

Migrating Mailing Lists to Google Groups

Jeremy Krovitz' Notes

Table of Contents

The Premise	2
Task	2
Background Research	3
Proof of Concept	3
Google Qwiklabs	3
Defining an API	3-4
Directory API	3-4
Observations from using the Google Sheets API	4
Questions for Denise and Answers	5
Questions and Answers on 06-13-2019	5
Questions and Answers on 06-21-2019	5
Research and Links to Articles Found	5-6
Things to Look Into	7
Should we host the test CSV files on a Linux server or a Windows server?	7
Mailing List on Educause	7
Replicating Real World Scenario: I am the developer and Denise is the client	7-8

The Premise

- Creating a script that would live within the infrastructure.
- Research **proof of concept**.
- Currently using mailman. Want to move mailing lists (mailing lists include all of the mailing lists on campus such as student mailing lists, advisor mailing lists, etc.) to Google Groups.
 - They want to accomplish this using the Google API. They need to put together a **proof of concept** that attempts to answer the following questions:
 - Is it technology feasible?
 - Could you even do it?
 - What needs would need to be met?
 - What would need to happen to move from mailman to google groups?
- Bring Denise a list of what would be needed to make the migration work.
 - Start with Google Qwiklabs.
 - Look at the Directory API and answer the following questions:
 - What is the Directory API for?
 - Why should I care?
 - How would I use it?

Task

Come up with a script for automating moving a file of names and email addresses to Google Groups. It should be an automated system that would probably back-up every night (similar to a cron job). Denise said that Google Tools should have a built in system for automating tasks (similar to cron).

First phase to focus on: just being able to get names and emails from a CSV file and be able to add the names and email addresses to an existing Google Group.

Background Research

Proof of Concept

"This process is designed to determine whether a software idea can be built in the real world, what technologies should be used in development, and whether the software is likely to be adopted by its intended users" ([MacPherson](#)).

Google Qwiklabs

Google Qwiklabs is a cloud computing platform.

Defining an API

An API is not a database or even the server. An API is the code that governs the access point(s) for the server. An API is not a database. It is an access point to an app that can access a database

Directory API

Look at the Directory API and answer the following questions:

- What is the Directory API for?

The Directory API is used for performing administrative tasks on users, groups, organizational units, and devices in your account.

- Why will the Directory API be important to use?

The Directory API is important because it provides a way of organizing an organization's chart and managing organizational structures programmatically. It would be useful for the Macalester mailing lists because it would allow for custom user profiles to be created to fit Macalester's organizational needs. For example, for an adviser's mailing list, the adviser needs to have an account identifying them as the adviser, so they are able to send one email to all of their advisees such as telling them they need to meet with them in order to get their pin number for registration. Furthermore, there are also various security measures in place, such as if a user is identified with a profile type of student, then they should not be able to modify the mailing list of their adviser.

- How would I use it?

Steps for using the Directory API:

1. Set up a new project in the Google API Console if one has not been setup already, and activate the **Admin SDK** service to use for this project.
2. Create a new Google Script.
3. Click on **Resources**. A dropdown will open.
4. Select **Advanced Google services...**
5. Click to toggle the Admin Directory API from off to **on**.
6. Press **OK**.

Observations from using the Google Sheets API

```
gc = gspread.authorize(credentials)
wks = gc.open('TestEmailList').sheet1
wks.get_all_records()
```

Looking at these three lines of code, `wks.get_all_records()` returns all cell values in the spreadsheet. The values are returned as a list of dictionaries. The header values in the spreadsheet make up the keys of each dictionary. The other rows of the spreadsheet make up the values for each dictionary. Therefore, each dictionary corresponds to one row of the spreadsheet.

Name	Email
Jeremy Krovitz	jkrovitz@gmail.com
Kam Poncius	kamponcius@gmail.com
Karen Davis	contactkarendavis@comcast.net



```
[{'Name': 'Jeremy Krovitz', 'Email': 'jkrovitz@gmail.com'},  
{ 'Name': 'Kam Poncius', 'Email': 'kamponcius@gmail.com'},  
{ 'Name': 'Karen Davis', 'Email': 'contactkarendavis@comcast.net'}]
```

Questions for Denise and Answers

Questions and Answers on 06-13-2019

- Thinking about the workflow for the script, you mentioned that the script should automate moving a file of names and email addresses to Google Groups. Should the script also automate the process of exporting the names and email addresses from mailman into the file that will be moved to Google Groups?

No, the names and email addresses will already have been exported from mailman into CSV files. The location of where the CSV files will be housed has not yet been determined. The names and email addresses will be in the following format:

Jeremy Krovitz <jkrovitz@macalester.edu>

- Is there a specific file format(s) that would be used? In my opinion Google Sheets makes logical sense. Should I be focusing on more than one file format?

CSV files will be used and will be the only format being used.

Questions and Answers on 06-21-2019

- Is the command `AdminDirectory.Groups.insert(group)` considered to be a *post* request? I'm wondering because I know there are implicit ways of making *post requests* where *post* is not specified explicitly such as `urlfetch` for *get and post*. I did some research online, but am having some difficulty finding the answer to this question. I'm thinking that since a JavaScript object rather than a JSON object was passed into the `insert` function, it would not be considered a *post* request since JSON notation is typically used when sending data to a server or retrieving data from a server. Because I did not use JSON, I am thinking the command `AdminDirectory.Groups.insert(group)` is not a post request. Am I on the right track here?

Research and Links to Articles Found

Research and Links found on June 27, 2019

The directory commands under the [Directory API section](#) of the Google Groups API POC Document lists commands but the commands listed are only for Java rather than JavaScript/Google Script. Here is the link where one of the commands is described in the documentation for implementing the Directory API in Java: https://developers.google.com/resources/api-libraries/documentation/admin/directory_v1/java/latest/com/google/api/services/admin/directory/Directory.Members.Insert.html. I know that you guys want to avoid using Java, but I just wanted to point this out, as the commands could easily be mistaken as commands that would work in a Google Script.

I may have found an example of a Google Group being created. I'm not entirely sure that it works. This is the link: <https://developers.google.com/admin-sdk/directory/v1/reference/groups/insert#auth>. This particular example uses the JavaScript Client Library, which it looks like is part of the Google Compute Engine: <https://cloud.google.com/compute/docs/tutorials/javascript-guide>. I doubt we intended to use Google Compute. The only reason I got to this was because when I was tinkering around with the code for creating a Google group, I got an error that said, "ReferenceError: 'gapi' is not defined."

This link has example functions using commands with the prefix AdminDirectory (a way of referring to the Directory API): https://developers.google.com/apps-script/advanced/admin-sdk-directory#list_all_users. So far two methods from this link have been successful including printing out members of a group and adding a member to a group. However, is not an example method using the command prefix AdminDirectory that would create a group. There is probably a way of using the command prefix AdminDirectory to create a group; I just have not yet figured it out.

I used the JSON formatting shown in this article <https://developers.google.com/admin-sdk/directory/v1/guides/manage-groups> as a guide to modifying the JavaScript object formatting in the function `createGroup()`, which was a copy of the function that inserted members into a group called `addGroupMember()`. In addition to modifying the JavaScript object formatting, I also modified the command with the AdminDirectory prefix for inserting members into the group which was "AdminDirectory.Members.insert(member, groupEmail)." In the `createGroup()` function this command was modified to be, "AdminDirectory.Groups.insert(group)."

<http://googleappscripting.com/> is a useful website for tutorials on Google Apps Script.

Research and Links found on 7/3/2019

This is a guide to writing formal requirements:

<https://drive.google.com/file/d/11sUmDOtK4YHLxAkXIKqW4UNsr2zQ-RR-/view?usp=sharing>

This article provides more detailed instructions and examples for writing functional requirements:

<https://www.altexsoft.com/blog/business/functional-and-non-functional-requirements-specification-and-types/>

Sample functional requirements specification that I found helpful (pages: 10 - 21):

https://docs.google.com/document/d/1UNd0LxN34rTk2_sCNWylDZOZV7R8KKEd-eVKJGUKmbA/edit#bookmark=id.v0i8zaadg5k1

Things to Look Into

Should we host the test CSV files on a Linux server or a Windows server?

(I think it would probably be the best idea to use the same type of server that Gabe uses since that is where the files will be hosted when the script is live. Denise thinks this is probably a good idea as well, but thinks we should still look into this.)

Update on June 24, 2019 - Assume that we will be using a Linux Server.

Mailing List on Educause

Go to Educause.org. They have a mailing list for Google Groups. Like a list-serve. See if I can join it. If I can't see if I can ask questions. (I should be able to because I have a Macalester email address).

I joined the mailing list.

Replicating Real World Scenario: I am the developer and Denise is the client

One thing to keep in mind is that it is impossible to have an exact answer about how long something is going to take you. It is important to consider:

1. The number of hours you will spend.
2. When the delivery date is.

To plan the number of hours in accordance with the delivery date, you must write up a formal list of requirements, often called the functional requirements (i.e. What is this thing going to do?). As a consultant, it is my job to put together a formal list of requirements to discuss with the client. It is important to add to the list items that are in scope (that are actually possible to get done by the deadline) and things that are out of scope. Items/tasks that are out of scope are things that you might eventually want to accomplish, but they are not the immediate focus.

Status reports are also very important. It is important to be in agreement with the client as to how often I should present them with a status report as well as the components that will be in the status report such as what I have accomplished, things that still need to be accomplished, and how much time I have spent so far.

It is important to be mindful of change requests. A change request is a separate document and is SUPER IMPORTANT. Change requests are not part of the original time estimate. If you get a change request from the client, you have to either change the amount of time the project will take or you have to decrease another functionality that you were initially going to build.

It is important to focus on what is appropriate to bill this customer for and what is not important to bill this customer for. It is important to keep track of time spent working on this project. It is important to think what are the phases of this project? It is important to plan it out in a way that you best work. It is important to be as straightforward as possible with the customer. Any

challenges or obstacles you see. Another thing that can be part of the status report are dependencies ie Waiting for Sean to spin up server... to be determined.

Status reports will be every Tuesday and Friday. How would you like those delivered? In an in-person meeting. Denise and I will set up recurring calendar events.

Replicating Real World Scenario: I am the developer and Denise is the client - July 2

Several ways of file feeds:

Append - Check if information has been added, if the record is already there don't add it, but it is not there then add it.

Update - Checks to see if the data exists. If the data is there, look to see if the data is the same or if it needs to be updated. For example, if somebody gets married, their email might be the same, but their last name may need to be updated.

Truncate - Wipes out the existing records in that group and does a batch upload of all of the new data. - very fast

The methods will be used on all different kinds of lists. Some will require append, some update, some both, etc.

How this will work is keeping track of who is already flagged to not receive while keeping track of other updates they may have, others who might need to be added or removed from the list, etc.

Look into ways to automate the process of running a script (similar to cron). It would happen nightly.

Print out status reports. Think about what would be helpful for Denise to know in order to keep the project moving.

In terms of the code, I can show that to Denise on the screen if it actually runs.

Status reports should include what's next

For the end of next week:

A working example of taking names and emails from the csv file and the names and emails to a test group.