

Digital Methods: Digital Ethnography and Computational Data Collection via APIs, Scraping, and Digital Archives.

Nicola Righetti

Summary

Digital Methods

- The term “digital methods” generally refers to a set of research techniques and approaches that aim to study and understand the digital realm as a new field of social inquiry.
- The goal of digital methods is to provide insights into the social and cultural phenomena that emerge in the digital age. These methods involve the use of digital data and digital tools for research purposes, often focusing on how people engage with and produce digital content and how this content shapes contemporary society.

Digital Methods between Computational and Qualitative Methods

- Digital methods are often computational methods, and can encompass a wide range of activities, such as web scraping, social media analysis, network analysis, and various forms of data mining and visualization.
- However, they can also be qualitative methods like digital ethnography (also called “netnography”).

Primary Methods for Acquiring Digital Data

API (1/3)

- An API, or Application Programming Interface, is a software intermediary that enables two applications to communicate with each other.
- APIs are available for most major social media platforms, offering a standardized way for applications to interact with these platforms.
- However, the data that can be collected through these APIs is subject to limitations and restrictions, which can vary depending on the specific platform and its policies.

API (2/3)

An API (Application Programming Interface) can be likened to a specialized webpage that you access using a unique key. This key serves as a form of identification and permission, allowing you to request specific data from the API.

Once you have access, you can make queries or requests to the API, asking for particular pieces of information. For instance, in the context of social media platforms, you might request data such as user profiles, posts, or metrics like engagement rates.

API (3/3)

The API processes these requests and returns the requested data, provided it falls within the scope of what is permissible and available through that particular API.

This process enables a streamlined and controlled method of accessing and retrieving data from complex systems like social media platforms.

Example of an API query (using [GDELT](#) open API):

[https://api.gdeltproject.org/api/v2/doc/doc?
query=climate%20change&mode=TimelineVol&STARTDATETIME=201701010](https://api.gdeltproject.org/api/v2/doc/doc?query=climate%20change&mode=TimelineVol&STARTDATETIME=201701010)

Web Scraping

- Web scraping is a technique used to automatically extract data from websites.
- It's particularly useful in situations where an API (Application Programming Interface) is not available, allowing for the collection of web data directly from the website's frontend.
- However, it's important to note that web scraping should be done in compliance with the terms of service and copyright laws of the website in question. This ensures legal and ethical data collection practices, respecting the rights of the website owners and content creators.

Web Scraping Example

For example, we might want to extract the title and text of this news article <https://edition.cnn.com/2024/01/13/politics/john-kerry-climate-envoy-biden/index.html>

```
news_title
1 \n      John Kerry to step down as US climate envoy\n

article_content
1 \n      \n      \n      CNN\n      — \n      \n\n      \n      US climate envoy John Kerry plans to step down from his post by this spring, a source close to Kerry confirmed to CNN.\n      \n      \n      Kerry is leaving the position after three years at the helm of US climate diplomacy under the Biden administration. The source close to Kerry said he will still attend the World Economic Forum in January and is expected to attend the Munich Security Conference in February.\n      \n      \n      The 80-year-old former secretary of state and Democratic nominee for president led US negotiations at three international climate summits, the most recent of which was held in Dubai.\n      \n      \n      Kerry specifically prioritized restarting international climate talks between the US and China, and was key in negotiating the November Sunnylands agreement, a wide-ranging climate agreement between the two countries before COP28.\n      \n      \n      Axios first reported Kerry was stepping down and intends to join the Biden campaign. CNN has reached out to the Biden campaign and the White House for comment.\n      \n      \n      Before joining the Biden administration, Kerry had long worked on climate issues. As secretary of state
```

Digital Ethnography (1/2)

- Digital ethnography, also termed as 'netnography', is a qualitative research methodology that adapts the traditional practices of ethnography to the study of the cultures and communities emerging through computer-mediated communications.
- It is distinct from conventional ethnography primarily in its focus; while traditional ethnography explores cultural dynamics through direct, often immersive interaction within physical spaces, netnography is dedicated exclusively to understanding the social dynamics and cultural norms of online communities.

Digital Ethnography (2/2)

- A pivotal ethical consideration in netnography is the necessity of obtaining informed consent from participants. This is crucial due to the public yet personal nature of online spaces.
- Researchers must navigate the complex terrain of public forums, social media, and other digital platforms, where the line between public and private information can be ambiguous. Hence, ensuring that participants are fully aware of the research's nature, scope, and intended use of the data collected is imperative for ethical netnographic practice.

