Technical Docs Review for ODK
25 November 2017

The ODK community has made a lot of progress improving user documentation, but Yaw wanted a professional, outside opinion on the current status of the new documentation set. This included hands-on testing of the Getting Started documentation and then reading every page in the new documentation set from a non-developer perspective to note anything that may be unclear, could be improved, or is missing.

Getting Started

To begin, I was asked to do the following:

> *https://docs.opendatakit.org is the new docs site and https://opendatakit.org/use is the old docs site that users are most familiar with. Use the new docs site to collect 1 text field, 1 picture, 1 geo coordinate in a single form on a mobile device running ODK Collect. Send that data back to ODK Aggregate and get that data out of ODK Aggregate into a CSV. Is this possible using only the new docs.*

This was a useful method of entering the review the same way a new ODK user would, by attempting a simple task to test the software system by following the documentation. I used the information in the Getting Started Guide at http://docs.opendatakit.org/getting-started/ for my attempt.

First, I installed ODK Collect on my Android phone. The process was quick and easy and well-documented. This took about 2 minutes.

Next, I installed ODK Aggregate locally on my workstation using the downloadable virtual machine. This was also quick and easy. I would like to suggest the following edit to the *readme.txt* file that accompanies the VM download. Excluding download time, this took about 3 minutes.

Step 6 currently reads:

> "6. Go to the web address shown in the VM boot screen. You will see the Aggregate login screen."

 I suggest it be edited to say this:

> "6. Go to the web address shown in the VM boot screen. You will see theAggregate main screen. Click the Log In link at the upper right to go to the Aggregate login screen."

I was not certain what the best way to create a form would be, but decided to start with ODK Build as it was recommended in the documentation as the quickest and easiest method.

I was unable to upload a form from ODK Build to Aggregate directly. I suspect this may be

because I was using a VM and accessing it via IP rather than a fully qualified domain name (FQDN), but I didn't bother to troubleshoot as other methods were mentioned in the Getting Started Guide.

I tried exporting from ODK Build to XLSForm. The export worked fine, but uploading into ODK Aggregate returned an error as this was not an accepted filetype. The filetype required is not mentioned in the docs.

I exported from ODK Build to XLS. The export and import into ODK Aggregate were both successful. Creating the survey took about 22 minutes.

Next, I worked on uploading the form to ODK Collect. This was fairly smooth and took about 5 minutes.

I was unclear what the specific server settings were that are required for the connection between ODK Collect on my mobile device and ODK Aggregate on my workstation. The settings are not detailed in the documentation and I used trial and error. The IP address was sufficient, but I did not know whether I needed to use admin credentials, create data collector credentials, or whether I could do this anonymously.

Completing data entry in the simple form in ODK Collect was smooth and only took a couple minutes.

Sending the finalized form to ODK Aggregate is where I ran into problems. While the connection between ODK Collect on the mobile device and ODK Aggregate on the workstation worked fine for uploading the form to the mobile device, I could not send the finalized form the other direction. Each attempt returned a 500 Internal Server Error.

I discussed this with Yaw and we decided that the progress to this point was sufficient for this documentation review, especially since 500 errors can be time-consuming to troubleshoot and the error is likely unrelated to the documentation being reviewed.

Overall, the information flow in the Getting Started Guide is good. Only the needed context is given and no more, preventing the reader who is new to the project from feeling overwhelmed. I like that extra detail is provided in Tip boxes, which most readers will presume to contain optional information which can be skipped, again preventing readers from feeling overwhelmed while also providing extra detail for the readers who seek it. The Tips are useful and clear. I like that the page starts with prerequisite knowledge and a high level view of the process, proceeds into clear and specific details, including an easy to understand description of multiple options when they exist, and ends with reference links for further reading and how to find help. Inline links throughout the page are useful and clear, but not distracting.

Other Docs

From here, I began a systematic review of every page of the new documentation. As a guide, I used the table of contents in the sidebar and read each page starting with the first listed and progressing down the list to the last listed. I did not test any of the procedures or other content from here on, but only read them for clarity and apparent completeness and accuracy. I looked for places where I perceived possible gaps (and found very few). Most of my comments surround

possible improvements to style and manner of expression and trying to enhance the clarity of the content.

Pages with specific comments are listed in this section of the report using a subheading that corresponds to the page title in the table of contents. Pages not mentioned in this section did not prompt specific comments during my review and are covered by the general comments that come later in my final summary. Also, pages reviewed earlier are likely to include comments that may apply to later pages. For the sake of brevity, those comments will not be repeated.

Welcome Page (not listed in ToC)

This is the top page of the documentation found at http://docs.opendatakit.org/. It is a good landing page. The content is short, clear, and gives clear signposts to guide readers who are seeking out specific content while also providing a brief introduction for the reader who does not yet know what they want.

ODK Collect

I like this overview page. The quick definition of the app is clear and useful. I especially like the inclusion of a typical pattern of use followed by the "you can also" and "along the way you might want to" sections. The links are clearly labeled and nothing is confusing. The page is informal without being cute or weird and without compromising professionalism.

Collect Installation and Setup

Including both methods of installation is good. Even better is that you begin with a clearly labeled recommended method that is quick and easy.

Notes, Tips, and Warnings are clearly labeled and include useful information. I like that the different types are used appropriately and give the reader precisely what is expected. I especially appreciate that the Warning gives both the "what" and the "why" without being verbose—that is not easy to do!

Headings are clear, consistent in language and style, and the heading hierarchy is obvious (for example, it is easy for the reader to determine that the multiple by-server-type connection methods are discrete options under the Connecting to a Server section).

It is my opinion that the procedures would be better served using ordered lists rather than unordered lists. Typically, unordered lists are best used for aggregated data surrounding a specific topic, such as "Prerequisites" or "Options" which do not have a ranking or required order. Procedural steps must be taken in a prescribed order; numbering them subconsciously communicates this to the reader. Numbering the steps also makes it easier to refer to them in comments or notes.

For options within steps (you can do this OR this), such as in Step 4 of the manual install procedure (In the download window...), a nested unordered list is appropriate within the procedure step.

I did not notice that entries in the table of contents expand to include headings and subheadings while on a page. That is a nice touch!

Collect Menus, Settings and Security

You need a comma after "Settings" in the title, to follow your Style Guide's stated custom of using the Oxford Comma. I wholeheartedly approve of the rule, by the way.

Good documentation generally includes three types of content:

- Contextual, which answers the "why" and gives the reader a broad understanding

- Procedural, which answers the "what" to give the reader guidance in accomplishing their tasks

- Reference, which gives data and information in greater detail

This documentation set has its content appropriately ordered, starting with contextual information, leading to procedural information, and scattering reference links where appropriate.

This page is the first example of pure reference material. It is precise, clear, detailed, and I could not find any content gaps. I like that General and Admin Settings each have their own section. The information is clearly organized and it is easy to find what I am looking for.

The writing on this page occasionally vacillates between typical American English usage and British English. For example, most of the documentation includes an article such as "an" or "the" when referring to a button or menu, as in "General Settings are accessible from the [icon] menu..." which is different from "You can import settings from a QR code saved on your device by clicking on **Select Code from SD Card** option" and "QR code is a JSON object..."

The way things are listed varies by section. In some places, you have an unordered list where each entry starts with the button or option followed immediately by a capitalized word and full sentence. In other places, you have an unordered list where each entry starts with a button or option followed by a colon, a lowercase word, and not a full sentence (but still ending with a period). Some lists mix these styles. Occasionally, entries do not fit either of these but use a style unique to that entry.

This page would benefit from an editorial read for consistency in the manner of expressing the otherwise clear information.

Form Management in Collect

It was in this page where I first noticed some style use inconsistency.

It is not clear to me why some text is in bold, such as "To **remove a response**, Long Press on the question label." Most documentation standardizes on using bold when referring to items the reader can see in the UI, such as "Press the **Any** key," however, you have the lovely button styling to serve this purpose.

It is not clear to me why some text is in italics, such as "Once you have reached the end of a form, you will have the opportunity to *Save* and *Exit* the form." Most documentation standardizes on using italics for terms that are being introduced to the reader for the first time or for denoting placeholder information that must be modified and entered by the reader, such as:

```
$ ls /home/username/appdir
```

Using Android Debug Bridge with Collect

The unordered list at the start of the page uses a style not yet seen in this documentation set. It is not a bad style, but it is not consistent with other unordered lists seen previously. The inconsistency is a bit distracting as I found myself wondering why it was being done differently and whether there is a style-based message being communicated that I need to know. This is a common issue in documentation sets written by multiple people, but one worth cleaning up.

The Note on this page is missing an article and missing punctuation.

Choosing a standard phrasing for introducing commands is common in documentation. What form you use is less important to the reader than whether it is used consistently. On this page, we find:

- Forms can be deleted by running:

- To download a completed form from the computer, run:

- Developers might also need to check the entries in the database from the computer. In such case pull the database file from the SD card and use any SQLite visualizer to view it. To pull the database into the computer, run:

- For taking a screenshot, run:

My personal preference is the second example, "To perform the task, run:"

When contextual information is useful before giving the command, as in the third example, I would separate the contextual information and command introduction into two paragraphs.

ODK Aggregate

I would love to see this page reorganized a bit to parallel the style and format of the ODK Collect page, with a bit more description, links, and a consistent format to the items in the list.

Aggregate Installation and Setup

The Tip boxes on this page have more information in them than on other pages. They feel long and complex. I suggest looking to see if some of the information can be moved out of the boxes. For example, the first Tip box does not really contain tips, but rather background information on options. Bullets in a box make the box feel cramped. The two bullets in the second box flow together well and have no need to be in list format.

It is interesting that your Linux desktop screen captures seem to have been taken using Ubuntu, but that you give install instructions that alternate between Apt and Yum, but never both side by side. For some things the reader is instructed to install using Apt commands, for others, Yum. Half of these install instructions will fail on Fedora and other rpm-based distributions and half will fail on Debian and other apt-based distributions.

Some screen captures include the browser frame in the image, some do not. At least one shows only half of the top bar of the browser, including bookmarks and the address entry box. Standardizing across the images will make the page look more professional and be less distracting to the reader.

Aggregate Deployment Planning

I like this page. One quibble: you are using an ordered list for Considerations. These are all equal in value and there is no specific order of steps being prescribed. I wonder if maybe the understanding of when to use each type of list got a little confused somewhere.

The short list at the bottom of the page is an exception, because it is enumerating the two modes of operation mentioned in the introduction to the list, so having items labeled 1 and 2 makes sense.

Aggregate Usage

One of the first things described on this page is the top of the Aggregate screen, mentioning three help buttons. I had no idea why there are three. I thought, "That's curious," and moved on. Quite a bit later in the page, in the Help Options section, the three buttons are each described in detail and I was pleased to discover each does something different. I suggest either moving the Help Options section up, so it is one of the first things a reader will encounter, or at least adding a comment and link down to the section from the initial mention of the three buttons.

Aggregate Limitations

This entire page is bullet points. I would change that from a list to paragraphs. The spacing between the paragraphs will be adequate to separate each limitation and it will look better.

ODK Briefcase

I would love to see this page reorganized a bit to parallel the style and format of the ODK Collect page, with a bit more description, links, and a consistent format to the items in the list.

Contributing

All of the pages in this section are clear, encouraging, and useful. Too few projects include this level of detail and clarity in their contributor's guides. Well done!

Developer Docs

I was amused by the mentions of Mufasa and the Circle of Life in the OpenRosa API docs. I don't know if the ODK team slipped that in or if it was inherited, but thanks for the Easter Egg.

Security and Privacy

This is the first time in the docs where I see a mention of the ODK software license. For me, licensing is important enough to deserve a brief mention on the welcome page. One sentence would be sufficient and would make open source proponents happy, such as "The ODK software is released under the Apache 2 License." with a link to this security and privacy page for further details.

Comparison of Old and New

When comparing the old documentation with the new, the only content I found that was not contained in both or did not have a clear "deprecated" notice was the information surrounding the 2.0 Tool Suite at https://opendatakit.org/use/2_0_tools/.

The new documentation is more clearly organized and therefore easier to follow. Much of the content is similar, so I'm going to guess there has been some copying and pasting during the transition and reorganization. That is common and acceptable. The next task will be to clean up the new documentation set to make it more consistent in the new style and more consistent in phrasing and manner of expression across the entire set.

Final Summary and Next Steps

Overall, my impression of the documentation is positive. I like the site style and theme as they make the content clear and easy to read without being distracting. The writing is consistently clear and precise. Short, direct sentences dominate and superfluous words are mostly absent. I found extremely few information gaps. From a purely technical standpoint, I believe the documentation is quite good.

The main recommendation I have is to spend some focused time standardizing the content across pages. Most pages are internally consistent, but there are content formats that change across pages. Here are two examples:

- Unordered lists are formatted in numerous ways across the site. Some suggestions: any time a list is used, it should be introduced with a sentence or a phrase, such as "The features of Briefcase include:". Never go straight from a heading to a list. If you feel the content is best left without an introduction, eliminate the bullets and use paragraph style.

- Headings are sometimes written in command form, "Install Collect," sometimes in gerund form, "Using (the) Command Line," sometimes in noun form, "Security and Privacy," and sometimes with abbreviations, "Intro to Forms in ODK." Pick one form and use it consistently for contextual and procedural pages, I suggest either the command or gerund form. The exception is for reference pages, which typically use a noun form, which is generally done here.

You use a lot of screen captures. This is useful for readers, but can be perilous, especially when including screen captures showing sources you do not control, such as AWS or GCE. I don't have any specific recommendations other than to remember that these must be audited regularly for accuracy. You may also want to consider whether some of these screen captures can be safely removed during these regular audits, as sometimes it is enough to use words, just to help reduce how much time is needed for writing maintenance. Obviously, I would only do this where you feel confident the loss of the screen capture will not negatively impact the reader.

To enhance your consistency, you should adopt a documentation style guide and have everyone use it. Where I work, we use the IBM Style Guide, which is print-only and not available online. Since ODK has such a close tie in with many Google services, perhaps the Google Developer Documentation Style Guide would be better. It is available for free at https://developers.google.com/style/. A good example of how the guide would help the ODK documentation is in the Procedures page at https://developers.google.com/style/procedures. As a side note, I did read your Style Guide at http://docs.opendatakit.org/style-guide/ before beginning my review. It is good, especially for the standards expressed, and covers many, but not all of the things I have mentioned in this report. I suggest keeping it while adding a sentence saying

something like, "Anything not covered in this Style Guide should then follow *styleguide_you_choose*."