Digital Trace Data - Group Exercise

Bamberg Summer Institute in Computational Social Science

Carsten Schwemmer, University of Bamberg

2019-07-30

Today we covered a broad range of topics including screen-scraping, working with APIs, and building apps and bots for social science research. Some of you have already had experience with some or even most of these issues– but for others, this may have been your first time collecting digital trace data.

This group exercise is designed to find a balance between practicing rudimentary skills (for those of you with little or no experience in this area) to cutting edge techniques (for those of you with extensive expertise in this area). As an added bonus, this exercise not only challenges you to practice your coding skills, but to think about how to ask questions that contribute new knowledge to social science theory as well.

Group exercise - tasks

1) Divide yourselves into groups of four by counting off in order around the room.
2) For 15 minutes, work together to identify a research question that you believe can be answered using some of the methods we discussed.
3) Identify a sampling frame to help you answer this question. For example, if your question is about politics, your sampling frame might be a list of elected officials on Twitter.
4) Use (legal) scraping techniques in order to collect the names of individual accounts, keywords, or topic areas to populate your sample.
5) Write code to collect data from each unit of analysis in your sample (e.g. tweets of elected officials).
6) Evaluate the strengths and weaknesses of the data you have collected.
7) Outline a hybrid research design (e.g. an app or a bot) that could be used to address the weaknesses of the data you collected, or otherwise improve your ability to answer the research question.

There is only one requirement: the group member with the least amount of experience coding should be responsible for typing the code into a computer.

You need not take the steps above in chronological order. However, after 2 hours you should be prepared to give a 5 minute presentation of your activities during the next two hours. Remember that these daily exercises are a way for you to explore possible topics for your group research projects (even if only by process of elimination), and to get to know each other better.

Questions?