Digital Trace Data in Social Science
Introduction

Dr. Taehee Kim
GESIS 2020 Workshop

University of Oldenburg

Introduction

Lecturer: Dr.Taehee Kim

Position and Contact

- Postdoc researcher ("Political System of Germany" team)
- email: taehee.kim(at)uni-oldenburg.de

Background:

- Ph.D. Political Science, Keio University, Japan
- M.Sc. Social and Economic Data Analysis, University of Konstanz

Projects using digital trace data:

- Hate speech spread in social media
- Conspiracy theory and extremists' movement in social media

Workshop Motivation

The digitalization and data revolution for the social sciences

Digital trace data are data documenting the activities, actions and interactions of users with digital devices or services (Howison et al., 2011)

- Nowadays, diverse kinds of digital trace data with a large volume have become available to researchers.
- It has a great potential for new approaches to social science questions.
- Vast textual corpora
- Alternatives to telephone or face-to-face surveys
- Users's interactions with various digital devices or services

The digitalization and data revolution for the social sciences

> Digital trace data are data documenting the activities, actions and interactions of users with digital devices or services (Howison et al., 2011)

- Nowadays, diverse kinds of digital trace data with a large volume have become available to researchers.
- It has a great potential for new approaches to social science questions.
- Vast textual corpora
- Alternatives to telephone or face-to-face surveys
- Users's interactions with various digital devices or services
  - ? How to obtain the data efficiently?
  - ? How to analyze the data in order to answer social science questions?

Available Sources

Twitter: APIs are available, open with some restrictions. Provides rich data, text, network, images, geo tag etc.

Facebook: Largely restricted unless you work with them.

Instagram: Great for research using images. You need to scrape.

Youtube: APIs are available, open.

Online communities, e.g., reddit, Wikipedia etc..

This workshop introduces Twitter API and provides multiple examples of collecting and analyzing Twitter data in R.

Time Schedule

7th December

10.00 - 11.30	Session
11.30 - 11.45	Break
11.45 - 13.00	Session
13.00 - 14.00	Break
14.00 - 15.30	Session
15.30 - 16.00	Break
16.00 - 18.00	Session

8th December

9.00 - 10.30	Session
10.30 - 10.45	Break
10.45 - 12.30	Session
12.30 - 13.30	Break
13.30 - 16.00	Session