustaining services based on EO data**. Crop monitoring and harvest prediction services enable farmers, traders, storage providers,

processers and other actors in the value chain to anticipate events in the coming harvest season. Satellite monitoring can be used to detect excessive rainfall or flooding of infrastructure. The satellite monitoring of crops may increase the confidence of microfinance companies to provide loans to smallholder farmers or to serve as an index for micro-insurance companies.

EARS – Earth Environment Monitoring (www.ears.nl) – is an example of a small innovative business that acts in this domain. They provide a crop monitoring service for Africa, which is based on hourly data from Meteosat (EUMETSAT). The visual and thermal infrared images are used to determine the crop water use and rainfall. These are then applied in drought monitoring, crop yield forecasting and river flow forecasting systems. Since 2009, EARS has developed a satellite-based drought and excessive rainfall index for insurance companies (based on 30 years of data).

Other index insurance pilots exist and the G4AW programme of the Dutch Government will result in more services (http://g4aw.spaceoffice.nl/en/). As a result, impact on services is emerging and we rank the current impact as 1, with a potential impact of 3. The impact on the value chain for smallholders has not been found (ranked 0), but potential impacts of flooding on the road network can be measured (ranked 1). For crop and yield estimates from space there is currently an issue of resolution as will be explained in the section below.

Impact on agricultural production

Direct monitoring and agricultural advice on the management practice in the field based on satellite information is difficult. There is a tension between the scale of observation (25–250 m) and the size of the cropping areas (which generally containing mixed crops). Direct advice based on satellite information on the production on the ground is in many cases difficult.

Despite this discrepancy, the first applications directly impacting smallholder agricultural production are emerging. For example eLEAF (www.eleaf.nl) is an advisory firm that operates in the transition area between RS science and operational applications. eLEAF is specialised in using satellite data (including open satellite data from Landsat and MODIS, Meteosat, etc.) for estimating crop water use and crop growth (biomass and yield) and has developed a satellite-based crop monitoring service, which is operational in a number of countries. In 2012, eLEAF was involved in a pilot project called SMART-ICT, funded by IFAD. The project included developing and using tools for smallholder farmers that can monitor plot specific information from satellite measurements and was tested in Egypt, Ethiopia and Sudan. Detailed and field-/crop-specific information (e.g. crop water use, irrigation requirement and biomass) were provided on demand through web platforms and SMS services.

Although more and more space and meteorological data are becoming openly available, their uptake and usage to enhance smallholder food and nutrition security is limited. This is partly because it takes specialist skills and knowledge to apply this data in a relevant and reliable manner, partly because the resolution of current generation of open satellite and meteorological data are too coarse in space and time. There are two programmes that aim to further stimulate the usage of open satellite data for smallholder food and nutrition security:

1) NASA and AGRA (Alliance for a Green Revolution in Africa) joined forces to get the data in the hands (and mobile phones) of people who could use it most. Currently, they are investigating ways to get satellite data to farmers and distributors through the mFarms platform (an ICT platform designed to help stakeholder in agricultural value chains communicate with each other efficiently, establish and maintain business relationships and manage the flow of goods and services among them). mFarms provides agricultural information via cell phones to their network – 80,000 farmers and thousands of other distributors, warehouses and more in 17 African countries (NASA, 2014b).
2) In 2013, the Dutch Government launched the Geodata for Agriculture and Water (G4AW) programme (http://g4aw.spaceoffice.nl/en/). Its objective is to improve food security in developing countries by

providing food producers with relevant information, advice or products facilitated by satellite information. Netherlands Space Office (NSO) is executing this programme, commissioned by the Dutch Ministry of Foreign Affairs.

As a result, we ranked the current impact of open space and meteorological data on smallholder production as 1, because some experiments have been carried out, and the potential impact as 2 with the current available sources, because of the spatial and temporal resolution of the available data. This may change if very-high-resolution sensors become available in openly accessible systems.

International organisations and partnerships

International organisations and partnerships are key drivers for the availability of open data in developing countries in two ways (Boyera and Iglesias, 2014). On the one hand, they encourage and facilitate open government movements by sponsoring or setting up projects and programmes and on the other hand, they make their own data resources available in open data format. The biggest player is the World Bank, active in all developing regions, leading a wide variety of sectorial initiatives such as open transport, open finance, open aid, open climate etc. World Bank was also the first large international organisation to open their data resources in 2010, starting with 2000 data sets. Currently all main international organisations make their global data sets and country reports available in open data format. Examples are: World Bank (http://data.worldbank.org/), www.opendataforafrica.org, FAO (http://faostat3.fao.org/faostat-gateway/go/to/home/E), UNEP (http://geodata.grid.unep.ch/), UN (http://data.un.org) and WTO (http://www.wto.org/english/res_e/statis_e/looking4_e.htm#summary)

Which data are being made available?

In general, data such as global- and country-level indicators, derived from governments, economical models and census are being made available. This type of data is very useful to monitor the state of the earth over time or to make a first characterisation of a country or compare between countries. A visualisation of such data is presented in the Figure 9. Different application programming interfaces (API) and other tools are being developed to make the data easily reusable by others (Figure 10).

Impact on the smallholder ecosystem

Impact on governance

Open data presented by international organisations has primarily an impact on governance. The data allows us to make a quick assessment of a country's statistics; to study changes over time; and to compare countries, including general agricultural indicators such as the amount of harvested or processed crops. The data allows governments, international organisations and NGOs to make policy decisions, especially at a higher level. If a country does not have reliable statistical or census agencies, these portals may contain the only available data sets about that country. They may also include data which is generally not collected by national governments but may be important to help solve the major global challenges such as the Millennium Development Goals, how to feed the world, address climate change, generate sustainable energy and address loss of biodiversity. These data are suitable to generate infographics that are used to inform decision-makers and the public.

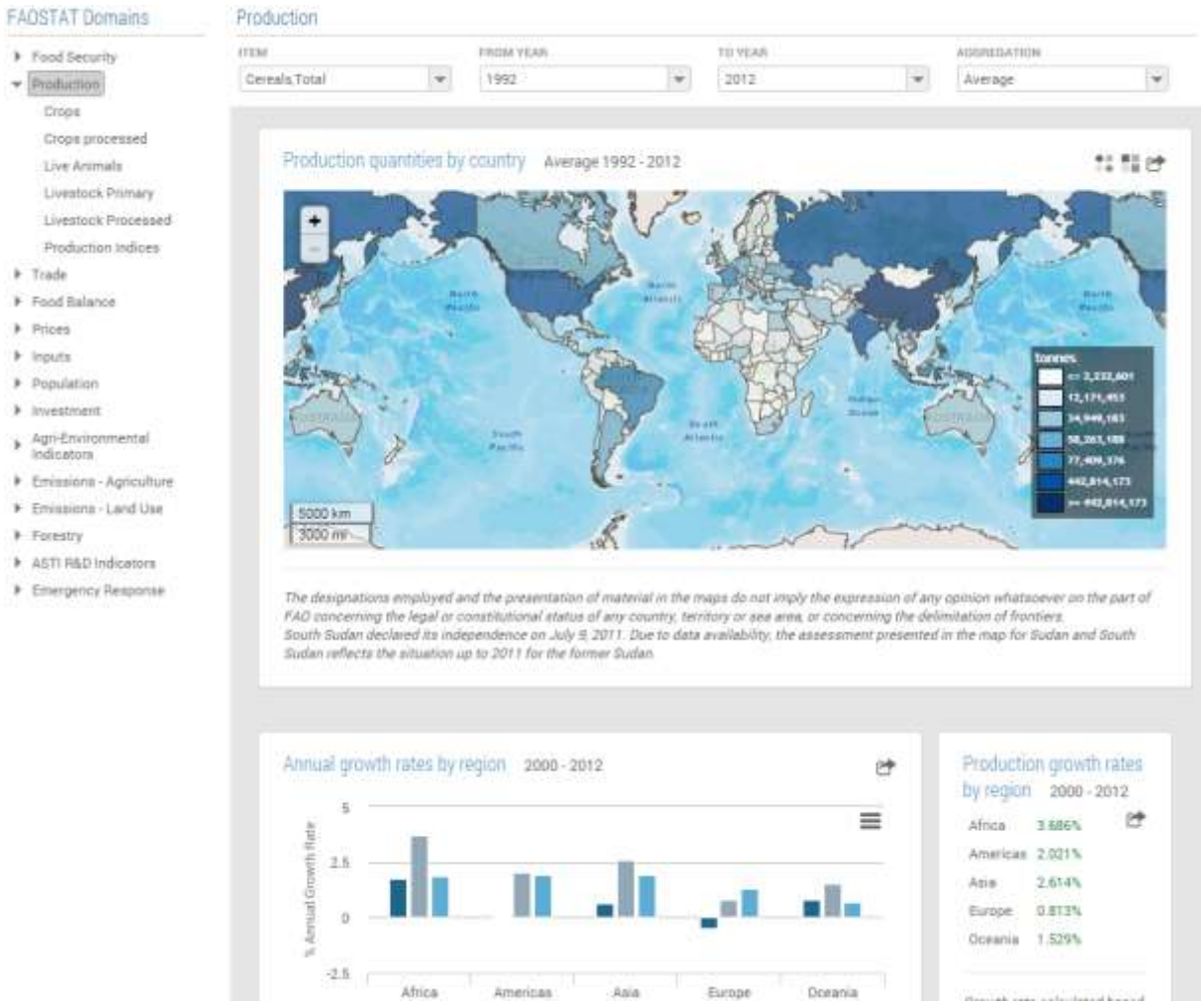

Figure 9. Visualisation of the total yearly cereal production of the countries in the world as typical example of an open data set from of the international organisations.

Source: http://www.foodsecurityportal.org/

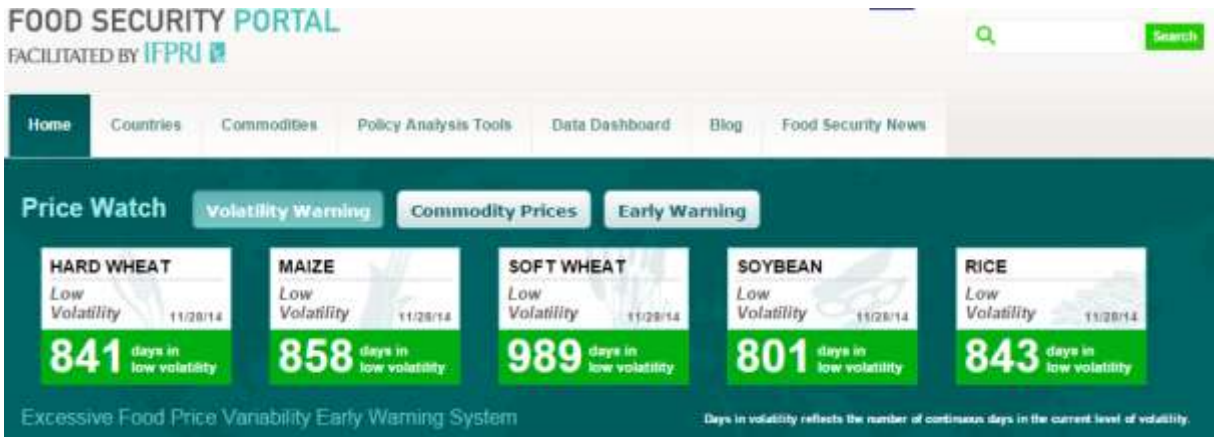

Figure 10: The IFPRI Food Security Portal as an example of the reuse of open data from the World Bank and other sources.

However, international organisations generate more impact when they open their raw data rather only the aggregated numbers. Many of these country-specific data sets will be based on data aggregated from regional, national and subnational levels. This is efficient in times when computational and storage capacity is low and may be the only way to analyse and comprehend large global data sets. But with each aggregation, information is lost. In recent times, computational and storage capacity have increase dramatically, and new ways of visualising and analysing large data sets are being discovered. Data from international organisations will become much more informative and generically applicable to a smallholder context as the raw, fine-grained data that form the basis for the global statistics become available.

Therefore we ranked the current impact of data from international organisations on governance as 2, with a potential to grow to 3 if more detailed data become available.

Impact on the value chain and sustaining services

No direct evidence for impact of open data presented by the international organisations has been found. The potential impact on the value chain and sustaining services is low. Most of the data have a high (regional, national, subnational) aggregation level and is not specific enough for direct application in the value chain or sustaining services. Exceptions to this are world and national commodities price information which is relevant for trade.

Therefore, we rank the current and potential impact of open data from international organisations on the value chain and sustaining services as 0.

Impact on agricultural production

No evidence for impact or potential impact on agricultural production could be found.

As a result, and by nature of the current data available, we do not see current or future impact on the food and nutrition security of the smallholder communities (ranked 0).

Science

The concept of open data first emerged in the scientific community with the establishment of the World Data Centres (WDC) in 1957. The objective of the WDCs was to minimise the risk of data loss and to maximise data accessibility (Wikipedia Open Data, 2014). Other advantages of open access to research data are that:

- research results based on data can be verified and critically examined;
- unnecessary duplication of research work can be avoided;
- data can be analysed comprehensively and made use of, e.g. in follow-up projects;
- the research process can be accelerated through data-sharing;
- new findings can be achieved by merging data from different sources;
- the merging of data brings an informal added value and yields higher-quality data products, e.g. indices and databases;
- data sets which are collaboratively assembled and jointly used are more cost-efficient;
- open access promotes reuse of data by the public and by industry (Open Access, 2014).

The first WDC system was founded in the United States, USSR and several other countries by the International Council of Science (ICSU) to archive and distribute data collected from the observational programmes of 1957–58 (The International Geophysical Year). Most of the data was exchanged on a barter system; non-members could acquire data at the basis of copying cost. However, the usage and up take of this ‘open’ data was limited and was only available to a selected number of individuals and organisations (Landers, 1979). Later, the data

became machine-readable and with the foundation of the internet in 1989, the potential for exchange of documents and information between individuals and organisations increased dramatically.

In 2004, the science ministers of all nations of the OECD (Organisation for Economic Co-operation and Development), which includes most developed countries, signed a declaration which essentially states that all publicly funded archive data should be made publicly available. Following a request and an intense discussion with data-producing institutions in Member States, the OECD published in 2007 the *OECD Principles and Guidelines for Access to Research Data from Public Funding* as a soft-law recommendation (OECD, 2007).

From the social sciences community, the Dataverse Network was created in 2006 at Harvard University (http://thedata.harvard.edu/dvn/). The Dataverse Network is an open-source application for sharing, citing, analysing and preserving research data. Its main goal is to solve the problems of data-sharing through building technologies that enable institutions to reduce the burden for researchers and data publishers, and incentivise them to share their data. By installing Dataverse Network software, an institution can host multiple individual virtual archives, called 'Dataverses' for scholars, research groups, or journals, providing a data publication framework that supports author recognition, persistent citation, data discovery and preservation. Examples of Dataverse Networks installed at universities and organisations, where data can be downloaded for reproducible research, are:

- Harvard Dataverse Network (http://thedata.harvard.edu/dvn/): which contains scientific data from all disciplines worldwide, including the world's largest collection of social science research data;
- Dutch Dataverse Network (https://dataverse.nl/dvn/): which contains the data from a number of Dutch universities;
- European Union Democracy Observatory (EUDO).

In 2008, the WDCs were reformed (together with the ICSU Federation of Astronomical and Geophysical data-analysis services) in the new World Data System (ICSU-WDS) to bring all scientific sources together under a single umbrella with same quality standard (ICSU-WDS, 2014). Accreditation criteria are set in terms of: scientific relevance, governance, data management, technical infrastructure and security, and open and equitable access to quality-assured scientific data, data services, products and information. All shared data will be made available with minimum time delay and at minimum cost. ICSU-WDS data portal gives access to data from members (currently more than 80 members), primarily active in the field of geophysical science: the WDC for Meteorology (USA, maintained by NOAA), the Ocean Data Portal (oceanographic data held by the IODE global network of 80 National Oceanographic Data Centres), the Global Change Master Directory (GCMD), or the Global Earth Observations System of Systems (GEOSS), etc.

There are many other data archives and repositories, some of which are specific to certain research communities. Databib and re3data contain a large searchable collection of online research data repositories.

This year, the next step in open science data development has been taken by launching the *Open Data Journal for Agricultural Research* (ODjAR) (http://library.wur.nl/ojs/index.php/odjar/). ODjAR aims to make open data publication more attractive for researchers while enabling authors to get scientific credit for their work through citations and digital object identifiers for future reference.

In the remainder of this section, we will discuss the open research data from a smallholder perspective, focusing on agricultural research data, using the CGIAR Consortium as an example.

Agricultural data

As a consortium of 15 international agricultural research centres, CGIAR developed its open data policy in 2013, which will be fully implemented in 2018. Many of the member organisations – AfricaRice, ICRAF, IFPRI,

Bioversity International, CIFOR, and CGIAR Research Program on Forests, Trees and Agroforestry – are also using Dataverse

In 2014, CGIAR launched the CGIAR Consortium Data Management System (CGDMS) (http://www.cgiar.org/resources/open/data-management-system/).

Some randomly chosen examples of open data sets from members of the CGIAR Consortium are:

- ILRI Data portal: ADA Nicaragua - Baseline Survey. The database of 169 households containing monitoring productivity of animals in DGEA1 (Germplasm for Dairy Development in East Africa) dairy cattle keeping households: calving information, milk production, etc. (http://data.ilri.org/portal/)
- IRRI Dataverse: SOCIO-ECONOMIC EVALUATION OF HYBRID RICE CULTIVATION IN THE PHILIPPINES Basic household and farm characteristics and input and output data in rice production such as; yield, fertiliser, insecticide, weedicide and other pesticides, and labour use. Data are also available on adoption of modern rice technologies, prices of input and output and farm wage rates. 1970 – 1971(http://irri.org/tools-and-databases/irri-dataverse).
- IRRI Dataverse: CENTRAL LUZON LOOP SURVEY: Basic household and farm characteristics and input and output data in rice production such as; yield, fertiliser, insecticide, weedicide and other pesticides, and labour use. Data are also available on adoption of modern rice technologies, prices of input and output and farm wage rates (http://irri.org/tools-and-databases/irri-dataverse).
- IWMI Water Data portal: District-wise crop area, production and yield for all crops growing in the region of 52 districts in Andra Pradesh, Maharashtra and Karnataka. Data were downloaded and processed by ACIAR project (http://waterdata.iwmi.org/DataArchive.php).

The 'standard' open science data portals are focused on sharing data within the research community to enable better science. Open agricultural research data is diverse in nature and generally very specific and focused on answering a certain research question. They tend to be collected in a limited time span. Only a few data sets are opened (151 in IRRI Dataverse and 30 in the ILRI data sets portal).

Some data sets are aimed at a more general application and are developed for the wider public.

- NEXTGEN Cassava project: 'Cassavabase' provides a 'one-stop shop' for cassava researchers and breeders worldwide, including genomic selection analysis tools and phenotyping tools.
- The 3000 Rice Genome project: The giga-data set contains the genome sequences (averaging 14× depth of coverage) derived from 3,000 accessions of rice with global representation of genetic and functional diversity. The challenge now is to comprehensively and systematically mine this data set to link genotypic variation to functional variation.
- RTB Atlas is an online mapping resource for the community of people working to improve roots, tubers and banana (RTB) crops. The platform helps scientists set priorities for interventions to improve production of cassava, yam, potato, sweet potato, banana and plantain and allows users to overlay 25 sets of variables onto a world map (including harvested area, potential yield, and yield gap).
- Global Yield Gap and Water Productivity Atlas aims to inform policy-makers about the difference in current average farm yields and the potential rain-fed and irrigated yield. Water productivity is used as indicator to express the efficiency in converting water to food.
- IWMI World Water and Climate Atlas gives irrigation and agricultural planners rapid access to accurate data on climate and moisture availability for agriculture. The atlas includes monthly and annual summaries for precipitation, temperature, humidity, hours of sunshine, evaporation estimates, wind speed, total number of days with and without rainfall, days without frost and Penman-Montieth reference evapotranspiration rates.
- The Integrated Database Information System (IDIS) is an online data-sharing platform that provides access to water, agriculture and environment scientific data to help researchers and their research partners

improve the water productivity. IDIS contains over 1 billion records, with a focus on IWMI and CPWF river basins.

Impact on the smallholder ecosystem

The direct applicability, and therefore impact, of the open science data sets in the smallholder ecosystem is limited. However in the long run, better scientific research on smallholder production and the smallholder ecosystem will have a large impact on the smallholders, but only if the new knowledge and insights generated in the science world find its way to other actors in the smallholder ecosystem, for example via extension workers. We ranked this indirect impact as 0, because it is not the result of open data development.

Another way is to develop specific data sets dedicated to actors outside the science community. Information systems such as Harvest Choices, Global Yield Gap and Water Productivity Atlas, IWMI World Water and Climate and IDIS aim for better governance, allowing planners to make better decisions on governance, investments or to optimise irrigation. Cassavabase, the Rice Genome project and the RTB Atlas are aimed at actors in the value chain facilitating breeders in crop improvement; these improved crops will impact the smallholder production directly by enabling to plant crops with specific characteristics. So far, no open data sets dedicated to agricultural production have been found, but IRRI provides the Rice Crop Manager app as a way to let smallholders benefit from scientific insights.

To increase the impact of scientific research, more dedicated open data sets need to be developed to facilitate knowledge- driven decision-making.

The current impact of dedicated data sets is ranked as 1 for the support services (breeders) and as 2 for governance. However, all four impact groups will benefit if more dedicated open data sets emerge (ranked as 3), although each of the actors within these impact groups will have specific questions and needs.

NGOs

Currently, there is a transition going on in the development sector towards 'open development'. This development is driven by the idea that transparency in development (Broek *et al.,* 2012):

- increases accountability of the development process in the South as well as in the North;
- improves allocation of scarce development resources in developing countries;
- increases impact of development in reducing poverty;
- improves lives in developing countries; and
- maintains domestic support for development at times of financial stringency.

At four High Level Forums for Aid Effectiveness (2002–2011), the international community committed itself to transparency and accountability as two cornerstones for effective development cooperation.

In March 2005, the Paris Declaration on Aid Effectiveness was signed. Donor governments, multilateral organisations, NGOs and partners agreed to work together to make developing countries more in charge of their own development processes, and to hold all stakeholders in the development process accountable for achieving concrete development results. Transparency is needed to achieve this goal. To demonstrate the commitment of NGOs to transparency and accountability the International NGO Accountability Charter was launched (INGO Accountability Charter, 2014). It aims to speed up progress in transparency among international NGOs by helping its members to establish a high quality accountability framework that helps them to report on transparency and effectiveness, among others. Increasingly, NGOs are making their project descriptions, development goals, activities and spending available in IATI Standard (IATI Standard, 2014).

Different governmental and private donors organisations (IDRC, DFID, Hewlett Foundation, Open Society Foundation) now also require open data publication from their beneficiaries (Boyera and Iglesias, 2014).

Which data are being made available?

As a result, most of the open data in the NGO world is focused on the project accountability data set. An XML structure describes the project in detail including the summaries, activities, the budget and the time schedule etc. These data are quarterly updated with the results of activities. Currently these types of data sets are used to:

- show projects on a website Cordaid (https://www.cordaid.org/en/projects/);
- show projects on a map (openaidsearch.org or http://mali.publishwhatyoufund.org/);
- gain insight into a country's aid portfolio (http://openaid.se/, http://openaid.nl/);
- provide internal intelligence (http://www.iatiregistry.org/publisher/oxfamgb) (Open for Change, 2014).

The NGO world is also experimenting with crowdsourcing techniques in order to monitor and evaluate these projects or to collect data. Examples are GoThree60 and Open RBF:

- In their pilot project, GoThree60 collects opinion about the function of a maternity health care clinic in Uganda to collect an independent opinion directly from the users (GoThree60, 2014).
- Open RBF take transparent financial monitoring a step further. For a health care centre in Burundi, Open RBF collects the improvement goals; monthly or quarterly health care providers enter their performance data in the system and the local community verifies service delivery and quality by providing independent feedback through interviews and mobile devices. The RBF system measures all these performance data against the agreed criteria and determines the amount of money to be allocated. This will then be transferred to the clinic or hospital. A fundamental aspect of Open RBF is that, once validated, all data are displayed on the internet for everyone to see and accessible through easy-to-read dashboards. This allows funders to see how resources are allocated, but more importantly, it encourages citizen engagement with a vital development issue. Through the Open RBF system, patients can follow the budgets, targets and performances of any clinic or hospital in real time. This information enables them to hold their governments to account for their efforts to provide better health care for all. The transparency offered by Open RBF can help to restore trust between citizens, local and national governments and private sector service providers (Cordaid, 2014).

Similar developments can be found in agriculture-focused projects. AgriTerra, an NGO that supports farmers' organisations in developing countries, currently works on different methodologies (on paper, via SMS, using special software on tablets) to enable individual farmers of households to record data at farm level, experimenting in Peru, the Philippines and some African countries. The data collected by the farmers can be used by farmers during the project to compare and to learn, by the farmer organisations to do a meta-analysis on their members and by AgriTerra for monitoring and evaluation of their project. The collected data at farm level is also relevant for third parties, e.g. labels such as UTZ may use these data to sustain their quality claim (AgriTerra, personal communication).

CABI's Plantwise project demonstrates how project data can be used to build up a detailed large-scale database of the spread of plant diseases in the world. Plantwise works with national governments to set up plant clinics where trained plant doctors provide farmers with practical, science-based diagnosis and advice to prevent and manage crop loss. Supporting this network of clinics, the Plantwise knowledge bank ensures an online and offline gateway to diagnostic services, pest tracking, and best-practice farmer recommendations specific to every country. The plant doctor reports back to the system the occurrence of diseases in the crops and within time a detailed map of global plant disease spread emerges (http://www.plantwise.org/).

Impact on the smallholder ecosystem

Impact on governance

NGOs become more transparent with the development and implementation of their open project data policies, explaining why, where, when and for what activities project money is being spent and what the result was. Smallholders, like other beneficiaries of NGOs, will benefit from this development as a result of better governance and money better spent.

An interesting development would be if NGOs were to open their data collected at grassroots level for monitoring and to evaluation and steer their projects. Currently, only a few structured large-scale data sets with a longer time span are available about smallholders at household level. Opening up this kind of data will reveal the detailed patterns behind large-scale census data collected by governments and other institutes. The monitoring and evaluation data may enable the understanding of the bigger picture supported by quantitative data, resulting in better policies.

However opening up farm household data is not an obvious development. It is the responsibility of the data owners – farmers, farmers' unions and/or NGOs – to do so. While doing this, farmers' rights need to be considered and maybe intermediate steps are needed such as anonymising or aggregating the data into relevant groups.

We ranked the current impact of open data on governance as 2, because the principle of open development is clearly emerging. Soon, this development will mature and have an impact rank of at least 3.

Impact on the value chain, sustaining services and agricultural production

A lot can be learned about the smallholder ecosystem from projects targeting agricultural production or the value chain; this can be used to further educate farmers, farmers' cooperatives or to design better projects. Currently, little of this information is available as open data. If data and results are shared more, the impact of these projects can become much larger and cross-fertilisation of projects, even across continents, can take place.

We ranked the current impact of NGO monitoring and evaluation data on the value chain, sustaining services and agricultural production as 0. However, potentially these data can provide a lot of information about the functioning of the smallholder ecosystem, and thus would be ranked 3.

Business sector

The United Nations Global Pulse states that open access to corporate data is the next frontier in the development of open data (Verhulst, 2014). Private sector companies accumulate a tremendous amount of data in their day-to-day operations. Market research, communications tracking, client relationship management and market activities generate a wealth of information, which tends to stay in the private domain (Responsible Data Forum, 2014). The corporate world in effect 'owns' terabytes of data and metadata, e.g. almost 7 billion telephone subscriptions are producing communication data every day, more than 1.82 billion people communicate on some form of social network and almost 14 billion sensor-laden everyday objects (trucks, health monitors, GPS devices, refrigerators, etc.) are connected and communicating over the internet, creating a steady stream of real-time, machine-generated data (Verhulst and Sangokoya, 2014). If this corporate data were made available in a de-identified and aggregated manner, researchers, public interest organisations and third parties would gain greater insights on patterns and trends that could help inform better policies and lead to greater public good. However, access to corporate data is sensitive and extremely limited. This is related to privacy issues, security and proprietary interests. The United Nations Global Pulse has the vision that the corporate world should not keep this data to themselves, but that it should be harnessed safely

and responsibly as a public good. The United Nations Global Pulse is trying to persuade the corporate world to start making data sets more openly available (Verhulst, 2014).

Also the private sector itself has a lot to gain from data philanthropy or corporate social responsibility addressing data. Sharing data may spark innovation, can be used to scout talent and can help to safeguard the client base. As an imaginary sample of the latter case, the data from a mobile operator operating in a developing country may contain the signals of misfortune of its clients discussing floods, crops failures or unaffordable price raises. By sharing and analysing this data with governments or NGOs, the misfortunes of the clients may be prevented (UN Global Pulse, 2013).

Which data are being made available?

A quick survey of the Responsible Data Forum (2014) identified six categories in which business are experimenting in sharing their data for the public good:

- academic research partnerships, in which corporations share data with specific universities and other research organisations;
- prizes and challenges, in which companies make data available to qualified applicants who compete to develop new apps or discover innovative uses for the data;
- trusted intermediaries, where companies share data with a limited number of known (often commercial) partners;
- APIs, which allow developers and others to access data for testing, product development, and data analytics;
- intelligence products, where companies share (often aggregated) data that provides general insight into market conditions, customer demographic information, or other broad trends;
- corporate data cooperatives or pooling, in which corporations group together to create 'collaborative databases' with shared data resources.

Which data are being made available?

In this document we will further focus on telecom and IT data, but global companies in agricultural inputs and food processing are interacting more and more with smallholders both as a market and as a resource.

Experiments based on telecom data

Analysing the spread of malaria

In Kenya, researchers from the Harvard School of Public Health have deducted the influence of human mobility on malaria spread based on all calls and text messages sent by the Kenyan mobile phone subscribers of SafariCom during a period 1 year and combined with detailed diseases data. The researchers could estimate the probability for each person in the data set to carry malaria parasites and build a map of parasite movements between 'source' areas (areas that act as reservoirs of disease) and 'sink' areas (areas that mostly receive disease) obtaining a better understanding of how a disease is spreading (Harvard, 2012).

Enforcing smallholder food security

In 2014 Orange launched the 'Data for Development Senegal', an innovation challenge on ICT big data for the purposes of societal development. For this challenge, Sonatel and the Orange Group are making anonymous data, extracted from the mobile network in Senegal, available to international research laboratories, as well as

data on hours of sunshine. Suggestions for applications in the agricultural domain can be found in Box 1 (Orange, 2014).

Box 1: Possible applications of telecom data in the agricultural domain as suggested by Orange in the 'Data for Development Senegal' Challenge

- analyses based on hours of sunshine:
  - exploration of explanatory factors and modelling of soil productivity and basic foodstuffs according to climatic conditions, as well as development of predictive models
  - anticipatory analyses of soil evaporation and the depth of wells and watercourses, modelling and anticipation of drought
  - exploration of factors explaining low production of rain-fed crops according to meteorological conditions
  - evaluation of the evaporative demand in bodies of water (holding tanks, agricultural reservoirs…)
  - correlation between hours of sunshine, temperature and conditions of conservation and quality of seeds
  - analyses of correlation between measurements of hours of sunshine and temperature and variation in basic foodstuff prices
- analyses based on mobile network use statistics:
  - optimisation of irrigation infrastructure in order to maximise the safe access to water and the productivity of irrigated land
  - optimisation of the efficiency of the harvesting infrastructure: especially for peanuts
  - optimisation of the location of warehouses for the various foodstuffs
  - impact analysis about the access to a local food supply
    - on the health of workers' families (see Health theme)
    - on population movements and insecurity
  - analyses of livestock migrations and method for minimising mortality
    - locating markets...
  - analysis of the impact of the seasons on rural populations
    - emigration of temporary workers
    - activity during the rainy season
    - migration in the case of drought
  - measurement of the impacts of fluctuation of prices of farming products in productive areas
  - modelling of possible impacts for roll-out of measuring methods
    - water level, crowdsourcing sensors…
  - impact analysis about a locust invasion
  - comparison measures with other emerging countries
  - farming yield forecasts (e.g. changes in call volume during periods of drought)
  - performance of the groundnut marketing campaign (e.g. telephone top-up amounts purchased during the marketing period in November–December in the groundnut growing region)
  - movements of herdsmen in northern and eastern Senegal (e.g. definition of transhumance routes from SIM cards).

Source: Orange (2014)

Measuring actual rainfall patterns

In the Netherlands, it has been demonstrated how rainfall pattern can be derived from commercial cellular communication networks (Figure 11, Overeem *et al.,* 2013). The potential of such networks is high, in particular in those parts of the world where networks of dedicated ground-based rainfall sensors is virtually absent and telecommunication networks are expanding, e.g. in most African countries.

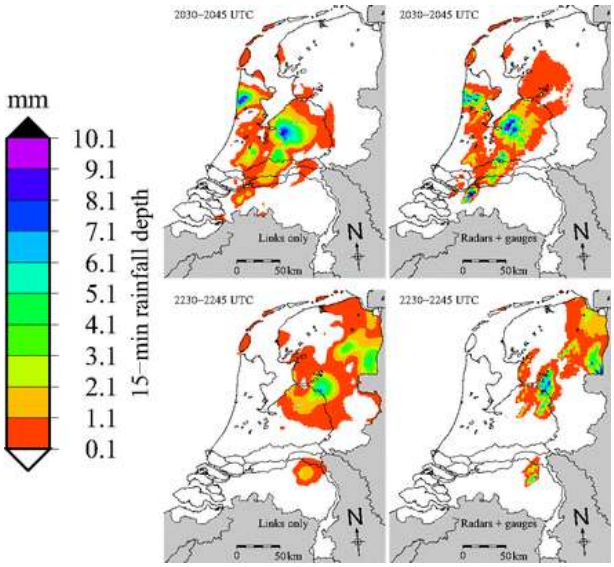

Figure 11: Space–time dynamics of 15-min rainfall depths (two panels per time step) from links (left) and radars plus gauges (right) for 10 September 2011, 2030–2045 and 2230–2245 hours UTC (validation).

Source: Overeem *et al.* (2013)

Experiments based on e-mail data

Analysing migration patterns

Researchers from the Max Plank Institute and Yahoo! Research have studied age and gender specific international migration patterns based on a large sample of Yahoo! e-mail messages, mapping anonymised e-mail users to the geographic locations using IP addresses. The findings were in line with existing administrative data sources, but provide new and rich information on mobility patterns and social networks of migrants (Zagheni and Weber, 2012).

Experiments based on data from mobile agricultural advise services

Early warnings for disease and prices spikes

Currently, more and more agricultural advisory services are being developed that can be accessed by mobile telephones (GSMA, 2014). Many of these applications ask the farmer or extension worker to enter data into the system in order to provide the right advice. Recently Palantir, a data mining company, analysed a data set of 1,000,000 requests from the Community Knowledge Worker programme of the Grameen Foundation. They were able to locate early outbreaks of animal diseases in Uganda and link these outbreaks to spikes in food prices (Palantir, 2012).

Large-scale agricultural and nutrition patterns

Similarly, at the Africa Open Data Jam (August 2014) IBM® used a cloud platform for cognitive computing to demonstrate an app that analyses SMS data to collect agricultural and nutrition information from farmers with survey questions and shared back aggregate information as open data (IBM, 2014).

Other experiments on corporate data-sharing

Early locust warning

The Disaster Charter is a cooperation of space companies and agencies that coordinated the of space facilities in the event of natural or technological disasters. An example is the usage of DMC data to predict locust breeding grounds and swarms in North Africa (DCMII, 2013).

Early disease warning

Google offers a service to visualise the search intensity for certain key words. Using this feature, Google is capable of following the outbreak of flu almost in real time. They discovered that there is close relationship between how many people search for flu-related topics and how many people actually have flu symptoms. Historical data demonstrate that the estimates based on Google search queries about flu closely match traditional flu activity indicators. Therefore the influenza can be used to detect outbreaks much faster than using traditional methods, enabling health professionals to better respond to seasonal epidemics and pandemics (Google, 2014b).

Impact on smallholder ecosystem

It can be concluded from the experiments described above that the potential impact of open business data coming from telecom and IT sources is very large.

- Mobile operators, in particular, have the opportunity to collect information on individuals at the grassroots level, almost in real time and in large volumes, even from remote areas. The mobile network has the potential to grow into a community sensor and, if governed with care, be used as a valuable monitor of rural life. The examples above do not only include the telecom data themselves, but also the weather sensors on the antennae. Even the microwave signals themselves can function as a sensor for rainfall.
- Related to the mobile operators are the operators of mobile services who provide (paid) information services to smallholders and other rural actors, e.g. price information or farm advice. In many case this is two-way communication, often related to agriculture and value chains. In the experiments mentioned above, Palantir demonstrates how this kind of data can be used to discover relevant patterns that are very relevant to smallholder communities.
- Data from these sources can have impact on an all four domains described. Data from mobile operators or mobile service providers can be used to:
    - Collect or validate national or regional statistics and census data, improving government data. In this way, baseline data or the effectiveness of a policy measure can be assessed in more reliable way, probably obtaining the results cheaper and in a more-timely manner then using traditional methods.
    - Validate global data sets. Many global data sets are now derived from satellite imagery with a coarse resolution of 1 km or more. Mobile data can be used to ground-truth and update these data sets or contextualise the data when applied a specific situation, especially in a smallholder context.
- Mobile data can be used to detect patterns and relations in smallholder practice that are now unknown because of lack of observations. Mobile data allow access to relatively cheap large-scale data sets over a long time span, which is currently unaffordable.
- Patterns detected in mobile data can be used to make agricultural advice services more specific for the context of a farm or region and using multiple sources the advice can be validated or improved over time when more data enters the system.
- New insights and information that can be derived from mobile data can be fed back into the farmer ecosystem.

However, the application of these data to the search for patterns that might benefit smallholder food and nutrition security is not common practice. There are limitations from the perspective of the private sector and from the point of view of the users of the networks and the services. Therefore, we ranked the current impact as 0, but the potential impact as 4. We gave business data the highest rank in this survey, because there are no other means to collect so much individual information at the grassroots level in remote areas. The experiments above demonstrate the potential of the data to provide impact. The question is if and how these data can be shared to the benefit of rural communities.

The UN is advocating the use of data from mobiles and mobile service operators to monitor the Millennium Development Goals (UN, 2014).

Synthesis and outlook of the impact of the open development on smallholder food and nutrition security

The current impact of open data on smallholder food and nutrition security is low

In general, the impact of open data in developing countries is low. There are various reasons for this, but this study shows that the data needed to have local impact is not there or not openly available.

- Government data are limited and are often outdated, too aggregated or unreliable.
- International organisations generally provide coarse data sets about country statistics, with limited direct relevance to smallholder communities.
- Agricultural research provides a limited number of research-oriented data sets which are difficult to apply.
- NGOs focus on open data for project accountability which is important for development but have limited application for food and nutrition security.
- Telecom and ICT businesses hold a wealth of data but do not regularly share this with other stakeholders.

As a result, only a few examples can be found of open data applications targeting food security issues directly in the smallholder ecosystem. In Figure 12 the applications mentioned in this report are overlaid on their application domain in the smallholder ecosystem. Most applications are targeting food security at a higher level and in the governance domain, a hand full applications target small holder food security and nutrition issues at the local level: in the service domain the value chain or to the farmers directly.

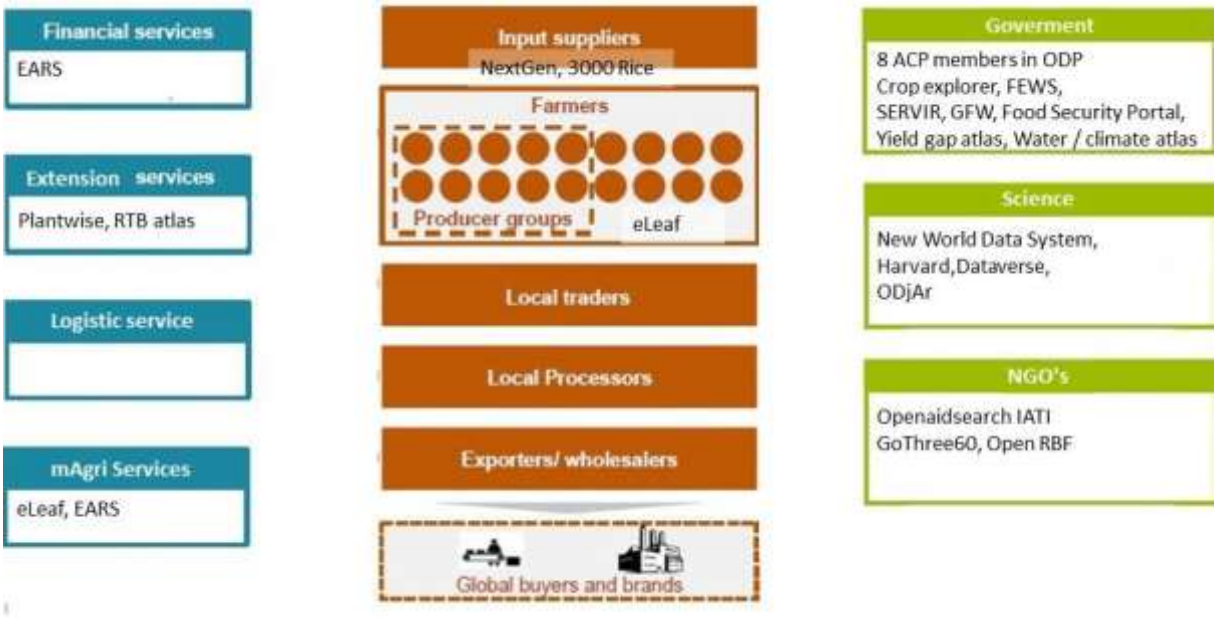

Figure 12. Examples of applications using open data impacting the smallholder ecosystem and its actors, overlaying the examples mentioned in this paper and their application domain.

Source: Freely adapted from Fonzi and Chau (2012).

Potentially there is a large impact of open data on smallholder food and nutrition security

Traditionally, the exchange of data and information between smallholder farmers and the outside world was difficult. This picture has changed with the emergence of mobile operators and ICT service providers in the rural areas of developing countries. Information exchange is now possible between all actors the smallholder ecosystem. The more information is exchanged between the different actors of the smallholder ecosystem, the more smoothly it can function. Mobile and mobile service are the key for open data to become useful to enhance food and nutrition security of smallholders by bringing information to the local level. However, in order to be relevant, in most cases this data needs to be locally relevant. Interestingly, the key to obtaining local relevant information is in the hands of the telecom and mobile service providers.

Business, in particular telecom and mobile service providers, should share their data about the smallholder ecosystem in an appropriate and effective way in order to make 'the big leap forward'.

If the data from the emergent mobile operators and mobile service providers become available in an appropriate way there is a lot to be learned about the functioning of smallholder ecosystems. This information must be used to further improve the services provided. Fine-grained data about the smallholder ecosystem can then be used in a variety of ways:

- National or regional statistics and census data can be collected or validated against mobile data at the lowest level. In this way, baseline data or the effectiveness of a policy measure can be assessed in a more reliable way, probably obtaining the results cheaper and in a more-timely manner than using traditional methods.
- Mobile data can be used to make agricultural advisory services more specific for the context of a farm or region and using multiple sources, the advice can be validated or improved over time when more data enters the system.

- Mobile data can be used to validate global data sets. Many global data sets are now derived from satellite imagery with a coarse resolution of 1 km or more. Mobile data can be used to ground-truth and update these data sets or contextualise the data when applied to a specific situation.
- Mobile data can be used to detect patterns and relations in smallholder practice which are now unknown due to lack of observations. Mobile data allows us to have relatively cheap, large-scale data sets over a long time span, which are currently unaffordable.
- New knowledge and information that can be derived from ICT data can be beneficial to actors in the farmer ecosystem.

Beside telecom providers, there are other businesses that have relevant data on smallholders. The international food processing industry and supply chain is increasingly doing business with smallholders, both as a source of commodities and as a market.

The question is, how should this data be shared and who is the 'owner' of this data: business, individual farmers, or farmer cooperatives? Before this ICT data can be shared, these issues need to be cleared up. The solution probably differs from case to case.

Other options to further improve the uptake and availability of open data for smallholder food and nutrition security

Possible improvements:

1) More, and more reliable, regional and local government data. Specifically smallholders can benefit from data about:
    - ownership and legal status of companies; input suppliers, traders, financial advisers
    - land ownership
    - (regional) trade statistics and prices
    - infrastructure and transport.
2) Data from international organisations would become much more informative and applicable to the smallholder context if the raw fine-grained data could become available rather than the aggregated data at country level.
3) A translation is needed for open research data towards open data that is applicable for other stakeholders.
4) By opening up the NGO monitoring and evaluation data, the fine-grained details behind the large-scale census data collected by governments and institutes becomes as visible as looking through a magnifying glass. These detailed patterns may lead to a better understanding of the bigger picture, supported by quantities data, especially when many NGOs are operating in the same country.

There is a trade-off between the aggregation level of data, the amount of information it contains and farmers' rights

Fine-grained data contain more information, and are more useful for different applications. However, there is a delicate balance between the benefits of open data and the risk of their potential misuse (Figure 13). Telecom data and monitoring data of NGO projects or data collected by farmers or farmers' organisations may contain data at the individual level. These data are valuable to better understand the farmer ecosystem, but also contain an inherent risk of misuse. Therefore, rules are needed to prevent misuse; ultimately it should be the 'data subjects' – the actors in the smallholder ecosystem – who should determine what and in what way data are shared or opened, taking into account the balance between information content and privacy. The UN

(2014) is proposing a set of 'Basic Principles for the Data Revolution for Sustainable Development'. These rules can be the starting point for a more rigorous discussion.

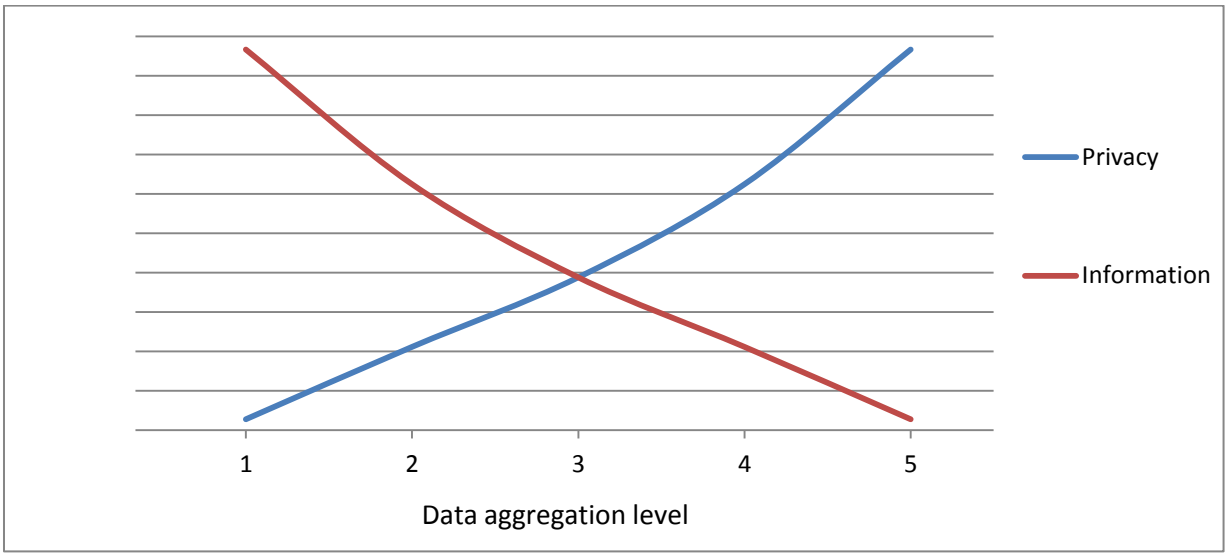

Figure 13. The more fine-grained data are, the more information they contain, the more useful they are for different applications. However there is a delicate balance between the benefits of open data and the risk of potential misuse.

References

Automatic Milking. 2014. *EU Project Automatic Milking.* Animal Sciences Group WUR. www.automaticmilking.nl (Accessed November 2014)

Boyera, S. and Iglesias, C. 2014. *Open data in Developing Countries: State of The Art.* Version 1.1 SBC4D for the Partnership for Open Data (POD). https://docs.google.com/a/okfn.org/document/d/1FMyfLu-jouL7j7Pw0kEwUn_B07aZ9IX3vlFGqPO0gX0/edit?pli=1#heading=h.94sw69toiyh4 (Accessed November 2014)

Broek van den, T., Rijken, M. and Oort van, S. 2012. *Towards Open Development Data. A Review of Open Development Data from an NGO Perspective.* TNO White Paper P10098. http://openforchange.info/sites/pelle.drupalgardens.com/files/TNO%20whitepaper%20-%20Towards%20open%20development%20data.pdf (Accessed November 2014)

CIARD. 2014a. CIARD Manifesto. *Towards a Knowledge Commons on Agricultural Research for Development*. http://www.ciard.net/about/manifesto (accessed November 2014)

CIARD. 2014b. Rights of Farmers for Data, Information and Knowledge. A CIARD E-Discussion. http://www.ciard.net/news-and-events/blog/rights-farmers-data-information-and-knowledge-ciard-e-discussion (accessed November 2014)

Civiccommons. 2014. Initiatives. http://wiki.civiccommons.org/Initiatives (Accessed November 2014)

Cordaid. 2014. *Open Development Movement Co-Creation Leads to Transformation.* Positioning Paper. May 2014.

CSRL. 2014. *SoilWeb: An online soil survey browser.* http://casoilresource.lawr.ucdavis.edu/drupal/node/902 (Accessed November 2014)

CTA. 2013. CTA website. http://www.cta.int/en/. (Accessed November 2014)

CTA/AgriHack. 2015a. CTA Open Data Baseline Study. Institutional Stakeholder Report.

CTA/AgriHack. 2015b. CTA Open Data Baseline Study. Developer Perceptions Report.

Davies, T. (ed.). 2013. *Open Data Barometer. 2013 Global Report.* http://www.opendataresearch.org/dl/odb2013/Open-Data-Barometer-2013-Global-Report.pdf (Accessed November 2014)

Davies, T. 2014a. *Open Data in Developing Countries – Emerging insights from Phase I*. Web Foundation. http://www.google.nl/url?sa=t&rct=j&q=&esrc=s&source=web&cd=2&cad=rja&uact=8&ved=0CCkQFjAB&url=http%3A%2F%2Fwww.opendataresearch.org%2Fsites%2Fdefault%2Ffiles%2Fpublications%2FPhase%25201%2520-%2520Synthesis%2520-%2520Full%2520Report-print.pdf&ei=4knJVJPuGMLEPKrngIgP&usg=AFQjCNHgnyhKptsRYUuSDxPWlzlHS9hIEw&bvm=bv.84607526,d.ZWU (Accessed November 2014)

Davies, T. 2014b. *Open Data Policies and Practice: An International Comparison.* Paper for European Consortium for Political Research Panel P356 – The Impacts of Open Data. http://ecpr.eu/Filestore/PaperProposal/d591e267-cbee-4d5d-b699-7d0bda633e2e.pdf (Accessed November 2014)

DCMII. 2013. *Satellite Imagery Helps Fight Locust Plagues in North Africa*. http://www.dmcii.com/?p=9191 (Accessed November 2014)

Lee, D., Cyganiak, R., Decker S. 2014. *Open Data Ireland: Best Practice Handbook*. Insight Centre for Data Analytics, NUI.

Deloitte. 2012. *Open growth. Stimulating the demand for open data in the UK.* http://www2.deloitte.com/uk/en/pages/deloitte-analytics/articles/stimulating-demand-for-open-data-in-the-uk.html (Accessed November 2014)

EC. 2011. *Open Data. An Engine for Innovation, Growth and Transparent Governance*. Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions. http://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52011DC0882 (accessed November 2014)

ESA. 2013. *Free Access to Copernicus Sentinel Satellite Data.* http://www.esa.int/Our_Activities/Observing_the_Earth/Copernicus/Free_access_to_Copernicus_Sentinel_satellite_ data (Accessed November 2014)

EUMETSAT. 2014. *The purpose of the forum is to reinforce the well established dialogue between EUMETSAT and the African user community to optimise the use of EUMETSAT satellite data and products throughout the continent*. http://www.eumetsat.int/website/home/AboutUs/InternationalCooperation/Africa/UserForum/index.html (Accessed November 2014)

Fonzi, C.J. and Chau, V. 2014. *Smallholder Impact and Risk Metrics: A Labyrinth of Opportunity.* Briefing 03. The Initiative For Smallholder Finance. http://www.globaldevincubator.org/wp-content/uploads/2014/03/Smallholder-Impact-and-Risk-Metrics-A-Labyrinth-of-Opportunity.pdf

GODAN (Global Open Data for Agriculture and Nutrition). 2014. *Statement of Purpose.* http://www.godan.info/statement.html (Accessed November 2014)

Google.org. 2014a. *Earth Engine. A Planetary-Scale Platform for Environmental Data & Analysis.* https://earthengine.google.org/#intro (Accessed November 2014)

Google.org. 2014b. *Explore flu trends around the world.* http://www.google.org/flutrends/intl/en_gb/ (Accessed November 2014)

GOThree60. 2014. *GoThree60*. http://www.gothree60.org/ (Accessed November 2014)

GSMA. 2014. *mAgri Deployment Tracker*. http://www.gsma.com/mobilefordevelopment/programmes/magri/tracker (Accessed November 2014)

Harvard T.H. Chan School of Public Health. 2012. *Using Cell Phone Data to Curb The Spread Of Malaria*. http://www.hsph.harvard.edu/news/press-releases/cell-phone-data-malaria/ (Accessed November 2014)

IBM. 2014. *Enabling the World: IBM Bluemix, the Cloud and Cool Apps*. https://www.ibm.com/developerworks/community/blogs/enablingwithbluemix/entry/let_us_go_change_the_world_with_openness_bluemix_goes_to_the_africa_open_data_jam?lang=en (Accessed 9 February 2015.

INGOaccountabilitycharter. 2014. *INGOaccountabilitycharter*. http://www.ingoaccountabilitycharter.org/ (accessed November 2014)

ISCU (International Soil Moisture Network)–World Data System. 2014. *Trusted Data Services for Global Science.* http://www.icsu-wds.org/ (Accessed November 2014)

ISMN (International Soil Moisture Network). 2014. *ISMN Homepage*. https://ismn.geo.tuwien.ac.at/ (Accessed November 2014)

Klein Tank, A.M.G., Jones, P.D. and Peterson, T.C. 2010. *Creating Surface Temperature Datasets to Meet 21st Century Challenges.* Met Office Hadley Centre, Exeter, UK, 7–9 September 2010 http://www.knmi.nl/~kleintan/www/white_paper5.pdf (Accessed November 2014)

Landers, J.F. 1979. 'Activities of the World Data Center-A.' *Reviews of Geophysics*, 17(7): 1901–1904. http://onlinelibrary.wiley.com/doi/10.1029/RG017i007p01901/abstract (Accessed November 2014)

Lokers, R. and Janssen, S. 2014. *Presentation: 'Putting Food Security Research in Practice at Wageningen UR: Exploiting global knowledge for food-security through cooperation, co-creation and sharing of information.'* ICRI 2014 (www.icri.eu), 3 April 2014

Lundqvist, A., Apers, P., Smeulders, A., Huizer, E. and Mandersloot, P. 2012 *Roadmap ICT for the Top Sectors.* p. 92 www.nwo.nl/binaries/content/.../roadmap-ict-for-the-top-sectors.pdf (Accessed November 2014)

McKinsey Global Institute. 2013. *Open Data: Unlocking Innovation and Performance with Liquid Information.* http://www.mckinsey.com/insights/business_technology/open_data_unlocking_innovation_and_performance_with_liquid_information. (Accessed November 2014)

Maru, A. 2014. *Rights of Farmers for Data, Information and Knowledge. A CIARD e-discussion.* http://www.ciard.net/news-and-events/blog/rights-farmers-data-information-and-knowledge-ciard-e-discussion (Accessed November 2014)

Mutuku, L. and Mahihu, C.M. 2014. *Open Data in Developing Countries: Understanding the Impacts of Kenya's Open Data Applications and Services.*

http://www.opendataresearch.org/sites/default/files/publications/ODDC%20Report%20iHub.pdf (Accessed November 2014)

NASA (National Aeronautics and Space Administration). 2014a. *Earth Science Data. Getting Earth Science Data.* http://science.nasa.gov/earth-science/earth-science-data/ (Accessed November 2014)

NASA (National Aeronautics and Space Administration). 2014b. *NASA Goddard to Bring Satellite Data to African Agriculture.* http://www.nasa.gov/content/goddard/to-bring-satellite-data-to-african-agriculture/#.VD_Jt0dBvqA (Accessed November 2014)

OECD (Organisation for Economic Co-operation and Development). 2007. *OECD Principles and Guidelines for Access to Research Data from Public Funding*. http://www.oecd.org/sti/sci-tech/38500813.pdf (Accessed November 2014)

OGP. 2014. http://www.opengovpartnership.org/ (Accessed November 2014)

Open Access to Scholarly Information. 2014. *Open Access to Data.* http://www.openaccess.nl/whatisopenaccess/openaccesstodata (Accessed November 2014)

Opendatabarometer. 2015. *Opendatabarometer*. Second Edition. http://opendatabarometer.org/. (Accessed January 2015)

Open Definition. 2014. *The Open Definition.* http://opendefinition.org/ (Accessed November 2014)

Open for Change. 2014. *7 Ways to Use IATI Data.* http://openforchange.info/content/7-ways-use-iati-data (Accessed November 2014)

Open Knowledge Foundation. 2012. *The Open Data Handbook.* opendatahandbook.org/ (Accessed November 2014)

Open RBF. 2014. *Managing Data for Smart Financing.* http://openrbf.org/ (Accessed November 2014)

Open Source. 2014. http://opensource.com/resources/what-open-source

Orange. 2014. *Data for Development: Senegal.* http://www.d4d.orange.com/en/agriculture (Accessed November 2014)

Overeem, A., Leijnse, H. and Uijlenhoet, R. 2013. 'Country-wide rainfall maps from cellular communication networks'. *PNAS,* 110(8): 2741–2745. http://www.pnas.org/content/110/8/2741.abstract (Accessed November 2014)

Palantir Blog. 2012. *Grameen Foundation & Palantir: Partners for Food Security.* https://www.palantir.com/2012/10/grameen-foundation-palantir-partners-for-food-security/ (Accessed November 2014)

Precision Agriculture. 2014. *Review: Crop Monitoring Systems.* http://precisionagricultu.re/review-crop-monitoring-systems/ (Accessed November 2014)

Responsible Data Forum. 2014. *Responsibility and Private Sector Data.* https://responsibledata.io/forums/responsibility-private-sector-data/ (Accessed November 2014)

RVO (Netherlands Enterprise Agency). 2014. https://mijn.rvo.nl/gecombineerde-opgave (Accessed November 2014)

Suber, P. 2014. *A Very Brief Introduction to Open Access.* http://legacy.earlham.edu/~peters/fos/brief.htm (Accessed 2014)

The *Guardian*. 2013. *How Might Open Data in Agriculture Help Achieve Food Security?* http://www.theguardian.com/global-development-professionals-network/2013/nov/25/open-data-food-security-agriculture (Accessed November 2014)

UN (United Nations). 2014. *A World That Counts. Mobilising the Data Revolution for Sustainable Development*. Report presented at request of the United Nations Secretary-General by the independent expert advisory group on a data revolution for sustainable development. November 2014. http://www.undatarevolution.org/report/ (Accessed 2014)

United Nations Global Pulse. *2014. Responsible Data Forum on Private Sector Data Sharing – Event Summary* http://www.unglobalpulse.org/RDF-private-sector-data-summary/ (Accessed November 2014)

USAID (United States Agency for International Development). 2014. *USDA Plant Hardiness Zone Map.* http://planthardiness.ars.usda.gov/PHZMWeb/ (Accessed November 2014)

Venkatalakshmi, B. and Devi, P. 2014. 'Decision support system for precision agriculture.' *International Journal of Research in Engineering and Technology,* 3(7): 849–852. http://www.academia.edu/8160091/DECISION_SUPPORT_SYSTEM_FOR_PRECISION_AGRICULTURE (Accessed November 2014)

Verhulst, S.G. 2014. *Mapping the Next Frontier of Open Data: Corporate Data Sharing.* http://www.unglobalpulse.org/mapping-corporate-data-sharing (Accessed November 2014)

Verhulst, S.G. and Sangokoya, D. 2014. *Open Up Corporate Data While Protecting Privacy.* The GovLab, New York University. http://www.openup2014.org/openup-corporate-data-protecting-privacy/ (Accessed November 2014)

Wassenaar, A. 2000. 'E-governmental value chain models: E-government from a business (modelling) perspective.' *In 11th International Workshop on Database and Expert Systems. Applications, 2000, 4–8 Sep 2000, London, UK* (pp. 289–293). http://doc.utwente.nl/55943/ (Accessed November 2014)

Wesolowski, W., Eagle, N., Tatem, A.J, Smith, D.L., Noor, A.M., Snow, R.W. and Buckee, C.O. 2012. 'Quantifying the impact of human mobility on malaria.' *Science,* 338(6104):267–270.

Wikipedia. 2014. *Open Data.* http://en.wikipedia.org/wiki/Open_data (Accessed November 2014)

Wikipedia. 2014. *Precision Agriculture.* http://en.wikipedia.org/wiki/Precision_agriculture (Accessed November 2014)

WMO (World Meteorological Organization). 1995. 'Resolution 40 (Cg-XII)'. *In WMO Policy and Practice for the Exchange of Meteorological and Related Data and Products Including Guidelines on Relationships in Commercial Meteorological Activities.* http://www.wmo.int/pages/about/Resolution40_en.html (Accessed November 2014)

World Bank. 2011. *ICT in Agriculture Sourcebook.* http://www.ictinagriculture.org/content/ict-agriculture-sourcebook (Accessed November 2014)

Wulder, M.A., Masek, J.G., Cohen, W.B., Loveland, T.R. and Woodcock, C.E. 2012. 'Opening the archive: How free data has enabled the science and monitoring promise of Landsat.' *Remote Sensing of Environment*, 122: 2–10. http://www.fs.fed.us/pnw/pubs/journals/pnw_2012_wulder.pdf (Accessed November 2014)

Zagheni, E. and Weber, I. 2012. 'You are where you e-mail: Using e-mail data to estimate international migration rates.' WebSci 2012, 22–24 June 2012, Evanston, IL, USA. (Accessed November 2014)

The Technical Centre for Agricultural and Rural Cooperation (CTA) is a joint international institution of the African, Caribbean and Pacific (ACP) Group of States and the European Union (EU). Its mission is to advance food and nutritional security, increase prosperity and encourage sound natural resource management in ACP countries. It provides access to information and knowledge, facilitates policy dialogue and strengthens the capacity of agricultural and rural development institutions and communities.

CTA operates under the framework of the Cotonou Agreement and is funded by the EU.

For more information on CTA visit, **www.cta.int**

Contact us
CTA
PO Box 380
6700AJ Wageningen
The Netherlands

Tel: +31 317 467100
Fax: +31 317 460067
Email: cta@cta.int

www.facebook.com/CTApage
@CTAflash