



API Extravaganza!

# Combining Google Analytics & ORCID APIs

ANDS Webinar, 22 Sept 2016

Liz Krznarich, Software Engineer/UI Designer, ORCID

e.krznarich@orcid.org <http://orcid.org/0000-0001-6622-4910>

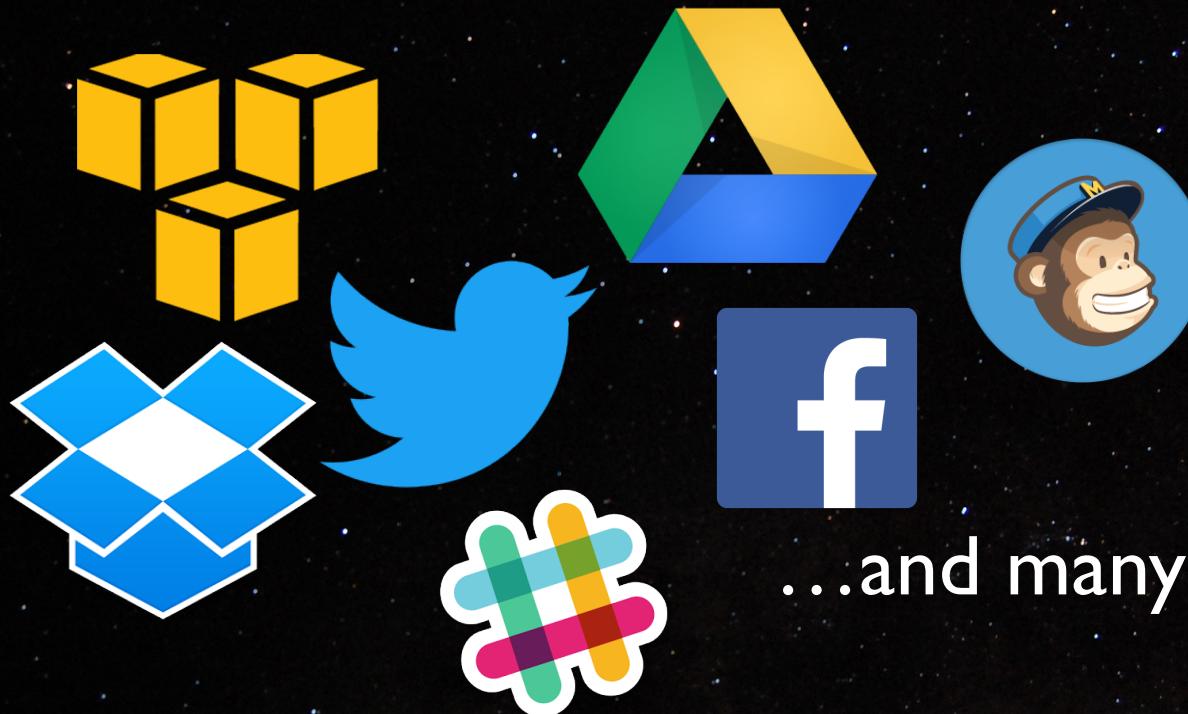
## Photos. Documents. Data. Apps.

We've got loads of stuff up there in the  
cloud, accessible at any time  
(This is awesome!)

...but sometimes one single source or app doesn't quite cut it. You need data/features from multiple sources.

**APIs to the rescue!**

Loads of APIs out there



...and many more

But first,

# What's an API??

- Application Programming Interface
- Allow computer applications to talk to each other
- Do things you can do in app user interface + more! (and faster)
- Combine data/functionality from multiple apps into your own custom app

Today's topics

Custom reporting with Google  
Analytics / Drive / Sheets APIs +  
ORCID APIs



# What are these tools?



**Analytics:** Track user behavior on your website



**Drive:** Cloud file storage/sharing



**Sheets:** Cloud spreadsheet app (lives in Drive)



**ORCID:** Persistent ID for researchers + digital record of scholarly contributions

## In this session we'll:

1. Query the Analytics API
2. Set up Google API credentials
3. Get Analytics data & upload to Drive
4. Set up ORCID API credentials
5. Get data from ORCID & add it to our Drive file (using the Sheets API)

# Pre-reqs: Build a website

<http://orcid.github.io/analytics-demo>

The Annals of ORCID

Home Analytics Get the code!

## Repository Stuff

**ORCID: a system to uniquely identify researchers**

Type: Publication (Journal article)  
Date: 1 October 2012  
Authors: Laurel L. Haak, Martin Fenner, Laura Paglione, Ed Pentz, Howard Ratner  
Identifier: 10.1087/20120404

[Download](#) [View](#)

**ORCID Public Data File 2015**

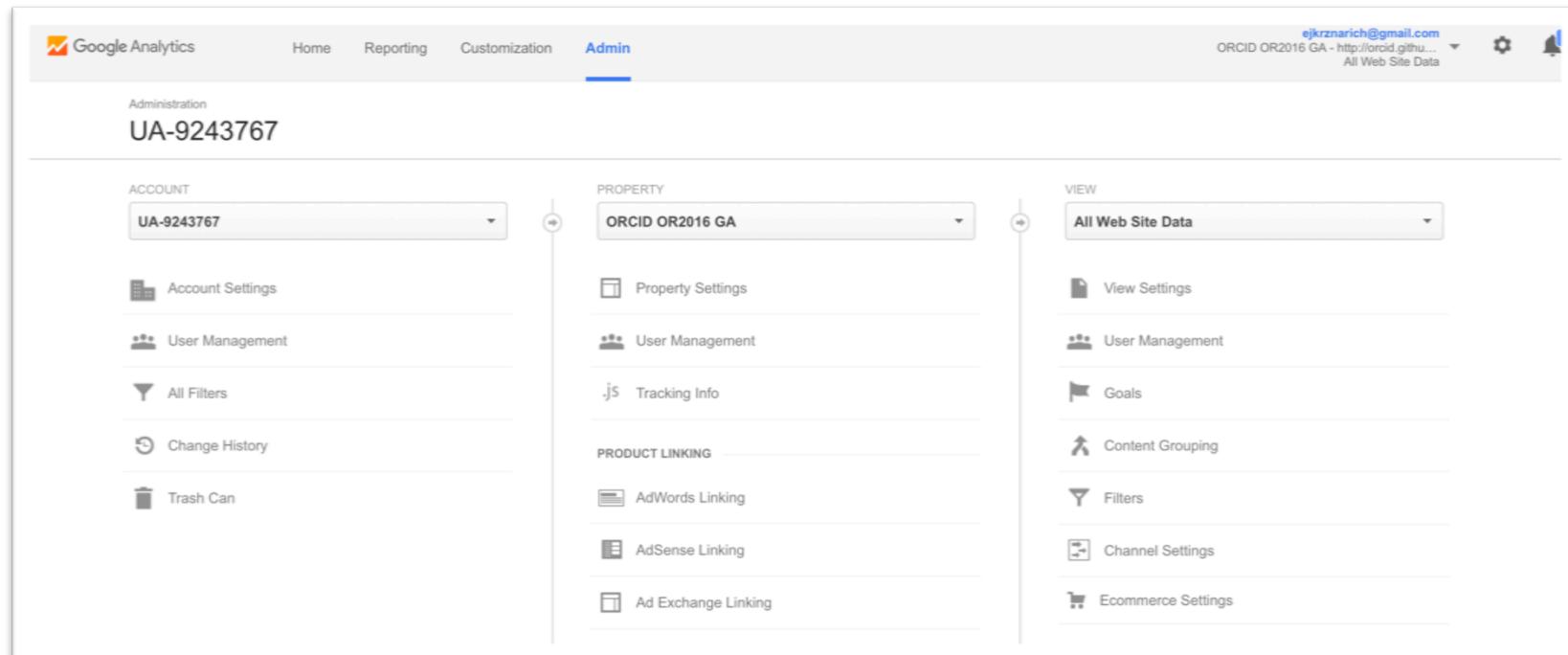
Type: Dataset  
Date: 23 October 2015  
Authors: Paglione, Laura; Peters, Robert; Wilmers, Catalina; Simpson, Will; Montenegro, Angel; Ramírez Monge, Fran; Tyagi, Shobhit; Krznarich, Elizabeth; Demeranville, Tom; Brown, Josh; Miyairi, Nobuko; Buys, Matthew; Cardoso, Ana; Sethate, Cheryl; Haak, Laurel  
Identifier: 10.6084/m9.figshare.1582705.v1

[Download](#) [View](#)

ORCID ID: THE ANNALS OF ORCID | PUBLIC DATA FILE 2015 | DOWNLOAD THE SOURCE CODE | CONFIDENTIAL (ATO: 001)

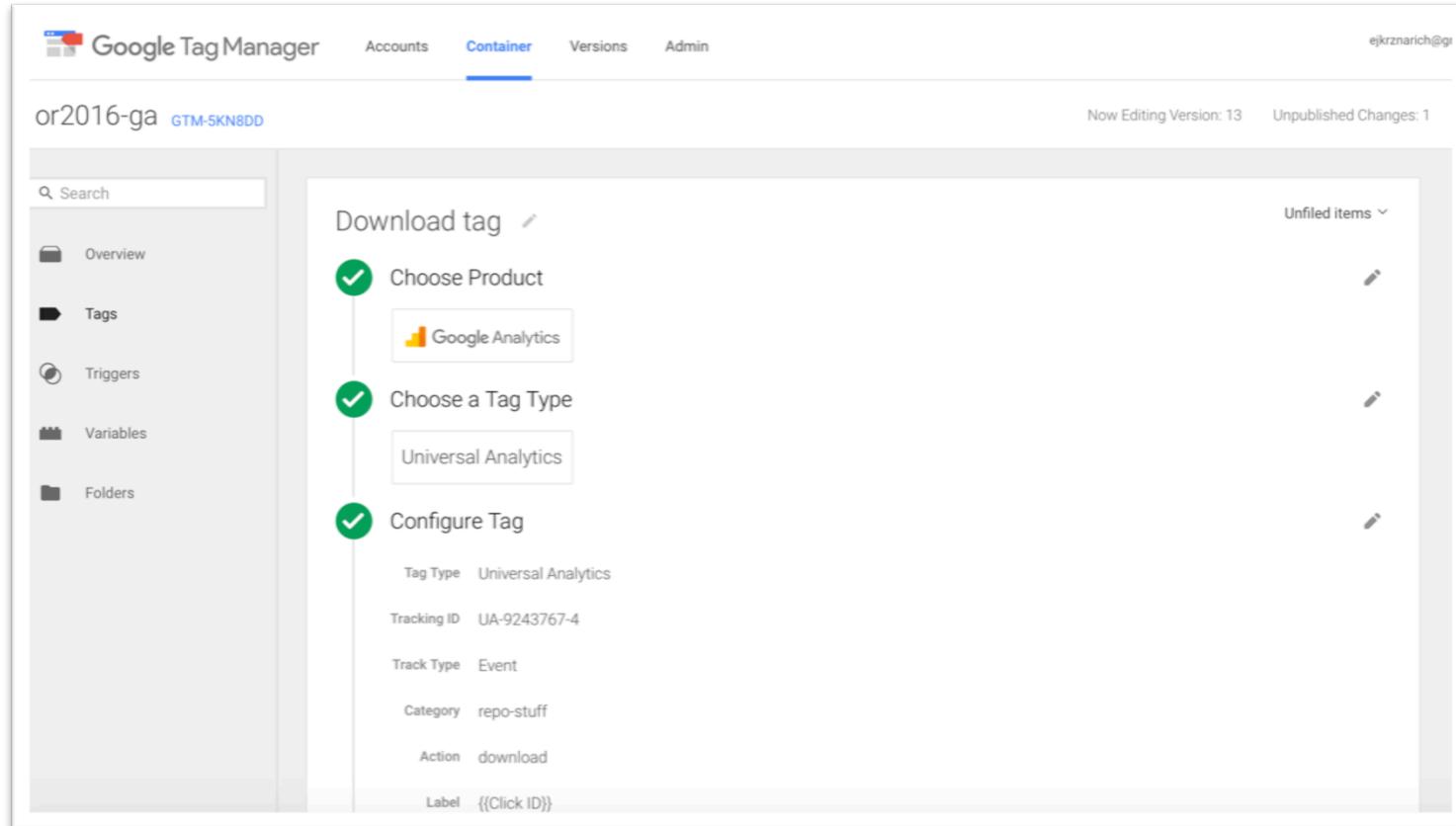
# Pre-reqs: Get a Google Analytics account, create a new project & set up tracking

<https://analytics.google.com>



# Google Tag Manager = easy customized tracking

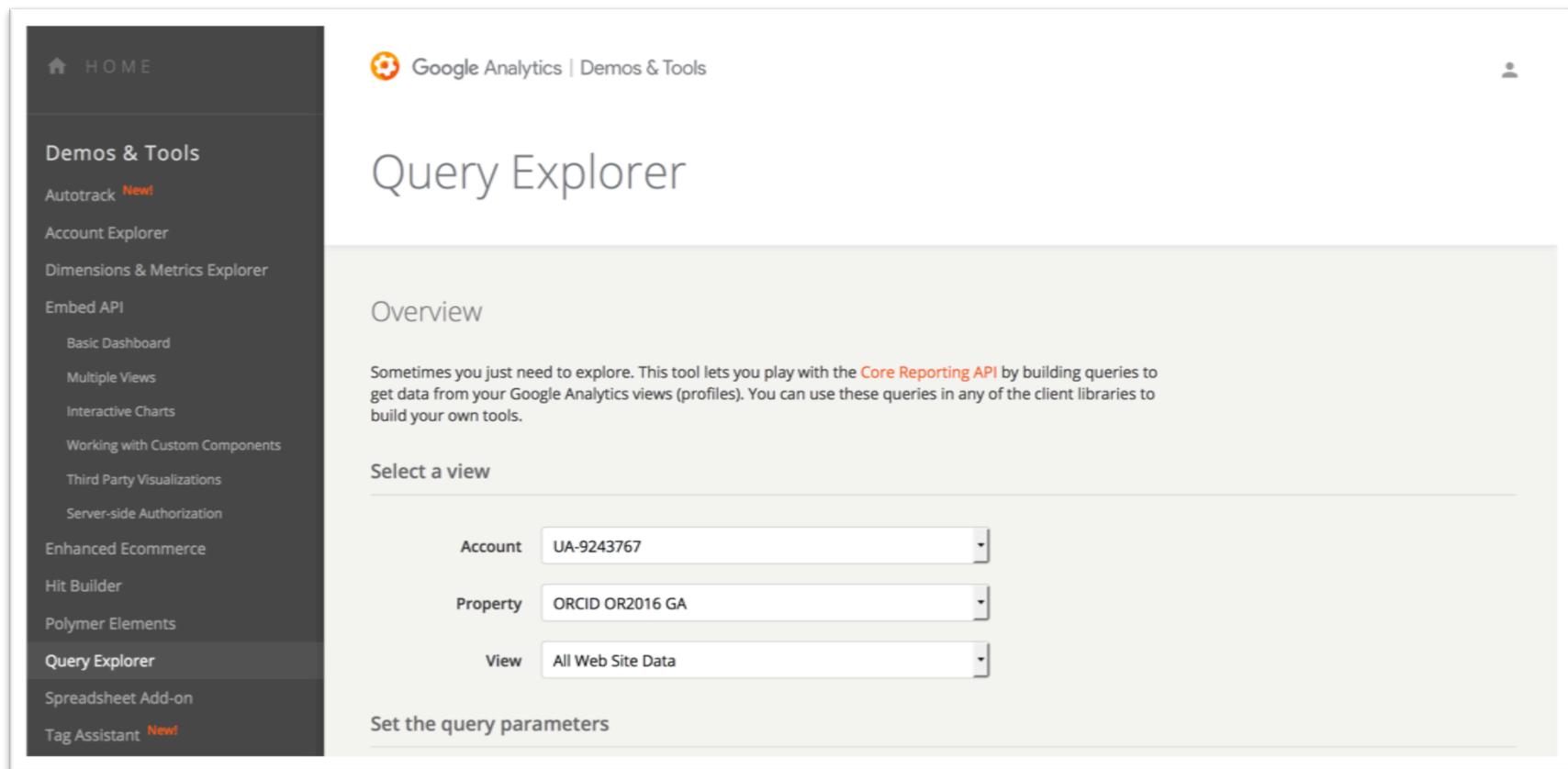
<https://tagmanager.google.com>



# I. Query the Analytics API

# Now for some fun with queries!

<https://ga-dev-tools.appspot.com/query-explorer/>



The screenshot shows the Google Analytics Query Explorer interface. On the left, a dark sidebar menu lists various tools: HOME, Demos & Tools (Autotrack, Account Explorer, Dimensions & Metrics Explorer, Embed API, Basic Dashboard, Multiple Views, Interactive Charts, Working with Custom Components, Third Party Visualizations, Server-side Authorization, Enhanced Ecommerce, Hit Builder, Polymer Elements, Query Explorer, Spreadsheet Add-on, Tag Assistant), and Demos & Tools (New!). The 'Query Explorer' item is highlighted. The main content area has a header 'Query Explorer' and a 'Google Analytics | Demos & Tools' logo. Below the header is a 'Overview' section with text explaining the tool's purpose: 'Sometimes you just need to explore. This tool lets you play with the Core Reporting API by building queries to get data from your Google Analytics views (profiles). You can use these queries in any of the client libraries to build your own tools.' Underneath is a 'Select a view' section with dropdown menus for 'Account' (set to UA-9243767), 'Property' (set to ORCID OR2016 GA), and 'View' (set to All Web Site Data). At the bottom is a 'Set the query parameters' section.

## Dimension & metrics & filters, oh my!

- **Dimensions:** How to break down the data (city, device)
- **Metrics:** What you're counting (clicks, views, etc)
- **Filters:** Limit data by specific criteria

### Analytics API Reference:

<https://developers.google.com/analytics/devguides/reporting/core/v3/reference>

### Dimensions & Metrics Reference:

<https://developers.google.com/analytics/devguides/reporting/core/dimsmets>

## 2. Set up Google API credentials

## Create a new Google Developer project

<https://console.developers.google.com/project>

New Project

Project name [?](#)

Your project ID will be `orcid-or2016-ga` [?](#) [Edit](#)

[Show advanced options...](#)

Please email me updates regarding feature announcements, performance suggestions, feedback surveys and special offers.

Yes  No

I agree that my use of any [services and related APIs](#) is subject to my compliance with the applicable [Terms of Service](#).

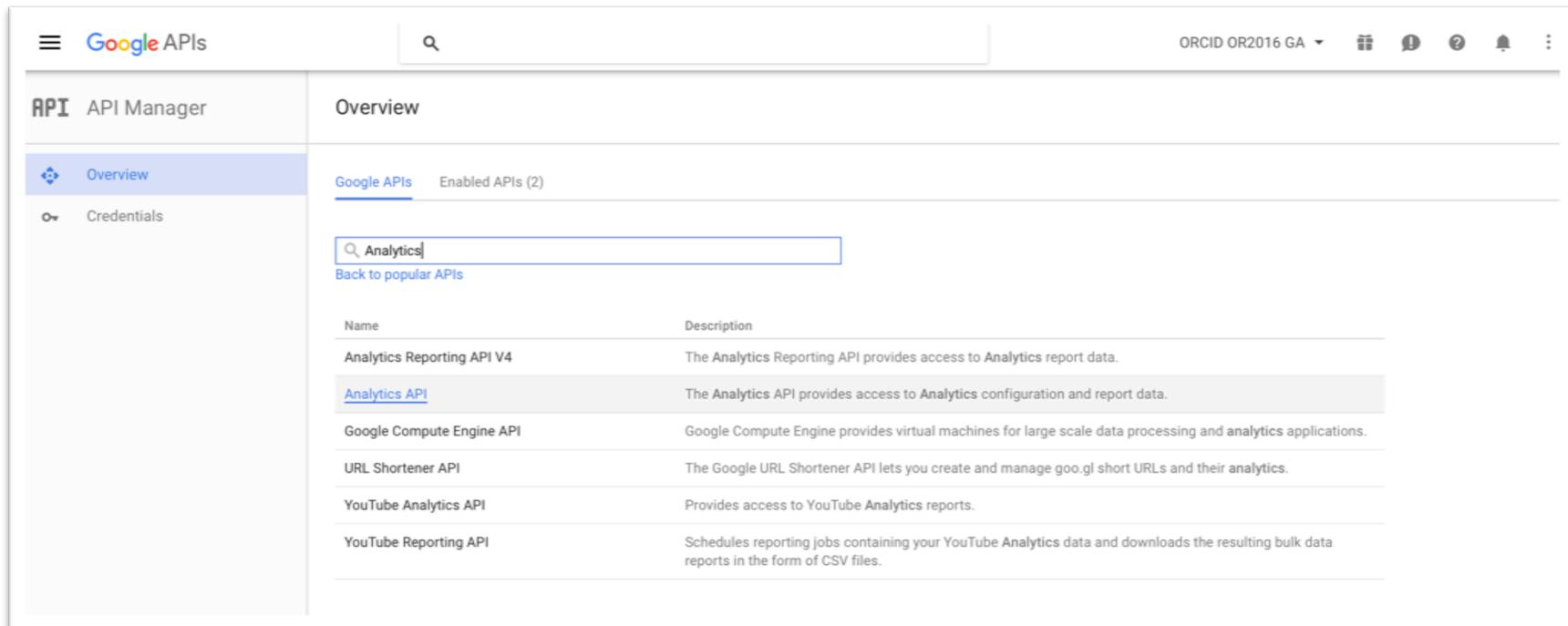
Yes  No

[Create](#) [Cancel](#)

Google Maps Roads API

# Enable APIs (Analytics & Drive)

<https://console.developers.google.com/apis/library>

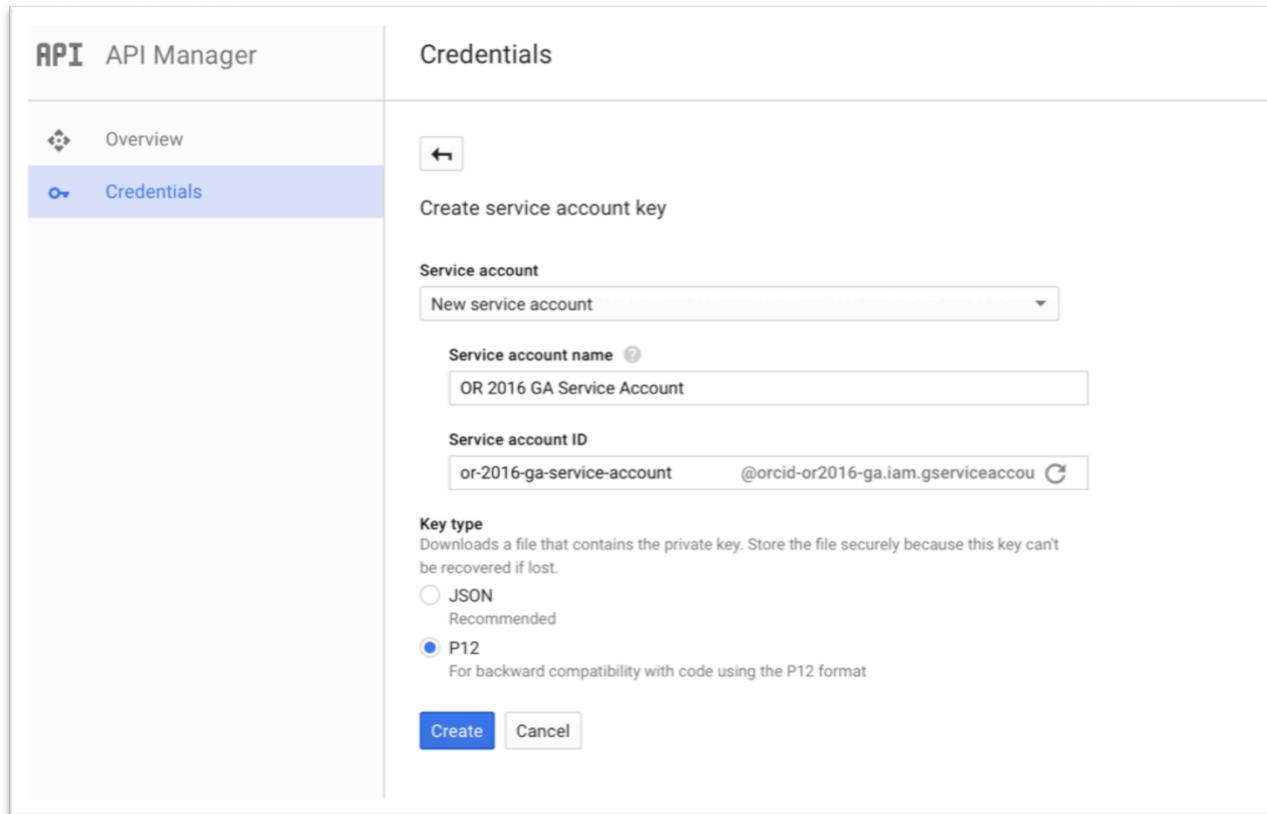


The screenshot shows the Google APIs console interface. The left sidebar has 'API Manager' selected. The main area is titled 'Overview' and shows 'Google APIs' selected. A search bar at the top has 'Analytics' typed into it. Below the search bar, there are two tabs: 'Google APIs' (selected) and 'Enabled APIs (2)'. A table lists the following APIs:

Name	Description
Analytics Reporting API V4	The Analytics Reporting API provides access to Analytics report data.
<a href="#">Analytics API</a>	The Analytics API provides access to Analytics configuration and report data.
Google Compute Engine API	Google Compute Engine provides virtual machines for large scale data processing and analytics applications.
URL Shortener API	The Google URL Shortener API lets you create and manage goo.gl short URLs and their analytics.
YouTube Analytics API	Provides access to YouTube Analytics reports.
YouTube Reporting API	Schedules reporting jobs containing your YouTube Analytics data and downloads the resulting bulk data reports in the form of CSV files.

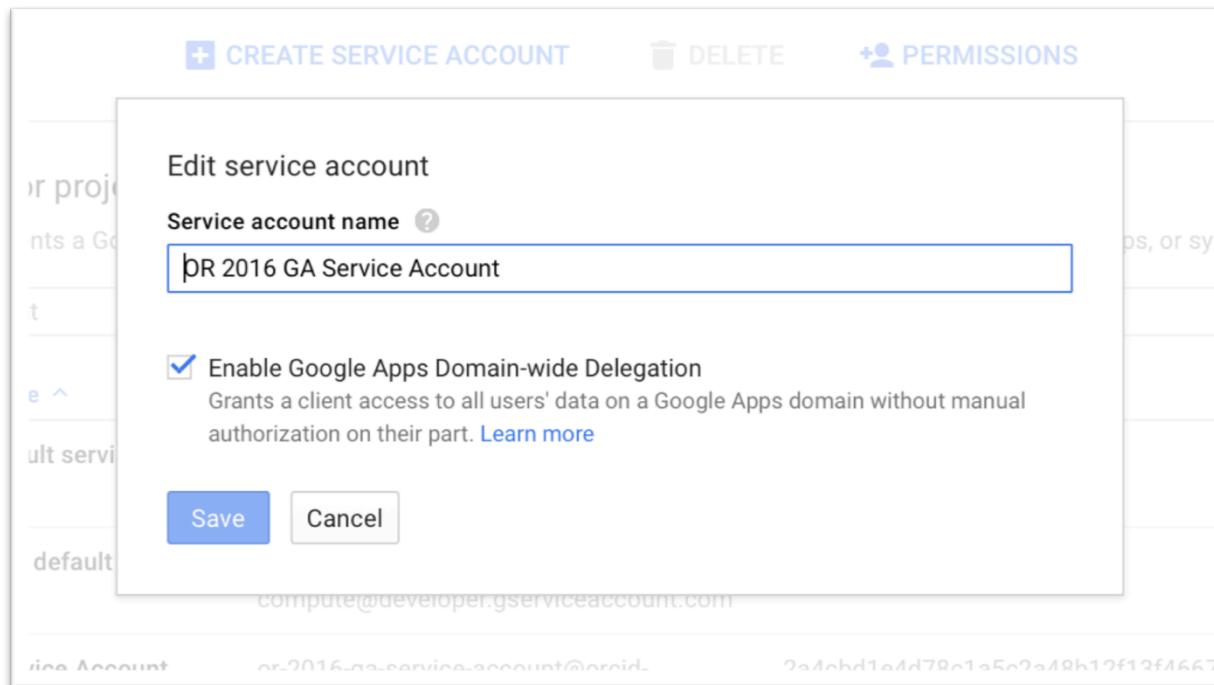
# Create Service Account & download PI2 key

<https://console.developers.google.com/apis/credentials>



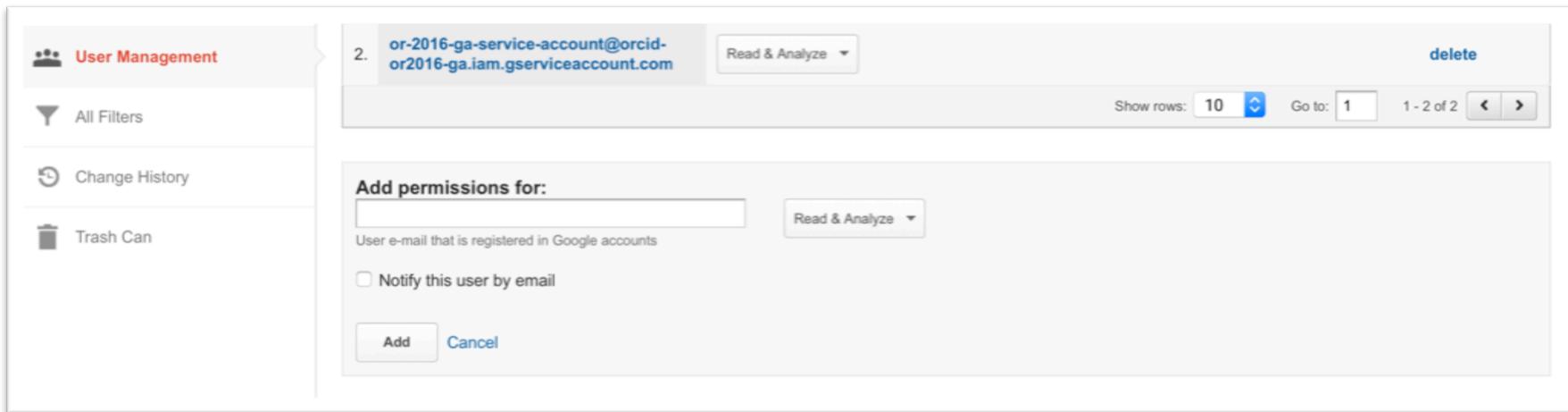
# Enable domain-wide access

<https://console.developers.google.com/iam-admin/serviceaccounts>



## Add Service Account to Analytics

<https://analytics.google.com/analytics/web/#management/Settings> > User Management

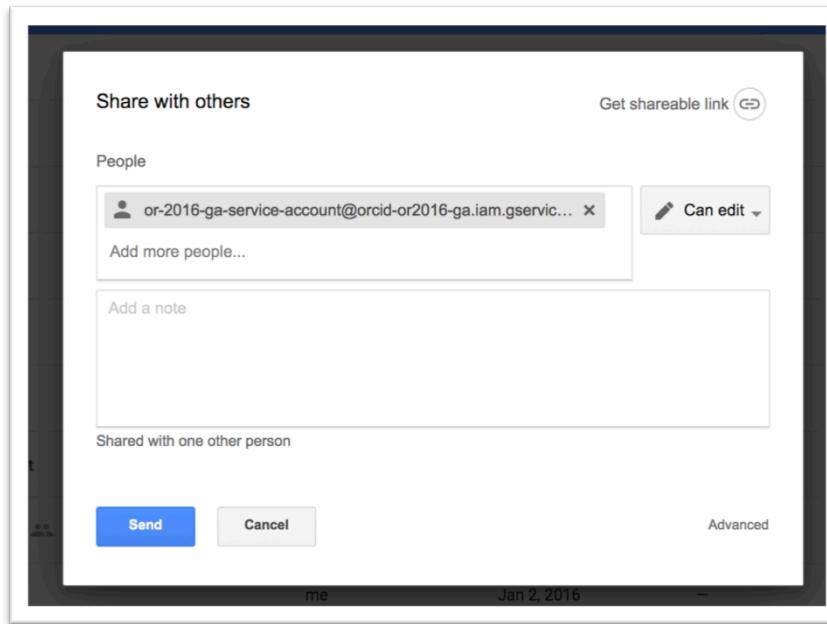


The screenshot shows the Google Analytics User Management interface. On the left, a sidebar has 'User Management' selected. The main area shows a list of users with one entry: 'or-2016-ga-service-account@orcid-or2016-ga.iam.gserviceaccount.com' with a 'Read & Analyze' dropdown and a 'delete' link. Below this is a 'Show rows: 10' and 'Go to: 1' pagination. A modal dialog is open over the list, titled 'Add permissions for:' with a text input field containing 'User e-mail that is registered in Google accounts' and a 'Read & Analyze' dropdown. It includes a checkbox for 'Notify this user by email' and buttons for 'Add' and 'Cancel'.

**Guess what?!** If you're using the DSpace Google Analytics module, you've already done steps 4-8, and you can use the same service account credentials for your custom applications!

# Create a new Drive folder & share it with Service Account

<https://drive.google.com/drive/my-drive>



Whew, that was tedious  
Time for some coding!

### 3. Get Analytics data & upload to Drive

- Import Google oauth2client & apiclient libraries
- Authenticate to Analytics, Drive & Sheets APIs
- Get Analytics data
- Upload Analytics data to Drive

# Authenticate to Analytics, Drive & Sheets APIs (Aaargh!!!)

(via Oauth2 SignedJwtAssertionCredentials)

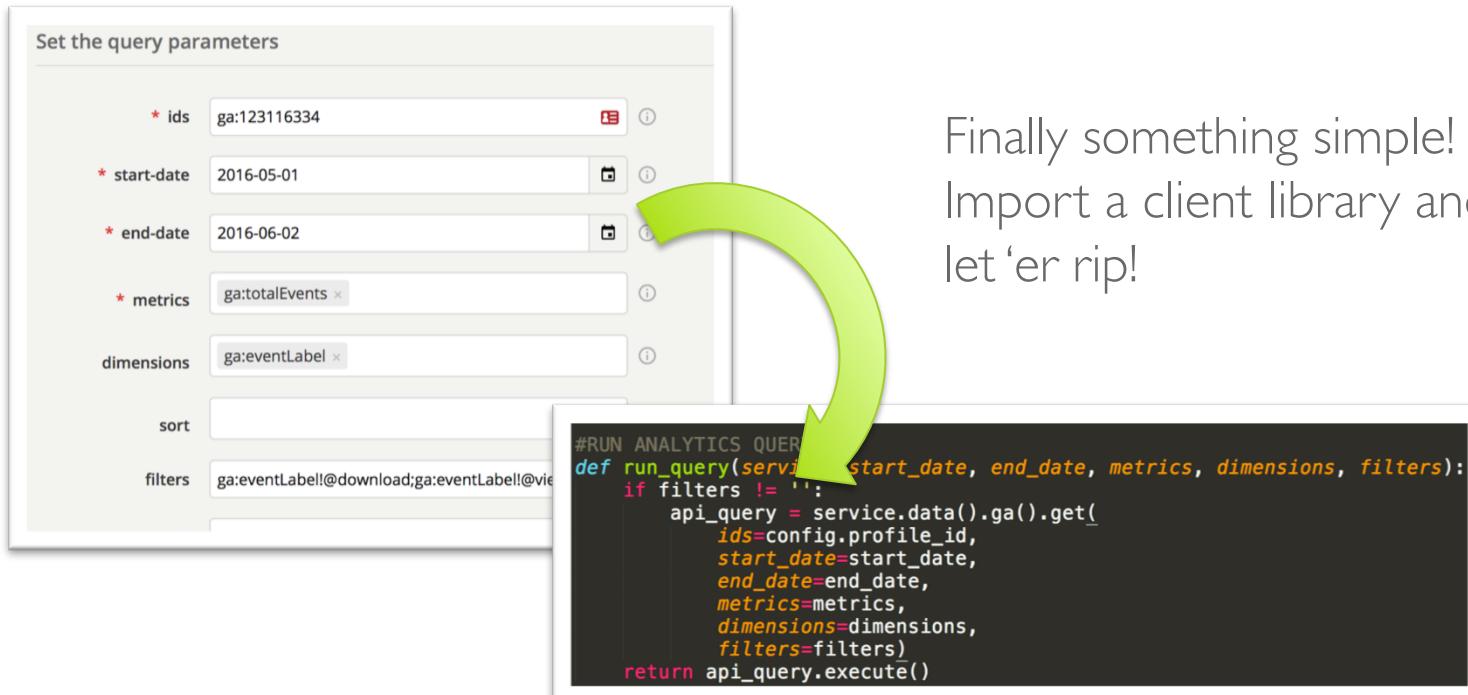
You could read this:

<https://developers.google.com/identity/protocols/OAuth2ServiceAccount#authorizingrequests>

Or just borrow someone else's code.

# Get Analytics Data

<https://developers.google.com/analytics/devguides/reporting/core/v3/coreDevguide>



The image shows a comparison between a user interface for setting Google Analytics query parameters and a corresponding Python code snippet. A large green curved arrow points from the UI on the left to the code on the right, indicating a direct correspondence or flow from configuration to execution.

**Set the query parameters**

- \* ids: ga:123116334
- \* start-date: 2016-05-01
- \* end-date: 2016-06-02
- \* metrics: ga:totalEvents
- dimensions: ga:eventLabel
- sort: (empty)
- filters: ga:eventLabel@download;ga:eventLabel@view

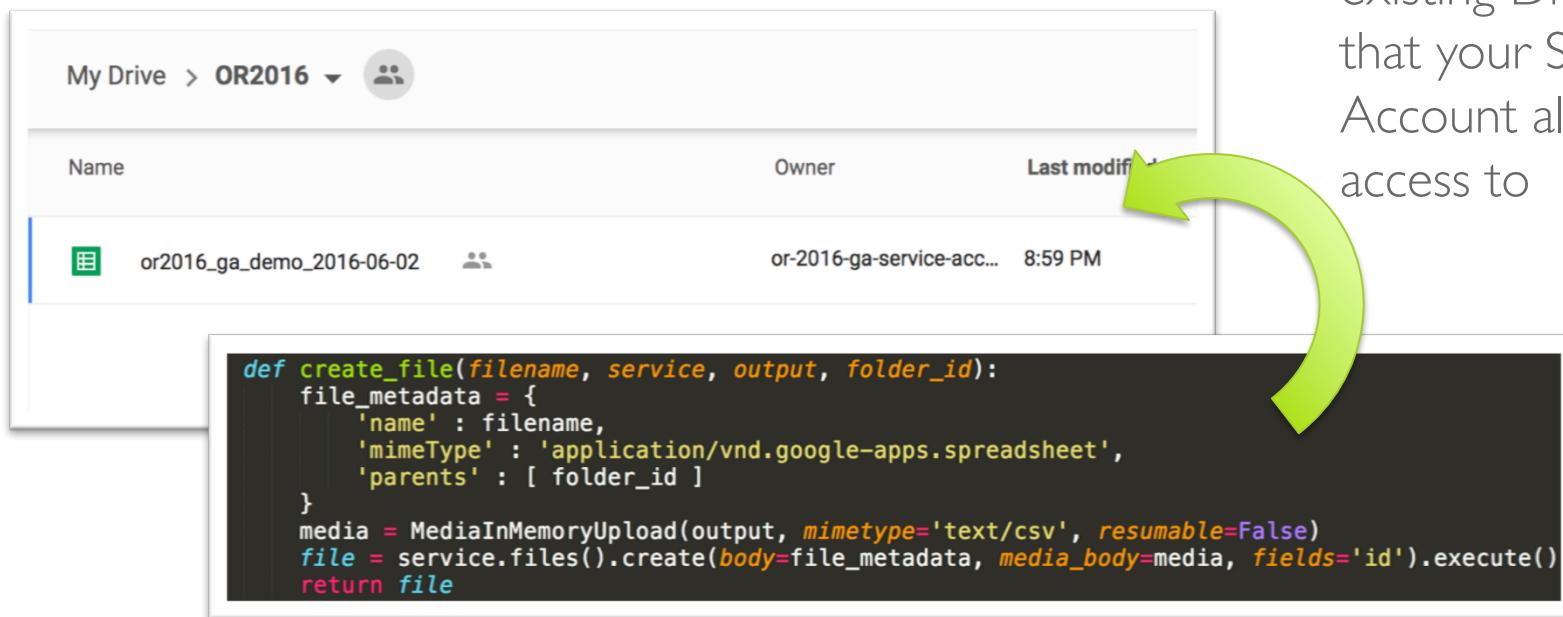
```
#RUN ANALYTICS QUERY
def run_query(service, start_date, end_date, metrics, dimensions, filters):
    if filters != '':
        api_query = service.data().ga().get(
            ids=config.profile_id,
            start_date=start_date,
            end_date=end_date,
            metrics=metrics,
            dimensions=dimensions,
            filters=filters)
    return api_query.execute()
```

Finally something simple!  
Import a client library and  
let 'er rip!

# Upload data to Drive

<https://developers.google.com/drive/v3/reference/files/create>

## files/create



My Drive > OR2016

Name	Owner	Last modified
or2016_ga_demo_2016-06-02	or-2016-ga-service-acc...	8:59 PM

```
def create_file(filename, service, output, folder_id):
    file_metadata = {
        'name' : filename,
        'mimeType' : 'application/vnd.google-apps.spreadsheet',
        'parents' : [ folder_id ]
    }
    media = MediaInMemoryUpload(output, mimetype='text/csv', resumable=False)
    file = service.files().create(body=file_metadata, media_body=media, fields='id').execute()
    return file
```

**Tip:** Start with an existing Drive folder that your Service Account already has access to

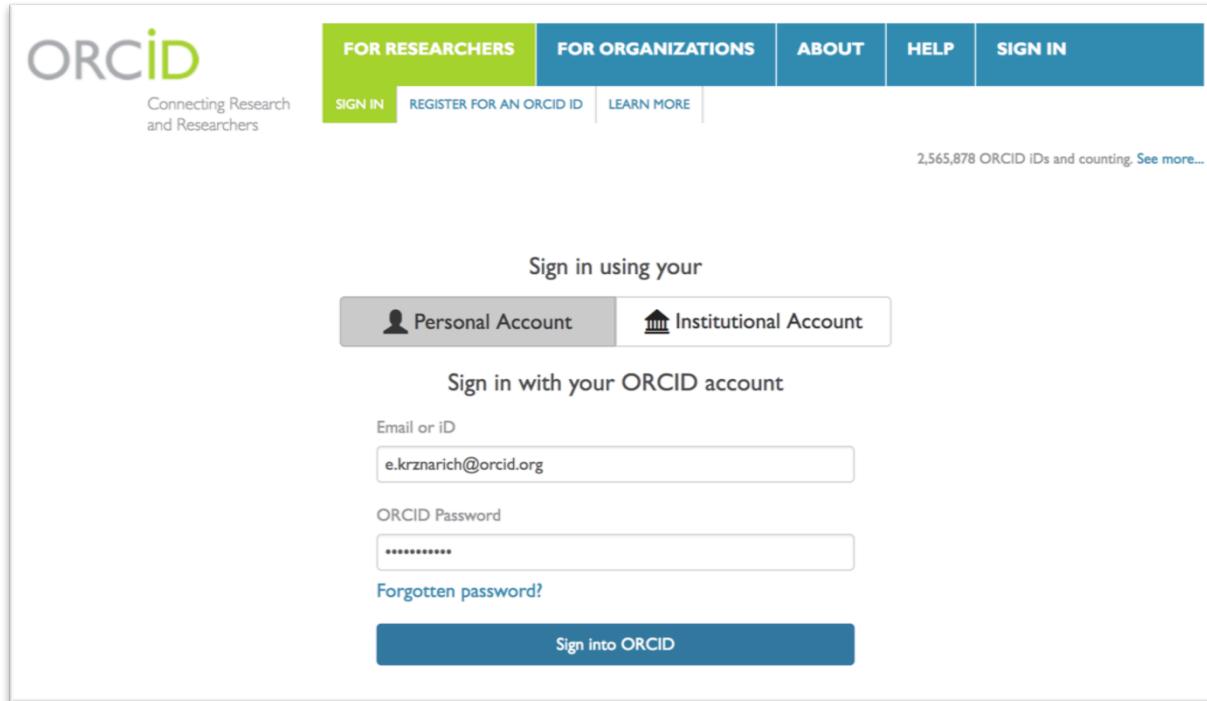
## 4. Set up ORCID API credentials

# ORCID Public vs Member APIs

API Type	Access	Features
Public	Freely available (ORCID iD required)	<ul style="list-style-type: none"><li>Get authenticated ORCID iDs from users</li><li>Search/retrieve public data</li></ul>
Member	Available to ORCID member organizations (sandbox test environment freely available)	<p><i>All Public API features, plus:</i></p> <ul style="list-style-type: none"><li>Search/retrieve "limited-access" data</li><li>Add to/update ORCID records</li><li>Register webhooks</li></ul> <p><i>Premium members only</i></p>

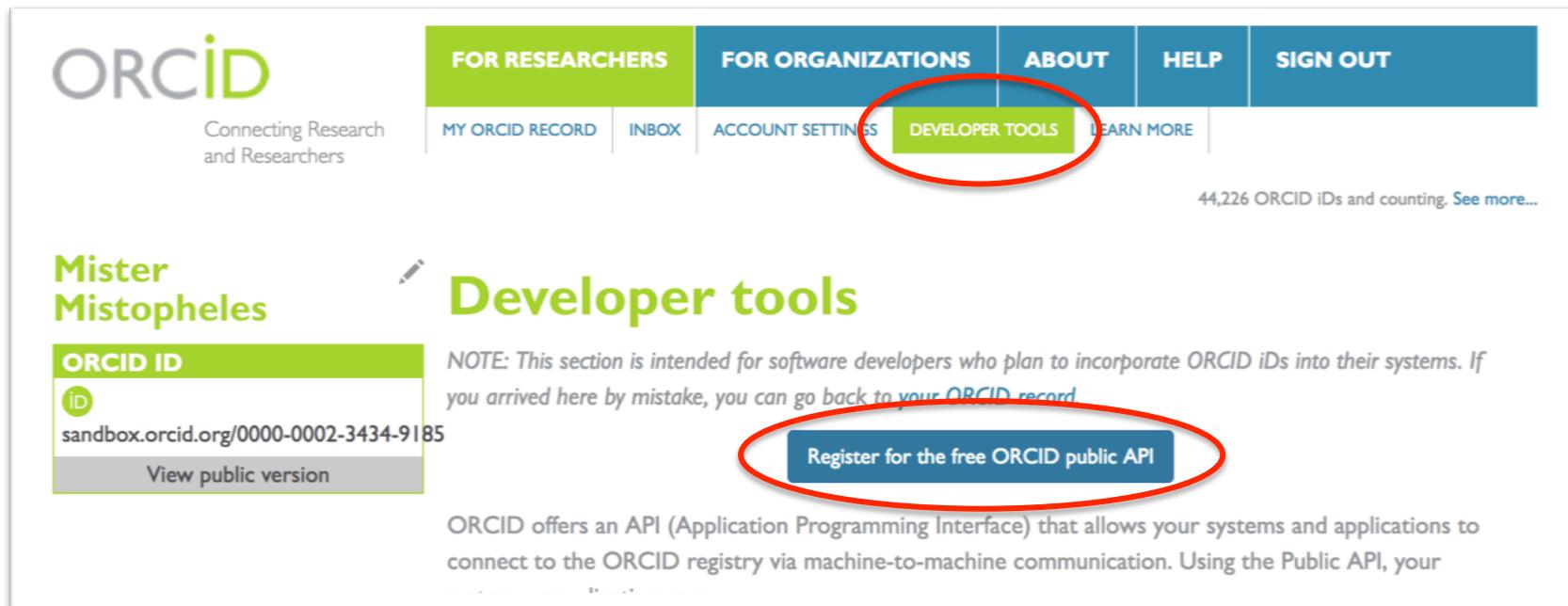
# Log in (or register for an ORCID iD)

<https://orcid.org/signin>



Click Developer Tools > Register...

<https://orcid.org/developer-tools>



ORCID

Connecting Research and Researchers

FOR RESEARCHERS FOR ORGANIZATIONS ABOUT HELP SIGN OUT

MY ORCID RECORD INBOX ACCOUNT SETTINGS DEVELOPER TOOLS LEARN MORE

44,226 ORCID IDs and counting. [See more...](#)

Mister Mistopheles

Developer tools

NOTE: This section is intended for software developers who plan to incorporate ORCID IDs into their systems. If you arrived here by mistake, you can go back to your [ORCID record](#).

Register for the free ORCID public API

ORCID offers an API (Application Programming Interface) that allows your systems and applications to connect to the ORCID registry via machine-to-machine communication. Using the Public API, your

# Enter your client app info

44,226 ORCID IDs and counting. [See more...](#)

**Mister  
Mistopholes**

**ORCID ID**

 [sandbox.orcid.org/0000-0002-3434-9185](https://sandbox.orcid.org/0000-0002-3434-9185) See the [Public client terms of service](#).

[View public version](#)

**Developer tools**

*NOTE: This section is intended for software developers who plan to incorporate ORCID IDs into their systems. If you arrived here by mistake, you can go back to [your ORCID record](#).*

You have enabled the public API

**Name of your application (Shown to users)**  
My fancy app

**Your website URL**  
<http://myfancysite.org>

**Description of your application**  
My fancy app will connect [myfancysite.org](#) to ORCID

**Redirect URIs**



[+ Add another redirect URI](#)



# Get your client ID & client secret

44,226 ORCID IDs and counting. [See more...](#)

**Mister  
Mistopheles**

**ORCID ID**

ID  
sandbox.orcid.org/0000-0002-3434-9185 [See the Public client terms of service.](#)

[View public version](#)

**Developer tools**

*NOTE: This section is intended for software developers who plan to incorporate ORCID IDs into their systems. If you arrived here by mistake, you can go back to your ORCID record.*

**My fancy app**

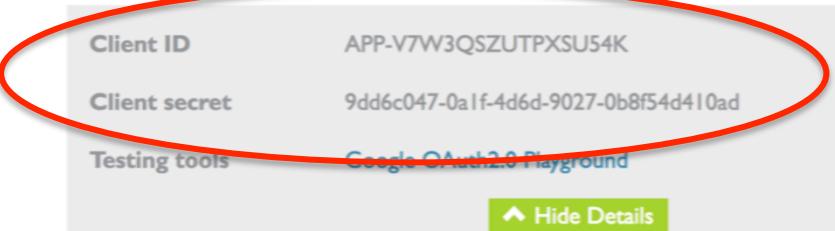
<http://myfancysite.org>

My fancy app will connect myfancysite.org to ORCID

**Redirect URIs:**

<b>Client ID</b>	APP-V7W3QSZUTPXSU54K
<b>Client secret</b>	9dd6c047-0a1f-4d6d-9027-0b8f54d410ad
<b>Testing tools</b>	<a href="#">Google OAuth 2.0 Playground</a>

[▲ Hide Details](#)



## 5. Get data from ORCID & add it to our Drive file (using Sheets API)

## Get data from ORCID

Now we'll use the credentials we created to query the ORCID API for information about the DOIs we're tracking in Analytics.

Specifically, we'll ask the ORCID API which ORCID iDs are connected to each of our DOIs.

## First, we need an access token

Valid for 20yrs – can be used for multiple searches

<https://members.orcid.org/api/tutorial-searching-data-using-api>

```
curl -i -L -H "Accept: application/json" \
-d "client_id=APP-01XX65MXBF79VJGF" \
-d "client_secret=3a87028d-c84c-4d5f-8ad5-38a93181c9e1" \
-d "scope=/read-public" \
-d "grant_type=client_credentials" \
"https://sandbox.orcid.org/oauth/token"
```

**NOTE! This is just a plain HTTP request...for a Python code example, see  
[github.com/lizkrznarich/orcid-demo](https://github.com/lizkrznarich/orcid-demo)**

Then, use the token to send queries

Solr query syntax - see

<https://members.orcid.org/api/tutorial-searching-data-using-api>

```
curl -H "Content-Type: application/orcid+xml" \
-H "Authorization: Bearer 5d0c7545-f59f-44c5-bcde-
f83438beb3ef" \
"https://pub.orcid.org/v1.2/search/orcid-bio/?
defType=edismax&q=digital-object-ids:%2210.1087/20120404%22"
```

**NOTE! This is just a plain HTTP request...for a Python code example, see  
[github.com/lizkrznarich/orcid-demo](https://github.com/lizkrznarich/orcid-demo)**

## Edit the new sheet to add your data

<https://developers.google.com/sheets>

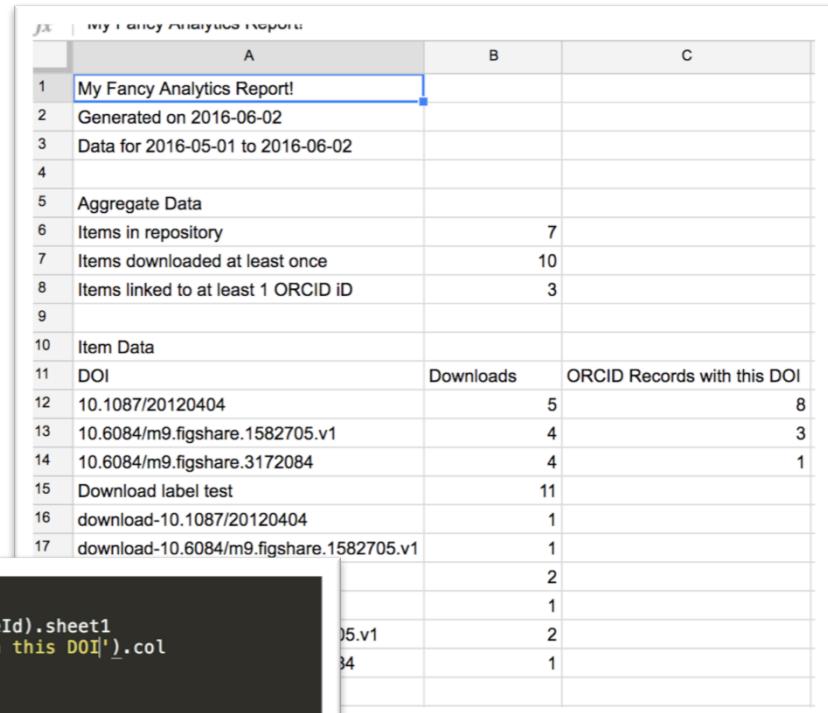
<http://gspread.readthedocs.io>

- Python gspread library makes life easy
- Sheets API v3 w/out a 3<sup>rd</sup> party library = frightening; v4 looks friendlier

# Edit the new sheet to add your data

1. Get the Drive file id
2. Open the file
3. Figure out which cells to edit & send the update request

```
def edit_spreadsheet(sheets_client, drive_file, orcid_data):  
    drive_file_fileId = drive_file.get('id')  
    drive_file_worksheet = sheets_client.open_by_key(drive_file_fileId).sheet1  
    orcid_record_col = drive_file_worksheet.find('ORCID Records with this DOI').col  
    doi_orcid_count = 0  
    for doi in orcid_data:  
        try:  
            doi_match_row = drive_file_worksheet.find(doi[0]).row  
            drive_file_worksheet.update_cell(doi_match_row, orcid_record_col, doi[1])  
            doi_orcid_count += 1  
        except:  
            pass  
  
    total_linked_orcid_row = drive_file_worksheet.find('Items linked to at least 1 ORCID iD').row  
    drive_file_worksheet.update_cell(total_linked_orcid_row, 2, doi_orcid_count)
```

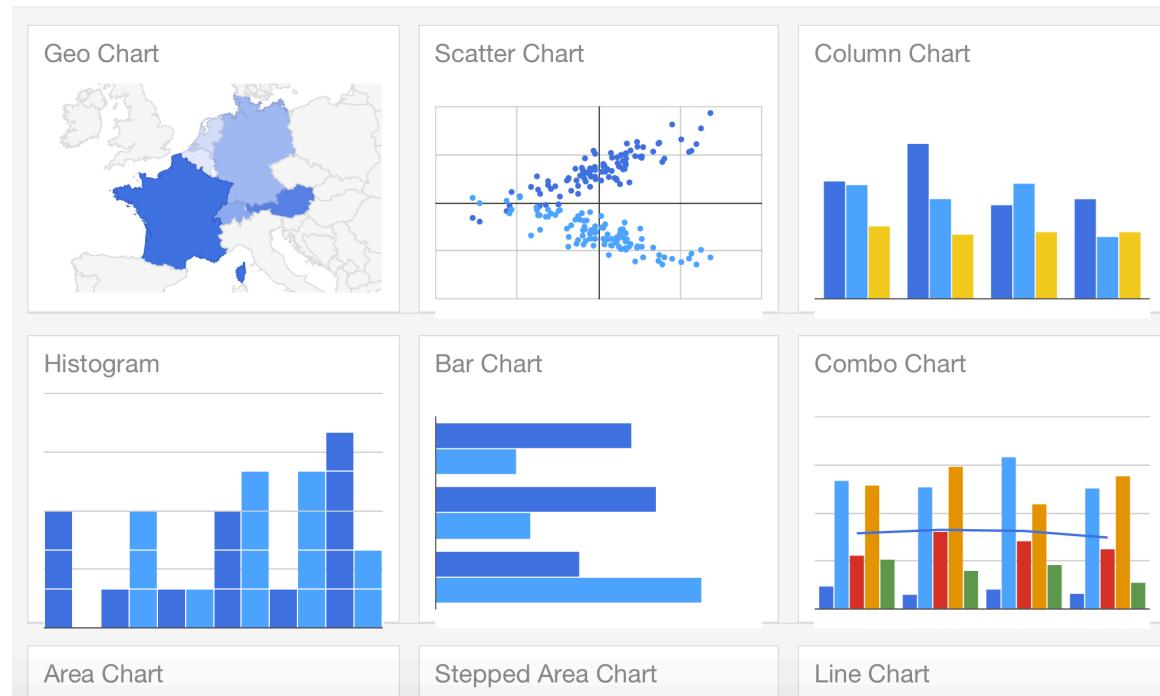


The screenshot shows a Google Sheets document with the following data:

	A	B	C
1	My Fancy Analytics Report!		
2	Generated on 2016-06-02		
3	Data for 2016-05-01 to 2016-06-02		
4			
5	Aggregate Data		
6	Items in repository	7	
7	Items downloaded at least once	10	
8	Items linked to at least 1 ORCID iD	3	
9			
10	Item Data		
11	DOI	Downloads	ORCID Records with this DOI
12	10.1087/20120404	5	8
13	10.6084/m9.figshare.1582705.v1	4	3
14	10.6084/m9.figshare.3172084	4	1
15	Download label test	11	
16	download-10.1087/20120404	1	
17	download-10.6084/m9.figshare.1582705.v1	1	
		2	
		1	
	05.v1	2	
	34	1	

Now, take that data and do cool stuff!  
(like charts with Google Charts API)

<https://developers.google.com/chart>



# Hot Tips!

- Tokens expire in 1hr (no refresh token...just get a new one)
- APIs are twitchy – expect sporadic errors when running lots of queries
- Sheets API is sloooooow – try to combine as many actions as possible into 1 request

# Get the code

Demo site

<http://orcid.github.io/analytics-demo>

Code

<https://github.com/lizkrznarich/orcid-demo>

# THANK YOU!

Questions? [e.krznarich@orcid.org](mailto:e.krznarich@orcid.org)

<https://github.com/lizkrznarich>

<http://orcid.org/0000-0001-6622-4910>