

Doing Big Data: methods and skills in the emerging field of data science

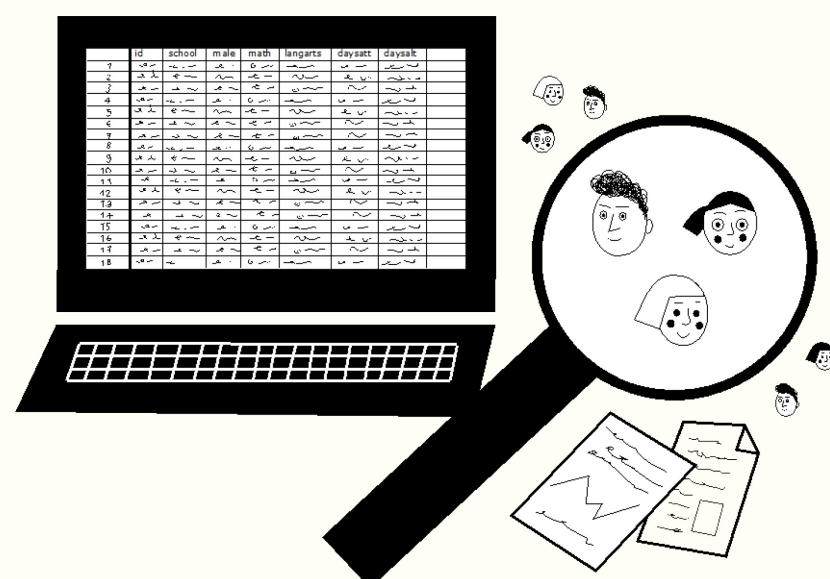
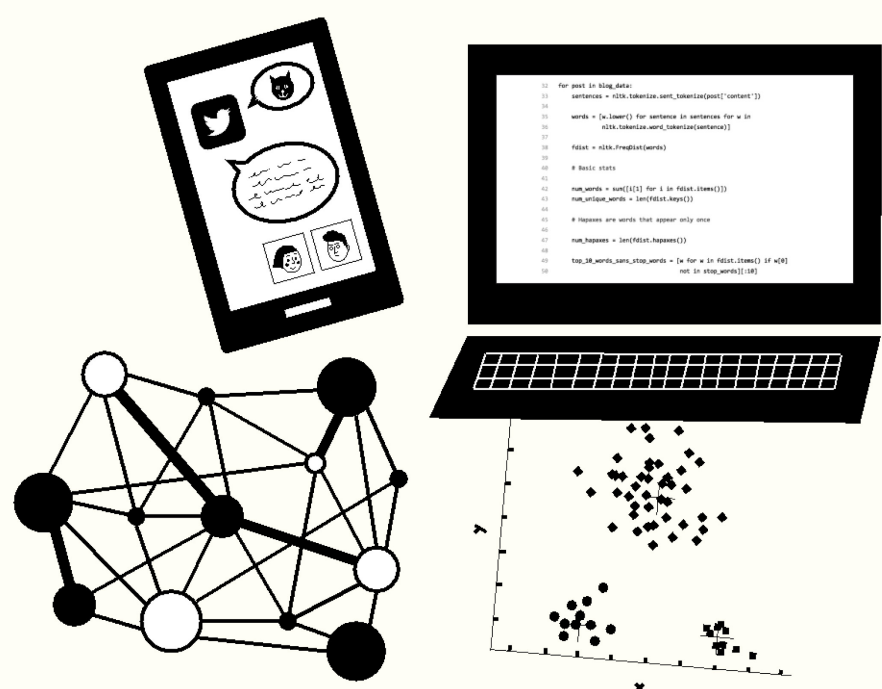
Philippe Saner
University of Lucerne, Dept. of Sociology

Challenges for the social sciences

- Implementation of sociological theories and methods in technical devices and algorithms (e.g. Google search, social media platforms)
 - New understanding/conception of the social and society through 'applied' disciplines (such as computer sciences, engineering etc.)?
 - Big Data driven analysis for optimization and predictive modeling gains in prominence in different social fields
 - Meanwhile diagnosis of a "crisis of empirical sociology" (Savage & Burrows 2007 and others)
- "Big Data is not about the Data" (King 2016), but also about the methods, tools and devices as well as the actors working with Big Data
- Social scientists need to understand the tools, methods and self-understandings of the professions that increasingly compete with social scientists to explain political, economic or cultural change

Context

Dissertation project within „Facing Big Data: methods and skills needed for a 21st century sociology” (PI: Prof. Sophie Mützel), funded by the Swiss National Science Foundation (National Research Programme 75 Big Data)



Data science as a research object

- An interdisciplinary field of people trained in statistics, mathematics, computer sciences, and engineering, using strong programming skills
- Data scientists work in various industries and academia and provide "data-driven" analyses; they predict and optimize performance and economic outputs
- For several years analysts problematize a considerable shortage of data scientists → rising salaries, but also high expectations ("jack of all trades")
- Hundreds of new educational programs have been established worldwide, both online (MOOCs) and offline (MA in Data Science), to meet the current high demand; political initiatives have been launched in many different countries to enhance data-intensive research infrastructure
- Data science as an organizational field is currently undergoing institutionalization (incl. professional training, research programs, scientific journals, conferences)



Objectives and research questions

1. Empirically examine the current state of methods, skills and tools used in training and in practice: How do they structure and shape the professional field as well as the practices of data scientists?
2. Investigate the ongoing institutionalization of data science as an organizational field: How do political and economic actors influence the establishment of educational and research programs and infrastructure related to data science?
3. Identify and describe boundary work (objects and crossings) between different disciplines: How are knowledge domains established or maintained? How do boundaries change between the different scientific disciplines involved?

Data and methods

Multimethod approach, combining 'traditional' social science methods with 'new' computational methods:

Data collection

- Online and document research, incl. text extraction (e.g. text books, syllabi, job ads)
- Ethnographic fieldwork
- Interviews with data science practitioners, training program managers, teaching staff, recruiters

Data analysis

- Textual analysis (e.g. text mining, topic modeling)
- Interpretive methods



Big Data
National Research Programme

Further information

Project „Facing Big Data: methods and skills needed for a 21st century sociology”
www.facingbigdata.ch

Contact details:
philippe.saner@unilu.ch
Twitter @ph_saner

