

Digital data collection

Erick Baumgartner Vrinda Kapoor

29 June, 2021

Digital data collection

- ▶ Computer-assisted Interviewing (CAPI and CATI) uses an electronic device like phone, tablet or laptop
- ▶ Avoids post-collection processing costs and time lag
- ▶ Access data in real time
- ▶ Design complex surveys, including logic skips, validation rules, double-entry checks, and pre-filling data
- ▶ Capture GPS and record surveys
- ▶ Better monitoring or surveyor performance

Types of digital data collection

- ▶ Digital survey platforms
 - ▶ Survey CTO (used by LEAP researchers)
 - ▶ Dimagi
 - ▶ KoBo toolbox
- ▶ Interactive Voice Resource System (IVRS)
- ▶ Web-based platforms
 - ▶ Virtual Lab (Facebook, Google)
 - ▶ Twilio (WhatsApp)
 - ▶ Qualtrics (SMS, email)
- ▶ Implicit Association Tests
 - ▶ Inquisit
 - ▶ Open Sesame

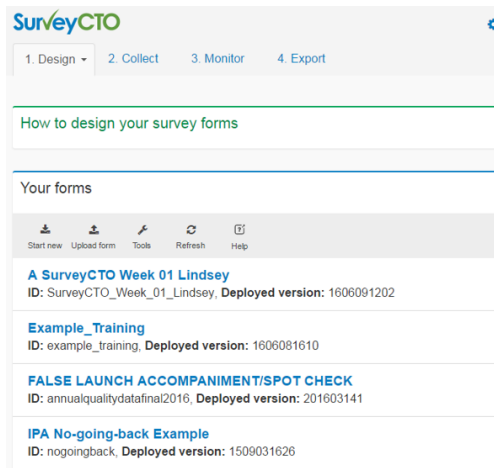
SurveyCTO

- ▶ Based on the Open Data Kit (ODK), an open-source platform
- ▶ SurveyCTO added improvements, documentation, encryption, and support
- ▶ Three components:
 - ▶ Server
 - ▶ Collect
 - ▶ Desktop

SurveyCTO Server

- ▶ Designing and uploading forms
- ▶ Uploading and attaching back-end data to the forms
- ▶ Basic data visualization
- ▶ Support community
- ▶ Preview of the uploaded form

SurveyCTO Server



The screenshot displays the SurveyCTO Server interface. At the top, the SurveyCTO logo is on the left, and a gear icon is on the right. Below the logo is a navigation bar with four steps: "1. Design", "2. Collect", "3. Monitor", and "4. Export".

Below the navigation bar is a section titled "How to design your survey forms".

Underneath is a section titled "Your forms". Below this title is a toolbar with five icons and their corresponding labels: "Start new", "Upload form", "Tools", "Refresh", and "Help".

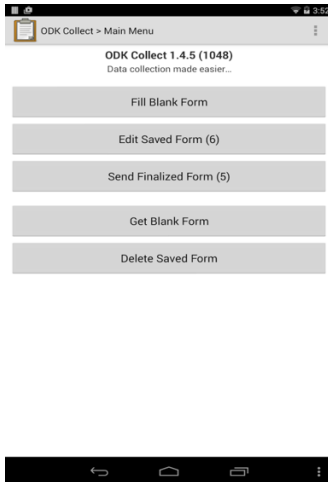
The "Your forms" section contains a list of four survey forms, each with a title and an ID:

- A SurveyCTO Week 01 Lindsey**
ID: SurveyCTO_Week_01_Lindsey, **Deployed version:** 1606091202
- Example_Training**
ID: example_training, **Deployed version:** 1606081610
- FALSE LAUNCH ACCOMPANIMENT/SPOT CHECK**
ID: annualqualitydatafinal2016, **Deployed version:** 201603141
- IPA No-going-back Example**
ID: nogoingback, **Deployed version:** 1509031626

SurveyCTO Collect

- ▶ Available on Android and iOS
- ▶ Collect data offline
- ▶ Upload surveys to the server
- ▶ Offline maps and GPS integration

SurveyCTO Collect



SurveyCTO Desktop

- ▶ Download encrypted datasets
- ▶ Multiple export options like automated dummy variables, long v/s wide dataset
- ▶ Auto-generate Stata do-files for labelling/cleaning
- ▶ Export GPS coordinates to Google Earth

SurveyCTO Desktop

SurveyCTO Desktop | My Desktop
SurveyCTO Desktop Edit Window Help



My Desktop

C:\Users\KApoor\SurveyCTO Desktop
Local Storage

Sync

Offline form tools

Local storage

SERVER SHORTCUTS

leapbocconi
leap@unibocconi.it

LOG IN

Forget this server

+ Add a shortcut

Desktop settings

Help and support

About SurveyCTO Desktop

+ Create a new desktop

SurveyCTO Desktop will process the raw data from your local storage and export it, according to these settings. **Note:** these are the export settings used for every export on this desktop.

Data format and processing

Submissions to include:

- Awaiting approval (pending)
- Approved
- Rejected

Export repeat groups in which format?

- Wide format
- Long format

Export to the following file type:

- .csv (universal)
- .sav (SPSS)
- .xlsx (Excel)

Responses to select_one and select_multiple fields:

- Export choice values (default)
 - Export choice values AND labels
- Export select_multiple responses as series of 1/0 columns?**
Automatically output select_multiple responses as a series of columns.

Treatment of enclosing groups in exports:

- Add groups to exported field names ("groupname-fieldname").
 - Exclude groups but still use them internally.
 - ignore groups so fields with same name export together (slowest).
- Override the default value separator in .csv files?**
- Export and merge to Excel files for some forms and datasets.**
- Export Google Earth .kml files for some forms.**
- Auto-run mail merge after export.**
If enabled, FORMID_template.docx files in your export directory will be auto-merged with your exported CSV files.
- Auto-run other processors after export.**
If enabled, all files in the "themenrnr" subdirectory of your export directory will be executed after every export.

Workflow

- ▶ Program survey using XLS form/ form designer
- ▶ Upload to SurveyCTO server
- ▶ Test it, test it, test it!
- ▶ Download form definition to devices using the Collect app
- ▶ Surveyors fill out and finalize forms
- ▶ Send forms back to SurveyCTO server
- ▶ Download dataset using SurveyCTO Desktop

Data collection

- ▶ Collection of data on field does not require net connectivity, open forms are sufficient
- ▶ Can be deployed remotely
- ▶ Can survey in remote areas
- ▶ Data transmission can be done at a later time, but good connectivity allows real time transmission as well
- ▶ If form changes, it has to be changed in all tablets

Features

- ▶ Double encryption of dataset
 - ▶ Cloud encryption
 - ▶ Public and private encryption key optional
- ▶ Required questions
- ▶ Constraints on questions like age, birth year, age of menarche
- ▶ Relevance (If no, then why?)

Features

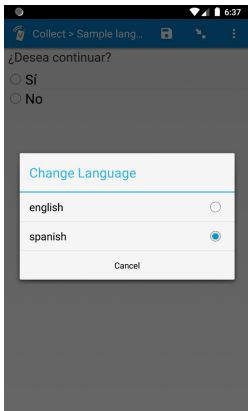
- ▶ Pre-loading data from baseline
- ▶ Logic checks (revenue, costs, profits)
- ▶ Skip patterns
- ▶ Repeat groups (crops, household members)
- ▶ Randomizing questions

Features

- ▶ Text audits
- ▶ Audio audits
- ▶ Speed limits (consent)
- ▶ GPS/location checks
- ▶ Digital signatures

Translations

Four languages automatic in every server: English, French, Spanish, and Hindi



Programming a survey

- ▶ Final(-ish) version of questionnaire ready
- ▶ Go to the design tab of the server
- ▶ Click 'start new'
- ▶ Give the form a name and ID
- ▶ Download XLS form template

XLS form

- ▶ Tabs
 - ▶ survey
 - ▶ choices
 - ▶ settings
 - ▶ some help tabs
- ▶ Question types
 - ▶ text
 - ▶ integer
 - ▶ select_one
 - ▶ select_multiple
 - ▶ note

XLS form

	A	B	C	E	F	G	H	I	
1	type	name	label	default	appearance	constraint	constraint message	relevance	di
2	start	starttime							
3	end	endtime							
4	deviceid	deviceid							
5									
6	note	intronote	Welcome to the sample form. Please swipe forward to continue.						
7									
8	select_one yesno	consent	Would you like to continue?						
9									
10	begin group	consented	Sample survey module		organized			\$(consent)=1	
11	text	name	What is your name?						
12	integer	age	How old are you?			.>3 and .<130	Please enter a valid age to continue.		
13									
14	note	confirmnote	Your name is \$(name) and your age is \$(age). Thank you.						
15	end group	consented							
16									

Navigation: survey | choices | settings | help-survey | help-choices | help-settings | [search]

Final thoughts

- ▶ Training enumerators on using tablets
- ▶ Bench-test a lot!
- ▶ Follow good variable naming conventions
- ▶ Add hints wherever possible
- ▶ Standardize values for 'don't know' or 'refuse to answer'

Thank you!