

# 4 Getting the Data We Need

---

# Table of Contents of the Module

<b>4</b>	<b>Getting the Data We Need</b>	<b>1</b>
<b>4 - 1</b>	Making decisions to get the data we need	7
<b>4 - 2</b>	Design a Bad Survey	10
<b>4 - 3</b>	Best Practices for Designing Surveys	13
<b>4 - 4</b>	Hands-on Review with External Data sets	15
<b>4 - 5</b>	Survey basics	18
<b>4 - 6</b>	Counting People - Handout	30
<b>4 - 7</b>	Household Survey Scenario	37
<b>4 - 8</b>	Using Spreadsheets Test	41
<b>4 - 9</b>	Mobile Data Collection and Data Protection	42

# Getting the Data We Need

---

The journey to getting the data you need requires a good start and planning. The purpose of this module is to guide teams or groups to ask the right questions at key moments in that journey. Collecting data yourself might seem easier, but there's lots of pitfalls, particularly if you go down the path of surveying. Identifying an external dataset might contain the data you need, but it was likely collected for a different purpose.

## Questions this module explores

- 1 What are decisions that will get you the data you need?
- 2 What's the right approach to surveying?  
What do you need to know about an external dataset?

## Learning Objectives

By the end of the module, learners will:

- ▶ Understand how to get the data they need and avoid collecting the data they don't
- ▶ Know when they should get the data themselves by collecting it, or go find an external data set.
- ▶ How to use an external data set.
- ▶ Be able to design surveys and analyse their effectiveness.

## Module Topics

- ▶ Planning to get your data
- ▶ Designing an effective survey
- ▶ Identifying and evaluating an external dataset.

## Recipes

### A suggested step-by-step process to achieve learning objectives

- ▶ Start by getting the team or your participants to review **Making decisions to get the data we need (4 - 1)** worksheet.
- ▶ If they've identified surveying as a way to get the data they need, get them to
  - a) **Design a Bad Survey (4 - 2)**
  - b) then compile **Best Practices for Designing Surveys (4 - 3)**
  - c) Design a survey with the **Household Survey Scenario (4 - 7)**
- ▶ If they've identified external dataset as a way to get the data they need, use **Hands-on Review with External Data sets (4 - 4)**
- ▶ Use **Generating a Data Quality Checklist (5 - 5)** as a final step

## Ingredients

Pick and choose ingredients to create your own recipe. Do you have an ingredient we're missing? Send an email to [data.literacy@ifrc.org](mailto:data.literacy@ifrc.org)

## Exercises

### Short, discrete social learning experiences

- ▶ **Making decisions to get the data we need (4 - 1)**
- ▶ **Design a Bad Survey (4 - 2)**
- ▶ **Best Practices for Designing Surveys (4 - 3)**
- ▶ **Hands-on Review with External Data sets (4 - 4)**
- ▶ **Household Survey Scenario (4 - 7)**
- ▶ **Using Spreadsheets Test (4 - 8)**

## Slide Decks

### Presentations to be used and/or adapted:

**Survey basics (4 - 5)**

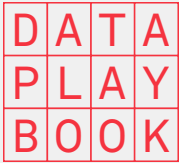
## Checklists/Handouts/Materials

### For documentation of essential elements of the learning experience

**Counting People - Handout (4 - 6)**

**Mobile Data Collection and Data Protection (4 - 9)**

## Next Steps



---

## Relevant modules in the Data Playbook

**Making Data Useful, Useable and Shareable (5)**

### Credit

---

Dirk Slater, IFRC V1 Sprint and Data Playbook Beta contributors

# 4 - 1 Making decisions to get the data we need

---

**This exercise will help a team understand how to get the data they need. It will work with any group that is working together to make decisions on how to get data.**

- ▶ **People:** 2 to five people
- ▶ **Time:** This exercise could take anywhere from 30 Minutes to three months depending on progress in their data project
- ▶ **Difficulty:** Medium
- ▶ **Virtual Materials:** virtual meeting platform, shared document/writing space
- ▶ **In Person materials:** Flipcharts/noteboards, sticky notes, markers

## Exercise

Keep asking the team these **10 questions** until you get the data you need!

	Questions	No	Yes
1	Do we know the question we want the data to answer?	No, Stop, go get clarity on the question and then resume.	Yes - go to question 2
2	Does the data we need exist somewhere?	No, go to question 3	Yes, go to question 8
3	Do we need to get it from people?	No, go to question 7	Yes, go to question 4
4	Are people happy to fill out a survey?	No, go to question 7	Yes, Are you sure? Have you surveyed them recently? Please ask question 4 again before going to question 5
5	Is there an existing survey that we could use or modify easily?	No, then go to question 5	Yes, Get that survey and ask are we sure we aren't asking too many questions?
6	Do we know how to design a survey?	No, review these exercises on survey design <b>Design a Bad Survey (4 - 2), Survey basics (4 - 5), and Best Practices for Designing Surveys (4 - 3)</b>	Yes, design and conduct the survey and go to question 10
7	What are the other sources for data? Consider satellite imagery, maps, case studies, reports, etc. After you've found it go to question 9.		
8	Is it in a repository like HDX?	No, go to question 7	Yes, the try <b>Hands-on Review with External Data sets (4 - 4)</b> and go to question 9
9	Does the dataset contain all the data we need to answer the question?	No, Are there other databases that might complement & answer the question?	Yes, go to question 10



	Questions	No	Yes
10	<p>Is the data we have gotten adequate and of quality? Does it actually do the job? Also see: Generating a Data Quality Checklist (Module 5)</p> <ul style="list-style-type: none"> <li>◦ Is it Accurate? No, go back to 2 Yes, then go to b</li> <li>◦ Is it complete? No, go back to 2 Yes, then go to c</li> <li>◦ Is it Consistent? No, go back to 2 Yes, then go to d</li> <li>◦ Is it Valid? No, go back to 2 Yes, then go to e</li> <li>◦ Is it Timely? No, go back to 2 Yes, <b>then you got the data you need!</b></li> </ul>		

Credit

Dirk Slater, Miki Tsukamoto, Eero Sario and Mununuri Musori

## 4 - 2 Design a Bad Survey

---

**This exercise is for any team or group of individuals that might be undertaking surveying as a way to get the data they need.**

- ▶ **People:** 4 to 16 people
- ▶ **Time:** 60 Minutes
- ▶ **Difficulty:** Easy
- ▶ **Virtual Materials:** virtual meeting platform, shared document/writing space
- ▶ **In Person materials:** Flipcharts/noteboards, sticky notes, markers

## Exercise

---

### Part 1: Explore (5 minutes)

Start the exercise by breaking the participants up into pairs or groups of three. Instruct them to share with each other their experiences of being surveyed over the last year (or recent time frame). What did they encounter in filling out a survey that they particularly didn't like? Ask them to list the things they didn't like out as 'elements' for their survey design.

---

### Part 2: Discuss (15 minutes)

After they have had 15 minutes to discuss and list elements - ask them to exchange and share elements with the other groups.

---

### Part 3: Design (20 minutes)

Instruct them to design a survey that will frustrate and annoy the respondents. They can choose their own elements or use the elements created in the other groups.

**Give them 20 minutes.**

At the end of 20 minutes, they can exchange their survey with other groups - ask them to rate which one they think is the worst survey.

---

### Part 4: Review (5 minutes)

To wrap up the exercise, ask them to review the elements they've identified. Ask: which elements could you improve and use in the next survey you design?

---

### Part 5: Conclude. (5 minutes)

Remind everyone that survey fatigue is a thing and ask: Do we really need a survey? Is there a more effective way to get the data we need?

## Extra Credit

Use with **Best Practices for Designing Surveys (4 - 3)** and **Household Survey Scenario (4 - 7)**

### Credit

---

This exercise was adapted from a concept developed by Mor Rubinstein. Contributors: Dirk Slater, Miki Tsukamoto, Eero Sario and Mununuri Musori.

## 4 - 3 Best Practices for Designing Surveys

---

**This exercise is for any group or team who are wanting to design surveys to get the data they need.**

- ▶ **People:** 4 to 16 people
- ▶ **Time:** 45 Minutes
- ▶ **Difficulty:** Easy
- ▶ **Virtual Materials:** virtual meeting platform, shared document/writing space
- ▶ **In-Person materials:** Flipcharts/noteboards, sticky notes, markers

## Exercise

---

### Part 1: Explore (5 minutes)

Break people into pairs or groups of three and get them to reflect on an experience they have recently had being surveyed. Were they good or bad experiences? Did they find the survey beneficial in any way? Was it informative in any way?

---

### Part 2: Review (25 minutes)

After they have completed conversations in their small groups, bring them back to plenary and review the **Survey basics (4 - 5)** Slide Deck.

Then get people to compile best practices, if online - have people enter their responses into a shared document. If in-person, have people write best practices on stickie notes and put them on a wall. Once they have finished compiling best practices, ask them to identify topics and categories that the best practices could be divided into and move accordingly.

---

### Part 3: Discuss (15 minutes)

After people have compiled the best practices, give them a few minutes to read them and then ask: How will you change your approach to survey design after reading what has been compiled?

---

### Extra Credit

Design a survey using the **Household Survey Scenario (4 - 7)**

### Credit

---

Based on a concept developed by Mor Rubinstein and designed by Dirk Slater

## 4 - 4 Hands-on Review with External Data sets

This exercise would be for a team who needs to understand if an external data set has the data they need or for training a group of individuals who want to learn about how to identify if an external data set is up to the task. A prerequisite for this exercise might be **Making decisions to get the data we need (4 - 1)** also in this module.

- ▶ **People:** 4 to 16 people
- ▶ **Time:** 45 Minutes
- ▶ **Difficulty:** Intermediate
- ▶ **Virtual Materials:** virtual meeting platform, shared document/writing space
- ▶ **In-Person materials:** Flipcharts/noteboards, sticky notes, markers

## Exercise

### Preparation:

If there is low or no access to the internet, facilitators should download a few example datasets for participants to use. There are datasets available on IFRC FDRS, IFRC Go, [HDX](#) (UN OCHA), and other data portals.

### Part 1: Explore (5 minutes)

What is the question we need to answer? Do we really understand the question and who is asking? Why are they asking this question?

### Part 2: Define (25 minutes)

Participants will require internet access for this step. Alternatively, the facilitator can download a few 'example' datasets on key topics such as climate, health, migration and/or crisis and disaster to use.

There are datasets available on IFRC FDRS, IFRC Go, HDX (UN OCHA), and other data portals.

### Part 3: Review (10 minutes)

Explore the dataset and ask the following questions:

- ▶ What was the purpose of this dataset? Is it aligned with our own?
- ▶ In terms of the question – is the dataset:
  - Accurate?
  - Complete?
  - Consistent?
  - Valid?
  - Timely?

If you've responded 'no' to any of the above then ask if there's another dataset that might do it better or be used to compliment this one. Or, do you need to build your own dataset?





## Part 4: Discuss (5 minutes)

Ask the group to reflect on what they need to consider or do before they identify and use an external dataset.

---

### Extra Credit

It might be helpful to use the **Making decisions to get the data we need (4 - 1)** exercise in conjunction with **Making Data Useful, Useable and Shareable (5)**.

### Credit

---

Dirk Slater

# 4 - 5 Survey basics

## Survey says...

---

### **Tell us about the last survey you participated in.**

- ▶ Was it what you expected?
- ▶ Was your privacy respected?
- ▶ Did you receive the results?
- ▶ What were the next steps? What was the impact?

## Surveys are:

A survey is a method of gathering information from a sample of individuals. Surveys are done to inform planning and decision-making.

There are many types of surveys. For example, a survey can include questionnaires or interviews to measure characteristics and/or attitudes of people. They can include qualitative and quantitative data.

# Planning Considerations

- ▶ Resources and appropriate timelines available to collect, analyse and use the data appropriately.
- ▶ Plan for processing the data
- ▶ Types of tools used to collect the data. Consider accessibility, social distancing, and other factors: online/offline survey?
- ▶ Is the methodology appropriate? Ethical? Data Privacy by Design? Minimization?
- ▶ Is the content appropriate?
- ▶ Are participants engaged in planning?

# Community/Audiences Considerations

- ▶ Audience/Planning outputs defined?
- ▶ Key informants engaged?
- ▶ Testing and Outreach planning?
- ▶ Consent obtained for fair use?
- ▶ Vulnerability assessed/Survey Literacy?
- ▶ Participants have access to response?
- ▶ Saturation/Fatigue/Duplication?
- ▶ Feedback loops for communities included?
- ▶ Is in-person a responsible way?



# Example of a Survey Design Plan (1)

Survey Timeframe			
Survey Management Team (SMT):			
Survey Team:			
Timing	Activity	Responsibility	Deliverable
1 week	<ul style="list-style-type: none"> <li>Appoint a survey management team</li> <li>Confirm budget available for survey.</li> <li>Review and finalize TOR</li> <li>Create dropbox for sharing background documents with survey team.</li> <li>Prepare draft timeline for survey process.</li> </ul>		Survey management team selected Approved budget available for survey ToR finalized and approved Dropbox created Draft Timeline
2 weeks	<ul style="list-style-type: none"> <li>Advertise ToR on several listservs (ALNAP, XCeval, MandE) to solicit applications for the survey team positions.</li> </ul>		Advertised ToR
1 week	<ul style="list-style-type: none"> <li>Shortlist/interview candidates and contracting of survey team</li> </ul>		Survey team selected and hired
1 week	<ul style="list-style-type: none"> <li>Survey consultants prepare inception report: include the proposed methodologies, data collection and reporting plans with draft data collection tools such as interview guides, the allocation of roles and responsibilities within the team, a timeframe with firm dates for deliverables, and the travel and logistical arrangements for the team.</li> </ul>		Inception report

# Example of a Survey Design Plan (2)

1 week	<ul style="list-style-type: none"> <li>Data collection</li> </ul>		Data collected via method most appropriate for field context Data monitoring and cleaning
2 weeks	<ul style="list-style-type: none"> <li>Analysis and draft report</li> </ul>		Draft report
1 week	<ul style="list-style-type: none"> <li>Debriefing with relevant stakeholders to check accuracy of data and gather further feedback from participants</li> </ul>		Input from relevant stakeholders
1 week	<ul style="list-style-type: none"> <li>Relevant stakeholders review draft report</li> </ul>		Any feedback sent on draft report
1 week	<ul style="list-style-type: none"> <li>Survey consultants incorporate comments and prepare final report <i>(this marks the end of the work for the survey team)</i></li> </ul>		Final report
1 week	<ul style="list-style-type: none"> <li>Final report is approved by relevant Management. If an IFRC secretariat-supported survey, it should be published on the IFRC Evaluation database. Go to: <a href="http://www.ifrc.org/en/publications-and-reports/evaluations/">http://www.ifrc.org/en/publications-and-reports/evaluations/</a></li> </ul>		Approved final report and payment of survey consultants



# Data Protection, Ethics, and Privacy by Design

- ▶ Rights of participants
- ▶ Protection from harm
- ▶ Collaborating with respondents
- ▶ Interviews guided by ethical responsibilities

# Engaging Participants

Community engagement should be inclusive. Organisations are accountable and transparent.

## Context

---

- ▶ Explain the purpose of the survey, length and roles
- ▶ Provide details on the organisation
- ▶ Explain how the information will be used

## Informed Consent

---

- ▶ Identify and discuss privacy considerations and risks
- ▶ Obtain informed consent before proceeding
- ▶ Participation is voluntary, questions may be optional

# Data Protection

Every project should be built with principles of privacy by design and protection from harm.

- ▶ Plan to collect only the data necessary to deliver the project goals.
- ▶ Curate Aggregated responses. Be aware that smaller samples in some contexts can be 'disaggregated' potentially causing harm.
- ▶ Be clear who has access to the data with clear roles, responsibilities, and procedures.
- ▶ All people engaged in the project are committed to confidentiality of participants and ensuring protection from harm.

# Be Ethical

Often people designing the survey may not be the interviewers:

- ▶ Plan for privacy and safety of the participant (respondents)
- ▶ Plan for the safety of the interviewers
- ▶ Train the interviewers on the best practices

Data managers may not be the same people who designed or conducted the survey:

- ▶ Instill privacy by design and responsible data use into all the information workflows.
- ▶ Provide context to inform the analysis and outputs.

(i) Note: especially during a pandemic, part of ethics is the decision if it is responsible to conduct the survey in-person.

# Thank you

Credit: Dirk Slater, Mununori Musori, Heather Leson, and Olaf Steenbergen

# 4 - 6 Counting People - Handout

## Summary

IFRC's data teams use best practices for counting people. From the recent COVID-19 response reporting to operational data to regular [Federation-wide Databank and Reporting System \(FDRS\)](#), it is a priority to standardize and adhere to best practices around counting people. This is essential for all reports and data analysis conducted by the FDRS-team and all others using FDRS-data. We have an IFRC document: "Technical Note: Counting People Reached". The following is an excerpt checklist. For more details, please see the [FDRS website](#) for further insights.



One of the main indicators for humanitarian action is “how many people were reached?” IFRC and National Societies report this annually for all our programs and activities.

## How can we count people reached?

This checklist handout will help participants consider how to incorporate best practices. Please fill this out based on the details you might have. Answer all the questions with a ‘privacy by design’ viewpoint.

## Things to consider

- ▶ Counting people reached is one type of measurement at IFRC. We use both quantitative and qualitative data to inform our work. This document pertains to improving quantitative data.
- ▶ This is a complex topic, but the excerpt should help inform how we might achieve data quality with a data standard framework.
- ▶ The Counting People Technical Note is used by the Monitoring & Evaluation practitioners at IFRC. To obtain this you can download a copy through the [Federation-wide Databank and Reporting System](#).

# Checklist

## Organisational Considerations for counting people reached

Item	Overall considerations at the National Society level for counting and reporting on people reached include:	Notes
1	Determine the overall purpose and scope of the people-reached reporting - at the project, program, country, regional or global level?	
2	At a minimum, ensure that counting and reporting on "people reached" meets the minimum National Society reporting requirements for FDRS.	
3	Establish data management systems that support systematic and reliable data collection and management that is responsible and realistic to the organisation. There is an increasing assortment of software solutions to support information management, from mobile data collection on handheld devices, ( e.g. ODK, Magpi/RAMP, KoBoToolbox) to organisation-wide, online management systems. At a minimum, Excel spreadsheets can be used.	
4	Promote understanding and commitment at all levels of data collection and management to support and sustain systematic and reliable counting of people reached. Tailor capacity and incentive building accordingly; in addition to sharing this technical note, consider learning opportunities ranging from in-person and online training to mentoring and direct technical assistance. Refer to the IFRC FDRS website and Regional PMER technical advisors for further guidance and resources. Incentive building also includes sharing and reporting back data to stakeholders to build understanding and an appreciation of its use.	



## Reporting on Multiple Projects/Programmes

Item	Key considerations for aggregating counts on people reached by multiple projects and programmes include:	Notes
5	Promote the use of systematic data entry forms/formats in the field that count people reached by service type, provider, delivery point and timeframe. This will support aggregating data at higher levels for reporting.	
6	Map the service delivery landscape, whether a city or whole country, to help identify and avoid potential double-counting. This typically involves a review of the project/program plans (frameworks) and consulting with managers to identify when certain target populations, services, or providers may overlap in time and place.	
7	Monitor data quality. Do not wait until preparing an annual report to discover that certain projects or programs did not count people reached properly, but be proactive and conduct field monitoring checks on data quality. Typically, such exercises will vary according to program area, and include quality assurance of processes that encompass more than people reached counts. Checklists Relating to Quality of Monitoring Information is a useful resource, but ultimately data quality monitoring will need to be tailored according to organisation and programme Area.	

## Counting and reporting at the project/programme level

Item	Key considerations for counting and reporting on people reached by a single project or programme include:	Notes
8.	Follow any specific reporting requirements and formats for people reached counts (i.e. see Box 2, above, on minimum reporting standards for people reached per the FDRS).	

Item	Key considerations for counting and reporting on people reached by a single project or programme include:	Notes
9	Counting and reporting of people reached should be planned as part of a coherent system to meet other project/program reporting needs. People reached is just one of an assortment of monitoring data needed for management decision-making and accountability. See the IFRC Project/Programme M&E Guide, Section 2.4 (p. 57) on information reporting and utilization.	
10	Use data entry forms/formats that support systematic recording and aggregation of people reached counts by service type, provider, delivery point and timeframe. This will also help identify and avoid double counting.	
11	Ensure human, material and financial resources are adequate and realistic for people reached reporting.	
12	Know in advance how people reached counts will be analysed and used and adapt data collection forms accordingly – for instance, vulnerability and capacity assessments (VCAs), baseline studies, emergency plans of action, the FDRS, etc.	
13	Ensure that data collection should be culturally appropriate, with attention to data collection teams that are representative of the population, linguistically competent, gender-balanced, and aware of cultural norms and taboos.	
14	Ensure people collecting and managing data are trained and prepared with competencies for data ethics, standards and 'do no harm' principles, such as informed consent, data accuracy, privacy and security.	
15	Identify and plan for collection of people reached data according to service type, with particular attention to direct and indirect counts accordingly – for instance, vulnerability and capacity assessments (VCAs), baseline studies, emergency plans of action, the FDRS, etc.	

## Direct counts of people reached

Item	Direct Counts of people reached	Notes
16	When necessary, use a tracking system to uniquely identify each individual receiving a service so that at the end of the reporting period there are accurate lists of people reached (by name and/or ID number).	
17	When possible, disaggregate people reached counts by sex, age, disability and any other relevant sociodemographic characteristics to inform analysis for effective service delivery.	
18	Determine whether using the average household size is necessary and advisable for counting people reached, with special consideration to inherent limitations in accuracy and disaggregated people reached data. If counting individuals in some instances AND households in other instances, be sure the counting does NOT overlap the different counting strategies.	
19	Determine whether catchment counts are reliable and accurate for counting people reached. Catchment counts can be used when the target population is likely to receive at least one service during the given time within the service delivery area.	

## Indirect counts of people reached

Item	Indirect counts of people reached	Notes
20	Counting people indirectly reached is inherently limited in accuracy and detail. Therefore...	

Item	Indirect counts of people reached	Notes
21	Carefully determine data sources to estimate indirect recipients with attention to reliability and credibility of counts.	
22	Do not estimate counts of indirect recipients when they receive services from direct recipients, unless there is a structured mechanism to ensure reliability of this process.	
23	Do not estimate counts of indirect recipients when they are indirectly reached by Federation Network messages or learning from another indirect recipient.	
24	Do not extrapolate and estimate counts of a national population as people indirectly reached unless there is substantial justification.	
25	It is sometimes acceptable to use the average household size or catchment populations to help estimate counts of people indirectly reached.	
26	Typically, it is not possible to disaggregate people “indirectly reached” because the service provider is not present to record demographic differences. However, there are exceptions when it may be justifiable to conclude about overall demographic characteristics for certain service recipients.	

# 4 - 7 Household Survey Scenario

---

**This exercise should work for any audience that needs to understand survey design coordination.**

- ▶ **People:** 4 to 24 people
- ▶ **Time:** 60 Minutes
- ▶ **Difficulty:** Intermediate
- ▶ **Virtual Materials:** virtual meeting platform, shared document/writing space
- ▶ **In Person materials:** Flipcharts/noteboards, sticky notes, markers

## Exercise

### Part 1: Explore (10 minutes)

Surveys are part of our work to engage with communities. This example “household survey scenario” is an example for a workshop. In your small groups, review the background details and clarify any issues/questions. Participants may need to make some assumptions to do the scenario. Document these and continue.

## Background

- ▶ Jalistan is a country situated near the equator with a population of 11,103,100 people, made up of many ethnic, tribal and linguistic groups.
- ▶ It is ranked 112 out of 177 on the UNDP human development index. However, the status of the constituent indicators (life expectancy, health and income) of the HDI in southern Jalistan are much poorer than the national average. This is due to over 30 years of intermittent internal conflict.
- ▶ The internal conflict has meant there has been little possibility of the government providing basic services in the southern regions of Jalistan. Following a negotiated peace settlement 1 year ago, many people displaced by the conflict are returning to their villages which they left, in some cases many years ago.
- ▶ The government is now fairly stable and quite well-received, as the overall security situation is quite optimistic.
- ▶ While all assessed villages perform badly on all development indicators, the situation in the villages around Akeret is worse, being devoid of most basic services. In some villages there are very few sources of safe water and in all of them, the schools are run down and understaffed.
- ▶ Rainfall is slightly erratic, over the past 20 years there has only been 3 years with seriously low rainfall.

**Project Goal:** Reduce death and illness related to Water and Sanitation related diseases in the targeted communities in Southern Jalistan.

**Implementer:** This project is entirely designed and carried out by the Jalistan Red Crescent Society.

**Duration:** The project runs from 1 January 2022 to 31 December 2023.

**Survey Objectives:** To collect primary data on a number of indicators related to water, sanitation and hygiene in villages located in Southern Jalistan using mobile technology. The survey findings will be used for village selection and for setting up of a monitoring and evaluation framework on water, sanitation and hygiene practices.

**Survey Methodology:** A quantitative household survey of water, sanitation and hygiene knowledge, attitudes and practices will be carried out in 4 villages in Southern Jalistan using mobile technology. The household interviews will be carried out by volunteers from the Jalistan Red Crescent Society. According to the sampling methodology, up to 100 households will need to be surveyed in order to make significant conclusions from the general observations of the targeted communities. Twenty volunteers (20) from the Jalistan Red Crescent Society will be carrying out the survey. These volunteers will be paired and efforts made to maintain a gender balance. It is expected that the training of volunteers and data collection will be done over a six-day period.

## Part 2: Survey Design (20 minutes)

Design survey questions by combining a scenario with an outcome. Note how your questions change as you switch scenarios with an outcome.

### Scenario 1:

This household is composed of 1 man, his wife and his 3 children in Kijereuk.

▶ **Parents:**

- Husband: Age 29
- Wife: Age 23

▶ **Three children:**

- Girl: 6 months.
- Boy: Age 7
- Girl: Age 8

### Scenario 2:

This household is composed of 1 widow and her daughter in Adong.

- ▶ Widow: Age 65
- ▶ Daughter: Age 35

**Goal: Reduce death and illness related to Water and Sanitation related diseases in the targeted communities**

**Outcome 1: Improved access to and use of sustainable sources of safe water in target communities**

**1.1** Community water points constructed or rehabilitated

**1.2** Community management of water points is improved

**Outcome 2: Improved access to and use of sustainable sanitation facilities among targeted communities**

**2.1** Sanitation facilities constructed

**2.2** Sanitation facility use is promoted

**2.3** Community management of sanitation facilities is improved

**Outcome 3: Enhanced practice of safe hygiene and sanitation in the household**

**3.1** Household knowledge increased on safe hygiene & sanitation

**3.2** Household training on safe hygiene and sanitation provided

## Part 3: Discuss (15 minutes)

In plenary, teams should share their survey plans and share any lessons/observations or insights.

## Extra Credit

Survey Basics} Slide Deck

### Credit

IFRC PMER Handbook



# 4 - 8 Using Spreadsheets Test

---

 [Download resources](#)

# 4 - 9 Mobile Data Collection and Data Protection

## Summary

---

**Data Protection is important for our work. This handout includes overall questions on how these will affect data and information workflows. This is a draft of key recommendations and some basic research. Please edit.**

# Recommendations

---

- 1 Review and update Standard Operating Procedures: By writing out the Standard Operating Procedures for ODK, Kobo etc, we will be closer to meeting the new guidelines. A policy is not enough anyway. We also need a proper workflow analysis (gaps, risks etc) to back-up the SIMs needs.
  - Every data set should have a 'version control and handoff procedure' (light weight). This will offset 'reuse of dataset' inquiry.
- 2 Training and Data Protection Guidance: It might be helpful to have a shared training or guidance document for all data and information workflows on data protection, even if we all work in different countries. This will show preparedness. Eg. the UK office gets only aggregated data via excel (email) spreadsheet from x deployment.
  - Also include guidance on de-identification of personal data, using pseudonymisation (masking) or anonymisation (aggregation, conversion, etc) of dataset. Other examples: Still images of an individual or community, video footage of an individual or community, DNA samples of an individual or community, and social security numbers.

Risks	Priority rating/Mitigation	Background link(s)	notes
Identify clear data governance: who is responsible for collecting, storing, processing and releasing personal data in the organisation?			
Tech setup -security, hosting, storage			
Lack of mapped technical workflow(s) to meet RESPONSIBLE DATA USE guidelines			
Data collection processes - consent, data minimization/mvp data set			
Collected data contains personally identifiable information (PII)			
Collected data contains Demographically Identifiable Information (DII)			
Missing business/ legal analysis on risks and preparation for RESPONSIBLE DATA USE			
Data controller does not manage this process			
Analysis on proprietary tool includes full data set			
Processing guidelines			
Processing by NS			
Processing by IFRC			

Risks	Priority rating/Mitigation	Background link(s)	notes
Processing by outside actor and out outside tool (academics/businesses)			
Transferring/Sharing - internal (EU)			
Transferring/Sharing - internal (Non-EU)			
Transferring/Sharing - external			
Inability to reach most vulnerable/do data and information workflows by adhering to Responsible Data Use/Data Protection regulations			
Disaggregation/Re-aggregation			
Archiving			

## Research Background

### ODK Docs

For the most up-to-date information please see the Security and Privacy section of the ODK Docs.

### Risks

- 1 Outside of usage analytics (typically opt-out) and crash reports (typically required), ODK software does not transmit or communicate any information (e.g., survey data) back to ODK's maintainers. When we do gather data, we default to anonymous or aggregate methods.
- 2 The software we have written does not have any mechanisms that might allow us to access or control your devices or systems. There is always the possibility that hackers can discover and exploit deficiencies or bugs in our software or in 3rd-party libraries to access or control your devices or systems.
- 3 "Central" is the ODK server now. See the ODK Security Guidelines. As well, keep informed on all relevant digital and data security, privacy and data protection guidelines of your organisation (IFRC and/or National Society).

- 4 With all 3rd party hosting services, you should expect your data to be viewable by the support staff of the hosting service. Different services go to differing lengths to restrict access to, encrypt, and/or secure the data and communications within their data centres.
- 5 We gather aggregate user behavior through Google Analytics and gather crash logs through Google Firebase Crashlytics. We use secure HTTPS communication to transfer this data to ODK's maintainers. Users may disable analytics in the settings of the application. Crash logging cannot be disabled.

## Handbook on Data Protection (ICRC, chapter 11)

### Key risks

- ▶ need for clear guidance on processing by humanitarian organisations of information extracted from the messaging app
- ▶ lack of awareness about the types of data they process
- ▶ metadata could be accessed and analysed by third parties and used by them in ways detrimental to the vulnerable populations.

## Towards a Secure Framework for mhealth - case study in mobile data collections

Gejibo, Samson Hussien - '[Towards a Secure Framework for mHealth. A Case Study in Mobile Data Collection Systems](#)', [Doctoral Thesis, Bergen Research Open Archive, 05/11/2015](#)

"For this work, we collaborated with the open-source MDCS, openXdata and Open Data Kit (ODK)."

### KoBo

- (i) P. 31 case studies considering security and privacy in data collection, transfers and archives. Very clear explanations of risks and security workflow gaps in international development

- ▶ [KoboToolbox data security measures: Keeping your data safe](#)
- ▶ [IFRC KoBo data protection information](#)

## Overview

There are three points of vulnerability for KoBo data: (a) on the handset, (b) in transmission, and (c) on the server. These are mitigated as follows:

On the handset, the data is available in clear text until it is transmitted, so we recommend that the handset itself is encrypted (a standard Android feature) so that unless an authorised user unlocks it, the data cannot be obtained. Once transmitted, the data is typically deleted from the handset.

In transmission, KoBo uses TLS by default.

On the server, the data is as secure as the server itself and is heavily dependant on the server administration and safety protocols and processes. Some agencies, including IFRC, have chosen to set up their own server partly to guarantee the security of their data to a higher level than that provided by the freely accessible servers.

As an advanced feature, KoBo also supports end-to-end encryption of data payloads so that the data packages are encrypted with a public key when the form submission is finalised on a handset and can only be decrypted by a matching private key held on an individual's local computer. This data is fully encrypted on the handset, in transmission and on the server and needs to be downloaded from the server to a local environment before its local decryption (this renders inaccessible any of the server-provided reporting and analysis options).

In practice, we've found that these points have always been sufficient to alleviate any data protection concerns on a technical level. Also, the actual data protection vulnerabilities are rarely technical but far more often due to poor form design. So we also recommend that anyone deploying any form (whether using KoBo or any other tool) carefully consider whether there is any real need to collect personally identifiable information in the first place. Unless they are explicitly doing beneficiary tracking (in which case the most critical point of vulnerability is the server database where the full information about beneficiaries is brought together, which is not KoBo), there is rarely a need to do so.

## Access to data

As a starting point, only the person who has created a project in KoBo has access to its data.

The only people who can access your data are the server administrators. However, they do not ever view your data (see contractual clauses below)

You can grant permissions to your project's data to additional KoBo users should you wish to do so.

## Security reviews

The ICRC conducted a security review of IFRC's KoBo server in May 2021 and found it secure for their purposes.

As a standard practice for externally hosted servers, the IFRC does not conduct penetration testing. However, given the wide use of the IFRC KoBo server and the nature of the data hosted on it, we would be happy to accept offers for penetration testing.

## Data protection contractual clauses between IFRC and KoBo Inc.

- ▶ Data shall only be stored, processed, sub-processed, backed-up, cached or otherwise hosted on servers in locations approved by the organisation.
- ▶ Under no circumstances shall US servers be used.
- ▶ The service provider shall at all times advise the organisation where data is held and processed, and consult with the organisation should a change in location be contemplated.
- ▶ The service provider shall involve sub-processors only with the consent of the organisation to both the entity concerned and the purpose of the sub-processing and comply with the contractual requirements even where data is processed by sub-processors.
- ▶ If the service provider receives a request for information of the organisation, it shall notify the organisation of such a request; in case of a non-disclosure order, the contract should require that the service provider asserts its contractual obligation to notify the organisation of a request for its information.
- ▶ Any information of the organisation processed by the service provider or any sub-processors remains the organisation's property and assets. Such information, including information about where data is stored or by whom it is processed, is confidential and the service provider must not disclose such information to any third party without prior written consent from the organisation. If the service provider receives a request for information from a State where organisation enjoys P&I, it shall explicitly assert organisation's privileges and immunities and state that data stored with the service provider constitutes property and assets that belong to the organisation and are subject to absolute immunity from search and seizure. Should the organisation not itself be in a position to do so, the service provider shall approach the relevant Ministry of Foreign Affairs, informing that Ministry of a request for information and asserting the organisation's P&I.



- ▶ In case of a request for information, and if privileges and immunities are not accepted by authorities, the CSP shall raise legal defences as instructed by the organisation. If the organisation cannot be notified of a request, the service provider shall raise all reasonable legal challenges available under the law applicable, including that of the State requesting information, to both the prohibition on notice and the legal demand to disclose the data.
- ▶ The contract is governed by Swiss law. The Parties shall not, under any circumstances refer to US law.

## Privacy Implementation Assessment template

[Towards a privacy impact assessment template for mobile health data collection systems](#)

Privacy IA template is typically structured in four parts:

- 1 Description of the application (i.e., Mobile Data Collection Systems - MDCS), in terms of its objective, requirements, users, stakeholders, application architecture, and data flows.
- 2 Identification of privacy threats with respect to a list of privacy targets embedded in the legal framework EU GDPR.
- 3 Proposal of technical and organisational controls for mitigating the identified privacy threats, i.e., counter measures
- 5 Documentation of the PIA regarding the MDCS being analysed

## Threat identification and selection of countermeasures

By looking into the MDCS context, a list of privacy threats can be postulated and associated to the aforementioned privacy targets. This threat identification analysis can be carried out by a group of experts, using brainstorm sessions and iterative reviews.

### Two preliminary examples of such threats are

#### Threat 1

User profiling (data minimization threat) The use of MDCS makes it fairly easy to link subject's data (i.e., patients or families), activities, kinship, demographics, and etc. User profiling is inherent in the health surveillance process. The further exploitation of data

relationships, creation of more complete profiles, might result in the use of personal data beyond the original purpose.

## Threat 2

Vague purposes (purpose biding threat): Vaguely defined purposes allow MDCS to be used for purposes not previously defined during the design stage. For example, do not follow the premises of meaningful use of medical data, and uses it for secondary purposes

## Why is data protection important to us?

- ▶ Humanitarian organisations collect and share more data than ever before
- ▶ Dignity and privacy of beneficiaries – privacy of home and family is recognized under international instruments
- ▶ Higher risk of harm to beneficiaries depending on their political, ethnic, religious affiliation or health status – the “do no harm” principle
- ▶ Comply or face risks of large fines up to 4% or 20M EUR
- ▶ Reputational and operational risks

## Data protection vs security vs responsibility..?

**“Data responsibility in humanitarian action is the safe, ethical and effective management of personal and non-personal data for operational response. It is a critical issue for the humanitarian system to address and the stakes are high.” [IASC](#)**

Legal Compliance	Privacy Policy	Terms and Conditions	Registry of processing	Lawyers and consultants
	DPIA	Internal policies	Notify DPA	

<b>Data Protection</b>	Consent Tracking	Audit Trail	Right to forget	You
	Identity	Breach Notification	Risk assessment	
	Data sharing policies	Data portability	Right to object	
<b>Data Security</b>	Authentication	Access Control	Secure transfert	IT and you
	Key Storage	Key Rotation	Search operation	
	Encryption at rest	Redundancy	API security	

## All you need to know...

- ▶ ...Is that personal data collected is accurate and relevant, and that the personal data is not misused, lost, corrupted, or improperly accessed and shared.
- ▶ Informed and freely given consent
- ▶ Gather only minimum amount of personal data
- ▶ Keep the data only as long as is needed
- ▶ Ensure access to personal data is controlled
- ▶ Use personal data only for its allowed purposes
- ▶ If in doubt, consult an expert!

## Refresher on GDPR

All EU NS are subject to GDPR, even when working outside the EU.

## Some key aspects of GDPR

Right to be informed (Ensure your users have been told what personal data you are collecting and how you will use it)

Right of access (If requested, provide a full copy of all their data within a reasonable time)

Right to rectification (If a user requests rectification, you must update/correct the data you hold about them without undue delay)

Right to erasure (to be forgotten) (When a user asks, you must delete all the data you hold on them without undue delay if there is no overarching reason to continue processing it. This is problematic if you are storing backups.)

## Sharing personal data outside your NS?

Anytime there is a contract, formal or otherwise where personal information will be shared, your NS should ask that the recipient of the personal data guarantee a minimum set of protections (GDPR Art. 28(3)):

- ▶ Only use the data according to your NS instructions
- ▶ Keep the data secure
- ▶ Assist your NS with any data protection requests you might receive
- ▶ Notify your NS of any breach immediately
- ▶ No subcontracting or data sharing without approval
- ▶ Delete and/or return data at the end of the contract

## Use Cases

(Answers provided by James De France)

## Case: Consent forms

**"We have refugees from different countries speaking different languages. Do we need to translate the consent forms to every language we encounter?"**

---

**“Yes, this would be ideal. If the person cannot understand why consent is being requested, then the consent is invalid, because it was not fully informed, nor freely given.”**

## Case: Device encryption

---

**“Should smartphones and laptops be encrypted in the field? Which resources are available to do these?”**

---

**“To the extent that it is possible, yes. Minimally, devices containing personal data should be password/fingerprint/passcode protected. Where possible, with the option to remotely wipe them if they are lost.”**

## Case: Cloud services

---

**“Can personal data be saved into a cloud service to which many people have access?”**

---

**“Yes. However, never make the link publicly accessible. Also, carefully review the terms of service for the cloud provider (run by legal). Only individuals who receive the link should be able to access the data, which should then also be password protected.”**

---

**“However, I would not use Google Docs to do this as Google has access to all data in its terms of service.”**

## Case: Sharing personal data

**"Sharing personal data of affected population with government? Other operational actors?"**

**"Should only happen after reviewing possible consequences and should be subject to a written agreement that restricts the use of the provided personal data. However, this also depends on any particular legal requirements in the country and whether the NS is the data controller."**

## Case: Data minimization

**"IFRC stance on data minimization? For example, when giving a kitchen set to a family, what data do you actually need?"**

**"This may depend on national law, donor or audit requirements. Only collect enough data to make sure that each family gets one kitchen set (if that is the objective)."**

### Resources

- ▶ ICRC [handbook](#) on data protection in humanitarian action (300+ pages)
- ▶ IFRC data protection [policy](#) (9 pages)
- ▶ IFRC data protection policy [guidance note](#) (2 pages)
- ▶ IFRC data protection [flyer](#) (2 pages)
- ▶ IFRC eLearning course [Targeted – An introduction to cyber security](#)
- ▶ IASC [operational guidance](#) on data responsibility in humanitarian action (37 pages)
- ▶ IFRC Data Playbook: [Responsible Data](#) (various resources)
- ▶ Data protection impact assessment (DPIA) [tool](#)

### Credit

Eero Sario, Dan Joseph, and James de France