

Hey everyone!

We're revising the slides

This version last updated Friday
4/28 at 5pm

You can use them now, but
remember to check back, they're
going to improve.

And remember the [GitHub](#)!

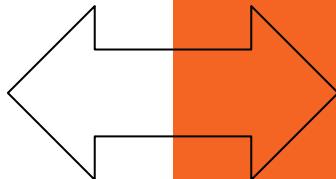
(Take home documentation *is*
coming, we promise)

Mac users

Please sit together on this side of the room

PC users

Please sit together on this side of the room



Everyone

Please connect to the wifi
Create a new folder named “vagrant” on your desktop
Copy contents of flash drive to this folder

There's an API for That

MARAC
April 20th, 2017

Lora J. Davis
Valerie Addonizio
Johns Hopkins University

Introduction

Workshop aims

1.

Practical, real life application

Info-packed take-home

Resource-packed GitHub

2.

Assurances:

You are not alone

I didn't get into this field for this
either (but some of you may have!)

This is difficult and frustrating

You didn't miss a flight; [the airport turned into a space launch while you were in the parking lot](#)

The 1-up learning experience

You can't learn it all and we can't teach it

Are you a 1?

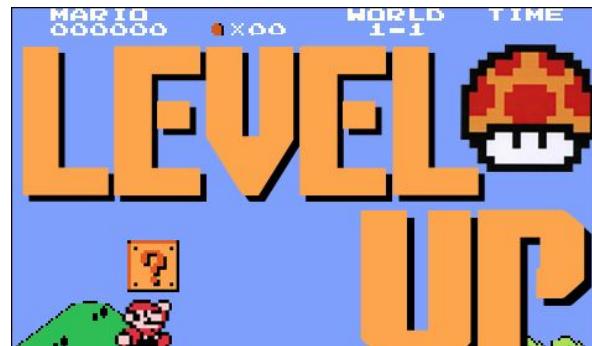
We hope you'll leave a
2.

Are you a 2?

We hope you'll leave a
3.

Are you a 3?

We hope to give you
new ideas, scripts, and
momentum.



The Odd Couple

We are stronger and smarter together

Lora

- Digital Archivist
- Never met a task she didn't want to solve with Python/Ruby
- Thinks hanging with the Bmore on Rails group sounds like a fun night out
- Helped migrate two institutions to ArchivesSpace and found it thrilling
- Enjoys craft beer, cross stitch, and the Pittsburgh Steelers
- Knows how to programmatically manipulate data
- CATS YAY!



Valerie

- (Former) Photo Archivist
- Never met a task she didn't want to solve with an MS Access database
- Knows rails are either “narrow gauge” or “standard” because ❤️ steam trains
- Helped migrate to ASpace and felt like a fish climbing a tree
- Enjoys geocaching, hiking, and planning ambitious camping trips
- Knows data modeling and systematic analysis of the *really hard* problems
- CATS YAY!



What does API stand for?

Application

As in a computer application, like Word or Chrome

Programming

As in computer “programming,” or taking steps to make a computer do something you want it to do

Interface

As in the place where two systems meet

What do APIs do?

As the prior slide suggests, APIs make it possible for **applications** to interact (or interface) with one another.

APIs are **not new**, and there are **many types** of APIs.

When you copy content from a Word document to your clipboard, then paste that content into an Outlook e-mail, it works because your computer operating system, which both your versions of Word and Outlook are programmed to run on, uses an API to **allow the interchange of information**.

APIs tell software developers the **rules of the road** that they must follow if they want their applications to play well with others.

That's not at all what I thought an API was!

Though anything that allows an interchange of information between two applications is *technically* a form of an API, what we typically mean today when we say “API” is a very specific thing.

That thing is a **web API**.

Ok, so what is a *web API*?

Complicated: A RESTful API is an **application program interface (API)** that uses HTTP requests to GET, PUT, POST and DELETE data.

Simple: You access it over the web, using URL-like directions, and are limited to 3-4 simple commands or activities.

For more: <http://searchcloudstorage.techtarget.com/definition/RESTful-API>

Extra nerdy sidebar:

- Web APIs also come in several flavors, including **SOAP** and **REST**.
- We're going to be exclusively working with **RESTful APIs** today, as they're far more prevalent in archives/libraries technologies.
- REST stands for “**representational state transfer**” and was defined in 2000 in a doctoral dissertation by Roy Fielding.
- REST essentially dictates how an application should be able to **textually interact** with a web service.

Vocabulary pitstop: API Terms

- **GET**, **POST**, and **DELETE** are the three cornerstone commands for a RESTful API
- We will use these terms throughout
- Think of them as View, Save, and of course, Delete
- All APIs allow GETs, some let you POST, and few allow you to Delete
 - ASpace does all three, but allows you to tailor permissions for each

I'm not an application, I'm an archivist!

Why should I care?

As librarians and archivists with collection descriptions and/or collections themselves on the web, you probably **do** care about being able to **access** and **meaningfully manipulate** textual data on the **web at scale**.

In many of the exercises we will work through together today, **you** are, in fact, one of the “applications” interfacing with web-based data.

API possibilities

Get data out → Do something to it → Put it back in

JSON
MARC21
Any standardized
data

Access
OpenRefine
XSLT
Custom script (your choice)
Find and Replace
Hand encoding, copy and
pasting, glue and popsicle
sticks, *whatever it takes!*

*needle coming off
the record*

This is the tough
part.

Questions???

Setting up your tech: A long-ish pitstop

(We promise you're in the right workshop)

Technical pitstop: The (FREE) Applications



Atom

- A text editor that is handy for interacting with JSON, scripts, and all sorts of structured data
- Can utilize additional packages to customize to your needs (e.g. a JSON “linter”)



Postman

- A GUI application for interacting with APIs



VirtualBox

- A virtualization application that lets you run another machine (a “guest”) within your existing computer (the “host”)
- We’re not using this in this workshop other than as a dependency for Vagrant, but you should check it out!



Vagrant

- A way to build and manage a virtual machine
 - We’re using it to set you all up with a test instance of ArchivesSpace
-

Technical pitstop: The Command Line

- While this is NOT the purpose of this workshop, we need to pause to learn a few very quick things about navigating the command line
 - For a fuller introduction to navigating the command line, see:
<https://www.codecademy.com/learn/learn-the-command-line>
- Common commands we will use today:
 - up arrow - shortcut to repeat the last command
 - ls - “listing,” or “tell me what’s in this directory”
 - (Note: This works for Windows users today because we’re using a Unix emulator, Cygwin. Natively, Windows users should use “dir.”)
 - cd - “change directory,” or “let’s move somewhere else”
 - pwd - “print working directory,” or “tell me exactly where I am”

Technical Pitstop: Scripting set-up

Technical Pitstop - Scripting set-up

We are about to:

- Show you a quick **shortcut**
- Get some important packages **installed** on your machines
- **Clone our GitHub repository** (download some scripts to your computer)

Caveats:

- This *is* scary, **we will lose some of you**
- You *do* need to know this, but **we have other tofu to fry**
- Use these slides if you need to set up your own workstations **back at the office**
- There is some other advice and more in-depth walkthroughs in the **take-home document**

Housekeeping:

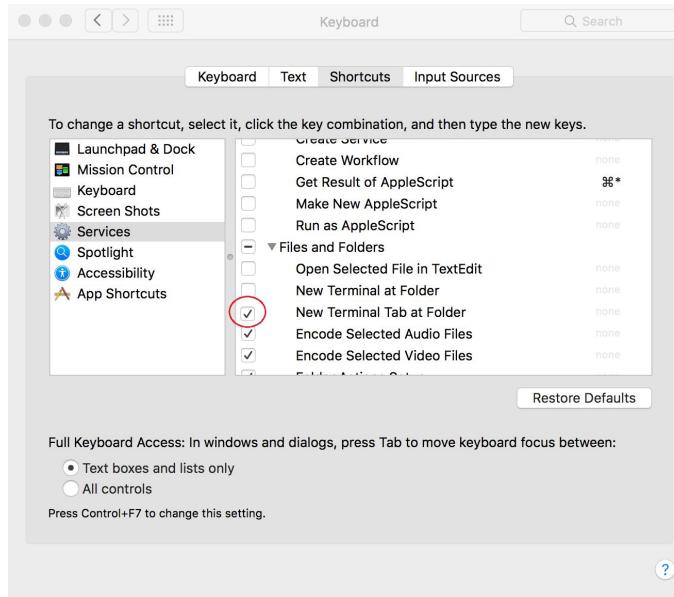
- With 30 people in this class, we've got 30 different environments to troubleshoot and 30 different opportunities to fail - **please be patient!**
- If, at any point for the remainder of the workshop, you need **assistance**, please place a **hot pink post-it** on the back of your laptop screen

Technical Pitstop - Scripting set-up

This is a handy shortcut that you'll need later

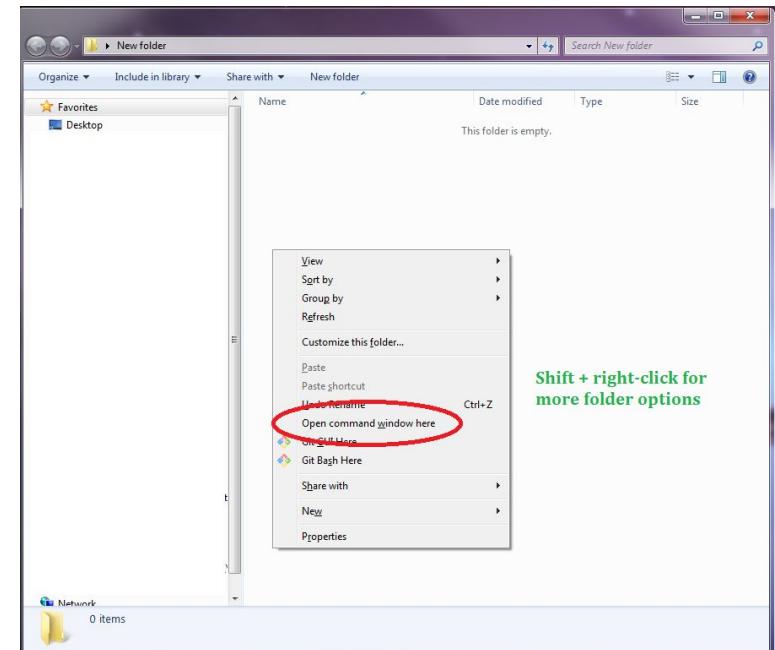
Mac

System Preferences > Keyboard > Shortcuts
tab > Services



PC

Shift + right click

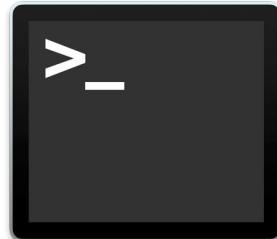


Shift + right-click for
more folder options

Technical Pitstop - Scripting set-up

Mac

1. Please open the terminal:
 - Use spotlight search to search for “Terminal”
 - OR, open your Applications folder, then open the Utilities folder. Open the Terminal application.



PC

1. Please navigate to <http://cygwin.com/install.html>
(that's c-y-g-w-i-n)
2. Download and bring up the installer



Technical Pitstop - Scripting set-up

Mac

1. Type `gcc --version` and hit enter

If you are prompted to install, hit Install!
Once complete, type `gcc --version` again.

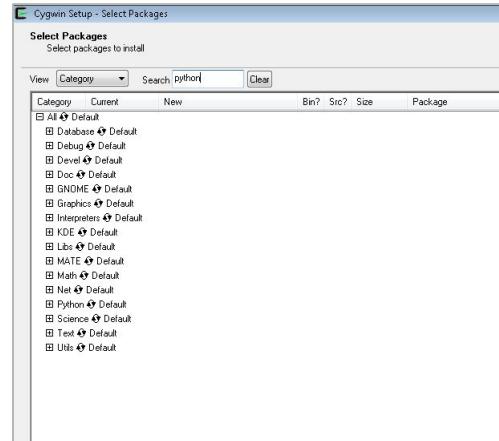
If you see the following, you're good to go.
Leave the terminal open.

```
$ gcc --version
# After installing, check the gcc --version again.
# You should get the below message:

Configured with: --prefix=/Applications/Xcode.app/Contents/Developer/usr --with-gxx-include-dir=/usr/include/c++
Apple LLVM version 8.0.0 (clang-800.0.38)
Target: x86_64-apple-darwin15.6.0
Thread model: posix
InstalledDir: /Applications/Xcode.app/Contents/Developer/Toolchains/XcodeDefault.xctoolchain/usr/bin
```

PC

1. Start downloading. Say yes to everything
2. Pick any site to install from and continue
3. Stop when you get to this screen:



Technical Pitstop - Scripting set-up

Mac

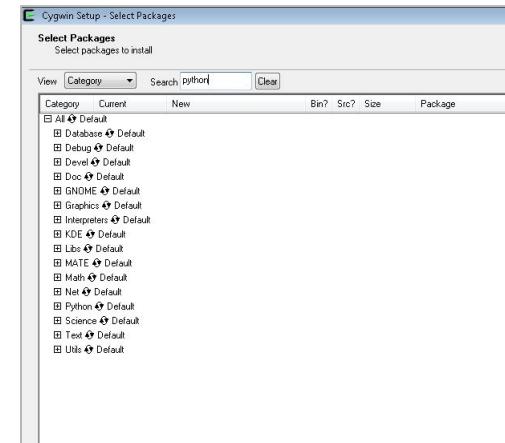
1. Open a browser and navigate to:
brew.sh
(yup that's a webpage)
2. Copy the long command and paste into terminal. Hit enter.



3. Enter your password if needed

PC

1. Start downloading. Say yes to everything
2. Pick any site to install from and continue
3. Stop when you get to this screen:



Technical Pitstop - Scripting set-up

Mac

1. Type `brew install python`
2. Type `python --version` and be sure

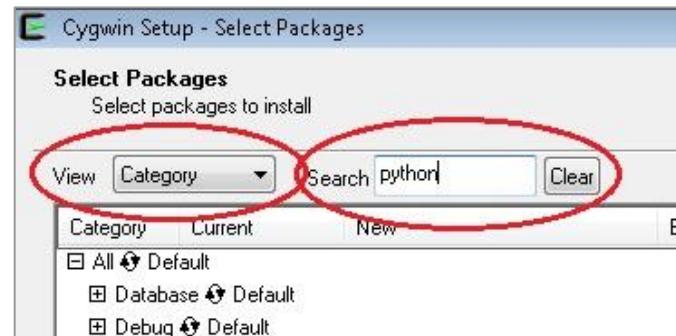
Python appears:

```
Loras-MacBook-Pro:~ lorajdavis$ python --version
Python 2.7.10
Loras-MacBook-Pro:~ lorajdavis$
```

3. Type `pip install requests`

PC

1. Look for this at the top
View: Category
Search for “python”



Windows users: Install packages

1. When you get to the Select Packages window, search for Python then *unskip*:
 - a. python2: Python 2 language interpreter **5,873k**
 - b. python2-requests: Python HTTP/1.1 request module **84k**
2. Search for git, expand Devel, *unskip*:
 - a. git: Distributed version control system **4,936k**
3. Search for openssh, under Net, *unskip*:
 - a. openssh: The OpenSSH server and client programs **750k**
4. Next > Next

<input checked="" type="checkbox"/> Net 	Skip	n/a	n/a	38k gnome-python: Python GNOME platform bindings (meta-package)
<input checked="" type="checkbox"/> Python 	Skip	n/a	n/a	33k gnome-python-desktop: Python GNOME Desktop bindings
	Skip	n/a	n/a	16k gnome-python-extras: Python GNOME extras bindings
	Skip	n/a	n/a	21k net-snmp-python: Net-SNMP (python)
	2.7.10-1	<input checked="" type="checkbox"/>	<input type="checkbox"/>	5,926k python: Python language interpreter
	Skip	n/a	n/a	5k python-avahi: mDNS/DNS_SD/Zeroconf implementation (Python)
	Skip	n/a	n/a	243k python-avogadro: Molecular editor and modeling system (Python)
	Skip	n/a	n/a	6k python-backports.ssl_match_hostname: SSL hostname verification
	Skip	n/a	n/a	61k python-beautifulsoup: Python HTML/XML parser
	Skip	n/a	n/a	49k python-bonobo2: Python bindings for the GNOME Bonobo library
	Skip	n/a	n/a	60k python-botan: C++ crypto library (Python bindings)
	Skip	n/a	n/a	1,211k python-boto: Python interface for AWS and other clouds
	Skip	n/a	n/a	172k python-bdbdb3: Python Berkeley DB bindings

Technical Pitstop - Scripting set-up

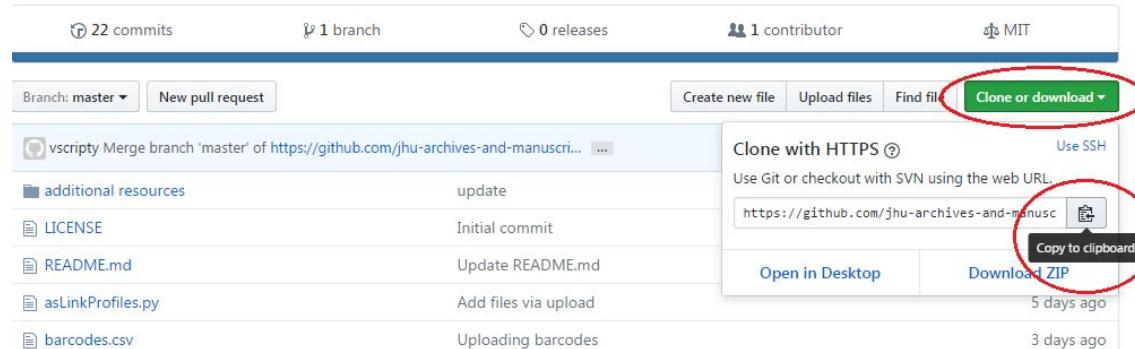
Everyone

1. Open Atom 
2. Along the menu bar, select File > Settings > Install > Install Packages > search for “beautify” and install
3. Check that it installed under Packages in the menu bar

Technical Pitstop - Scripting set-up

Everyone

1. Navigate to GitHub.com
2. Search for “MARAC API Workshop”
3. **Bookmark! You will need this again.**
4. Then click the **green** button, and copy the path to the clipboard



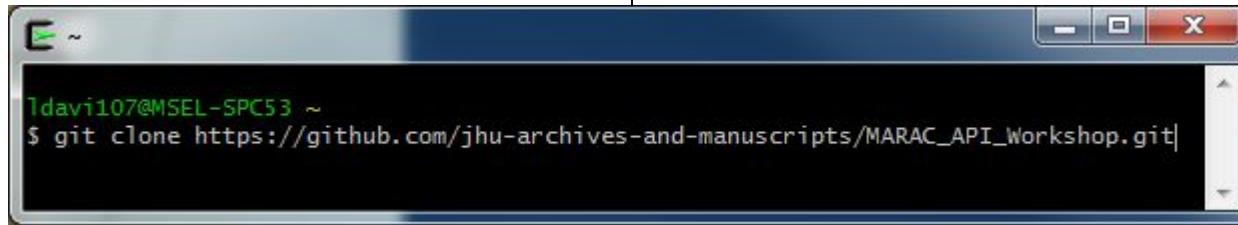
Technical Pitstop - Scripting set-up

Mac

1. Ctrl+click on Desktop > New Terminal at Folder
2. Type `git clone` [ctrl-click > paste]

PC

1. Open cygwin
2. Type `git clone` [right-click > paste]



Now you have a folder (either on your Desktop, or in your Cygwin home directory) that contains all the materials you'll need for the rest of today's workshop.

This folder, titled “**MARAC_API_Workshop**,” is a *direct clone* of what you see in your browser on [github.com](https://github.com/jhu-archives-and-manuscripts/MARAC_API_Workshop).



JUST
breathe

15 minute break!

GET

API possibilities

Get data out → Do something to it → Put it back in



GET with GUI - ProPublica

GET through a GUI

Scenario: A user wants access to data about every non-profit involving animal welfare in Nonprofit Explorer at ProPublica.

ProPublica.org is a research-based journalism site that provides its data to other reporters and the public

1. Navigate to ProPublica.org
2. Click Data
3. Click Nonprofit Explorer
4. Do a search for “animal,” which will bring up a list of animal-based charities
5. Look at the address bar

The screenshot shows the ProPublica Data section. At the top, there's a navigation bar with links like Home, Investigations, Data (which is highlighted with a red circle), MuckReads, Get Involved, About Us, and social media links. Below the navigation is a search bar with the placeholder "Search ProPublica".

The main content area features four news applications:

- Workers' Compensation Reforms by State**: An interactive map showing changes in workers' compensation laws from 2001 to 2005 across different states.
- Employers Complain of Rising Premiums, But Workers' Comp Is at 25-Year Low**: A line graph showing premium trends over time.
- Workers' Comp Benefits: How Much is a Limb Worth?**: An illustration of two human figures with a dollar sign between them.
- Nonprofit Explorer**: A section with a green triangle icon and a table showing Form 990 Returns for the year 2011, with a total revenue of \$10,142,780.

Below these are sections for "Selected Interactive News Applications" and "The ProPublica Nerd Blog".

Selected Interactive News Applications

- MACHINE BIAS**
Chicago Area Disparities in Car Insurance Premiums
by Al Shaw, Julia Angwin and Jeff Larson
ProPublica, April 5, 4:59 a.m.
Some car insurers charge higher premiums in Chicago's minority neighborhoods than in predominantly white neighborhoods with similar risk of accidents.
- Trump's Changing Trust, Annotated**
by Al Shaw and Derek Krautz
ProPublica, April 3, 7:59 a.m.
Here's an annotated look at notable changes to the Donald J. Trump Revocable Trust's certification document.

ProPublica Data Store
Download or Purchase the Data Behind Our Journalism

FDA Data (Free)
Because ProPublica needs to make sure our journalism is accurate, we've created a free dataset of FDA drug approvals and other information.

The ProPublica Nerd Blog

<https://projects.propublica.org/nonprofits/api/v2/search.json?q=animal>

https://projects.propublica.org/nonprofits/api/v2	The address of the API. All requests must start with a similar string in order to point the request to where it needs to go. It essentially reads: go here.
/search.json	The search method set by the API. By appending this to the above URL, it essentially reads: go here, search, output that search in JSON.
?	Question marks denote that whatever follows is a parameter. Adding this to the above essentially reads: go here, and search, and use the following parameters. You can use more than one parameter.
q=	The parameter. The ProPublica API defines q as “A keyword search string. Searches using this parameter will search (in order) organization name, organization alternate name, city.”
animal	Any keyword supplied by the user. Go here, and use the keyword search parameter to search for <i>animal</i> , which will output as JSON.

ProPublica API documentation: <https://projects.propublica.org/nonprofits/api>

Vocabulary pitstop: GUI

- GUI (gooey) stands for Graphic User Interface: *every program* you use has a GUI
- But in the programming/scripting world, there is also the command line/terminal
- We will be using both: Postman is a GUI

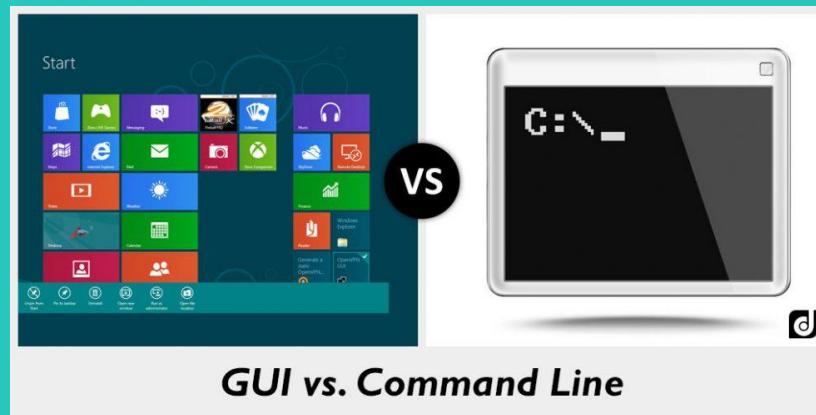


Image from <http://www.differencebtw.com/difference-between-gui-and-command-line/>

GET through a GUI - ProPublica

The screenshot shows the Postman application interface. At the top, there is a header bar with a search field containing "http://projects.proput" and a "+" button. To the right of the search field are environment dropdowns set to "No Environment" and several icons for settings and sharing.

The main workspace contains a request configuration card. The method is set to "GET" with a dropdown arrow. The URL input field contains the value "http://projects.propublica.org/nonprofits/api/v2/search.json?q=animal", which is highlighted with a green rectangular selection box. To the right of the URL are buttons for "Params", "Send" (which is circled in green), and "Save". Below the URL, there are tabs for "Authorization", "Headers", "Body", "Pre-request Script", and "Tests".

At the bottom left, there is a "Type" dropdown set to "No Auth".

<http://projects.propublica.org/nonprofits/api/v2/search.json?q=animal>

Web search versus API

Search for Nonprofit Data

Enter a nonprofit's name, a keyword, or city

animal

State

Any State

Major nonprofit categories

Any Category

Org. Type

Any Type

SEARCH

Advanced Search

Examples: ProPublica, Research or Minneapolis

2435 organization results for animal. Results are ordered by relevance.

Note: Our results **only** include organizations that filed a tax return — for any fiscal year — during the 2012-2014 calendar years.

1 2 3 4 ... 23 24 25 Next » Last »

Company Name	City	State	NTEE Classification	Org. Type
ANIMAL ALLIANCE	WOODLAND HLS	CA	Animal Protection and Welfare ↳ Animal-Related	501(c)(3)
ANIMAL ANGELS	JACKSBORO	TX		501(c)(3)
ANIMAL APPEAL	SHARPSVILLE	PA	Boys Clubs ↳ Youth Development	501(c)(3)

You don't get different *results*, you get the same results in a different format.

```
"total_results": 2435,
"organizations": [Array[100]
  -0: {
    "ein": 731663130,
    "strein": "73-1663130",
    "name": "ANIMAL ALLIANCE",
    "sub_name": "ANIMAL ALLIANCE",
    "city": "WOODLAND HLS",
    "state": "CA",
    "ntee_code": "D20",
    "raw_ntee_code": "D20",
    "subseccd": 3,
    "has_subseccd": true,
    "have_filings": true,
    "have_extracts": true,
    "have_pdfs": true,
    "score": 443.46564
  },
  -1: {
    "ein": 731663130,
    "strein": "73-1663130",
    "name": "ANIMAL ALLIANCE",
    "sub_name": "ANIMAL ALLIANCE",
    "city": "WOODLAND HLS",
    "state": "CA",
    "ntee_code": "D20",
    "raw_ntee_code": "D20",
    "subseccd": 3,
    "has_subseccd": true,
    "have_filings": true,
    "have_extracts": true,
    "have_pdfs": true,
    "score": 443.46564
  }
],
```

Vocabulary pitstop: JSON

- JSON (jason) is the most typical data transmission standard in APIs
- It is lightweight and easy to read and NOT scary
- Consists of key-value pairs, “key”: “value”

```
<unittitle>Johns Hopkins University library records</unittitle>
```

```
“Title”: “Johns Hopkins University library records”
```

GET through a GUI - ProPublica

So taken to extremes, I wonder what would happen if you copied the ProPublica search address into Postman and hit Send....

What about a Google search? Look at the address bar.

Amazon search? Address bar.

Your bank account?? Andress bar!

APIs, like wizards and geocaches, are all around you.

Questions???

GET with GUI - Chronicling America

Web search versus API

Scenario: You wish to link to every digitized edition of a certain newspaper hosted in Chronicling America.

1. Navigate to ChroniclingAmerica.loc.gov
2. Search for “the times dispatch” in quotes

The screenshot shows the Chronicling America homepage. At the top, there's a search bar with the query "the times dispatch" highlighted by a green oval. Below the search bar, the results page displays three newspaper front pages from different publications: "The Ogden Standard", "The Brattleboro Daily Reformer", and "Hickory Daily Record". Each newspaper is shown with its title, location, and the number of pages. On the left side of the results page, there's a sidebar with links like "About Chronicling America", "About the Site and API" (which has a green arrow pointing to it), and "More Resources".

CHRONICLING AMERICA
Historic American Newspapers

Search Pages Advanced Search All Digitized Newspapers 1789-1924

All states + from 1789 to 1924 + "the times dispatch" GO

Pages Available: 11,845,995

100 Years Ago Today: 4/18/1917 (93 issues)

About Chronicling America
About the Site and API 
Recommended Topics
Help

More Resources

- › National Digital Newspaper Program
- › NDNP Award Recipients
- › Newspaper and Current Periodicals Reading Room
- › Ask LC Newspaper & Current Periodicals Librarian
- › Historic Newspapers on Flickr (part of the LC Flickr Commons photostream)

The Ogden Standard
NAVAL GUNS HEARD NEAR BOSTON
The Ogden standard. (12pp.)
Ogden City, Utah

The Brattleboro Daily Reformer
(8pp.)
Brattleboro, Vt.

HICKORY DAILY RECORD
The Hickory daily record. volume (4pp.)
Hickory, N.C.

Web search versus API

Scenario: You wish to link to every digitized edition of a certain newspaper hosted in Chronicling America.

1. Click any record
2. Click All Issues



Web search versus API

This lists every issue, but not that helpful. I wonder if there's another way.

Browse Issues: The times dispatch.																																																																																																																																										
Richmond, Va. (1903-1914)																																																																																																																																										
Browse Issues About Libraries that Have It MARC Record																																																																																																																																										
Issues for: 1903 ▾																																																																																																																																										
Show all front pages																																																																																																																																										
<p>Single edition: dates in bold.</p> <p>Multiple editions: dates in <i>bold italic</i>.</p>																																																																																																																																										
January, 1903		February, 1903		March, 1903		April, 1903																																																																																																																																				
<table><thead><tr><th>S</th><th>M</th><th>T</th><th>W</th><th>T</th><th>F</th><th>S</th></tr></thead><tbody><tr><td></td><td></td><td></td><td></td><td>1</td><td>2</td><td>3</td></tr><tr><td>4</td><td>5</td><td>6</td><td>7</td><td>8</td><td>9</td><td>10</td></tr><tr><td>11</td><td>12</td><td>13</td><td>14</td><td>15</td><td>16</td><td>17</td></tr><tr><td>18</td><td>19</td><td>20</td><td>21</td><td>22</td><td>23</td><td>24</td></tr><tr><td>25</td><td>26</td><td>27</td><td>28</td><td>29</td><td>30</td><td>31</td></tr></tbody></table>		S	M	T	W	T	F	S					1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	<table><thead><tr><th>S</th><th>M</th><th>T</th><th>W</th><th>T</th><th>F</th><th>S</th></tr></thead><tbody><tr><td>1</td><td>2</td><td>3</td><td>4</td><td>5</td><td>6</td><td>7</td></tr><tr><td>8</td><td>9</td><td>10</td><td>11</td><td>12</td><td>13</td><td>14</td></tr><tr><td>15</td><td>16</td><td>17</td><td>18</td><td>19</td><td>20</td><td>21</td></tr><tr><td>22</td><td>23</td><td>24</td><td>25</td><td>26</td><td>27</td><td>28</td></tr></tbody></table>		S	M	T	W	T	F	S	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	<table><thead><tr><th>S</th><th>M</th><th>T</th><th>W</th><th>T</th><th>F</th><th>S</th></tr></thead><tbody><tr><td>1</td><td>2</td><td>3</td><td>4</td><td>5</td><td>6</td><td>7</td></tr><tr><td>8</td><td>9</td><td>10</td><td>11</td><td>12</td><td>13</td><td>14</td></tr><tr><td>15</td><td>16</td><td>17</td><td>18</td><td>19</td><td>20</td><td>21</td></tr><tr><td>22</td><td>23</td><td>24</td><td>25</td><td>26</td><td>27</td><td>28</td></tr></tbody></table>		S	M	T	W	T	F	S	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28																					
S	M	T	W	T	F	S																																																																																																																																				
				1	2	3																																																																																																																																				
4	5	6	7	8	9	10																																																																																																																																				
11	12	13	14	15	16	17																																																																																																																																				
18	19	20	21	22	23	24																																																																																																																																				
25	26	27	28	29	30	31																																																																																																																																				
S	M	T	W	T	F	S																																																																																																																																				
1	2	3	4	5	6	7																																																																																																																																				
8	9	10	11	12	13	14																																																																																																																																				
15	16	17	18	19	20	21																																																																																																																																				
22	23	24	25	26	27	28																																																																																																																																				
S	M	T	W	T	F	S																																																																																																																																				
1	2	3	4	5	6	7																																																																																																																																				
8	9	10	11	12	13	14																																																																																																																																				
15	16	17	18	19	20	21																																																																																																																																				
22	23	24	25	26	27	28																																																																																																																																				
May, 1903		June, 1903		July, 1903		August, 1903																																																																																																																																				
<table><thead><tr><th>S</th><th>M</th><th>T</th><th>W</th><th>T</th><th>F</th><th>S</th></tr></thead><tbody><tr><td></td><td></td><td></td><td></td><td>1</td><td>2</td><td></td></tr><tr><td>3</td><td>4</td><td>5</td><td>6</td><td>7</td><td>8</td><td>9</td></tr><tr><td>10</td><td>11</td><td>12</td><td>13</td><td>14</td><td>15</td><td>16</td></tr><tr><td>17</td><td>18</td><td>19</td><td>20</td><td>21</td><td>22</td><td>23</td></tr><tr><td>24</td><td>25</td><td>26</td><td>27</td><td>28</td><td>29</td><td>30</td></tr><tr><td></td><td></td><td></td><td></td><td></td><td></td><td>31</td></tr></tbody></table>		S	M	T	W	T	F	S					1	2		3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30							31	<table><thead><tr><th>S</th><th>M</th><th>T</th><th>W</th><th>T</th><th>F</th><th>S</th></tr></thead><tbody><tr><td></td><td></td><td>1</td><td>2</td><td>3</td><td>4</td><td></td></tr><tr><td>7</td><td>8</td><td>9</td><td>10</td><td>11</td><td>12</td><td>13</td></tr><tr><td>14</td><td>15</td><td>16</td><td>17</td><td>18</td><td>19</td><td>20</td></tr><tr><td>21</td><td>22</td><td>23</td><td>24</td><td>25</td><td>26</td><td>27</td></tr><tr><td>28</td><td>29</td><td>30</td><td></td><td></td><td></td><td></td></tr></tbody></table>		S	M	T	W	T	F	S			1	2	3	4		7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30					<table><thead><tr><th>S</th><th>M</th><th>T</th><th>W</th><th>T</th><th>F</th><th>S</th></tr></thead><tbody><tr><td></td><td></td><td>1</td><td>2</td><td>3</td><td>4</td><td></td></tr><tr><td>5</td><td>6</td><td>7</td><td>8</td><td>9</td><td>10</td><td>11</td></tr><tr><td>12</td><td>13</td><td>14</td><td>15</td><td>16</td><td>17</td><td>18</td></tr><tr><td>19</td><td>20</td><td>21</td><td>22</td><td>23</td><td>24</td><td>25</td></tr><tr><td>26</td><td>27</td><td>28</td><td>29</td><td>30</td><td>31</td><td></td></tr></tbody></table>		S	M	T	W	T	F	S			1	2	3	4		5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	
S	M	T	W	T	F	S																																																																																																																																				
				1	2																																																																																																																																					
3	4	5	6	7	8	9																																																																																																																																				
10	11	12	13	14	15	16																																																																																																																																				
17	18	19	20	21	22	23																																																																																																																																				
24	25	26	27	28	29	30																																																																																																																																				
						31																																																																																																																																				
S	M	T	W	T	F	S																																																																																																																																				
		1	2	3	4																																																																																																																																					
7	8	9	10	11	12	13																																																																																																																																				
14	15	16	17	18	19	20																																																																																																																																				
21	22	23	24	25	26	27																																																																																																																																				
28	29	30																																																																																																																																								
S	M	T	W	T	F	S																																																																																																																																				
		1	2	3	4																																																																																																																																					
5	6	7	8	9	10	11																																																																																																																																				
12	13	14	15	16	17	18																																																																																																																																				
19	20	21	22	23	24	25																																																																																																																																				
26	27	28	29	30	31																																																																																																																																					
September, 1903		October, 1903		November, 1903		December, 1903																																																																																																																																				
<table><thead><tr><th>S</th><th>M</th><th>T</th><th>W</th><th>T</th><th>F</th><th>S</th></tr></thead><tbody><tr><td></td><td></td><td>1</td><td>2</td><td>3</td><td>4</td><td>5</td></tr><tr><td>6</td><td>7</td><td>8</td><td>9</td><td>10</td><td>11</td><td>12</td></tr><tr><td>13</td><td>14</td><td>15</td><td>16</td><td>17</td><td>18</td><td>19</td></tr><tr><td>20</td><td>21</td><td>22</td><td>23</td><td>24</td><td>25</td><td>26</td></tr><tr><td>27</td><td>28</td><td>29</td><td>30</td><td></td><td></td><td></td></tr></tbody></table>		S	M	T	W	T	F	S			1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30				<table><thead><tr><th>S</th><th>M</th><th>T</th><th>W</th><th>T</th><th>F</th><th>S</th></tr></thead><tbody><tr><td></td><td></td><td>1</td><td>2</td><td>3</td><td>4</td><td></td></tr><tr><td>4</td><td>5</td><td>6</td><td>7</td><td>8</td><td>9</td><td>10</td></tr><tr><td>11</td><td>12</td><td>13</td><td>14</td><td>15</td><td>16</td><td>17</td></tr><tr><td>18</td><td>19</td><td>20</td><td>21</td><td>22</td><td>23</td><td>24</td></tr><tr><td>25</td><td>26</td><td>27</td><td>28</td><td>29</td><td>30</td><td>31</td></tr></tbody></table>		S	M	T	W	T	F	S			1	2	3	4		4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	<table><thead><tr><th>S</th><th>M</th><th>T</th><th>W</th><th>T</th><th>F</th><th>S</th></tr></thead><tbody><tr><td></td><td></td><td>1</td><td>2</td><td>3</td><td>4</td><td>5</td></tr><tr><td>6</td><td>7</td><td>8</td><td>9</td><td>10</td><td>11</td><td>12</td></tr><tr><td>13</td><td>14</td><td>15</td><td>16</td><td>17</td><td>18</td><td>19</td></tr><tr><td>20</td><td>21</td><td>22</td><td>23</td><td>24</td><td>25</td><td>26</td></tr><tr><td>27</td><td>28</td><td>29</td><td>30</td><td>31</td><td></td><td></td></tr></tbody></table>		S	M	T	W	T	F	S			1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31									
S	M	T	W	T	F	S																																																																																																																																				
		1	2	3	4	5																																																																																																																																				
6	7	8	9	10	11	12																																																																																																																																				
13	14	15	16	17	18	19																																																																																																																																				
20	21	22	23	24	25	26																																																																																																																																				
27	28	29	30																																																																																																																																							
S	M	T	W	T	F	S																																																																																																																																				
		1	2	3	4																																																																																																																																					
4	5	6	7	8	9	10																																																																																																																																				
11	12	13	14	15	16	17																																																																																																																																				
18	19	20	21	22	23	24																																																																																																																																				
25	26	27	28	29	30	31																																																																																																																																				
S	M	T	W	T	F	S																																																																																																																																				
		1	2	3	4	5																																																																																																																																				
6	7	8	9	10	11	12																																																																																																																																				
13	14	15	16	17	18	19																																																																																																																																				
20	21	22	23	24	25	26																																																																																																																																				
27	28	29	30	31																																																																																																																																						

GET with GUI - Chronicling America

Does Chronicling America have an API we can use to access this information we're seeing in our browsers?

YES!

Back to Postman!

About the Site and API

Introduction

Chronicling America provides access to information about historic newspapers and select digitized newspaper pages. To encourage a wide range of potential uses, we designed several different views of the data we provide, all of which are publicly visible. Each uses common Web protocols, and access is not restricted in any way. You do not need to apply for a special key to use them. Together they make up an extensive application programming Interface (API) which you can use to explore all of our data in many ways.

Details about these interfaces are below. In case you want to dive right in, though, we use HTML link conventions to advertise the availability of these views. If you are a software developer or researcher or anyone else who might be interested in programmatic access to the data in Chronicling America, we encourage you to look around the site, "view source" often, and follow where the different links take you to get started. When describing Chronicling America as the source of content, please use the URL and a Web site citation, such as "from the Library of Congress, Chronicling America: Historic American Newspapers site".

For more information about the open source Chronicling America software please see the [LibraryOfCongress/chronam](#) GitHub site. Also, please consider subscribing to the [ChronAm-Users](#) discussion list if you want to discuss how to use or extend the software or data from its APIs.

The API

Jump to:

- [Search](#) the newspaper directory and digitized page contents using OpenSearch.
- [Auto Suggest](#) API for looking up newspaper titles
- [Link](#) using our stable URL pattern for Chronicling America resources.
- [JSON](#) views of Chronicling America resources.
- [Linked Data](#) views of Chronicling American resources.
- [Bulk Data](#) for research and external services.
- [CORS](#) and [JSONP](#) support for your JavaScript applications.

Searching the directory and newspaper pages using OpenSearch

The [directory of newspaper titles](#) contains nearly 140,000 records of newspapers and libraries that hold copies of these newspapers. The title records are based on MARC data gathered and enhanced as part of the NDNP program.

GET with GUI - Chronicling America

Scenario: You wish to link to every digitized edition of a certain newspaper in Chronicling America.

The screenshot shows a REST client interface with the following details:

- URL: `http://chroniclingamerica`
- Method: `GET` (highlighted with a green oval)
- Request URL: `http://chroniclingamerica.loc.gov/lccn/sn85038615.json` (highlighted with a green rectangle)
- Environment: `No Environment`
- Params: None
- Send button: `Send` (highlighted with a green oval)
- Status: `200 OK`
- Time: `1425 ms`

`http://chroniclingamerica.loc.gov/lccn/sn85038615.json`

GET with GUI - Chronicling America

Scenario: You wish to link to every digitized edition of a certain newspaper in Chronicling America.

The screenshot shows a POSTMAN interface with the following details:

- Method:** GET
- URL:** <http://chroniclingamerica.loc.gov/lccn/sn85038615.json>
- Authorization:** No Auth
- Status:** 200 OK
- Time:** 1228 ms
- Body:** (Pretty, Raw, Preview, JSON) - The JSON response is displayed below.

```
1 {  
2   "place_of_publication": "Richmond, Va.",  
3   "lccn": "sn85038615",  
4   "start_year": "1903",  
5   "place": [  
6     "Virginia--Richmond"  
7   ],  
8   "name": "The times dispatch.",  
9   "publisher": "Times-Dispatch Co.",  
10  "url": "http://chroniclingamerica.loc.gov/lccn/sn85038615.json",  
11  "end_year": "1914",  
12  "issues": [  
13    {  
14      "url": "http://chroniclingamerica.loc.gov/lccn/sn85038615/1903-01-27/ed-1.json",  
15      "date_issued": "1903-01-27"  
16    },  
17    {  
18      "url": "http://chroniclingamerica.loc.gov/lccn/sn85038615/1903-01-28/ed-1.json",  
19      "date_issued": "1903-01-28"  
20    }  
21  ]  
22}
```

Re-purposing API data

```
1  [
2    {
3      "place_of_publication": "Richmond, Va.",
4      "lccn": "sn85038615",
5      "start_year": "1903",
6      "place": [
7        "Virginia--Richmond"
8      ],
9      "name": "The times dispatch.",
10     "publisher": "times-Dispatch Co.",
11     "url": "http://chroniclingamerica.loc.gov/lccn/sn85038615.json",
12     "end_year": "1914",
13     "issues": [
14       {
15         "url": "http://chroniclingamerica.loc.gov/lccn/sn85038615/1903-01-27/ed-1.json",
16         "date_issued": "1903-01-27"
17       },
18       {
19         "url": "http://chroniclingamerica.loc.gov/lccn/sn85038615/1903-01-28/ed-1.json",
20         "date_issued": "1903-01-28"
21     },
22     {
23       "url": "http://chroniclingamerica.loc.gov/lccn/sn85038615/1903-01-29/ed-1.json",
24       "date_issued": "1903-01-29"
25     },
26     {
27       "url": "http://chroniclingamerica.loc.gov/lccn/sn85038615/1903-01-30/ed-1.json",
28       "date_issued": "1903-01-30"
29     },
30     {
31       "url": "http://chroniclingamerica.loc.gov/lccn/sn85038615/1903-01-31/ed-1.json",
32       "date_issued": "1903-01-31"
33     },
34     {
35       "url": "http://chroniclingamerica.loc.gov/lccn/sn85038615/1903-02-01/ed-1.json",
36       "date_issued": "1903-02-01"
37     },
38     {
39       "url": "http://chroniclingamerica.loc.gov/lccn/sn85038615/1903-02-02/ed-1.json",
40       "date_issued": "1903-02-02"
41     },
42     {
43       "url": "http://chroniclingamerica.loc.gov/lccn/sn85038615/1903-02-04/ed-1.json",
44       "date_issued": "1903-02-04"
45     },
46     {
47       "url": "http://chroniclingamerica.loc.gov/lccn/sn85038615/1903-02-05/ed-1.json",
48       "date_issued": "1903-02-05"
49     },
50     {
51       "url": "http://chroniclingamerica.loc.gov/lccn/sn85038615/1903-02-06/ed-1.json",
52       "date_issued": "1903-02-06"
53     },
54     {
55       "url": "http://chroniclingamerica.loc.gov/lccn/sn85038615/1903-02-07/ed-1.json",
56       "date_issued": "1903-02-07"
57     },
58     {
59       "url": "http://chroniclingamerica.loc.gov/lccn/sn85038615/1903-02-08/ed-1.json",
60       "date_issued": "1903-02-08"
61     }
62   ]
63 }
```

A	B
1 url	date_issued
2 http://chroniclingamerica.loc.gov/lccn/sn85038615/1903-01-27/ed-1.json	1903-01-27
3 http://chroniclingamerica.loc.gov/lccn/sn85038615/1903-01-28/ed-1.json	1903-01-28
4 http://chroniclingamerica.loc.gov/lccn/sn85038615/1903-01-29/ed-1.json	1903-01-29
5 http://chroniclingamerica.loc.gov/lccn/sn85038615/1903-01-30/ed-1.json	1903-01-30
6 http://chroniclingamerica.loc.gov/lccn/sn85038615/1903-01-31/ed-1.json	1903-01-31
7 http://chroniclingamerica.loc.gov/lccn/sn85038615/1903-02-01/ed-1.json	1903-02-01
8 http://chroniclingamerica.loc.gov/lccn/sn85038615/1903-02-03/ed-1.json	1903-02-03
9 http://chroniclingamerica.loc.gov/lccn/sn85038615/1903-02-04/ed-1.json	1903-02-04
10 http://chroniclingamerica.loc.gov/lccn/sn85038615/1903-02-05/ed-1.json	1903-02-05
11 http://chroniclingamerica.loc.gov/lccn/sn85038615/1903-02-06/ed-1.json	1903-02-06
12 http://chroniclingamerica.loc.gov/lccn/sn85038615/1903-02-07/ed-1.json	1903-02-07
13 http://chroniclingamerica.loc.gov/lccn/sn85038615/1903-02-08/ed-1.json	1903-02-08
14 http://chroniclingamerica.loc.gov/lccn/sn85038615/1903-02-10/ed-1.json	1903-02-10
15 http://chroniclingamerica.loc.gov/lccn/sn85038615/1903-02-11/ed-1.json	1903-02-11
16 http://chroniclingamerica.loc.gov/lccn/sn85038615/1903-02-12/ed-1.json	1903-02-12
17 http://chroniclingamerica.loc.gov/lccn/sn85038615/1903-02-13/ed-1.json	1903-02-13
18 http://chroniclingamerica.loc.gov/lccn/sn85038615/1903-02-14/ed-1.json	1903-02-14
19 http://chroniclingamerica.loc.gov/lccn/sn85038615/1903-02-15/ed-1.json	1903-02-15
20 http://chroniclingamerica.loc.gov/lccn/sn85038615/1903-02-17/ed-1.json	1903-02-17
21 http://chroniclingamerica.loc.gov/lccn/sn85038615/1903-02-18/ed-1.json	1903-02-18
22 http://chroniclingamerica.loc.gov/lccn/sn85038615/1903-02-19/ed-1.json	1903-02-19
23 http://chroniclingamerica.loc.gov/lccn/sn85038615/1903-02-20/ed-1.json	1903-02-20
24 http://chroniclingamerica.loc.gov/lccn/sn85038615/1903-02-21/ed-1.json	1903-02-21
25 http://chroniclingamerica.loc.gov/lccn/sn85038615/1903-02-22/ed-1.json	1903-02-22
26 http://chroniclingamerica.loc.gov/lccn/sn85038615/1903-02-24/ed-1.json	1903-02-24
27 http://chroniclingamerica.loc.gov/lccn/sn85038615/1903-02-25/ed-1.json	1903-02-25
28 http://chroniclingamerica.loc.gov/lccn/sn85038615/1903-02-26/ed-1.json	1903-02-26
29 http://chroniclingamerica.loc.gov/lccn/sn85038615/1903-02-27/ed-1.json	1903-02-27
30 http://chroniclingamerica.loc.gov/lccn/sn85038615/1903-02-28/ed-1.json	1903-02-28

Converting these JSON search results to a CSV (spreadsheet) took less than 10 seconds.

Questions???

GET with GUI - Twitter

Using an API to collect records

@JohnsHopkins

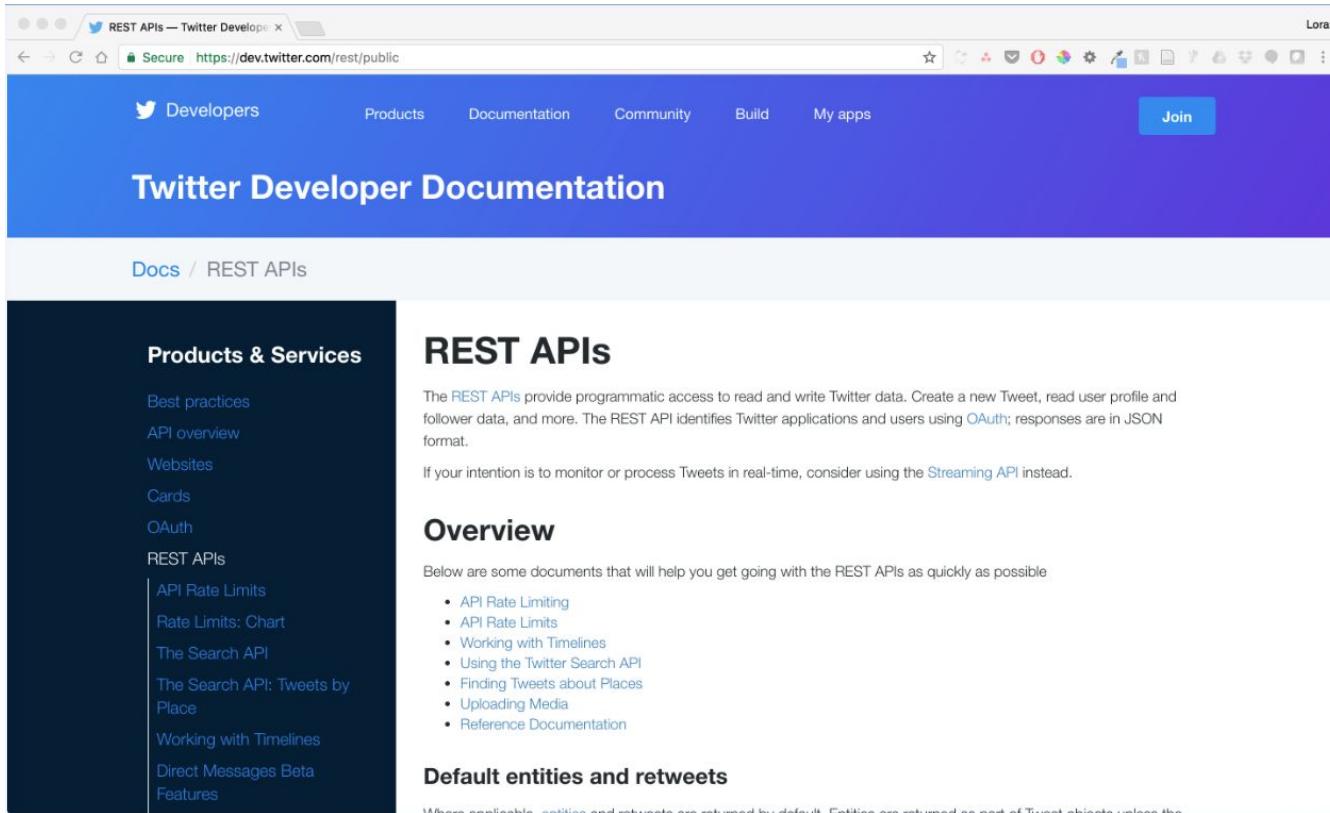
Scenario: As your university's records manager, you wish to regularly capture Tweets mentioning your university.

The screenshot shows a Twitter search interface with the query 'johnshopkins' in the search bar. The 'LATEST' tab is selected. The results list several tweets from various users:

- March For ScienceCLE** (@ScienceMarchCLE) - 28m ago: Great recognition for a world-class scientist from @JEL! hub.jhu.edu/2017/04/17/rac... #ScienceMarchCLE @JohnsHopkins
- Johns Hopkins biologist, geneticist Rachel Green...** (@JohnsHopkins) - 3 hours ago: Johns Hopkins biologist and geneticist has been researching the cellular structure for nearly three decades hub.jhu.edu
- GalenusDixit** (@GalenOfPergamum) - 35m ago: Today, I am giving a paper on so-called "folk" and "popular" medicine in the ancient world in the history of medicine dept. @JohnsHopkins
- Hopkins Engineering** (@HopkinsEngineer) - 1h ago: #ICYMI: Photographer captures the complex work materials scientists bit.ly/2nGivds via @HubJHU @JohnsHopkins
- Going to extremes: Photographer captures the co...** (@JohnsHopkins) - 1h ago: MICA photographer Jay Gould creates arts based on his HEMI artist-in-residence experience hub.jhu.edu

Three specific mentions of '@JohnsHopkins' are highlighted with red circles.

Using an API to collect records



The screenshot shows a web browser displaying the Twitter Developer Documentation for the REST APIs. The URL is <https://dev.twitter.com/rest/public>. The page has a blue header with the Twitter logo and navigation links for Developers, Products, Documentation, Community, Build, My apps, and Join. Below the header is a large blue banner with the text "Twitter Developer Documentation". The main content area has a dark blue sidebar on the left containing links for Products & Services, Best practices, API overview, Websites, Cards, OAuth, REST APIs, API Rate Limits, Rate Limits: Chart, The Search API, The Search API: Tweets by Place, Working with Timelines, Direct Messages Beta, and Features. The main content area has a white background with the title "REST APIs". It describes the REST APIs as providing programmatic access to Twitter data and identifies them as using OAuth and JSON format. It also mentions the Streaming API for real-time monitoring. Below this is an "Overview" section with a list of documents for getting started, including API Rate Limiting, Working with Timelines, Using the Twitter Search API, Finding Tweets about Places, Uploading Media, and Reference Documentation. At the bottom is a section titled "Default entities and retweets". A small note at the very bottom states: "Where available, mention and retweets are returned by default. Entities are returned as part of Tweet objects' values.".

Using an API to collect records

The screenshot shows a web browser displaying the Twitter Developer Documentation at <https://dev.twitter.com/rest/public/search>. The page title is "The Search API — Twitter Dev". The main content area is titled "The Search API". It describes the API as part of Twitter's REST API, allowing queries against recent or popular Tweets. It notes that the API is focused on relevance and not completeness, and provides a detailed reference for the `GET search/tweets` endpoint. A section titled "How to build a query" explains how to test a query by visiting `twitter.com/search` and copying the URL. A red circle highlights the URL `https://api.twitter.com/1.1/search/tweets.json?q=%40twitterapi`.

The page includes a sidebar with links to "Products & Services" such as Best practices, API overview, Websites, Cards, OAuth, REST APIs, API Rate Limits, Rate Limits: Chart, The Search API, The Search API: Tweets by Place, Working with Timelines, Direct Messages Beta, Features, and Collections.

The Search API

The Twitter Search API is part of Twitter's REST API. It allows queries against the indices of recent or popular Tweets and behaves similarly to, but not exactly like the Search feature available in Twitter mobile or web clients, such as [Twitter.com search](#). The Twitter Search API searches against a sampling of recent Tweets published in the past 7 days.

Before getting involved, it's important to know that the Search API is focused on relevance and not completeness. This means that some Tweets and users may be missing from search results. If you want to match for completeness you should consider using a [Streaming API](#) instead.

A detailed reference on this API endpoint can be found at [GET search/tweets](#).

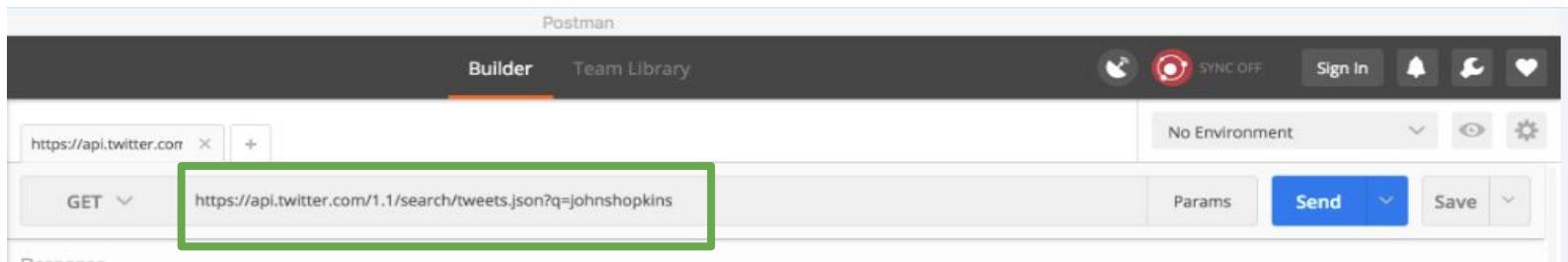
How to build a query

The best way to build a query and test if it's valid and will return matched Tweets is to first try it at [twitter.com/search](#). As you get a satisfactory result set, the URL loaded in the browser will contain the proper query syntax that can be reused in the API endpoint. Here's an example:

1. We want to search for Tweets referencing @twitterapi account. First, we run the search on [twitter.com/search](#)
2. Check and copy the URL loaded. In this case, we got: <https://twitter.com/search?q=%40twitterapi>
3. Replace "<https://twitter.com/search>" with "<https://api.twitter.com/1.1/search/tweets.json>" and you will get:

<https://api.twitter.com/1.1/search/tweets.json?q=%40twitterapi>

Using an API to collect records



<https://api.twitter.com/1.1/search/tweets.json?q=johnshopkins>

Let's give it a try!

Using an API to collect records

The screenshot shows the Postman application interface. At the top, the title "Postman" is visible, followed by tabs for "Builder" and "Team Library". Below the tabs, the URL "https://api.twitter.com" is entered in the address bar, which is highlighted with a green box. To the right of the URL are buttons for "Sign In", "SYNC OFF", and various notification and settings icons.

In the main workspace, a "GET" request is selected, and the URL "https://api.twitter.com/1.1/search/tweets.json?q=johnshopkins" is displayed in the request field. To the right of the URL are buttons for "Params", "Send", and "Save".

The "Headers" tab is active, showing a single header entry: "Key" (New key) and "Value" (value). There are buttons for "Bulk Edit" and "Presets" below this table.

The "Body" tab is active, showing the response body in JSON format. The response status is "Status: 400 Bad Request", "Time: 29 ms", and "Size: 360 B".

The response body is:

```
1+ [
2+   "errors": [
3+     {
4+       "code": 215,
5+       "message": "Bad Authentication data."
6+     }
7+   ]
8+ }
```

The message "Bad Authentication data." is highlighted with a red box.

Using an API to collect records

Authentication on all endpoints

Applications must authenticate all requests with OAuth 1.0a or Application-only authentication. This allows us to prevent abusive behavior, and it also helps us to further understand how categories of applications are using the API. We apply this understanding to better meet the needs of developers as we continue to evolve the platform.

Source: <https://dev.twitter.com/rest/public>

Using an API to collect records

So, we must **authenticate** in order to use Twitter's search API.

The type of authentication Twitter requires is **OAuth**, an open protocol for authentication (see: <https://oauth.net>)

Postman will help walk us through OAuth1.0 authentication, but you must have a Twitter developer account in order to do so!

Show and tell!

The screenshot shows the Postman application interface. At the top, the URL is set to `https://api.twitter.com`. Below the URL, there is a dropdown menu for the 'Authorization' tab, which is currently set to 'No Auth'. A dropdown menu is open, showing options: 'No Auth', 'Basic Auth', 'Digest Auth', 'OAuth 1.0', 'OAuth 2.0', 'Hawk Authentication', and 'AWS Signature'. The 'OAuth 1.0' option is highlighted. In the main body area, there is a JSON response shown in a 'Pretty' format. The response indicates a 'Bad Authentication' error with code 215. The status bar at the bottom right of the Postman window shows 'Status: 400 Bad Request' and 'Time: 29 ms'.

```
1 - {  
2 -   "errors": [  
3 -     {  
4 -       "code": 215,  
5 -       "message": "Bad Authentication"  
6 -     }  
7 -   ]  
8 - }
```

Using an API to collect records

In the end, we've got JSON. Though, be wary of the Twitter **developer license** you signed!

```
1  {
2    "statuses": [
3      {
4        "created_at": "Tue Apr 18 14:16:03 +0000 2017",
5        "id": 854337735129092096,
6        "id_str": "854337735129092096",
7        "text": "RT @ScienceMarchCLE: Great recognition for a world-class scientist from #CLE! https://t.co/i2Gtx37EMr #ScienceMarchCLE @JohnsHopkins",
8        "truncated": false,
9        "entities": {
10          "hashtags": [
11            {
12              "text": "CLE",
13              "indices": [
14                72,
15                76
16              ]
17            },
18            {
19              "text": "ScienceMarchCLE",
20              "indices": [
21                102,
22                118
23              ]
24            }
25          ],
26          "symbols": [],
27          "user_mentions": [
28            {
29              "screen_name": "ScienceMarchCLE",
30              "name": "March For ScienceCLE",
31              "id": 824422865206472706,
32              "id_str": "824422865206472706",
33              "indices": [
34                3,
35                19
36              ]
37            },
38            {
39              "screen_name": "JohnsHopkins",
40              "name": "Johns Hopkins U.",
41              "id": 14441010,
42              "id_str": "14441010",
43              "indices": [
44                119,
45                132
46              ]
47            }
48          ]
49        }
50      }
51    ]
52  }
```

F. Be a Good Partner to Twitter

1. Follow the guidelines for using Tweets in broadcast if you display Tweets offline.
2. If you provide Content to third parties, including downloadable datasets of Content or an **API that returns Content**, you will only distribute or allow download of Tweet IDs and/or User IDs.
 - a. You may, however, provide export via non-automated means (e.g., download of spreadsheets or PDF files, or use of a “save as” button) of up to 50,000 public Tweets and/or User Objects per user of your Service, per day.
 - b. Any Content provided to third parties via non-automated file download remains subject to this Policy.

Questions???

Scripting

Scripting - Why?

Using a GUI application like **Postman** to interact with APIs can be a great way to *learn, explore, and troubleshoot*, but ultimately you'll hit a brick wall, because:

- It takes an **awful lot of clicks** to get out a small amount of data (relatively speaking)
- If you want to get multiple full records OUT you've got to run a GET as **many times** as there are records you want to retrieve
- While you can POST many one-off changes using a GUI like Postman, you can rarely get a GUI to make **intelligent, iterative POSTs at scale**
- **Manually authenticating** is a pain
- Though we told you that you will be sometimes playing the role of “application” in this API world, you don't *always* want to **be the application!**

Scripting - How?

Yes, this is a huge barrier to entry for most users, but it can be mitigated:

- We (defined here as both **archivists** and **developers**) are a **community** that likes sharing!
 - Frankly, if you're sitting down to write scripts from scratch, you're **doing it wrong**
- There is no “**one right language**” to make this work
 - If you have *any* prior knowledge of a particular scripting language, **start there**
 - All the scripts you will use in this workshop are **Python** because: 1) Python (and, to a lesser degree, Ruby) is Lora's preferred hammer, and 2) unscientifically speaking, it seems that Python is the preferred language of archivists (which means there's more to steal/borrow)
 - But, if you want, you can use a **Ruby** or **Perl** or **PHP** or **JavaScript** shaped hammer!
- The Internet is full of **helpful advice**!
 - Just don't feed the trolls

Scripting - No, really, *how*?

Remember all the legwork you did both at home and during the early part of this workshop? You've:

- Installed applications, including the **text editor Atom**
- Installed (or located) a **shell**, namely *Terminal* (Mac) or *Cygwin* (Windows)
- Installed (or confirmed installation of) **python**

Guess what? You've set up a **python development environment** already! Good work!

With that work complete, for the remainder of this workshop you should only need to type `python [name of script here].py` into Terminal/Cygwin, and you'll be **executing Python scripts!** Just remember:

- You should be located in the same directory as the script (and any files it is reliant on) before you type your command (you can always `ls` to confirm the script is there!)

For more, see: <http://www.shubhro.com/2014/05/29/development-environment/> and/or
<http://python-guide-pt-br.readthedocs.io/en/latest/starting/install/osx/> (Mac specific)

GET with Script- ProPublica

GET with a script - ProPublica

Let's return to our ProPublica example from earlier.

Scenario: A user wants **access to data** about every non-profit involving animal welfare in Nonprofit Explorer at ProPublica.

But, this time, let's assume that this user isn't content just copying/pasting pages of data out of Postman. This researcher has **big plans** for this dataset, and wants a **standalone JSON file** instead!

Scenario, v. 2: A user wants a **JSON file** containing ProPublica Nonprofit Explorer data about every non-profit involving animal welfare.

GET with a script - ProPublica

Mac

1. In the Finder navigate to your MARAC_API_Workshop directory
2. Ctrl+click the MARAC_API_Workshop directory, and select “New Terminal at Folder”
3. Type `ls` and examine the contents of that folder
4. Type `python proPublica.py`
5. A new file called “proPublicaRecord.json” should now be in your MARAC_API_Workshop directory
6. Open “proPublicaRecord.json” with Atom and take a look!

PC

1. Open Cygwin
2. Type `cd MARAC_API_Workshop` to enter the MARAC_API_Workshop directory
3. Type `ls` and examine the contents of that folder
4. Type `python proPublica.py`
5. A new file called “proPublicaRecord.json” should now be in your MARAC_API_Workshop directory
6. Open “proPublicaRecord.json” with Atom and take a look!

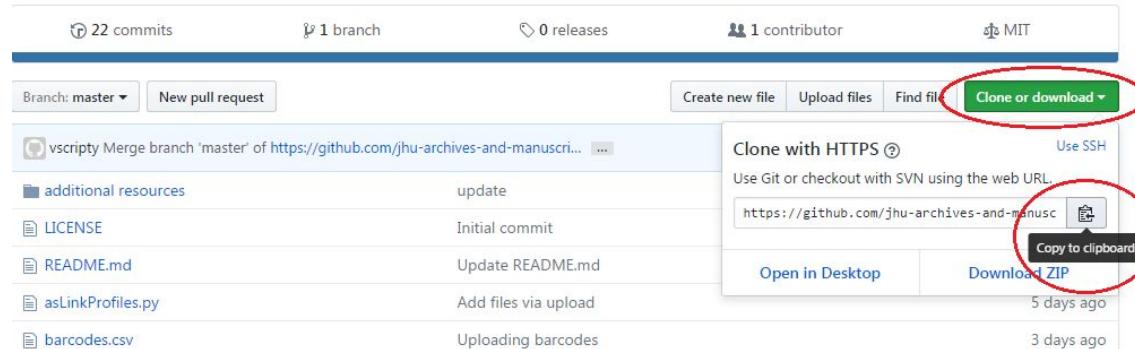
Questions???

Lunch!

Technical Pitstop - Scripting set-up

Everyone

1. Navigate to GitHub.com
2. Search for “MARAC API Workshop”
3. **Bookmark! You will need this again.**
4. Then click the **green** button, and copy the path to the clipboard



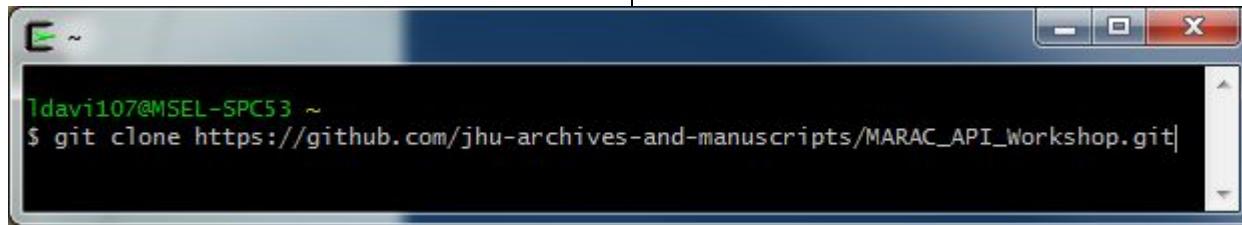
Technical Pitstop - Scripting set-up

Mac

1. Ctrl+click on Desktop > New Terminal at Folder
2. Type `git clone`, leave a space, then [ctrl-click > paste], and hit enter

PC

1. Open cygwin
2. Type `git clone`, leave a space, then [right-click > paste], and hit enter



Now you have a folder (either on your Desktop, or in your Cygwin home directory) that contains all the materials you'll need for the rest of today's workshop.

This folder, titled “**MARAC_API_Workshop**,” is a *direct clone* of what you see in your browser on [github.com](https://github.com/jhu-archives-and-manuscripts/MARAC_API_Workshop).

Technical Pitstop - Scripting set-up

Mac

1. Does your Desktop now have a folder called MARAC_API_Workshop on it?

PC

1. Does C:\cygwin64\home\[username] have a folder named MARAC_API_Workshop?

Load, GET, and compare - VIAF

API possibilities

Get data out → Do something to it → Put it back in



We are
here

Load, GET, and compare - VIAF

Scenario: You have an existing spreadsheet containing a number of organizational names that are either subjects or creators of some of your collections. Now, you want to take this manually-made spreadsheet and actually do some authority control work!

Load, GET, and compare - VIAF

1. Open Excel/Numbers (or your preferred spreadsheet program)
2. Type “name” (lower-case) in the first column, first row
 - a. You can inspect the file “organizations.csv” in GitHub if you want to see a sample
3. Type several corporate names of your choosing (universities, businesses, etc.) in the subsequent rows of column A
4. Save the file as a .csv with the title “organizations.csv” in the MARAC_API_Workshop directory
 - a. Warning: You will be asked to overwrite an existing file named “organizations.csv” since we’ve saved a sample there as a just-in-case. Please overwrite this file.
5. From Terminal/Cygwin type `python viafReconciliationCorporate.py`

Load, GET, and compare - VIAF

Uh oh...

Technical Pitstop - Installing extra packages

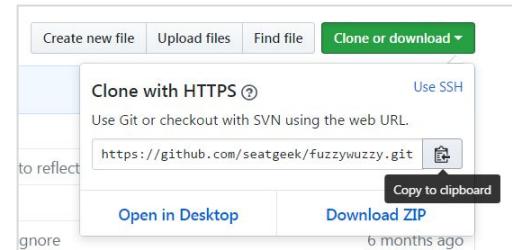
Everyone

Downloading the fuzzy wuzzy python package:

1. Google “fuzzy wuzzy github” and it should be the first result
2. Click the green “Clone or Download” button, and copy the path to the clipboard
3. Confirm that terminal/cygwin is still in the MARAC_API_Workshop folder
4. Type `git clone {paste path}`, which should look like

```
git clone
```

```
https://github.com/seatgeek/fuzzywuzzy.git
```



Technical Pitstop - Installing extra packages

Everyone

Installing the fuzzy wuzzy python package:

1. Type `cd fuzzywuzzy` to enter the newly created fuzzywuzzy directory
2. Type `ls` to see what's in the directory and note the script "setup.py"
3. To execute that setup script, type `python setup.py install`
4. The package is **installed!**
5. Type `cd ..` to return you to the MARAC_API_Workshop directory
6. Try the script again! `python viafReconciliationCorporate.py`

Load, GET, and compare - VIAF

Let's try this again!

1. Type `python viafReconciliationCorporate.py`
2. Success!
3. There is a newly created “`viafCorporateResults.csv`” file in your `MARAC_API_Workshop` directory

Mac users: Desktop

Windows users: `C:\cygwin64\home\[username]\MARAC_API_Workshop`

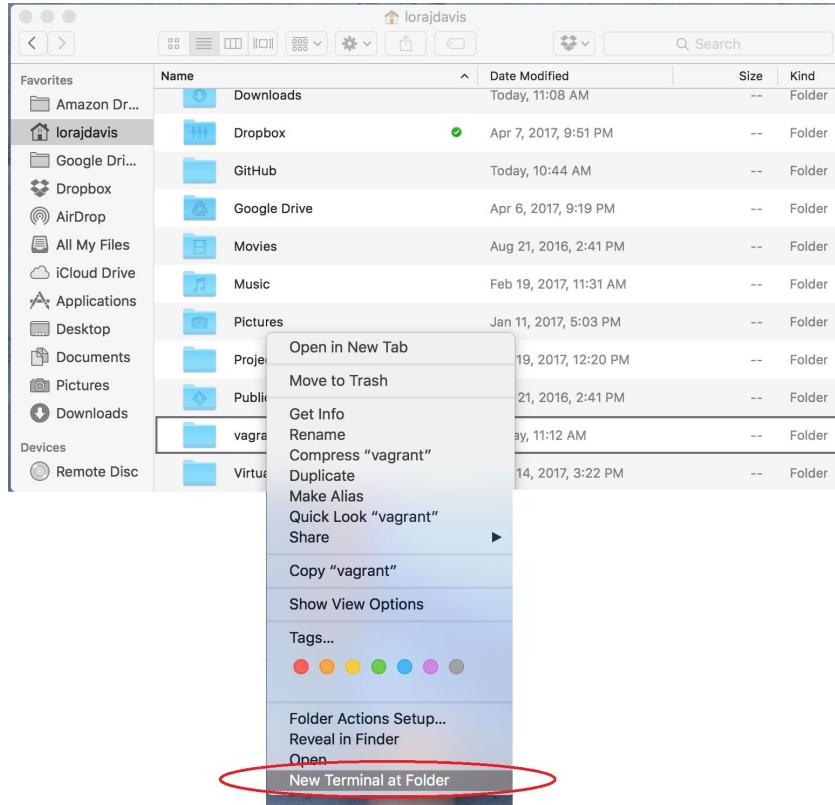
4. Let's open and inspect it!

Questions???

Technical Pitstop: vagrant install and vagrant up

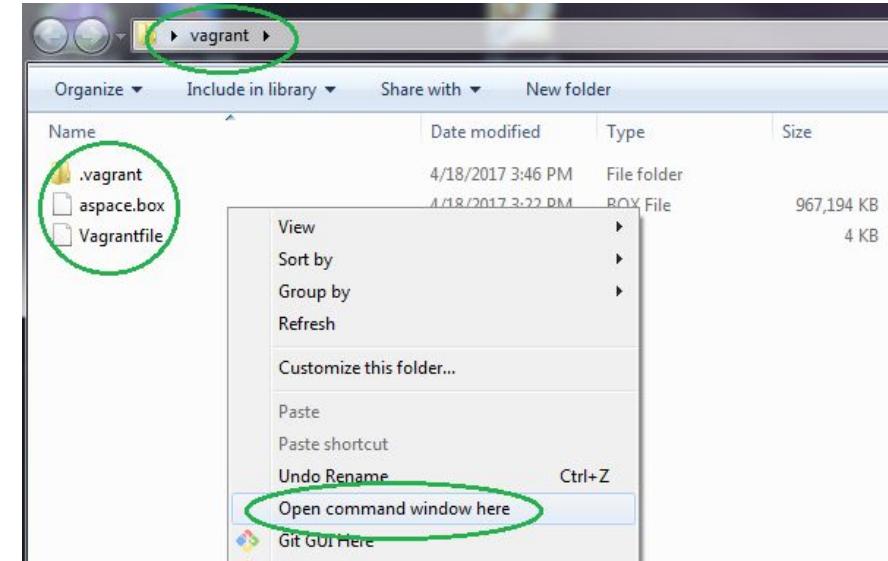
(this is super awesome)

Mac

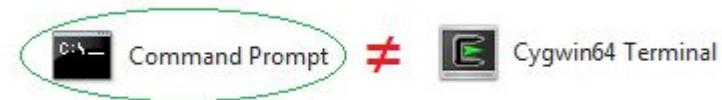


PC

Navigate to your new vagrant folder on the desktop. Open the command prompt in that folder with shift+Right-click:



Note: The command prompt is not Cgywin, we specifically mean the command prompt on this slide:

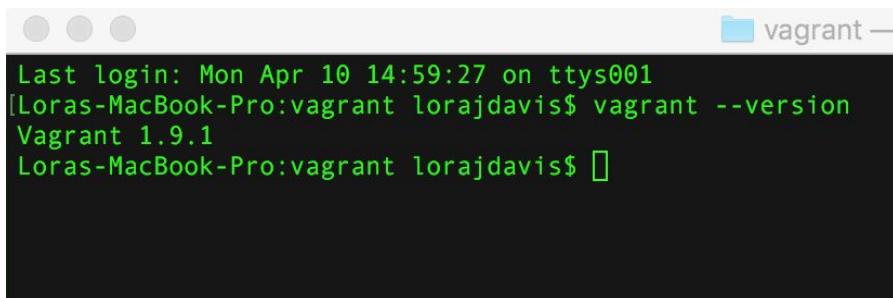


Technical Pitstop - Vagrant

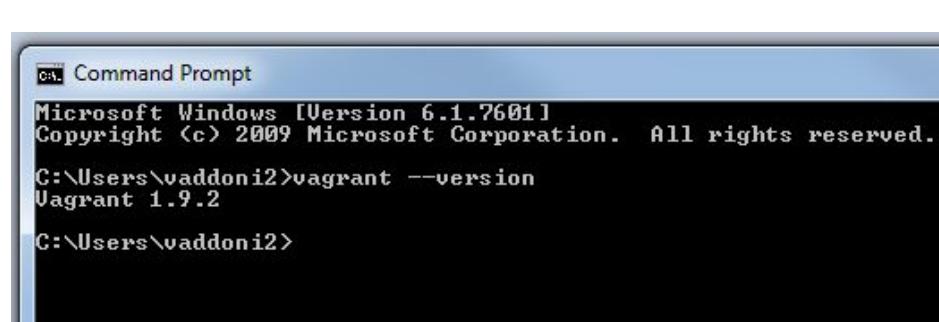
Mac

PC

vagrant --version



```
Last login: Mon Apr 10 14:59:27 on ttys001
[Loras-MacBook-Pro:vagrant lorajdavis$ vagrant --version
Vagrant 1.9.1
Loras-MacBook-Pro:vagrant lorajdavis$ ]
```



```
Command Prompt
Microsoft Windows [Version 6.1.7601]
Copyright (c) 2009 Microsoft Corporation. All rights reserved.

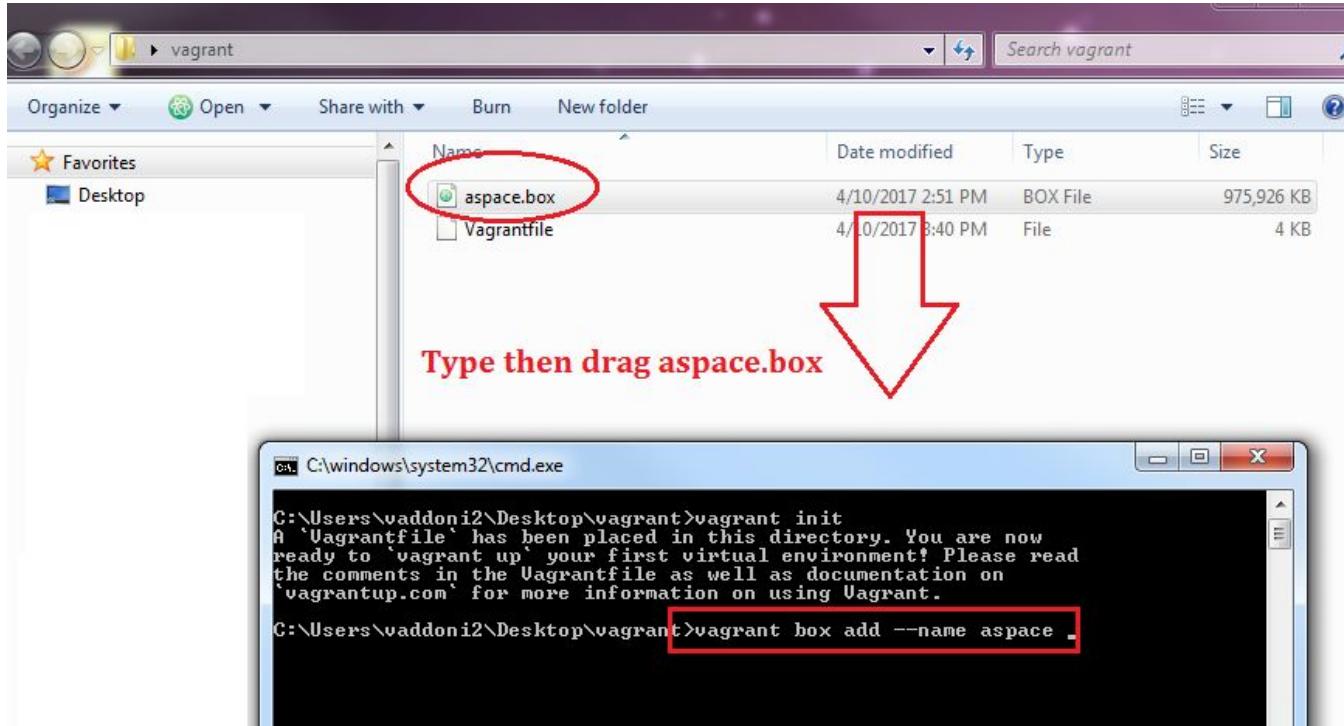
C:\Users\vaddoni2>vagrant --version
Vagrant 1.9.2

C:\Users\vaddoni2>
```

If you see any number that doesn't start with "1.9" or if you get an error altogether please use your post-its, or shout out!

Everyone

Type **vagrant box add --name aspace** and leave a space, then drag:
(see next slide to confirm before hitting enter)



Everyone

It should look like this now. If it does, hit Enter:



vagrant — -bash — 120x40

```
Last login: Mon Apr 10 11:19:13 on ttys000
[Loras-MacBook-Pro:vagrant lorajdavis$ vagrant init
A `Vagrantfile` has been placed in this directory. You are now
ready to `vagrant up` your first virtual environment! Please read
the comments in the Vagrantfile as well as documentation on
`vagrantup.com` for more information on using Vagrant.
[Loras-MacBook-Pro:vagrant lorajdavis$ ls
Vagrantfile      aspace.box
Loras-MacBook-Pro:vagrant lorajdavis$ vagrant box add --name aspace /Users/lorajdavis/vagrant/aspace.box]
```

Technical Pitstop - Initialize Vagrant

Mac

PC

vagrant init aspace

```
Last login: Mon Apr 17 21:30:39 on ttys002
[Loras-MacBook-Pro:vagrant lorajdavis$ vagrant --version
Vagrant 1.9.1
[Loras-MacBook-Pro:vagrant lorajdavis$ vagrant box add --name aspace /Users/lorajdavis/vagrant/aspace.box
=> box: Box file was not detected as metadata. Adding it directly...
=> box: Adding box 'aspace' (v0) for provider:
box: Unpacking necessary files from: file:///Users/lorajdavis/vagrant/aspace.box
=> box: Successfully added box 'aspace' (v0) for 'virtualbox'!
Loras-MacBook-Pro:vagrant lorajdavis$ vagrant init aspace
```

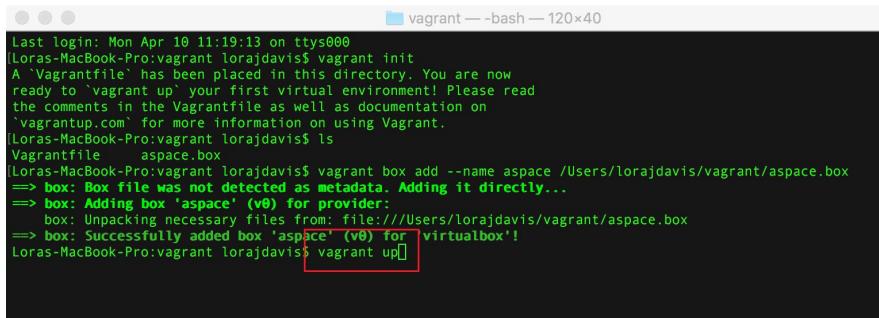
```
C:\Users\vaddoni2\Desktop\vagrant>vagrant init aspace
A 'Vagrantfile' has been placed in this directory. You are now
ready to 'vagrant up' your first virtual environment! Please read
the comments in the Vagrantfile as well as documentation on
'verbanup.com' for more information on using Vagrant.
```

Technical Pitstop - Vagrant up

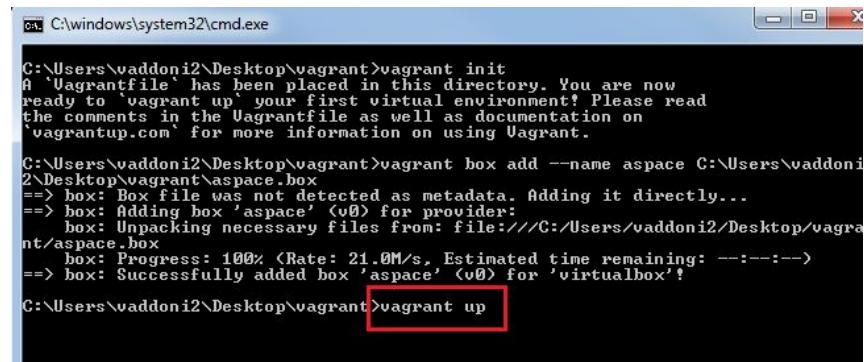
Mac

PC

vagrant up



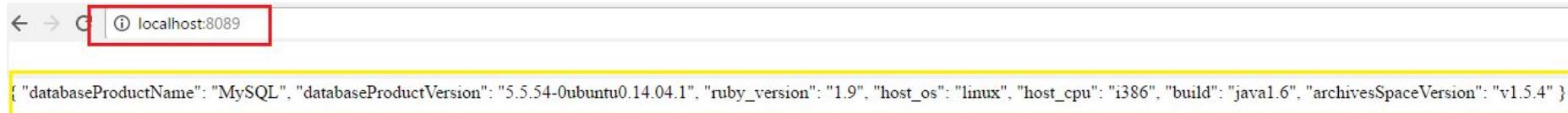
```
Last login: Mon Apr 10 11:19:13 on ttys000
[Loras-MacBook-Pro:~]lorajdavis$ vagrant init
A 'Vagrantfile' has been placed in this directory. You are now
ready to 'vagrant up' your first virtual environment! Please read
the comments in the Vagrantfile as well as documentation on
'vagrantup.com' for more information on using Vagrant.
[Loras-MacBook-Pro:~]lorajdavis$ ls
Vagrantfile  aspace.box
[Loras-MacBook-Pro:~]lorajdavis$ vagrant box add --name aspace /Users/lorajdavis/vagrant/aspace.box
==> box: Box file was not detected as metadata. Adding it directly...
==> box: Adding box 'aspace' (<v0>) for provider:
    box: Unpacking necessary files from: file:///Users/lorajdavis/vagrant/aspace.box
==> box: Successfully added box 'aspace' (<v0>) for 'virtualbox'!
[Loras-MacBook-Pro:~]lorajdavis$ vagrant up
```



```
C:\Windows\system32\cmd.exe
C:\Users\vaddoni2\Desktop\vagrant>vagrant init
A 'Vagrantfile' has been placed in this directory. You are now
ready to 'vagrant up' your first virtual environment! Please read
the comments in the Vagrantfile as well as documentation on
'vagrantup.com' for more information on using Vagrant.
C:\Users\vaddoni2\Desktop\vagrant>vagrant box add --name aspace C:\Users\vaddoni2\Desktop\vagrant\aspace.box
==> box: Box file was not detected as metadata. Adding it directly...
==> box: Adding box 'aspace' (<v0>) for provider:
    box: Unpacking necessary files from: file:///C:/Users/vaddoni2/Desktop/vagrant/aspace.box
    box: Progress: 100% <Rate: 21.0M/s, Estimated time remaining: ---:-->
==> box: Successfully added box 'aspace' (<v0>) for 'virtualbox'!
C:\Users\vaddoni2\Desktop\vagrant>vagrant up
```

Technical Pitstop - Check Vagrant backend

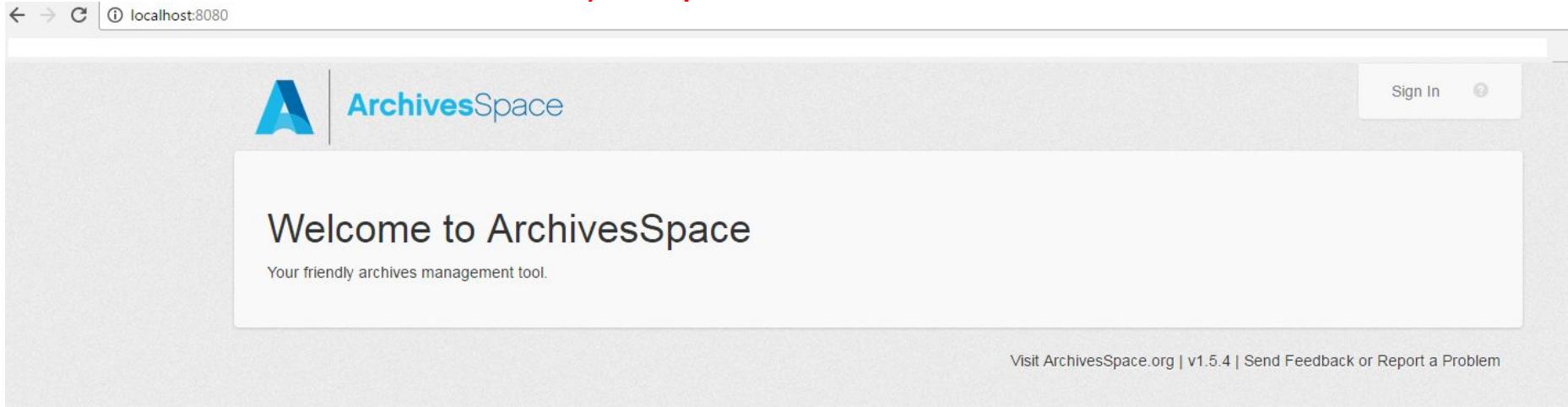
Open browser, navigate to:
http://localhost:8089



```
{ "databaseProductName": "MySQL", "databaseProductVersion": "5.5.54-0ubuntu0.14.04.1", "ruby_version": "1.9", "host_os": "linux", "host_cpu": "i386", "build": "java1.6", "archivesSpaceVersion": "v1.5.4" }
```

Technical Pitstop - Check Vagrant staff interface

Try: <http://localhost:8080>



Now you have a personal ArchivesSpace! It's yours and it's re-useable! Look around.

Username: admin
Password: admin

POST:
ArchivesSpace

ArchivesSpace!

Scenario: You have successfully migrated into -- or have begun to use -- an instance of ASpace at your institution, but...

- There's all this new functionality, what do I do with it?
 - I don't have barcodes for my containers, or I have faux codes
 - I do have barcodes, but they're not in ArchivesSpace. How do I get them in without ruining a student worker's semester?
- There are new fields where there were no fields before
 - I'd love to use URIs for Agents, but that's a lot of work
 - BARCODES, again with the barcodes
- We didn't use Archivists' Toolkit for accessions, how do I get them in now?
- Suppressing and unsuppressing, publishing and unpublishing, and how do I publish everything but not *those* things?

As some of you know, it's a huge undertaking and you might have dozens/hundreds/thousands of old and new problems.

ArchivesSpace!

Scenario: You have successfully migrated into -- or have begun to use -- an instance of ASpace at your institution, but this is a short workshop, so here are our problems for today:

1. We don't have **container profiles** in ArchivesSpace and would like to, so we need to create some
2. In following the migration instructions for 1.5, we had to add **faux codes**; we'd like to use our *actual* barcodes
3. Now that we have container profiles, we need to link them to *actual top containers*

Extra archivistry sidebar

- These are, in fact, all problems we've addressed (or are addressing) at Johns Hopkins. And this is exactly how we did (or will be) solving these issues.
- If you switch out "container profiles" for "agent records" or "subject headings" or "digital objects," the steps are similar and will likely transfer. Namely:
 - Create new records
 - Modify existing records
 - Link records

API possibilities

Get data out → Do something to it → Put it back in



We are
here

POST to AS with GUI-Container profiles

Container profiles

What's a container profile?

ASpace offers “container modeling” for the first time in the archives world.

Every type of box (ex. record carton) in your library gets its own record (a profile), which records its height, width, and depth. This helps calculate space on a huge scale, and is a game-changer for some repositories.

So, we have boxes o'plenty ----->

But to use this feature, we need to get their profiles into AS.



POST to AS with GUI

Before we start posting to AS, we need to authenticate, so let's do that and try a GET first:

The screenshot shows the Postman application interface. At the top, there is a toolbar with buttons for 'POST' (selected), 'Params', 'Send', and 'Save'. Below the toolbar, the URL 'http://localhost:8089/users/admin/login' is entered. Underneath the URL, there are tabs for 'Authorization', 'Headers (1)', 'Body' (which is selected and highlighted with a red border), 'Pre-request Script', and 'Tests'. To the right of these tabs are 'Cookies' and 'Code' buttons. In the 'Body' section, there are three radio buttons: 'form-data' (disabled), 'x-www-form-urlencoded' (selected and highlighted with a green border), and 'raw' and 'binary' (disabled). Below the body type selection, there is a table with two columns: 'Key' and 'Value'. A single row is present with the key 'password' and the value 'admin'. The 'password' key has a checked checkbox next to it. A 'Bulk Edit' button is located at the bottom right of the table.

Endpoint: <http://localhost:8089/users/admin/login>

password: admin

POST to AS with GUI

The screenshot shows the Postman application interface. At the top, there is a header bar with a 'GET' button (circled in green), a URL field containing 'localhost:8089/users/admin/login', and tabs for 'Authorization', 'Headers (1)', 'Body', 'Pre-request Script', and 'Tests'. Below this, there is a dropdown menu for 'Type' set to 'No Auth'. The 'Body' tab is selected, showing options for 'Pretty', 'Raw', 'Preview', and 'JSON' (with a dropdown arrow). The JSON response body is displayed as:

```
1 {  
2   "session": "525f7eb786396736a7722347e87d9a98b3bb2918addae5bacdd7916b7b4f5077",  
3   "user": {  
4     "lock_version": 3,  
5     "username": "admin",  
6     "name": "Administrator",  
7     "is system user": true.  
8 }
```

A red box highlights the session value '525f7eb786396736a7722347e87d9a98b3bb2918addae5bacdd7916b7b4f5077'. To the right of the JSON preview, a red text overlay reads: 'Copy just the value and switch to GET'.

POST to AS with GUI

GET <http://archivesspace01.mse.jhu.edu:8089/repositories/3/resources/3>

Authorization Headers (2) Body Pre-request Script Tests

Key Value

Content-Type application/x-www-form-urlencoded

X-ArchivesSpace-Session ed9936142c2bd42139ba4daa8d9a1d20dae04fd16abc5c8fcc38accf382b0bbb

Body Cookies Headers (6) Tests

Pretty Raw Preview JSON

```
1 "lock_version": 5,
2 "title": "University history scrapbook collection",
3 "...."
```

Key: X-ArchivesSpace-Session

Endpoint: <http://localhost:8089/repositories/1/resources/1>

POST to AS with GUI - Container profiles

1. Navigate to the directory with our cloned GitHub repo

Mac users: Desktop

Windows users:

C:\cygwin64\home\[username]\MARAC_API_Workshop

2. Open “containerProfiles.json” with Atom
3. Here are the profiles, ready to go
4. Copy, and go back to Postman

```
[{  
    "name": "Flat box01",  
    "extent_dimension": "width",  
    "height": "3",  
    "width": "12",  
    "depth": "16",  
    "dimension_units": "inches",  
    "jsonmodel_type": "container_profile"  
},  
{  
    "name": "Flat box02",  
    "extent_dimension": "width",  
    "height": "3",  
    "width": "21",  
    "depth": "25",  
    "dimension_units": "inches",  
    "jsonmodel_type": "container_profile"  
},  
{  
    "name": "Flat box03",  
    "extent_dimension": "width",  
    "height": "3",  
    "width": "9",  
    "depth": "11",  
    "dimension_units": "inches",  
    "jsonmodel_type": "container_profile"  
},  
]
```

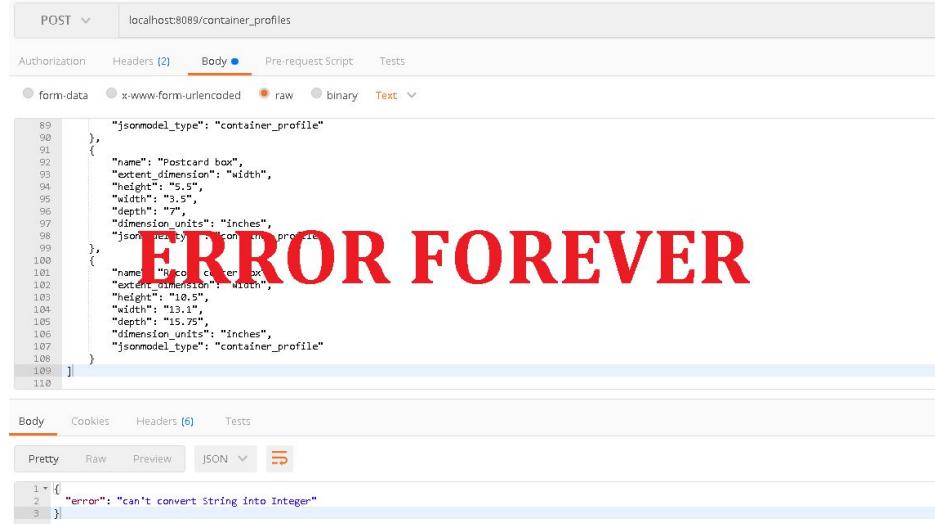
POST to AS with GUI - Container profiles

The screenshot shows the Postman application interface. At the top, there is a 'POST' button and a URL field containing 'localhost:8089/container_profiles'. Below the URL, there are tabs for 'Authorization', 'Headers (2)', 'Body (✓)', 'Pre-request Script', and 'Tests'. Under the 'Body' tab, there are five options: 'form-data', 'x-www-form-urlencoded', 'raw', 'binary', and 'Text'. The 'raw' option is selected and highlighted with a green circle. A 'Paste JSON' button is visible next to the raw input area. The main body area contains a JSON array with three objects, each representing a container profile:

```
[{"name": "Flat box01", "extent_dimension": "width", "height": "3", "width": "12", "depth": "16", "dimension_units": "inches", "jsonmodel_type": "container_profile"}, {"name": "Flat box02", "extent_dimension": "width", "height": "3", "width": "21", "depth": "25", "dimension_units": "inches", "jsonmodel_type": "container_profile"}, {"name": "Flat box03", "extent_dimension": "width", "height": "3", "width": "12", "depth": "16", "dimension_units": "inches", "jsonmodel_type": "container_profile"}]
```

Endpoint: http://localhost:8089/container_profiles

POST to AS with GUI - Container profiles



The screenshot shows a POST request to `localhost:8089/container_profiles`. The request body contains the following JSON:

```
89 },
90     "jsonmodel_type": "container_profile"
91 },
92     "name": "Postcard box",
93     "extent_dimension": "width",
94     "height": "5.5",
95     "width": "3.5",
96     "depth": "7",
97     "dimension_units": "inches",
98     "jsonmodel_json": "...."
99 },
100 [
101     {
102         "name": "Photo album box",
103         "extent_dimension": "width",
104         "height": "10.5",
105         "width": "13.1",
106         "depth": "15.75",
107         "dimension_units": "inches",
108         "jsonmodel_type": "container_profile"
109     }
110 ]
```

The response body shows an error message:

```
1 [
2     {
3         "error": "can't convert String into Integer"
4     }
5 ]
```

Don't hate us: you cannot post multiple records through the GUI
This frustrating exercise will save you a month
(use your month wisely - take a vacation from ArchivesSpace)

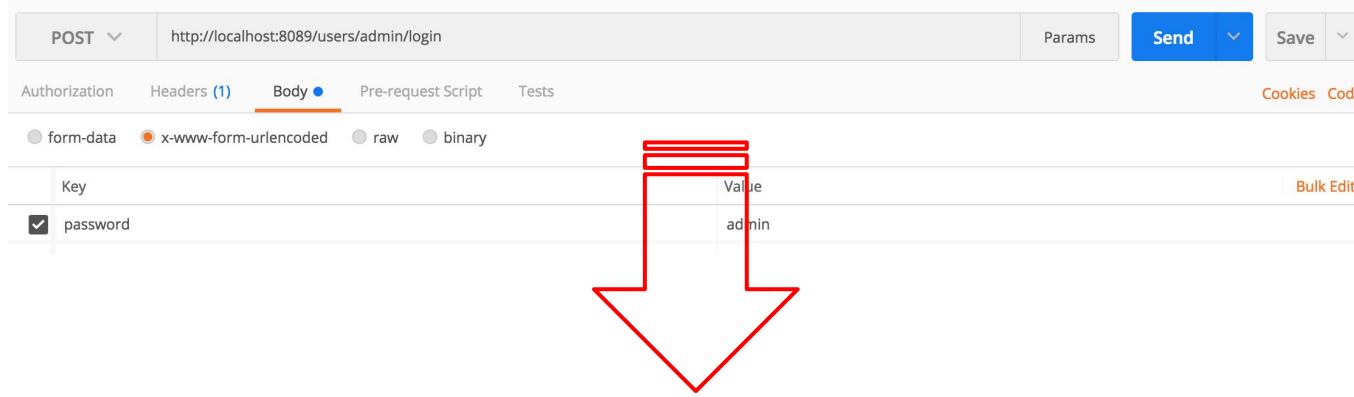


JUST
breathe

POST to AS with script- Container profiles

POST to AS with script

Before we start posting to AS, we need to authenticate, so how do we do that with scripts?



The screenshot shows the Postman interface with a POST request to `http://localhost:8089/users/admin/login`. The 'Body' tab is selected, showing a single form-data entry: 'password' with value 'admin'. A red arrow points from this entry down to a code editor window.

Key	Value
password	admin

Below the code editor, a red arrow points up to the 'password' entry in the Postman body table.

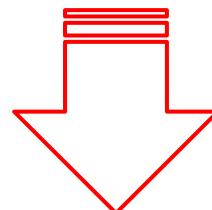
```
secrets.py
1 baseURL='http://localhost:8089'
2 user='admin'
3 password='admin'
```

```
secrets.py
1 baseURL='http://localhost:8089'
2 user='admin'
3 password='admin'
```

POST to AS with script

“Keep it secret, keep it safe.” - Gandalf

```
secrets.py  
1 baseURL='http://localhost:8089'  
2 user='admin'  
3 password='admin'
```



*This means no manual authenticating!
Learn to script just for that and call it a win!*

```
auth = requests.post(baseURL + '/users/' + user + '/login?password=' + password).json()  
session = auth["session"]  
headers = {'X-ArchivesSpace-Session': session}
```

POST to AS with script- Container Profiles

Mac

1. In the Finder navigate to your MARAC_API_Workshop directory
2. Ctrl+click the MARAC_API_Workshop directory, and select “New Terminal at Folder”
3. Type `ls` and examine the contents of that folder
4. Type `python postContainerProfiles.py` (case sensitive!)
5. Navigate back to AS in your browser (<http://localhost:8080>)

PC

1. Open Cygwin
2. Type `cd MARAC_API_Workshop` to enter the MARAC_API_Workshop directory
3. Type `ls` and examine the contents of that folder
4. Type `python postContainerProfiles.py` (case sensitive!)
5. Navigate back to AS in your browser (<http://localhost:8080>)

ArchivesSpace!

Scenario: You have successfully migrated into -- or have begun to use -- an instance of ASpace at your institution, but this is a short workshop, so here are our problems for today:

1. ~~We don't have container profiles in ArchivesSpace and would like to, so we need to create some~~
2. In following the migration instructions for 1.5, we had to add **faux codes**; we'd like to use our *actual* barcodes
3. Now that we have container profiles, we need to link them to *actual top containers*

POST to AS with script- Edit barcodes

Barcodes/top_containers

Répète: ASpace offers “container modeling” for the first time in the archives world.

Every type of box (ex. record carton) in your library gets its own record (a profile), which records its height, width, and depth.

Every actual box in your collections *also* gets a record, and this is called a top container. Simply put, this is the thing you put a number on: Box 1.

So, your archives might have hundreds or thousands of boxes called “Box 1”

Barcodes make that sane for AS. Hence, AS 1.5 requires some sort of unique code in every top container record.

Container 1: [31151030080422]		Top Container
Container Profile	Record center box	
Indicator	1	
Barcode	31151030080422	
Exported to ILS	Not exported	
Legacy	False	
Restricted?		

Barcodes/top_containers

Every marathon runner and every top_container must have a unique ID to participate.



Barcodes/top_containers

1. Navigate back to ASpace in your browser (<http://localhost:8080>)
2. Browse > Resources > Gérard Defaux papers > View > expand Research Materials > click on any file > scroll down to Instances > see fake barcode
3. These are the barcodes generated by the barcoder plugin. Hopkins has thousands of them.
4. Navigate to our [GitHub](#) and look at *barcodes.csv*

Barcodes/top_containers

If you're in ASpace, you will have some version of this problem, which is why we're featuring it.

Your top containers might:

- Have barcodes already! Well, this is still a lesson in editing records
- Have “faux codes,” like the ones in the AS vagrant
- Have nothing, and you have no idea where to start. We’ll have to refer you to the [AS 1.5 instructions](#) and [this plugin](#) by Chris Fitzpatrick

So let's fix our problem and imagine it working at scale.

POST to AS with script- Edit barcodes

Mac

1. In the Finder navigate to your MARAC_API_Workshop directory
2. Ctrl+click the MARAC_API_Workshop directory, and select “New Terminal at Folder”
3. Type `ls` and examine the contents of that folder
4. Type `python postBarcodes.py` (case sensitive!)
5. Navigate back to AS in your browser (<http://localhost:8080>)

PC

1. Open Cygwin
2. Type `cd MARAC_API_Workshop` to enter the MARAC_API_Workshop directory
3. Type `ls` and examine the contents of that folder
4. Type `python postBarcodes.py` (case sensitive!)
5. Navigate back to AS in your browser (<http://localhost:8080>)

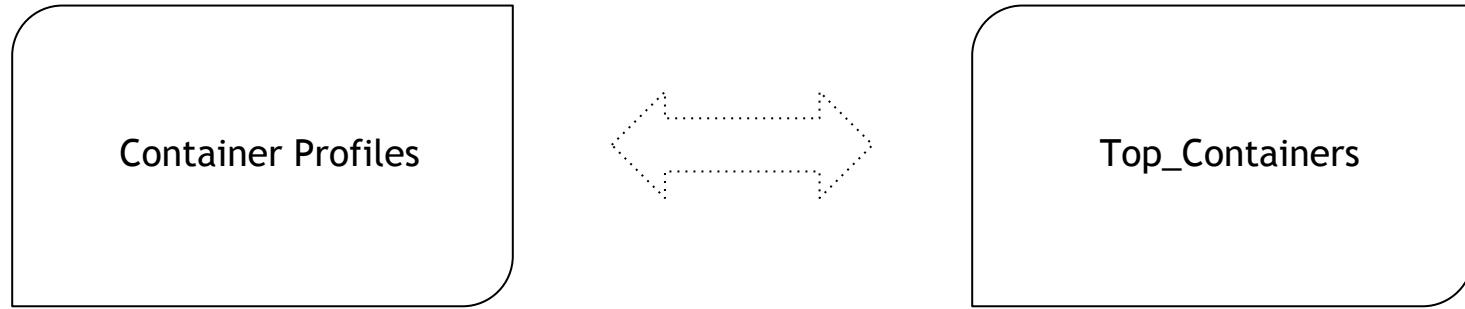
ArchivesSpace!

Scenario: You have successfully migrated into -- or have begun to use -- an instance of ASpace at your institution, but this is a short workshop, so here are our problems for today:

1. We don't have ~~container profiles~~ in ArchivesSpace and would like to, so we need to create some
2. In following the migration instructions for 1.5, we had to add ~~faux codes~~; we'd like to use our *actual* barcodes
3. Now that we have container profiles, we need to link them to *actual top containers*

POST to AS with script- Link profiles

Linking profiles to containers



POST to AS with script - Linking profiles

1. Type `ls` and examine the contents of that folder
2. Type `python asLinkProfiles.py` (case sensitive!)
3. You will be prompted for a resource id and a container id... how do you determine what you need to know?
4. Let's return to ASpace real quick

POST to AS with script - Linking profiles

The interface – just another lens on the same data – is helpful for constructing API requests.

View a resource record for its resource number:

A screenshot of the ArchivesSpace interface. The title bar says "ArchivesSpace | Resource". The address bar shows the URL "localhost:8080/resources/1#tree::resource_1". A red circle highlights this URL. Below the address bar, the ArchivesSpace logo and the word "ArchivesSpace" are displayed. The main content area shows a breadcrumb navigation: "Home / Resources / Gérard Defaux papers". Underneath, there's a tree view with "Gérard Defaux papers" expanded, showing "Research Materials" as a child node. At the bottom, there are "Browse" and "Create" dropdown menus, a "Search All Records" input field, and a "Record center box" containing "Basic Information" and "Edit" buttons.

View a container profile for its profile number:

A screenshot of the ArchivesSpace interface. The title bar says "ArchivesSpace | Container Profile". The address bar shows the URL "localhost:8080/container_profiles/12". A red circle highlights this URL. Below the address bar, the ArchivesSpace logo and the word "ArchivesSpace" are displayed. The main content area shows a breadcrumb navigation: "Home / Container Profiles / Record center box". Underneath, there's a "Record center box" containing "Basic Information" and "Edit" buttons. To the right, a separate panel titled "Record center box" shows the "Basic Information" tab.

POST to AS with script - Linking profiles

1. Type `ls` and examine the contents of that folder
2. Type `python asLinkProfiles.py` (case sensitive!)
3. You will be prompted for a resource id and a container id... how do you determine what you need to know?
4. Let's return to ASpace real quick

ArchivesSpace! You did it!

Scenario: You have successfully migrated into -- or have begun to use -- an instance of ASpace at your institution, but this is a short workshop, so here are our problems for today:

1. We don't have ~~container profiles~~ in ArchivesSpace and would like to, so we need to create some
2. In following the migration instructions for 1.5, we had to add ~~faux codes~~; we'd like to use our *actual* barcodes
3. Now that we have ~~container profiles~~, we need to link them to *actual* ~~top~~ containers

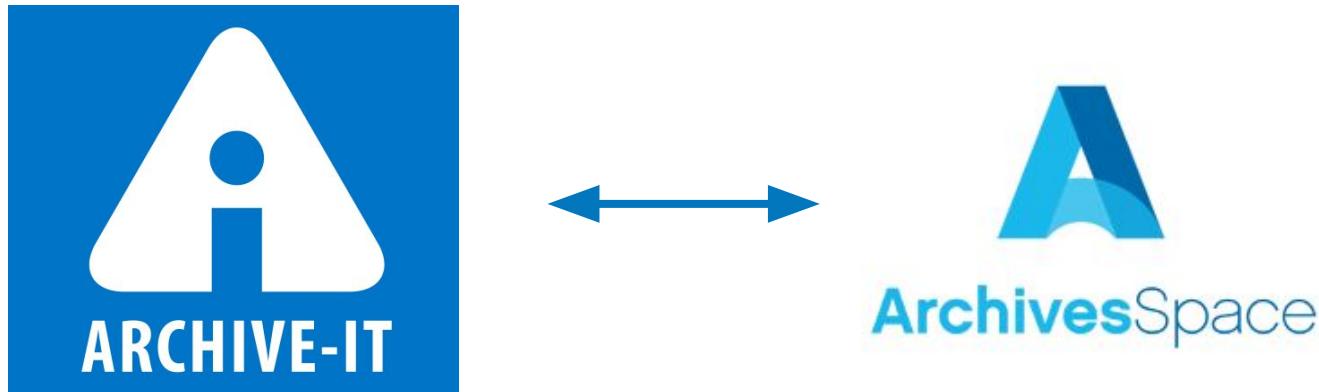


15 minute break!

GET and POST across two applications with Python

App-to-app Communication

Scenario: As your university's web archivist, you wish to make your Archive-It web crawls accessible to users who access your collections via ArchivesSpace without having to individually create digital objects every time you run a new Archive-It crawl.



App-to-app Communication



ArchivesSpace

1. In ArchivesSpace, navigate to the “Records of the Johns Hopkins University Library” resource
2. Expand Subgroup 12: Library Website
3. Click on library.jhu.edu
4. Note that archival object’s level

App-to-app Communication

We now know that we can access ArchivesSpace's archival object records via the **ArchivesSpace API**, right?



In fact, with a decent enough search we could probably even have a script return JUST those archival objects with the level "**Web archive**."

Since we're going to want to keep programmatically working with/altering this data after we find it, we'll use a **Python script**, instead of Postman to run this search.

```
39  # search AS for archival_object's with level "Web archive"
40  query =
41  •   '/search?page=1&filter_term[]={"primary_type":"archival_object"}&filter_term[]={"level":"Web"
42  •   "archive"})'
43  ASoutput = requests.get(baseURL + query, headers=headers).json()
44  print 'Found ' + str(len(ASoutput['results'])) + ' archival objects with the instance type "Web
45  •   archive."'
```

Code snippet from archivelt.py

App-to-app Communication



The screenshot shows the Archive-It homepage. At the top left is the Archive-It logo (a white 'A' inside a blue square). To its right are four navigation links: HOME, EXPLORE, LEARN MORE, and CONTACT US. Further to the right are social media icons for Facebook, Twitter, and WordPress, followed by a 'Login' button. Below the navigation is a large text block: 'The leading web archiving service for collecting and accessing cultural heritage on the web' with the note 'Built at the Internet Archive'. To the right of this text is a small icon of a classical building.

Narrow Your Results

Type of Collecting Organization

Sort By: Count | (A-Z)

Colleges & Universities

Collecting Organization

Sort By: Count | (A-Z)

Johns Hopkins University (11)

Explore All Archives

Items in the archive are listed below. Narrow your results at left, or enter a search query below to find a collecting organization, collection, site, specific URL or to search the text of archived webpages.

Collection Name : Johns Hopkins University web collection 

The following results were found for the term(s): library.jhu.edu

- 11 Sites were found.
- Additional results for library.jhu.edu may be found by searching [within the page text](#).

Sites

Search Page Text

Page 1 of 1 (11 Total Results)

Sort By: Best Match | Title (A-Z) | Title (Z-A) | URL (A-Z) | URL (Z-A)

URL: <http://library.jhu.edu>

Collection: Johns Hopkins University web collection

Organization: Johns Hopkins University

Captured 43 times between Aug 20, 2010 and Feb 28, 2017



App-to-app Communication

Does **Archive-It** have an **API** we can use to access this information we're seeing in our browsers?



YES!



Wayback Machine APIs

The Internet Archive Wayback Machine supports a number of different APIs to make it easier for developers to retrieve information about Wayback capture data.

The following is a listing of currently supported APIs. This page is subject to change frequently, please check back for the latest info.

Updated on September, 24, 2013

Wayback Availability JSON API

This simple API for Wayback is a test to see if a given url is archived and currently accessible in the Wayback Machine. This API is useful for providing a 404 or other error handler which checks Wayback to see if it has an archived copy ready to display. The API can be used as follows:

<http://archive.org/wayback/available?url=example.com>

which might return:

```
{  
  "archived_snapshots": {  
    "closest": {  
      "available": true,  
      "url": "http://web.archive.org/web/20130919044612/http://example.com/",  
      "timestamp": "20130919044612",  
      "status": "200"  
    }  
  }  
}
```

App-to-app Communication



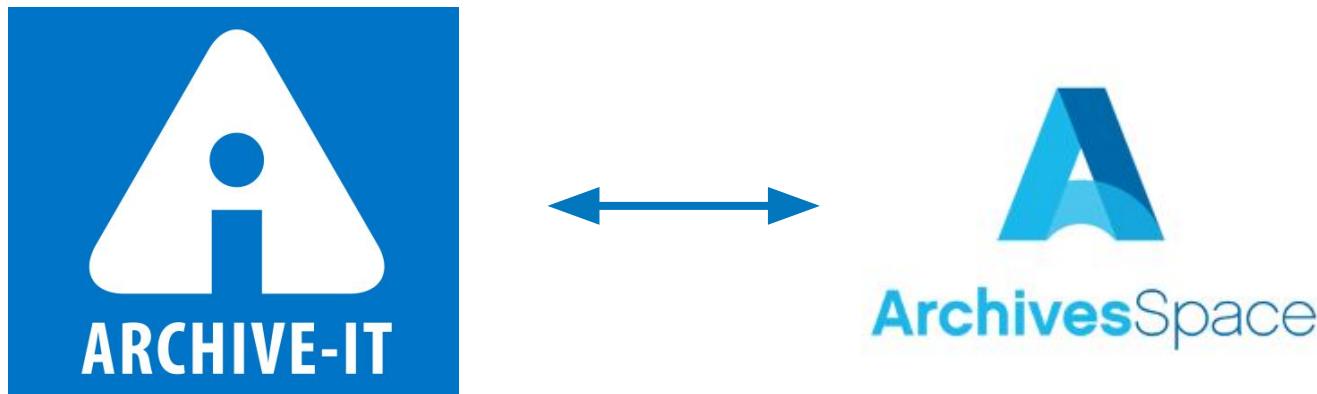
With the right amount of trial and error, we can also get information about our Archive-It holdings out of the Archive-It API with a **Python script** as well!

```
67     for crawl in crawlList:  
68         doid = 'https://wayback.archive-it.org' + '/' + archiveit_coll + '/' +  
69         crawl['timestamp'] + '/' + crawl['original']  
70         query = '/search?page=1&filter_term[]={"primary_type":"digital_object"}&q=' + doid  
71         existingdOID = requests.get(baseURL + query, headers=headers).json()  
72         if len(existingdOID['results']) > 0:
```

Code snippet from archivelt.py

App-to-app Communication

1. In Terminal/Cygwin run `python archiveIt.py` and let's see what happens!
2. Go check out that “Records of the Johns Hopkins University Library” resource record once again.



User stories

Can the API create reports?

- Yes... but an API is the wrong tool for reporting
- Like using a jack-hammer indoors: yes it will work, but it will be more effort with the wrong tool
- For AS users in particular:
 - Wait for reporting to improve, because it will
 - But in the meantime (or for more customized reporting) we suggest connecting AS to MS Access. If you don't know Access, it's easier to learn that than learning to script just for reports.
 - You can find explanatory slides in the workshop GitHub > additional resources
 - Credit to Nancy Enneking, Head of Institutional Records at the Getty and Celia Caust-Ellenbogen, Friends Historical Library of Swarthmore College for the method

Can APIs improve my agent records?

- Yes!
- Hopkins made changes to almost all of our Agent and Subject records through the API
- We added VIAF ids to Agents and converted LCSH to FAST
- You already have improved Corporate Names from earlier
- Let's put those to work!

POST with script - VIAF

Post from a CSV - VIAF

```
1  [
2    {
3      "place_of_publication": "Richmond, Va.",
4      "lccn": "sn85038615",
5      "start_year": "1903",
6      "place": [
7        {
8          "name": "The times dispatch.",
9          "publisher": "times-Dispatch Co.",
10         "url": "http://chroniclingamerica.loc.gov/lccn/sn85038615.json",
11         "end_year": "1914",
12         "issues": [
13           {
14             "issue": [
15               {
16                 "url": "http://chroniclingamerica.loc.gov/lccn/sn85038615/1903-01-27/ed-1.json",
17                 "date_issued": "1903-01-27"
18               },
19               {
20                 "url": "http://chroniclingamerica.loc.gov/lccn/sn85038615/1903-01-28/ed-1.json",
21                 "date_issued": "1903-01-28"
22               },
23               {
24                 "url": "http://chroniclingamerica.loc.gov/lccn/sn85038615/1903-01-29/ed-1.json",
25                 "date_issued": "1903-01-29"
26               },
27               {
28                 "url": "http://chroniclingamerica.loc.gov/lccn/sn85038615/1903-01-30/ed-1.json",
29                 "date_issued": "1903-01-30"
30               },
31               {
32                 "url": "http://chroniclingamerica.loc.gov/lccn/sn85038615/1903-01-31/ed-1.json",
33                 "date_issued": "1903-01-31"
34               },
35               {
36                 "url": "http://chroniclingamerica.loc.gov/lccn/sn85038615/1903-02-01/ed-1.json",
37                 "date_issued": "1903-02-01"
38               },
39               {
40                 "url": "http://chroniclingamerica.loc.gov/lccn/sn85038615/1903-02-02/ed-1.json",
41                 "date_issued": "1903-02-02"
42               },
43               {
44                 "url": "http://chroniclingamerica.loc.gov/lccn/sn85038615/1903-02-04/ed-1.json",
45                 "date_issued": "1903-02-04"
46               },
47               {
48                 "url": "http://chroniclingamerica.loc.gov/lccn/sn85038615/1903-02-05/ed-1.json",
49                 "date_issued": "1903-02-05"
50               },
51               {
52                 "url": "http://chroniclingamerica.loc.gov/lccn/sn85038615/1903-02-06/ed-1.json",
53                 "date_issued": "1903-02-06"
54               },
55               {
56                 "url": "http://chroniclingamerica.loc.gov/lccn/sn85038615/1903-02-07/ed-1.json",
57                 "date_issued": "1903-02-07"
58               }
59             ],
60             "date_founded": "1903-01-01"
61           }
62         ]
63       }
64     ]
65   ]
```

A	B
1 url	date_issued
2 http://chroniclingamerica.loc.gov/lccn/sn85038615/1903-01-27/ed-1.json	1903-01-27
3 http://chroniclingamerica.loc.gov/lccn/sn85038615/1903-01-28/ed-1.json	1903-01-28
4 http://chroniclingamerica.loc.gov/lccn/sn85038615/1903-01-29/ed-1.json	1903-01-29
5 http://chroniclingamerica.loc.gov/lccn/sn85038615/1903-01-30/ed-1.json	1903-01-30
6 http://chroniclingamerica.loc.gov/lccn/sn85038615/1903-01-31/ed-1.json	1903-01-31
7 http://chroniclingamerica.loc.gov/lccn/sn85038615/1903-02-01/ed-1.json	1903-02-01
8 http://chroniclingamerica.loc.gov/lccn/sn85038615/1903-02-03/ed-1.json	1903-02-03
9 http://chroniclingamerica.loc.gov/lccn/sn85038615/1903-02-04/ed-1.json	1903-02-04
10 http://chroniclingamerica.loc.gov/lccn/sn85038615/1903-02-05/ed-1.json	1903-02-05
11 http://chroniclingamerica.loc.gov/lccn/sn85038615/1903-02-06/ed-1.json	1903-02-06
12 http://chroniclingamerica.loc.gov/lccn/sn85038615/1903-02-07/ed-1.json	1903-02-07
13 http://chroniclingamerica.loc.gov/lccn/sn85038615/1903-02-08/ed-1.json	1903-02-08
14 http://chroniclingamerica.loc.gov/lccn/sn85038615/1903-02-10/ed-1.json	1903-02-10
15 http://chroniclingamerica.loc.gov/lccn/sn85038615/1903-02-11/ed-1.json	1903-02-11
16 http://chroniclingamerica.loc.gov/lccn/sn85038615/1903-02-12/ed-1.json	1903-02-12
17 http://chroniclingamerica.loc.gov/lccn/sn85038615/1903-02-13/ed-1.json	1903-02-13
18 http://chroniclingamerica.loc.gov/lccn/sn85038615/1903-02-14/ed-1.json	1903-02-14
19 http://chroniclingamerica.loc.gov/lccn/sn85038615/1903-02-15/ed-1.json	1903-02-15
20 http://chroniclingamerica.loc.gov/lccn/sn85038615/1903-02-17/ed-1.json	1903-02-17
21 http://chroniclingamerica.loc.gov/lccn/sn85038615/1903-02-18/ed-1.json	1903-02-18
22 http://chroniclingamerica.loc.gov/lccn/sn85038615/1903-02-19/ed-1.json	1903-02-19
23 http://chroniclingamerica.loc.gov/lccn/sn85038615/1903-02-20/ed-1.json	1903-02-20
24 http://chroniclingamerica.loc.gov/lccn/sn85038615/1903-02-21/ed-1.json	1903-02-21
25 http://chroniclingamerica.loc.gov/lccn/sn85038615/1903-02-22/ed-1.json	1903-02-22
26 http://chroniclingamerica.loc.gov/lccn/sn85038615/1903-02-24/ed-1.json	1903-02-24
27 http://chroniclingamerica.loc.gov/lccn/sn85038615/1903-02-25/ed-1.json	1903-02-25
28 http://chroniclingamerica.loc.gov/lccn/sn85038615/1903-02-26/ed-1.json	1903-02-26
29 http://chroniclingamerica.loc.gov/lccn/sn85038615/1903-02-27/ed-1.json	1903-02-27
30 http://chroniclingamerica.loc.gov/lccn/sn85038615/1903-02-28/ed-1.json	1903-02-28

Way back during the Chronicling America exercise, we showed you how to convert JSON to CSV...

Post from a CSV - VIAF

	total_results	organizations_ein	organizations_strein	organizations_name	organizations_sub_name	organizations_city	organizations_state	organizations_ntee_code	organizations_raw_ntee_code	organizations_
001	2435	731663130	73-1663130	ANIMAL ALLIANCE	ANIMAL ALLIANCE	WOODLAND HLS	CA	D20	D20	3
002		752461428	75-2461428	ANIMAL ANGELS	ANIMAL ANGELS	JACKSBORO	TX	null	null	3
003		232896507	23-2896507	ANIMAL APPEAL	ANIMAL APPEAL	SHARPSVILLE	PA	O21	O21	3
004		43654364	04-3654364	ANIMAL ASSISTANCE	ANIMAL ASSISTANCE	E BRUNSWICK	NJ	D20	D20	3
005		880099106	88-0099106	ANIMAL FAIR	ANIMAL FAIR	SEALDALE	MO	D20	D20	3
006		752391827	75-2391827	ANIMAL HAVEN	ANIMAL HAVEN	DALLAS	TX	null	null	3
007		953793138	95-3793138	ANIMAL HELPINE	ANIMAL HELPINE	MORONGO VLY	CA	D200	D200	3
008		570921521	57-0921521	ANIMAL MISSION	ANIMAL MISSION	COLUMBIA	SC	D200	D200	3
009		462455876	46-2455876	ANIMAL NETWORK	ANIMAL NETWORK	LAS VEGAS	NV	D20	D20	3
010		18462882	01-8462882	ANIMAL ORPHANAGE	ANIMAL ORPHANAGE	ORONO	ME	D200	D200	3
011		454902341	45-4902341	ANIMAL PAD	ANIMAL PAD	SAN DIEGO	CA	D20	D20	3
012		818679304	81-0679304	ANIMAL ADVOCATES	ANIMAL ADVOCATES	BARNWELL	SC	D11	D11	3
013		20469900	82-0469900	ANIMAL ALLIES	ANIMAL ALLIES	MANCHESTER	NH	D20	D20	3
014		680630714	68-0630714	ANIMAL BALANCE	ANIMAL BALANCE	EL CERRITO	CA	Z99	Z99	3
015		953967863	95-3967863	ANIMAL COMPANIONS	ANIMAL COMPANIONS	PIPE CREEK	TX	null	null	3
016		237237862	23-7237862	ANIMAL FUND	ANIMAL FUND	GREENBRAE	CA	D813	D813	3
017		204444813	20-4444813	ANIMAL LIFELINE	ANIMAL LIFELINE	WARRINGTON	PA	D20	D20	3
018		680208668	68-0208668	ANIMAL PLACE	ANIMAL PLACE	GRASS VALLEY	CA	D200	D200	3
019		880610723	88-0610723	ANIMAL UNION	ANIMAL UNION	DALY CITY	CA	P80	P80	3
020		260796104	26-0796104	ANIMAL ADVOCATES	ANIMAL ADVOCATES	WESTCLIFFE	CO	D20	D20	3
021		223451569	22-3451569	ANIMAL ALLIES	ANIMAL ALLIES	EWING	NJ	D200	D200	3
022		571098821	57-1098821	ANIMAL ALLIES	ANIMAL ALLIES	SPARTANBURG	SC	D200	D200	3
023		208502915	20-8502915	ANIMAL ANGELS	ANIMAL ANGELS	MEDIALPOLIS	IA	D20	D20	3
024		262933425	26-2933425	ANIMAL DIPLOMACY	ANIMAL DIPLOMACY	CAMBRIDGE	MA	D20	D20	3
025		10714781	01-0714781	ANIMAL TRACKS	ANIMAL TRACKS	ACTION	CA	D34	D34	3
026		454027119	45-4027119	ANIMAL AID	ANIMAL AID	ATHENS	TN	D40	D40	3
027		272034873	27-2034873	ANIMAL WATCH	ANIMAL WATCH	BOULDER	CO	D30	D30	3
028		481287089	48-1287089	ANIMAL ADVOCATES	ANIMAL ADVOCATES	LOS ANGELES	CA	D20	D20	3
029		890530102	89-0530102	ANIMAL FOLKS	ANIMAL FOLKS	SAINT PAUL	MN	D20	D20	3
030		371768388	37-1768388	ANIMAL FOUNDATION	ANIMAL FOUNDATION	HERMANN	MO	P29	P29	3
031		203596003	20-3596003	ANIMAL MAGIC	ANIMAL MAGIC	BELLEVILLE	MI	D20	D20	3
032		953171867	95-3171867	ANIMAL SANITARIANS	ANIMAL SANITARIANS	THOUSAND PLMS	CA	D200	D200	3

```
- "organizations": [Array[100]
  -0: {
    "ein": "731663130",
    "strein": "73-1663130",
    "name": "ANIMAL ALLIANCE",
    "sub_name": "ANIMAL ALLIANCE",
    "city": "WOODLAND HLS",
    "state": "CA",
    "ntee_code": "D20",
    "raw_ntee_code": "D20",
    "subeccd": 3,
    "has_subeccd": true,
    "have_filings": true,
    "have_extracts": true,
    "have_pdfs": true,
    "score": 443.46564
  }
  -1: {
    "ein": "752461428",
    "strein": "75-2461428",
    "name": "ANIMAL ANGELS",
    "sub_name": "ANIMAL ANGELS",
    "city": "JACKSBORO",
    "state": "TX",
    "ntee_code": null,
    "raw_ntee_code": null,
    "subeccd": 3,
    "has_subeccd": true,
    "have_filings": true,
    "have_extracts": true,
    "have_pdfs": true,
    "score": 443.46564
  }
  -2: {
    "ein": "232896507",
    "strein": "23-2896507",
    "name": "ANIMAL APPEAL",
    "sub_name": "ANIMAL APPEAL",
    "city": "SHARPSVILLE",
    "state": "PA",
    "ntee_code": "O21",
    "raw_ntee_code": "O21",
    "subeccd": 3,
    "has_subeccd": true,
    "have_filings": true,
    "have_extracts": true,
    "have_pdfs": true,
    "score": 443.46564
  }
]
```

...well, we can also convert a csv (like the file that resulted from running `python viafReconciliationCorporate.py`) to JSON!
And ArchivesSpace likes eating JSON!

Post from a CSV - VIAF

Before posting these VIAF corporate entities into ArchivesSpace, we must first add VIAF as a valid “name source” in ArchivesSpace

1. Go to the ArchivesSpace staff interface at: <http://localhost:8080>
2. In the top right select “System > Manage Controlled Value Lists”
3. In the drop-down you’re provided, select “Name Source (name_source)”
4. In the middle right click “Create Value”
5. Name this value “vaf”
 - a. Note: The punctuation here is important since it must match what is in our Python script. All lowercase, no spaces.
6. Click “Create Value”

Post from a CSV - VIAF

With the VIAF name source added to ArchivesSpace, next:

1. Confirm you still have a file called “viafCorporateResults.csv” in your MARAC_API_Workshop directory
2. From within the MARAC_API_Workshop directory in Terminal/Cygwin execute `python postVIAFOrganizations.py`
3. Go back to the ArchivesSpace staff interface at: <http://localhost:8080>
4. In the top left click “Browse > Agents”
5. Voila!

Icing and Advice

Icing: Interpreting (ASpace) API endpoints

<http://archivesspace.github.io/archivesspace/api>

GET /repositories/:repo_id/resources/:id

Description

Get a Resource

Parameters

Integer id – The ID of the record

Integer repo_id – The Repository ID – The Repository must exist

[String] resolve – A list of references to resolve and embed in the response

Returns

200 – (:resource)

```
curl -H "X-ArchivesSpace-Session: $SESSION" "http://localhost:8089/repositories/:repo_id/resources/1"
```

Icing: Interpreting (ASpace) API endpoints

The interface – just another lens on the same data – is helpful for constructing API requests.

Determining the repository number:

http://archivesspace.fakeu.edu/repositories/3

Special Collections

Repository is Currently Selected

Determining an agent number:

archivesspace.fakeu.edu/agents/agent_person/87

Abbe, Cleveland, 1838-1916

Basic Information

Agent Type: Person
Publish: True

Created by admin 2016-05-13 1

Dates of Existence

Existence 1838 – 1916

Sample endpoint from documentation: http://localhost:8089/repositories/:repo_id/resources/1

Example “fake” endpoint that mimics real life: <http://archivesspace.fakeu.edu:8089/repositories/3/resources/1>

http://localhost:8089	The address of your instance of ASpace. You will ONLY replace “local host,” the colon and port number remain. EX. http://archivesspace.fakeu.edu:8089
<code>/repositories</code>	The presence of “repositories” here means that this endpoint is repository-specific. Some non-repo specific requests in AS are for Agents and Access Points, which span all of AS. EX. http://archivesspace.fakeu.edu:8089/agents
<code>:repo_id</code>	The presence of this colon means this value will be unique to your institution. How can you determine the repository number? You can use the repo endpoint, or, from within AS navigate Systems > Manage Repositories > select repository > and look at the address bar. EX. http://archivesspace.fakeu.edu:8089/repositories/3
<code>/resources</code>	Other examples are /accessions or /top_containers. EX. http://archivesspace.fakeu.edu:8089/repositories/3/accessions
<code>/1</code>	The first resource. How can you determine resource numbers? Navigate to the resource in the interface and its number will be in the address bar. EX. http://archivesspace.fakeu.edu:8089/repositories/3/resources/1

Icing: What IS GitHub anyway?

The least most helpful thing you'll hear is, "It's in our GitHub!"

If you're serious about learning to script, you should watch the 10 million GitHub intro videos on YouTube

Even casual users will benefit from using other people's scripts (that's how devs work!)

Let's go look at our repo together, we made it for you!

https://github.com/jhu-archives-and-manuscripts/MARAC_API_Workshop

Icing: There will ALWAYS be “gotchas”

We purposely made you “fail” a few times today. Get used to it!

- You WILL not succeed on the first try.
- You WILL hit unanticipated snafus, oftentimes due to data models and/or poorly written documentation (aka, due to no fault of your own!).
- You WILL be fitter, happier, and more productive if you start building a community now and asking questions.

Icing: A frequent ASpace “gotcha”

it·er·a·tion

īdē' rāSH(ə)n/

noun

1. the repetition of a process or utterance.
 - repetition of a mathematical or computational procedure applied to the result of a previous application, typically as a means of obtaining successively closer approximations to the solution of a problem.

Lock version - a value that incrementally increases every time an AS record is altered. In practice, this means work cannot and should not continue on the data in question, i.e. your team has to stop work

```
"lock_version": 5,  
"title": "university history scrapbook collection",  
"publish": true,
```

Icing: A frequent ASpace frustration

Session time and page limits

- Ever been timed out of your bank account? Frustrating but vital
- The amount of time you have after authenticating is called “session time”
- ASpace default is very short
- Ask your ASpace tech person to up the session time in the AS config; we’ve provided instructions in the take home document (which really will happen, we promise)

Icing: What about XYZ application?

In your pre-workshop surveys, many of you (understandably) noted applications we didn't get a chance to cover today, notably:

- ILSes like [Voyager](#);
- Digital exhibition applications like Omeka;
- Digital collections/repository applications like ContentDM or Islandora; and,
- Cloud-based sharing applications like Dropbox/Drive/Box.

While we didn't work through exercises with these applications today, hopefully you now know the **steps to take** to do future API work of your own, namely:

- Research the API, including authentication requirements and endpoint documentation;
- Play around in Postman (or another GUI API application);
- Determine whether your desired tasks can be accomplished through the GUI, or if you need a scripting language;
 - If the latter, determine if someone else has already tackled your task (for example, on GitHub)
- Iteratively test (in a *non-Production environment!*)
- Profit!

Wrap up!

Shutting down your Vagrant box

Do the following in Terminal/the command prompt:

1. Type `vagrant destroy` and confirm the command by typing `y` when prompted
 - a. “Destroying” is equivalent to hitting “shut down” on your computer; the computer is still sitting on your desk with your work saved, but it is powered off.
2. Type `vagrant box remove aspace` to remove the box completely
 - a. If “destroying” a box is like shutting down, “removing” a box is like taking a time machine back to the day you first cut the shrink wrap on your brand new computer. All work you have done in the virtual environment will be lost, but, once the box is re-added and upped again, you will be back to the exact version we shared with you here.

Thanks!

The Vagrant developed for this workshop was adapted by Lora Davis from **Dallas Pallen's Archivagrant**. See his blog post for more information:

<http://archival-integration.blogspot.com/2016/01/archivesspace-vagrant-archivagrant.html>

Additional Vagrant advice was offered by **Francis Kayiwa**.

The inspiration (if not the flawed code) for the ArchivIt-ArchivesSpace exercise comes from the fantastic work of **Greg Wiedeman**. Check it out on GitHub:

https://github.com/UAlbanyArchives/ASpace_WebArchives

Thanks to **Eric Hanson**, Metadata Librarian at JHU, for his troubleshooting and all-around good spirits (YAY ERIC!)

Several of these concepts and/or exercises were tested on JHU Sheridan Libraries' staff, including NDSR Resident **Elizabeth England**.

Anyone who has ever shared help/advice/support on blogs/listservs/bar stools who are too numerous to fully name (but we'll try in person if you ask us to!).

Technical pitstop: Command line bootcamp

- A few very simple command line commands are necessary in this workshop
 - But more important is being able to use them effectively
 - Mac users, and PC users in Cygwin, will be using the same commands
-

Command line bootcamp: Navigation

Where are you, and where do you want to go?

Mac

PC

1. In the Finder navigate to your MARAC_API_Workshop directory
2. Ctrl+click the MARAC_API_Workshop directory, and select “New Terminal at Folder”

1. Open Cygwin

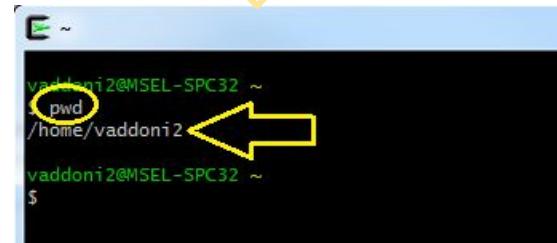
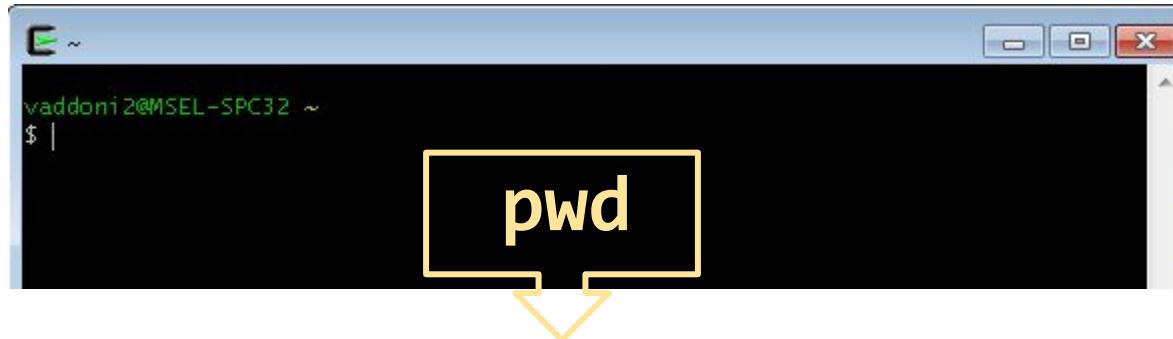
Command line bootcamp: Navigation

Where are you, and where do you want to go?

PC

Here is yonder cygwin prompt. Your first instinct may be to ask, “But where am I?”

Where am I? = “print working directory” = `pwd`



Command line bootcamp: Navigation

Where are you, and where do you want to go?

Where am I? = “print working directory” = `pwd`

What else is here? = “list” = `ls`

Move from here to there = “change directory” = `cd` [type the name of the directory]

Move up one level = `cd ..`

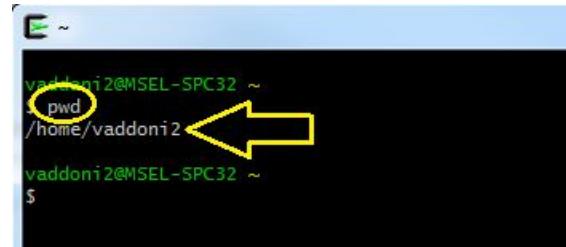
Command line bootcamp: Navigation

Where are you, and where do you want to go?

PC

But where is that? This is non-intuitive, but *you're already in C:\cygwin* because you're using the cygwin window, so

This location:



```
vaddoni2@MSEL-SPC32 ~  
$ pwd  
/home/vaddoni2
```

A screenshot of a terminal window titled 'vaddoni2'. The prompt shows the user is at the root directory (~). The command 'pwd' is entered, and the output is '/home/vaddoni2'. A yellow circle highlights the command 'pwd', and a yellow arrow points from it to the word 'pwd' in the question above.

Is this location:



Command line bootcamp: Navigation

Where are you, and where do you want to go?

Mac users

Also try out `pwd`, you should be on your desktop.

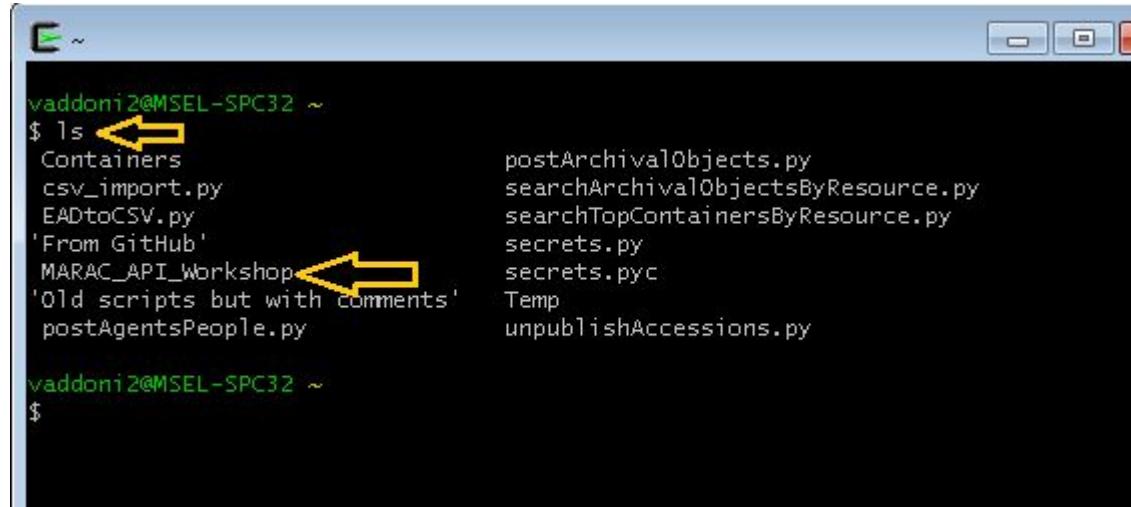
Command line bootcamp: Navigation

Where are you, and where do you want to go?

Now let's move *from* where you are *into* the MARAC API Workshop clone folder:

So let me make sure that folder is really here. I check by listing the contents of the directory with `ls` (that is LS).

What else is here? = “list” = `ls`



```
vaddoni2@MSEL-SPC32 ~
$ ls
Containers
csv_import.py
EADtoCSV.py
'From GitHub'
MARAC_API_Workshop
'Old scripts but with comments'
postAgentsPeople.py

postArchivalObjects.py
searchArchivalObjectsByResource.py
searchTopContainersByResource.py
secrets.py
secrets.pyc
Temp
unpublishAccessions.py

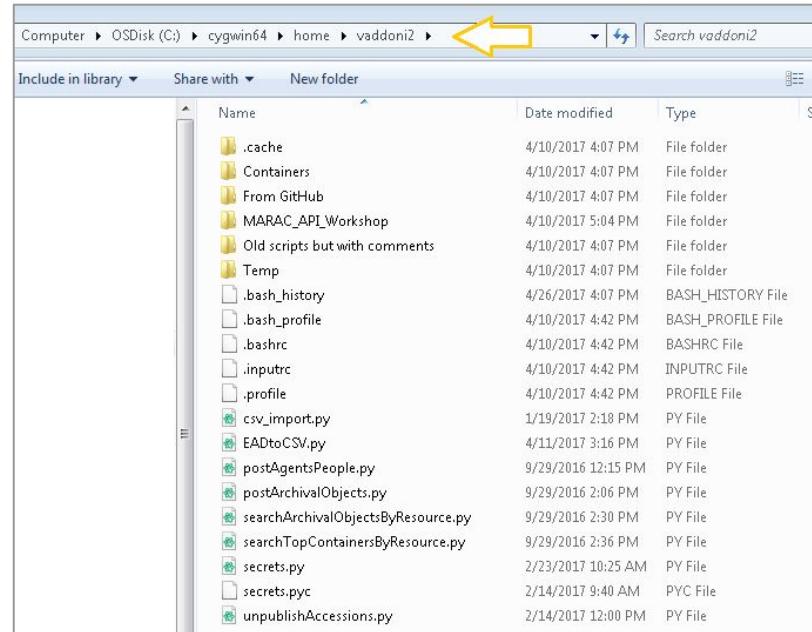
vaddoni2@MSEL-SPC32 ~
$
```

Command line bootcamp: Navigation

Where are you, and where do you want to go?

PC

Look, that is the same list of contents that I see if I navigate to C:\cygwin64\home\vaddoni2 in Windows:



Name	Date modified	Type
.cache	4/10/2017 4:07 PM	File folder
Containers	4/10/2017 4:07 PM	File folder
From GitHub	4/10/2017 4:07 PM	File folder
MARAC_API_Workshop	4/10/2017 5:04 PM	File folder
Old scripts but with comments	4/10/2017 4:07 PM	File folder
Temp	4/10/2017 4:07 PM	File folder
.bash_history	4/26/2017 4:07 PM	BASH_HISTORY File
.bash_profile	4/10/2017 4:42 PM	BASH_PROFILE File
.bashrc	4/10/2017 4:42 PM	BASHRC File
.inputrc	4/10/2017 4:42 PM	INPUTRC File
.profile	4/10/2017 4:42 PM	PROFILE File
csv_import.py	1/19/2017 2:18 PM	PY File
EADtoCSV.py	4/11/2017 3:16 PM	PY File
postAgentsPeople.py	9/29/2016 12:15 PM	PY File
postArchivalObjects.py	9/29/2016 2:06 PM	PY File
searchArchivalObjectsByResource.py	9/29/2016 2:30 PM	PY File
searchTopContainersByResource.py	9/29/2016 2:36 PM	PY File
secrets.py	2/23/2017 10:25 AM	PY File
secrets.pyc	2/14/2017 9:40 AM	PYC File
unpublishAccessions.py	2/14/2017 12:00 PM	PY File

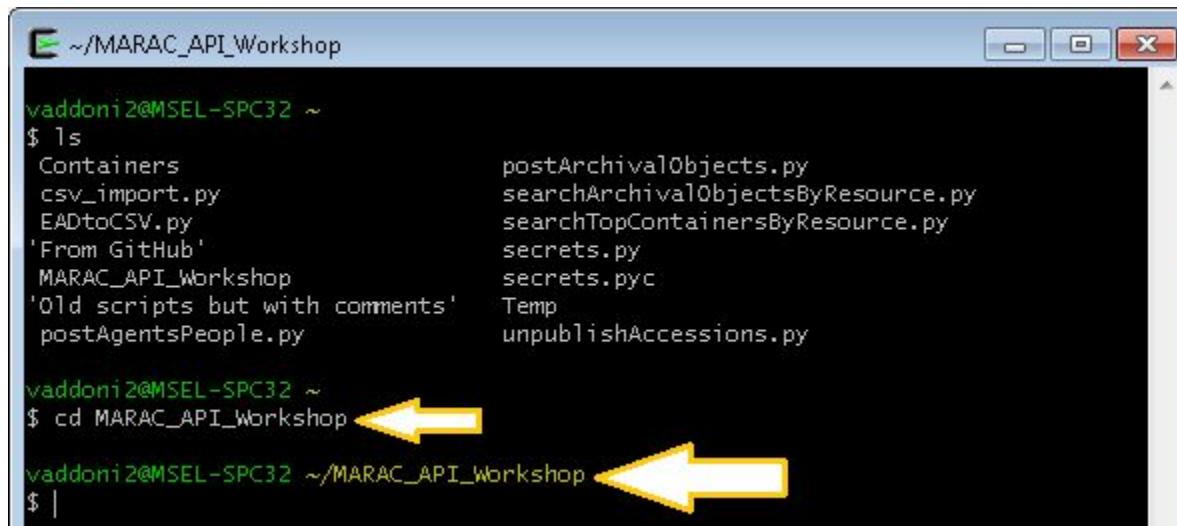
Command line bootcamp: Navigation

Where are you, and where do you want to go?

PC

Now I'm going to move *from* the Cygwin home directory *into* the MARAC API Workshop clone folder:

To move into that folder I'll type `cd` (change directory), leave a space, and then the name of the directory I want to go into: `cd MARAC_API_Workshop`



```
vaddoni2@MSEL-SPC32 ~
$ ls
Containers          postArchivalObjects.py
csv_import.py       searchArchivalObjectsByResource.py
EADtoCSV.py        searchTopContainersByResource.py
'From GitHub'
MARAC_API_Workshop secrets.py
'Old scripts but with comments'
postAgentsPeople.py secrets.pyc
Temp
unpublishAccessions.py

vaddoni2@MSEL-SPC32 ~
$ cd MARAC_API_Workshop
vaddoni2@MSEL-SPC32 ~/MARAC_API_Workshop
$ |
```

Command line bootcamp: Navigation

Where are you, and where do you want to go?

PC

Happily, the directory you're in now is more obvious with that handy yellow text.

So remember:

- The path in Windows is: C:\cygwin64\home\vaddoni2\MARAC_API_Workshop (but this varies by user)
- And the *same path* in Cgywin looks like the new prompt, below:



```
vaddoni2@MSEL-SPC32 ~/.MARAC_API_Workshop $ |
```

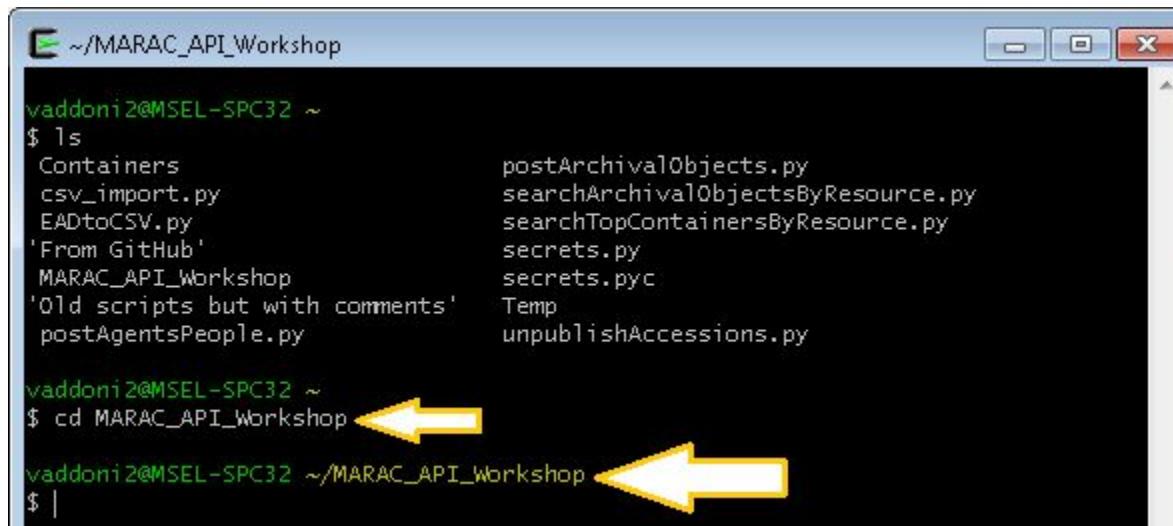
Command line bootcamp: Navigation

Where are you, and where do you want to go?

PC

Now I'm going to move *from* the MARAC API Workshop clone folder back to the Cgywin home directory:

Why? To teach you that ..To move into that folder I'll type `cd` (change directory), leave a space, and then the name of the directory I want to go into: `cd MARAC_API_Workshop`



The screenshot shows a terminal window with the following content:

```
vaddoni2@MSEL-SPC32 ~
$ ls
Containers          postArchivalObjects.py
csv_import.py       searchArchivalObjectsByResource.py
EADtoCSV.py        searchTopContainersByResource.py
'From GitHub'
MARAC_API_Workshop secrets.py
'Old scripts but with comments' secrets.pyc
postAgentsPeople.py Temp
unpublishAccessions.py

vaddoni2@MSEL-SPC32 ~
$ cd MARAC_API_Workshop
vaddoni2@MSEL-SPC32 ~/MARAC_API_Workshop
$ |
```

Two yellow arrows point from the bottom right towards the command `cd MARAC_API_Workshop` in the history, indicating the target of the navigation instruction.