Python & Web APIs

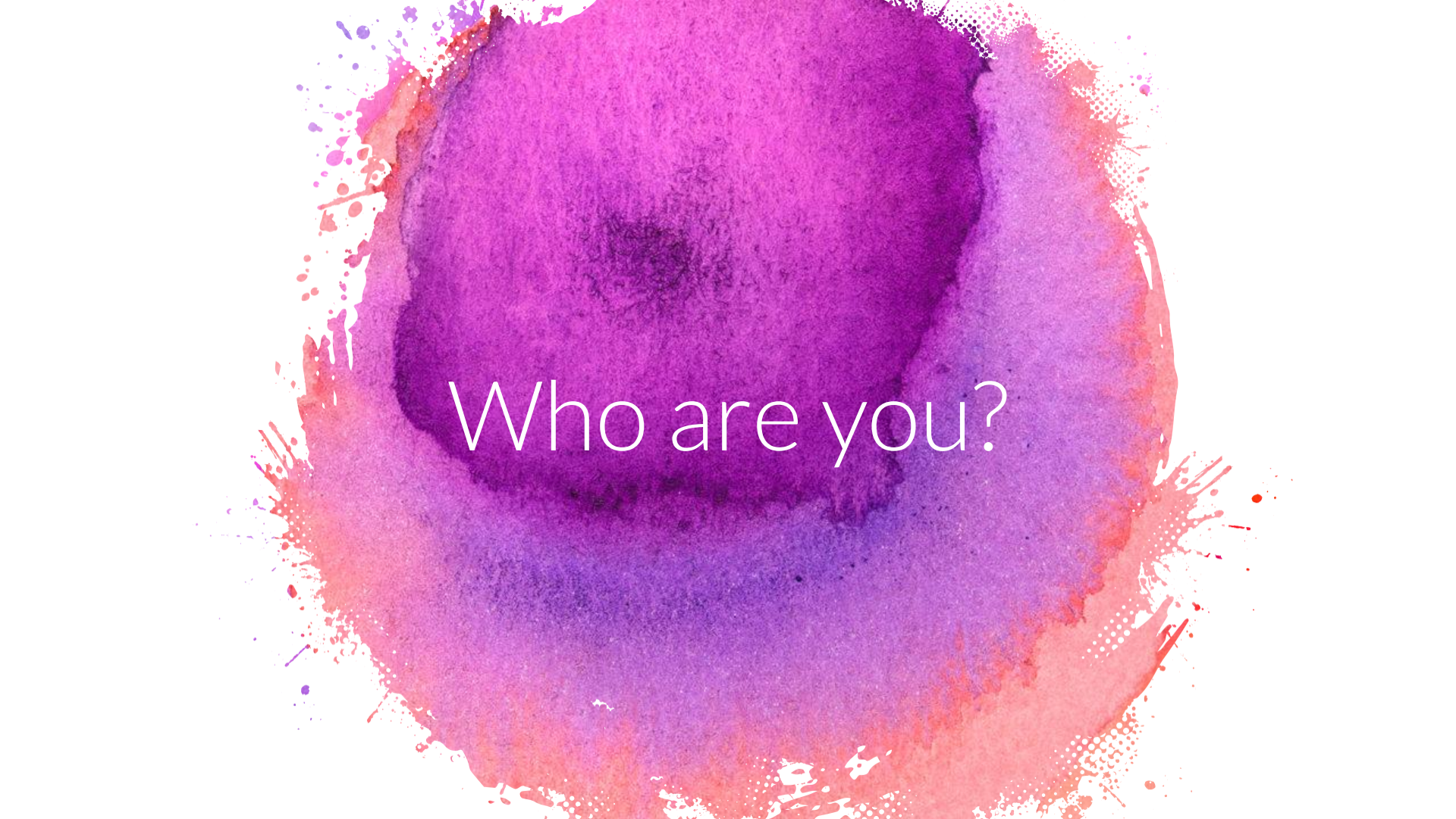
Cole Crawford
Humanities Research Computing Specialist

Jeremy Guillette
Instructional Technologist (FAS Academic Technology)

What are we doing?

- Introduction
- Finding APIs and interpreting documentation
- API demo in Python Notebook
  - Querying API
  - Using queries to enrich data
  - Using data to create content via API
- Resources for building APIs
- Questions


Who are we?


Who are you?

What's an API?

From Wikipedia:

In computer programming, an application programming interface (API) is a set of subroutine definitions, protocols, and tools for building application software. In general terms, it's a **set of clearly defined methods of communication** between various software components. A good API makes it easier to develop a computer program by providing all the building blocks, which are then put together by the programmer.

Stories

GET /articlesearch.json

Article Search

Article Search requests use the following URI structure:

Hide details ↑ | Try it out →

Parameters

q string

Location: query `?q=xyz`

Search query term. Search is performed on the article body, headline and byline.

fq string

Location: query `?fq=xyz`

"Filtered search query using standard Lucene syntax.

The filter query can be specified with or without a limiting field: label.

See Filtering Your Search for more information about filtering."

begin_date string

Location: query `?begin_date=xyz`

"Format: YYYYMMDD

Restricts responses to results with publication dates of the date specified or later."

end_date string

Location: query `?end_date=xyz`

"Format: YYYYMMDD

Restricts responses to results with publication dates of the date specified or earlier."

sort string

Location: query `?sort=newest`

"By default, search results are sorted by their relevance to the query term (q). Use the sort parameter to sort by pub_date."

Allowed values are:

- newest

Responses

200

The docs requested by the article search.

Schema Example

```
▼ {
  response: ▼ {
              docs: ▼ [
                      ▼ {
                        web_url:          string
                        snippet:          string
                        lead_paragraph:   string
                        abstract:         string
                        print_page:       string
                        blog:             ▼ [
                                            ▼ {
                                            }
                                          ]
                        source:           string
                        headline:         ▼ {
                                            main:   string
                                            kicker: string
                                          }
                        keywords:         ▼ {
                                            rank:  string
                                            name:  string
                                            value: string
                                          }
                        pub_date:         string
                        document_type:    string
                        news_desK:        string
```

Request

- Query parameters

Response

- Formatted output
  - JSON
  - XML

What's an API? (contd.)

For an online data source, an API gives you a consistent means of **requesting** data, and a consistent **format** for the data that you receive.

Why Python?

- Simple syntax
- Powerful tools
  - Modules
  - Notebooks
- "Pythonic"

Finding APIs

- APIs are the backbone of the interactive web
- Places to start:
  - https://www.programmableweb.com/
  - https://rapidapi.com/
  - https://apis.guru/browse-apis/
  - https://github.com/toddmotto/public-apis

Reading API Docs

- Writing good documentation is time consuming
  - Learn to read spotty or bad documentation
  - Email devs if you need to
  - Reading API Documentation
- Watch out for versioning and deprecation
- Look for return format

JSON:

```
{
  "status":"This is a Twitter post",
  "lat": 37.222,
  "long": -82.1234
}
```

XML:

```
<?xml version="1.0" encoding="utf-8" ?>
<Update>
  <Status>"This is a Twitter post"</Status>
  <Latitude>37.222</Latitude>
  <Longitude>-82.1234</Longitude>
</Update>
```

Harvard Art Museums

- Documentation: https://github.com/harvardartmuseums/api-doc
- Register: https://www.harvardartmuseums.org/collections/api
- API App Examples
  - http://apps.harvardartmuseums.org/art-explorer/
  - http://apps.harvardartmuseums.org/museum-explorer/
  - http://apps.harvardartmuseums.org/suns-explorer/

GeoNames

- All Documentation: http://www.geonames.org/export/web-services.html
- Search Documentation: http://www.geonames.org/export/geonames-search.html
- API Key Signup: http://www.geonames.org/login

Workshop Resources

- GitHub: https://github.com/artshumrc/python-web-apis/tree/master
  - Download, or
  - `` `git clone https://github.com/artshumrc/python-web-apis.git ` ``
  - Use the right branch!