Statistical Journal of the IAOS 38 (2022) 163–180
DOI 10.3233/SJI-210908
IOS Press

Facilitating data use for decisionmaking: 50x2030's approach

Paul Winters[a,*], Emily Hogue[b], Michael Steiner[c], Chista Keramati[d], Maryam Rokhideh[d] and Sarah Eissler[d]
[a]*Keough School of Global Affairs, University of Notre Dame, South Bend, Indiana, USA*
[b]*Independent Consultant*
[c]*50x2030 Data Use Component Manager, International Fund for Agricultural Development, Rome, Italy*
[d]*50x2030 Data Use Component, International Fund for Agricultural Development, Rome, Italy*

Abstract. Data are a critical input into policy and can help central actors make informed and evidence-based decisions. Yet, data within a country are not always presented and analyzed in a manner that facilitate their use. Analysis of data use rarely moves beyond a supply and demand framework and fails to recognize that the data ecosystem within a given country is complex and includes a host of actors. Recognizing this complexity, the 50x2030 Data Initiative is exploring a new approach to understanding and facilitating data use. This paper provides an overview of the methodological approach to understanding data use. It examines early country experiences with the approach in order to ultimately help countries reach the end goal of data-driven decisionmaking in agriculture and rural development. Focusing on results from Cambodia, Uganda, and Georgia, the paper presents lessons on the methodological approach as well as findings on the constraints that are most hindering data use in the examined countries.

Keywords: Data use, 50x2030 Initiative, agricultural data, agricultural statistics, data ecosystem

1. Introduction

Data alone cannot solve development problems since people in governments, civil society and the private sector are ultimately the central actors in transforming data into useful information that can improve livelihoods and lives [1]. Yet, data are a critical input into policy and can help those central actors make informed and evidence-based decisions. In fact, country-level availability of data on the Millennium Development Goals indicators has been shown to be associated with better development outcomes [2]. And two-thirds of leaders responding to the *Listening to Leaders Survey* report using domestic development data in their decisionmaking, particularly national statistics [3].

Using data appropriately is critical to agriculture and central to the achievement of *Sustainable Development Goal 2: Zero Hunger* (SDG2); agriculture is closely linked to food security (SDG2.1), productivity and income of small-scale producers (SDG 2.3), and sustainable and resilient agricultural practices (SDG 2.4). Since 2015, progress in reducing hunger has stalled and gotten worse under COVID-19 with an estimated 720 to 811 million people facing hunger in 2020, and an estimated 30% of the global population, or 2.7 billion people, not being able to afford a healthy diet [4]. Further, the recently published Sixth IPCC Assessment Report concludes that virtually every region of the world will experience concurrent effects of climate change, challenging the world's resilience and adaptation capacity. It also notes, from 2007 to 2016, agriculture, forestry, and other land use activities contributed 23% of the total net anthropogenic emissions of GHGs, second only to energy combustion [5]. In response to these global challenges, the United Nations convened a Food System Summit in September 2021 highlighting the

*Corresponding author: Paul Winters, Keough-Hesburgh Professor of Global Affairs, Keough School of Global Affairs, University of Notre Dame, South Bend, Indiana, USA. Tel.: +1 240 476 1911; E-mail: pwinters@nd.edu.

© 2022, International Bank for Reconstruction and Development / The World Bank, Food and Agriculture Organization of the United Nations and International Fund for Agricultural Development This is an Open Access article distributed under the terms of the Creative Commons Attribution-Non Commercial License (CC BY-NC-SA 3.0 IGO).

importance of these issues. It is increasingly clear that averting hunger requires data and synthesis [6]. But it also requires that those data are used for decisions that seek to address these issues.

This assumes, of course, that data are available and useful to decisionmakers within a country and are presented in a manner that facilitate their use. Yet, evidence from low- and middle-income countries suggests that data provided by National Statistical Offices (NSOs) (i) often focus on international partners rather than governments and domestic users, (ii) use dissemination strategies that are not in line with what governments want, and (iii) are not as easy to use and as accessible as possible [7]. The emphasis in promoting data is often on data production and a standard model that sees data through a supply and demand framework, which, while useful for identifying barriers, is potentially oversimplified and not capturing does not capture the nuanced roles of different actors [8,9]. The current approach needs to be rethought and evolve in new directions.

One starting point is understanding the data ecosystem, which is the community of organizations and individuals that engage with data, the data assets with which they interact, and the rules and structures that govern their interactions. The data ecosystem within a given country is complex including a host of actors beyond NSOs and government decisionmakers. This includes knowledge intermediaries, referred to as brokers, translators, etc., who ultimately promote communication and knowledge-sharing between researchers and decisionmakers to facilitate data use [7,10,11]. To facilitate the use of data by decisionmakers requires understanding of the data ecosystem, how it operates, and where barriers and constraints may limit effective data use [12].

Enhancing the data ecosystem is the approach taken by the 50x2030 Data Initiative, which seeks to increase the capacity of 50 low and lower middle-income countries to produce, analyze, interpret, and apply data to decisions in the agricultural sector [13,14]. The Initiative is a partnership between those 50 countries, several donors, and three implementing partners: the Food and Agriculture Organization of the United Nations (FAO), the International Fund for Agricultural Development (IFAD) and the World Bank. Each implementing partner provides leadership in three distinct areas: Data Production (FAO), Data Use (IFAD) and Methods and Tools Development (World Bank).

The Initiative's Data Use activities seek to strengthen a country's data ecosystem by improving capacities, communications, policies, and practices [13]. As a first step, this requires an understanding of the data ecosystem and data use within a given country [12]. The objective of this paper is to describe the methods employed by 50x2030 to understand data use, provide examples of how this methodological approach was implemented in the few initial countries, and to draw lessons for the Initiative and others for future efforts to understand and expand data use. Toward this end, Section 2 provides the framing for the approach taken, Section 3 the methods used to assess that framing, Section 4 examples from the implementation of the methodological approach in first few countries, Section 5 an overview of findings on data use, and Section 6 the implications and way forward for understanding data use.

2. Data use and decisionmaking

Given the desire to enhance the use of evidence in policy decisions, an emerging literature has focused on understanding the limitation of current approaches and identifying what alternative approaches might work best given these limitations. As noted in the introduction, a key aspect of this perspective is to move beyond a supply and demand framework, which does not capture the range of different actors in the data ecosystem nor their nuanced roles [8,9]. This may require thinking beyond a "pipeline model" which sees evidence going from researchers to policymakers, to a greater understanding of policy processes more generally and where evidence is used in this process [15].

A starting point for considering data use is recognizing that data currently being generated at the country level is being used. For example, an analysis of seven NSOs in low- and middle-income countries shows clear evidence that NSO websites and data portals are widely used to access data and that search engines drive interested parties to NSO websites [16]. Further, respondents from the *Listening to Leaders Survey*, which includes responses from leaders from the government, the private sector, CSO and development partners, indicate that they do employ evidence to diagnose problems, set priorities, and design or inform implementation strategies (Masaki et al., 2017). Further, 85 percent of respondents in the Masaki et al. study note that government data are the most commonly used [3].

The data use issues highlighted in the literature are less about whether data are seen as valuable and more about the appropriateness of the data for domestic use and for decisionmaking. For example, Sethi and Prakash find that NSO officials surveyed are most likely

to see international development organization as their most important (60%) and most frequent user (64%) followed by research organizations (58% and 59% respectively) [9]. Government officials only come after these in priority. If the primary audience is not governments or other domestic decisionmakers, this potentially skews what is collected, presented and analyzed to these other audiences. Sethi and Prakash also find that when NSO and Ministry officials are asked about how to encourage data use, both NSO and Ministry officials note that data should be easier to access, easier to use, and easier to obtain on websites [9]. These limitations inhibit data use.

When the right data are available to decisionmakers, it has proven to be valuable. Synthesizing the results of 31 randomized control trials, BenYishay and Parks find that the provision of location-specific data to public officials can improve resource allocation and service delivery outcomes, particularly when the information that public officials receive is legible, actionable, and inclusive of both aid flows and population needs [17]. Yet, the evidence suggests that development data are often insufficiently granular, of questionable accuracy, too infrequent and not timely, and often fragmented across numerous agencies in incompatible formats inhibiting its value [1,18]. When assumptions around the usefulness of collected data breakdown, there is a risk of producing data graveyards of unused data, instead of data for action [18].

Facilitating data use then requires recognizing that data are being used, but that it could be improved through a greater understanding of the links between data and decisionmaking and, more broadly, the policy decision process. A key part of understanding these links is to understand the data ecosystem and the barriers and constraints within that system [19]. In developing guidelines for enhanced statistical capacity, Paris 21 highlights that analyzing the data ecosystem is a key starting point for facilitating change [12]. The focus on data ecosystems has been driven by political/institutional initiatives to promote data use as well as the emergence of digital technologies which facilitates greater access and allows for integration of distinct data sets [11]. It recognizes the key role of intermediaries or translators in the data ecosystem, which act as knowledge brokers to facilitate data use [7,10].

Building on this recent literature, this is the methodological approach taken by the Data Use component of the 50x2030 Data Initiative [20]. In the next section, we discuss the approach in detail. This is followed by a discussion of insights from the initial use of the approach in the first few countries in which the data ecosystem assessment has been implemented.

3. Methodological approach

Following the guidance provided in Paris 21 [12], the 50x2030 Agricultural Data Use Assessment identifies the stakeholders and structures of the data ecosystem, how they interact, and where enablers and constraints to data use reside in the system. The assessment uses a participatory approach that engages key stakeholders of the 50x2030-supported survey program, both from the government and external entities, from designing the data collection methods through to the validation of results. Within the assessment process, data are used to create a data ecosystem map that illustrates key findings on the various relationships between stakeholders, data flows, and the main constraints to data use within the system. The final assessment report and map is a tool to prioritize needs and to identify interventions that have the potential to enhance stakeholder collaboration and ensure the effective use of 50x2030 survey data in decisionmaking.

3.1. Overview of data use framework

As outlined in the 50x2030 Data Use Initiative Guide [20], the data use assessment examines data use within a country's data ecosystem, defined as the community of actors, stakeholders, and entities who engage with data, the data assets (data sets, data products, platforms, tools, technologies) with which they interact, and the rules, norms, and structures that govern those interactions (policies, cultures, organizational structures, etc.). The assessment aims to identify enabling or constraining factors for data use within the ecosystem, guided by a data use conceptual framework presented in Table 1. This framework enables the data use assessment team to understand, examine, prioritize, and address factors that enable data use, or inversely, constraint data use if absent in the country context.

Based on this framework, in each country, the assessments are designed to answer four research questions on:

1. the current uses of agricultural data and perceived benefits;
2. the key factors constraining data use;
3. human and institutional capacities that need to be strengthened to promote data use; and
4. interaction/collaboration between ecosystem stakeholders and possible means to improve it.

The assessment methodology uses three complementary methods to answer those questions:

Table 1
The 50x2030 data use conceptual framework

Data use factor	Definition	Sub-factors
Demand	Decisionmakers want to use data to make decisions	Decisionmakers believe that data-based decisions are beneficial Political and cultural beliefs and norms enable data use
Expertise	Decisionmakers know what data are needed and how to use it	
Availability	Data are produced and exist	
Access	Data intermediaries and decisionmakers have access to available data	
Awareness	Data intermediaries and decisionmakers know data are available and accessible	
Utility	The data are relevant and useful to decisionmakers to inform necessary decisions	Data are the types (variables, disaggregation, periodicity, etc.) needed for targeted decisions Data are in needed formats and products Information (analyzed data) is in needed formats and products
Trust	Data intermediaries and decisionmakers trust the data and believe they are useful and accurate	Data quality due to weak capacity Political interference in data Competition with other data sources

1. a review of existing material;
2. quantitative data collection and analysis through online surveys; and
3. qualitative data collection and analysis through key informant interviews (KIIs) and open-ended questions in the online surveys.

As part of answering the research questions, particularly question (4), a data ecosystem map is created. The data ecosystem map seeks to identify the relationships and level of interaction/collaboration between different stakeholder groups and the various data users, data intermediaries, and decisionmakers in the system. It provides a visual representation of the data ecosystem and how value is shared between stakeholders, including the data assets (data sets, data products, platforms, tools, technologies), and the norms and structures that govern those interactions (policies, cultures, organizational structures, etc.). Data ecosystem maps can help to identify where exactly there are gaps and barriers to data use in the system and how to promote a collaboration between actors to ensure access. Creating a visual map that illustrates how data are being accessed, used, and shared by stakeholders facilitates discussions on ways to improve data use by stakeholders themselves. While it may make sense to conduct a broader data ecosystem mapping that covers all sectors, the data ecosystem mapping conducted for this analysis focuses on the agricultural data ecosystem to ensure its analytical utility for 50x2030. Broader efforts may be ultimately more cost effective provided that they ensure sufficient sectoral detail.

3.2. Implementing the data use framework

To initiate the assessment, the Data Use team holds a meeting with the 50x2030 Country Coordination Group (CCG) to outline the process and methods, and to gather feedback on any needed modifications to standard data collection methods. The 50x2030 CCG is a group of representatives from the partner government (MoA, NSO, other agencies), 50x2030 implementers, development partners, and other key organizations working in agricultural data in the country. The primary role of the CCG is to guide and help facilitate the work of 50x2030 in country.

Following that meeting, members of the CCG are asked to offer recommendations on relevant existing materials related to agricultural data to review and to provide input on the standard survey questionnaire, interview protocol, and proposed lists of respondents for surveys and KIIs in the country. With the CCG's input, the review is conducted and the data collection instruments are revised, finalized, and sent for translation to the local language, if needed.

The review of existing materials covers documents on national data laws, agricultural policies and plans, previously conducted assessments/reviews on agricultural data use, and stakeholder websites to create a description of the country's agricultural data ecosystem. This description includes statistical capacity, data laws/policies, key stakeholders, data needs for policies and monitoring frameworks, data assets and sources, and data technologies. The review informs the adaptation of the survey questionnaire and interview protocol to each specific country context.

Quantitative data collection and analysis are based on a written online survey disseminated to stakeholders using Google Forms. The assessment process was designed in 2020 during the COVID-19 pandemic, making online surveys the most viable choice. The online format generally worked very well and so became the

standard approach. The starting point is a standardized questionnaire that is aligned to the seven factors of the Data Use Framework noted previously (see Table 1) and adapts the questionnaire to the national context. The surveys include questions that request either a quantitative/categorical or qualitative/open-ended response. The quantitative data inform indicators under factors of the Data Use Framework and help identify the most binding constraints and enabling factors for data use in the data ecosystem.

Qualitative methods use key informant interviews (KIIs) and open-ended survey questions to delve deeper into critical questions, providing rich explanatory information on current data use, the barriers to it, and possible solutions to overcome those barriers. The interview approach is semi-structured and organized into sections based on the research questions. Interviewers follow a guide that identifies the objectives of each section and provides suggested questions, ultimately relying on the interviewer's judgement to determine which questions to ask and in what order. The approach is designed to produce the maximum amount of quality data that allows for contextual and cultural adaptation for each country through a conversational approach that leverages the respondent's areas of knowledge and interest or concern.

Qualitative data from the KIIs are supplemented by open-ended questions in the quantitative online survey. Approximately 20% of the survey questions in the standard questionnaire are open-ended and designed to allow respondents to provide more explanatory and descriptive information on their responses to categorical or ordinal questions. While qualitative data from the surveys is typically briefer and more concise, it has the value of drawing from a larger sample.

To understand and map the data ecosystem in each country, network questions are drawn predominately from the online surveys and supplemented by the KIIs. Analysts used the software, Kumu to visualize and map the data ecosystem [21]. The end result provides a visual that can be discussed with stakeholders and be the basis for identifying constraints to enhancing the data ecosystem. Thus, in each country, a preliminarily data ecosystem map is used as a participatory tool during the validation workshop and subsequent planning meetings to generate helpful insights to guide interventions and activities. To address the barriers to data use, country participants in the planning workshop examine the major constraints as indicated by the findings and provide ideas of interventions that can help reduce the constraints and get the ecosystem to a shared vision or the ideal state. Together with the country stakeholders and Data Use Assessment team, the principles of the ideal state are identified and agreed upon in each country.

3.3. Sampling

The samples for both the survey and KIIs are designed to cover three types of stakeholders – producers, intermediaries, and decisionmakers – which are defined as follows:

- A *data producer* is an individual or entity that produce data following the steps of data collection, data curation and preparation, and data dissemination.
- A *data intermediary* is an individual or entity that takes existing summary reports, summary tables, and microdata sets and adds value to them by conducting and interpreting analyses to answer questions and make recommendations for action.
- A *decisionmaker* is an individual or entity that applies the data to answer questions and inform decisions related to programs, policies, or investments.

Both the survey and interview samples are designed to also cover five general sectors:

1. the public sector (NSO, Ministry of Agriculture, and other relevant ministries);
2. development sector;
3. research and academic sector;
4. private sector; and
5. civil society, including the media.

Sampling for the online survey is purposive and targets 100–500 potential respondents identified by the 50x2030 CCG as stakeholders in the ecosystem. The sample and list of potential respondents is adapted to respond to preliminary knowledge of the ecosystem. For example, countries with a stronger statistical system and longer history of data-related development interventions, such as Uganda, will likely yield a larger potential sample than countries that have not historically received as much support, such as Cambodia and Georgia. Additionally, the sample is adapted to cover country-specific agricultural issues, such as including private sector actors working in commodities that are prioritized through national agricultural policies.

Selection for the KIIs uses the same initial sample collected for the online survey. Approximately 20–24 knowledgeable and purposively selected stakeholders are sampled for KIIs to provide expert knowledge and perspectives on data use issues. However, in all countries to date, the Data Use team prioritized selecting

higher-level government officials for a KII to ensure their perspectives were included in the assessment, as they were determined less likely to complete an online survey.

3.4. Data collection logistics

The qualitative and quantitative data collection is typically conducted simultaneously and lasts from one to six months, depending on the schedules and events in the country. Surveys are managed by members of the Data Use team. KIIs are conducted by contracted local firms with expertise in qualitative research and topical knowledge on agriculture and/or data systems.

The mode of KIIs varies across countries and is dependent both on pandemic-related factors, such as restrictions related to COVID 19, and cultural factors, such as the level of interviewers' and respondents' preference for meeting in person or using a virtual platform. In Cambodia, nearly all KIIs were conducted in person in December 2020–January 2021, which was a less intensive time of the pandemic for Cambodia, whereas all KIIs in Georgia had to be conducted virtually due to COVID 19. There have been no notable differences in the quality of data through in-person or virtual delivery.

In all countries where English is not an official language, instruments are finalized in English and then translated into the official national language. Online surveys are distributed in both English and the national language, as some respondents may be foreign nationals who speak more English than the national language. KIIs are audio recorded, with the respondents'permission, and transcribed into English by a local firm. The transcribed KIIs serve as the raw data for analysis by the Data Use team. If the respondent does not authorize audio recording the interview, interviewers and notetakers take extensive notes to provide for analysis. Local contacts are consulted if any terms or translations are not clear to ensure data are properly interpreted.

Lastly, incentive gifts can be used to show appreciation and compensate respondents for their time, if deemed culturally appropriate. In Cambodia, the CCG and local research firm advised to provide incentive gifts to both people being interviewed and internet survey respondents. For the online survey, respondents were offered the choice of a $5 USD phone card to respondents'phone service provider of choice, or to have a $5 donation made in their name to a local COVID relief fund. KII respondents received either a $10 phone card or a $10 donation made in their name to a local COVID relief fund. The use of a small local handicraft was considered as an incentive gift for KII respondents, but not used since nearly all interviews were conducted virtually. Local contacts advised not to provide incentive gifts in Uganda and Georgia, as they were not considered appropriate.

3.5. Data analysis

Data analysis begins with distinct analytical processes for the quantitative and qualitative data. Analyzed data are then later synthesized in the last step of the analytical process to formulate robust answers to each research questions. A detailed description of the steps for the qualitative, quantitative, and mixed-method synthesis of analysis is provided in Fig. 1. As noted, the methodological approach is reviewed and adapted to each country context based on input from CCG and the literature review.

Data analysis from the online questionnaire involves two distinct analytic approaches based on the type of survey questions. For closed-ended questions, summary tables and figures are created using pivot table and chart tools available in Google Sheets or Excel. Responses to these questions are cross tabulated mainly based on the respondents' stakeholder type or their sector (both self-identified). Open-ended (descriptive) survey questions are analyzed qualitatively, using software and tools such as NVivo,[1] Excel, and Microsoft Word. Distinct quotes or statements that provide key insights to an issue are pulled out to be used to illustrate critical points. When open-ended response questions can be grouped and categorized into themes, the data are organized into summary tables to show the incidence of certain themes using NVivo software. Key insights, summary tables, and figures are organized in one document under each survey question. This document will be used by the team as a resource for future reference. Next, findings from the online questionnaire are synthesized under research questions before triangulating findings with the qualitative findings and the written report.

The KIIs are analyzed guided by qualitative thematic analysis techniques using a codebook that includes a mix of deductive and inductive codes. Deductive codes are those that are identified as relevant concepts or themes from existing literature, whereas inductive codes are identified by reviewing the raw qualitative data and

[1] NVivo by QSR International is a software program used to manage, process, analyze, and visualize qualitative (i.e., non-numerical) data.

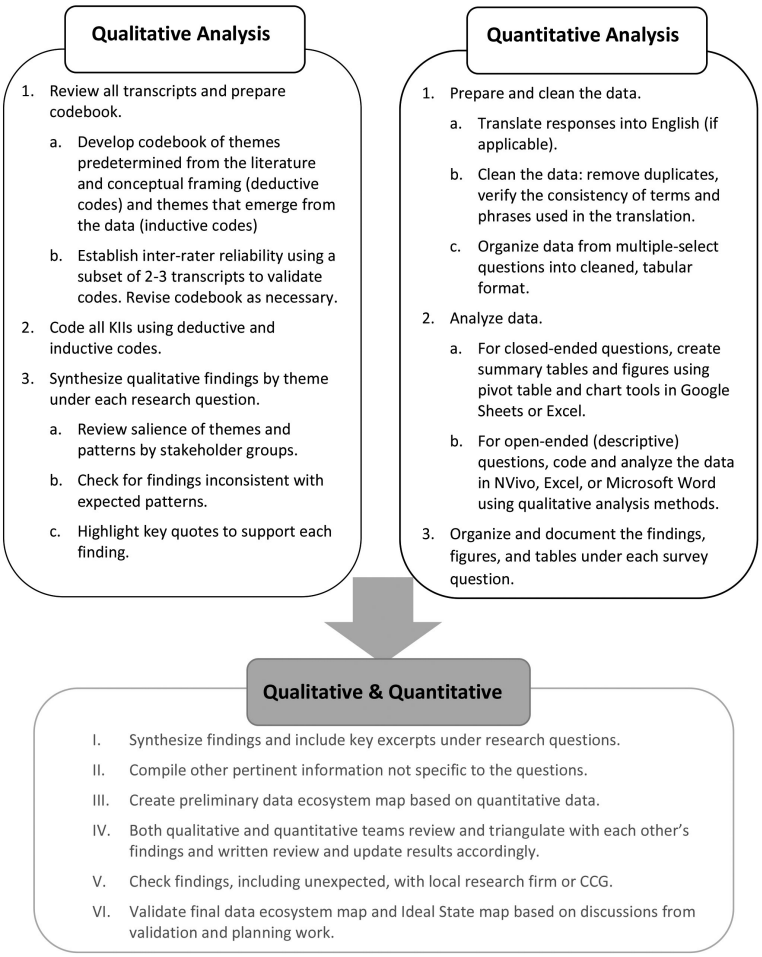

Fig. 1. Overview of mixed methods approach.

selecting emergent themes that were not previously determined from the existing literature. As a first step, analysts jointly develop the codebook (consisting of a list of codes with associated definitions) with deductive codes that align directly with the research questions and the seven factors from the developed Data Use conceptual framework. After cleaning and reviewing all transcripts, the analysts jointly determine inductive codes, or those that emerge from repeated themes throughout the data. This step enables the qualitative analysis to capture emergent themes that are relevant for each country context. The analysts establish inter-rater reliability prior to coding all transcripts using a subset of two to three transcripts for each country. Establishing inter-rater reliability ensures that the analysts apply the codes consistently while analyzing all transcripts, a critical step for establishing the reliability of the analytic process. The analysts then code all transcripts using NVivo software. After all KIIs are coded, the analysts synthesize key findings based on most salient codes and major themes and use key excerpts or quotes to exemplify findings.

A mixed method approach is used to synthesize both qualitative and quantitative findings, wherein the quantitative or qualitative data form the foundation of the answer to each research question. Depending on the quality of the quantitative and qualitative data related to the research question, each type of data may form the primary basis for the answer to the research question while data from the other method is used to supplement, explain, or illustrate the answer. For example, quantitative data provides the strongest answer to the research question on which factors pose the greatest constraints and, so, are used to create the primary answer, while qualitative data are used to explain or elaborate on the answers. For the question on capacities, qualitative data often proves stronger to identify the most salient capac-

Table 2
Application of data use component in Cambodia, Uganda and Georgia

	Cambodia	Uganda	Georgia
Dates of field work	December 2, 2020–February 25, 2021	April 30, 2021–June 17, 2021	August 6–December 15, 2021
Partners	1). Nuppun Research and Consulting (conducted KIIs) 2). National Institute of Statistics of Cambodia (NIS) 3). Cambodian Ministry of Agriculture, Forestry and Fisheries (MAFF)	1). Associates Research Trust Uganda (conducted KIIs) 2). Uganda Bureau of Statistics (UBOS) 3). Uganda Ministry of Agriculture, Animal Industry and Fisheries (MAAIF)	1). ACT Research Georgia (conducted KIIs) 2). National Statistics Office of Georgia (GEOSTAT) 3). Ministry of Environmental Protection and Agriculture of Georgia (MEPA)
# key informant interviews	23	25	23
# respondents completing online questionnaire	49	100	66
Notes/observations	– KIIs conducted in English or Khmer – One interviewer from Nuppun conducted all KIIs – Nearly all KIIs were audio recorded – Letter endorsing the survey by IFAD Cambodia Country Director was sent to all people being interviewed and all people receiving the Internet Survey. Letter was in English and Khmer. – An email message was sent to all people to be interviewed by IFAD HQ staff, requesting their cooperation for an interview. Email message included letter of endorsement, and informed people they would be contacted by NUPPUN. NUPPUN followed up to arrange time for an interview. – An email message was sent by IFAD HQ staff to people receiving the Internet Survey, requesting them to complete the survey by clicking on a link in the message. Email messages included letters of endorsement in English and Khmer. The email message was in English and Khmer. People could complete the Internet Survey in either English or Khmer. – People who were interviewed and people responding to the Internet Survey were offered an incentive of either a telephone card (of their choice of phone companies) or a donation in their name to a Cambodia COVID relief fund	– KIIs conducted in English – One interviewer from Associates Research conducted all KIIs – Nearly all KIIs were audio recorded – Letter endorsing the survey was sent to all people being interviewed and people receiving the Internet Survey. Letter was signed by IFAD Director of RIA, and Director of East and Southern Africa Division. – An email message was sent to all people to be interviewed by IFAD HQ staff, requesting their cooperation for an interview. Email message included letter of endorsement, and informed people they would be contacted by Associates Research. Associates Research followed up to arrange time for interview. – An email message was sent by IFAD HQ staff to people receiving the Internet Survey, requesting them to complete the survey by clicking on a link in the message. Email messages included letters of endorsement in English. The email message was in English, and the Internet Survey was in English. – No incentive gifts were offered to either interview respondents or Internet Survey respondents – In-depth qualitative data collection training conducted for interviewers by 50x2030 qualitative analysis team.	– KIIs conducted in English or Georgian – Three interviewers from ACT conducted KIIs – Nearly all KIIs were audio recorded – Letter endorsing the survey by IFAD Georgia Country Director was sent to all people being interviewed and all people receiving the Internet Survey. Letter was in English and Georgian. – An email message was sent to all people to be interviewed by IFAD HQ staff, requesting their cooperation for an interview. Email message included letter of endorsement, and informed people they would be contacted by ACT. – An email message was sent by IFAD HQ staff to people receiving the Internet Survey, requesting them to complete the survey by clicking on a link in the message. Email messages included letters of endorsement in English and Georgian. The email message was in English and Georgian. People could complete the Internet Survey in either English or Georgian. – No incentive gifts were offered to either interview respondents or Internet Survey respondents – In-depth qualitative data collection training conducted for interviewers by 50x2030 qualitative analysis team.

ity problems and can be supported by quantitative data to strengthen the answer.

4. Application to Cambodia, Uganda, and Georgia

The 50x2030 method was first implemented in Cambodia followed by Uganda and Georgia. The implementation phase in these three countries spanned from 2020 to 2021. Table 2 presents the timeline of activities in each country together with the lists of implementing partners, numbers of KIIs conducted, numbers of online survey respondents, as well as other notes and observations for each country.

In Cambodia, where there is a limited history of collecting agricultural data, the purposive sample of potential survey respondents included 100 individuals out of whom 49 people (about 50%) responded to the survey.

Respondents represented all five targeted sectors. The number of responses for each of the five sectors (public, development, research and academic, private, and civil society including the media) were relatively small in Cambodia, so results for some sectors were limited.

To enhance the results for each of the five sectors, the target number of completed internet surveys was increased in other countries, from 50 in Cambodia to 100 or more in other countries. In an attempt to get a greater number of responses. In Uganda and Georgia, where agricultural survey data have been available for longer periods of time, the purposive samples included about 300 to 400 number of potential respondents. In creating a larger list, names were received from a variety of sources and names, email addresses and phone numbers were harder to verify. Hence, more survey responses were received, but the response rate was lower at 25–30 percent.

The KIIs in Cambodia were successful in revealing interesting findings around data use that reinforced and elaborated on the quantitative findings. However, the analysis team noted that stronger training for data collectors was required to produce more robust results. Specifically, in-depth training with the local research firm in each country is needed to cover topics around best practices for qualitative data collection (including the use of probing questions), the key themes of the research study, strong notetaking skills, and strong facilitation skills. The analysis team implemented these trainings with Uganda and Georgia. Implementing the trainings prior to data collection yielded much improved interview transcripts with more detail elicited from respondents.

Results from the mixed method assessment were presented to members of the CCG and other key stakeholders in each country in a validation workshop, a critical step to validate the findings of the assessment. Two separate data validation sessions were planned for Cambodia, but because of scheduling conflicts, only one 2.5 hour session was held, which was determined to be too short of a time. A full half day session was held in Uganda and Georgia to allow for adequate time for discussion of results from the assessment.

Some high level government officials in Cambodia were sensitive about indications of problems with data access, even though the assessment and other workshop participants indicated that access was a big issue. Because of limited time in the validation workshop in Cambodia, discussion focused on key constraining factors, such as issues around access to data. In the validation workshop held in Uganda, the Data Use team presented positive results first prior to focusing on areas for improvement, to minimize any potential political sensitivities. There was no disagreement about the findings, and a good discussion. This appears partly due to the manner in which the results were presented and that there was not a Senior-level official that dominated the conversation as occurred in Cambodia.[2] The focus of the presentation and composition of the workshop participants is critical.

While the Cambodia survey experienced the lowest total number of responses, it had the highest response rate, which may have been linked to the more focused roster of potential survey respondents or the provision of an incentive gift. In Cambodia, the list of contacts for the survey was smaller than other countries and consisted of people known to work with agricultural information. More work is needed on how to improve the survey response rates in each country.

For the Cambodia data ecosystem map exercise, different map ideas were drawn focusing on data exchange between stakeholders, including data flows, data needs, and constraints to data use. Analysts used Kumu software to design an ecosystem map to visualize the distribution of the main constraints to data use for each targeted sector [21]. The Cambodia map (Fig. 2) used inference and proxy data on sources and types of data from the survey data to present the (seven conceptual) factors impeding the uptake and use of data for each sector.

Given that the initial mapping strategy and methodology was still being refined and not clearly determined yet, the Cambodia data ecosystem map focused solely on the relationship between the Cambodian National Institute of Statistics (NIS) and five stakeholder groups. Furthermore, the initial data instruments did not include the necessary questions to gather data on relationships between stakeholders nor how data was shared in the data ecosystem. Subsequent data collection tools were revised and expanded the framework to include questions on relationships between stakeholders in the ecosystem. Specific network questions were added in the updated Georgia online survey instrument to obtain information about relationships between stakeholders. It is critical to verify if this approach provides a map which provides a stronger picture of the data ecosystem and engagements within it.

The overall approach taken by 50x2030 to understand data use seems to have been successful at pro-

[2]The validation workshop for Georgia has been delayed due to logistical challenges and is currently scheduled for February 2022.

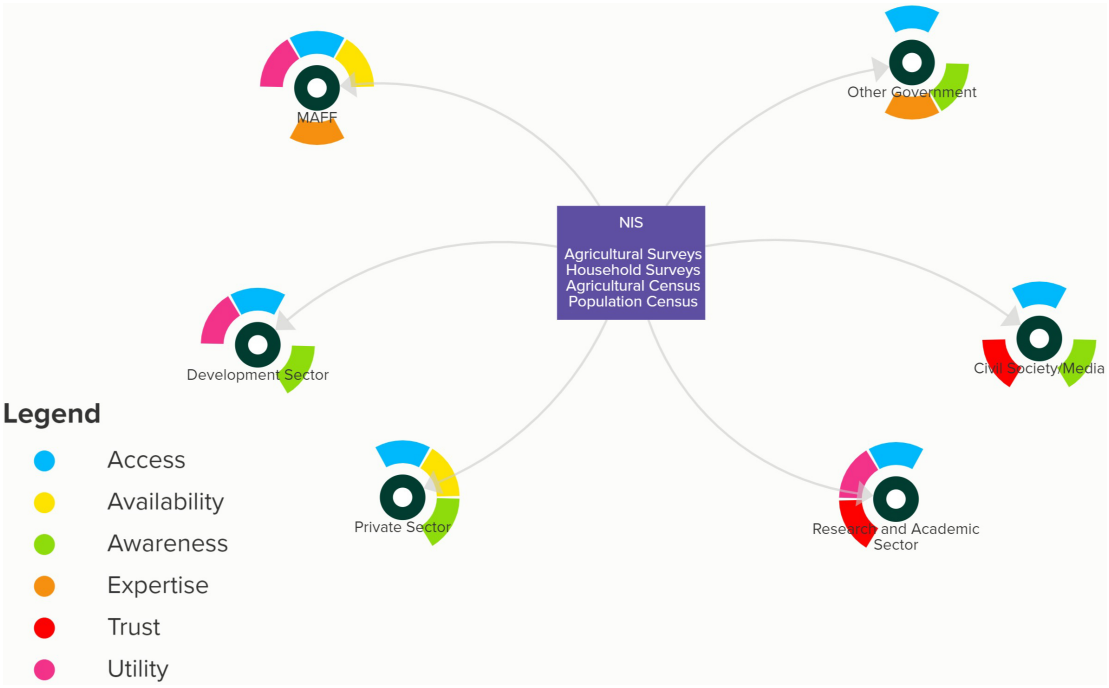

Fig. 2. Cambodia data ecosystem map.

viding insights into data use within the countries. The mixed-method approach together with data validation workshops with key stakeholders provided opportunities for triangulating and validating the findings as well as for offering nuance to data use dynamics in the ecosystem. For example, the mixed method approach enabled the Data Use team to determine in what ways certain factors were constraining the use of data in a particular country context and how those constraints could be addressed.

The qualitative analytic approach has worked well in applying a systematic approach to analyze the qualitative data from different country contexts in a way that enables direct comparison of findings while also allowing for flexibility to capture emergent themes specific to each country context. Specifically, the codebook includes both deductive and inductive codes so that all countries' qualitative sample is analyzed using the same deductive codes (i.e., those that are predetermined from the conceptual data use framework and the four research questions). Each country's qualitative sample is then reviewed, and the analysts determine the inductive codes (i.e., those that emerge from the data) to apply to that specific country's sample. The same qualitative approach has been applied in all three countries.

Both the online survey and KII guide ask specific information about the respondent to categorize responses by type of respondent, sector, and stakeholder group. Initially, the Cambodia KII guides did not include a section to ask information about the respondent, but during analysis, it was determined this information was necessary and the local data collection firm provided this information for each respondent. For the Uganda and Georgia guides, a section was added to gather information on the respondent, specifically on their role, position, organizational affiliation, sector, and stakeholder group. By analyzing both the qualitative and quantitative data by sector and stakeholder group, the Data Use team can observe patterns between sectors or stakeholder groups for a more nuanced understanding about how constraints or enabling factors to data use are experienced in the country context. For example, in Uganda, academic and government sectors did not experience similar constraints of access to needed data than did civil society and the private sector.

Additionally, the use of a semi-structured KII guide with sufficient probing questions allowed for more in-depth exploration of the different constraining and enabling factors of data use and helped to explain findings and patterns revealed from online survey data.

However, based on the implementation of the method in the first three countries, a few key lessons on implementing the methods emerged.

1. *Questions need to be carefully constructed.*

After the first round of data collection in Cambodia and Uganda, it was clear that both the qualitative and quantitative instruments needed to be revised. This was mainly to ensure that all four research questions would be sufficiently answered in both approaches. Initial data collection, particularly in Cambodia, found that the survey instrument elicited information primarily geared towards answering the second research question on constraints to data use, while the interview guide did not have sufficient prompts to answer two research questions focused on enabling collaboration between stakeholders and the different capacities needed to improve data use. The revised interview guides used in the Georgia assessment included specific topic areas to ensure questions were asked of respondents to provide direct insights to needed capacities, as well as ways to motivate or improve collaboration/interaction among stakeholders for data use.

In addition, the initial instruments did not include nuanced or detailed questions nor did it provide sufficient clarity on key terminology. For example, questions about access to data in the online questionnaire delivered in Cambodia did not distinguish between micro-level data and more disaggregated data. This was revised for the survey questionnaire delivered in Uganda and Georgia to include two distinct sets of questions that asked respondents about access to disaggregated data versus raw data (microdata). With respect to terminology, the term "microdata" was unfamiliar to most respondents in different country contexts, therefore the term "raw data (microdata)" was used to offer more clarity.

A revised approach was implemented in Georgia that addressed these issues and included a new set of questions to sufficiently answer all four research questions and elicit detailed responses. Specifically, a matrix question was added in the survey instrument to gather data on the relationships and interactions in the data ecosystem for both the Data Ecosystem mapping as well as the research question on interaction/collaboration around data use. With regards to the interview guide, many questions and prompts were added to gather detailed information on all four research questions as well. For comparison purposes, the initial interview guide used in Cambodia included five main sections and objectives as opposed to seven sections in the updated interview guide implemented in Georgia. The new approach can be seen in Table 3.

2. *A larger sample of responses is needed to ensure adequate representation from each of the five sectors of users.*
   The original sampling target in Cambodia was 40–60 responses for the online survey and the push for data collections in Cambodia stopped after obtaining 49 responses. However, analysis of the Cambodia data revealed the importance of analyzing results from each of the five sectors of users. Yet, the sample from each of the five sectors was insufficient to draw clear conclusions. This indicated a need for a larger overall sample for the online survey including a greater number of respondents from each of the five sectors. The new target number of responses for the online survey is 100 with broader representation from the five sectors.
3. *In depth training for qualitative work is needed.*
   The qualitative data collection experience in Cambodia highlighted the critical need for in depth training of interviewers to yield robust qualitative data for analysis. In Uganda and Georgia, a revised approach was implemented to include two consecutive trainings on qualitative data collection. First, a two-day, in-depth training on the qualitative data collection approach and the specific interview protocol was provided to the in-country teams by the 50x2030 Data Use qualitative analysis team. Second, after the in-country teams completed the test KIIs, the senior researcher specialist reviewed the test interview transcripts and provided a debriefing session to go over any issues, refine as needed, and ensure that the interviewers were fully comfortable with the interview approach. Furthermore, detailed instructions were added to the interview protocol to guide in-country interviewers. Additionally, specific instructions and guides were sent to in-country teams based on the interviewers' needs after the debriefing session.
4. *Data ecosystem mapping requires careful consideration*.
   While creating a map to illustrate connections in the data ecosystem appears valuable, it requires careful consideration and a clearly defined strategy. Although the objective of the data use assessment is to examine data use within a country's data ecosystem, which includes the entire community of actors, stakeholders, and entities

Table 3
Revised sections and objectives for Georgia interview guide

No.	Section name	Objective
A	Respondent Information	*To ensure the relevant background information on each respondent is recorded systematically for analysis.*
B	Data Availability, Awareness, and Access	*To elicit information on to what extent data that the stakeholders need is available, whether they have access to needed data, and whether they are aware of existing data sources in the country. (Note that awareness will be assessed by comparing objective information gleaned through the desk review and other interviews versus what the respondent believes to be true.)*
C	Current Use of Data	*To elicit information on if and how agricultural data are currently used by the respondent and/or in the country.*
D	Data Demand and Expertise	*To elicit information on the extent to which decisionmakers and data users want data to use to make their decisions and whether they know how to use it.*
E	Data Utility	*To elicit information on whether generated data and data products are relevant and useful to decisionmakers and how data could be made more useful to decisionmakers.*
F	Trust in Data	*To elicit information on whether data are trusted and believed to be accurate and if not, why not.*
G	Interaction	*To understand specifics about how and with whom respondents interact with around data use in the ecosystem and to better understand challenges and benefits of that interaction, and what needs to be done to improve collaboration.*

who engage with data, the data assets (data sets, data products, platforms, tools, technologies) with which they interact, and the rules, norms, and structures that govern those interactions (policies, cultures, organizational structures, etc.), the initial data collections, specifically in Cambodia and Uganda, did not gather sufficient data to examine the entire community of stakeholders, their interactions, and the rules and norms that govern those interactions.

The first data ecosystem map that was completed for Cambodia, therefore, included only the main stakeholders and the constraints they faced. After many deliberations, the mapping approach was expanded and revised to document and visualize the actors, their interactions, and the governing norms in the entire country data ecosystem. This would be referred to as a current state map or the current status of the data ecosystem based on preliminary findings. A subsequent map would be devised in a participatory approach whereby the in-country stakeholders would identify the rules, norms, and structures that govern those interactions (policies, cultures, organizational structures, etc.) in order to come up with a shared vision or ideal state of the country data ecosystem. Then, in the validation and planning workshops, in country stakeholders come up with concrete interventions and activities to reach the targeted principles or ideal state.

5. Overview of findings from Cambodia, Uganda, and Georgia

The summary of findings based on each of the four research questions are presented in Table 4 for Cambodia, Uganda, and Georgia.

In Cambodia, the survey and KIIs were conducted before the release of the results from the 2019 Cambodia Agricultural Survey. Thus, predictably, the main constraint reported by interview and survey respondents was access to data. Respondents also reported low levels of awareness and trust in the data as the next most constraining factors to data use. Other factors, such as the utility of data, data availability, and the expertise of data users (or decisionmakers) were also reported as constraints, albeit with less emphasis compared to access, awareness, and trust. However, due to the interconnected nature of the different factors of the data use framework, it is expected that with more data becoming accessible in Cambodia, the need for improving expertise and the utility of data would also increase.

In Uganda, three main factors contributed to achieving more nuanced findings compared to Cambodia: 1) the refinement of the survey and interview tools, 2) offering in-depth training for enumerators, and 3) bigger sample size for the online questionnaire. In terms of findings, demand for data was ranked equally high by both Ugandan and Cambodian respondents in their respective countries. Unlike the respondents to the Cambodia survey, however, the Ugandan respondents reported the expertise of data users as a constraint even though Uganda has had a longer history of data availability and use compared to Cambodia. This difference in findings between the two countries further highlights the expectation that once more data are available in a particular country, the need to enhance expertise would also become more apparent to stakeholders.

As in Cambodia and Uganda, awareness and access were the top two constraints to data use in Georgia.

Table 4
Key results from data use assessment

No.	Research question	Cambodia	Uganda	Georgia
1	*How are agricultural and rural data currently being used and for what purposes? What are the potential development benefits to the agricultural and rural sectors of the country if the survey data are appropriately used for decisionmaking in policy, programs, and investments?*	– Throughout the assessment, respondents described common ways in which stakeholders use data in Cambodia: to inform strategy, program, and project design; statistical research and analysis; national and local policy development; budget formulation; and investment decisions (in order from most to least common uses). Only 2% of respondents said that they did not use agricultural and rural data produced in Cambodia. – The main benefits identified by respondents were very similar to the uses described above. Using data to inform those actions generate improved products and results.	– Most demand for agricultural and rural data in Uganda among respondents is for "Strategy, program, or project design" (68%), "Project or policy impact assessment" (67%) and "Budget formulation" (66%). These uses are followed by "Statistical Research or Analysis", "Market Research or Analysis", "Investment decisions", "National policy development", and "Local policy development" (in order from most to least commonly mentioned uses). 8% of the total survey respondents have declared that they do not use agricultural and rural data produced in Uganda. – Benefits of data use include prioritizing and tailoring interventions by using evidence-based decisions, which helps to mitigate wasting resources, such as funding and time, on unproven interventions.	– Most demand for agricultural and rural data in Georgia among respondents is for "Strategy, program, or project design" (58%), "Statistical research or analysis" (48%), "Project or policy impact assessment" (44%), and "Market research or analysis" (42%). These uses are followed by "Investment decisions" (23%), "National policy development" (21%), and "Local policy development" (20%). The least amount of demand is for "Budget formulation" which was selected only by 9% of the respondents. Finally, 6% of the total survey respondents have declared that they don't use agricultural and rural data produced in Georgia. Benefits of data include evidence-based decision-making, setting priorities, raising and allocating resources effectively, and increasing agricultural production.
2	*What are the key factors or sub-factors from the Data Use Framework) that constrain the use of data or will constrain the use of data if not addressed?*	– The identified top constraints to data use were (in order from most to least constraining): 1. Access to data. 2. Awareness about the availability of data. 3. Trust in the data. 4. Utility of data–esp. the adequacy of the types of data (e.g. useful variables, disaggregation, periodicity) and data formats. 5. Availability of data. 6. Expertise of data users (knowing what data they need and how to use it).	– The identified top constraints to data use were (in order from most to least constraining): 1. Awareness about the availability of data. 2. Access to needed data (esp. microdata). 3. Utility of data–esp. the adequacy of the types of data (most prominent among them were improved levels of disaggregation, timeliness, useful variables/indicators, and improved periodicity) as well as better data formats. 4. Expertise of data users (knowing what data they need and how to use it). 5. Trust in the data (esp. due to the perception that data are altered by political interference). 6. Availability of data.	– The identified top constraints to data use were (in order from most to least constraining): 1. Awareness about the availability of data. 2. Access to needed data (esp microdata). 3. Utility of data–esp. the adequacy of the types of data (e.g. timeliness and disaggregation) and data formats. 4. Availability of data. 5. Trust in data. 6. Expertise of data users (knowing what data they need and how to use it).

Table 4, continued

No.	Research question	Cambodia	Uganda	Georgia
3	*Which human and institutional capacities most need to be enhanced to improve data use?*	– The most salient capacity needs relate to various points in the data cycle and include capacities: 1. to produce needed data types, 2. to produce high-quality data, 3. to analyze and interpret data, 4. to maintain current systems for data access and information sharing.	– The most salient human capacity needs relate to all points in the data cycle 1. to improve data collection, management, analysis, validation, and synthesis, research design and communication, 2. to strengthen awareness and sensitization for the value and need to use data in decisionmaking. Institutional capacities needed include – producing up-to-date data – operating a centralized, user-friendly platform to access and share data across stakeholders – standardized and validated data collection protocols – dedicated budgets to facilitate capacity improvements.	– The most salient human capacity needs include capacity-building and training in basic data literacy, computer skills, data interpretation, analysis, and visualization, modeling, forecasting, and applying new technologies to improve data quality. – Another salient need was to increase public awareness to promote data-based decisionmaking and the purpose of data collection and data importance. Institutional capacities needed include – Sharing information on available data and methodologies for data collection would help increase reliability and therefore data use. – The need for more human and financial resources; adequate pay and incentives for data producers for quality data collection. – Digitalization and use of technology in the production and analysis process. – Capacities needed include higher level of research and analysis in all sectors, including the public and private sectors and civil society.
4	*In the particular cultural context, what incentives would motivate stakeholders to engage or work together to strengthen the ecosystem and their exchanges within it?*	– *Main incentive:* 1. Demand for reliable and accurate data. – *Main disincentives:* existing constraints to data use such as, 1. lengthy bureaucratic processes 2. perceived unwillingness to share data 3. low levels of data utility and/or accuracy 4. lack of transparency about mechanisms to access official data	– *Main incentives:* 1. improving communication between stakeholders about data use activities to increase awareness 2. improving capacity of data collectors and analysts to improve the use of and trust in reliable data 3. strengthening the relationship between data producing organizations and the media to use and report on agricultural data 4. establishing formal channels or agreements for data use between stakeholders (such as MOUs or technical committees) at all levels of government. – *Main disincentives:* 1. competition between data producing institutions 2. issues around proprietary and data ownership for data producing organizations.	– *Main incentive:* 1. Demands for government initiated and managed coordination mechanism to better communicate, share, and make accessibly available data for all interested data users across stakeholder groups. 2. Strengthening the capacity across all stakeholder levels to produce, analyze, interpret, and use agricultural data and statistics. 3. Respondents are incentivized to collaborate to improve the quality of their data to make informed decisions. There is a high value placed on informed decisionmaking with quality and accurate data. – *Main disincentives*: 1. Lack of awareness of or transparency in how to access available data.

Table 4, continued

No.	Research question	Cambodia	Uganda	Georgia
	Notes/Observations	– The survey and KIIs were conducted before the release of the results from the 2019 Cambodia Agricultural Survey in March 2021.	– The Uganda survey was slightly modified by adding questions that would help gather more nuanced answers (e.g. with respect to organizations where stakeholders would get their data from, access to aggregate vs. microdata, utility needs for different data sources, etc.). – The bigger sample for Uganda allowed for a more nuanced analysis per sector (especially private sector). – Ugandan survey respondents ranked expertise among one of the key constraining factors to data use. Whereas for Cambodian respondents, expertise was ranked at a lower level of priority despite Cambodia's more limited access to data compared to Uganda.	– Georgia has a short history of having statistically reliable data including microdata, which is very different from other countries such as Uganda where the World Bank LSMS-ISA has been taking place for over 10 years. The lack of awareness of data and limited access to data contributes to the situation where Georgia has varying levels of expertise and experience in preforming detailed research and analysis.

Georgia, like Cambodia, has a short history of conducting statistically reliable agricultural surveys, with a short time period of having microdata available for analysis. Uganda has been conducting surveys for a longer period of time. Georgia has very small number of government staff for analyzing data prepared by data producers and preparing analytical reports for decision makers, and also a small number of analysts outside of the government with experience analyzing data and preparing research papers. Expertise in data production, analysis, and interpretation in Georgia varies, as there exist many professionals with high levels of expertise but also a wide range of staff with general knowledge and understanding of methods and statistics, but lacking in-depth analytic skills. The need for more data by disaggregated crops, agricultural units, and geographic areas is important for Georgia, as for most countries.

Overall, the findings about data use in the first three countries were:

1. *Users note the benefits of agricultural and rural data.*
   In all countries studied, data users report that they use agricultural data for broader strategies, and for program and project design. Data use for local policy development, and to some extent national policy development, has been ranked as a less common use of data a result consistent with the literature which notes that it is often insufficiently granular for this purpose. Depending on the country context, other uses of data (such as statistical research or analysis, and budget formulation) have received different rankings by respondents. Respondents have cited that main benefits of using data are to make informed and evidence-based decisions to prioritize and tailor interventions, which helps to mitigate wasting resources, such as funding and time, on unproven interventions.
2. *Yet, there are a number of key constraining factors to data use.*
   Across the three countries, demand has been ranked the strongest enabling factor to data use, indicating high demand for agricultural data. Access to and awareness of data were determined as the strongest constraining factors to data use in Cambodia, Uganda, and Georgia. Wherever access is reported as a constraint to data use, awareness has either preceded or followed as another constraint – indicating the interconnected nature of these two factors.
3. *Human and institutional capacities that need to be enhanced to improve data use.*
   The most salient human capacity needs relate to improving skills and capacities at all points in the data cycle. Specifically, these include increased skills and capacities to improve data collection, management, analysis, validation, and synthesis, research design and communication, and strengthened awareness and sensitization for the value and need to use data in decision-making. Across the three countries, needed institutional capacities include sufficient budgets and funding dedicated to improving data use needs, standardized approaches to collect, analyze, and disseminate needed data, and sufficient systems to share and disseminate data for public access.
   A key capacity limitation for data use in particular is the lack of knowledge brokers/translators, particularly in countries like Cambodia and Georgia where agricultural data collection is just starting. This means that little relevant analysis is done outside of the government since there is limited capacity to perform the right type of analysis. Since inside the government, there is also very little capacity to conduct analysis, this means that knowledge brokers role will need to be developed. Even though Uganda is better off than Cambodia and Georgia in that it has history of having data, government staff and outside researchers and analysts need to expand their work considerably.
   Specific capacities that are linked to data use include the need for the right type of data in terms of disaggregated data and microdata as well as frequently updated data. It was noted that agricultural activities are seasonal, yet the available data are often outdated and cannot be used to make timely decisions and implement projects. Another major institutional capacity needed to improve data use includes improving current data sharing and knowledge-sharing systems and procedures to ensure free and open access between stakeholders as well as the appropriate communication tools to disseminate data widely. User-friendly formats, language, communication methods are linked to increasing data use across all levels of stakeholders (i.e. farmers and government leaders).
4. *Incentives need to be put in place to motivate stakeholders to work together.*
   The main incentives that motivate stakeholders to work together to strengthen the data ecosystem essentially would require removing or reducing existing constraints to data use. These incentives include improving communication and coordina-

tion between stakeholders about data use activities to increase awareness and building the capacity of data collectors to improve trust in the reliability of the data. While trust was not ranked as a critical constraint in any of the three countries, respondents did share concerns about trust in the reliability and validity of government-produced data. In Uganda, these concerns focused on concerns of government staff's capacity to collect, manage, and produce reliable and up to date figures that could be used to effectively inform decisionmaking. In Cambodia, these concerns centered more on a nebulous understanding that government data can often be skewed to show more favorable results than what may reflect reality, and that government staff may also lack capacity to accurately collect, manage, and produce reliable data. Issues of trust are typically stronger expressed in the qualitative data than the quantitative. Improving trust in data by sharing data collection method notes and increasing transparency around how the data were collected, managed, analyzed and produced would facilitate an increase in trust among stakeholders to use data. Likewise, existing constraints to data use (if left unaddressed) would disincentivize stakeholders to work together. Other issues, such as competition between data producing institutions and issues around proprietary and data ownership for data producing organizations, may disincentivize stakeholders to work together or interact on data use.

6. Implications and way forward

To ensure that agricultural data being generated by NSOs is adequately being used for decisionmaking within countries requires understanding data use within those countries. This paper offers a methodological approach for understanding data use which includes a review of the literature along with quantitative and qualitative data collection. In doing so, the approach seeks to build on the recent literature on data use and move beyond a supply and demand framework for data use towards an approach that focuses more broadly on the data ecosystem, how it operates and where barriers and constraints may limit effective data use.

While a number of key lessons have emerged to improve the methodological approach and will continue to do so as the methodology is developed, overall the approach provides new insights into data use, which may be used to facilitate the strengthening of the data ecosystem. A key next step is to move from understanding data ecosystems and their constraints and to seek concrete actions which will strengthen the system and facilitate decisionmaking. The collected evidence suggests that access and awareness are key constraints as are human and institutional capacities. Incentives also need to be adjusted. Moving forward then requires simultaneously considering how to improve on the methodological approach and garner greater insights into data use, while developing a concrete action plan for strengthening the system.

Acknowledgments

The authors are grateful to members of the 50x2030 Program Management Team for helpful feedback on the 50x2030 Data Use Guide, to CABI for guidance and input on the Cambodia data ecosystem mapping, to Christian Chihababo, Elizabeth Nderitu, Nelson Kolliesuah, Giulia Salvi, Cheryl Wagah, and Wilbon Yegon for research support, and to Christine Kobusingye for support in collecting and analyzing data for the project and Stephen Katz for communications and management support.

References

[1] World Bank. *World Development Report 2021: Data for Better Lives*. Washington, DC: World Bank. 2021. doi: 10.1596/978-1-4648-1600-0.

[2] Jacob A. Mind the gap: Analyzing the impact of data gap in Millennium Development Goals' (MDGs) indicators on the progress toward MDGs. *World Development*. 2017; 93: 260–278.

[3] Masaki T, Custer S, Eskenazi A, Stern A, Latourell R. *Decoding data use: How do leaders source data and use it to accelerate development*? Williamsburg: AidData at the College of William & Mary. 2017.

[4] FAO, IFAD, UNICEF, WFP, WHO. *The State of Food Security and Nutrition in the World 2020. Transforming food systems for affordable healthy diets*. Rome, FAO. 2021.

[5] IPCC. *Climate Change 2021: The Physical Science Basis*. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change [Masson-Delmotte V, Zhai P, Pirani A, Connors SL, Péan C, Berger S, Caud N, Chen Y, Goldfarb L, Gomis MI, Huang M, Leitzell K, Lonnoy E, Matthews JBR, Maycock TK, Waterfield T, Yelekçi O, Yu R, and Zhou B (eds.)]. Cambridge University Press. In Press. 2021.

[6] Porciello J, Tufan H, Njuki J, Winters P, Mabaya E, Coffman R. Comment "Averting hunger in sub-Saharan Africa requires data and synthesis". *Nature*. 2020; 584: 37–40. doi: 10.1038/d41586-020-02281-w.

[7] Sethi T, Prakash M. *Counting on Statistics: How can national statistical offices and donors increase use*? Williamsburg, VA: AidData at William & Mary. 2018.

[8] Development Initiatives. Data use: an overview of conceptual and practical approaches. Bristol, UK: Development Initiatives. 2018. http://devinit.org/wp-content/uploads/2018/07/data-use-an-overview-of-conceptual-and-practical-approaches.pdf [Accessed 17 September 2021].

[9] Dufief E, Dorward N, Simons R, Welford K. With Publication Comes Responsibility: Using open data for accountability in Benin and Tanzania – A discussion paper. 2017. Available at: http://www.publishwhatyoufund.org/wpcontent/uploads/2017/09/With-Publication-Brings-Responsibility-A-discussion-paper.pdf [Accessed 17 September 2021].

[10] Poirrier C. Scoping Study: Evidence Translators' Role in Evidence-Informed Policymaking. Washington, DC: Results for Development. 2018. https://r4d.org/resources/scoping-study-evidence-translators-role-in-evidence-informed-policymaking/ [Accessed 17 September 2021].

[11] Oliveira MIS, Lima GDFB, Lóscio BF. Investigations into Data Ecosystems: A systematic mapping study. *Knowledge and Information Systems*. 2019; 61(2): 589–630.

[12] Paris 21. Guidelines for Developing Statistical Capacity A Roadmap for Capacity Development 4.0 Paris: OECD. 2020.

[13] 50x2030 Data Initiative. *The 50x2030 Initiative: Bringing Together Committed Partners to Fill the Agricultural Data Gap*. Rome: 50x2030 Data Initiative. 2021a. https://www.50x2030.org/sites/default/files/files/50x2030-2020-Brochure.pdf [Accessed 22 September 2021].

[14] Villarino et al., forthcoming. From agricultural statistics to zero hunger: How the 50x2030 Initiative is closing data gaps for SDG2 and beyond. This Special Section.

[15] Oliver K, Lorenc T, Innvær S. New directions in evidence-based policy research: A critical analysis of the literature. *Health Research Policy and Systems*. 2014; 12(1): 1–11.

[16] Swanson E, Pittman A. *Measuring Data Use: An Analysis of Data Portal Web Traffic*. Open Data Watch. 2018.

[17] BenYishay A, Parks B. Can Providing Local Data on Aid and Population Needs Improve Development Decision-Making? A Review of Recent Experimental Evidence. AidData at William & Mary. 2019.

[18] Custer S, Sethi T, eds. Avoiding Data Graveyards: Insights from Data Producers and Users in Three Countries. Williamsburg, VA: AidData at the College of William & Mary. 2017.

[19] Open Data Institute. *Mapping Data Ecosystems*. ODI. 2020.

[20] 50x2030 Data Initiative. *A Guide to Promoting Data use under the 50x2030 Initiative*. Rome: 50x2030 Data Initiative. 2021b. https://www.50x2030.org/sites/default/files/resources/documents/2021-01/Methodological%20Guidance%20on%2050x2030%20Data%20Use%20Activities_15January2021_CLEAN.pdf [Accessed 22 September 2021].

[21] Downing T, Goentzel J, Wetmore F. Guide to Shock Maps in Kumu. USAID: Uganda Feed the Future Market System Monitoring Activity. 2020. https://hdl.handle.net/1721.1/127277.