OpenDataKit
magnifying human resources through technology

ABOUT   USE   DOWNLOAD   HELP   PARTICIPATE

Aggregate

ODK Aggregate provides a ready-to-deploy server and data repository to:

- provide blank forms to ODK Collect (or other OpenRosa clients),
- accept finalized forms (submissions) from ODK Collect and manage collected data,
- visualize the collected data using maps and simple graphs,
- export data (e.g., as CSV files for spreadsheets, or as KML files for Google Earth), and
- publish data to external systems (e.g., Google Spreadsheets or Google Fusion Tables).

ODK Aggregate can be deployed on Google's App Engine, enabling users to quickly get running without facing the complexities of setting up their own scalable web service. ODK Aggregate can also be deployed locally on a Tomcat server (or any servlet 2.5-compatible (or higher) web container) backed with a MySQL or PostgreSQL database server.

Installation¶

1. Please read all the instructions and notes before beginning.
2. Make sure you have used ODK Collect and are familiar with how it works.
3. Next, try the ODK Aggregate demo server to explore the core functionality.
4. Decide whether to install a cloud instance or a local instance. We strongly recommend you try an App Engine cloud instance first. If you wish to host locally, see Aggregate Deployment Planning.

Local hosting implies that you are taking ownership of the off-site back-up and restoration of your data and are documenting the steps necessary to return your systems to operation in circumstances that might include a full hardware failure or the destruction of your facility. You must also plan for the security of your data and systems. And finally, it requires that you configure your network routers (e.g., WiFi router). We would encourage you to seek assistance from your local computer-technical-support community before proceeding. The set-up of the ODK Aggregate web server and database are very easy in comparison.

The Importance of Upgrading¶

It is important to upgrade to newer ODK Aggregate versions as they come out (not necessarily immediately, but this should be something you do at least once a year).

There are several reasons for this:

- **Security vulnerabilities** -- we (and Google) are constantly upgrading the libraries we use with newer, safer, versions. The older your software, the greater the number of vulnerabilities in it.
- **Hosting revisions** -- Google AppEngine is a managed environment, unlike, say, AWS or other "bare-box" hosting services. Google is continuously updating features and removing support for older features in this environment. If you don't upgrade, there may not be an upgrade path that works due to these changes -- unlike "bare-box" hosting, you, and the ODK team, only have partial control over the software and hardware environment.
- **Performance revisions** -- as we find performance issues and address them, the tools get better and faster.
- **Enhanced capabilities** -- the form-processing library (javarosa) has roughly-annual updates to add new functions (e.g., sin(), cos()) and occasionally data types. And new features are slowly added to ODK Aggregate, too.

Installing on App Engine (Cloud)¶

1. Make sure Java 7 or higher is installed on the computer you plan to use. If it is not, download and install it; if that page does not have a link to older versions, Java 7 is also available here. If you are using MacOSX, it may require special care and attention. See MacOSX Java install and MacOSX Java FAQ.

2. You will need a Gmail account to use App Engine; this Gmail account will be the owner of the Google Cloud Platform "project" under which your App Engine will execute. Google Cloud Platform hosts projects under these terms. There is no cost to set up these projects and App Engine projects that are lightly used will incur no charges; when you exceed the free-usage levels that Google allows, your App Engine server will stop accepting new requests for the remainder of a 24-hour period that resets at midnight in the Los Angeles, CA timezone. See usage fees for ways to minimize usage.

3. To set up a Google Cloud Platform project, go to Google Cloud Platform and click on "Console" (this shows as "My Console" in this earlier image capture):
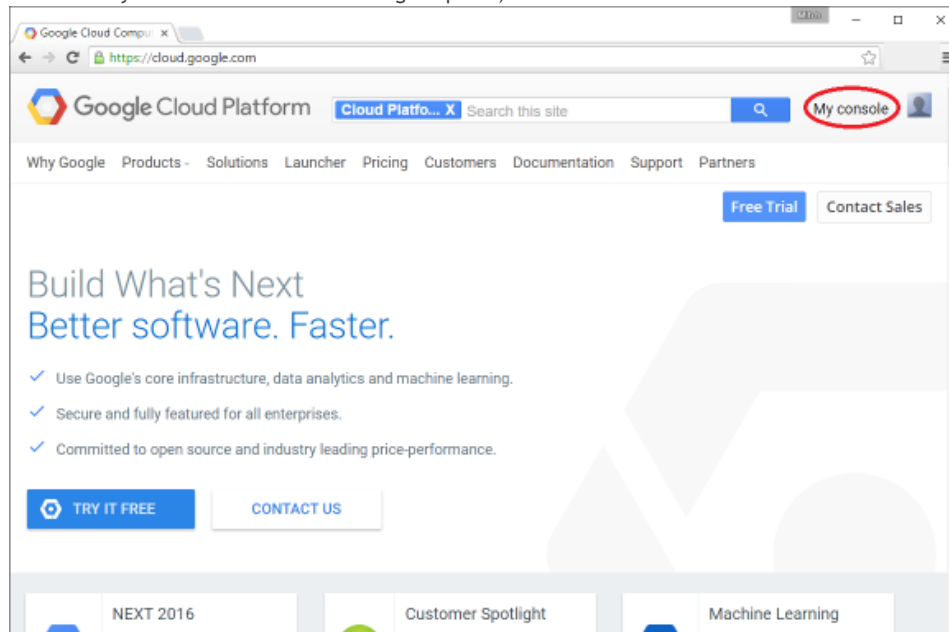

4. If you have never configured a Google Cloud Platform project, click on "Create an Empty Project":
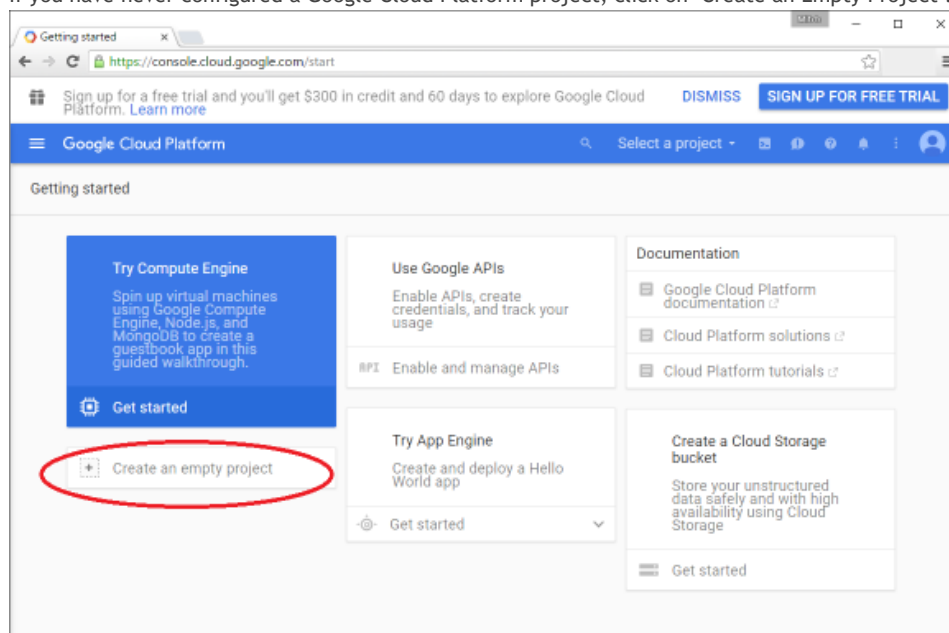

Otherwise, this link will open onto either a page with a "Create Project" button and a table listing all of your projects, or it will open into one of your existing projects. In the later case, click on that project name to access a drop-down menu and choose "Create Project" from there. E.g., if clicking "My Console" opened onto the project "msundt-agg-test3", you would click on that to access

a menu where you can chose "Create a project...":
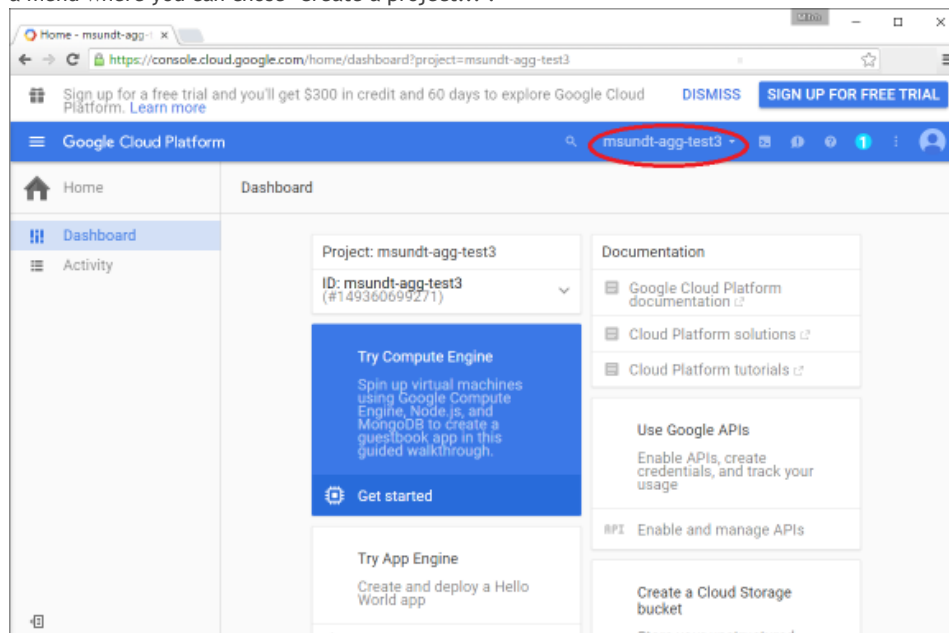
Due to site updates, this no longer appears on the upper right. It appears on the upper left, next to the words Google Cloud Platform.

5. On the project-creation pop-up dialog, type in a project name that makes sense to you. Then choose "Edit" to edit the project id. The project id will be the first part of the URL to your ODK Aggregate server (i.e., your server will be referenced with a URL of the form `https://project_id.appspot.com`). You may want to use a project id that combines your organization name and the name of your data collection group or project. You may also need to accept Google's terms-and-conditions, as indicated in the screen-shot below:
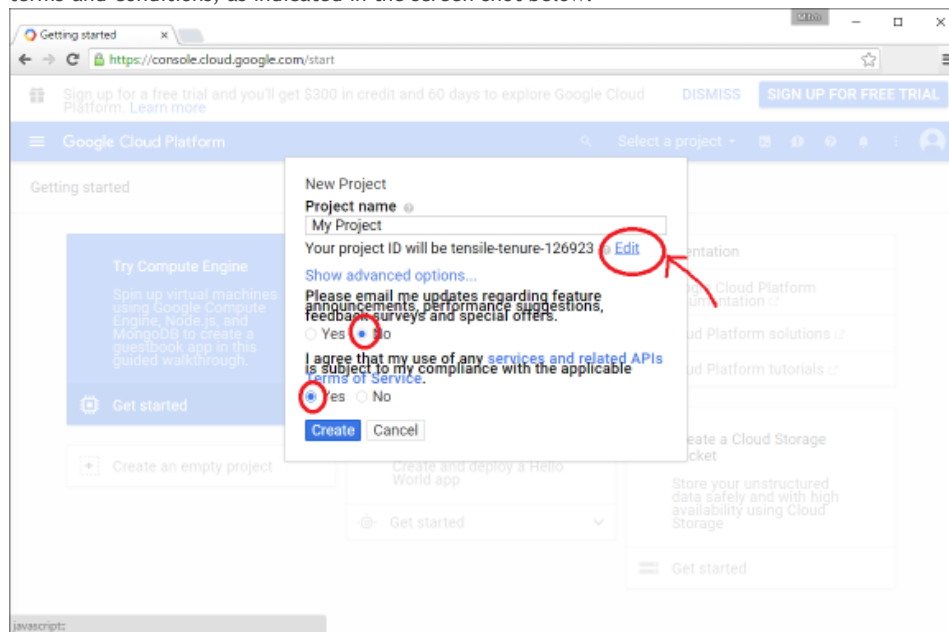

6. Upon creating the Google Cloud Platform project, you will be on an empty-project screen, showing the project name in the upper-right side of the screen. In this case, I named my project and

project id the same - "msundt-agg-test3":
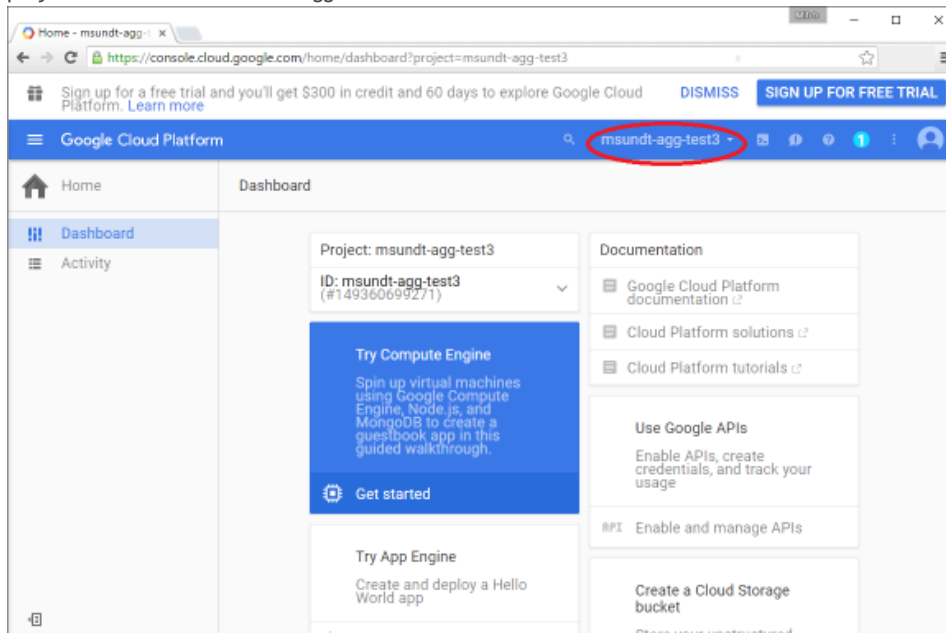
This now appears on the upper left next to Google Cloud Platform, rather than the upper right, due to site updates.
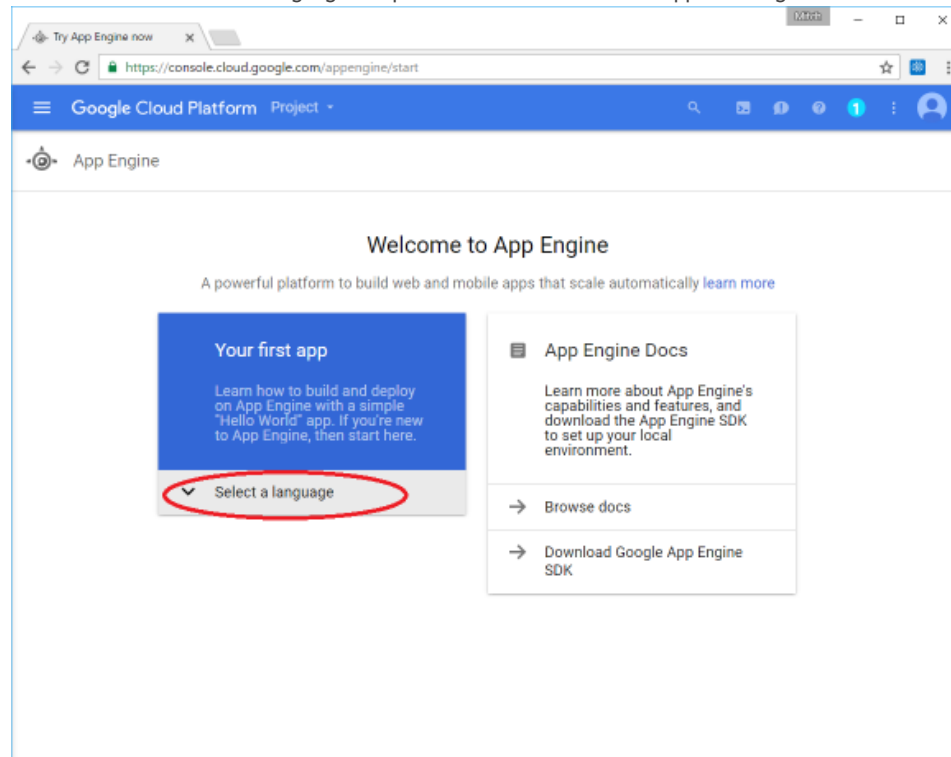
7. Now, click on the menu icon (three horizontal bars) to the left of "Google Cloud Platform" in the upper left side of the screen:
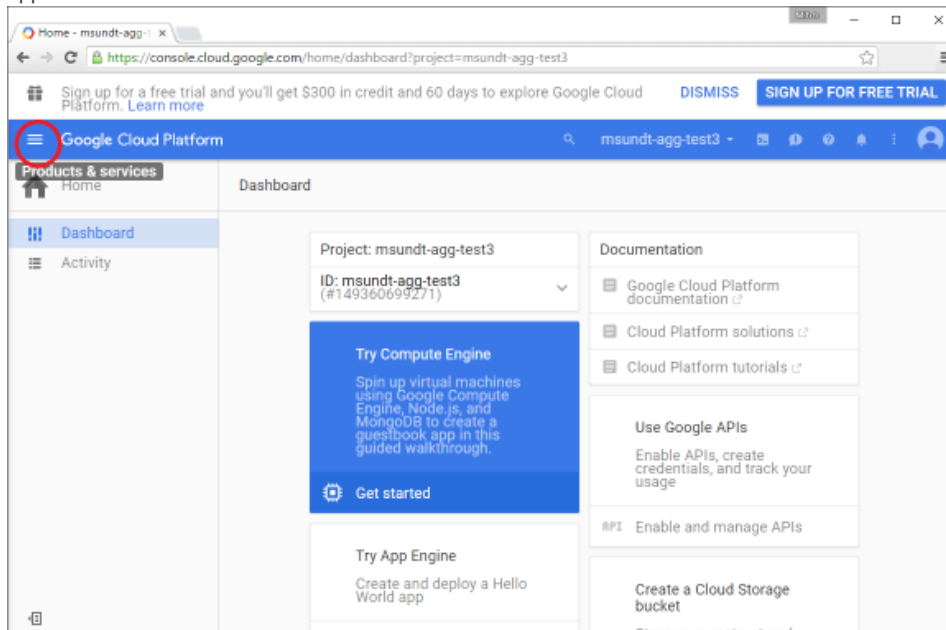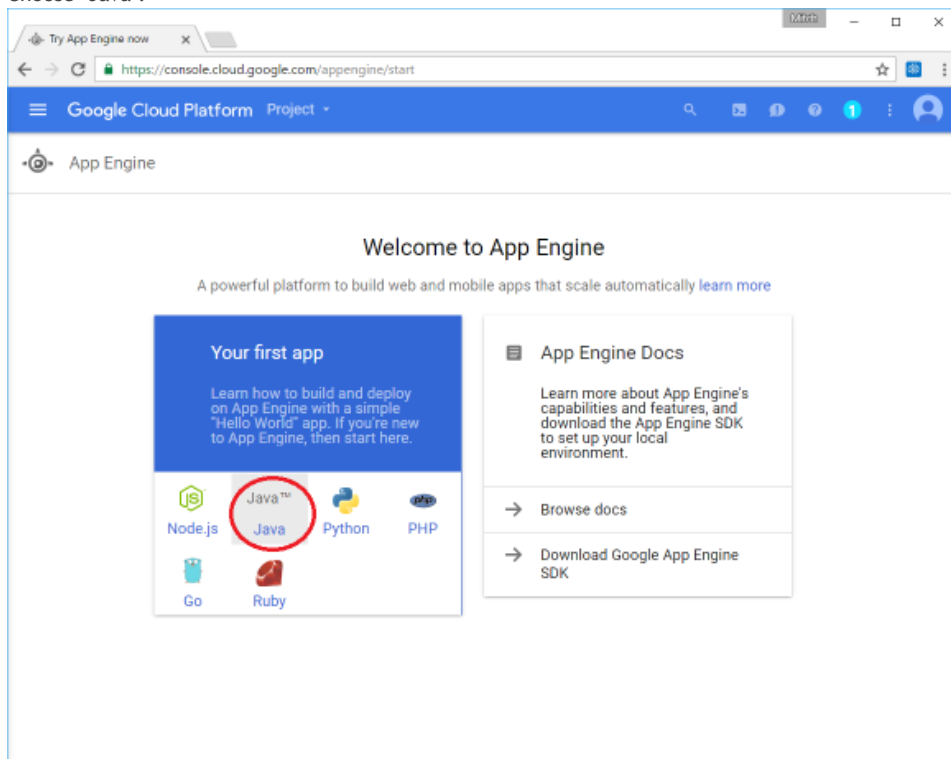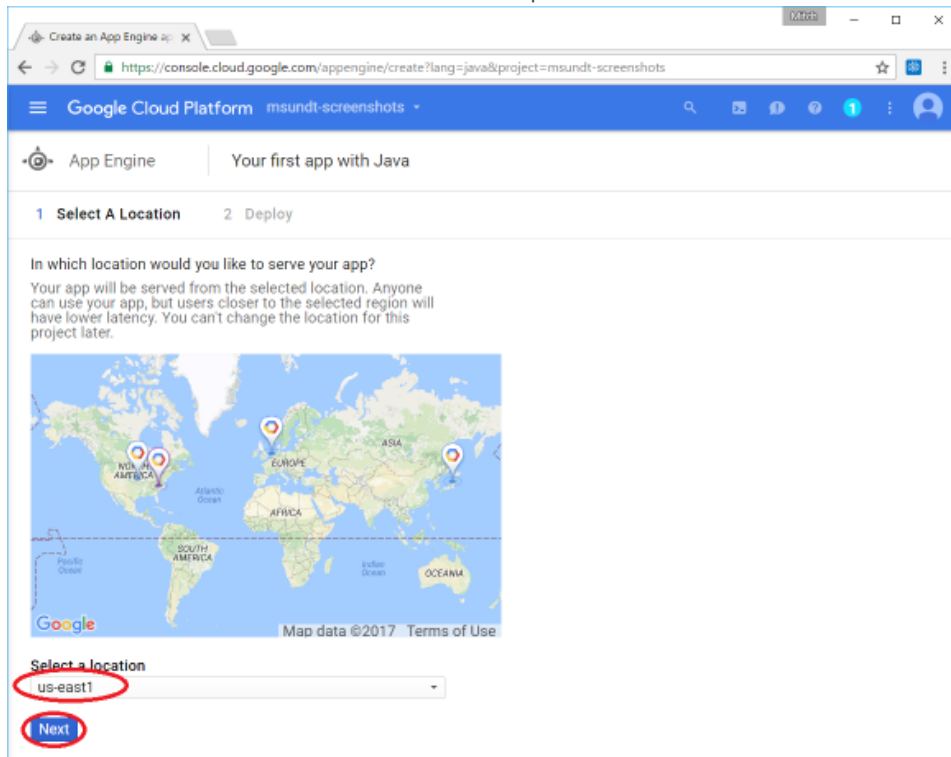
and select "App Engine":
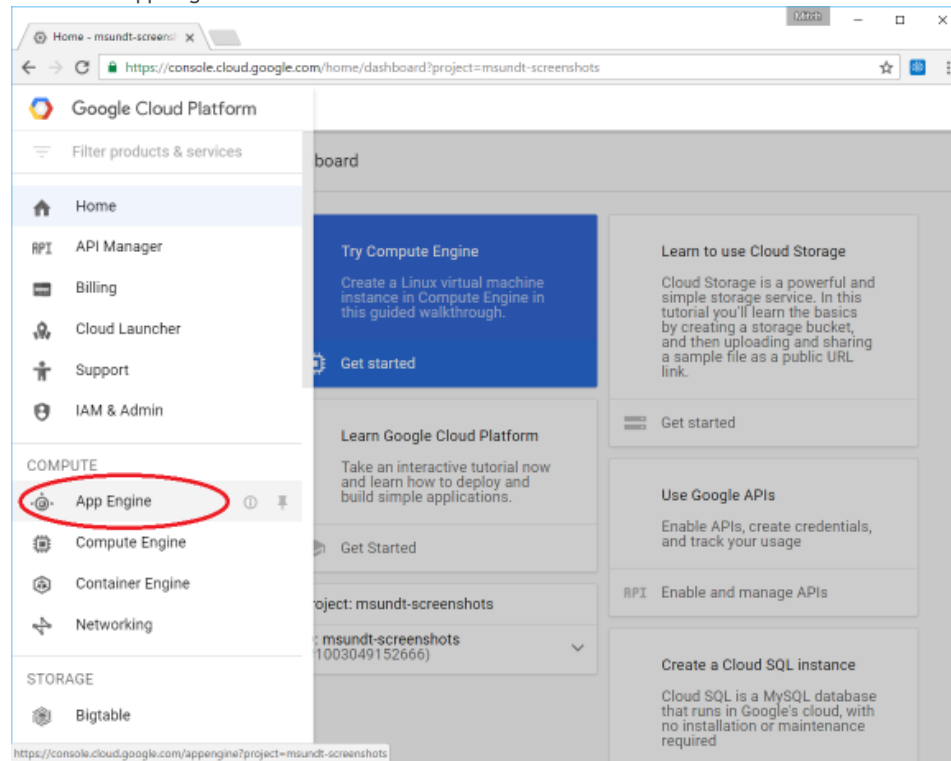
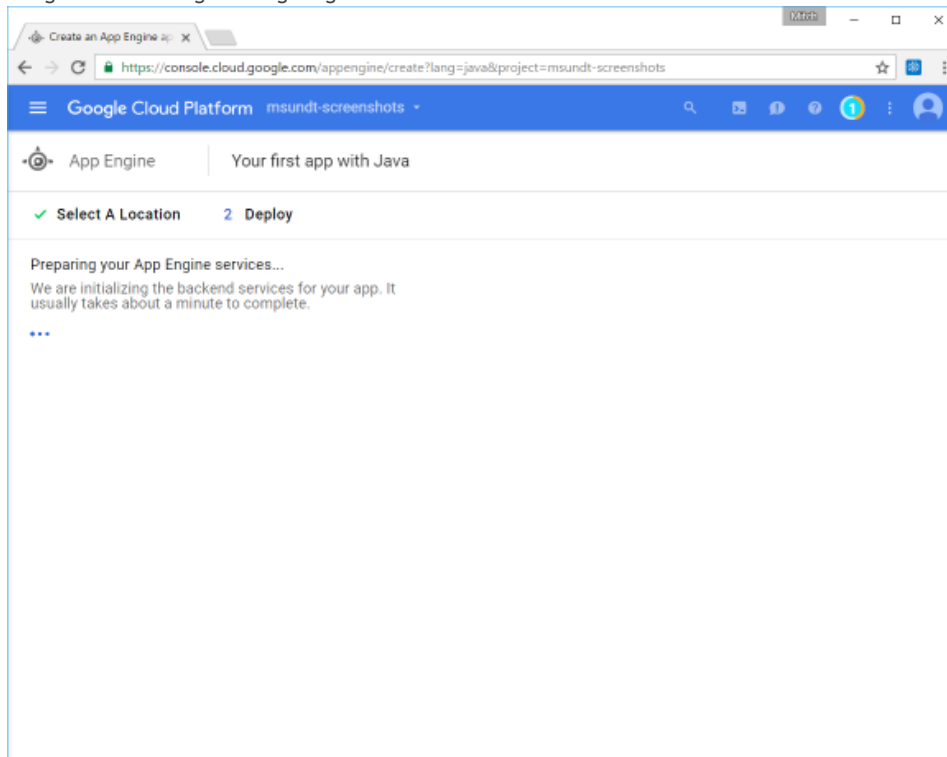8. And click on the "Select a language" dropdown under the "Your first app" heading:
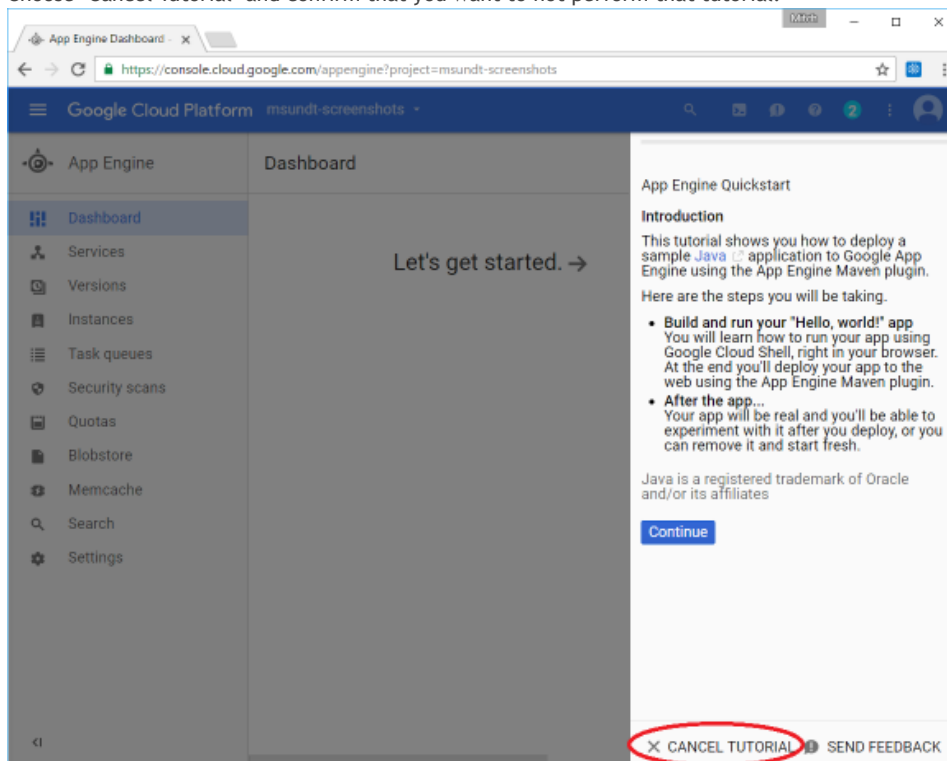
9. Choose "Java":

10. Select the datacenter location where this server will operate and click Next:

11. Google will then begin configuring the server:

When this completes, you will be directed to begin a tutorial to install a sample application. Choose "Cancel Tutorial" and confirm that you want to not perform that tutorial:

12. Download ODK Aggregate v1.N.N. Select the latest release for your operating system. These downloads are wizard-based installers for the various operating systems. You can verify the download using SHA256 signatures as described at the top of the downloads page. If you are running OSX, you must unzip the downloaded file before running the installer within it. If you are on MacOSX Mountain Lion or onward, you will need to fiddle with GateKeeper settings in order to run the installer. On Windows 10, you will need to approve running an unsigned installer. On Linux, you will need to change the downloaded file's permissions to enable running it as a program. Double-click the downloaded file to run it.

13. The installer will guide you through configuring ODK Aggregate for App Engine and then launch an upload tool to upload this configured ODK Aggregate to App Engine. NOTE: Beginning with Java 7

During the installation, the setup wizard will ask you to select an output parent directory. Any is fine, e.g., Documents.
It will also ask you to choose your platform. Choose Google App Engine.

Update 51, there are security level settings described here that may prevent the upload tool from running. A reported work-around is to add the file: path (e.g., `file:///`) to the Exception Site list. After you run the installer, you should remove that exception, as it defeats the purpose of that security measure. If you need to install a pre-1.4.8 version of ODK Aggregate, see the workaround here. The installer does not modify your system in any way; it only writes files to the folder you specify. The upload tool does write to one file in your home directory, but also, otherwise, does not modify your system. If the installer does not run on one computer, try running it on a different computer. If necessary, you can then copy the directory that the installer created to another system and run the upload tool from there.

14. The upload tool looks like this:

ODK Aggregate App Engine Updater - Version 2016-04-21

Email of Google Cloud Platform account (the owner of your App Engine):

Code from Google's account authorization web page (granting access for Google App Engine appcfg):

After entering the Email address, Click on "Get Token."

You will be redirected to a browser to obtain a code from Google's account authorization web page.
Copy that code into the pop-up window that will appear, and click OK to proceed.

IMPORTANT NOTE: The code is several characters wider than the text area that displays it on Google's web page. Please make sure you copy-and-paste the entire code into the pop-up window (and not just what is visible).

Output:

Delete Token   Get Token   Upload ODK Aggregate   Rollback

Enter the gmail account that you specified in step 2, above. This will enable the `Get Token` button.

15. Click the `Get Token` button. At this point, two things will happen.

1. Your default browser will open to a Google site (accounts.google.com) where you are asked to choose a gmail account (select the account from step 2), and then approve allowing "Google App Engine appcfg" to View and manage your AppEngine instances and datastores:

<div style="border:1px solid red; color:red;">
You may create any super-user name you would like for your Aggregate account. You do not have to have registered it anywhere before.
To find your Google AppEngine ID, on the Google Cloud Platform page, click on the name of your project (just to the right of the Google Cloud Platform title); the application id appears in gray, in all lower-case letters and with dashes in place of spaces, just to the right of the title.
</div>
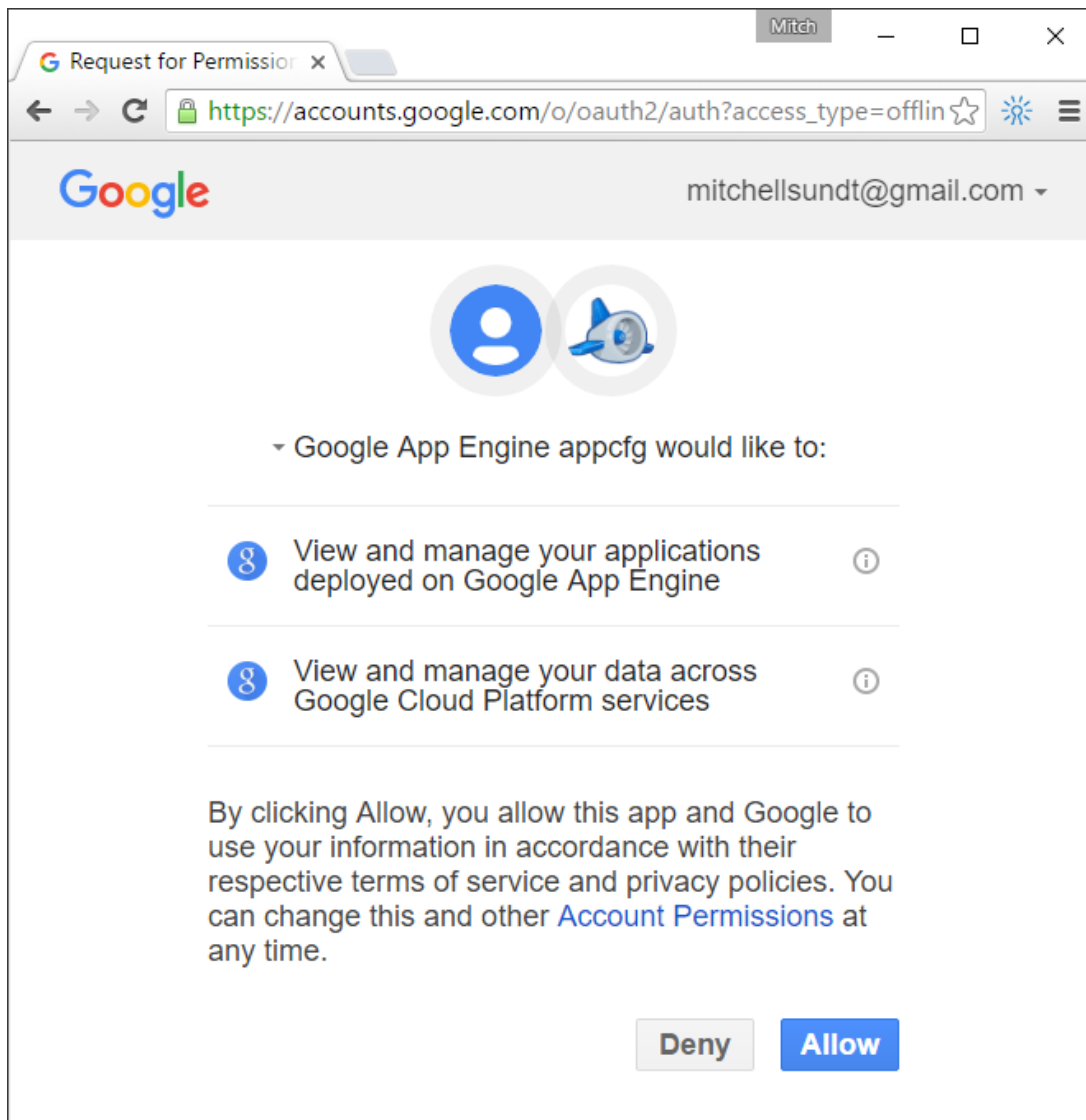
Click `Allow`. This will take you to a screen with instructions to copy a code.

2. At the same time, a pop-up dialog should be displayed by the upload tool. Regrettably, this seems to be suppressed about half the time, apparently because of the delay in opening the Google site. The pop-up looks like this:
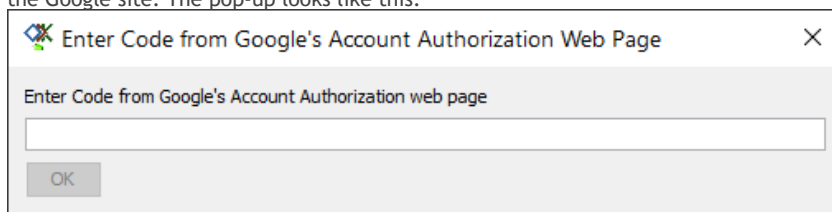

If the pop-up dialog does not show, close the upload tool and open a file browser or Finder window on the directory you specified for the installer to place its files. Navigate into the `ODK Aggregate` directory. If on Windows, double-click the `ODKAggregateAppEngineUpdater.jar` file. If on Mac OSX, double-click the `uploadAggregateToAppEngine.app` file. If on Linux, open a bash shell and run `uploadAggregateToAppEngine.sh`. These should all re-launch the upload tool. Re-enter the e-mail address, and once again click `Get Token`. The pop-up dialog should now appear.

This is wherever you JUST saved ODK in Step 13.

16. Copy the code from the browser into the upload tool's pop-up dialog and click `OK`. Note that the text box on Google's site is not as wide as the code; be sure to copy the entire code. If the code is accepted, the output should look something like this:

```
getToken!:  Jul 15, 2016 9:46:10 AM com.google.appengine.repackaged.com.google.api.client.extensions.java6.auth.oauth2.FileCredentialStore

getToken!:  WARNING: unable to change file permissions for everybody: C:\Users\Admin\.appcfg_oauth2_tokens_java

getToken :  Please enter code: Reading application configuration data...

getToken!:  Jul 15, 2016 9:46:14 AM com.google.apphosting.utils.config.IndexesXmlReader readConfigXml

getToken!:  INFO: Successfully processed ./LegacyRemoval\WEB-INF/datastore-indexes.xml

getToken :

getToken :

getToken :  Beginning interaction for module default...
```

```
getToken :  max_blob_size: 32000000

getToken :  max_file_count: 10000

getToken :  max_file_size: 32000000

getToken :  max_files_to_clone: 2000

getToken :  max_total_file_size: 9223372036854775807

getToken :  Success.

getToken :  Cleaning up temporary files for module default...

getToken!:

getToken :

status :  Action Succeeded!
```

If the output does not look like that, you may have delayed too long between getting the code and pasting it into the tool. Click `Delete Token` and try again.

If you see a failure with a message of the form:

```
getToken :  403 Forbidden

getToken :  You do not have permission to modify this app (app_id=u's~your-app-id').

getToken :

getToken!:  You do not have permission to modify this app (app_id=u's~your-app-id').

getToken!:  This is try #3

getToken :  Please see the logs [/tmp/appcfg1625673990974710084.log] for further information.

getToken :

getToken!:

status :  Action failed or aborted by user
```

Then it is likely that you have several different gmail accounts and Google has gotten confused during the token-issuing process. In this case, when the browser window opens, before selecting an account, copy the URL, open a Private Browsing or Incognito Window in your browser, and past the URL into that. Then proceed to get the token, etc. This should fix this issue.

17. Now click `Upload ODK Aggregate`.

    This will spew a very long list of progress messages into the Output window. The `listBackends :` and `deleteBackendBackground :` sections may report "500 Internal Server Error" and Severe errors, and Warnings about the use of Backends, a deprecated feature. This is expected.

    Toward the bottom, the `update :` section should not report errors and at the end, a `status : Action Succeeded!` line should be written. This indicates that the upload completed successfully.

    An example of the scrolling progress messages produced by a successful upload is here.

18. Once the updater script has run and uploaded the ODK Aggregate configuration to App Engine, return to the Google Cloud Platform console. With the console displaying your project, click on the menu icon (three horizontal bars) to the right of "Google Cloud Platform" in the upper left side of the screen:
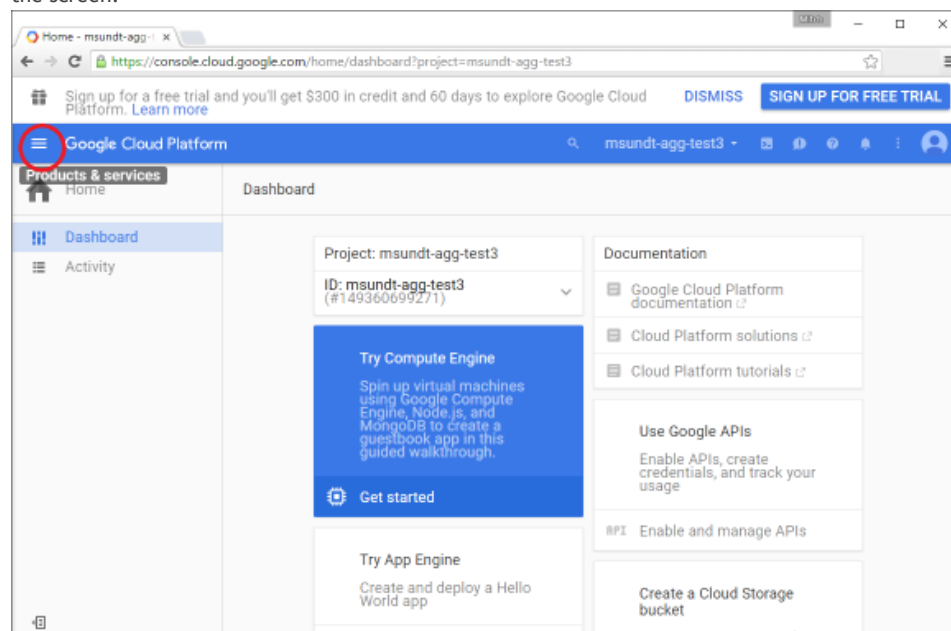
19. From this menu, select "App Engine":
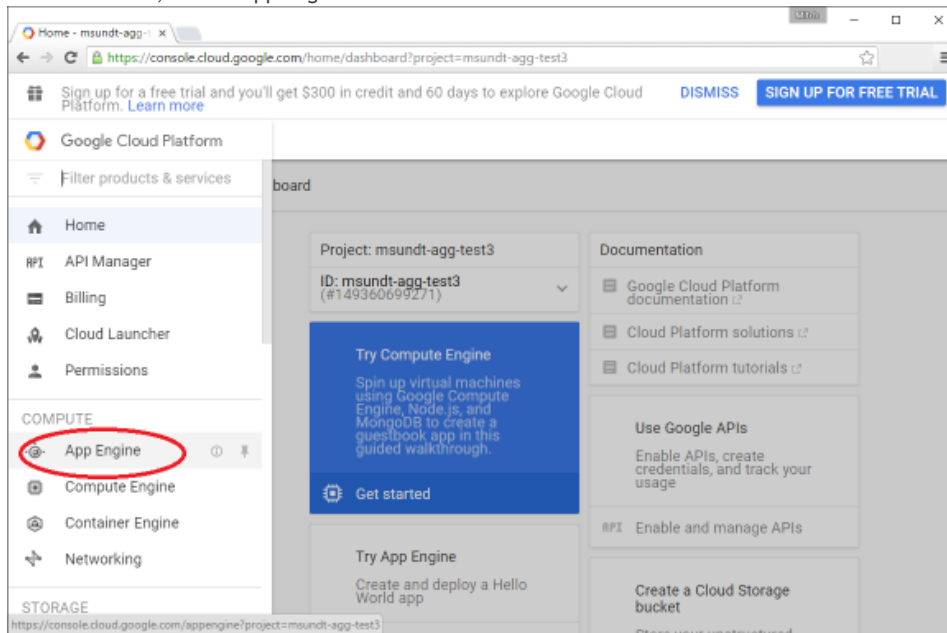
20. This will display the App Engine dashboard. You can confirm the upload by verifying that the list of modules contains both "default" and "background". If you need to set up billing, the settings to do that are at the bottom of this screen.
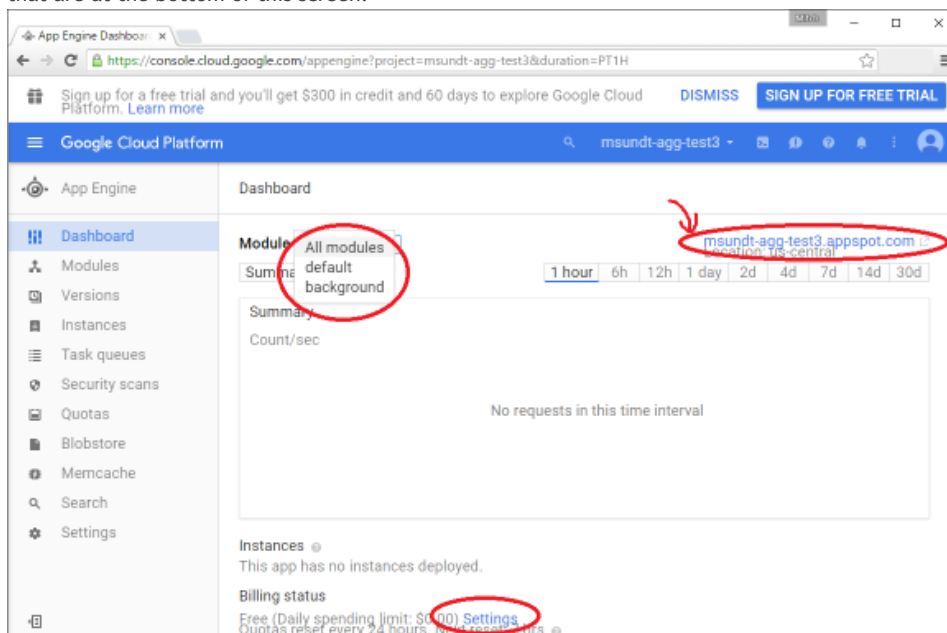
> This doesn't say Module anymore. It now says Services. Everything else applies though.

21. If you click on the project id (e.g., msundt-agg-test3.appspot.com) URL in the above screen, you will go to your ODK Aggregate server:
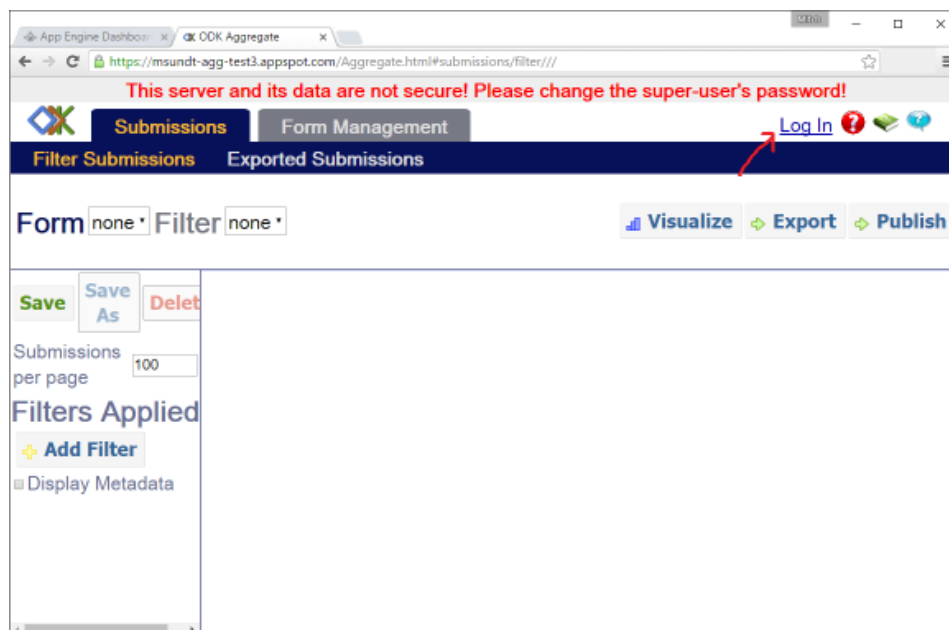
From here, you can click on "Log In" to log in, enter the ODK Aggregate username (superuser) that you specified within the installer (the initial password for this username will be `aggregate`) and access the site administration screens for your server. See Using the Application below.

Installing VM (Local or Cloud)¶

The ODK Aggregate VM is a fully-configured install of Aggregate that you can run on any computer. It requires very little setup, works well without Internet connectivity, and gives you complete control over your data collection campaign. Download the ODK Aggregate VM to begin.

Installing on Tomcat (Local or Cloud)¶

To run on ODK Aggregate on a Tomcat server backed with a MySQL or PostgreSQL database, see Aggregate Tomcat Install.

NOTE: The Visualization and Mapping features of ODK Aggregate do not work when the server is configured to run under certain non-US, non-English locales. We believe this occurs because, in those locales, numbers are formatted with a comma as the decimal point, and that this confuses the visualization software we are using. If you are having issues with these features not working, reconfigure your Tomcat server to run in the US English locale.

Installing on AWS (Cloud)¶

To run on ODK Aggregate on a Linux micro-instance on the Amazon Web Services EC2 infrastructure, see Aggregate AWS Install.

NOTE: The Visualization and Mapping features of ODK Aggregate do not work when the server is configured to run under certain non-US, non-English locales. We believe this occurs because, in those locales, numbers are formatted with a comma as the decimal point, and that this confuses the visualization software we are using. If you are having issues with these features not working, reconfigure your Tomcat server to run in the US English locale.

Using the Application¶

1. Beginning with ODK Aggregate 1.3.2, upon the initial installation of the server, it is configured to allow unauthenticated (*anonymousUser*) submissions from ODK Collect and unauthenticated browser access to the submissions and forms management functionality of ODK Aggregate. When the URL to the ODK Aggregate server is first opened, you will be presented with the application page showing the *Submissions* and *Form Management* tabs.

Scroll to Using the Application unless you're considering the server or virtual machine approach (see our main instruction document for more details)

2. In April 2015, the use of Google e-mail accounts for accessing the site (via *Sign in with Google*) stopped working (Google turned off that functionality). If you have an existing site running an old version of ODK Aggregate that does not have ODK Aggregate usernames configured for website access (and offers a *Sign in with Google* sign-in choice), you will need to upgrade to regain access to it.

3-5 are asides about the tabs. Focus on 6.

3. Submitted data, once in ODK Aggregate, can be viewed, exported, mapped and deleted.

4. Use the *Add New Form* button on the *Form Management* tab to upload a new form definition to ODK Aggregate.

5. View data submitted from ODK Collect on the *Submissions* tab.

6. If the *Site Admin* tab is not visible, click the *Log In* link in the upper right corner of the screen to be presented with the *Log onto Aggregate* screen. Choose the *Sign in with Aggregate password* button and enter the super-user username you specified within the installer. The initial password for this account is `aggregate` When signing in with this method, if you do not enter the password correctly, you may need to close all your browser windows and quit your browser before you can try again.

7. If you have not yet changed your super-user password to something other than `aggregate`, the server will display **This server and its data are not secure! Please change the super-user's password!** at the top of the web page. Please visit the *Permissions* sub-tab under the *Site Admin* tab to change this user's password.

8. If the instance name of the server changes (the installer asks for this name), then the passwords for all ODK Aggregate usernames will be cleared (preventing their use) and the super-user username's password will be reset to `aggregate` and the above message will also be displayed. In this case, you should log in, change the super-user's password, and change the passwords for all of your ODK Aggregate usernames.

9. Use the *Permissions* sub-tab under the *Site Admin* tab to restrict who can download forms or submit data from ODK Collect. Do this by creating an ODK username and password and granting it *Data Collector* privileges. This username and password can then be entered into ODK Collect's settings page. ==When restricting access you must also remove the *Data Collector* privilege from the *anonymousUser*.== Remember to click *Save Changes* to make these changes take effect. Conversely, granting the *Data Collector* privilege to the *anonymousUser* enables anyone to submit data to your ODK Aggregate server.

10. Use the *Permissions* sub-tab under the *Site Admin* tab to specify additional usernames with browser access to the server. For each user you add, select whether they have access to the submitted data (*Data Viewer* privileges), the ability to upload forms and export or publish data (*Form Management*), or the ability to manage site access and users (*Site Admin*) privileges. Remember to click *Save Changes* to make these changes take effect. And, for each username you define, remember to *Change Password* to assign a password; by default, usernames are created with unusable passwords. Granting any of these privileges to the *anonymousUser* enables browser access to these functions without first logging in.

11. Refer to the inline help in ODK Aggregate (near the login button) for more detailed instructions.

From here down, advanced users only

If you have many form definitions on your server, you may get better performance and reduce the likelihood for data corruption if you increase the size of your server (this also applies to Tomcat deployments but Tomcat server size is configured differently and is not discussed here).

To change Google App Engine configuration, you must edit the configuration files produced by the installer and re-run the uploader script to push the changes to Google's servers. There are two server settings that can be changed -- the web server size and the background server size.

Data corruption is generally caused by the premature termination of an action (e.g., saving of a submission) because it took longer than the allotted time. The likelihood of data corruption occurring is tied to the quantity of form definitions on the server, the size of the individual submissions, the number of devices simultaneously submitting data, and the speed of the network. Increasing the web server size enables it to handle larger workloads faster, which can reduce the likelihood of hitting this time limit thereby avoiding data corruption.

For data corruption caused by slow network speeds, you might also be able to change more aspects of the App Engine configuration (specified in these files) to make your web server always-available and to replace it with a Bx instance that does not have an automatic request time limit (the documentation

https://opendatakit.org/use/aggregate/

These two are important, especially if you'll be working with research assistants. Per the details on the Site Admin tab (read these!): you can create an ODK account for RAs and invite them.
Note that the site administrator CANNOT be a data collector. You will have to create a second account to enable yourself to collect data. Create this account, give it all permissions (from collector through site administrator), and click Save Changes. Then change the password for that account and click Save Changes again.
Once that account is created, click Log Out (top right-hand corner), close the tab, and close your browser window. Then re-open your browser, navigate to your appspot.com URL (remember, it's [your project name].appspot.com), and log in to the account you just created. This is the account you should use to load forms, collect data, etc!

provided by Google is currently unclear on whether this is still possible with the new services constructions).

Web Server Size

The web server handles all browser interactions and all data-submission and form-download requests from ODK Collect and ODK Briefcase. Increasing the size of the web server should reduce the likelihood of data corruption if it is not caused by slow network speeds. To change the Google App Engine web server size, go to the folder you specified to the installer. Within that folder, navigate to:

```
ODKAggregate/default/WEB-INF
```

Within that directory, there will be a file `appengine-web.xml`. Open that file in a text editor like Notepad++ or Notepad. The file contents will look something like:

```
<appengine-web-app xmlns="http://appengine.google.com/ns/1.0">
    <application>opendatakit-simpledemo</application>
    <module>default</module>
    <version>1</version>

        <instance-class>F2</instance-class>
  ...
```

To change the size of the web server, replace **F2** with a different instance class size. There are several different instance classes available. **Select from among the instance classes beginning with the letter F.** See instance classes for their descriptions or search for "Google AppEngine instance classes standard environment" on the web.

Then re-run the upload tool within the `ODKAggregate` folder either by double-clicking the "ODKAggregateAppEngineUpdater.jar" file (Windows), or double-clicking the "uploadAggregateToAppEngine.app" file (Mac OSX), or double-clicking the "uploadAggregateToAppEngine.sh" file (linux). Once you have uploaded these changes to App Engine, your server will be running on the instance size that you have specified.

Background Server Size

App Engine deployments use a "background" copy of the website to process long-running actions like generating CSV and KML files for export and for publishing all accumulated data to an external server. If you experience difficulty exporting to CSV or KML, the size of that server will also need to be updated. In that case, go to:

```
ODKAggregate/background/WEB-INF
```

Within that directory, there will be a slightly different file with the same `appengine-web.xml` filename. Open that file in a text editor like Notepad++ or Notepad. The file contents will look something like:

```
<appengine-web-app xmlns="http://appengine.google.com/ns/1.0">
    <application>opendatakit-simpledemo</application>
    <module>background</module>
    <version>1</version>

        <instance-class>B2</instance-class>
  ...
```

To change the size of the server, replace **B2** with a different instance class size. There are several different instance classes available. **Select from among the instance classes beginning with the letter B.** As before, see instance classes for their descriptions or search for "Google AppEngine instance classes standard environment" on the web.

And, as above, re-run the upload tool to make these changes take effect on Google's servers.

Useful Notes¶

- ODK Aggregate install for Windows from Geo for Good 2015
- The developer wiki has release notes and tips on Tomcat troubleshooting (and App Engine troubleshooting).
- Google App Engine servers may be located anywhere in the world. Depending on the sensitivity of the data and specific storage rules/restrictions, the server infrastructure may not have all necessary security precautions (such as encryption). It is the organization's responsibility to

research and comply with applicable laws and regulations before storing data on Google App Engine. The organization is also responsible for taking the appropriate security precautions and educating users that the information will be available to the organization and stored on Google servers as specified in the Google App Engine Terms of Service.

- Looking for other servers? Try ODK Briefcase (for local offline storage), FormHub (for free data hosting on the cloud), DataHQ (like Aggregate but in Python), RapidSMS (for SMS integration) or OpenMRS (for medical records).

About
Overview
Tools
Deployments
Research
Press
Team
Contact
Security and Privacy
Blog

Use
Overview
XLSForm
Build
Collect
Aggregate
  Data Transfer
  Deployment Planning
  Tomcat Install
  OAuth2 Service Account
Form Uploader
Briefcase
Validate
Sensors
  Zebra Printer driver
2.0 Tool Suite
  Current Distribution
  Previous Distribution
  Older Versions

Download
Overview
Aggregate
Collect
Briefcase
Validate
Sensors and Printers
Survey and Tables
Other
Unsupported

Help
Overview
FAQ
Form Design
  Binding
  Body
  Guidelines
  Data Preloading
  Cascading Selects
  External Itemsets
  External Apps
  OpenStreetMap
  iText
  Examples
  XLSForm
  Gotchas and Workarounds
Encrypted Forms
Image-based Forms
Verifying Downloads
Training Guides
Help For Hire

Participate
Overview
Code
Donate