

GETTING SMART WITH SURVEYS: USING COMPUTERS, TABLETS OR SMART PHONES FOR HOUSEHOLD SURVEYS



Tablet-/smart phones-based surveys are better than paper-based surveys as they make it easy to monitor data collection on a real time basis. Also, it saves time in data entry as the collected data is directly retrieved in the form of Excel sheets. In this blog Aditya KS, Subash SP and Bhuvana N discuss

survey data collection using computers, tablets and smart phones. A detailed step-by-step explanation on using Kobo Tool Box – an open access resource for collecting survey data using tablets/mobile phones – is also provided here.

CONTEXT

Surveys are the bread and butter for any type of researcher in the social sciences; be they students (Masters and PhD) or researchers, or organizations, such as the National Sample Survey Office (NSSO), which carry out large surveys for tracking different socio-economic characteristics. Research in the field of social science is mostly observational (except when it uses experimental approaches) and a researcher has no control over the events. The data is collected as they are observed and analyzed to draw conclusions, and the researcher has to depend on either secondary or primary data to draw inferences. As most of the secondary data sources are either macro level aggregates or not suitable to address the specific research question under consideration, primary surveys are inevitable in most research projects. Primary surveys offer flexibility to the researcher to understand different socio-economic and cultural perspectives related to the research problem.



From the perspective of a social scientist in Agriculture (Agricultural Extension or Agricultural Economics), primary surveys are a strategic tool for assessing the demands, aspirations, priorities and challenges of farmers and other stakeholders in agriculture. They are also important for understanding the target population while planning extension intervention or designing policies. Further, surveys are equally important in assessing the impact of interventions (technologies or policies).

SURVEYS

The history of survey-based data collection goes back to Seebohm Rowntree's first poverty survey in 1899 at York, England (Careletto and Gourlay 2019). Over the years household surveys evolved rapidly. Simultaneously, the evolution of technological aids in surveys – to make it easier and reduce errors – are also progressing. In the next section, we talk a bit about the problems of paper-based surveys before discussing how to do a survey without paper (super cool and ecofriendly too!!).



A paper-based survey in the mid-hills of Nepal

Paper-based survey

Traditionally, surveys use paper-based questionnaires. A set of questions (Structured or semi-structured) are developed based on the research problem and printed on paper (a lot of sheets). Enumerators (surveyors) go to the field and record responses on the paper sheets, which later gets digitized through a data entry process. This is a cumbersome process as it takes quite a lot of time and has several problems and challenges associated with it. For instance, data entry can take a long time and many human errors can creep in. Furthermore, real time monitoring is difficult and often the errors in surveys are detected only after the data is entered and tabulated. By that time, it is too late to do anything about it. Using computer/tablet or mobile-assisted surveys can help to overcome these. We will discuss the advantages of computer-based surveys vis à vis paper-based ones in the later part of this blog.

Computer-Aided Personal Interviews (CAPI)

Technological advancements have changed the way surveys are carried out now. Computer-Aided Personal Interviews (CAPI) is one such technological advancement in data collection (satellite and big data are the other advancements). CAPI is a face-to-face data collecting method in which the enumerator uses a small computer, or a tablet, or a smart phone to collect the data. Availability of

hardware (like inexpensive computers, tablets and smart phones) and software (both paid and Open Access) has made it easier to carry out CAPI surveys. This is helpful for two very important reasons:

1. No need for data entry process to digitize the information as in paper-based surveys;
2. Real time data monitoring to minimize errors.

There are many software tools available that are customized to carry out face-to-face surveys. There is a wide variety of software tools to choose from, and they can be used by people with no programming knowledge (like us!). They are increasingly being used by researchers in organizations, such as the World Bank, and CGIAR institutes.

The latest Period Labour Force Survey (PLFS) carried out by NSSO used CAPI for data collection. ICAR-National Academy of Agricultural Research Management (NAARM) and PJTSAU, Hyderabad, undertook a consumption survey of Telangana State using CAPI in 2017 (Kumar et al. 2017). There are several advantages and disadvantages with CAPI which we will discuss here. (The advantages and disadvantages given here are based on DIME 2020 and our own experience.)

Advantages

- The collected data gets digitized immediately and sent to the server. In places with good networks we could carry out high frequency checks (cross checking the responses on a daily basis). Real time monitoring is really easy unlike paper surveys, a person sitting in the office can cross-tabulate all the responses and test for data consistency.
- Enumerators can be monitored in case of larger surveys. The start time, end time, GPS location can be automatically recorded and supervisors/researchers get to see it. Even the interviews can be randomly recorded (without the knowledge of enumerators) and later verified by supervisors/researchers.
- It's easy in case of questions with skip logic – goes directly to a different set of questions based on the response to a previous question (e.g., If yes then...? And if No then...?).
- Do calculations and conversions easily by using inbuilt calculators (for instance *Bigha* to Hectare, and so on).
- Collect data, such as images, farm plot sizes (GIS-based plots), and qualitative data (record statements).
- Avoid errors in data collection (missing questions) and data entry (most common errors).
- Data validation can be incorporated, for example, farming experience cannot be more than the age of the farmer. Such conditions can be imposed while preparing the survey itself to avoid the errors.

In a nutshell, CAPI offers many advantages in terms of features to reduce data collection errors. Irrespective of these advantages, there are certain disadvantages, rather challenges, in using CAPI.

Disadvantages

- Respondents may not be comfortable with CAPI-based survey. They get suspicious and often distracted seeing the gadgets. (At some rural villages in Byndoor district of Karnataka, our enumerators were forcibly locked up in a room and we had to get police help to get them out. The community was not very open to outsiders and using tablets made them even more suspicious. Gram panchayat election was also close at hand!- Don't worry such situations won't happen all the time.)



Satellite image-based plot data

- Need trained enumerators (longer training periods) with a bit of knowledge on handling gadgets.
- Difficult to carry computers and tablets in crime-ridden areas (Risk of theft is there - in one of the surveys it happened to us).
- They also require electricity (back up batteries can help) and good network connectivity (although data could be uploaded to a server once a week or later in a place with good connectivity).
- Developing the questionnaire in CAPI is time-consuming compared to paper-based questionnaire (though the overall time saved is unquestionable).
- Language restrictions (unlike paper-based surveys, CAPI has limited options for developing the questionnaire in local languages).

Software and Hardware essentials to carry out CAPI survey

We have different software available for CAPI surveys. A list of commonly used ones are given in Table 1. Each of them have their own set of advantages and disadvantages. A detailed discussion on each of them is beyond the scope of this blog. Keeping in mind that many of our readers would like to know about the open source options to carry out CAPI surveys, we have provided a detailed step-by-step guide to use Kobo Tool Box, a free and open access software to carry out CAPI surveys. However, if the researcher has money to go for paid software, he can always buy one.

Before starting our discussion on Kobo, let us try to understand a few points that must be kept in mind while choosing the software and hardware for CAPI. Some of the important questions to be asked while selecting software are:

- What kind of data is needed (text, pictures, audio)?
- How are they managed (does it require its own server?)?
- What is the output file format (most of them have multiple options)?
- Does it have language support (native language questionnaire)?

Similarly, while choosing the hardware (tablet/computer/smartphone), first, the hardware should be compatible with the software requirements; and it also calls for a good quality camera (if pictures are to be recorded). Further the size of the screen depends on the length of the questionnaire (for shorter questionnaires smart phones are fine, but for larger questionnaires tablets with 7-inch screens are preferable). The hardware should also have enough internal memory (8 GB preferably) and external storage (SD cards), the gadget should also be GPS enabled with better accuracy (10-15 meters), and good battery life.

It is a good practice to purchase tablets at the institute level. However, there is also an option to rent them. If the organization is involved in frequent surveys, it can also purchase cloud storage to store all the survey data at one place. However, the point is that, under resource constraints, android mobile phones are enough to carry out a CAPI survey. The table below provides a snapshot of a few software options for CAPI surveys.

Table 1. Software for CAPI

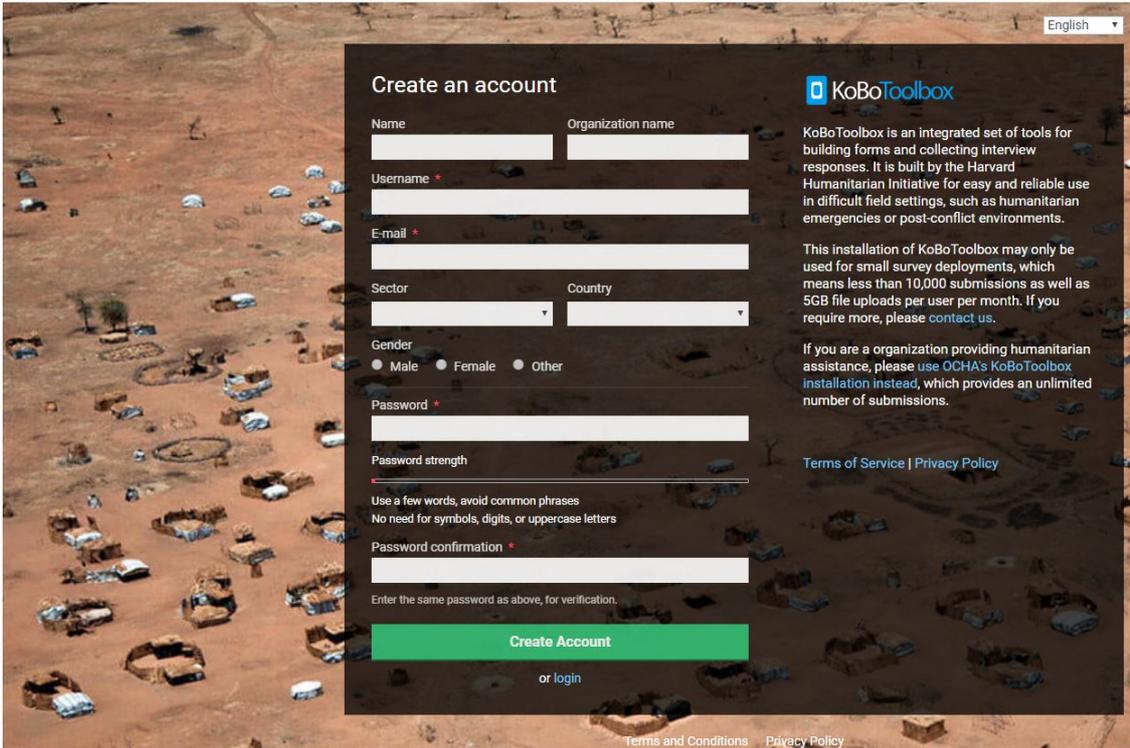
No	Software	Developer	Access option	Link
1.	Blaise	Statistics, Netherlands	Restricted	https://blaise.com/
2.	SurveyCTO	Dobility Inc.	Paid	https://www.surveyccto.com/
3.	CSPPro	United States Census Bureau	Open Access	https://www.census.gov/data/software/cspro.html
4.	Dooblo	Dooblo Ltd., Israel	Paid	https://www.dooblo.net/
5.	SurveyBe	Economic Development Initiatives Limited, UK	Paid	https://surveybe.com/
6.	SurveySolutions	The World Bank	Open Access	https://mysurvey.solutions/
7.	Kobo Tool Box	Harvard Humanitarian Initiative	Open access (Researchers)	https://www.kobotoolbox.org/
8.	Open Data Kit (ODK)	University of Washington's Department of Computer Science and Engineering	Open access	https://docs.opendatakit.org/odk-x/survey-intro/

As stressed earlier, even though many paid software options are available, in this blog we will focus on a very commonly used one – Kobo Tool Box – a free and open access software for CAPI survey.

Kobo Tool Box

Kobo Tool Box is an open access toolset for collecting survey data using mobiles or tablets. In this section we will elaborate on how to use Kobo Tool Box with a detailed step-by-step guide. Relevant screenshots are also provided.

- **Step 1:** Register as a researcher at Kobo on the following link <https://www.kobotoolbox.org/>. This offers 10,000 submissions per month with 5 GB of storage space, which is sufficient for most of the surveys done in academia. After registering, please note the username carefully, you will need it later.



kf.kobotoolbox.org/accounts/register/#

English

Create an account

Name Organization name

Username *

E-mail *

Sector Country

Gender
 Male Female Other

Password *

Password strength
Use a few words, avoid common phrases
No need for symbols, digits, or uppercase letters

Password confirmation *

Enter the same password as above, for verification.

[Create Account](#)

[or login](#)

KoBoToolbox

KoBoToolbox is an integrated set of tools for building forms and collecting interview responses. It is built by the Harvard Humanitarian Initiative for easy and reliable use in difficult field settings, such as humanitarian emergencies or post-conflict environments.

This installation of KoBoToolbox may only be used for small survey deployments, which means less than 10,000 submissions as well as 5GB file uploads per user per month. If you require more, please [contact us](#).

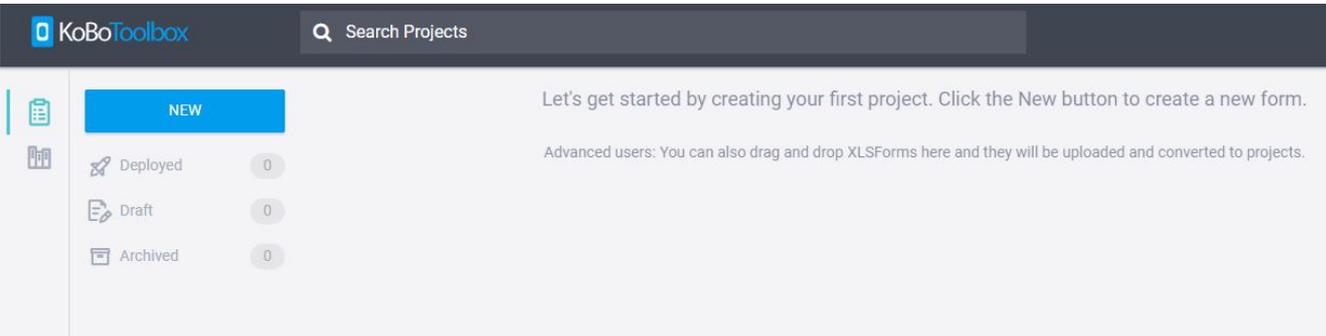
If you are an organization providing humanitarian assistance, please use OCHA's [KoBoToolbox installation](#) instead, which provides an unlimited number of submissions.

[Terms of Service](#) | [Privacy Policy](#)

[Terms and Conditions](#) | [Privacy Policy](#)

Step 1

- **Step 2:** Login to your account and click on New to start preparing the survey schedule. Give a project name, select the suitable discipline and select the country. Click 'Create Project' to proceed to the next step.



KoBoToolbox Search Projects

NEW

Deployed 0

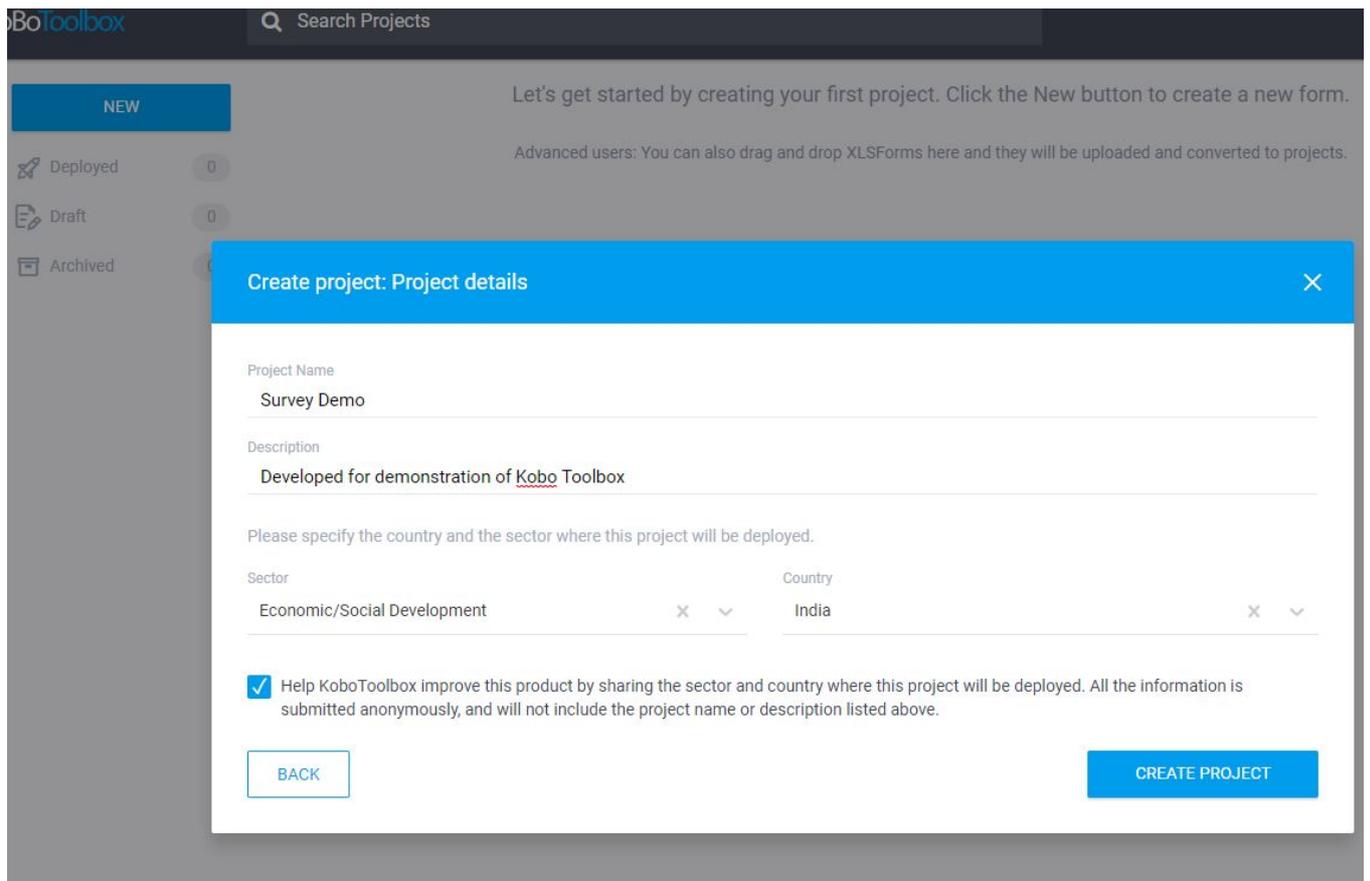
Draft 0

Archived 0

Let's get started by creating your first project. Click the New button to create a new form.

Advanced users: You can also drag and drop XLSForms here and they will be uploaded and converted to projects.

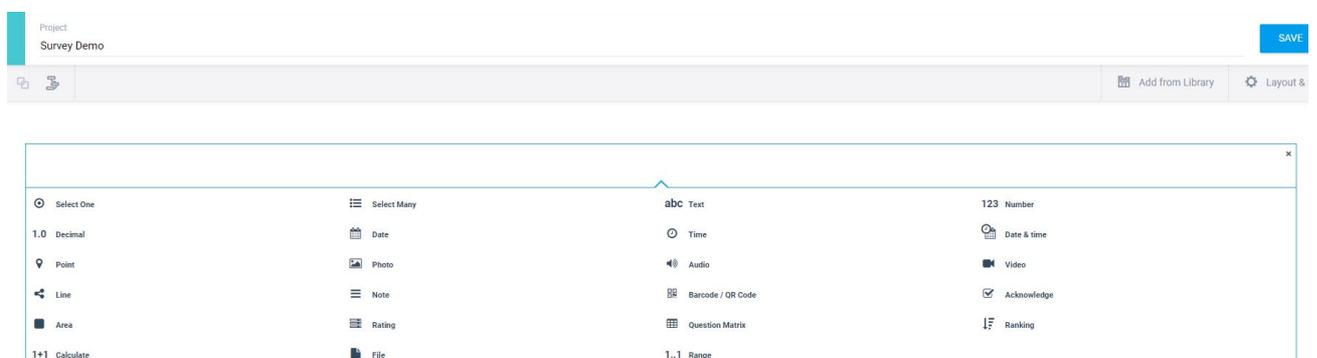
Step 2



- **Step 3:** Click on +Add Question and select the question type. There are different types of questions available - text, numerical, select one, select many, decimal, rating, ranking, grid, date and time, etc. According to the expected type of answer, select the right question type.



Step 3



- For example, family size will be expressed in terms of whole number, hence the question type has to be numeric. In the question 'Major Occupation', the right question type is select one,

where either 'Agriculture' or 'Non Agriculture' has to be selected. In the 'Secondary Occupation' question, farmer can have more than one secondary occupation, hence, select many is the right choice.

The screenshot shows a survey interface with two questions. The first question, 'Major occupation', has two radio button options: 'Agriculture' and 'Non Agriculture'. The second question, 'What are the secondary occupation', has four checkboxes: 'Employed elsewhere', 'Agricultural Wages', 'Processing and Value addition', and 'Others'. Both questions include a 'Question hint' and a '+ Click to add another response...' link.

- Skip logic is another attractive feature. Based on the response to the previous question, you can impose a condition to display or skip a particular question. For instance, the question, 'mention the other occupation' is displayed only if the respondent has selected the 'other occupation' in the previous question. Or else that question will be skipped. This can save a lot of time. For instance, consider a survey involving two crops. If the farmer is not growing wheat, all the questions pertaining to wheat will be skipped if this logic function is appropriately used.

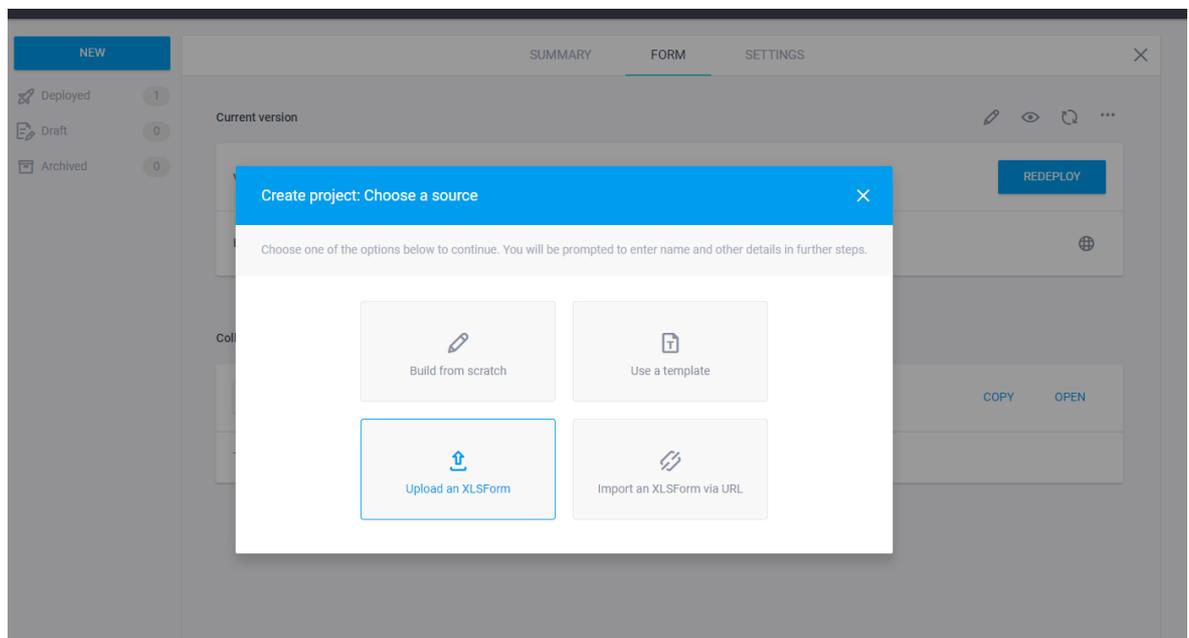
The screenshot shows the 'Skip Logic' settings for a question. The settings are displayed in a table-like structure with a 'Settings' header. The 'Skip Logic' row is highlighted, showing a green bar for '+ Add a condition' and a blue bar for 'Manually enter your skip logic in XLSForm code'.

- Another useful function is Validation criteria. You can restrict the response values within a limit to avoid data entry errors. For example, the question is 'What is the farm gate price of rice in Rs/kg?' The answer has to be less than 50 even by conservative estimates. But, the enumerator could get confused and enter 1200 (considering it as Rs/quintal). To avoid such errors, we can use validation criteria to limit the response to less than 50. If the enumerator enters any value

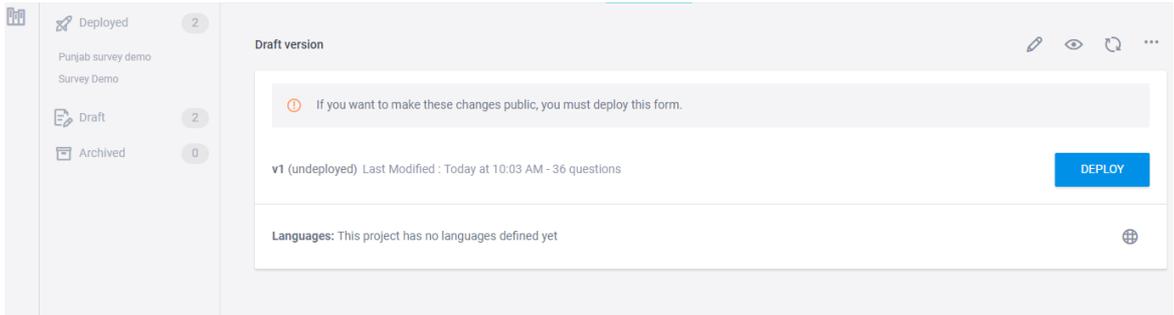
higher than that an error message will pop up. We can even customize the error message to remind him that the price is in Rs/kg, not in Rs/quintal.

The screenshot displays the KoboToolbox interface for editing a questionnaire. The top section shows a question titled "What are the secondary occupation" with a "Question hint" and a list of response options: "Employed elsewhere", "Agricultural Wages", "Processing and Value addition", and "Others". Below this, there is a section for "Validation Criteria" with a "Settings" sidebar on the left. The sidebar includes "Question Options", "Skip Logic", and "Validation Criteria". The "Validation Criteria" section is active, showing a green bar for "+ Add a condition" and a blue bar for "Manually enter your validation logic in XLSForm code". An "Error Message:" field is visible at the bottom right of the validation criteria section.

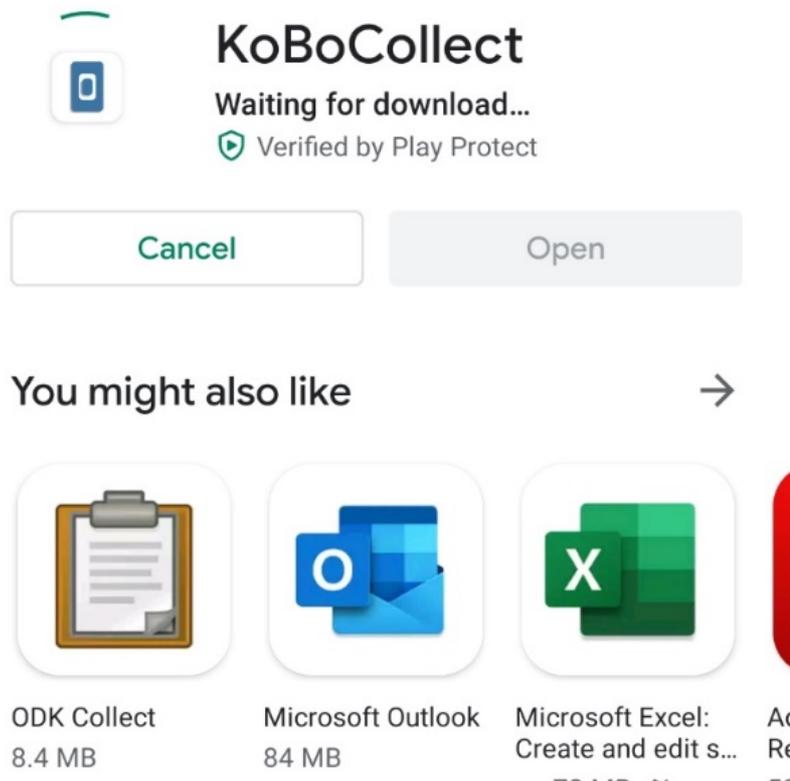
- Making the questionnaire directly in the Kobo Tool box can be time consuming and repetitive. The alternative is to prepare the questionnaire in 'Open Data Kit (ODK)' format and upload it directly to Kobo. Understanding the ODK format at the beginning can be a bit tricky, but it will save time in the long run. However, explaining the ODK format is beyond the scope of this blog. One sample ODK file is anyhow provided in the link below, which can be directly uploaded to Kobo (link).



- Once the questionnaire is ready, first preview it. If satisfied, then deploy the questionnaire.



- **Step 4:** Next step is to download the 'KoBoCollect' in all the devices which will be used for data collection (preferably Android, in IOS devices Web Form can work).

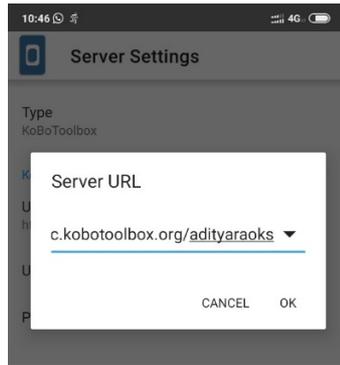


Step 4

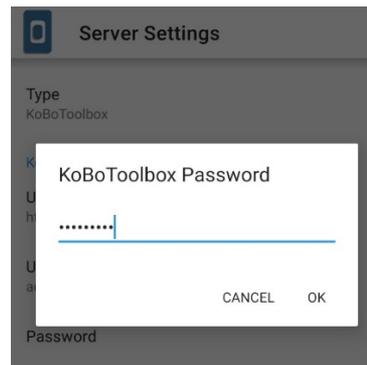
- **Step 5:** Once installed, go to the general setting - server. Modify the server URL as <https://kc.kobotoolbox.org/username> - here in the place of username, you have to input the username that you had used to create the survey schedule. For instance, in the demo survey, the username used is 'adityaraoks'. So the URL is edited to include adityaraoks at the end.



Step 5a

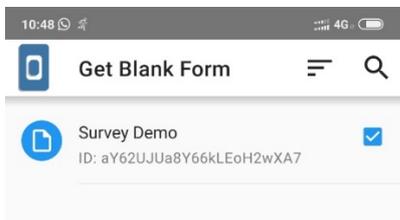


Step 5b

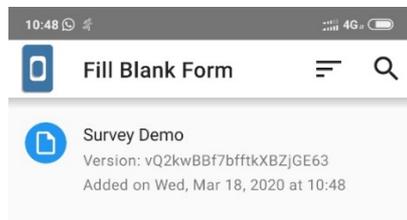


Step 6

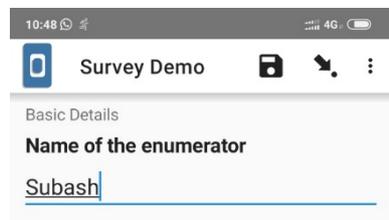
- **Step 6:** Enter the username and password of the Kobo account which is used to generate the survey. This is a onetime process and it won't ask for username and password again.
- **Step 7:** Click on the 'Get Blank Form' tab. All the surveys deployed by the respective Kobo account will be displayed. Select the survey which you want to fill.
- **Step 8:** Now go to Fill blank form. Here you will see the questionnaire that you had developed.
- **Step 9:** Now you can see all the questions - answer them till you reach the end of the survey.
- **Step 10:** Click on Save and Exit.



Step 7



Step 8



Step 9a

- **Step 11:** Once you have the internet, click on the send finalized form option. The Survey data collected will be sent to the server, which can be immediately accessed. So, the survey can be conducted even when there is no internet connection. Finished surveys can be sent at the end of the day once you have the internet connection.

10:13 4G

Survey Demo

Basic Details
Group

Age of the farmer
25.0

Family Size
40

Major occupation
Agriculture

What are the secondary occupation
Processing and Value addition, Others

* If other, mention
Contractor

Farming experience

Go To Start Go To End

Basic Details

Date and time of the Survey

Select date

Mar 10, 2020

Select time

10:49

Step 9b

Survey Demo

You are at the end of Survey Demo.

Name this form
Survey Demo

Mark form as finalized

Save Form and Exit

Step 10

10:13 4G

Main Menu

KoBoCollect v1.25.1
Part of KoBoToolbox
Based on ODK Collect

Fill Blank Form

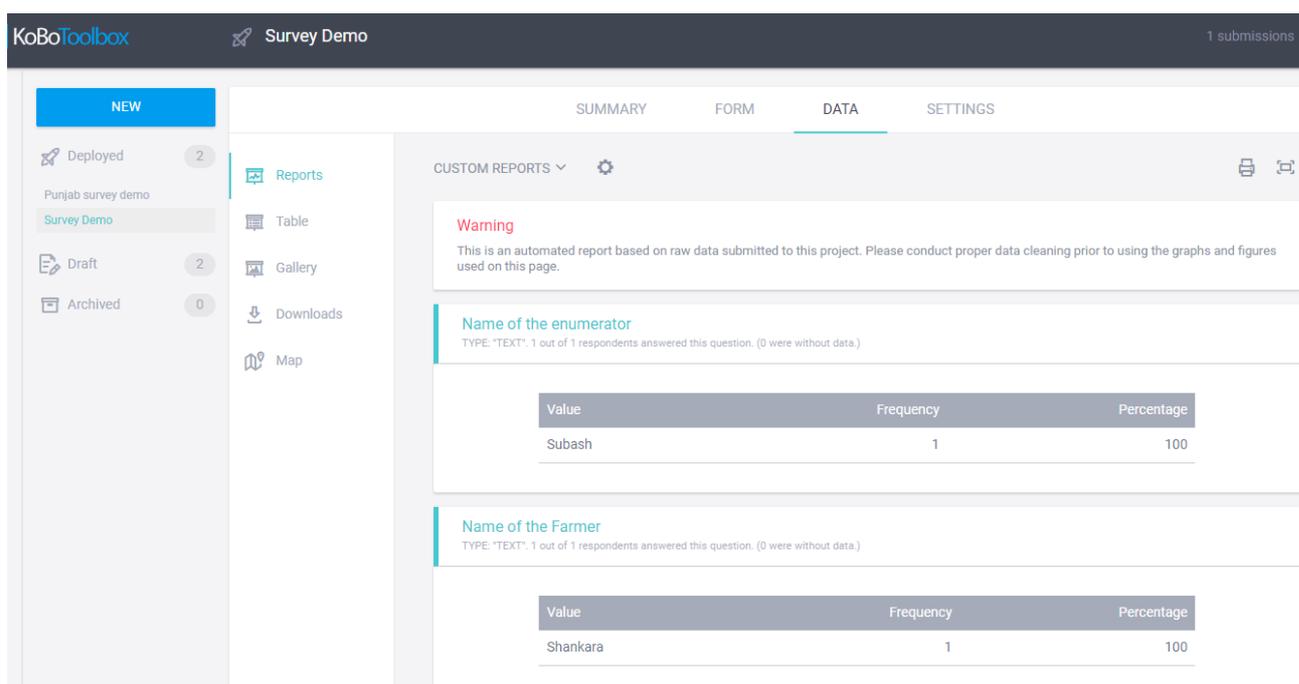
Edit Saved Form (1)

Send Finalized Form (1)

View Sent Form

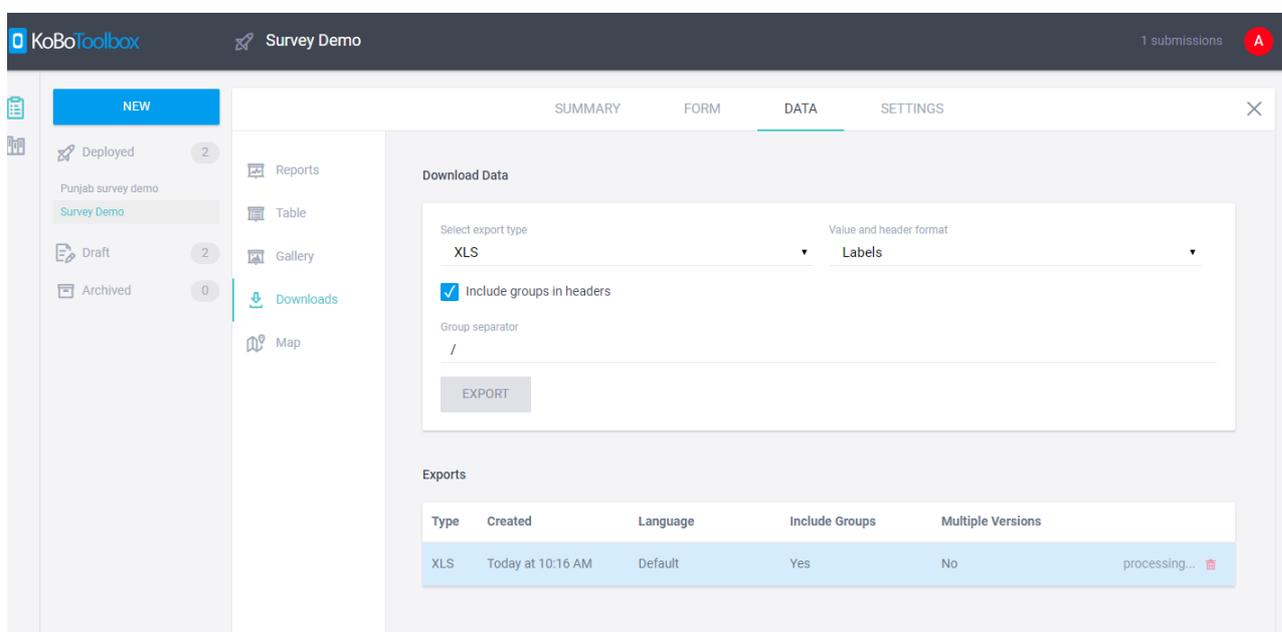
Step 11

- **Step 12:** Next step is to access the collected data from the server. Once the form is sent by the enumerator, we can access the data by clicking on data tab. Custom graphs and tables are displayed here.



Step 12

- **Step 13:** The data can be downloaded in XLS format. The person monitoring the survey can check for the consistency of data by tabulating or summarizing the relevant information. Tabulating the data by enumerator can also indicate a few enumerator-specific errors in data entry. Immediate feedback can be given to the respective enumerator about the error.



Step 13

Here are a few tips from our personal experience:

- Predict the common data collection/entry mistakes. Even pre-testing can be used for this purpose. Then use validation criteria to minimize them.

- If you wish to collect the data in the form of a table, see below

Sl. No. of person	Age	Education	Employment

Customizing table format in Kobo is difficult. It is better to group the questions – Age, Education, and Employment and repeat the groups as many times as you wish. You can use skip logic to regulate the number of times the question is repeated.

- Pre-testing of the questionnaire in CAPI is a must. Sometimes the skip logic may not work or some question/s may be incorrectly displayed or ordered. The errors can be minimized by identifying them during pre-testing.
- Tabulate the data as frequently as possible during real time monitoring. Tabulate responses by enumerator and observe for patterns in response. For instance, in a question to rank the constraints faced by farmers, if an enumerator follows a typical pattern like 3, 2, 4 and 1 for all the farmers, then there is a possibility that he might be filling responses on his own without even asking. You can talk to the enumerator about it and verify the details.
- On similar lines, another common error is ‘average value of response’. Once an enumerator finishes maybe 5 or 10 surveys, he gets a vague idea of how much fertilizer is used, what is the seed rate, etc. There is a possibility that he starts to enter those average values for all the future respondents. So, closely observe for such errors.

CONCLUSION

CAPI is emerging as an alternative to paper-based surveys. Though this method is very common across CGIAR institutes and other international research institutes, the use of these in National Agriculture Research and Education System (NARES) in India is low, to the best of our knowledge. **(We request those readers who have used CAPI for their survey to share their findings and experience in the comment session.)** It’s been taught in extension courses (E-extension, Advanced Research Methods) for quite some time. A common notion is that CAPI is only for larger surveys with huge funding, which as discussed is not a fact. There were several reasons, such as lack of accessible hardware, software and issues with user-friendliness, which kept them beyond the reach of researchers and students in NARES. Availability and prevalence of smartphones/ cheaper tablets and open access software which are user-friendly are opening up those doors. We strongly feel that it’s time for us to get smarter and embrace CAPI as it offers many advantages over paper-based surveys. In the initial phases this could pose some challenges, but as the saying goes ‘life without challenges is boring’!

References

- Careletto C and Gourlay S. 2019. A thing of the past? Household surveys in a rapidly evolving (agricultural) data landscape: Insights from the LSMS-ISA. *Agricultural Economics* 50(S1):51-62.
- DIME. 2020. Computer-Assisted Personal Interviews (CAPI).
[https://dimewiki.worldbank.org/wiki/Computer-Assisted_Personal_Interviews_\(CAPI\)](https://dimewiki.worldbank.org/wiki/Computer-Assisted_Personal_Interviews_(CAPI))
- Kumar S, Kumar R, Seema, Dhandapani A, Sivaramane N, Meena PC and Radhika P. 2017. Food consumption pattern in Telangana State - 2017. Hyderabad, India: ICAR-National Academy of Agricultural Research Management.
- KoBoToolbox. Available online on <https://www.kobotoolbox.org/>. Accessed on 19-03-2020.

Aditya KS, Scientist, Division of Agricultural Economics, ICAR-Indian Agricultural Research Institute, New Delhi. Email id: adityaag68@gmail.com;

Subash SP, Scientist, ICAR-National Institute of Agricultural Economics and Policy Research, New Delhi. Email id: subashspar@gmail.com and

Bhuvana N, PhD Scholar, Department of Agricultural Extension, PJTSAU, Hyderabad. Email: bhuvana.rao7@gmail.com

**AESA Secretariat: Centre for Research on Innovation and Science Policy (CRISP),
Road No 10, Banjara Hills, Hyderabad 500034, India
www.aesanetwork.org Email: aesanetwork@gmail.com**