Data Collector Workflows

ODK is often described as a data collection tool but more broadly, ODK tools and forms help encode workflows or sequences of steps to achieve goals. One of the outputs of these workflows is generally data. Data collection can either be the primary action (for example in agricultural field mapping) or a way to confirm that the primary action was completed and to assess its quality (for example in vaccination campaigns).

ODK provides rich functionality that can be mixed and matched without a prescribed structure. This approach allows for great flexibility but also places responsibility on the project designer to define a structure to support their work. The most important step of any project is to carefully consider its goals and the human workflows to achieve them. Then, ODK features can be selected to support that workflow.

This section presents common workflow patterns and the ODK functionality that can be used to support them. The techniques and ideas described can be mixed and matched to support your project's goals.

Single encounters (cross-sectional)

The simplest workflows involve collecting data at one point in time. Data is then submitted for reporting and analysis. These types of workflows would traditionally have been carried out with a paper form. For example:

- Spontaneous observations: e.g. anyone who visits a nature center is encouraged to note each time they see an animal
- Cross-sectional studies: e.g. sampling criteria are defined for stray dogs to be captured and tested for various diseases
- Periodic reports: e.g. during an election, one person per election site is responsible for sending periodic updates about compliance with protocols
- Discovery: e.g. one person is asked to capture information about all health facilities in their town

A single ODK form can take the place of a paper form and augment the workflow with required responses, questions that conditionally show and hide, data validation, and more. This replacement can be made with minimal or no training of field staff.

Note fields and hints are particularly helpful in single form workflows. They explain important parts of the workflow and help users remember their goals. They also prompt the user to perform actions outside of the form such as walking to a quiet part of the room, making sure that the respondent is comfortable, etc. These are all workflow steps that you should consider and design.

- Example of a form supporting a single encounter

Encounters with known entities

Many workflows are centered around a person, place or thing – a specific entity. It's common for those entities to already be known. For example, a municipality may have an existing list of health facilities and want to visit each one to capture structured data.

The known entities can be represented in a CSV file attached to a form as a dataset. In the form, a select one question can be used to identify which entity the current submission is about. The form can use any data known about the entity to show and hide certain questions, personalize the script, verify new data and more. If the entities have known locations, they can be shown on a map.

- Example entity list
- Example of a form supporting an encounter with a known entity

The entity list does *NOT* get updated automatically as entities are selected. Data collectors use `View Sent Form` or notes to keep track of work they have completed. Alternatively, you can help data collectors keep track of completed work by filtering the list based on some information such as where the data collector is located. You can also use repeats to represent entities and ensure each one is only selected once.

Assigning entities to data collectors

kind of assignment by adding a column to your entity list and filtering the list according to the data collector identity.

- Example entity list with assignment
- Example encounter form with assignment

Multiple encounters with the same entity (longitudinal)

Workflows can involve repeated encounters with the same entity at different points in time. For example:

- Longitudinal studies typically involve a baseline data capture about a subject, some intervention and then follow-up data capture at one or more additional points in time
- Field inspections involve visiting an entity after an intervention or periodically to complete a structured report
- Monitoring and evaluation (M&E) of various kinds of programs involves periodically capturing information about entities that the program is intended to have impact on
- Treatment of an illness in a person typically involves diagnosing the illness, taking some treatment action, and then periodically checking on the person until the illness is resolved

Many disciplines have terminology such as "case" or "incident" to describe a person, place, thing or situation that needs assessment followed by a series of coordinated interventions. In particular, case management is a term commonly used in public health. Examples of cases are:

- a person with HIV
- a person suffering from chronic food insecurity
- a dog with a guinea worm infection

Other related types of workflows that involve multiple encounters are issue tracking, incident management, customer relationship management.

When picking from the techniques described below, consider:

- how much time passes between encounters
- who will perform the different steps of the workflow
- whether information needs to be shared between different workflow steps

Multiple encounters in a single submission

Multiple encounters can be captured by a single form submission if the encounters can be carried out using a single device and happen close in time without a need for analysis between workflow steps.

For example, a form definition could have a section for doing initial assessment of a dog suspected to have guinea worm, another section for selecting a treatment and carrying it out, another section for follow-up and a final section for certifying that the worm has exited. This whole workflow could be carried out by a single person over a series of days and then submitted at the end.

In the example above, the person would start with a new blank form when they first hear about the suspected guinea worm case. They would fill out intake information and then close the form without finalizing it. When they are ready for the next step of the workflow, they would go to Edit Saved Forms and use the instance name to find the appropriate submission to continue filling out. They could then use the hierarchy view to find the section to fill out.

When designing these kinds of forms, thoughtful group names and instance names can help with navigation. Both can include status information that help data collectors take the appropriate next step.

- Example of a single form supporting multiple encounters

There are limitations to representing multiple encounter workflows with a single form:

- A single device must be used to perform the whole workflow.
- There’s no opportunity to clean, monitor or report on data between steps of the workflow.
- If the workflow needs to be interrupted, finding the correct filled form to edit and the correct question to jump to can be tedious (instance name and groups help with this).
- A single long form can be hard to test and troubleshoot.

Multiple encounters across different submissions

Many workflows can't be captured by a single form submission because:

- Different steps of the workflow are completed by different people and/or at different times
- Data is needed between steps for things like reports, monitoring, cleaning, assignment, eligibility determination

Workflows can be split across multiple submissions of the same or different forms. The submissions are linked during analysis by a common entity identifier. The identifier can also optionally be used to look up data about the selected entity.

The sections below describe some common tools and patterns for capturing multiple encounters across different submissions. As long as multiple submissions each have one field representing the same identifier, they can be linked no matter what their structure is so many variations are possible.

Using barcodes to link encounters

Barcodes are ideal for uniquely identifying entities because they can be consistently scanned without concern for data entry errors.

- Example of a form for an initial encounter
- Example of a form for a follow-up encounter

If barcodes aren't available or practical, you can use other kinds of unique identifiers such as phone numbers or names that will be entered manually. Be careful about identifiers that may not be unique, that may change, or that can be lost.

Once submissions are received, they can be linked through the unique identifier. For example, this can be done using Excel functions such as vlookup() or index() and match().

Using the sample forms above, part of the workflow happens outside of the forms: users are responsible for knowing when to use the registration form and when to use the follow-up form.

When designing a workflow, you have to make a choice between multiple smaller, simpler forms or fewer bigger forms with more logic. It's usually simpler to author and verify small forms but may require more training to ensure users pick the right one. Long forms with rich logic can be very useful for guiding users through a linear workflow but can be hard to verify and become hard to navigate if jumping around is necessary.

- Example of a single form that will be filled at each encounter

Using data from previous encounters

For many workflows, it's not sufficient to guarantee that records can be linked in analysis, it's also important to have access to some previously-known information at time of follow-up. In simple workflows, data flows one way: from data collectors to the server. More complex workflows often need two-way data flow: data collectors send data to the server and the server provides updated data back to them.

As described in the section on encounters with known entities, you can attach datasets to forms to represent known entities and information about them. These datasets may come from a registration workflow managed by an ODK form or from an external system.

If entities are not known ahead of time, a multiple encounter workflow starts with a registration phase (also known as enrollment, intake or discovery). This phase captures a unique identifier for each entity. It can also optionally capture some unchanging information about the entities and some baseline data. The data from the registration phase is cleaned and the columns that will be needed for follow-up are made into a CSV that will be attached to follow-up forms.

- Example CSV from registration form submissions

Follow-up forms are then the same as encounters with known entities.

If you have barcodes as described above, the follow-up form can use the barcode ID instead of a select one to look up the entity:

- Example of a form for a follow-up encounter with barcode lookup

There are many variations on baseline/follow-up workflows possible. For example, a user with a Zebra printer could generate a QR code with some entity information as part of registration.

- Example of a registration form that prints a barcode
- Example of a follow-up form that pulls information from the barcode

Multiple observations with the same data

Many workflows involve collecting general information such as data collector name, date, region, etc, and then collecting several observations or reports. This pattern can either be represented as one submission per observation or as a single submission with a repeat to represent all of the observations.

When using one submission per observation, last-saved can be used to reduce how much duplicated information is entered.

- Example of a form that uses last-saved to reduce entry of duplicated data

When using repeats, data that is the same for all observations is captured outside the repeat. Each observation is represented by one repetition.

- Example of a form that uses repeats to reduce entry of duplicated data

Advantages of repeats

- Less need to navigate around Collect: users swipe forward in a form and are prompted to capture a new observation.
- Logic can be used between repeat instances (this can be done with either an external app (e.g. counter) or last-saved but neither supports submission edits). For example:
  - Keep capturing plants until you have observed at least 10 of the same species
  - Sequentially number households

Disadvantages of repeats

- Form can get slow or hard to navigate as the number of repeat instances increases
- Data can't be downloaded or viewed until all observations are captured and the full form is sent
- The repeats and the top-level data will generally need to be joined in analysis (see the tip in the repeats section)

Location-centric workflows

Many workflows have a geospatial component to them. For example, a forestry workflow may involve periodically visiting and assessing specific old-growth trees.

The form submission map for a form can be used as the landing screen for location-centric workflows. It can help plan travel and display completed work.

Other tools for location-centric workflows include select one from map and offline layers. External apps can also be used for things like wayfinding and navigation.

- Example form that launches Organic Maps for navigation

Using Collect settings for workflow support

ODK Collect is highly configurable and different combinations of settings will better support some workflows than others. Once you design an ideal workflow for the tasks you need to accomplish, we recommend looking at the available settings and considering which could help support your workflow.

Consider what you want the user experience to be before entering your form, during form filling and after completing the tasks represented by the form. If you change default settings, we recommend you configure one device as desired and create a QR code to configure others.

Below are a few examples of settings that can help support common workflows.

Before form entry

When Collect connects to an ODK Central server, it exactly matches the form list from the Central project it connects to. Users will always see all of the forms available to them from Fill Blank Form.

For some workflows, it may be preferable to let users manually download the specific form(s) they need and then

delete it and perform another manual download of other forms once they enter a different phase of the workflow.

During form entry

Many workflows must be completed in a specific order. By default, users can use the hierarchy view to jump around within a given form, temporarily skipping over required values or going back to edit previously-entered values.

You can enforce a linear flow through a form using Form Entry Settings from the restricted settings. Linear flow gives the form designer more control than an interface that encourages jumping around. For example, a linear flow makes it possible to force double entry or verification of existing data without the possibility to edit it.

If you know your users have small devices, train them to swipe between questions and remove navigation buttons to reclaim vertical space.

After form entry

By default, form data, even after it has been submitted, stays on users' devices and can be viewed from `View Sent Form` as well as the submission map. This can act as a reference for users or help with troubleshooting issues. If you are collecting sensitive data or wish to save device storage space, you may want to enable the Delete after send setting.

You also have different options when users reach the end of your form. By default, forms are marked as finalized which means that they are ready for submission (manual or automatic) as soon as Collect can connect to the server. Users are also shown a checkbox to save without finalizing. You can change the default and/or hide this option. This can be useful for workflows that include a supervisor check before sending data, for example.