Online course learning objectives
Utilizing big data is becoming increasingly important in social research, but it brings an array of ethical challenges and research design elements to consider. On this course, learners will gain an understanding of the emerging field of social data science and take your first steps into the big data-driven approach to research, learning from recent examples of social data science publications and projects.

By the end of this course, learners will be able to:
- Understand the relationship between empirical research, theory generation and testing.
- Define and formulate research problems, questions and hypotheses to be tested.
- Understand the rationale for using qualitative or quantitative research methods and the integrated or complementary nature between different methods in mixed methods research designs.
- Understand different forms of sampling, sampling error, and case selection, and their potential implications when interpreting findings.
- Apply concepts of generalisability, validity, reliability, and replicability.
- Understand ethical aspects of social data science and how to cope with them.

Language: English
Time to complete: 6 hours
Instructor: Dr. Taha Yasseri

Online course full syllabus

MODULE ONE: WHAT IS SOCIAL DATA SCIENCE?

Topic one: What is science?
- History of science and scientific methods.
- How science leads to knowledge generation.
- What are the elements of scientific methods?

Topic two: How can we communicate observations?
- Observation and measurement as the first steps in scientific activities.
- Finding measurable parameters, variables, quantifying.
- Phenomenology and its importance.

Topic three: How can we test our theories through experimentation?
- Randomized controlled trials.
- Quasi and natural experiments.
- Observational study.

Topic four: Challenges when studying humans
- Intrinsic differences between human systems and natural systems.
- Agency and adaptivity.
- Ethical aspect of studying humans.
MODULE TWO: HOW TO DETERMINE DESIGN ELEMENTS AND DATA SOURCES OF A RESEARCH PROJECT

Topic five: What are the boundaries of your research?
- How to answer questions that determine the research boundaries: what, when, where, who, and why.

Topic six: What is the scale and time resolution of your research?
- Units of analysis: individuals, groups, societies.
- Units of time: seconds, minutes, years, life-time, history.

Topic seven: What is big data and how does it differ to traditional data collection?
- What are found data and big data?
- Sampling techniques and self-selection bias.

Topic eight: What is consent in data collection?
- Implicit and explicit consent.
- Integrity and ethical consequences.

MODULE THREE: WHAT METHODS CAN WE USE IN SOCIAL DATA SCIENCE?

Topic nine: Quantitative methods from data science
- Data mining.
- Text mining.
- Social network analysis.

Topic ten: Qualitative methods from social science and mixed methods
- Interview and focus groups.
- Qualitative questionnaire.
- Qualitative coding and content analysis.
- How to mix qualitative and quantitative methods.

MODULE FOUR: WHAT CAN BE LEARNT FROM YOUR RESEARCH?

Topic eleven: From research question to hypothesis
- How to develop a good research question.
- The role of hypothesis generation and testing in research design.

Topic twelve: Validity
- Internal and external validity.
- Sampling bias.
- Correlation vs. causation.
- Generalisability.
Topic thirteen: Reproducibility and replicability
- Open science and data sharing.
- Documentation and transparency.

MODULE FIVE: WHAT IS ETHICAL SOCIAL DATA SCIENCE?

Topic fourteen: Ethical considerations in social data science
- Privacy.
- Recruitment.
- Consent.
- Data management.
- Data security.