CTA
Working Paper
17/09

Farmer profiling: Making data work for smallholder farmers

Stéphane Boyera, Chris Addison and Chipo Msengezi

Series: ICTs for agriculture

CTA Working Paper 17/09 | November 2017

Farmer profiling: Making data work for smallholder farmers

Stéphane Boyera

SBC4D

Chris Addison

CTA

and

Chipo Msengezi

CTA

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 the authors

Stéphane Boyera is a pioneer in web and mobile technologies for social and economic development. Since 2005, Stéphane has been exploring the potential of mobile services to support under-privileged communities in the developing world. His work started at the World Wide Web Consortium, as the founder and manager of the Mobile Web for Social Development Initiative. He subsequently played an important role in the creation of the World Wide Web Foundation. Since then, Stéphane has been leading SBC4D, a software and consulting firm specialised in ICT and mobile for development and in open data in the developing world. For more information on SBC4D, visit www.sbc4d.com

Chris Addison is Senior Programme Coordinator for Data4Ag at CTA. The Data for Agriculture (Data4Ag) project focuses on data use to benefit smallholder farmers. Chris has worked in the ICT and knowledge management (KM) for development sector for the last 18 years and as director of the non-profit One World Europe. He contributed to the OpenAire Open Data report as joint author of the Agriculture chapter. Whilst at IFPRI, he commissioned the conversion of the Global Hunger Index data to linked open data and is currently working on a project to publish the CTA archive as a linked open dataset.

Chipo Msengezi has over eight years' experience in planning and conducting programmes in Sub-Saharan Africa to enhance capacity amongst the research and education communities in the latest information tools and advocate for the adoption of new technologies that drive development in Africa. At CTA, Chipo is responsible for the coordination of the GODAN Action Project, which aims to strengthen data users, producers and intermediaries to engage effectively with open data and maximise its potential for impact in the agriculture and nutrition sectors.

About CTA Working Papers

CTA's Working Papers present work in progress and preliminary findings and have not been formally peer reviewed. They are published to elicit comments and stimulate discussion. Any opinions expressed are those of the author(s) and do not necessarily reflect the opinions or policies of CTA, donor agencies, or partners. All images remain the sole property of their source and may not be used for any purpose without written permission of the source. This document has been produced with the financial assistance of the European Union. The contents of this document are the sole responsibility of CTA and can under no circumstances be regarded as reflecting the position of the European Union.

Please address comments on this Working Paper to Chris Addison (addison@cta.int).

Contents

List of figures

List of acronyms

ACP	African, Caribbean and Pacific Group of States
API	application programming interfaces
DEM	digital elevation model
DSM	digital surface model
DTM	digital terrain model
EMMS	electronic membership management system
FO	farmer organisation
GHG	greenhouse gas
GIS	geographic information system
GPS	global positioning system
ICT	information and communication technology
ID4D	identity for development
IoT	internet of things
IVR	interactive voice response
M&E	monitoring and evaluation
ROSCA	rotating savings and credit association
SaaS	software as a service
SACCO	savings and credit cooperative organisation
SDG	sustainable development goal
USSD	unstructured supplementary service data

Introduction

The study presented in this report was commissioned by the Technical Centre for Agricultural and Rural Cooperation[1] (CTA) as a member of the Global Open Data for Agriculture & Nutrition[2] (GODAN) initiative, and was conducted by SB Consulting[3] (SBC4D). The objective of the research is to understand the role of farmer organisations (FO) and cooperatives in the agriculture data ecosystem.

These organisations have long been recognised to play an important role in society that translates into the improvement of living conditions of their members, particularly the low-income earning population. More than 40% of households in Africa are member of a cooperative society ([ILO-2000]) and the cooperative movement is Africa's biggest non-governmental organisation. The key question this report explores is the role of these organisations in the emergent "data revolution." How can they ensure that this data revolution benefits their members and the smallholder farmers in general, and at the same time contribute to the revolution by providing valuable information to policy makers or other stakeholders of the ecosystem?

This research expands and complements other GODAN research published recently, such as CTA's research on Open Data and Smallholder Food and Nutritional Security ([Jellema, Meijninger, Addison – Feb2015]) and the Agriculture Open Data Package ([Jellema 2016]).

This document includes 3 main sections. The first one provides the overall context of the research, including an overview of smallholder farmers' role and challenges, the methodology followed in this research, and the scope of the study. The second section details for each stage of the crop cycle, the findings in terms of data use, needs, and challenges. The third section is a summary of the findings in the form of recommendations for farmer organisations and cooperatives in terms of farmer profiling activity. Finally, the document contains an acknowledgement section, a conclusion and a series of annexes including the bibliography and references used.

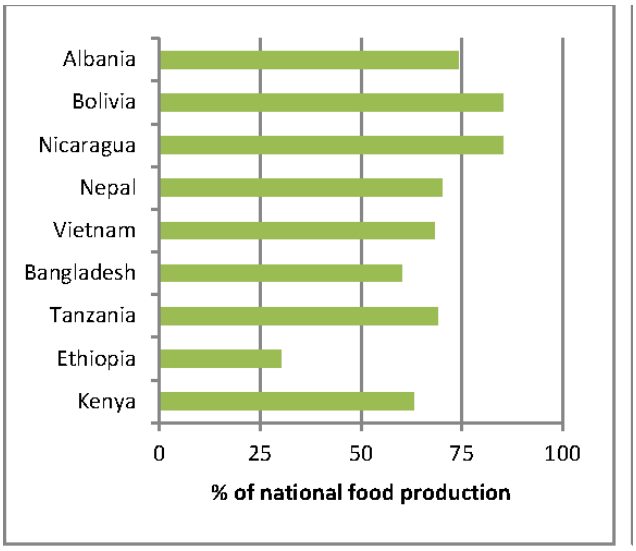
[1] http://www.cta.int
[2] http://www.godan.info/
[3] http://www.sbc4d.com

Section I: Context

Background: Smallholder farmers' role and challenges

As the main role of FOs and cooperatives is to support their members who are smallholder farmers, it is essential to understand the role of these actors in the overall economy and the challenges they are facing, to understand the potential impact on the data ecosystem. This section provides a quick data snapshot of smallholder farmers. The information presented below is mainly coming from the FAO Smallholder Farmers' Dataportrait[4] ([Rapsomanikis 2015]).

Smallholder farmers represent the biggest employment sector in rural areas of the developing world and they are also the most important contributors to the global food production. More than 90% of the farms in the world are family farms; they produce 80% of the food and they operate 75% of the farm land ([FAO-SOFA-2014]). Seventy-two per cent of farms are less than 1 ha ([FAO-SOFA-2014]). The figure below shows data from FAO Smallholder Farmers' Dataportrait ([Rapsomanikis 2015]).

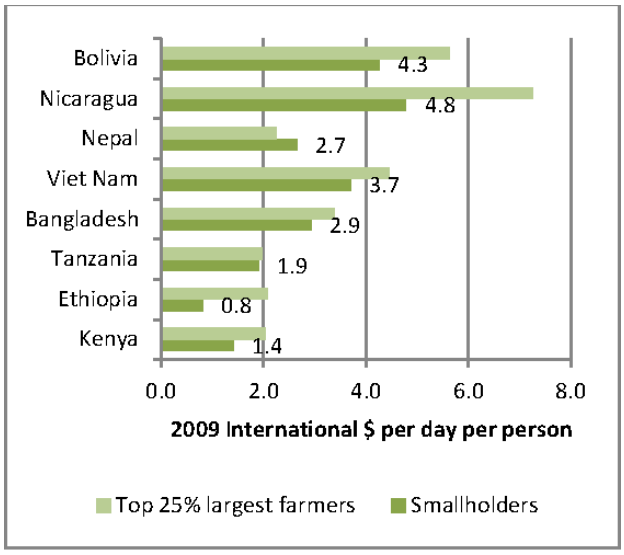


Source: Smallholder Farmers' DataPortrait.

Figure 1: Proportion of national food production by smallholders

However, at the same time, smallholder farmers are also part of the poorest category of most developing nations. The figure below ([Rapsomanikis 2015]) shows that smallholders earn less than US$2 a day in many countries in Sub-Saharan Africa, and sometimes even less than US$1 (e.g. in Ethiopia).

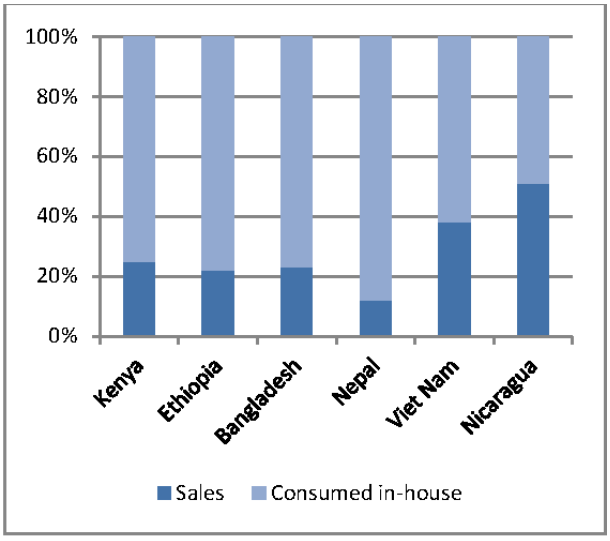
[4] http://www.fao.org/family-farming/data-sources/dataportrait/farm-size/en/



Source: Smallholder Farmers' DataPortrait.

Figure 2: Income, US$ per person and per day

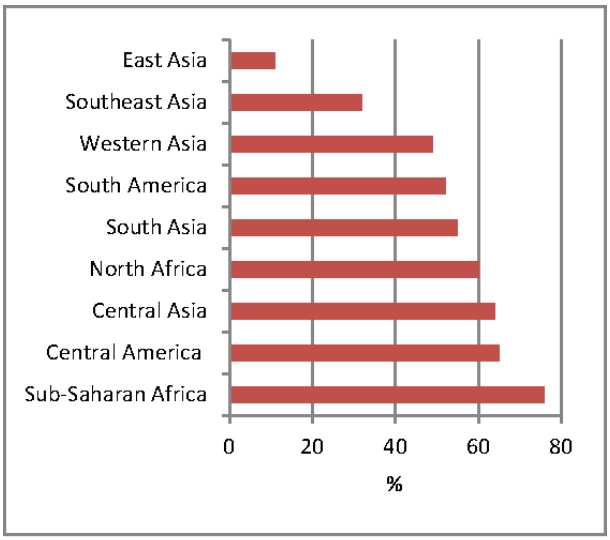
There is unlikely one specific challenge or reason that can explain this situation. However, two data points are interesting to consider. The first is related to the income generated by yields. Income in terms of US dollars comes from trading. The figure below ([Rapsomanikis 2015]) shows the percentage of traded/sold yield versus consumed yield by the famer household.



Source: Smallholder Farmers' DataPortrait.

Figure 3: Smallholder agricultural production sold in markets and consumed in-house

Figure 3 shows that in Sub-Saharan African countries, just twenty percent of the production is sold. It is likely that the more yields are sold, the higher the income for households is. The second interesting data point is related to crop productivity. The figures below from various sources characterise crop productivity in Sub-Saharan Africa. They show that yields in Africa could be increased by up to 75% compared to the best practices.



Source: FAO (2011).

Figure 4: Percentage of yield gaps per region

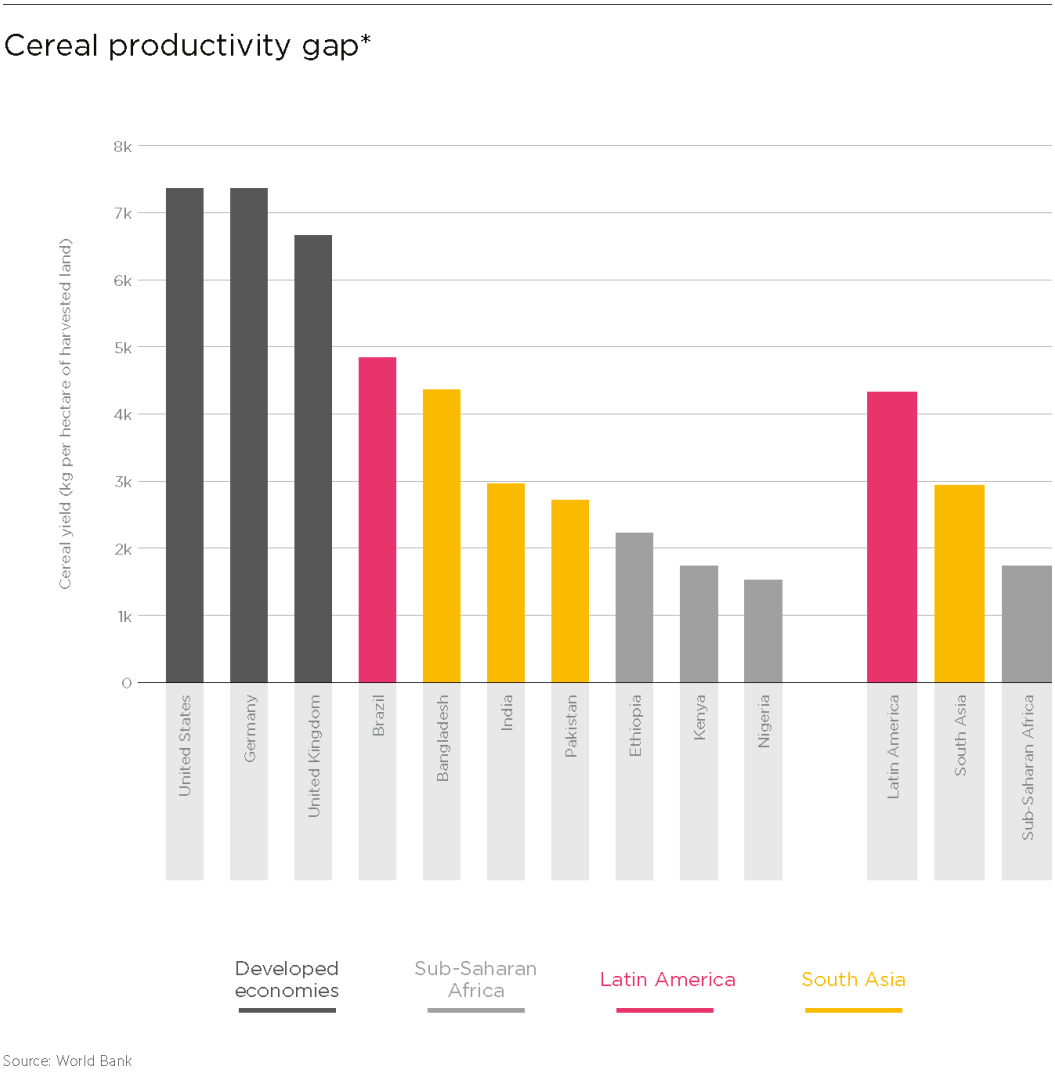


Figure 5: Cereal yield gap (GSMA Agricultural Value-added Services Toolkit 2.0 2016 [GSMA-MAGRIVAS-2016])

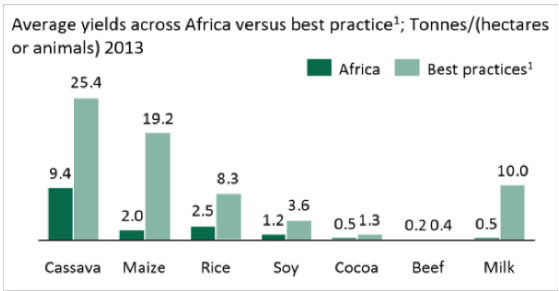


Figure 6: Average yields across Africa versus best practice
(ADB Feed Africa: Strategy for Agricultural Transformation in Africa 2016-2025 [ADB-FEEDAFRICA-2016])

Based on these two data points, it is likely that if the yields increase, not only will the food insecurity risk decrease but also the proportion of the crops sold will also increase, leading to more income for smallholder farmers. The key question is to understand the root cause of the low yield. The literature details a lot of possible reasons, from lack of access to specific instruments (credit, insurance), to issues related to climate change that make traditional ancestral knowledge not as efficient as it used to be, to access of appropriate information to take informed decisions at the right time. That said, the hypothesis behind the "data revolution" is that the provision of and access to information and data by all stakeholders in the value chains will help solve a significant part of these issues. For example, understanding the cause of underperformance of some crops will lead policy makers to put in place appropriate legislations, subsidy schemes and interventions to address these issues. Access to detailed field information will help credit companies build reliable credit profiles and deliver loans to smallholder farmers more easily. At the same time, if farmers, at each stage of the crop cycle, can access timely actionable information, they will be able to take informed decisions on the best way to get the most of their fields in a sustainable, eco-friendly way.

For most stakeholders, the equation resides on the mash-up of global data (satellite images, research studies, databases with information about crops, seeds, pests and diseases, etc.) with farmer-level (credit records, field ownership documentations, etc.) and field-based information (soil information, geographic location, state of the fields, crops, etc.) to determine the appropriate information to take a decision. The results for the farmer are both the availability of new products to support their production (credit, insurance, etc.) and the availability of timely information to support decision-making. The figure below summarises this view.

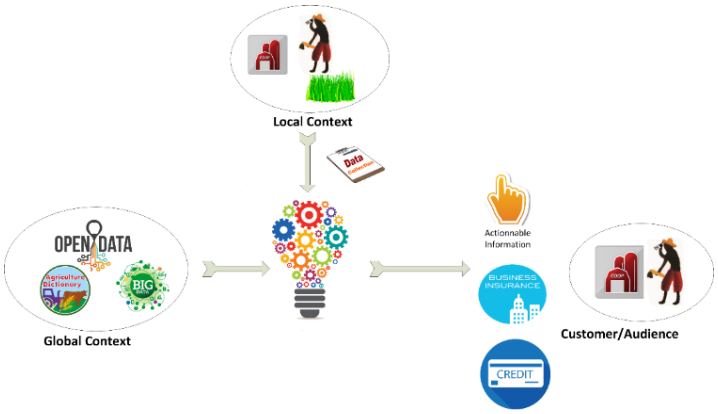


Figure 7: Mashing-up various sources of information to deliver actionable information or key instruments to farmers

The angle of attack of this research is to look at both the global and local context, to understand the various types of data that are potentially useful to support farmers, either directly or indirectly by supporting organisations in charge of providing specific instruments (credit, insurance, etc.) to farmers. In the analysis, we will cover the potentially useful datasets, the data that should be collected in the field as well as the technologies that can be used to collect this data, and the technologies that can be used to then disseminate the results to farmers.

Our focus in this study is on data collected at the field level, and that could be considered as the "farmer profile", i.e. the list of information collected about the farmer, the farm and the attached fields. This information is essential for various stakeholders in the farmer context. It allows them to provide farmers with the right instruments and/or the right actionable information at each stage of the crop cycle to increase yields and market access for products.

Scope of the study

The previous section described the global approach of the study. However, to make a deep analysis of each stage of the crop cycle, and to have not too high-level output, but rather precise recommendations, we focus our study on specific elements. The first one is on the type of products. This study focuses on crops and will therefore not cover livestock and aquaculture. There are a series of new developments in these sectors like the recently launched Fisheries Transparency Initiative[5] (FiTI) that aims toward publishing numerous data on fisheries and fishery industries. However, the life cycle of these products as well as the challenges of the value chains are slightly different compared to the crop life cycle.

In the same way, we focus on seasonal crops, and we do not cover multi-year crops, including trees. In this area, there are several innovations, but the challenges are of a different nature compared to seasonal crops. We will also not cover traceability as part of this study, except under the lens of certification. Traceability is a process that takes place during each step in the value chain, while this study is mainly focusing on the farmer level. We will therefore cover the requirements for farmers to acquire certifications, but no other steps of traceability.

In terms of geography, the primary focus of this study is on Sub-Saharan Africa, and developing regions in general. A few examples and research exploited in this study cover the African, Caribbean and Pacific (ACP) region.

Finally, while this study is executed in the context of GODAN, the focus is not primarily on open data, but more specifically on the gap between opening data and exploiting these data to design and deliver innovative services to improve farmers' lives either directly (services targeting farmers) or indirectly (better policies, etc.).

Methodology

The methodology used in this study is structured around three elements:

- The first phase focused on desk research. This phase consisted of the review of studies on two main subjects:

[5] http://fisheriestransparency.org/

- Open data and agriculture: This included the research commissioned under the GODAN initiative, as well as the various studies and data published by international organisations such as the Food and Agriculture Organization of the United Nations (FAO) and CTA. See the section on open data below.
- ICT for agriculture: We reviewed research and compendia on the role of ICT in agriculture, as well as numerous examples of services across the world with a focus on Sub-Saharan Africa. This allowed us to have a global view on technologies used for data collection and service delivery, as well as information collected on the ground, related challenges and strategies developed to address them.

We also included reviews of farmer registry initiatives or more globally identity initiatives at local, national, regional and international levels in the desk research.

- The second phase was dedicated to interviews and surveys. We conducted a series of individual interviews with experts in the field, farmer organisations as well as specific service providers, to understand the type of data they use, the impact and the challenges (see the acknowledgement section for the list of interviewees).
  We have also used a CTA survey in Madagascar covering 100 farmer organisations, to gather specific information related to farmer profiling to inform our research (see Annex 2 for the list of questions).
- The third phase was dedicated to the compilation of all elements gathered during the first two phases to build a global overview presented in Section II, as well as a series of recommendations that are presented in Section III.

Open data and agriculture overview

This section describes the list of datasets that we have considered when reviewing farmer data needs during each stage of the crop cycle. This list is mainly based on the Agriculture Open Data Package ([Jellema 2016]) plus a few additions from other sources, such as the presentations made at the big data session of the USDA/USAID International Food Assistance and Food Security Conference ([USDA/USAID 2016]). The list of datasets is not exhaustive, and it was not part of the focus of this study to explore this dimension, but instead built on the work already done.

As part of the analysis presented in the second section of this report, we have also characterised datasets depending on their use. For example, the same datasets might be useful at different stages of the crop cycle, but with different requirements. One such example is market prices. Market prices are useful when selecting the crop to grow, if there is a time series showing the recent evolution of prices in the past years. Market prices are also useful at the selling stage, but at this stage, the dataset must have near real-time information to be useful.

The table below lists the datasets we have considered, and describes their content.

Data category		Key datasets	Definition
Government, agricultural law and regulations	1	(Phyto)sanitary regulations (list of quarantine organisms, etc.)	Phytosanitary quarantine measures promoting protection of plants from penetration, occurrence and distribution of pests, illnesses and weeds.
	2	Environmental regulations	Self-explanatory - these are regulations/rules and requirements that generally cover the following: Pollution control: regulating how much pollution (chemicals or other undesirable materials such as "heat", "suspended particulates") a facility releases. In practice this is an amalgam of state and federal statutes, regulations, and common-law principles covering air pollution, water pollution, hazardous waste, the wilderness, and endangered wildlife.
	3	Subsidy schemes	An agricultural subsidy is a governmental subsidy paid to farmers and agribusinesses to supplement their income, manage the supply of agricultural commodities, and influence the cost and supply of such commodities. In some countries subsidies apply to specific lands or are based on no specific land size.
	4	Import/export regulations	Regulations governing trade in food and agricultural products.
Official records	5	Land registration	Land registration generally describes systems by which matters concerning ownership, possession or other rights in land can be recorded (usually with a government agency or department) to provide evidence of title, facilitate transactions and to prevent unlawful disposal.
	6	Licensed organisations (corporations, businesses, NGOs)	Organisations in the profit and not for profit sector certified and registered to work in the agriculture sector/supply chain.
	7	Import/export tariffs	Tariffs governing trade in food and agricultural products.
	8	Permitted crop protection products	Certified products that are legally permitted.

Government finance data	9	Agricultural subsidy expenditure (direct payments, product support, tariffs etc.)	Agricultural subsidy is a governmental subsidy paid to farmers and agribusinesses to supplement their income, manage the supply of agricultural commodities, and influence the cost and supply of such commodities.
	10	Agriculture-related tax income	Crop and related income/payments that are taxable. This also includes all farm income, commodity, credit/loan included as income.
	11	Penalties given to agricultural actors	This is usually related to the oversight of federal/state agencies' use of penalties set out in the agriculture and agri-food policies. These may include cases involving agriculture, food and animal handling and transportation, violation for bringing animal or plant products into the country without permission or where a producer, transporter or handler of animals or plants violates set standards. These may also include other penalties such as polluting common water bodies, use of unauthorised fertilisers and pesticides, burning of residual crops etc., depending on the law of the land.
	12	Investment in research and education (extension, research institutes, professional training and universities)	Investments in agriculture research, training, dissemination, extension etc., primarily conducted by universities and agriculture departments.
Rural development project data	13	General project information, including financial data, location, beneficiaries, activities	These datasets relate to project specific information that may include financial data, geographic location, activities and linked outcomes and impact. This may also include data on how the land was used and for which crop till the harvesting stage.
	14	Project output, outcome and impact	
	15	Project baseline and survey data	
	16	Project documents	
Land use and productivity data	17	Land use data	
	18	Cultivated areas	
	19	Current crop in the fields	
	20	Harvested crop	
	21	Crop types	

Value chain data	22	Profiles of different value chain actors and organisations [1] Farm data, e.g., farming system, crops, land area, farm income, household composition, farm employment, farm holder's age, fertiliser use, etc. [2] Cooperatives [3] Trade [4] Processors, e.g. type, size, turnover, capital, investments, environmental transparency indicators etc. [5] Retail	Datasets on farming system, crops, land area, farm income, household composition, farm employment, farm holder's age, fertiliser use, etc. This may also include data held by government and non-government organisations, cooperatives and various other trade bodies.
	23	(Food) product data, e.g. food nutritional value, food composition, origin of produce, environmental factors, time and location of production, etc.	Datasets mostly used for marketing purposes that may include origin of produce and other attributes that may certify a way of growing and harvesting
	24	(Safety) inspection results	Audit reports of safety inspections carried out in the agriculture supply chain
	25	Certification	There are many different types of certifications. Farmers who wish to achieve certification status in a specific production or management practice must meet and maintain a certification programme's unique set of standards or requirements. Certified farms usually brand and market their products with the programme's certified label. Therefore, certification dataset = data on different types of certifications. These may be related to certification status in a specific production or management practice and must meet and maintain a certification programme's unique set of standards or requirements. Certified farms usually brand and market their products with the programme's certified label. http://www.standardsmap.org/ https://www.nal.usda.gov/afsic/farm-certifications

Infrastructure data	26	Road network and conditions	Datasets related to the road network infrastructure and its state.
	27	Road maintenance schedule	Road maintenance schedule.
	28	Public transport	Data related to public transport - this may include timetables, route information, etc.
	29	Waterways	Waterways connectivity information.
	30	Internet connectivity map	Internet connectivity availability information.
	31	Mobile connectivity map	Mobile connectivity availability information.
Market and price data	32	Global food prices	Datasets related to global food prices.
	33	National stock exchange prices	Datasets related to national stock exchange prices.
	34	Regional market prices	Datasets related to regional market prices.
	35	Local market prices	Data related to local market prices.
	36	Location of national markets	Location information for major national markets.
	37	Location of regional markets	Location information for major regional markets.
	38	Location of local markets	Location information for local markets.
	39	Import/export volume	Data related to the quantum of exports and imports.
Meteorological data	40	Short-term weather forecast	Weather forecast related to a very specific time-period. This may be daily or weekly forecasts.
	41	Seasonal weather forecasts (3-6 months ahead)	Six monthly weather forecast data.
	42	Real-time observations	Real-time data as observed daily.
	43	Historic archives of observations	Archived data of weather conditions across a region.
	44	Historical simulated weather from re-analysis	Use of data to simulate a weather condition based on archived data.
	45	Climatological observations	Data related to climatic conditions of a particular region.
	46	Climatological reference data	Historical climate data. For example, the U.S. Climate Reference Network (USCRN) is a systematic and sustained network of climate monitoring stations with sites across the conterminous U.S., Alaska, and Hawaii. These stations use high-quality instruments to measure temperature, precipitation, wind speed, soil conditions, and more.
	47	Climate zones	Information on climate zones. This may include latitude, terrain, and altitude, as well as nearby water bodies and their currents.

	48	Climate change predictions	Socioeconomic scenarios are used by analysts to make projections of future greenhouse gas (GHG) emissions and to assess future vulnerability to climate change. Producing scenarios requires estimates of future population levels, economic activity, the structure of governance, social values, and patterns of technological change. Economic and energy modelling (such as via the World3 or the POLES models) can be used to analyse and quantify the effects of such drivers.
Elevation data	49	Digital elevation model	A digital elevation model (DEM) is a digital model or 3D representation of a terrain's surface — commonly for a planet (including earth), moon, or asteroid — created from terrain elevation data. There is no universal usage of the terms digital elevation model, digital terrain model (DTM) and digital surface model (DSM) in scientific literature. In most cases the term digital surface model represents the earth's surface and includes all objects on it. In contrast to a DSM, a DTM represents the bare ground surface without any objects like plants and buildings.

	50	Elevation maps	The elevation of a geographic location is its height above or below a fixed reference point, most commonly a reference geoid, a mathematical model of the earth's sea level as an equipotential gravitational surface. Elevation, or geometric height, is mainly used when referring to points on the earth's surface, while altitude or geopotential height is used for points above the surface, such as an aircraft in flight or a spacecraft in orbit, and depth is used for points below the surface. A topographical map - or an elevation map is the main type of map used to depict elevation, often through use of contour lines. In a geographic information system (GIS), digital elevation models are commonly used to represent the surface (topography) of a place, through a raster (grid) dataset of elevations. Digital terrain models are another way to represent terrain in GIS.
	51	Height points	On an elevation map this is a contour layer with height info. E.g. the height from contour layer to the point data for each location.
	52	Slope	Reference definition: Esri Support GIS Dictionary (Euclidean geometry). The incline, or steepness, of a surface. Slope can be measured in degrees from horizontal (90), or percent slope (which is the rise divided by the run, multiplied by 100). A slope of 45 degrees equals 100% slope.

	53	Aspect	Reference definition Arc GIS: Aspect identifies the downslope direction of the maximum rate of change in value from each cell to its neighbours. It can be thought of as the slope direction. The values of each cell in the output raster indicate the compass direction that the surface faces at that location. It is measured clockwise in degrees from 0 (due north) to 360 (again due north), coming full circle. Flat areas having no downslope direction are given a value of -1. With the aspect tool, one can do the following: - Find all north-facing slopes on a mountain as part of a search for the best slopes. - Calculate the solar illumination for each location in a region as part of a study to determine the diversity of life at each site. - Find all southerly slopes in a mountainous region to identify locations where the snow is likely to melt first as part of a study to identify those residential locations likely to be hit by runoff first. - Identify areas of flat land.
	54	Catchments	In human geography, a catchment area is the area from which a city, service or institution attracts a population that uses its services. For example, a school catchment area is the geographic area from which students are eligible to attend a local school. Governments and community service organisations often define catchment areas for planning purposes and public safety such as ensuring universal access to services like fire departments, police departments, ambulance bases and hospitals.

	55	Drainage	Drainage is the natural or artificial removal of surface and sub-surface water from an area. The internal drainage of most agricultural soils is good enough to prevent severe waterlogging (anaerobic conditions that harm root growth), but many soils need artificial drainage to improve production or to manage water supplies.
	56	Erosion susceptibility	Erosion susceptibility is defined by the predisposition (of a land unit) to erode, preparatory factors (such as the removal of forest), the likelihood and severity of an erosion event, and the consequences of an erosion event.
Hydrological data	57	Location of water sources	Dataset related to the location of water bodies. These may include information on ponds, wells, stop dams, rivulets, rivers, etc.
	58	Flood zones	Flood zones are land areas identified by the federal and state authorities in terms of its risk of flooding.
	59	Historical records on flooding	Archived data of flooding in a specific geographic area.
	60	Real-time water levels	Data on water levels of a specific water body.
	61	Water quality	Data on water quality – this may include the chemical, physical, biological, and radiological characteristics of water. It is a measure of the condition of water relative to the requirements of one or more biotic species and/or to any human need or purpose. It is most frequently used by reference to a set of standards against which compliance can be assessed. The most common standards used to assess water quality relate to health of ecosystems, safety of human contact, and drinking water.
	62	Water tables	Data related to the upper surface of the zone of saturation. The zone of saturation is where the pores and fractures of the ground are saturated with water.
	63	Water management	Data related to water use efficiency and its sustainability.

Soil data	64	Soil maps	A soil map is a map, i.e. a geographical representation, showing diversity of soil types and/or soil properties (soil pH, textures, organic matter, depths of horizons, etc.) in the area of interest. It is typically the end result of a soil survey inventory, i.e. soil survey. Soil maps are most commonly used for land evaluation, spatial planning, agricultural extension, environmental protection and similar projects. Traditional soil maps typically show only general distribution of soils, accompanied by the soil survey report. Many new soil maps are derived using digital soil mapping techniques. Such maps are typically richer in context and show higher spatial detail than traditional soil maps. Soil maps produced using (geo)statistical techniques also include an estimate of the model uncertainty.
	65	Soil samples	Soil sample datasets that may include information on nutrient content, composition, and other characteristics such as the acidity or pH level.
	66	Soil classifications	Soil classification is the separation of soil into classes or groups, each having similar characteristics and potentially similar behaviour. A classification for engineering/data purposes should be based mainly on mechanical properties, e.g. permeability, stiffness, strength.
Production advice data	67	Data on cultivars, landraces and farmer varieties including new releases	Data on [1] Cultivars = Plant varieties that have been produced in cultivation by selective breeding; [2] Landraces = Domesticated, locally adapted, traditional varieties of a species of animal or plant that have developed over time, through adaptation to its natural and cultural environment of agriculture and pastoralism, and due to isolation from other populations of the species; and other farmer varieties.

	68	Crop selection advice including new releases	Advisory data on crop selection including information on new crop options.
	69	Crop calendars	The crop calendar is a tool that provides timely information about seeds to promote local crop production. It contains information on planting, sowing and harvesting periods of locally adapted crops in specific agro-ecological zones.
	70	Intercropping, relay cropping, rotation	Data on [1] Crop rotation = The system of varying successive crops in a definite order on the same ground, especially to avoid depleting the soil and to control weeds, diseases, and pests; [2] Intercropping = Multiple cropping practice involving growing two or more crops in proximity. The most common goal of intercropping is to produce a greater yield on a given piece of land by making use of resources or ecological processes that would otherwise not be utilised by a single crop; [3] Relay cropping = Essentially a special version of double cropping, where the second crop is planted into the first crop before harvest, rather than waiting until after harvest as in true double-cropping.
	71	Resource-related farm advice	Data related to crop selection, crop and land management as typically found in extension services information.
	72	Fertiliser recommendations	Data related to fertiliser recommendation. This usually sets out fertiliser recommendations for agricultural and horticultural crops.
Disease and pest management data	73	Occurrences and distribution of plant pests	This data usually consists of type of pests and the associated diseases and toxicology data including measures to mitigate the same.
	74	Treatment of pests and diseases	
	75	Recognition of pests and diseases	
	76	Biology of pests and diseases	
	77	Toxicology or plant protection measures	

Section II - Study findings

This section presents the study findings. In the first part, we will introduce the crop cycle we adopted, then we will present for each stage the specific data needs that can support farmers, and finally we will summarise our findings related to ICT usage across the various stages and activities. Each stage of the cycle follows the same structure that introduces the farmer objectives, the relevant global data that can enable the implementation of the objectives, and the relevant farmer and farm data required to exploit the global data. Such mapping can also be used by service providers to better understand farmer needs and customise their offerings.

Crop cycle

As presented in the previous section, our approach is based on the analysis of data needs at each stage of a crop cycle. Each crop and product have their own cycle and the literature introduces many different models for a crop cycle. The one we have adopted and that is summarised in the figure below is inspired from the crop cycle used in [DELOITTE-2012] and has been adapted during the research based on our findings in terms of data needs. This proposed cycle does not fit all crops, but is generic enough to provide a framework that can then be instantiated for specific cases.

Figure 8: Crop cycle adopted for the research

Key datasets and farmer profile information mapping

In this section, we explore the mapping between key datasets presented in the first part of this document (Open data and agriculture overview) and farmer profile information. The objective here is to identify information available at the farmer/farm level that is essential to exploit datasets and extract information that is relevant to the specific farmer. Such mapping enables us to identify the various elements of the farmer profile that are useful at different stages of the crop cycle. Such mapping can also be used by service providers to better understand farmer needs and customise their offerings.

It is important to note that there are some common elements in the farmer profile, such as communication and personal information that appear across the life cycle. These are basic data elements required to (a) identify the farmer and (b) enable and establish a communications link with the farmer. Because of this fundamental requirement, this profile information is not repeated across sections.

In the table below, the first column depicts the crop life cycle stage and the intended activity (farmer activity or objective for an advisory service provider). The second column correlates from the available datasets and lists the ones that correspond to the profile mapping. The last column highlights the key profile information that is essential to exploit datasets, and extract relevant information to support farmers with their activities.

Service/activity objective	Key datasets	Farmer profile mapping
Phase: Land and crop selection [1] Determining ownership/rights to use the land	1. Land registration	On the farmer's side, information to map datasets is: [1] Farm registration number: In some countries, farm registration includes ownership of land [2] Field information o Location: The location of the field can be used to identify the owner if a national database is available o Land title: If the farmer has a land title, details of land are usually available of the certificate
[2] Ensure land use rights to cultivate the (agriculture) land	1. Land use data	To know whether a piece of land can be cultivated, a farmer should investigate for each of his/her field: o Location: The location of a field can be used to check on a national database of land use whether the land can be cultivated o Land title: In some countries, the land title includes the land tenure
Phase: Land and crop selection **Activity/service**: Identify the crops and the varieties to grow based on legal framework and subsidy schemes	1. (Phyto)sanitary regulations (list of quarantine organisms, etc.) 2. Environmental regulations 3. Subsidy schemes 4. Permitted crop protection products 5. Land use data 6. Current crop in the fields 7. Harvested crop 8. Crop types	Information required to identify crops that could be used based on regulation and based on subsidy scheme are: [1] Location: o Farmer and farm location: This information may allow specific subsidies o Field information: Based on the location and size, fields may have restrictions on crop, and/or may lead to access to specific subsidies [2] Financial data o Income level: Some subsidies are linked to the farmer's income
Phase: Land and crop selection **Activity/service**: Identify the	1. Seasonal weather forecasts (3-6 months ahead) 2. Weather: Real-time observations 3. Climatological reference data 4. Climate zones 5. Soil maps and soil samples	The selection of crops depends on field information: o Location: This is essential to determine weather forecast o Elevation: Some crops have requirements on altitude and on the shape of the field o Size: Some crops have constraints on size (maximum or minimum size)

crops and the varieties to grow based on local conditions such as climate zone; agroecological zone; weather forecast, soil or global appropriateness of the field (e.g. flooding risks)	6. Data on cultivars, landraces and farmer varieties, including new releases 7. Crop selection advice, including new releases 8. Crop calendars 9. Intercropping, relay cropping and rotation 10. Elevation maps 11. Height points 12. Slope 13. Aspect 14. Catchments 15. Drainage 16. Erosion susceptibility 17. Location of water sources 18. Flood zones	○ Soil: This is critical to identify which crops can be grown ○ Crop history: Crop history may have an impact on soil and therefore on crop selection
Phase: Land and crop selection **Activity/service**: Identify the crops and the varieties to grow based on the availability of inputs	1. Subsidy schemes 2. Profiles of different value chain actors and organisations a. Farm data, e.g. farming system, crops, land area, farm income, household composition, farm employment, farm holder's age, fertiliser use, etc. b. Cooperatives c. Trade d. Processors, e.g. type, size, turnover, capital, investments, environmental transparency indicators, etc. e. Retail 3. Certification 4. Fertiliser recommendations 5. Occurrences and distribution of plant pests 6. Treatment of pests and diseases 7. Recognition of pests and diseases 8. Biology of pests and diseases 9. Toxicology or plant protection measures	The selection of crops depends on the local availability of inputs (seeds, treatments, fertilisers). This is linked to: [1] Field information ○ Location ○ Size [2] Local availability: Farmer's linkage to potential provider of input ○ Cooperatives/production cluster ○ Agro-dealer linkages

Phase: Land and crop selection **Activity/service**: Identify the crops and the varieties to grow based on the availability of other elements essential for cultivation (water, machines, materials, transformation/ processing equipment, etc.)	1. Land registration 2. Erosion susceptibility 3. Location of water sources 4. Flood zones 5. Historical records on flooding 6. Real-time water levels 7. Water quality 8. Water tables 9. Water management	The selection of crops by the farmer in this category depends on: [1] Field information o Location o Size o Soil: Water usage depends on the state of the soil [2] Farm details o Manpower on the farm (availability of labour force) o Equipment ▪ Planting ▪ Harvesting ▪ Post-harvesting
Phase: Land and crop selection **Activity/service**: Identify the crops and the varieties to grow based on the potential of the crop/variety (length of the cycle, number of cycles/year,	1. Food product data, e.g. food nutritional value, food composition, origin of produce, environmental factors, time and location of production, etc. 2. Data on cultivars, landraces and farmer varieties, including new releases 3. Crop selection advice, including new releases 4. Crop calendars 5. Intercropping, relay cropping and rotation 6. Profiles of different value chain actors and organisations a. Farm data, e.g. farming system, crops, land area, farm income, household composition, farm employment, farm holder's age, fertiliser use, etc. b. Cooperatives	The length of a cycle as well as the number of cycles and global output (yield, income) depend on: [1] Field information o Location o Size [2] Certification information: Which certification the farmer is targeting for a given field [3] Materials and resources on the farm o Manpower on the farm (availability of labour force) o Equipment [4] Post-harvest information o Warehouse access and characteristics o Processor linkages ▪ Position of processing equipment for transport

nutritional value, output)	c. Trade d. Processors, e.g. type, size, turnover, capital, investments, environmental transparency indicators, etc. e. Retail 7. Certification 8. Global food prices 9. National stock exchange prices 10. Regional market prices 11. Local market prices 12. Location of national markets 13. Location of regional markets 14. Location of local markets 15. Import/export volume	○ Post-harvesting processing materials ○ Post-harvesting transformation ○ Storage capacity and characteristics ○ Transport service availability [5] Selling opportunities the farmer has access to ○ Cooperatives/production cluster membership ○ Markets ○ Agribusinesses linkages
Credit services	1. Subsidy schemes 2. Land registration 3. Agricultural subsidy expenditure (direct payments, product support, tariffs, etc.) 4. Project documents 5. Land use data 6. Cultivated areas 7. Current crop in the fields 8. Harvested crop 9. Crop types 10. Certification 11. Climate change predictions 12. Seasonal weather forecasts (3-6 months ahead) 13. Weather: Real-time observations 14. Climatological reference data 15. Climate zones 16. Digital elevation model 17. Elevation maps	Farmers have access to credit based on information similar to the information they used to select the crop and evaluate the output: [1] Field information ○ Location ○ Size ○ Elevation ○ Soil ○ Land tenure/land title [2] Crop data ○ Seeds ○ Variety ○ Cost [3] Farm details ○ Manpower on the farm (availability of labour force) ○ Equipment ▪ Planting ▪ Harvesting

	18. Height points 19. Slope 20. Aspect 21. Catchments 22. Drainage 23. Soil maps 24. Soil samples 25. Soil classifications 26. Crop selection advice, including new releases 27. Intercropping, relay cropping and rotation 28. Resource-related farm advice 29. Fertiliser recommendations	▪ Post-harvesting [4] Production information o Yield forecast o Qualification and certification [5] Post-harvest information o Warehouse access and characteristics o Processor linkages ▪ Position of processing equipment for transport o Post-harvesting processing materials o Post-harvesting transformation o Storage capacity and characteristics o Transport service availability [6] Selling opportunities the farmer has access to o Cooperatives/production cluster membership o Markets o Agribusinesses linkages [5] Farmer's financial data o Income o Subsidy programme and amount [6] Insurance information o Insurance details and conditions of realisation [7] Credit information and account data o Credit records o Farm business plan o SACCOs/ROSCAs membership o Credit allocation per activity [8] Financial instrument to receive credit and to pay reimbursement o Bank account(s) o Mobile money account(s)

Insurance services	1. Phytosanitary regulations – list of quarantine organisms, etc. 2. Investment in research and education (extension, research institutes, professional training and universities) 3. Current crop in the fields 4. Flood zones 5. Climate change predictions 6. Seasonal weather forecasts (3-6 months ahead) 7. Weather: Real-time observations 8. Climatological reference data 9. Climate zones 10. Resource-related farm advice 11. Fertiliser recommendations 12. Occurrences and distribution of plant pests 13. Treatment of pests and diseases 14. Recognition of pests and diseases 15. Biology of pests and diseases 16. Toxicology or plant protection measures	Access to insurance services is linked to information about first evaluating the risks, and then identifying means to access the service (payment of the insurance, and reimbursement in the event of covered risks): [1] Field information o Location o Size [2] Crop data (crop, varieties, seeds) [3] Qualification and certification information [4] Farmer's linkage to agro-business to identify insurance delivery options o Cooperatives/production cluster membership o Markets farmers have access to o Agro-businesses linkages [5] Financial instrument to receive credit and to pay reimbursement o Bank account(s) o Mobile money account(s)
Crop-cultivation and harvesting: Planting phase	1. Harvested crop 2. Crop types 3. Certification 4. Data on cultivars, landraces and farmer varieties, including new releases 5. Crop selection advice, including new releases 6. Fertiliser recommendations 7. Climate change predictions 8. Seasonal weather forecasts (3-6 months ahead) 9. Weather: Real-time observations 10. Climatological reference data 11. Climate zones	The farmer's objective is to determine the best planting date. This is largely dependent on weather forecast linked to the crop specificities, and the work needed. At the farmer's level this is linked to: [1] Field information o Location o Size [2] Crop data (crop, varieties, seeds) [3] Certification information: Certification usually requires a specific planting approach (number of seeds, etc.) [4] Farm details o Manpower on the farm (availability of labour force) o Planting equipment

Crop-cultivation and harvesting: Field management and monitoring	1. Certification 2. Fertiliser recommendations 3. Occurrences and distribution of plant pests 4. Treatment of pests and diseases 5. Recognition of pests and diseases 6. Biology of pests and diseases 7. Toxicology or plant protection measures	Advisory services depend on all the data related to the field and crop: [1] Field Information o Size o Location o Elevation o Soil [2] Crop Information o Seeds o Variety o Date of planting o Spacing o Intercropping information [3] Past production activities o Previous crop treatments for pests and diseases o Fertiliser application o Other activities such as weeding [4] Certification data [5] Farm details o Manpower on the farm o Equipment
Crop-cultivation and harvesting: Harvesting phase	1. Current crop in the fields 2. Harvested crop 3. Crop road network and conditions 4. Public transport types 5. Intercropping, relay cropping and rotation 6. Resource-related farm advice 7. Short-term weather forecast 8. Seasonal weather forecasts (3-6 months ahead) 9. Real-time observations	The farmer's objective is to determine the best harvesting date. This is largely linked to weather forecast in relation to the crop specificities, the work needed, the availability of post-harvest services (transport, transformation, storage) and the market demand. At the farmer's level this is linked to: 1] Field information o Location o Size [2] Crop data (crop, varieties, seeds) [3] Certification information [4] Farm details

		○ Manpower on the farm (availability of labour force) ○ Planting equipment [5] Post-harvest information: ○ Warehouse access and characteristics ○ Processor linkages ▪ Position of processing equipment for transport ○ Post-harvesting processing materials ○ Post-harvesting transformation ○ Storage capacity and characteristics ○ Transport service availability [6] Selling opportunities the farmer has access to ○ Cooperatives/production cluster membership ○ Markets ○ Agribusinesses linkages
Post-harvest processing phase	1. Profiles of different value chain actors and organisations a. Farm data, e.g. farming system, crops, land area, farm income, household composition, farm employment, farm holder's age, fertiliser use, etc. b. Cooperatives c. Trade d. Processors, e.g. type, size, turnover, capital, investments, environmental transparency indicators, etc. e. Retail 2. Road network and conditions 3. Public transport types 4. Resource-related farm advice 5. Certification	After harvesting, farmers must decide what to do with yields, what the processing opportunities are, what the certification constraints are, and what the market opportunities are. At the farmer's level this is linked to: [1] Production information ○ Yield ▪ Volume ▪ Date of harvest ▪ Grades/quality of the production [2] Post-harvest information ○ Warehouse access and characteristics ○ Processor linkages ▪ Position of processing equipment for transport ○ Post-harvesting processing materials ○ Post-harvesting transformation ○ Storage capacity and characteristics ○ Transport service availability [3] Certification information

Marketing and selling phase	1. Profiles of different value chain actors and organisations a. Farm data, e.g. farming system, crops, land area, farm income, household composition, farm employment, farm holder's age, fertiliser use, etc. b. Cooperatives c. Trade d. Processors, e.g. type, size, turnover, capital, investments, environmental transparency indicators, etc. e. Retail 2. Internet connectivity map 3. Mobile connectivity map 4. Global food prices 5. National stock exchange prices 6. Regional market prices 7. Local market prices 8. Location of national markets 9. Location of regional markets 10. Location of local markets 11. Import/export volume	Selling depends on the product available (volume, quality), how to conduct financial transactions and how to deliver products to a buyer. Selling can be direct, or through e.g. group selling. At the farmer's level this is linked to: [1] Production information o Yield ▪ Volume ▪ Grades/quality of the production o Production information ▪ Treatments ▪ Fertiliser ▪ Activities o Planting/harvesting/post-harvesting material used [2] Post-harvesting activities o Treatments o Transformation o Storage o Transport [3] Transport service availability and cost [4] Certification data [5] Financial instrument o Bank account(s) o Mobile money account(s) [6] Selling opportunities the farmer has access to o Cooperatives/production cluster membership o Markets o Agribusinesses linkages

Information and communication technologies

While it was not the primary focus of this study, the analysis of numerous ICT for agriculture systems has led us to identify various technologies used to acquire field information and disseminate information towards the farmer. The tables below summarise these findings. In the same way, the last sub-section identifies the various options we found in our research in terms of technologies, systems and approaches used by FOs when implementing an electronic membership system.

ICT to acquire data

The table below lists different ICT technologies used to acquire data and their scope of usage. This list comes from exhaustive desk research based on the following publications: [CTA-Apps4ag], [CTA-ARDYIS 2012], [Rahman-Fong 2016], [Pye-Smith 2015], [DELOITTE-2012], [ICT Update-82] and [Compton 2016].

Note that the absence of a cross in a cell means that we have not identified any example in this category (e.g. use SMS/IVR (interactive voice response)/USSD (unstructured supplementary service data) to acquire information about crop selection or harvesting). This however does not mean that the channel is inappropriate, or that it is not recommended to use this channel.

Technology	Land and crop selection	Credit	Insurance	Planting	Field management	Harvesting	Processing	Marketing and selling
Satellite imagery	**X**		**X(4)**	**X**	**X**	**X**		
Drone imagery	**X**		**X(4)**		**X**	**X**		
Sensors and IoT (Internet of Things) technologies				**X**	**X**	**X**		
Smartphone apps (survey tools, GIS mapping)	**X**	**X**	**X(1)**		**X**	**X**	**X(1)**	**X(1)(2)(3)**
SMS		**X(1)**	**X(1)**		**X**		**X(1)**	**X(1)(2)(3)**
USSD		**X(1)**	**X(1)**		**X**		**X(1)**	**X(1)(2)(3)**
IVR		**X(1)**	**X(1)**		**X**		**X(1)**	**X(1)(2)(3)**
Mobile money		**X(a)(b)**	**X(a)(b)**				**X(a)**	**X(a)(b)**

(1) For farmers to submit requests
(2) For market information collection
(3) To propose products
(4) Weather

(a) Using mobile money to receive money
(b) Using mobile money to pay for service

ICT to disseminate information

The table below lists different ICT technologies used to disseminate information to farmers and their scope of usage. The list comes from a compilation of the same desk research as the previous section on ICT to acquire data. In our analysis below, the use of video (e.g. Digital Green) is different than the use of TV (e.g. Shamba Shape Up[6] and TV Koodo[7]). The same distinction is valid for delivery channels for audio content and radio broadcast.

Note also that the absence of a cross in a cell means that we have not identified any example in this category (e.g. promotion of agriculture insurance over TV), but it does not mean that the channel is inappropriate, or that it is not recommended to use this channel.

Technology	Land and crop selection	Credit	Insurance	Planting	Field management	Harvesting	Processing	Marketing and selling
Radio	**X**	**X(1)**	**X(1)**	X	X	X	X	X
TV	**X**			X	X			
Video	**X**			X	X	X	X	
Audio	**X**				X			
Smartphone apps	**X**	**X(1)**	**X(1)**	X	X	X	X	X
SMS	**X**	**X(1)**	**X(1)**	X	X	X		X
USSD		**X(1)**	**X(1)**		X	X		X
IVR	**X**	**X(1)**	**X(1)**	X	X	X		X
Web	**X**	**X(1)**	**X(1)**	X	X	X	X	X
Calling centre	**X**	**X(1)**	**X(1)**	X	X			X
Mobile money		**X(a)(b)**	**X(a)(b)**					X

(1) Raising awareness/ subscribing

(a) Using mobile money to receive money

(b) Using mobile money to pay for service

[6] http://www.shambashapeup.com/

[7] https://www.youtube.com/watch?v=3BFGH8YNClY

ICT to manage farmer profiles

As part of this study, we investigated current technologies and platforms that farmer organisations are using to track and maintain their member profiles, and how they collect and update them. We also used the 'Enhancing development through cooperativesthroug[8]) initiative led by OCDC[9], OXFAM[10], CIAT[11] and CTA which organised large scale African cooperatives surveys to have a sense of current usage of ICT in general and for farmer profiling in particular. The results of these surveys are presented in the last part of this section.

ICT technologies for farmer profiling

The study identified four main technologies/tools/approaches currently used for managing farmer profiles:

- Spreadsheet software (e.g. MS Excel, Google Sheet)
- Online services using the Software as a service (SaaS) model[12]
- Off-shelf customisable products
- Specific tailored solutions

NB: the aim of this study was not, in any way, to do an extensive comparison of existing platforms on the market and compare them, but instead to identify the various categories available and pick a couple of examples in each for evaluation.

Spreadsheet software

Spreadsheet software is usually the very first step for a cooperative to digitise their paper listing of members. This solution presents several interesting features:

- Cost: Some tools are free to use (e.g. Google Sheets), some are part of a usual computer package (e.g. Excel in Office). In all cases, the investment is low.
- Connectivity: Software runs on a local machine and does not require connectivity. Customers can access and exploit their user base even during network failure, or even without electricity or laptop with batteries.
- Reliability and maintenance: Spreadsheet software is very stable software and usually bug free. There is no need for maintenance or upgrade.
- Scalability: Spreadsheets can manage millions of records without trouble.
- Extensibility: The extension of profiles is very easy, just adding more columns will expand the profile. This is by far the most flexible and adaptive solution.
- Capacity building: The use of spreadsheets for basic operations like entering profiles or searching information is relatively easy. Moreover, the development of capacities on using spreadsheets can largely benefit organisations such as cooperatives in all their activities (e.g. financial management, etc.).

[8] http://edc.cta.int/
[9] http://www.ocdc.coop/index.html
[10] http://www.oxfam.org.uk/
[11] https://ciat.cgiar.org/
[12] https://en.wikipedia.org/wiki/Software_as_a_service

- Power user: There is no limitation in the exploitation of data, the development of analytics (graphics, complex visualisation, etc.) or the mash-up with other data sources (e.g. with financial data). However, the use of advanced functions requires deep skills and capacities in the specific software.
- Mobile use: Most modern software (e.g. Excel 2016, Google Sheets) offers mobile versions of their tools that work offline. Data files can be shared and synchronised using cloud storage such as Google Drive, OneDrive and Dropbox.

This approach also presents a series of limitations:

- Security model: Although some software offers security because rights can be limited to certain users, the options are limited and not very flexible. The second issue is the ease to modify or clear cells without notice or by mistake. Critical data can be erased without notice.
- Backup: The backup is usually manual, and the versioning of files is very limited.
- Multi-user usage: Spreadsheets are not multi-user and in particular the offline/online synchronisation is at best problematic, but in most cases destructive (one version replaces a previous one and changes made between two synchronisations are lost).
- Modularity: It is difficult to share only part of the data with others such as a field agent. The database is usually in a single file and monolithic.
- Error checking: There is no easy way to build entry validation, and do basic checks (e.g. valid phone numbers, valid ID numbers, etc.). It is possible to design an application for data entry in almost all spreadsheet software, but this is more about designing a specific application (the last category in this section).
- Mobile use: The management of a large spreadsheet on a small screen is not easy and not user friendly.
- Capacities: As mentioned earlier, while spreadsheet software is very powerful, the use of advanced functionalities (e.g. visualisation, pivot table, etc.) requires specific skills that are not easy to acquire.
- Search functionalities: Multi-criteria search, while possible, is difficult to implement and requires advanced skills.

Online services using the Software as a service (SaaS) model

The SaaS model is becoming very popular, as internet coverage and bandwidth increase all over the developing world. Various types of platforms with different functionalities can fall under this category, as they have similar pros and cons. Examples of such platforms include Digital Green CoCo[13], Farmerline Mergdata[14], Jokalante Platform[15], VOTO Mobile[16], and engageSPARK[17]. It is interesting to note that there are also national-level platforms whose aim

[13] https://www.digitalgreen.org/
[14] http://farmerline.co/
[15] http://jokalante.com/
[16] https://www.votomobile.org/
[17] https://www.engagespark.com/

is to serve various organisations across countries. For example, the case of the Fiji Crop and Livestock Council[18] (FCLC), serving public agencies as well as farmer associations in Fiji.

These platforms present several interesting features:

- Costs: Usually, the cost is based on a subscription package, on usage (e.g. the number of minutes of communication used), or on a mix of these two. In all cases, setup costs are usually very low (up to zero, e.g. for engageSPARK), and monthly costs depend on usage. This is by far the cheapest option for small-scale cooperatives.
- Integration: Most platforms offer a complete integration of services, in particular for communication services. This includes the provision of phone numbers (IVR/SMS), sometimes even with nationwide short codes. Customers through one provider can acquire a complete bundle.
- Maintenance: Maintenance as well as data security (backups) are managed by the platform and transparent to the user.
- Security model: Some platforms (e.g. Jokalante) offer advanced security models where different users of the same organisation have different rights. This is usually configurable.
- Multi-user usage: By design, these platforms are multi-user, and can be accessed by various people from very different locations.
- Scalability: The infrastructure (storage, power, bandwidth) usually grows as the number of users grows. As a shared resource, on the model of cloud services, customers are charged only for the portion they use, and this can grow without limits.
- Capacity building: The platforms are web-based and usually very easy to use.
- Power user: Most of these platforms offer powerful features (search functions, analytics, etc.) that are very easy to use.
- Data integrity: Data entry is usually protected with a series of validations that ensure data integrity, and prevent the entry of invalid information.

This approach also presents a series of limitations:

- Extensibility: The biggest issue with SaaS platforms is their flexibility. In the context of this study, while some platforms offer customisation for user profiles (e.g. Jokalante, engageSPARK), the flexibility is very limited, and limited to unitary information (e.g. a field with one possible value). In terms of service, new services become available based on the platform provider roadmap. Customers can request features, but they have relatively little power on the development.
- Export: In most cases, data is maintained in an internal proprietary format. Some of the platforms have output formats (usually Excel or CSV), but this is limited. None of the platforms seem to offer application programming interfaces (API) access to enable e.g. a third-party service provider to access and exploit data available on the platform in real-time.

[18] http://www.fclc.org.fj/

- Languages: Platforms usually support only a very limited number of languages (usually English or French, rarely both, sometimes Spanish, more rarely Portuguese) that can make their access difficult in countries for which they were not designed.
- Connectivity: The platforms require good stable bandwidth for their users. In case of network problems, customers can access and use their user base and associated services.
- Costs: Most of the existing platforms have business models linked to usage. As the usage grows, the cost grows exponentially. This is particularly the case for communication platforms (e.g. engageSPARK, Twilio[19], etc.) that charge a premium cost for communication time, leading to very high costs for high usage.

There is no mention of mobile usage in the above analysis, mostly because it depends on the platform itself and whether it supports offline features and mobile design. Some platforms do support mobile offline (e.g. Jokalante, Digital Green, Farmerline), some do not (e.g. engageSPARK, VOTO Mobile).

It is also important to note that we were not able to conduct a deep analysis on data policy adopted by various platforms. At least, and at a first glance, there is not a clear data policy publicly available on the sites of the platforms we explored apart from the FCLC one[20]. It is therefore hard to know what platform owners can do with data entered by their customers.

Finally, it is important to consider the risks associated with the SaaS model. The biggest potential issue is the viability of the provider that may stop its service at any time.

Off-shelf customisable products

Such platforms are relatively similar to the previous category, but instead of being used through a SaaS model, they are available as software packages to be installed at the customer's location (or on a cloud hosting). Some platforms are available under the two models (e.g. FCLC or Jokalante). Compared to the SaaS model, these platforms present specific pros and cons:

- Infrastructure and integration: These platforms require integration at the customer's premises. The customer therefore must pay for the hosting infrastructure as well as the maintenance (backup, security, etc.).
- Costs: The front costs are usually higher for the license of the platform and the setup costs, including hosting costs. However, on the longer term, the operational costs, particularly in the case of communication services, are usually far lower than on a SaaS model, and become cheaper as a longer-term solution or for high usage.
- Data protection: The customer is the only one having access to the data and there is no risk of misuse by a third party.
- Customisation: These products are usually more flexible than the SaaS ones. It is possible to:
    - adapt the profile and customise it to the customer needs
    - add specific modules (mobile modules, mapping, etc.) at the setup time or later in the process
    - increase functionalities as needs appear at the customer side.

[19] https://www.twilio.com/
[20] https://admin.fclc.org.fj/css/fclc_privacy_20160128.pdf

- Connectivity: Platforms can be installed at the premises of the customer, and can even be connected to a communication infrastructure (SMS/IVR) without any internet connection. This element can be critical for rural setup that may have no or instable connections.

Specific tailored solutions
This section concerns platforms and products specifically developed for a customer. Compared to the previous sections, this solution is obviously the most flexible one, but also by far the costliest one in terms of money and time. Such an approach is unlikely going to be the best option, except for large organisations with specific needs, willing to develop their own platform for specific services.

It is important to note that over time, platforms are spawning over the three last categories. An online platform is available also as a stand-alone customisable product, in which it is possible to add specific modules tailored to a given customer and which will be developed for that purpose (e.g. as part of the third category).

It is also critical to note that most of the platforms mentioned in this section offer some functionalities for farmer profiling, but were not primarily developed for that purpose. Most of the platforms are communication platforms (e.g. engageSPARK, VOTO Mobile). Some platforms offer other services such as data collection (e.g. Jokalante, Farmerline), agri-specific services like production forecast (e.g. Farmforce[21]), weather forecast (Farmerline), market price information (FCLC) or business matching (FCLC). Finally, some platforms (e.g. Grameen Foundation[22]) integrate dedicated software for profiling (Salesforce[23] in the case of Grameen Foundation). In that regard, it is not appropriate to compare a spreadsheet approach which is only for profiling, with multi-services platforms that exploit farmer profiles to deliver higher-value services for the farmer organisations. In Section III of this paper, under "Value of electronic membership management system", the value for profiling is detailed, and provides a way for farmer organisations to select the most appropriate option based on the benefits they want to capture. Finally, it is important to note that it is relatively easy to migrate from one platform to another as they all offer, at a first glance, minimal import and export functionalities.

Profile update processes

During the study, we have also investigated the processes and approaches used by organisations profiling farmers to collect and update profile information. We have identified two main strategies:

- Self-registering/self-updating: Farmers are the ones registering to the system using various channels (SMS/USSD/IVR or office visits) and provide information. Then, when their information changes (e.g. they change their phone number) or when new information is available (e.g. planting information, harvesting information, product

[21] http://www.farmforce.com/
[22] http://www.grameenfoundation.org
[23] https://www.google.com/earth/outreach/stories/grameen.html

availability, etc.) they update their information. This approach is the easiest to implement for farmer organisations. However, the main issue is the incentive for farmers to register and update their profile. If there is no direct measurable outcome or clear benefit for a farmer, it is unlikely that he or she will make the effort, particularly if there is a cost associated to the action (SMS/phone call/travel to the office).

- Regular update/census: The second strategy for collecting and updating profiles is through field surveys via an existing network of extension agents or via dedicated activities with external enumerators. Profile collections and updates usually happen on two occasions:
    - At a specific date
        - Beginning of a season/crop cycle: For organisations interested in supporting farmers across the crop cycle, the profile collection/update happens at the beginning of the season before planting.
        - Harvest time: For organisations mainly interested in marketing, the profile collection/update happens usually at the harvest time to evaluate available products.
    - During specific events
        - Setup of an electronic membership system: To feed a new ICT system, a large-scale survey is executed to collect profiles.
        - Non-working phone number: For most communication platforms, the phone number is the most important piece of information of the profile that links farmers and their organisations. Platforms detect invalid/non-functioning numbers and can trigger an update of the profile (sending an agent to the place, or calling a neighbour).
        - Project-specific interventions: As part of their monitoring and evaluation processes, and to capture their beneficiaries' information, organisations use project opportunities to capture profiles.

These data collection activities are executed by the organisations' existing extension agents network if such a network exists, or outsourced to external enumerators for the task.

It is obviously possible to mix the two approaches (self-update and organisation-driven large-scale collection), and to implement various strategies where farmers can for example register directly into the system, or be reached through a census exercise. Then farmers can update their information directly or be reached when there is a problem (e.g. non-working phone numbers). Based on the interviews, it is clear that collected information has a relatively short lifetime, and profiles can get outdated quickly, particularly the information regarding phone number, availability of products, and farmer needs (input, transport, etc.). To tackle this issue, it is essential to put update processes in place, and to consider the possible incentives for farmers at the design phase.

It is interesting to note that digitisation processes of paper files seem to provide poor results. For example, RESOPP[24], a network of farmers and pastoral organisations in Senegal, had about 50,000 members in their paper files, and a large-scale field survey that covered 18,000 farmers showed that more than 60% of the information on paper was inaccurate (mostly farmers who died or left the region). The electronic membership database setup therefore exploited the results of the survey instead of the paper information.

EDC surveys output

This section presents the results of the analysis of three large-scale surveys conducted by the 'Enhancing development through cooperativesthroug[25]) initiative. EDC is led by OCDC[26], OXFAM[27], CIAT[28] and CTA and among their activities is the organisation of surveys of African cooperatives in 2016 and 2017 in Malawi[29] (September 2016), Uganda[30] (May 2016) and Madagascar[31] (February 2017).

During each event, around 100 cooperatives were surveyed. During the first two events, there were no specific questions about the study developed in this document, but only a couple of questions about e-mail and web presence. The last event in Madagascar had a dedicated section about farmer profiling. The list of questions that were added to the survey is listed in Annex 2.

Communication channels and ICT usage

The graphic below shows the percentage of cooperatives with a phone number, e-mail address, web presence and using an ICT platform. The first column shows the percentage of cooperatives covered by a mobile signal (either in a stable way or intermittently). It is interesting to note that the infrastructure is not (or no more) a problem (almost all cooperatives are covered by a phone network). In the same way, on average, almost 80% of the cooperatives have a phone number. Internet usage is still low (only 33% of cooperatives have an e-mail address but with major differences between countries, and only 6% on average have a web presence), but ICT usage is relatively important, likely due to the increasing penetration rate of smartphones (on average 53% of cooperatives are using an ICT platform).

[24] http://www.resopp-sn.org/
[25] http://edc.cta.int/
[26] http://www.ocdc.coop/index.html
[27] http://www.oxfam.org.uk/
[28] https://ciat.cgiar.org/
[29] https://sites.google.com/site/edcooperatives/collaborative-edc/10--malawi
[30] https://sites.google.com/site/edcooperatives/collaborative-edc/7-uganda-2016
[31] http://edc.cta.int/2017/03/21/malagasy-cooperatives-to-play-a-new-role-in-agricultural-and-rural-development/

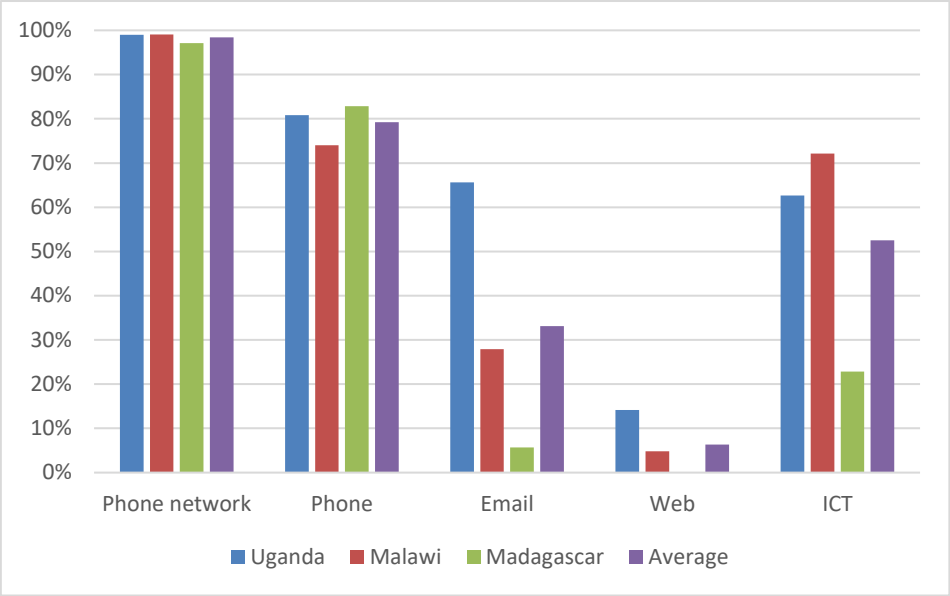


Figure 9: ICT channels and cooperatives in Uganda, Malawi and Madagascar

Farmer profiling

In this section, we only represent data for Madagascar (105 respondents). Given that the country has the poorest ICT usage in the previous figure, the numbers below might not be fully representative of Sub-Saharan countries. The main objective of this exercise was to: 1) evaluate the number of cooperatives and FOs profiling their members; 2) identify the information collected and their usage; 3) have a view on the use of ICT for profiling activities; 4) briefly capture the use of open data by cooperatives and FOs.

Concerning profiling activities, the figure below shows that almost 75% of surveyed cooperatives and FOs are profiling their members (i.e. collecting more than just identity information).

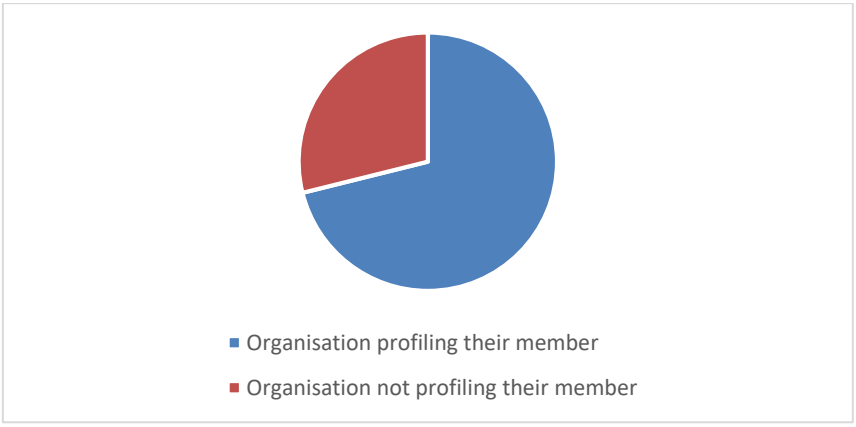


Figure 10: Percentage of FOs and cooperatives profiling their members

The figure below shows the type of information collected.

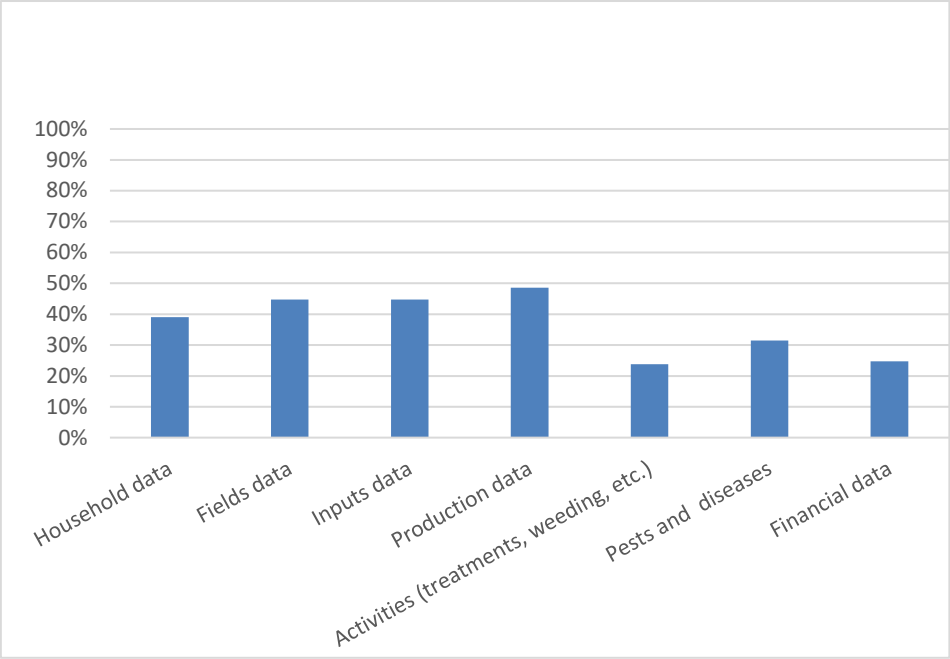


Figure 11: Information collected during profiling

It is interesting to note that there is not one core set of data that is collected by all, but there is instead a homogeneous dispersion among FOs and cooperatives. This means that there is not one specific block of information that should be the core of an electronic membership management system (EMMS), but instead that an EMMS must be flexible and adapt to the organisation needs.

In the same way, we also find similar relatively homogeneous disaggregation of exploitation of collected data presented in the figure below.

Figure 12: Exploitation of profile information

Group selling is slightly higher than other use of information. This matches the previous figure showing a slightly greater focus on production data. The use of data for extension services is slightly lower because it is not a usual function of cooperatives and FOs.

In terms of technology usage, the figure below demonstrates well the state of ICT at the cooperatives and FOs level.

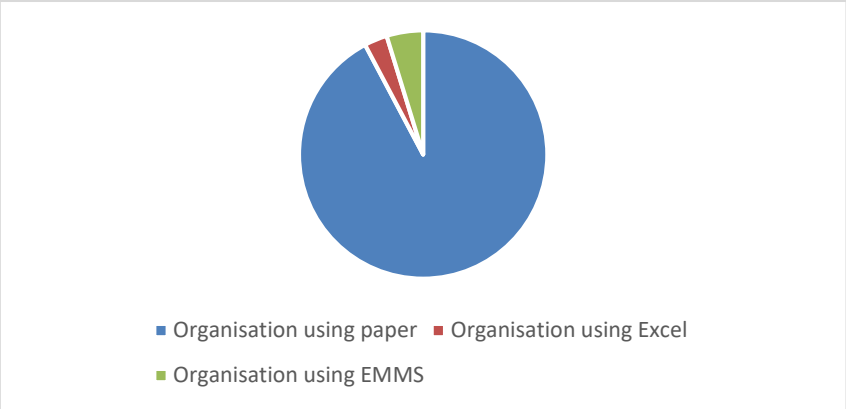


Figure 13: Use of ICT for profile management

Close to 97% of cooperatives and FOs are still using paper for profile management. This is homogeneous with the global ICT data we presented in the previous section, where Madagascar was the country with the lowest use of ICT at all levels (e-mail, web and ICT platform). This ratio still shows an important need for promoting EMMS both for managing profiles but also to ease the data collection. The figure below is also interesting and shows that cooperatives and FOs are conducting data collection more than once a year for more than 50% of them.

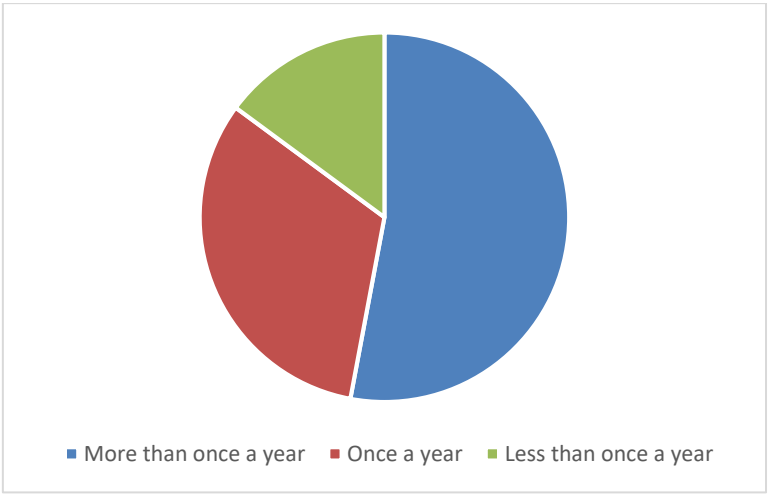


Figure 14: Frequency of profile information update

In such a context, the use of ICT tools would not only ease the update but also the use of information.

Finally, the results about open data are not very meaningful. Twenty per cent of the surveyed organisations mentioned the use of some kind of national open datasets. However, it is difficult to exploit this data point as respondents may not be aware of open data and may refer to paper-based data. Moreover, open data is not developed in the country (no initiative at the government level, no coverage in any open data index) and therefore the awareness as well as usage is low.

Section III – Recommendations

This section is a summary and compilation of the study findings organised to answer the two main questions that farmers' organisations and cooperatives ask themselves: Why should we profile our members? What should we collect and why? Apart from these 2 questions, this section also includes a part on data ownership and privacy protection that is a growing concern for all stakeholders.

Value of electronic membership management system

The first key question with regards to profiling is to understand the rationale and the value of the exercise. This value depends on the stakeholders, and we have identified three perspectives to consider: Cooperatives and farmers' organisations, farmers themselves, and policy makers.

Cooperatives and farmers' organisations perspective

Cooperatives and FOs have the highest probability to take advantage of a deep understanding and knowledge of their members. An EMMS enables FOs and cooperatives to know their members in detail: Who they are, what they do, where they live, what they produce, etc. This information is essential for many reasons:

- Planning and strategy: FOs can plan their services, their interventions and their areas of investment based on real data. They will be able to identify areas where they could expand, or places where there are specific opportunities in terms of production or selling. A deep understanding of its memberships allows organisations to define their roadmap and identify new services. It also allows them to have financial forecast and evaluate with high precision potential markets for various services. This information will allow FOs and cooperatives to plan their activities based on real facts and data.
- Easier membership management: Using an EMMS helps managing memberships for all internal activities, such as payments or elections and votes. The use of such platforms helps FOs save time and money for those tasks.
- Easier communication: The capture of communication details, in particular phone numbers, allows using communication platforms that automate sending of information in various formats (voice, SMS, etc.). The use of such platforms allows better and more regular communication between the central organisation and its members. Using new communication channels enables organisations to:
    - better understand needs and demands from their members
    - better understand constraints and pain points
    - query their members on specific topics and get their feedback.
- Greater opportunities to identify and put in place new services: A better knowledge and understanding of its membership enables organisations to identify new targeted value-added services. These services not only provide new benefits to the members, but also increase the value of membership, and enable FOs and cooperatives to recruit new members. Among potential new services, ICT services developed by third-parties have

a specific place. The availability of a maintained EMMS is critical for ICT service providers. It saves them huge costs, making services more affordable and at the same time sustainable. To reach this goal, those services must be designed jointly between cooperatives and FOs, service providers and farmers. It is important to note that while many articles highlight the opportunity for farmers and cooperatives to sell data and generate income from this activity (see e.g. [Fedor 2016]), at the same time, there is a growing number of ICT service providers failing to make ICT services sustainable due to the cost vs revenue to serve farmers. It is therefore essential to have partnerships, and a clear benefit-sharing approach between cooperatives/FOs and ICT service providers.

- Greater power in advocacy: A deep knowledge of its membership allows FOs and cooperatives to have a stronger voice in advocacy. At a basic level, an organisation with an EMMS can prove their membership and demonstrate the number of people it is representing, who they are and where they are. This gives power to their voice based on their representativeness. At a more advanced level, an organisation can exploit its membership to inform policy makers in various ways.
    - Simulating the impact of proposed measures: Based on farmers' information, it will be possible for an organisation to measure or simulate the impact of new measures (e.g. a new subsidy scheme) and define its position with regards to proposed measures based on real data.
    - Executing quick surveys to get members' opinion on specific topics: Organisations can mobilise their members and collect their opinions on specific topics. Here again, such a process will help an organisation to define its position and defend it based on real data, and based on a clear mandate from their members.
- New sources of revenue: The farmers' profiles are a potential source of revenue for third-party activities such as:
    - research
    - market surveys
    - advertisement.

  However, it is important to note that farmers must agree with the use of their data and their participation in such activities. See the section on data ownership and privacy protection on that matter.

Farmers' perspective

While cooperatives and FOs have large interest in profiling, similar opportunities exist at the farmer's level. The implementation of an EMMS will provide a series of benefits to farmers:

- Access to new value-added services: The biggest opportunity provided by an EMMS is to enable third-party service providers and ICT service providers to design and deploy services for farmers. The use of profiles enables service providers to design services that can adapt to farmer specific needs.

- Better communication with the FO/cooperative: Like for FOs and cooperatives, the use of communication platforms together with an EMMS enables farmers to more easily communicate with their organisation(s) and have access to critical information on a regular basis.
- Greater global visibility/ability to voice concerns: The use of an EMMS and communication platforms also allows farmers to participate in global dialogues and have their voice heard on specific topics that matter to them.

It is important to note that all the benefits mentioned above are potential benefits, and largely depend on the use and the tools FOs and cooperatives are putting in place together with the EMMS. A platform whose aim would be only to gather farmer profiles would be useful for FOs and cooperatives, but not much for farmers. This is a very important point to consider, as the lack of direct interest or incentives has an impact on the willingness and effort farmers will make to give information and update their profiles. Without an incentive, it is unlikely that famers will support the exercise, leading to profiles becoming obsolete quickly. This risk was clearly highlighted in a few interviews.

Policy makers' perspective

Policy makers and public authorities have substantive interest in the information stored in farmers' profiles. At a basic level, it is important for public authorities to ensure that they talk to the most representative organisations. As mentioned earlier, the use of an EMMS enables organisations to prove their representativeness.

However, the potential greater impact is at the data level. The information stored at the profile level has a critical value for policy makers and for a lot of public agencies (National Statistical Office, etc.) in terms of measurement and planning. An EMMS provides disaggregated data at a hyperlocal level. Those data point mainly to the agriculture sector, but given the importance of agriculture in rural areas, it also provides a lot of information on most households. The data stored in a profile, after anonymization (see the discussion on this topic in the data access paragraph later in this section), could contribute to many data that public authorities are interested in:

- Agriculture core datasets such as land usage, production, etc.
- Measurement of public policies impact: e.g. reach and impact of subsidies schemes
- Sustainable Development Goals (SDG[32]): In Annex III, we have conducted a deep analysis of SDGs and listed targets that could be instantiated using farmer profile information
- General household data: Education, household composition, income, land ownership, etc.

All this data could be used either as a source of information or as a verification mechanism compared to official data collection processes.

[32] https://en.wikipedia.org/wiki/Sustainable_Development_Goals

It is important to note that this data can have an impact both at a local level through sharing with local government agencies, as well as at a national level through aggregations. Such anonymized data is also of high importance for researchers, for example to adapt and improve theoretical models such as weather forecast models, crop models, etc.

Farmer profile structure

This section summarises and structures the findings from Section II on the key information and fields of a farmer profile. The proposed profile is structured around two levels:

- A section level that groups information of the same nature
- A field level that describes a specific element of information

In this paragraph, the heading represents section level information and the fields are listed under it. Note that the information below is a compilation of analysis of existing platforms and interviews. The proposed sections and fields have been cited or are currently used.

Personal information

This section is the very basic part of the profile and lists the information about the farmers to identify:

- Last name and first name
- Father's name: In some cultures and countries, the father's name is essential to identify the individual
- Identity number: Note that not all countries have IDs for their citizens
- Gender

Apart from this basic information, other interesting data includes:

- Birthday: The age of the farmer is useful information for investment or to access specific subsidies. Depending on the country, the exact date of birth is not known, and storing the birth year is the only reliable information. It is important to note that some platforms store the age or age group of the farmer. While supposedly giving the same information, as age evolves over time, the capture date must be recorded, and the field computed at each access, which is problematic. It is recommended to store only the date of birth.
- Main/second occupation: It is important to know whether the farmer has any other, secondary occupation.
- Income: Information about income is useful for many purposes: Access to credit, access to specific subsidies, monitoring and evaluation (M&E) purposes, etc. However, this is information that is difficult to capture as people are reluctant to provide it for various reasons, from tax purposes to privacy. One option is to define income groups to capture useful information.
- Subsidies: What are the subsidy programmes that the farmer uses and what is the income provided by the programme? This is critical information not only to access credit, but also to identify programmes that the farmer should have access to but is not using.
- Education level: This is usually useful for M&E processes.

Apart from the farmer, information about the household can be interesting, mainly for accessing specific subsidies, or for M&E. The basic information is to know the number of individuals in the household. Then the same details for each member could be collected, but this is not directly in line with the objectives of an EMMS to support agriculture and smallholder farmers.

Communication information

Communication information could be part of personal information, but given the details that could be useful, we decided to separate it in a dedicated section. This section covers all information to interact with the farmer either directly or through broadcast media. It includes the following information:

- Phone(s): The phone number is usually the most important piece of information to interact with the farmer. It is important to note a couple of points:
    - Farmers usually have more than one phone number. All should be recorded, with the default/preferred clearly stated.
    - The concept of phone sharing is still largely prevalent in developing regions, sometimes at the household level, sometimes at a larger scale. It is essential that this feature is correctly implemented.
- E-mail: While e-mail is still not used very much, usage is growing, and the information should be captured.
- Social networks: As for e-mail, social networks like WhatsApp have a growing set of users.
- Preferred communication channel: When a farmer has e-mail, a social network, and some phone numbers, it is important to rank those various options.
- Language(s): What are the various languages/local dialects that the farmer understands? It is important to note a few points:
    - Sometimes farmers understand different languages, but prefer one, or are more fluent in one. This information is useful to track.
    - Understanding and speaking are two different concepts. Sometimes, some farmers can understand a language without being able to speak in this language.
    - In the same way, some people can understand a language orally, but not in writing. This information is also essential for written communication (e.g. SMS).
- Radio/TV information: For cooperatives and FOs planning to use broadcast media, it is essential to capture some information:
    - Whether the farmer has access to a radio or TV
    - At what time he or she listens to the radio or TV
    - What the preferred stations are.
- Use of phone/mobile phone literacy: The type of phone (basic, feature phone/smartphone) and the use of other features than voice (e.g. SMS, apps) are critical information to identify the most powerful channel accessible to a given farmer.
- Maximum number of messages to receive per week: Some organisations capture this information to define their campaign and the level of communication.
- Willingness to use pay-for services: Many interviewees mentioned this element as critical to define the business model of ICT services.

- Push or pull service: This field tracks whether a farmer prefers to receive information about specific topics, or if he or she prefers to contact a service to get this information. This is related to the above field for service providers to design their services based on user preferences.

Location

This is also an integral part of farmers' personal information, but includes different elements:

- Administrative location: Depending on the country, this is composed of two to four information elements from region to district to county to commune. The lists are usually normalised by the administration and uniquely qualify a place that covers a few villages.
- Other address details: While the administrative location delimits a region, it does not qualify the exact place of the farm. Other address information is needed:
    - Village
    - Address: In many rural places, address has no meaning, and just the village and the farmer identity are sufficient to locate the farm.
- GPS coordinates of the farm: When available, this is the most precise information to complement the administrative location information. It is important to note that in several countries, it is very difficult to derive the administrative location from the GPS. The administrative location is critical in many cases for subsidies, relation with extension agents, etc.

The three blocks of information mentioned above are the core information that should be collected for all profiles. Some of the fields may not be necessary depending on the targeted services (e.g. information about radio or TV), but the three sections will appear in all EMMSs. The sections that are listed are more service specific and may or may not be relevant to a given FO or cooperative depending on its objectives and targeted activities.

Financial instrument

Information about financial instruments available at the farmer level is critical for services that involve money transfers (e.g. credit, insurance or subsidies payments):

- Bank account(s): The exact bank account details should not be stored for security reasons, but whether the farmer has a traditional bank account and in which bank branch is useful for service providers.
- Mobile money account(s): In the same way, knowing whether the farmer has access to mobile money and on which phone number is essential for the implementation of services requiring money transfers.

Farm details

Information about the farm as an enterprise is critical to identify specific needs and interventions to support the activities. Key information identified is:

- Registration number: If the farm is a formal business, it usually has a registration number that can be useful for various activities such as traceability.
- Manpower on the farm (availability of labour force): This information is critical to identify existing resources for activities.

- Equipment: In some cases, specific activities require specific equipment. In some other cases, the availability of equipment reduces the manpower need. Knowing in detail this information has impact on the overall business model of the farm. Usually, the equipment is linked to specific activities happening at specific stages:
    - Planting
    - Harvesting
    - Post-harvesting.
- Extension agent(s): The (list of) extension agent(s) associated with the farm/farmer.

Note that the farm as a business is characterised by its financial data (turnover, benefit, etc.). As part of this study, we focused on smallholder farmers where in most cases, the financial data of the farm is the financial data of the farmer. We have therefore merged the two, and financial data of the farmer is covered in various blocks (see personal information for income, credit information or production for costs).

Qualification and certification data

Qualification and certification apply to either the farm or the farmer. Most certification requires a training first. However, some training does not lead to any certification. This information is critical for many purposes. First, most certifications have regulations on various activities from planting to applying treatment to harvesting. Extension services must adapt to these constraints. Then, certifications provide added-value to the end product, and this is critical for the marketing activities. Finally, knowing the certifications a farmer has enables him to more easily access other certifications. This, for example, is the objective of a service like Sustainability Map[33] that lets a farmer know, based on their current certifications and the ones they want to reach, the set of modules they must follow. The information required for qualification and certification is:

- Training/certification name/label
- Training/certification date
- Training/certification institution

Field information

In many cases, a farmer manages more than one field in different places, or even if it has one piece of land, the space is split into sections with different crops. Core field information is:

- Location: The location is critical for many services such as insurance, weather forecast, soil mapping, etc. There are different ways of capturing the location of a field:
    - Administrative location (see the location section above) complemented by a detailed description: This may give an exact location of the field, but this is hardly exploitable automatically for field level weather forecast, satellite image analysis, soil mapping, etc. This information is usually available on the land title, when such a paper exists.
    - A GPS point in the field: This will give the exact location of the field that can then be exploited by other services mentioned above.

[33] http://sustainabilitymap.org/home

    - A map (geo-fencing) of the field: While such a detail is not necessary for the location, if such mapping is done, it can be exploited for the location of the field.
- Size: The size of the field is the second critical information for forecasting yield, insurance, etc. The size can be captured in different ways:
    - Land title: Usually, a land title includes the size of the field.
    - Evaluation: The farmer or an extension agent evaluates the size of the field by looking at it.
    - Map (geo-fencing) of the field: This is the most precise way to get the exact size of the field.
- Map (geo-fencing) of the field: As mentioned above, the map of the field can be useful for size and location, but it has also other applications, such as traceability.
- Elevation: The shape of the field is potentially useful depending on the culture. The information could be qualitative (a label such as flat, hilly, etc.) or could be in the form of a shape file.
- Soil: Soil information is critical to identify the best crops and varieties. The information about crops can be derived from a soil mapping database based on the location, and in that case, there is no need to store the information on the profile. But when such mapping is not available or when analysis has been conducted on the field, this information is critical to support advisory services.
- Land title: The legal rights for the farmer to use the field are critical to protect his/her work. Ensuring that the field belongs to the farmer and can be exploited is critical to secure the farmer and define the rules on the exploitation of the field. As mentioned above, the land title can also provide useful information such as the location or the size.
- Crop history: Keeping records of the crops that were grown on the field is also valuable information to ensure that soils are regenerating well using different crops over time.

The information in the next sections (crop information, production information, insurance information, weather information) is also related to each field, but is put in separate sections for clarity.

Crop information

Information about crops that were planted in each field is critical for many services from advisory services to marketing. Relevant information includes:

- Crop
- Variety
- Seeds: Type of seeds such as certified seeds.

Note that we focused in this study on seasonal crops. For longer term crops, like trees, the age/date of planting is also important information.

Production information

This section covers all the information about production. Depending on the objectives of the cooperatives or the FOs, only some elements might be interesting to capture.

- Certification: Which certification is targeted for a given field?
- Planting information: There is little information related to planting that is useful for the forecast of yields (date and volume):

    - Date of planting: This is critical for forecasting the harvest date.
    - Spacing: This is important for some crops to forecast the yield.
    - Intercropping information: This information is important for treatment to apply, or to apply a different model for yield forecast for some crops.
    - Equipment used: This information is helpful as constraints for some certifications, or as hints for yield (regularity of planting, etc.).
    - Amount of seeds: This is helpful to evaluate costs.
- Activities information: For certification purposes, or to have a deep understanding on the events that impact yields, it could be useful to track activities across the production. This includes:
    - Treatments used: What was used on a given field, when and how (quantity and method of application)?
    - Fertiliser: What was added to a given field, when and how (quantity and method of application)?
    - Extension service interventions: When did an extension agent visit the field and the intervention he conducted? This is useful to evaluate the impact of extension agents, and monitor the agent activities.
        - Extension agent monitoring with GPS coordinates.
    - Pest and disease attacks: When did an attack take place? This should be related to the treatment section. This section may also include specific diagnosis (e.g. lack of nutriments, water stress). This is useful to correlate activities and yield.
    - Activities and events: Specific activities like weeding can be tracked. This is useful to correlate activities and yield.
    - Water usage: Some interventions look at reducing water usage, or some varieties of specific crops are known to consume less or more water. This information should therefore be monitored.
- Weather data: Information related to weather can be tracked to monitor climate change, as well as to adapt advisory services to the conditions. This information is also important to relate with the volume of yields as well as to check and update a weather forecast model. The information is collected regularly and tracked in the form of dated values. The tracking can be manual or automatic through sensors, for example. Key information includes:
    - Rainfall
    - Temperature
    - Hygrometry.
- Yield: Information about yields is essential for all activities upward the value chain, from transport to transformation to marketing.
    - Forecasted volume: It might be useful to keep track of forecasted yields to evaluate the accuracy of the model used. The forecast may evolve over the crop cycle based on various events (pest and disease attacks, weather, etc.) and the evolution of the forecast should also be tracked
    - Real volume
    - Date of harvest
    - Equipment used for harvest
    - Grades/quality of the production.

- Post-harvest data: Post-harvest data are essential for the exploitation of the yield. There are two types of information: The context in which the farmer operates and that is global to the farm, and the information about a specific series of activities that is associated with a given yield in each field.
    - Post-harvest context:
        - Warehouse access and characteristics
        - Processor linkages: The processors that the farmers are linked to. To support the identification of transport needs, it is important to track the position of processing equipment
        - Post-harvesting processing materials: Materials and equipment that the farmer can use
        - Storage capacity and characteristics
        - Transport linkages: The transport services that the farmer can use and their characteristics.
    - Post-harvesting activities
        - Treatment
        - Equipment used
        - Transformation
        - Storage
        - Transport.

The previous sections (field information, crop information, and production information) are essential to track, forecast and manage yields and associated services (advisory services, etc.) from planning to the final product. The sections below focus on the financial aspects of the farm operations: business, credit and insurance information.

Business information

Business information is a critical element for marketing and selling yields or transformed products. This information describes linkages between the farmer and other key stakeholders in the value chain for conducting their business. It includes the following elements:

- Cooperatives/production cluster membership: This field lists the linkage with organisations that can sell farmer's products.
- Markets: It is critical to know to which markets a farmer is linked. This information can be used for many purposes such as the evaluation of transport costs, but also to support market price information (providing information to farmer's relevant markets only) or to identify existing opportunities.
- Agribusinesses linkages: The lists of the farmer's agribusinesses linkages can be used for different services from buying products (inputs, seeds, treatment) to selling yields.

Apart from global information, season specific information is also useful:

- Total amount of products sold: This information has to be disaggregated per crop/field/grade/certification to have a clear view of output of each category and income generated
- Prices sold.

Credit information

Credit information is critical to support access to credit. Note that the information listed below complements other information scattered across the various elements of the profile. For example, the next section is about insurance, and the fact that a farmer has an insurance is a factor that is critical for credit firms and micro-finance institutions. In the same way, the records of past yields on the same fields are critical elements to evaluate the quality of the business plan.

- Credit records: The past records of the farmer in terms of credits and reimbursement.
- Farm business plan: This is a critical piece of information to identify credit needs in terms of amount and timing based on the crop cycle. The building of the business plan is based on a clear identification of all costs related to input buying (seed, treatment, fertiliser), activities cost (equipment renting, labour force hiring, transport, transformation, marketing, etc.) and the forecasted income based on yields. The business plan should highlight not only the volume, but also the timing of the credit.
- SACCOs/ROSCAs membership: Savings and credit cooperative organisations (SACCOs) and rotating savings and credit associations (ROSCAs) are both a potential means for delivering credit and a place hosting credit records for a given farmer.
- Active credit information: Active credits information is essential to identify the ability of a farmer to get more credit. It is also a way to forecast payments based on credit allocation per activity.

Insurance information

The insurance information details the insurance that the farmer has contracted. It includes:

- Field(s) covered
- Risk(s) covered: This includes the risks and the period
- Insurance company
- Cost
- Amount repaid if one of the risks covered happens.

It is critical to note that the recommendation is not to implement the complete profile presented above. The aim of this section is to define a global framework for cooperatives and FOs to conduct a detailed needs analysis using the mapping presented in Section II to identify the information relevant to the organisation, based on the targeted services and targeted use of the profile information. In the interviews with service providers, they acknowledged that profile mapping is currently missing from their decision matrix and can be a very useful addition for mapping risks.

The profile above is the largest set of useful information that we could identify as part of this study, but it is very unlikely that all sections are relevant for a given organisation. It is also very likely that the profile evolves over time as the organisation develops more activities. It is therefore essential for organisations to invest in platforms that can easily expand over time, and include new blocks of information.

However, it is also critical to consider the previous sub-section discussing the value of an EMMS. Some of the key impacts rely not only on the information stored, but also on some other features of the EMMS itself, its ability to give selected access to information, to support communication, or to support anonymization and publication to open data.

Finally, it is also interesting to note that all platforms reviewed during the study are not primarily EMMSs. They are mainly communication platforms with developed profiling functionalities. None of them would be adequate for implementing a complete profiling solution as described in this section. To the best of our knowledge, there is not yet an EMMS solution that could then be used and integrated in other services (e.g. communication services). The design of an EMMS-only software package with a well-documented set of APIs to enable integration in other services is an interesting direction to follow to support FOs and cooperatives effectively.

Data ownership and privacy protection

This section aims to cover the issues related to data ownership and farmer privacy protection. This is a domain that is currently attracting significant attention and research. On the research side, [Sykuta 2016], [Parkhurst 2015], [Ferris-Rahman 2016] and [DeBeer 2016] present a good overview of the issues. On the technical side, the Open Ag Data Alliance (OADA[34]) works towards technical standards and protocols. It was not the aim of this study to contribute to this area of research. However, given the information stored in farmers' profiles, the question of data ownership and privacy protection cannot be eluded.

In this document, we are promoting a model with three layers, where cooperatives and FOs are at the centre of the operations to collect profile information on one side and manage access for third party service providers, and ensure anonymization and aggregation for policy makers or for publication as open data on the other side. The concerns are therefore potentially at two levels:

- Between farmers and cooperatives/FOs
- Between cooperatives/FOs and service providers

Concerning the first point, there are a few issues that need to be considered. The first one is related to privacy protection and data usage. It is essential to provide clear guidance, as well as clear data policy for farmers to understand how their data will be used. Who will have access to which part of the information, and for which purpose? This aspect is not only ethically important, but it is also important for farmers to understand the opportunities and therefore communicate the right information and keep their profile up to data.

The second issue is the benefit-sharing between cooperatives and farmers. Profiles have potential market value for service providers that will be able to exploit them. Cooperatives and FOs are likely going to make businesses and deals with those service providers and it is important that these benefits are shared with the farmers. In a cooperative structure, managed

[34] http://openag.io/

by farmers themselves, this should not be a problem. However, if farmers can capture and understand this opportunity, this is another incentive for providing information and keeping the information up to date.

The third issue is a grey area related to technology. New technologies such as sensors or drone imaging are generating massive amounts of data. This data is usually generated and hosted by the technology providers, but this data is about farmers' fields. Here again, before implementing technology-centric solutions, it is essential to clarify who owns the data, and ensure that there is a fair share of data ownership. A more exhaustive presentation of legal issues related to this topic can be found in [Rasmussen 2016].

Finally, it is important to note that sometimes farmers shift from one cooperative to another, or belong to multiple cooperatives. To facilitate these use cases, farmers should have the ability and rights to give access to their profile stored in any FO/cooperative they want. This should be part of the agreement between the cooperative and the farmer.

Concerning the second point of links between cooperatives/FOs and third party service providers, it is disappointing and worrying to see that there is not a clear data policy for most SaaS platforms we investigated. For those providing such a data policy, the content is usually very worrying, enabling the service provider to reuse the data collected for other purposes like surveys, market analyses, etc. It is therefore highly recommended that cooperatives/FOs adopt protective data policy in agreements with service providers. Templates and models should be developed and made available to cooperatives/FOs.

The second key issue is farmer privacy protection. EMMSs must provide modular access to profile information, and give access to the subset of information that is relevant only to the service the providers offer. This is a technical issue that EMMSs must offer to ensure maximum data protection.

In the same way, the publication of anonymized aggregated data is a very important feature for many stakeholders such as policy makers or researchers. It is essential that EMMSs integrate this feature natively. As a side note, it is important to differentiate open data access that protect farmers' privacy, and access to private personal information (e.g. name, location, phone numbers, etc.). Access to private information enables the design of specific high-value tailored services. Access to open data is critical for global knowledge. EMMSs should offer both options to capture the full potential of profile information.

Finally, it might be interesting to question the proposed layered model, where cooperatives and FOs are at the centre of this process. Another possible option would be to design an approach where farmers are at the heart of the model, and in charge of providing and updating profile information. From our perspective, this approach sounds very challenging to implement in the developing world for two main reasons:

- Farmers in rural areas have both equipment and capacities issues that will make it difficult for them to provide information in a digital format. Perhaps this issue will disappear over time with the increasing penetration rate of smartphones, but it is likely that this situation will occur for at least a decade.

- As mentioned earlier in this paragraph, the success of an EMMS largely relies on farmers understanding the benefits and the value, and understanding the incentives to provide and keep information up to date. This requires direct connection, and a massive communication campaign that only organisations at the very local level can implement. For that reason, cooperatives are best placed to conduct these tasks.

An example of such an approach is the worldwide-level farmer profile initiative, called Blue Number[35], launched in September 2015 at the United Nations (UN) general Assembly as a partnership between GS1[36] and the International Trade Centre[37] (ITC). Unfortunately, after the initial interest and a couple of pilots, the initiative has now transformed into a separate organisation (the Blue Number foundation[38]) with GS1 dropping its involvement, and the focus slightly changing. At this stage, as far as we know and based on publications, there is not much work and interest around this initiative anymore. There were not any real successes, or success stories. In the same way, there was no real investment in promotion and outreach that lead to farmers spontaneously registering. It is also important to note that there is no clear data policy linked from the platform. Finally, the information was limited to standards and certification mainly, with a limited set of information. From our perspective, it is unlikely that such an approach will develop further and become a success in the near future.

Another possible approach, a kind of mix between the two mentioned in this section, could be to have a decentralised system, where various parts of the profile information are managed by different entities, some more at the farmer level and some, related to agriculture activities, at the cooperatives/FOs level. In this category, new initiatives around digital identity, while embryonic at this stage, present a lot of opportunities. The domain of research, called ID4D (identity for development) is actively explored by various actors such as the World Bank[39] and GSMA[40].

The overall idea is that many developing countries are lacking a well-established identity infrastructure (see [WB_ID2016] for countries ranking in terms of ID schemes). The absence of such a scheme creates major issues for citizens to access e.g. education, or subsidies schemes and for public policies in terms of planning. Instead of aiming toward putting in place a card-based identification, which is very challenging, the overall idea of ID4D is to directly jump on digital identity that could be used both for real life identification and online identification. The digital identity will be generated or updated at the time of sim card acquisition and carried by citizens on their mobile phones. This approach is well documented in [WB_GSMA_DI_2016], [WB_ID4D_2016] and [Kende_McMillan_Theodorou_2016].

In the framework of this study, while this approach is not yet widely deployed and only a couple of pilots are being implemented, this is surely something to follow. Indeed, the digital identity

[35] http://www.intracen.org/uploadedFiles/intracenorg/Content/Redesign/Events/29%20Blue%20Number-web.pdf
[36] http://www.gs1.org/
[37] http://www.intracen.org/
[38] http://www.bluenumber.org/?lang=en
[39] http://www.worldbank.org
[40] http://www.gsma.com

will carry personal information of farmers plus will also keep up to date their phone numbers. It will also be an easy to use authentication scheme to access ICT services. Such an approach will allow farmers to manage themselves their core identity and personal information through a complete ecosystem, and decide to who they want to give access. At the same time, EMMSs could leverage such an authentication scheme already implemented, and cooperatives and FOs will not have to collect and maintain this part of the profile that is critical for communication.

Conclusion

In the first section of this report, the analysis of smallholder data points shows that there is a huge gap in terms of productivity for most crops. This gap is also a tremendous opportunity to support smallholder farmers and increase their income and food security. There is not one issue but a series of challenges at each stage of crop cycles, and the provision of the right information or the right instrument (credit, insurance, etc.) at the right time can help farmers bridging (part of) the gap. This research demonstrates that the biggest value is on the most specialised services to individual farmers. But the delivery of such specialised services at each stage of the crop cycle requires capturing a very large set of information, and the mash-up of this information with also very large sets of external datasets. It is not possible to consider setting up from scratch a system whose aim would be to assist farmers at every step in the most efficient way. From our perspective, and based on the analysis of various farmer organisations at different stages of the development of their electronic membership databases, we believe that FOs and cooperatives should have a step-by-step approach, starting with basic profiling to know exactly who their members are and where they are, then enabling easy communication between members and between the FO and their members to ease the flow of information in both directions, before moving up to value-added services. In this staged approach, at the most detailed level, profiles will have to be adapted to the specificities of each value chain. Each crop or transformed product has specific characteristics in terms of lifetime, in terms of requirements for transport or storage, etc. The farmer profiling exercise should be adapted to specific value chains and products to provide the biggest opportunities in terms of value-added services. The proposed profile in this study comes from a deep analysis of a wide variety of initiatives that cover very different products. Some elements are surely irrelevant to many crops, some are likely missing, because they are specific. However, it is likely that the blocks of information that constitute a profile are exhaustive.

While, from a data ownership and privacy protection perspective, FOs and cooperatives seem to be the best actors to protect farmers, at this point in time, very few of them are engaging in such profiling activities. The reason is twofold: 1) FOs and cooperatives do not seem to be aware of the opportunities and benefits provided by EMMSs; 2) There is not any easy-to-use, easy-to-configure software package dedicated to this task. These two points should be addressed as a priority to promote the setup of such systems.

From a public policy perspective, the deployment of such platforms can have a huge impact for governments to understand where farmers are, what they do, who the representative organisations are, and what is happening on the ground. Such an approach provides huge opportunities to implement evidence-based policy making, as well as collect and monitor a massive number of SDGs indicators across multiple goals. Such a decentralised model, where FOs are the key nodes to collect and manage farmer's data is likely to be the most scalable and affordable approach. Given these opportunities, governments should support and promote the setup of such systems within FOs.

Finally, one opportunity that is just starting to emerge is the so-called ID4D concept, and the discussion around electronic ID. This is surely an element to consider as part of this ecosystem, and that could largely ease the tasks of FOs, ensure farmers' ownership of their own data and at the end benefits all actors from farmers till governments. However, this opportunity is unlikely to be widely deployed in the next few years, and the value of electronic membership databases could already be realised today.

Acknowledgements

We would like to thank the following people whose contribution to the project was invaluable: Ben Addom, Isolina Boto, Gian Nicola Francesconi, Ken Lohento, Isaura Lopes-Ramos and Giacomo Rambaldi from CTA, Nupur Parikh from Technoserve, Mary Goretti and Stephen Muchiri from the East Africa Farmers Federation (EAFF), Ishmael Sunga from the Southern African Confederation of Agricultural Unions (SACAU), Rose Funja from Agrinfo-Tanzania, Sharifa Chikaha from the Eastern Africa Grain Council, D. Busiega from ACRE and Charles Sumbwe from MGen.

We would also like to thank the respondents to the CTA survey in Madagascar as part of the Enhancing Development through Cooperatives (EDC) initiative. Their responses to our questionnaire provided valuable information and contributions to the study. We also like to thank Liva Rakotoarisoa, agro-economist, who digitalised all answers of the survey in Madagascar.

Annex 1 – Bibliography

- [ADB-FEEDAFRICA-2016] ADB Feed Africa: Strategy for Agricultural Transformation in Africa 2016-2025 https://www.afdb.org/fileadmin/uploads/afdb/Documents/Generic-Documents/Feed_Africa-_Strategy_for_Agricultural_Transformation_in_Africa_2016-2025.pdf
- [Compton 2016] Godan Success Stories, Issue 1, August 2016, Godan http://www.godan.info/documents/godan-success-stories-issue-1
  [CTA-Apps4Ag] CTA Apps4Ag database https://apps4ag.org/
- [CTA-ARDYIS 2012] Emerging Voices in ICT and Agriculture, published as part of the ARDYIS project, CTA, 2012 https://publications.cta.int/media/publications/downloads/1703_PDF.pdf
- [DeBeer 2016] Ownership of Open Data: Governance Options for Agriculture and Nutrition, Jeremy de Beer for GODAN, www.godan.info/file/46928/download?token=zaFhiMIZ
- [DELOITTE-2012] eTransform Africa: Agriculture Sector Study - Sector Assessment and Opportunities for ICT, Deloitte Consulting, February 2012, http://siteresources.worldbank.org/EXTINFORMATIONANDCOMMUNICATIONANDTECHNOLOGIES/Resources/282822-1346223280837/Agriculture_FullReport.pdf
- [FAO-SOFA-2014] The State of Food and Agriculture 2014, FAO, 2014 http://www.fao.org/publications/sofa/2014/en/
- [Fedor 2016] Farmers Tap New Source Of Income: Big Data, Liz Fedor, Twin Cities Business, May 2016, http://tcbmag.com/Industries/Food/Farmers-Tap-New-Source-Of-Income-Big-Data
- [Ferris-Rahman 2016] Responsible Data in Agriculture, Lindsay Ferris, Zara Rahman, September 2016, GODAN, http://www.godan.info/file/46934/download?token=Co5JuSLO
- [GSMA-MAGRIVAS-2016] GSMA Agricultural Value-added Services Toolkit 2.0 2016
- [ICT Update-82] Drones for Agriculture, ICT Update issue 82, April 2016 http://ictupdate.cta.int/category/issues/82-drones-for-agriculture/
- [ILO-2000] Jürgen Schwettmann, International Labour Office (ILO), Cooperatives and Employment in Africa, 2000, https://www.microfinancegateway.org/sites/default/files/mfg-en-paper-cooperatives-and-employment-in-africa-2000.pdf
- [Jellema, Meijninger, Addison – Feb2015] Andre Jellema, Wouter Meijninger and Chris Addison, Open Data and Smallholder Food and Nutritional Security, CTA Working Paper 15/01, February 2015, https://publications.cta.int/media/publications/downloads/1842_PDF.pdf?
- [Jellema 2016] Agriculture Open Data Package, http://agpack.info/ accessed January 2016, Andre Jellema, lead author, GODAN
- [Kende_McMillan_Theodorou_2016] Michael Kende, Rory Macmillan and Yiannis Theodorou, Regulatory and policy trends impacting Digital Identity and the role of mobile – Considerations for emerging markets – GSMA – October 2016

http://www.gsma.com/mobilefordevelopment/wp-content/uploads/2016/10/Regulatory-and-policy-trends-impacting-Digital-Identity-and-the-role-of-mobile.pdf

- [Parkhurst 2015] Robert Parkhurst, Environmental Defense Fund, Who will protect farmers' privacy in the big data boom?, blog post, November 2015, http://blogs.edf.org/growingreturns/2015/11/10/who-will-protect-farmers-privacy-in-the-big-data-boom/
- [Pye-Smith 2015] Climate solutions that work for farmers, Charlie Pye-Smith, CTA, 2015 https://publications.cta.int/media/publications/downloads/1867_PDF_PBBJWiT.pdf
- [Rahman-Fong 2016] Innovate for Agriculture – Young ICT entrepreneurs overcoming challenges and transforming agriculture, Reem Rahman, Jocelyn Fong, CTA 2016, https://publications.cta.int/media/publications/downloads/1924_PDF_ILGFWF6.pdf
- [Rapsomanikis 2015] FAO Smallholder Farmers' Dataportrait http://www.fao.org/family-farming/detail/en/c/385074/
- [Rasmussen 2016] From Precision Agriculture to Market Manipulation: A New Frontier in the Legal Community, Minnesota Journal of Law, Science & Technology Vol. 17 Issue 1, 2016 http://scholarship.law.umn.edu/mjlst/vol17/iss1/9/?utm_source=scholarship.law.umn.edu%2Fmjlst%2Fvol17%2Fiss1%2F9&utm_medium=PDF&utm_campaign=PDFCoverPages
- [Sykuta 2016] Big Data in Agriculture: Property Rights, Privacy and Competition in Ag Data Services, IFAMA (International Food and Agribusiness Management Review), Special Issue - Volume 19 Issue A, 2016, http://www.ifama.org/resources/Documents/v19ia/320150137.pdf
- [USDA/USAID 2016] Proceedings of the USDA/USAID International Food Assistance and Food Security Conference, October 2016, De Moines, Iowa, USA https://agrilinks.org/library/event-resources-usdausaid-international-food-assistance-and-food-security-conference Big data session: https://agrilinks.org/sites/default/files/resource/files/Big%20Data%20Presentations_0.pdf
- [WB_GSMA_DI_2016] Digital Identity: Towards Shared Principles for Public and Private Sector Cooperation – A joint World Bank Group – GSMA – Secure Identity Alliance Discussion Paper, July 2016 http://www.gsma.com/mobilefordevelopment/wp-content/uploads/2016/07/Towards-Shared-Principles-for-Public-and-Private-Sector-Cooperation.pdf
- [WB_ID2016] Country ranking on identity scheme (31st Jan 2016): http://databank.worldbank.org/data/download/IDS/WB_ID4D_dataset.xlsx
- [WB_ID4D_2016] Identification for Development – Strategic Framework – January 2016 http://pubdocs.worldbank.org/en/179901454620206363/Jan-2016-ID4D-Strategic-Roadmap.pdf

Annex 2 – Questionnaire to farmer organisations in Madagascar

Are you profiling your members (i.e. collecting a series of information about the member apart from usual identities)? **Collectez-vous des informations sur vos membres en dehors de leur identité (nom, prénom et lieu de résidence) ?**	0. No 1. Yes **0. Non** **1. Oui**
What is the type of information collected? **(Multiple answers are possible)** **Quels types d'information collectez-vous ?** **(plusieurs choix possible)**	1. Farmer data (phone number, fields, location, GPS coordinates, etc.) 2. Field data (soil, field map and size, history of crops, etc.) 3. Input data (crop to be planted, inputs/fertiliser needs, etc.) 4. Production data (crop, size, yields, grade of product, timing, etc.) 5. Activity data (action in the field like weeding, fertiliser/pesticide application, etc.) 6. Pest and disease data (event occurring and treatment) 7. Financial data (credit record, credit needs, income, etc.) 8. Others_____ **1. Données sur l'exploitation agricole et le paysan (numéro de téléphone, champs, adresse détaillée, coordonnées GPS, etc.)** **2. Données sur les champs (type de sols, cartographie des champs et taille, historique des cultures, etc.)** **3. Données sur les intrants agricoles (semences à planter, besoins en fertilisants, etc.)** **4. Données concernant la production (produits cultivés incluant la variété et la catégorie, taille des parcelles, taille supposée et mesurée des récoltes, calendrier du cycle de production, etc.)** **5. Données sur l'activité (recensement des actions sur les parcelles comme**

	défrichage/désherbage, application de pesticides, etc.) **6. Données sur les maladies (maladies détectées sur les plants et traitements)** **7. Données financières (historique de crédit, besoin en crédit, revenu, business plan de l'exploitation etc.)** **8. Autre à préciser ci-dessous** _____
What is the use of the information collected? **(Multiple answers are possible)** **Quelle utilisation faites-vous des informations collectées ?** **(plusieurs choix possible)**	1. Credit 2. Insurance 3. Group buying for inputs/seeds/etc 4. Group selling/business matching 5. Extension services, 6. Others (traceability, statistics, etc.) Credit **1. Crédit** **2. Assurance** **3. Achats groupés pour les intrants agricoles** **4. Ventes groupées et mise en relation avec des acheteurs** **5. Service de vulgarisation et bonnes pratiques agricoles** **6. Autres (traçabilité, statistiques, etc.) à préciser ci-dessous** _____
Are you using national databases or other government information systems (national ID cards system, census, etc.) to gather information about your members? **[If not skip next question]** **Utilisez-vous des bases de données nationales ou des informations gouvernementales (registre des cartes nationales d'identité, données de recensements, etc.) pour enrichir les données sur vos membres ?** **[si vous répondez non, passez la question suivante]**	1. No 2. Yes **1. Non** **2. Oui**
What is the type of data you pull and for what purpose? **(Multiple answers are possible)** **Quels types de données récupérez-vous et pour quel(s) usage(s) ?** **(plusieurs choix possible)**	1. Personal and family information 2. Field/land information (ownership, etc.) 3. Credit information (credit record) 4. Other information_____ **1. Information personnelle et familiale** **2. Information sur les parcelles (certificats de propriété etc.)**

	3. Information sur les crédits en cours et passés (historique de remboursements etc.) **4. Autre à préciser ci-dessous** _____
How is the information managed? **Comment gérez-vous les données collectées ?**	1. Paper 2. Excel 3. Farmer membership management software **1. Sur support papier** **2. Sous Excel** **3. En utilisant un logiciel dédié de gestion de membres**
How often is the information updated? **Avec quelle périodicité mettez-vous à jour les informations sur vos membres ?**	1. A few times a year when needed (beginning of a crop cycle, etc.) 2. Once a year 3. Less often but systematically through a kind of census process 4. Rarely and not systematically **1. Plusieurs fois par an quand c'est nécessaire (début d'un cycle agricole, etc.)** **2. Une fois par an** **3. Moins souvent mais de façon systématique en utilisant un processus de type recensement** **4. Rarement et pas de façon systématique sur l'ensemble des membres**

Annex 3 – Sustainable Development Goals and farmer profile information

In this section, we identify the Sustainable Development Goals[41] and targets that could be instantiated through the exploitation of farmer profile information. The mapping is presented in the table below with goals, targets, and corresponding profile information.

SDG	Target	Profile information
1 - No poverty	1.1 By 2030, eradicate extreme poverty for all people everywhere, currently measured as people living on less than US$1.25 a day.	Farmers income Access to specific subsidy schemes
2- Zero hunger	2.3 By 2030, double the agricultural productivity and incomes of small-scale food producers, in particular women, indigenous peoples, family farmers, pastoralists and fishers, including through secure and equal access to land, other productive resources and inputs, knowledge, financial services, markets and opportunities for value addition and non-farm employment.	Production information (yields, input usage, etc.) Credit information Farmers income Field information (land rights)
	2.4 By 2030, ensure sustainable food production systems and implement resilient agricultural practices that increase	Production information (yields, input usage, etc.)

[41] https://sustainabledevelopment.un.org/?menu=1300

	productivity and production, that help maintain ecosystems, that strengthen capacity for adaptation to climate change, extreme weather, drought, flooding and other disasters and that progressively improve land and soil quality.	Weather information Disaster information Field information (soil)
	2.5 By 2020, maintain the genetic diversity of seeds, cultivated plants and farmed and domesticated animals and their related wild species, including through soundly managed and diversified seed and plant banks at the national, regional and international levels, and promote access to and fair and equitable sharing of benefits arising from the utilization of genetic resources and associated traditional knowledge, as internationally agreed.	Crop information
	2.c Adopt measures to ensure the proper functioning of food commodity markets and their derivatives and facilitate timely access to market information, including on food reserves, in order to help limit extreme food price volatility.	Market information Selling information
8 – Decent work and economic growth	8.2 Achieve higher levels of economic productivity through diversification, technological upgrading and innovation, including through a focus on high-value added and labour-intensive sectors.	Production information (use of equipment)
	8.10 Strengthen the capacity of domestic financial institutions to encourage and expand access to banking, insurance and financial services for all.	Credit information Insurance Information
	8.a Increase aid for trade support for developing countries, in particular least developed countries, including through the Enhanced Integrated Framework for Trade-Related Technical Assistance to Least Developed Countries.	Access to specific subsidy scheme
10 – Reduce inequalities	10.1 By 2030, progressively achieve and sustain income growth of the bottom 40% of the population at a rate higher than the national average.	Farmers income
12 – Responsible consumption	12.4 By 2020, achieve the environmentally sound management of chemicals and all wastes throughout their life cycle, in accordance with agreed international frameworks, and significantly reduce their release to air, water and soil in	Production information: use of treatments, fertiliser and equipment

and production	order to minimize their adverse impacts on human health and the environment.	
	12.8 By 2030, ensure that people everywhere have the relevant information and awareness for sustainable development and lifestyles in harmony with nature.	Advisory service information
13 – Climate action	13.1 Strengthen resilience and adaptive capacity to climate-related hazards and natural disasters in all countries.	Production information Weather information Disaster information
	13.3 Improve education, awareness-raising and human and institutional capacity on climate change mitigation, adaptation, impact reduction and early warning.	Production information Weather information Disaster information
	13.b Promote mechanisms for raising capacity for effective climate change-related planning and management in least developed countries and small island developing states, including focusing on women, youth and local and marginalized communities.	Production information Weather information Disaster information
15 – Life on land	15.3 By 2030, combat desertification, restore degraded land and soil, including land affected by desertification, drought and floods, and strive to achieve a land degradation-neutral world.	Field information
	15.6 Promote fair and equitable sharing of the benefits arising from the utilisation of genetic resources and promote appropriate access to such resources, as internationally agreed.	Farmers income Selling information
	15.8 By 2020, introduce measures to prevent the introduction and significantly reduce the impact of invasive alien species on land and water ecosystems and control or eradicate the priority species.	Production information (including pest and disease attacks, and other activities such as weeding)
17 – Partnerships for the goals	17.11 Significantly increase the exports of developing countries, in particular with a view to doubling the least developed countries' share of global exports by 2020.	Certification data Selling information

The Technical Centre for Agricultural and Rural Cooperation (CTA) is a joint international institution of the African, Caribbean and Paci c (ACP) Group of States and the European Union (EU). Its mission is to advance food security, resilience and inclusive economic growth in Africa, the Caribbean and the Paci c through innovations in sustainable agriculture.

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