



Exercise 5: Research on Digital Media & APIs

Exercise 5 for the lecture 'Foundations of Data Science'

Prof. Dr. Karsten Donnay, Assistant: Philipp Kling



This session covers

- Accessing APIs



**University of
Zurich** ^{UZH}

Department of Political Science

Accessing APIs



Application programming interfaces (APIs)

- Websites sometimes provide access to (parts of) their content via an Application programming interface (API)
- The motivation for websites is often of commercial nature: providing access to their content makes developing applications that request information from that website easier, thus, potentially increasing its publicity
- **Advantages:**
 - Structured data formats are returned (often JSON)
 - Legal access to data
- **Disadvantages:**
 - Content retrieval dependent on rate limits.
 - Content is dependent on the provider (we only get what is being provided)
 - Commercial motivation of provider sometimes not what's optimal for researchers



Application programming interfaces (APIs)

Step	Task	Sources
1	Get an overview of the API. What does it provide? Which ratelimits are queries subject to?	Read the documentation of the API
2	Is there already an R package that enables access to the API? If yes, follow its instructions. If not:	
3	If needed: create an account or register an application for the API	
4	Determine authentication procedure: does the API use OAuth or can we send requests with URLs.	ROauth, httr
5	Send query, retrieve information, transform it into R content	xml2, jsonlite