SAINT LUCIA

World Bank

Survey Solutions

Session IVa. Use of Technology for Data Capture

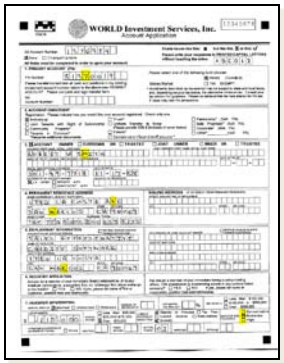
Survey Solutions For Census 2020

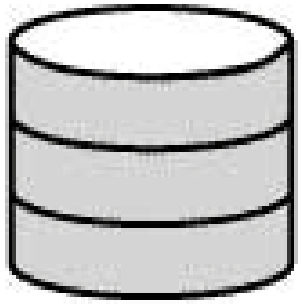
COMPUTER-ASSISTED PERSONAL and WEB INTERVIEWING

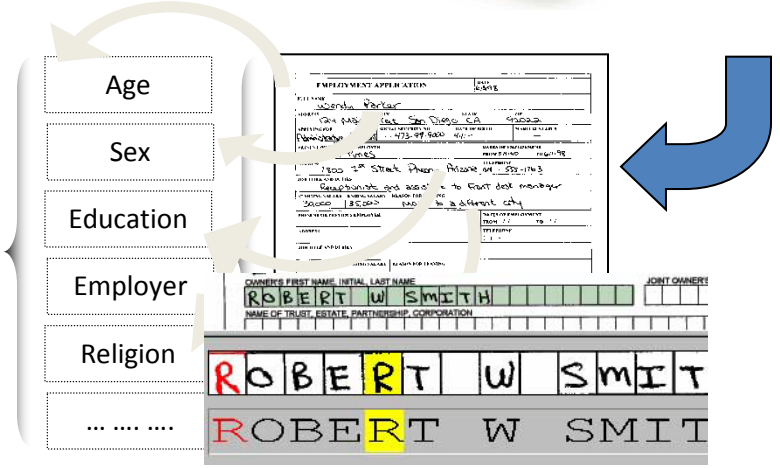
Edwin St Catherine, Director of Statistics, Saint Lucia

FROM PAPER TO TABLET DATA COLLECTION

Automated Paper Processing and Business Process







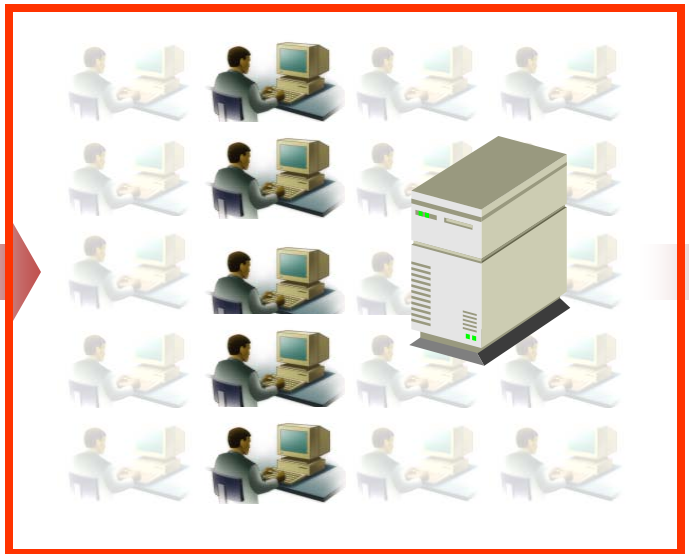
Where's the Improvement?

Same Data Entry Scenario – With Scanning Software

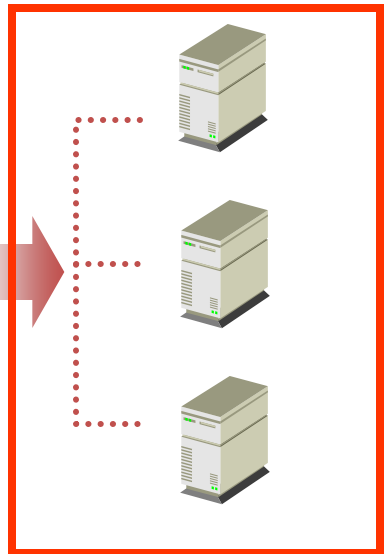
Automatically converts census forms to digital information, stores and index the form in dynamically created folders by Parish and ED, and sends the data to enterprise systems

Original Forms
(Paper or electronic)

Received Forms
Processed daily



15 Fewer Data Entry Clerks
75% Reduction cost and time



Immediate
Access to Data

Benefits of CAPI

Improve timeliness of data collection

Ensure data quality and comparability

Allow collection of new types of information/data

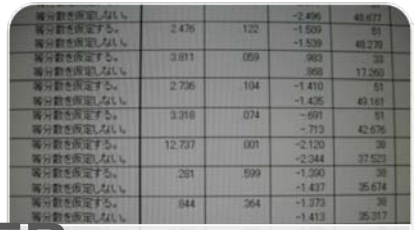
Cost-effective, sustainable solution for NSOs

Results:

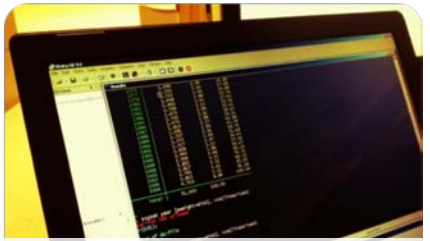
Data collection

CAPI/WEB

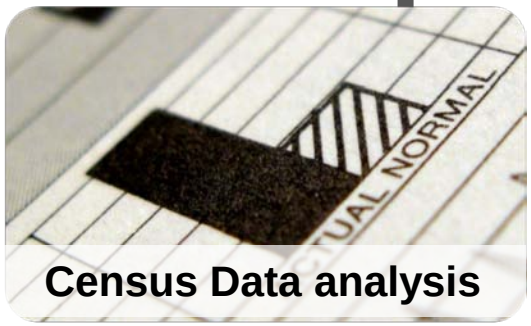
Data entry



Data cleaning



Data processing



Census Data analysis

Policy decisions

SPSS/SAS/STATA/ CSPRO/R

From data collection to policy: 4-5 months

CAPI System Requirements

- Simple yet flexible system for the non-expert users. Typical clients – National Stat Offices
- Functionality for
  - **data capturing:** entering data on a tablet
  - **census management**: managing teams of enumerators
  - **data management:** data aggregation, versioning, reporting
- Tablet-based with ability to display and navigate through multi-level large questionnaires.
- Support of complex validation algorithms.
- Cost effective system that can be used and supported by NSOs without external TA.

Survey Solutions: hybrid approach

- Sustainable, low-cost system for NSO
- Simple, flexible interface for questionnaire development and testing.
- Tablet interface allows easy navigation through complex questionnaires.
- Standardized survey management protocol based on the best practices of data collection
- Intuitive, informative survey status reporting, survey maps.
- Yet, powerful language for data validation and control of questionnaire flow.

Main differences from other systems

- **Out-of-the-box solution for survey data collection**: data capturing, *data management, and survey management*. No software on the market provides such a package. All other system focus mostly on data capturing.

- **Minimum TA; Focus on Capacity Building**: Survey Solutions is designed to minimize the TA. Lowest learning curve. Other systems require constant and significant TA. Expert versus User centered approach.

- **Designed for large surveys:** Survey Solutions is specifically designed for LSMS and HBS-types of surveys: Nested Rosters, Cascading and Linked questions, Roster-specific validations. Online collaboration. Web Interviewing for difficult to reach and more sophisticated respondents.

- **Data security:** Survey Solutions allows storing data on the local servers of NSO thus complying with the local data privacy and anonymity laws.

Survey Solutions: data capturing

- Large questionnaires: several thousand questions
- All standard types of questions:
  - text, numeric, date; multi-choice; dynamic lists;
  - Linked questions: Whose cow is it? User can select from the list of household members.
  - GPS location; Time; Barcode; Binary files (pictures)
  - **HQ/Supervisor-filled questions**
- Rosters: can be generated from:
  - Fixed lists; dynamic lists; numeric; multi-choice questions.
  - **Nested rosters with unlimited degree of nesting**
- Interviewer Comments on a question and interview
- Question Instructions

Para Data: Adaptive Survey Design

- **Improve data quality by correcting survey process during the field operation.**
- System records all events with timestamps that happen on a tablet: Data entry, Data correction, Responsibility changes, etc.
- Analysis of time per interview, time per question, section.
- Changes in productivity over time, for different interviewers, teams
- Quality control, monitoring and evaluation

Audit Trail in Each Data Case

- Accurate Tracking of Data on each household

Household Data

000d9f007a73492ead8df3451d267ede.tab

Person Data

00ad1581885b404a875483dc6ac838b4.tab

Households

Id	disctrict	enumeration_district	ref_week	month	year	address	community	town	district	household_no	GeoLocation__Latitude	GeoLocation__Longitude	GeoLocation__Accuracy	GeoLocation__Altitude	GeoLocation__Timestamp
001c105eb3054f3c84e0baea1484f9a5	2	2304	5/13/2017	5	2017	Bagatelle	Bagatelle	Castries	2	16	14.00038	-60.9833	20	-10.6099	2017-05-16T15:15:31
009c9795edfb4947b3fdca61927335d2	8	9801	5/6/2017	5	2017	jetrine saltibus	jetrine	laborie	Laborie	51	13.78323	-61.0063	9	103	2017-06-30T18:22:00
00ad1581885b404a875483dc6ac838b4	9	113	5/2	5		Ca ner t	Cant en t	ux f		115	13.75567	-60.9776	4	24	2017-05-10T23:09:40
00c5476754934bcb8460af606f4d0140	10	12109	7/8/2017	7	2017	viger	viger	micoud	10	33	13.78704	-60.9325	6	67	2017-07-25T21:00:34
00e6c037d7c5408eb8198ae96a5c71c9	10	12101	5/6/2017	5	2017	desruisseaux	desruisseaux	micoud	10	99	13.80612	-60.9462	48	141.4383	2017-05-27T13:59:18
0188c26c0ca24467bb344e214a4c0eba	1	501	7/29/2017	7	2017	Faux Au Faux	Castries	Castries	1	40	14.00877	-60.9953	20	84.71567	2017-08-27T19:48:16
0189e6bc56ab44b794fd43725e44f549	2	1301	9/30/2017	9	2017	Morne Du Don	Morne Du Don	Castries	2	75	14.01151	-60.9801	7	67	2017-11-04T16:12:55

Automatic Management of links between records is a powerful management feature which avoids the creation of duplicates and manages the unique tracking of records

Persons

Id	Relations_to_head_of_HH	sex	Resident_of_SL	ParentId1
0	1	1		1001c105eb3054f3c84e0baea1484f9a5
1	3	2		1001c105eb3054f3c84e0baea1484f9a5
2	7	1		1001c105eb3054f3c84e0baea1484f9a5
3	5	2		1001c105eb3054f3c84e0baea1484f9a5
0	1	1		1009c9795edfb4947b3fdca61927335d2
1	2	2		1009c9795edfb4947b3fdca61927335d2
2	3	1		1009c9795edfb4947b3fdca61927335d2
0	1	1		100ad1581885b404a875483dc6ac838b4
1	2	2		10 a 15 1 5b404a875483dc6ac838b4
2	3	1		100ad1581885b404a875483dc6ac838b4
3	3	2		100ad1581885b404a875483dc6ac838b4
0	1	1		100c5476754934bcb8460af606f4d0140
0	1	2		100e6c037d7c5408eb8198ae96a5c71c9
1	5	2		100e6c037d7c5408eb8198ae96a5c71c9
0	1	1		10188c26c0ca24467bb344e214a4c0eba
1	2	2		10188c26c0ca24467bb344e214a4c0eba
2	3	2		10188c26c0ca24467bb344e214a4c0eba
0	1	2		10189e6bc56ab44b794fd43725e44f549
1	3	1		10189e6bc56ab44b794fd43725e44f549
2	5	2		10189e6bc56ab44b794fd43725e44f549

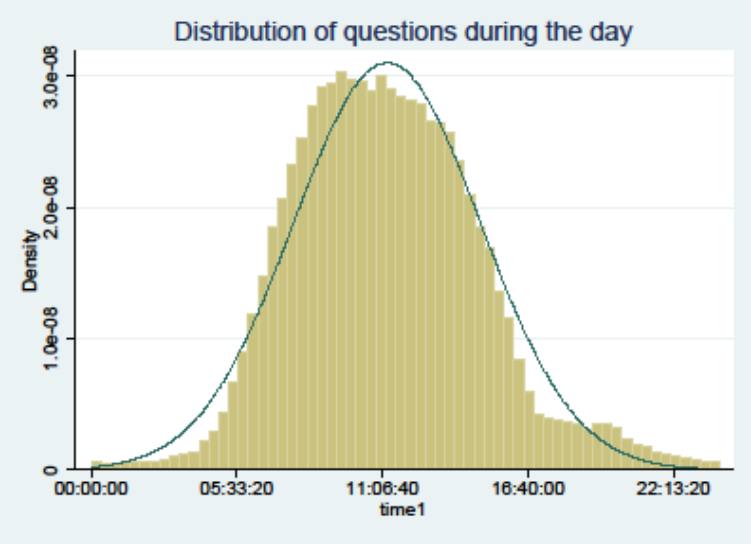
Household Questionnaire Capture Log on Tablet

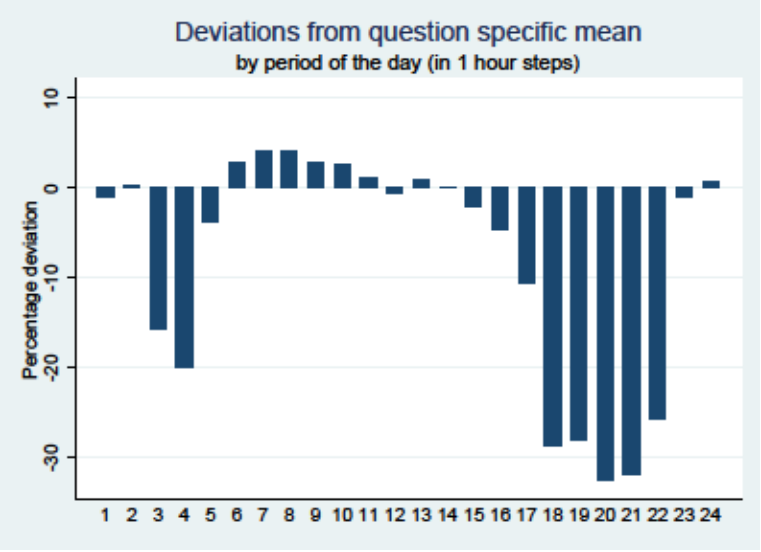
• AnswerSet	05/05/2017	13:36:51	disctrict	9	
• AnswerSet	05/05/2017	13:36:51	enumeration_district		11204
• AnswerSet	05/05/2017	13:36:51	subsample	5042b	
• AnswerSet	05/05/2017	13:36:51	ref_week	5/20/2017	
• AnswerSet	05/05/2017	13:36:51	district	Vieux-Fort	
• AnswerSet	05/05/2017	13:36:51	household_no	98	
• QuestionDisabled				Males_Left_SL_to_live_abroad	
• QuestionDisabled				Females_left_SL_to_live_abroad	
• QuestionDisabled				Food_Insecure_NoofTimes	
• QuestionDisabled				indicate_if_makeshift	
• QuestionDisabled				other_type_ofdwelling	
• QuestionDisabled				future_monthly_rent	
• QuestionDisabled				monthly_rent	
• SupervisorAssigned	05/05/2017	13:36:51			
• InterviewerAssigned	05/18/2017	13:13:02			
• ReceivedByInterviewer	05/18/2017	21:19:32			
• AnswerSet	05/23/2017	20:56:45	GeoLocation	13.745046,-60.9645087[24.1809997558594]0	
• AnswerSet	05/24/2017	13:24:57	address	Cedars Heiht	
• AnswerSet	05/24/2017	13:25:06	community	"Cedars "	
• AnswerSet	05/24/2017	13:25:12	address	Cedars Height	
• AnswerSet	05/24/2017	13:25:15	community	Cedars Height	
• AnswerSet	05/24/2017	13:25:20	town	Vieux fort	
• Completed	05/24/2017	13:25:33	No contact		
• ReceivedBySupervisor		05/24/2017	13:33:52		
• ApproveBySupervisor		05/24/2017	14:44:48	no contact	
• ApproveByHeadquarter		06/16/2017	17:56:43		

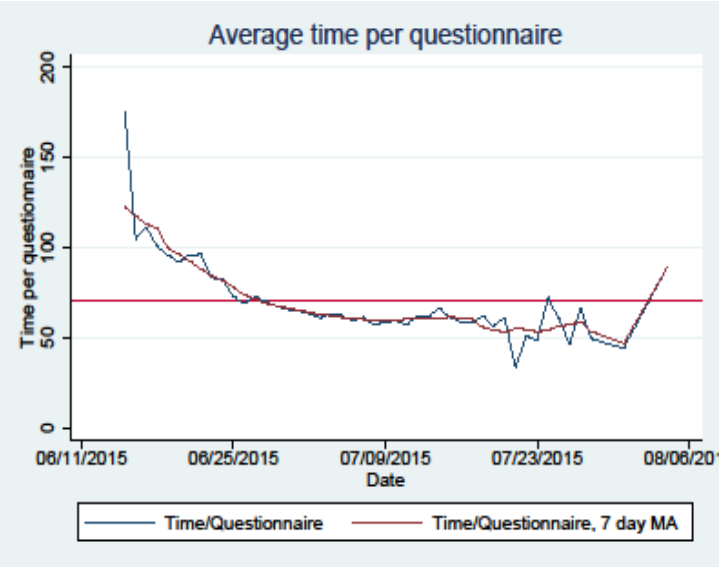
Person Questionnaire Capture Log on Tablet

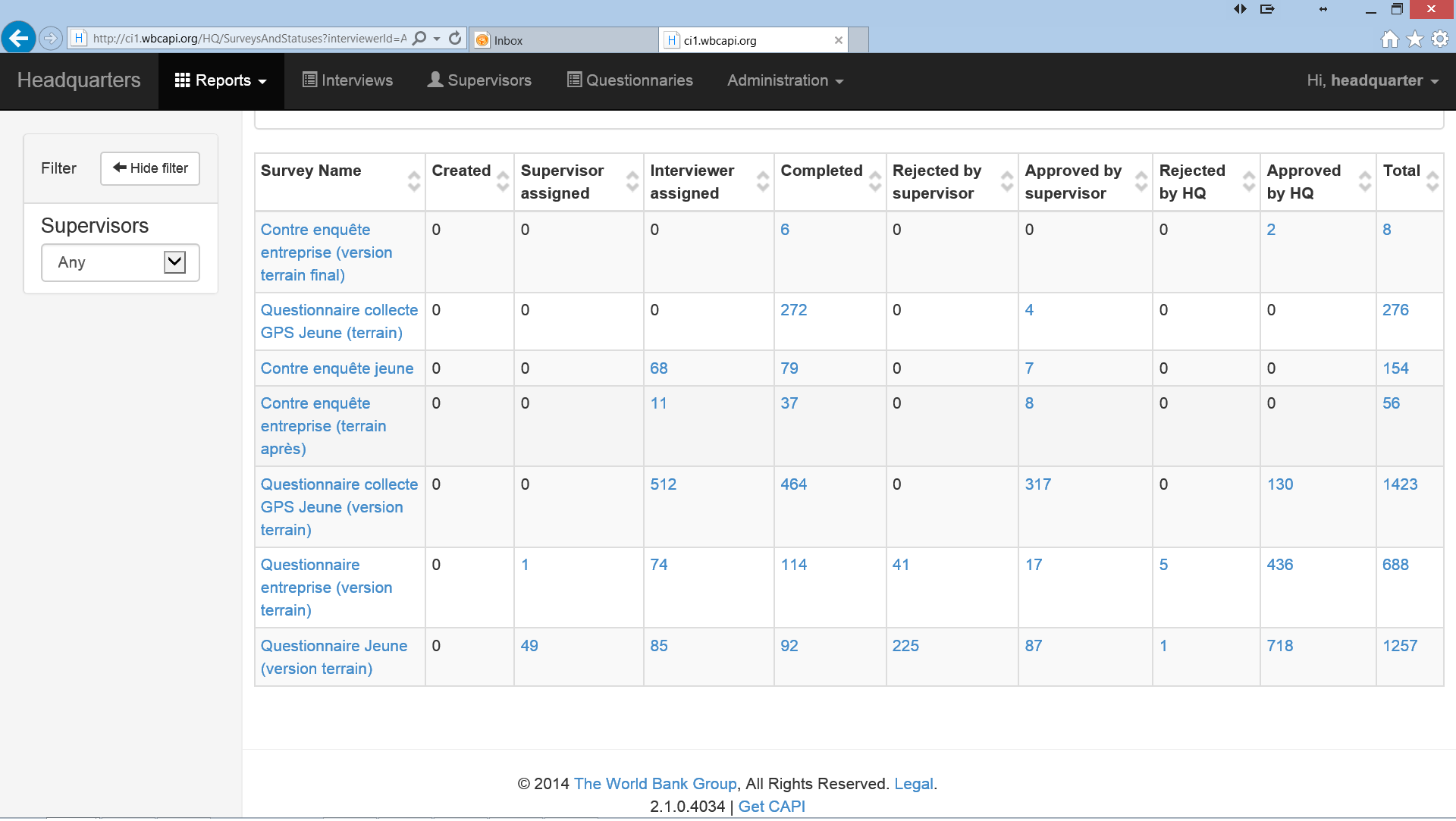
•	AnswerSet	05/05/2017	13:36:50	disctrict	9	
•	AnswerSet	05/05/2017	13:36:50	enumeration_district	11305	
•	AnswerSet	05/05/2017	13:36:50	subsample	5041b	
•	AnswerSet	05/05/2017	13:36:50	ref_week	5/6/2017	
•	AnswerSet	05/05/2017	13:36:50	district	Vieux-Fort	
•	AnswerSet	05/05/2017	13:36:50	household_no	115	
•	QuestionDisabled				Males_Left_SL_to_live_abroad	
•	QuestionDisabled				Females_left_SL_to_live_abroad	
•	QuestionDisabled				Food_Insecure_NoofTimes	
•	QuestionDisabled				indicate_if_makeshift	
•	QuestionDisabled				other_type_ofdwelling	
•	QuestionDisabled				future_monthly_rent	
•	QuestionDisabled				monthly_rent	
•	SupervisorAssigned	05/05/2017	13:36:50			
•	InterviewerAssigned	05/05/2017	14:03:00			
•	ReceivedByInterviewer	05/05/2017	17:23:54			
•	AnswerSet	05/10/2017	23:02:01	GeoLocation	13.75566509,-60.97762324[4]24	
•	AnswerSet	05/10/2017	23:03:27	phone		
•	AnswerSet	05/10/2017	23:03:34	contact		
•	AnswerSet	05/10/2017	23:03:44	Number_of_ppl	4	
•	AnswerSet	05/10/2017	23:03:46	Qty_persons_live_abroad	0	
•	AnswerSet	05/10/2017	23:03:47	ChildUnder5Death	2	
•	AnswerSet	05/10/2017	23:03:48	Food_Insecurity	2	
•	AnswerSet	05/10/2017	23:04:02	name	Fake Name	0
•	AnswerSet	05/10/2017	23:04:06	Relations_to_head_of_HH	1	0
•	AnswerSet	05/10/2017	23:04:08	sex	1	0
•	AnswerSet	05/10/2017	23:04:09	Resident_of_SL	1	0
•	AnswerSet	05/10/2017	23:04:16	level_of_education	2	0
•	AnswerSet	05/10/2017	23:04:18	health	1	0
•	AnswerSet	05/10/2017	23:04:19	health	2	0
•	AnswerSet	05/10/2017	23:04:25	age	52	0
•	AnswerSet	05/10/2017	23:04:29	Highest_level_of_exam	1	0
•	AnswerSet	05/10/2017	23:04:29	Highest_level_of_exam	2	0
•	AnswerSet	05/10/2017	23:04:33	Highest_level_of_exam	1	0
•	AnswerSet	05/10/2017	23:04:38	training_received	1	0
•	AnswerSet	05/10/2017	23:04:39	Method_of_training	2	0
•	AnswerSet	05/10/2017	23:04:44	Working_history_for_past	52	0
•	AnswerSet	05/10/2017	23:04:47	economic_activity	1	0
•	AnswerSet	05/10/2017	23:04:51	multiple_job_holding	2	0
•	AnswerSet	05/10/2017	23:04:56	usual_working_time	40	0
•	AnswerSet	05/10/2017	23:04:58	actual_hrs	40	0

Paradata: Improving data quality



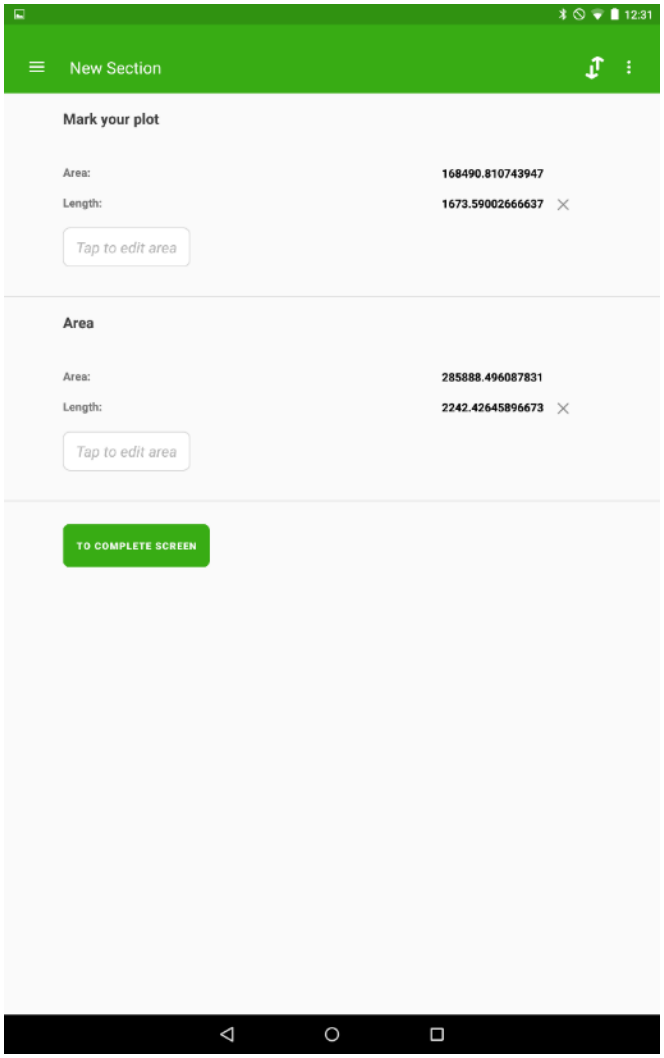


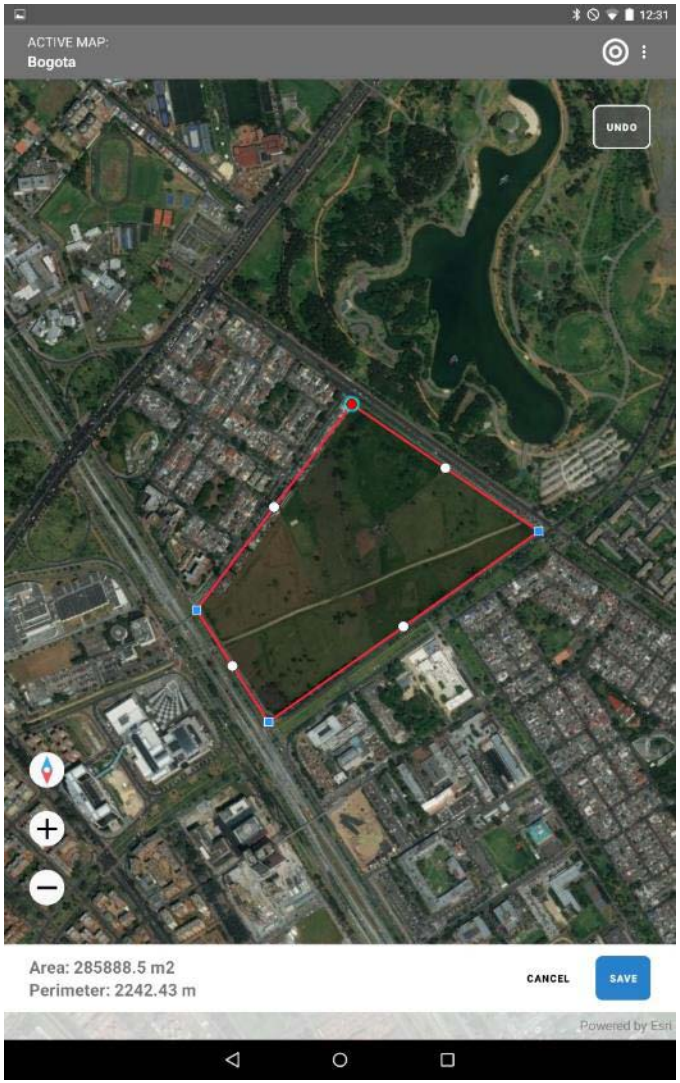


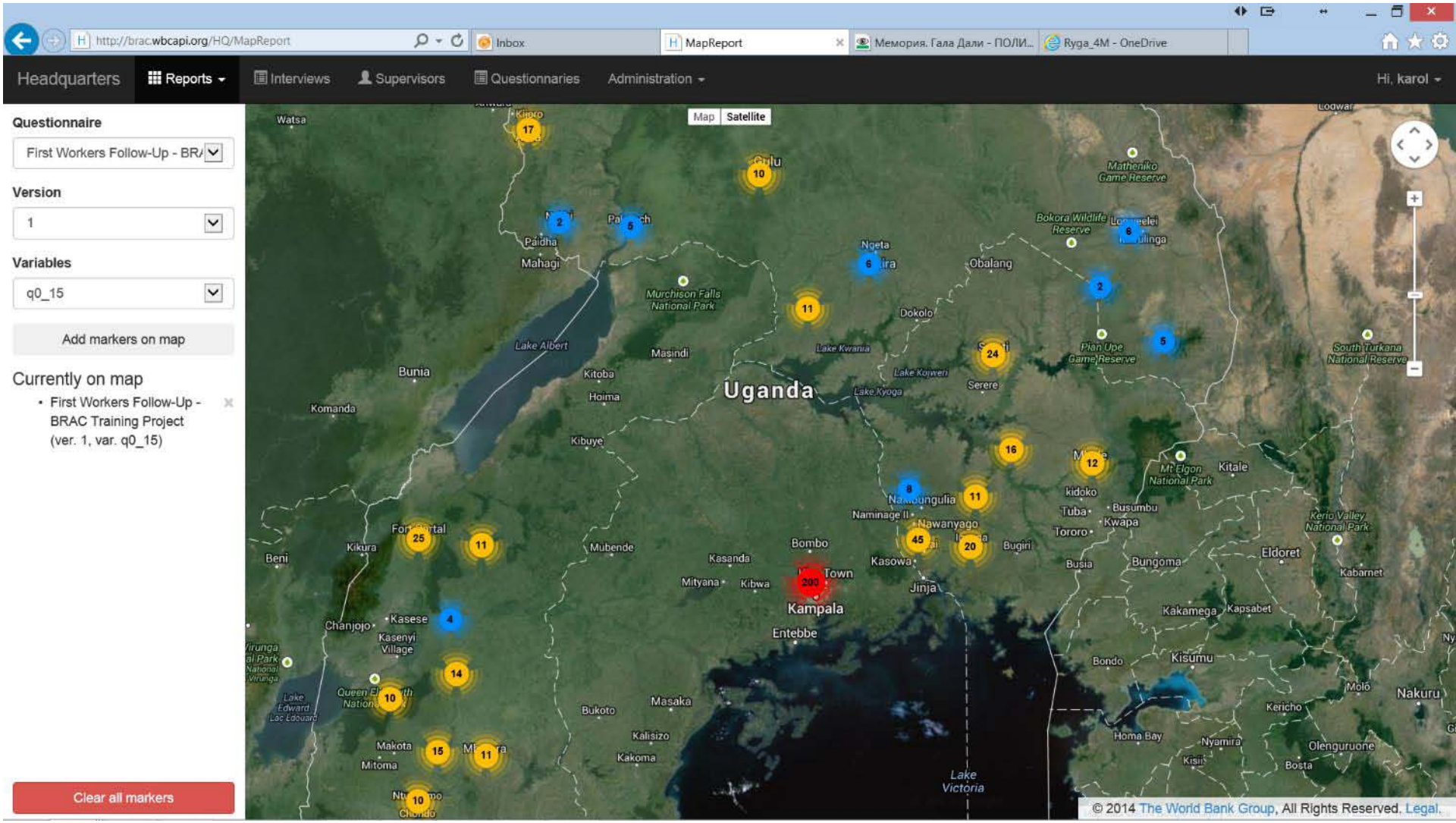


Real time status of interviews, process flow can be altered through system API

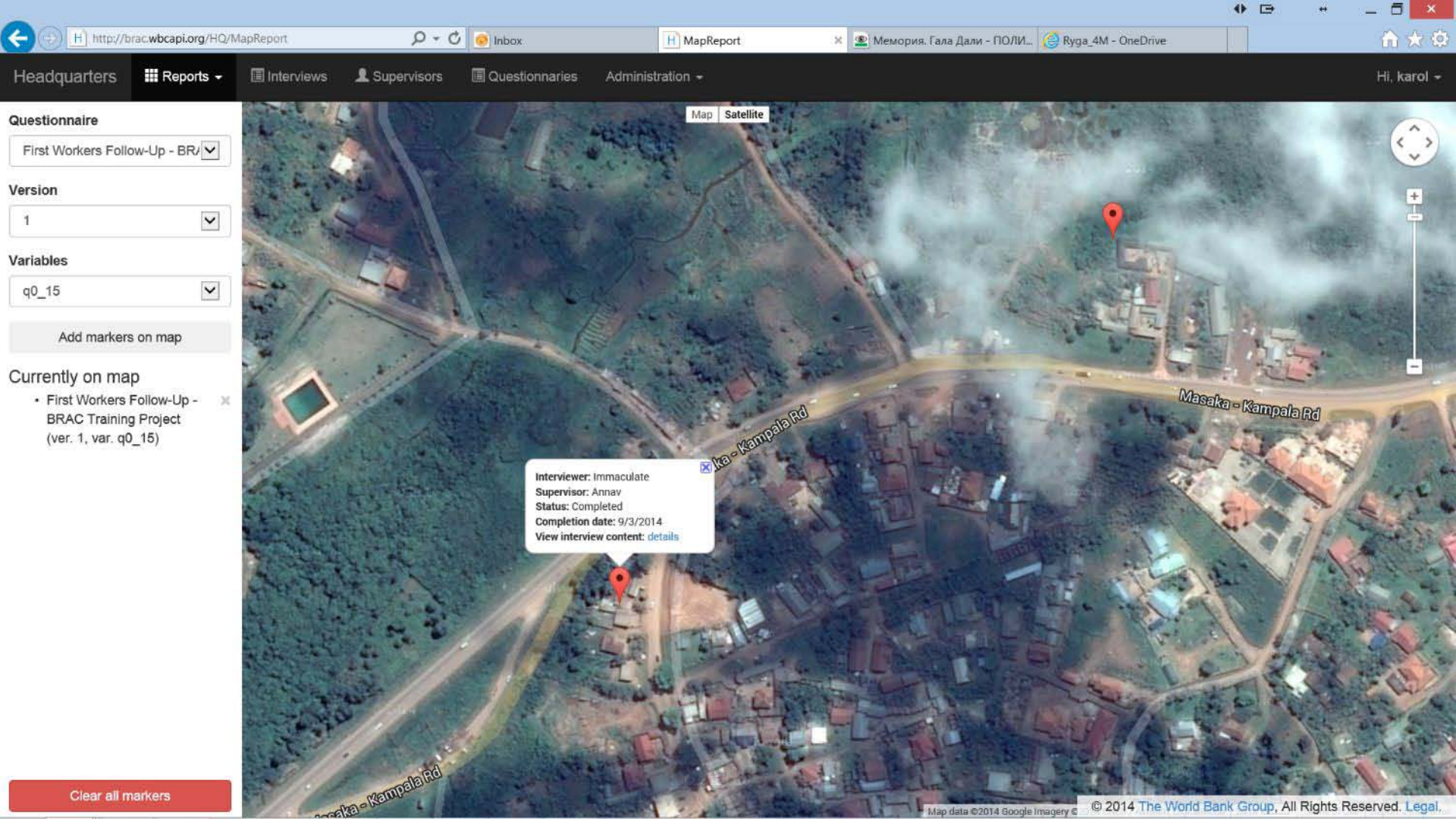
ESRI Map Integration on GIS For Sub-Meter Accuracy for GPS Location







Map of the survey



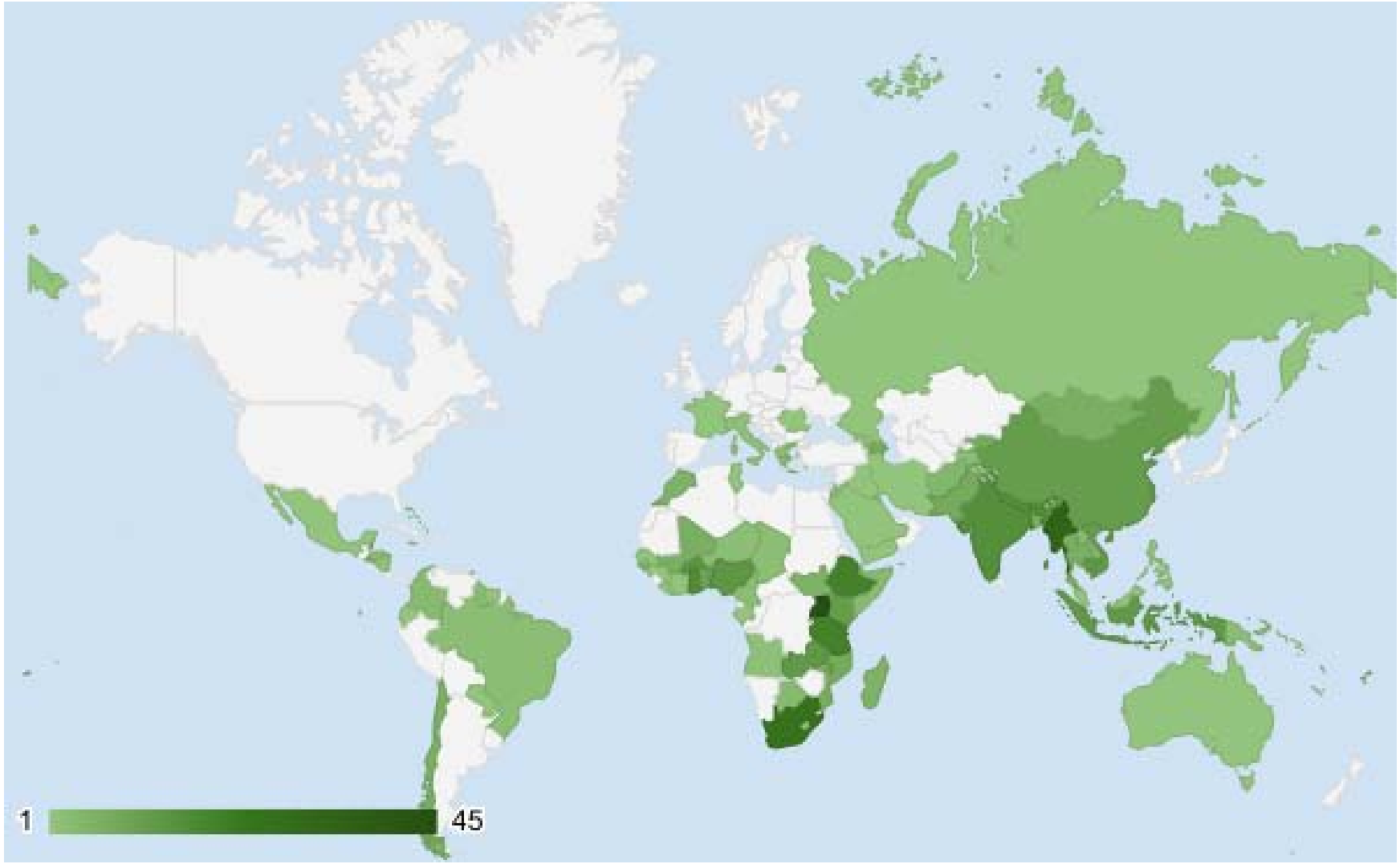
Monitor the survey by checking the GPS location of where and when the interview took place.

Collaboration with AWS and Facebook to Reduce Costs of Data For Census/Surveys

- Sampling based on population estimates using high resolution satellite images and ML
- Anthropometric measurements using image recognition
- Verification of individuals in panel surveys using face recognition and face aging algorithms.
- Experimentation in determining the reliability of answers based on facial expressions.
- ML for non standard units
- Sign language recognition

SURVEY PIPELINE: *700+ NATIONAL SURVEYS IN 109 COUNTRIES*

7,500,000+ FACE TO FACE INTERVIEWS



SURVEY/CENSUS SCOPE

- **Census: South Pacific Countries, Fiji etc Saint Lucia in Planning Phase**

- **Survey types**: LSMS, HBS, LFS, Enterprise Survey, EDU, Health

- **Large Clients:** India NSSO; Stats SA; Thailand NSO; FAO; IFAD; OECS; AfDB; IADB; OECS; IFS, LSTD, …, BRAC; OPM; Mathematica

- **Largest survey:** SA – 2,600,000 households, 12,000 enumerators, 500 supervisors, 60K questionnaires per day
- **Questionnaire:** Malawi – 3,000+ questions, 94 rosters

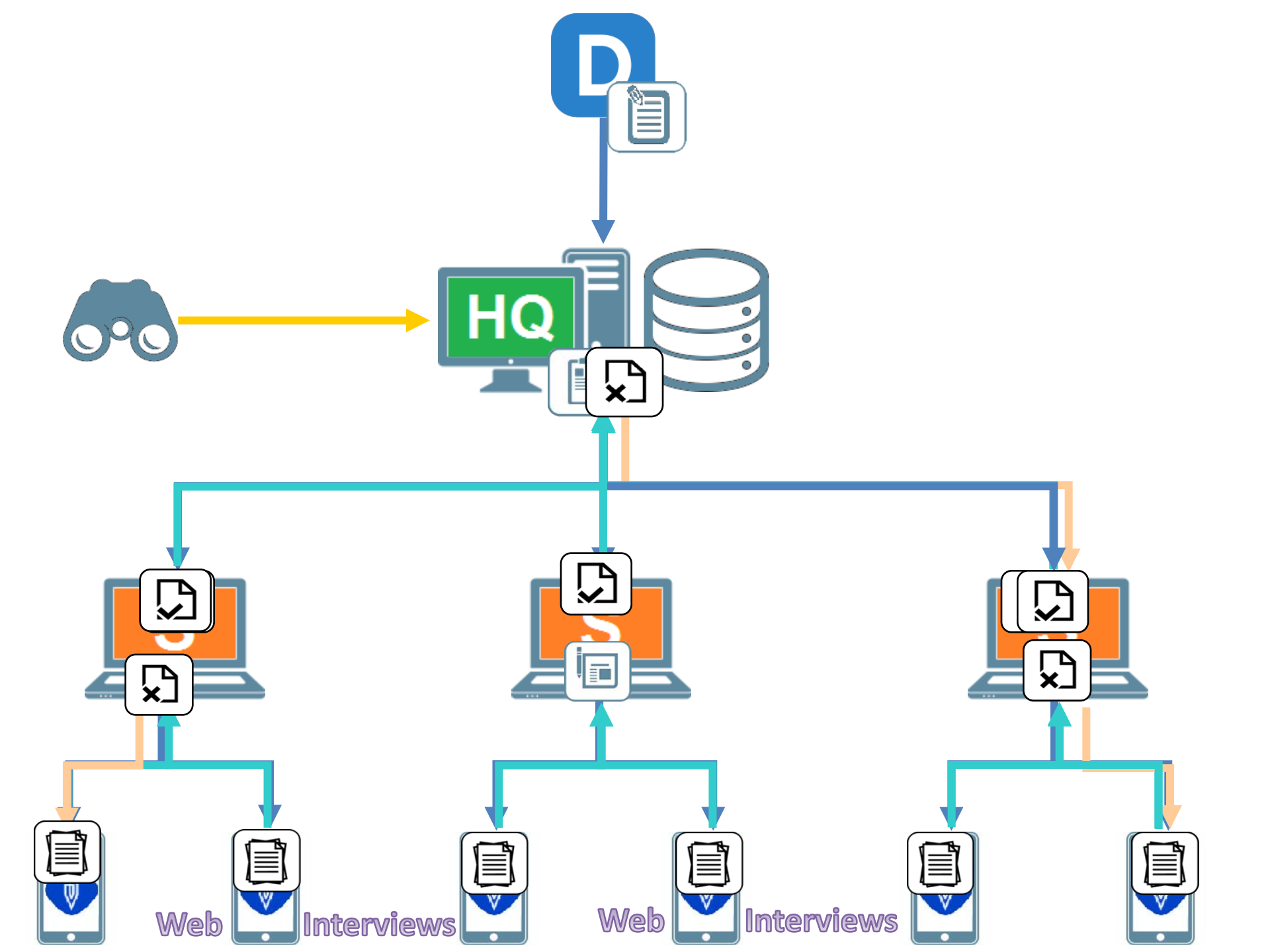
- **Other large surveys**:
  - Indonesia Sakarnas 200,000
  - Ethiopia, Uganda, Malawi LSMS
  - Nigeria LSMS
  - WAEMU Household surveys in 12 countries of West Africa
  - Many others

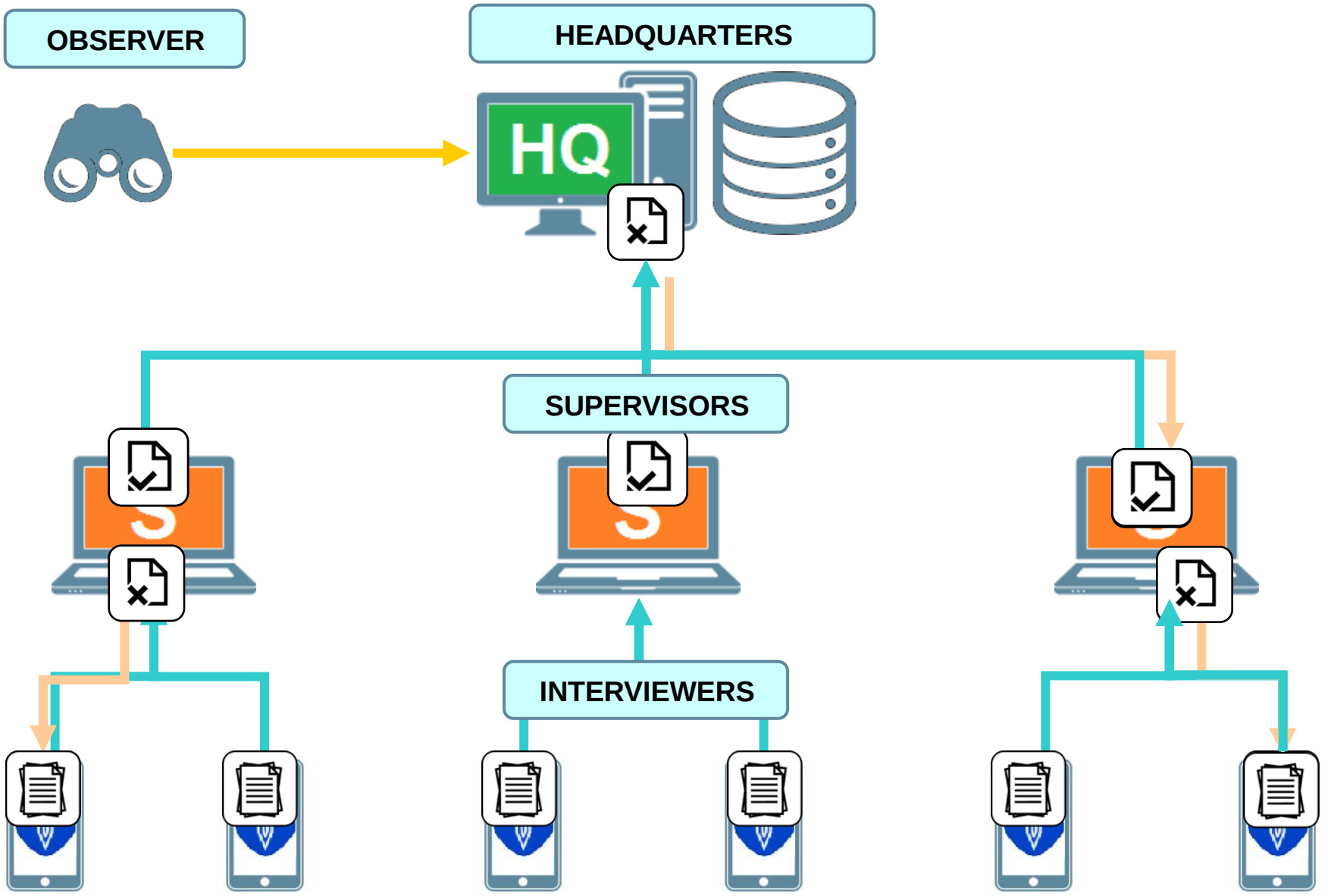
PROCESS OF IMPLEMENTATION IN ST LUCIA

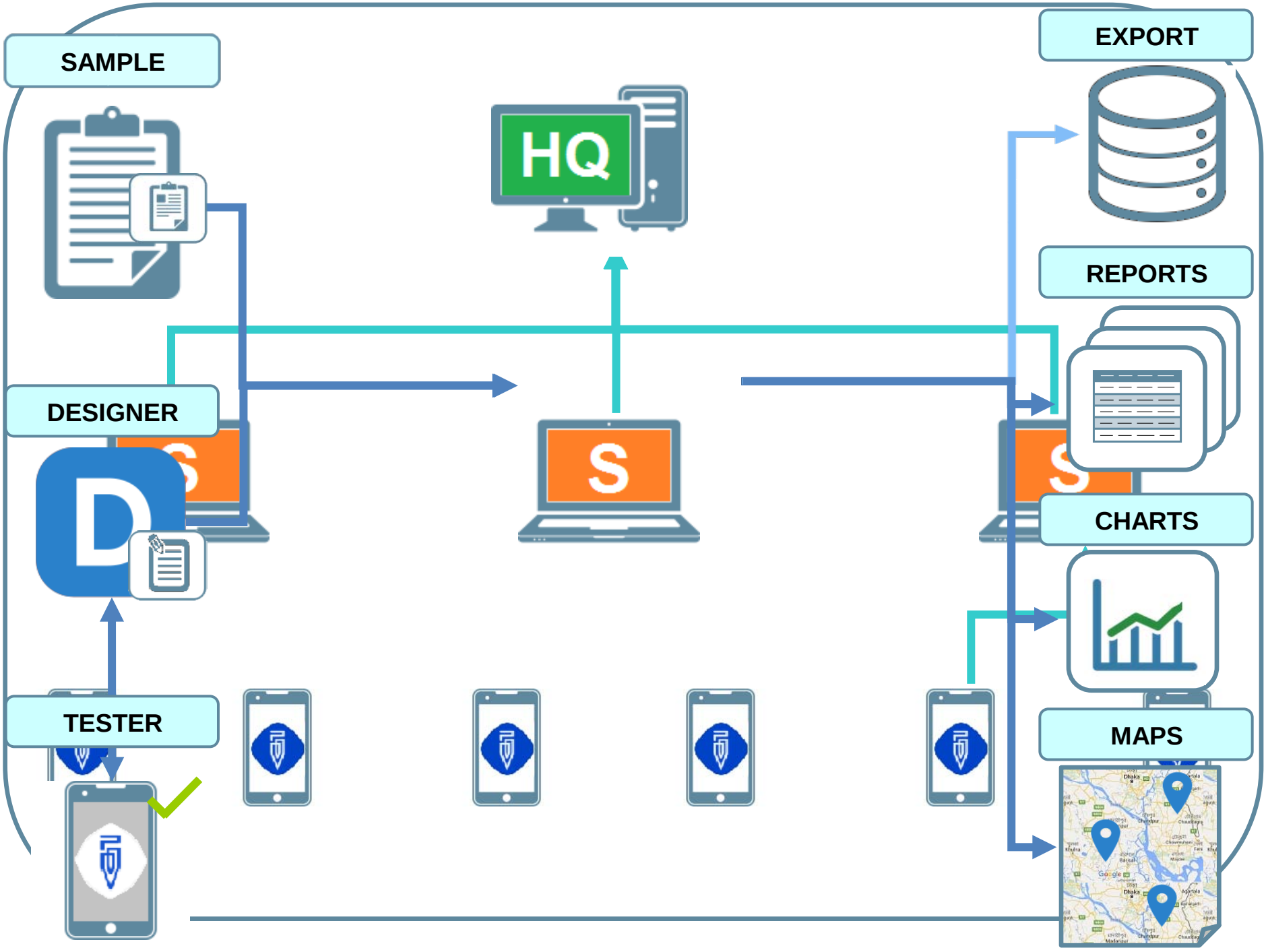
- From January 2013, full implementation of Survey Solutions v5.5
  - Quarterly Labour Force Survey
- The SLC/HBS 2016 is implemented using Survey Solutions
  - Completed a video in collaboration with the FAO/GIS to promote the use of Survey Solutions

STARTUP OF CAPI IMPLEMENTATION PROCESS

- CSO Trinidad expected to commence in April 2016.
- Suriname: Request made for Training to the World Bank Computational Tools Division
- St Vincent and the Grenadines/Grenada/St Kitts and Nevis/Antigua and Barbuda: Use of Survey Solutions in the LFS upcoming.
- Belize – LFS and Pilot Testing for HBS


HQ
Web Interviews
Web Interviews


OBSERVER
HEADQUARTERS
HQ
SUPERVISORS
INTERVIEWERS


SAMPLE
HQ
EXPORT
REPORTS
DESIGNER
CHARTS
TESTER
MAPS

Main Issues

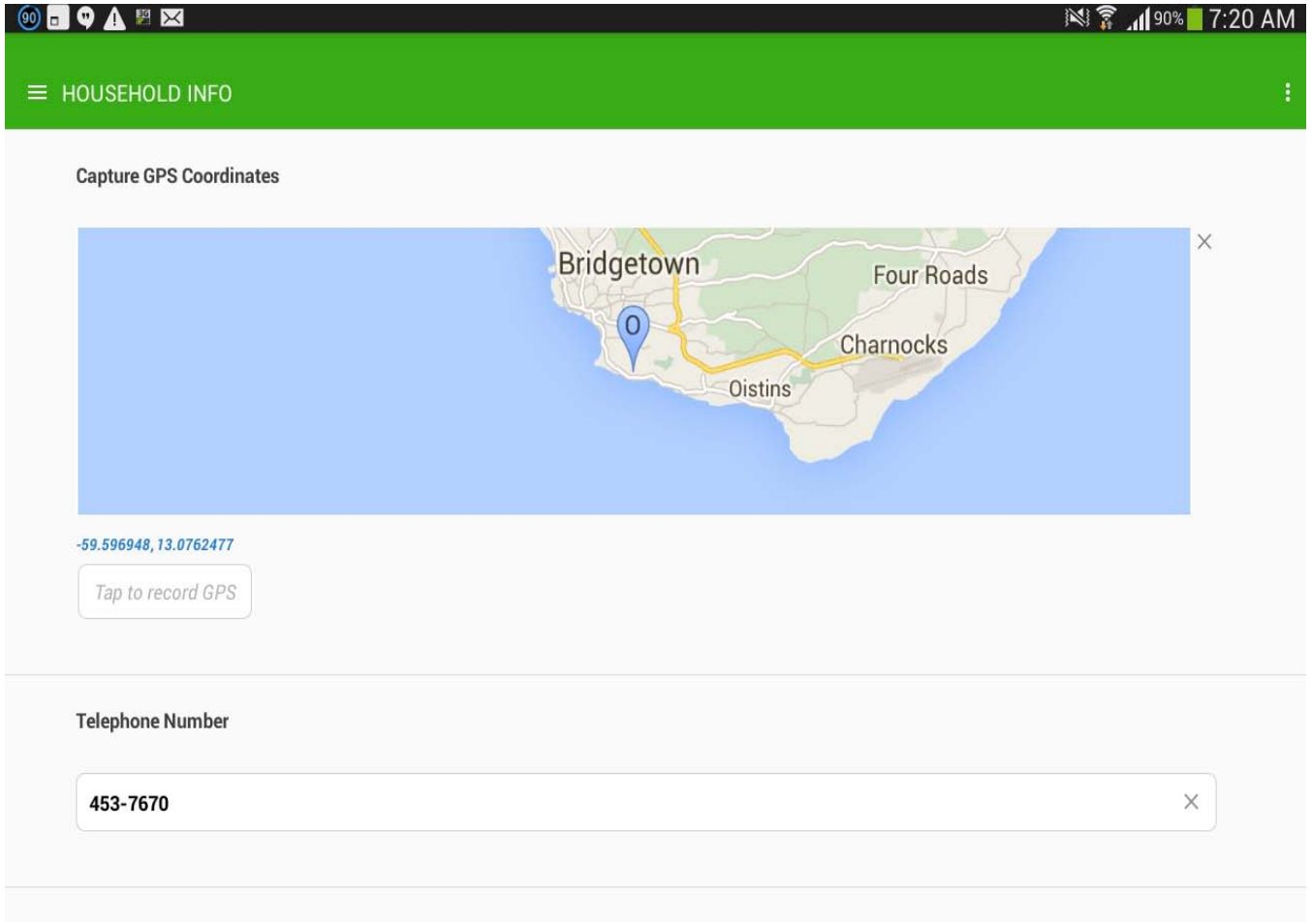
- Number of tablet devices to purchase for a de-jure census to be completed within four weeks
  - No of households to be surveyed
  - Final length of Questionnaire
  - Monitoring reports to be produced using data submitted on each EA
  - Elimination of control forms used in paper environment

- Items to preload on Headquarters Module and Tablet
  - Household Address list Census Questionnaire – Created Dynamically or preloaded
  - Pre-loaded showing routes to each address using Open GIS map.me or Google Maps, actions with missing addresses
  - Maps of shape files to constrain enumerators within enumeration boundaries
  - Census Questionnaire with full list of validations
  - 95% of validations can be pre-loaded on the tablets, 5% post census validation especially hot-decks type imputation validations

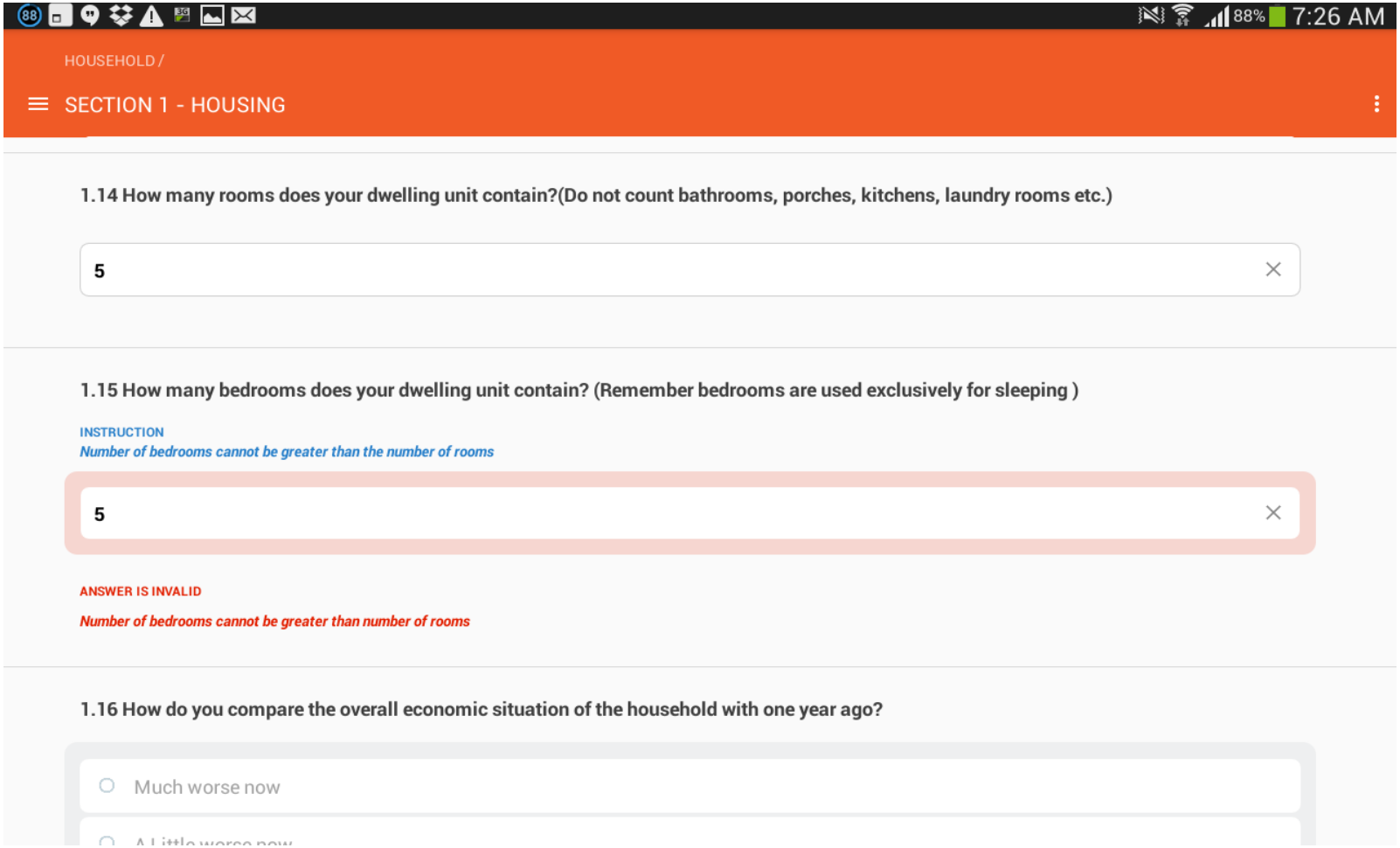
- Headquarter module of the survey requires use of desktop or laptop computer. Supervisor must demonstrate ability to use a brower and must own a laptop or desktop computer. May need to have some backups.

- Decide on rules for data submission to HQ
  - At Least once every other day to ensure supervisor can review and reject any forms

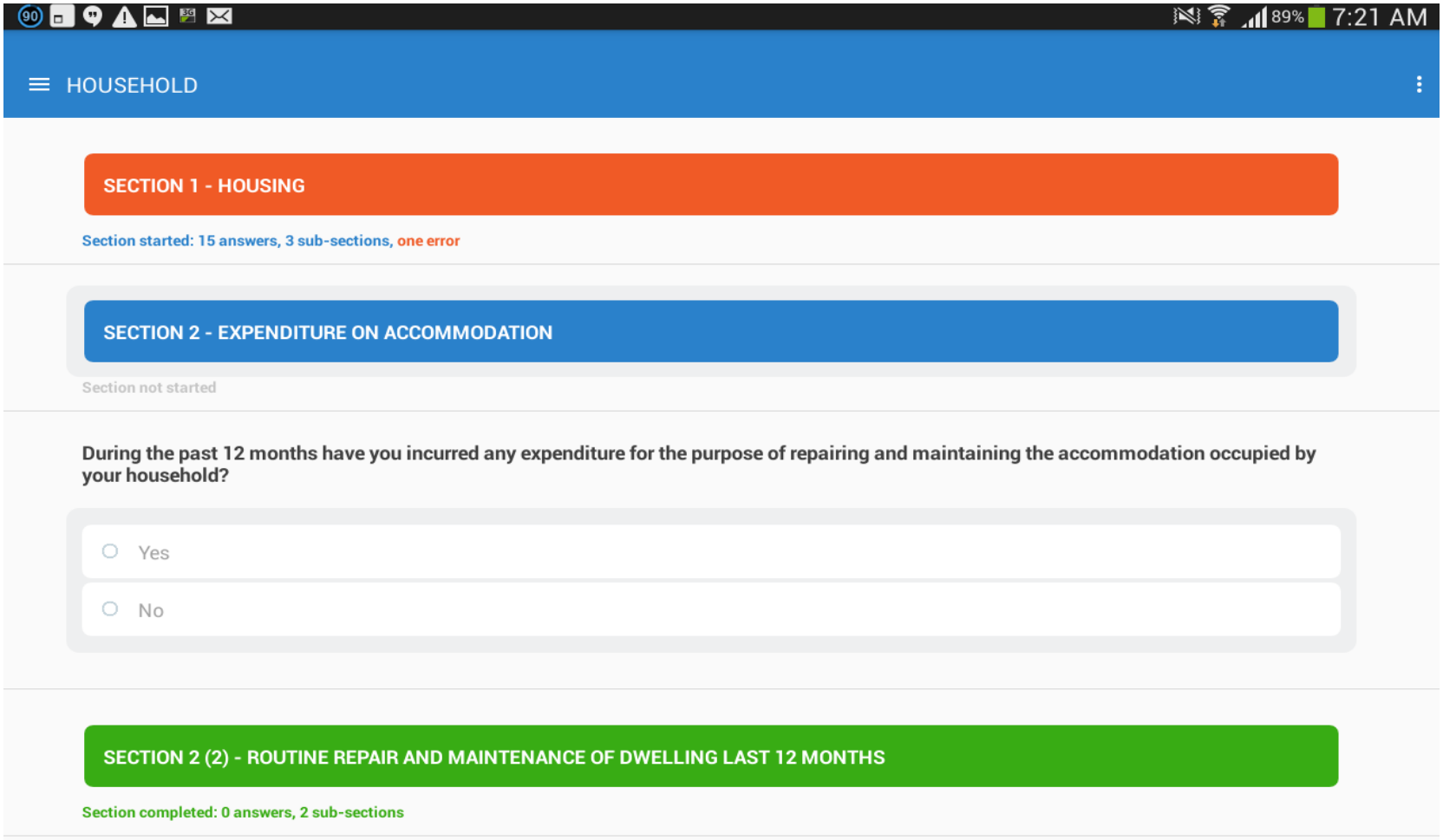
Tablet Application - Control



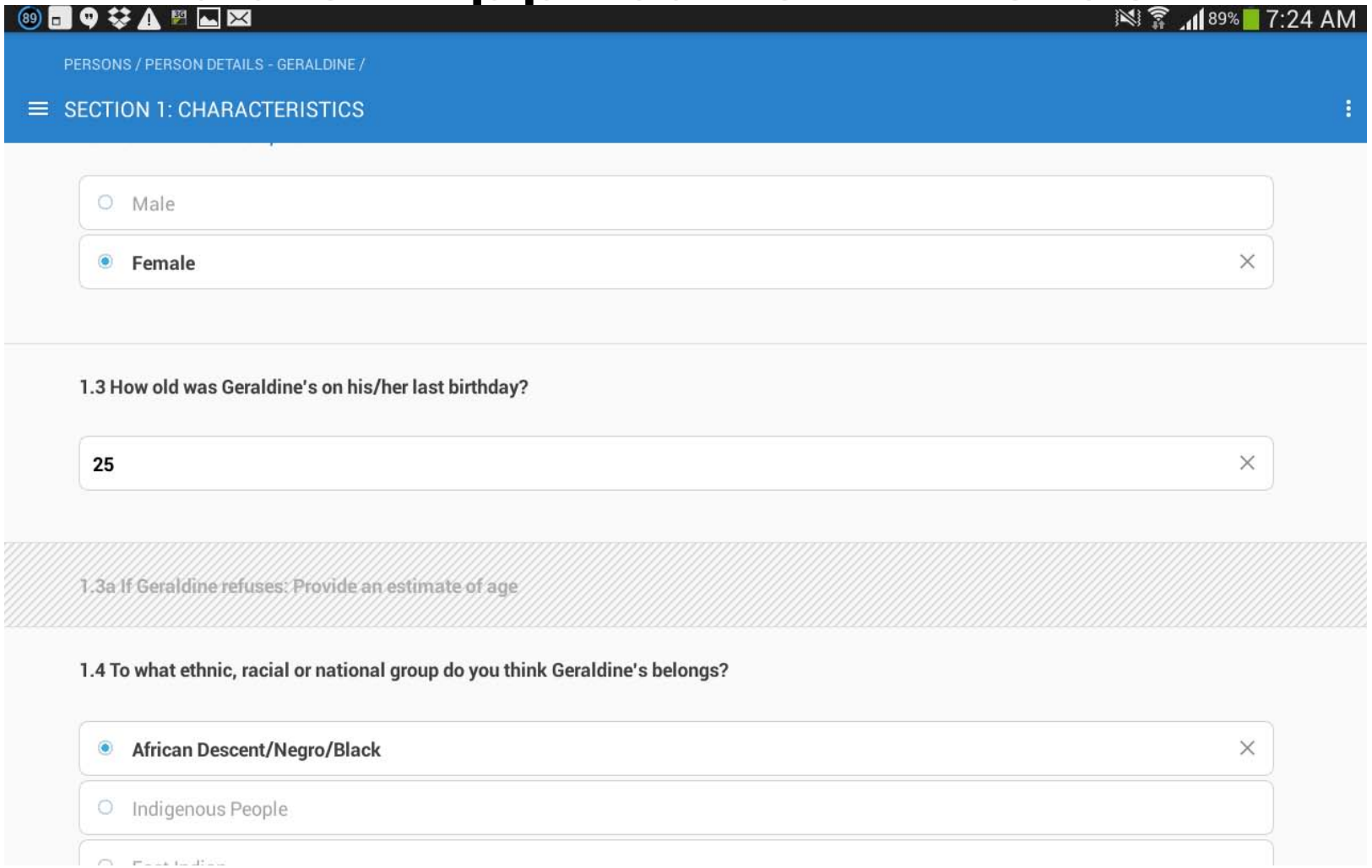
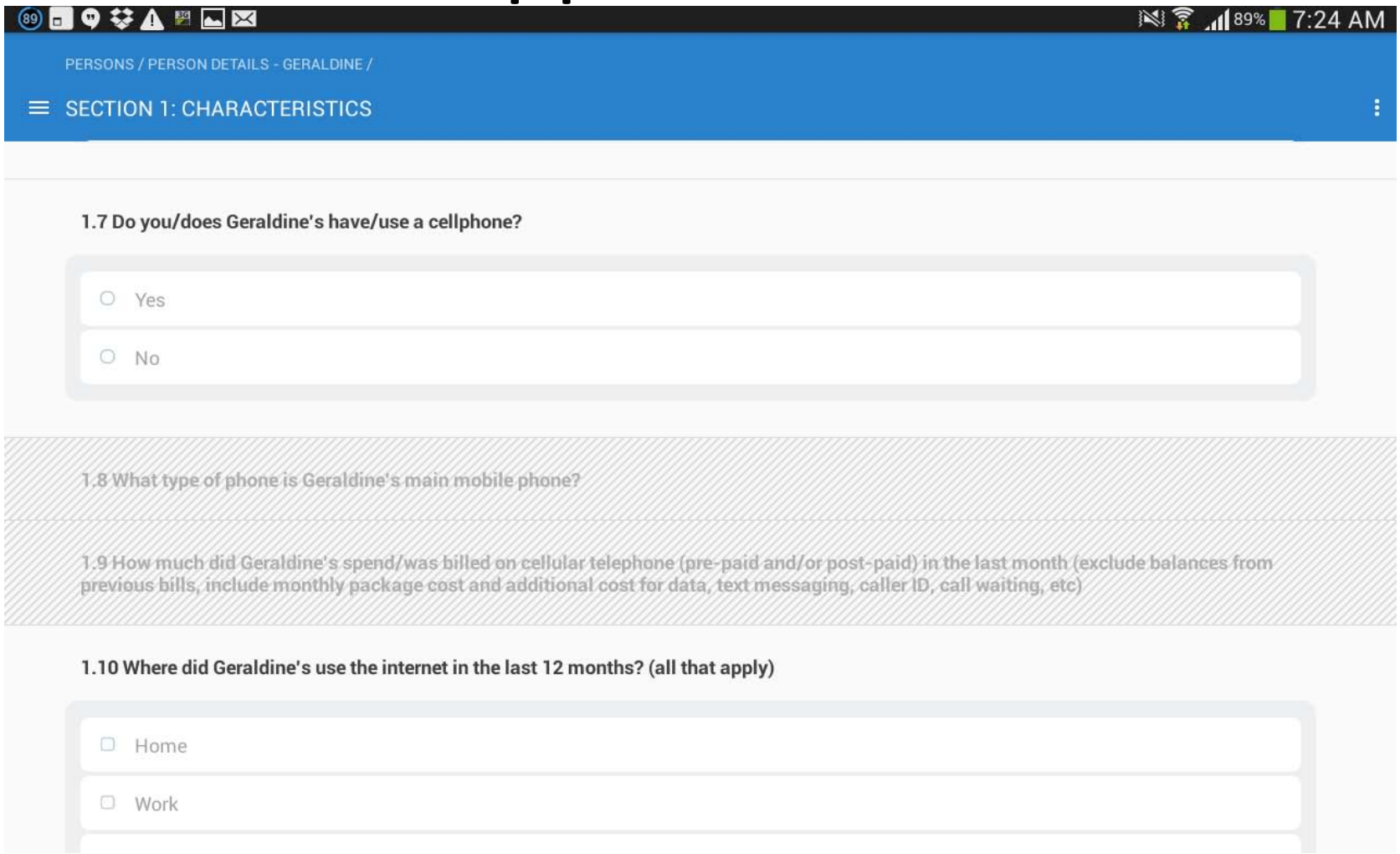
Tablet Application - Housing



Tablet Application - Housing



Tablet Application - Person



Tablet Application - Person