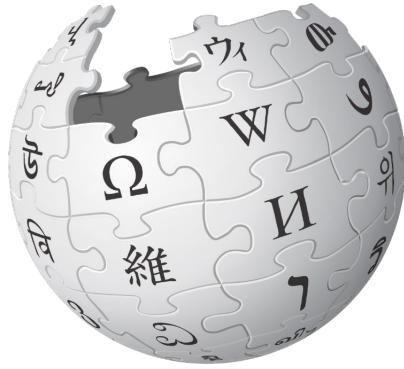




Helge Marahrens & Anne Kavalerchik
Indiana University Bloomington
Department of Sociology
Email correspondence to: hmarahre@iu.edu

Workshop in Methods (WIM)

APIs & Webscraping for Social Scientists



The New York Times



Science

CLIMATE | SPACE & COSMOS | HEALTH | TRILOBITES | SCIENTETAKE | OUT THERE



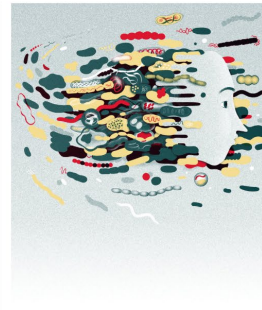
ARNO BURGI/PICTURE ALLIANCE, VIA GETTY IMAGES

TRILOBITES

Seeking Superpowers in the Axolotl Genome

The smiling salamanders can regrow most of their body parts, so researchers are building improved maps of their DNA.

1d ago • By STEPH YIN



SEAN MCSORLEY

MATTER

Germs in Your Gut Are Talking to Your Brain. Scientists Want to Know What They're Saying.

The body's microbial community may influence the brain and behavior, perhaps even playing a role in dementia, autism and other disorders.

1d ago • By CARL ZIMMER

A Closer Look at the Polar Vortex's Dangerously Cold Winds

Chicago will be as cold as the Arctic on Wednesday. We'll show you why.



11h ago • By YULIYA PARSHINA-KOTTAS, KARTHIK PATANJALI, JEREMY WHITE, BENJAMIN WILHELM and EMAN GROTHJAN

This Is Your Brain Off Facebook

Planning on quitting the social platform? A major new study offers a glimpse of what unplugging might do for your life. (Spoiler: It's not so bad.)



1h ago • By BENEDICT CAREY

Option 1: API

- Organized
- Easier Access
- Limited
 - Not all information sources have APIs
 - Information restricted by whoever maintains API

Option 2: Webscraping

- Flexible
- Difficult
- Understand html
 - e.g. where NYT article saves title information
- Ethical concerns



API (Application Programming Interface)

- A set of protocols and routines for building and interacting with software applications.
 - tool that allows computers to exchange information
 - tool that allows you easy access to new datasets

Examples:

<https://www.propublica.org/datastore/api/propublica-congress-api>

<https://developer.nytimes.com/docs/archive-product/1/overview>

<https://www.census.gov/data/developers/data-sets/acs-1year.html>





INDIANA UNIVERSITY BLOOMINGTON