

**Biagio Aragona.** Assistant Professor of Sociology.  
Department of Social Sciences, University of Naples Federico II.

Course: *Data and new data for social sciences and statistics*

### Goals

The course aims to present and discuss the different definitions and uses of the term data within the social sciences and statistics. It starts by distinguishing between the three terms data, information and knowledge. Furthermore, the definitions of the term “data” given in sociology, statistics and computer science are presented and critically discussed. The course continues by introducing some classifications of data according to several criteria: research stages (microdata, macrodata and metadata); operational definition (numerical data-vs-textual data); time (cross-sectional data-vs-longitudinal data); institutional settings (process produced data, open data), data management (structured, semi-structured, unstructured). The course ends by addressing the social and technological changes that have brought about the emergence of new data (big data, big corpora and linked data) and presenting their main features (velocity, variety, volume, unobtrusiveness, etc.) and some of their classifications: (automated data, directed data, voluntary data; internet of things, sensors, transactional data and social media data).

After this course Ph.D. students should be able to:

- recognize the difference between information and data;
- address their research interests to the different kinds of data;
- address their research interests to the different kinds of new data;
- understand the socio-technical changes that have signed the passage from data-scarce to data-intensive societies.

### Topics & Readings

*Class 1:* Data and information: definitions and discussion. Microdata, macrodata and metadata. Numerical and textual data. Cross-sectional data and longitudinal data. Process produced data and open data.

*Class 2:* Socio-technical changes and the emergence of new data. The main features of big data. Classifications of new data.

### Key readings

1. Kitchin R. (2014). *The Data Revolution: Big Data, Open Data, Data Infrastructures and Their Consequences*. London: Sage;

### Suggested additional readings

1. Amaturio E., Aragona B., Grassia M., Lauro C. Marino M., (eds) (2017). *Data Science and Social Research: Epistemology, Methods, Technology and Applications*, Heidelberg: Springer. (only Section 4);
2. Aragona, B. (2016). Big data o data that are getting bigger?. *Sociologia e ricerca sociale*, 42-53, Doi: 10.3280/SR2016-109005;