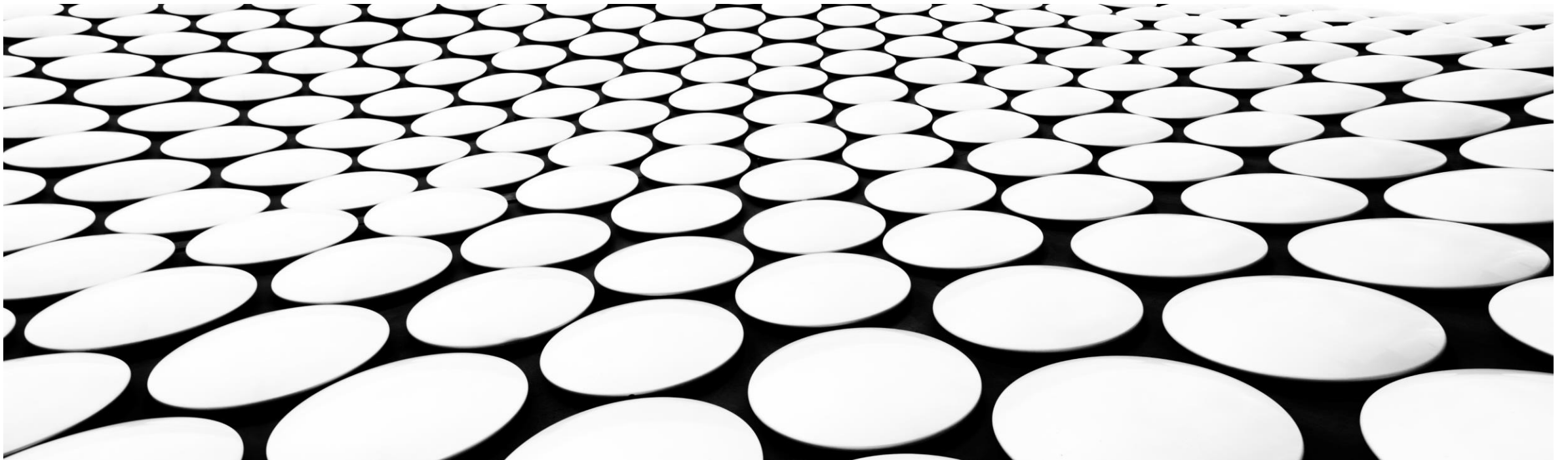# WHY USE APIS FOR SOCIAL SCIENCE RESEARCH?

ASPEN RUSSELL, CCSS DATA SCIENCE FELLOW, CORNELL UNIVERSITY

JONATHAN CHANG, CCSS SENIOR DATA SCIENCE FELLOW, CORNELL UNIVERSITY

CORNELL CENTER FOR SOCIAL SCIENCES (CCSS)

The Cornell Center for Social Sciences provides a welcoming environment for everyone embracing all backgrounds or identities. All instructors and attendees agree to abide by our community norms. We encourage the following behaviors in our workshops:
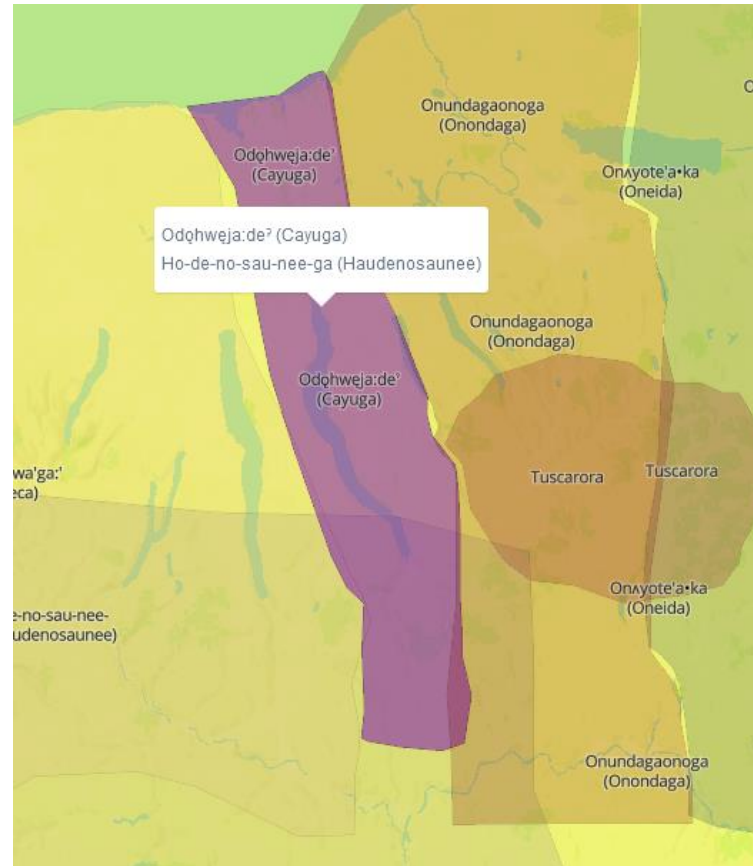
- Respect differing viewpoints and ideas

- Share your own perspectives and ask any questions

- Accept constructive criticism

- Use welcoming and inclusive language

- Show courtesy and respect for all instructors and attendees

If you believe that an instructor or attendee has violated the code of conduct, please report the violation to ccss-researchsupport@cornell.edu. We take all reported incidents seriously.

LAND ACKNOWLEDGEMENT

Cornell University is located on the traditional homelands of the Gayogo̱hó꞉nǫ' (the Cayuga NatiCCSS-RS@cornell.eduon). The Gayogo̱hó꞉nǫ' are members of the Haudenosaunee Confederacy, an alliance of six sovereign Nations with a historic and contemporary presence on this land. The Confederacy precedes the establishment of Cornell University, New York state, and the United States of America. We acknowledge the painful history of Gayogo̱hó꞉nǫ' dispossession and honor the ongoing connection of Gayogo̱hó꞉nǫ' people, past and present, to these lands and waters.

- American Indian & Indigenous Studies Program (AIISP), Cornell University

Here for more on the history of Cornell's violent, colonial formation, the movement to return native lands, and about the AIISP program at Cornell.
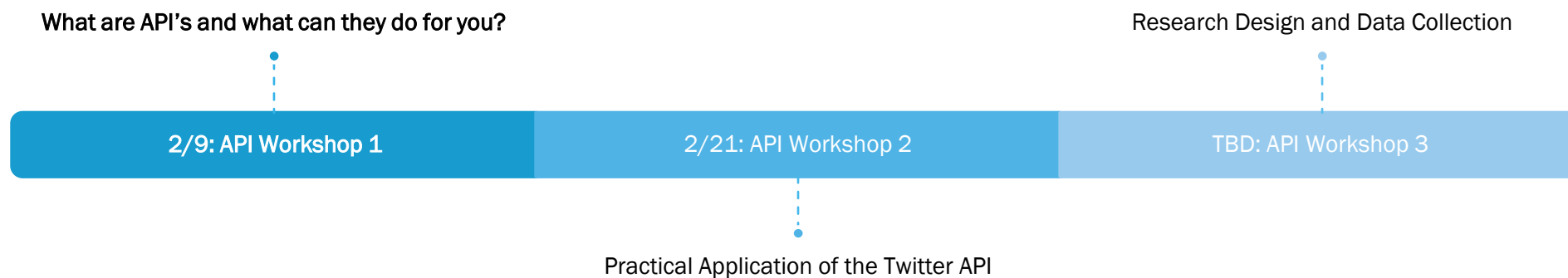
ABOUT THE PRESENTERS:

Aspen Russell

- PhD Student in Information Science
- Norms in Online Communities
- Hate speech, harassment, and misinformation
- Has worked with: Reddit, Twitter, Facebook, arXiv, YouTube, and more.
- Moonlights as a Dungeon Master

Jonathan Chang

- PhD Candidate in Computer Science
- Online content moderation
- Making online interactions healthier
- Has worked with: Reddit, Facebook, Wikipedia
- Slowly grinding the new level cap in Pokémon Go

THE API WORKSHOP SERIES

What are API's and what can they do for you?

Research Design and Data Collection

| 2/9: API Workshop 1 | 2/21: API Workshop 2 | TBD: API Workshop 3 |

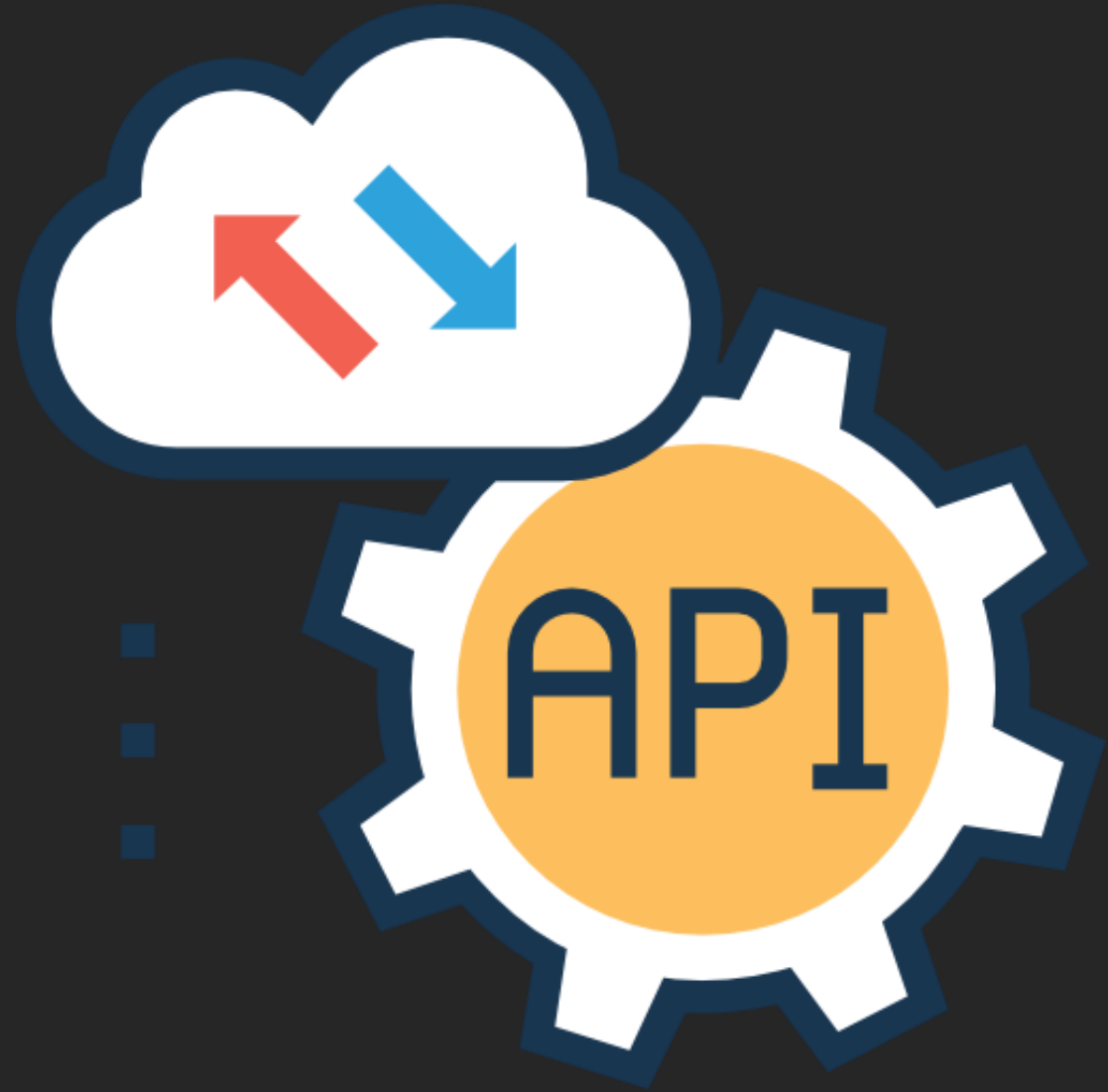Practical Application of the Twitter API

WHY APIS?

What if you had behavioral evidence of a hypothesis, theory, or phenomena you cared about?

What if you had 1000x as many examples of that behavior?

What if you had access to context surrounding that behavior?

EXAMPLES OF API DATA USE

Social Media
- User timeline of posts (change over time?)
- Hashtag posts (perhaps an activist movement)
- Tweets from a period of time (election)
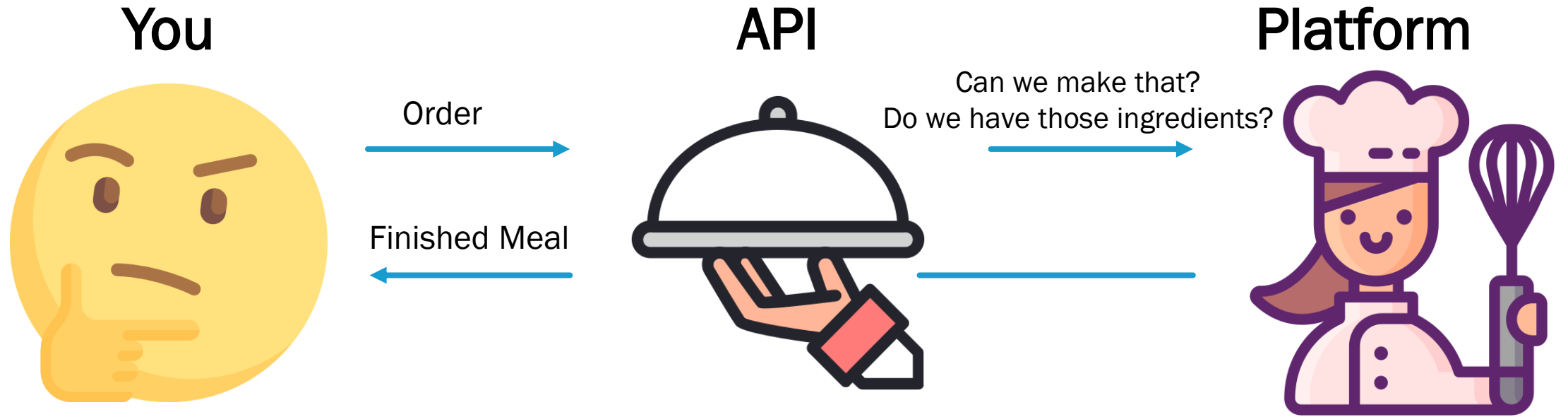- Investigating comments before and after a moderation change

Traditional Text (books, academic articles, etc)
- Co-authorship networks
- Publication reach based on identity
- Influence networks (who started what)
- Change over time (how do words change in meaning)

WHAT BROUGHT *YOU* HERE TODAY?
(BREAKOUT ROOMS)

WHAT *IS* AN API?

You

API

Platform

Order →

Can we make that?
Do we have those ingredients? →

← Finished Meal

**LIVE API DEMO:
REDDIT API VIA PUSHSHIFT**

WHAT'S OUT THERE? PROGRAMMING PACKAGES/TOOLS

	Reddit	Facebook	Twitter	YouTube
R	RedditExtractoR	FORT API	academictwitteR	tubeR
Python	PRAW: The Python Reddit API Wrapper	FORT API	search-tweets	python-youtube
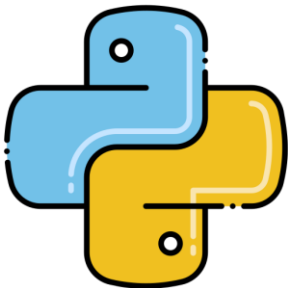
REDDIT EXTRACTOR

The Query ->

```
library(tidyverse)
library(RedditExtractoR)

# To extract URLs and thread titles from a subreddit --------------------

conspiracy <- reddit_urls(search_terms = "karen",
                          subreddit = "conspiracy",
                          cn_threshold = 5, # minimum of 5 comments
                          page_threshold = 6, # increasing the page threshold will return more
point)
                          sort_by = "comments") # sorting by comments here

conspiracy2 <- reddit_urls(search_terms = "karen",
                           subreddit = "conspiracy",
                           cn_threshold = 5,
                           page_threshold = 6,
                           sort_by = "relevance") # searching again, this time by relevance
```
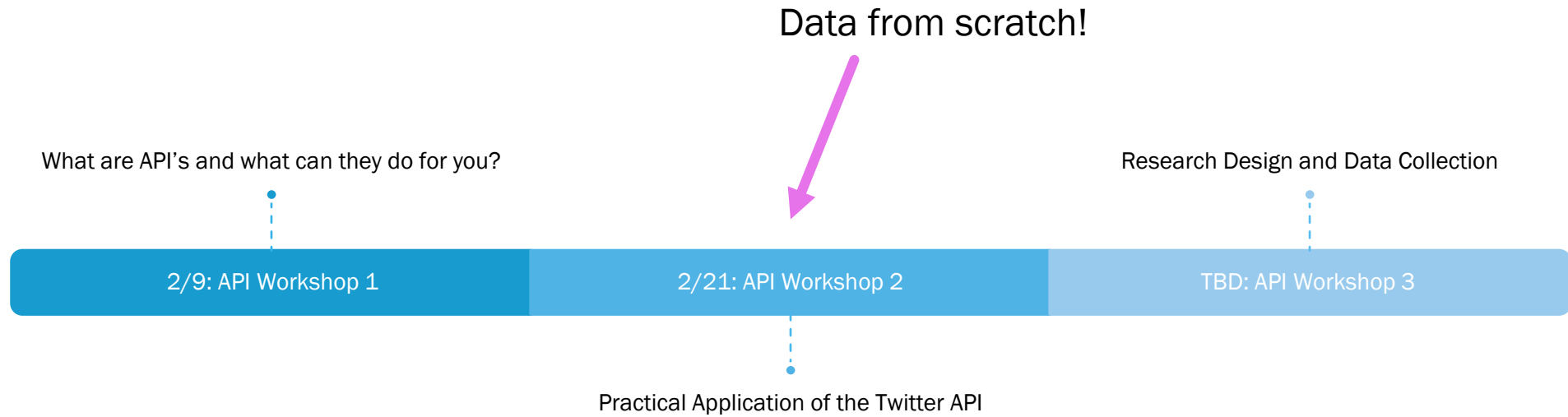
<- The Output

	date	num_comments	title	subreddit	URL
1	17-06-15	800	10 Reasons Why 9/11 Was Demonstrably An Inside Job	conspiracy	http://www.reddit.com/r/conspiracy/comments/3a67mc/10_...
2	02-10-21	661	No one is anti vax lol. It□s a trigger word. The media is clear...	conspiracy	http://www.reddit.com/r/conspiracy/comments/pzpvg6/no_...
3	27-08-20	386	yes karen. i want to enjoy the world	conspiracy	http://www.reddit.com/r/conspiracy/comments/ihbp0v/yes_...
4	13-08-21	398	Obey or you are a Karen. The hive mind operation. Exploit t...	conspiracy	http://www.reddit.com/r/conspiracy/comments/p3qtwn/ob...
5	19-04-21	306	None of this makes sense and it□s on purpose	conspiracy	http://www.reddit.com/r/conspiracy/comments/mtzax2/non...
6	06-05-20	299	Our generation got royally fucked by the political correctnes...	conspiracy	http://www.reddit.com/r/conspiracy/comments/gedhqb/our...
7	23-04-20	292	You know you are living in the "Upside Down" when...	conspiracy	http://www.reddit.com/r/conspiracy/comments/g6xcmv/yo...
8	14-04-20	207	Stay home <8d□<8	conspiracy	http://www.reddit.com/r/conspiracy/comments/g1ctag/stay...
9	26-08-21	190	We call on the Reddit Moderators to stop the rampant virtu...	conspiracy	http://www.reddit.com/r/conspiracy/comments/pboqu1/we...
10	03-12-21	169	NZ doctor tricks woman into getting 2nd vaccine shot (desp...	conspiracy	http://www.reddit.com/r/conspiracy/comments/r88wuo/nz_...

WHAT'S NEXT FOR THE API SERIES?

Data from scratch!

What are API's and what can they do for you?

Research Design and Data Collection

| 2/9: API Workshop 1 | 2/21: API Workshop 2 | TBD: API Workshop 3 |

Practical Application of the Twitter API

Q&A
(THANK YOU!)

Aspen Russell
Information Science PhD Student,
Cornell University
aspen@infosci.cornell.edu
@InfoTisayentist on Twitter

Jonathan Chang
Computer Science PhD Candidate,
Cornell University
jpchang@cs.cornell.edu
u/GradientDissent on Reddit

ATTRIBUTIONS

- Icons sourced from flaticon.com and from these artists: freepik, flaticon, becris, Yellowline, vidyavidz, Pixel perfect, and Prosymbols