Ideational Foundations of Computational Social Science (CSS): Syllabus

Junprof. Dr. Andreas Jungherr

Winter 2018, University of Zürich

Institut für Politikwissenschaft
Affolternstrasse 56
8050 Zürich

Email: andreas.jungherr@gmail.com

Office hours: Wednesdays 16:00–17:00

1 Course Description

The course examines the methodological foundations of various popular approaches in computational social science (CSS). Over the last ten years, social scientists have found themselves confronting a massive increase in available data sources. In the debates on how to use these new data, the research potential of "computational social science" has featured prominently. While a lot of attention in the current debate has featured in the adaptation of computational methods new to the social sciences, questions about how to interpret the outputs of these methods meaningfully with regard to substantial questions in the social sciences has fallen short. The current state of the field could be described as method-heavy and methodology-thin. In this course, students will be introduced to a set of novel approaches prominently used in computational social science. We will focus on the roots of these methods, underlying assumptions, and consequences thereof with regard to connecting findings to substantial research interests. We will discuss among other approaches the use of digital trace data, social network analysis, predictive modeling.

Learning Objectives: Students are familiar with a selection of the currently dominant approaches in computational social science (CSS) and are able to interpret findings of these methods meaningfully in the context of substantial questions.

2 Requirements

2.1 Regular and active participation

The course features the discussion of the required readings. To benefit from this, students are expected to read the texts listed as required readings before each session and actively participate in the discussion for each session. Participation will not be graded.

Don't be scared off by the *Background Readings*. Those are meant to offer you a starting point if you decide to dig deeper into one of the topics discussed during the

course. They might offer a promising starting point for you, once you start thinking about your term paper.

Also not part of the course, but maybe helpful to get you started on implementing some of the approaches discussed here, are the readings in the *How To* sections. Here, you find helpful tutorials and introductions to the practical use of software and methods discussed in the sessions.

2.2 Presentation

Students will be asked to present a research paper during one of the topical session. During the first session, each student will be assigned a text from the listed readings for presentations. In preparing the presentations please use the following guiding questions:

1. What is the research question?

2. What are the concepts used in the study?

3. What are the mechanisms proposed in the study?

4. What are the hypotheses? How are they linked with concepts and mechanisms under study?

5. What empirical approach do the authors take? What is the data in use? How are the data analyzed? Does this seem appropriate?

6. What are the results and how are they connected with concepts and mechanisms under study?

7. How does the study related to the topics discussed in the required readings for the respective session?

8. How convincing do you find the arguments presented by the authors?

Beyond these guiding questions, please keep the following considerations in mind in preparing your presentation:

- Please plan your presentation to take between 10-15 minutes;

- Please prepare a slide deck with a presentation program of your choice (except for Prezi);

- Please prepare a handout of one to two pages for your fellow students, summarizing the main points of your presentation;

- Upload the handout on the day your presentation is due to the assignment folder on the course's OLAT repository. Use the following template for the filename `"your_last_name-handout..."`;

- The presentation will be graded and contribute 30% to your final grade;

- Please make sure to see me at least once during office hours to discuss the topic and scope of your presentation.

2.3 Term Paper

Following the course, students will be asked to hand in a term paper. The aim of this paper is for you to independently discuss an essay question focusing on the potential, adequate use, and interpretative challenges in the use of selected computational methods. Essay questions will be handed out during the second half of the course.

For the term paper, please adhere to the following guidelines:

- Style requirements:

    - Font: Times New Roman, 12pt;

    - Line-separation:1.5;

    - Page borders: 2.5 cm left and right, 2cm above and below;

    - Page set: Block;

    - The first line of each paragraph is indented.

- Citation Style: Please follow the citation convention of the American Political Science Review (APSR) available at
  `http://www.apsanet.org/APSR-Submission-Guidelines-August-2016;`

- Cover page: University, department, course title, paper title, name, Matriknr., semester count, study program, and e-mail-address;

- Length: ca. 3000 words +/-10%

- Deadline: Please return the paper on the date specified by the department (March 03, 2019) electronically at andreas.jungherr@gmail.com and by hardcopy at my University of Konstanz address (Box 85, Universitätsstrasse 10, 78464 Konstanz–D). The date is mandatory and can only be extended in case of officially certified illness;

- Use the following template for the filename `"your_last_name-paper..."`.

- The term paper will be graded and contribute 70% to your final grade;

- Please make sure to see me at least once during office hours to discuss the topic and scope of your term paper.

3 Course Outline

Class will meet at the following times and locations:

Wednesdays 14:15–15:45
(BIN-0-K.11)
Exception: On Wednesdays, October 24 and November 14 we will meet in Room AFL-F-121)

3.1   Week 1: Introduction (September 19)

Introduction, housekeeping, and assignment of presentations.

3.2   Week 2: No meeting (September 26)

3.3   Week 3: Computational Social Science (October 3)

Required Reading:
– Scott A. Golder and Michael W. Macy. 2014. "Digital Footprints: Opportunities and Challenges for Online Social Research". *Annual Review of Sociology* 40:129–152. doi:`10.1146/annurev-soc-071913-043145`.
– Rob Kitchin. 2014. "Big Data, new epistemologies and paradigm shifts". *Big Data & Society* 1 (1): 1–12. doi:`10.1177/2053951714528481`.

How To:
– Wes McKinney. 2017. *Python for Data Analysis: Data Wrangling with Pandas, NumPy, and IPython.* 2nd ed. Sebastopol, CA: O'Reilly Media.

Background Readings:
– Gary King. 2011. "Ensuring the data-rich future of the social sciences". *Science* 331 (6018): 719–721. doi:`10.1126/science.1197872`.
– David Lazer et al. 2009. "Computational Social Science". *Science* 323 (5915): 721–723. doi:`10.1126/science.1167742`.
– David Lazer and Jason Radford. 2017. "Data ex Machina: Introduction to Big Data". *Annual Review of Sociology* 43:19–39. doi:`10.1146/annurev-soc-060116-053457`.

3.4   Week 4: No meeting (October 10)

3.5   Week 5: Digital Traces (October 17)

Required Reading:
– Chapter 2: Observing Behavior (pp.13–77) in Matthew J. Salganik. 2017. *Bit By Bit: Social Research in the Digital Age.* Princeton, NJ: Princeton University Press.

Background Readings:
– Matthew J. Salganik. 2017. *Bit By Bit: Social Research in the Digital Age.* Princeton, NJ: Princeton University Press.

How To:
– Pascal Jürgens and Andreas Jungherr. 2016. *twitterresearch.* `https://github.com/trifle/twitterresearch`.
– Matthew A. Russell. 2018. *Mining the Social Web.* 3rd ed. Sebastopol, CA: O'Reilly Media.

– Zachary C. Steinert-Threkeld. 2018. *Twitter as Data*. Cambridge, UK: Cambridge University Press.

Presentations:
– James Howison, Andrea Wiggins, and Kevin Crowston. 2011. "Validity Issues in the Use of Social Network Analysis with Digital Trace Data". *Journal of the Association for Information Systems* 12 (12): 767–797.

– Andreas Jungherr. 2019. "Normalizing Digital Trace Data". In *Digital Discussions: How Big Data Informs Political Communication*, ed. by Natalie Jomini Stroud and Shannon C. McGregor, 9–35. New York, NY: Routledge.

3.6 Week 6: Measurement (October 24)

Required Reading:
– Robert Adcock and David Collier. 2001. "Measurement Validity: A Shared Standard for Qualitative and Quantitative Research". *American Political Science Review* 95 (3): 529–546.

Background Readings:
– David J. Hand. 2004. *Measurement Theory and Practice: The World Through Quantification*. London, UK: Wiley.

– Joel Michell. 1999. *Measurement in Psychology: A Critical History of a Methodological Concept*. Cambridge, UK: Cambridge University Press.

Presentations:
– Andreas Jungherr et al. 2017. "Digital Trace Data in the Study of Public Opinion: An Indicator of Attention Toward Politics Rather Than Political Support". *Social Science Computer Review* 35 (3): 336–356. doi:`10.1177/0894439316631043`.

– Nils B. Weidmann and Sebastian Schulte. 2016. "Using night light emissions for the prediction of local wealth". *Journal of Peace Research* 54 (2): 125–140. doi:`10.1177/0022343316630359`.

3.7 Week 7: Prediction (October 31)

Required Reading:
– Jake M. Hofman, Amit Sharma, and Duncan J. Watts. 2017. "Prediction and explanation in social systems". *Science* 355 (6324): 486–488. doi:`10.1126/science.aal3856`.

– Marc Keuschnigg, Niclas Lovsjö, and Peter Hedström. 2018. "Analytical sociology and computational social science". *Journal of Computational Social Science* 1 (1): 3–14. doi:`10.1007/s42001-017-0006-5`.

Background Readings:
– Sandra González-Bailón. 2017. *Decoding the Social World: Data Science and the Unintended Consequences of Communication*. Cambridge, MA: The MIT Press.

– Peter Hedström. 2005. *Dissecting the Social: On the Principles of Analytical Sociology.* Cambridge, UK: Cambridge University Press.

Presentations:
– Harald Schoen et al. 2013. "The power of prediction with social media". *Internet Research* 23 (5): 528–543. doi:`10.1108/IntR-06-2013-0115`.

3.8 Week 8: No meeting (November 7)

3.9 Week 9: Classification & Scoring (November 14)

Required Reading:
– Danielle Keats Citron and Frank Pasquale. 2014. "The Scored Society: Due Process fo Automated Predictions". *Washington Law Review* 89 (1): 1–33.

Background Readings:
– Ajay Agrawal, Joshua Gans, and Avi Goldfarb. 2018. *Prediction Machines: The Simple Economics of Artificial Intelligence.* Boston, MA: Harvard Business Review Press.

– David L. Poole and Alan K. Mackworth. 2017. *Artificial Intelligence: Foundations of Computational Agents.* 2nd ed. Cambridge, UK: Cambridge University Press.

– Stuart Russell and Peter Norvig. 2016. *Artificial Intelligence: A Modern Approach.* 3rd ed. Boston, MA: Pearson Education.

How To:
– Brett Lantz. 2015. *Machine Learning with R.* Sebastopol, CA: PACKT Publishing.

– Sebastian Raschka and Vahid Mirjalili. 2017. *Python Machine Learning: Machine Learning and Deep Learning with Python, scikit-learn, and TensorFlow.* 2nd ed. Birmingham, UK: PACKT Publishing.

Presentations:
– Chapter 3 (pp.34–61) in Andrew Guthrie Ferguson. 2017. *The Rise of Big Data Policing: Surveillance, Race, and the Future of Law Enforcement.* New York, NY: New York University Press.

3.10 Week 10: Text as Data (November 21)

Required Reading:
– Justin Grimmer and Brandon M. Stewart. 2013. "Text as Data: The Promise and Pitfalls of Automatic Content Analysis Methods for Political Texts". *Political Analysis* 21 (3): 267–297. doi:`10.1093/pan/mps028`.

Background Readings:
– Andrew Piper. 2018. *Enumerations: Data and Literary Study.* Chicago, IL: The University of Chicago Press.

– Peter D. Turney and Patrick Pantel. 2010. "From Frequency to Meaning: Vector Space Models of Semantics". *Journal of Artificial Intelligence Research* 37:141–188. doi:`10.1613/jair.2934`.

How To:
– Benjamin Bengfort, Rebecca Bilbro, and Tony Ojeda. 2018. *Applied Text Analysis with Python: Enabling Language Aware Data Products with Machine Learning.* Sebastopol, CA: O'Reilly Media.

– Julia Silge and David Robinson. 2017. *Text Mining with R: A Tidy Approach.* Sebastopol, CA: O'Reilly Media.

Presentations:
– Christopher Bail. 2014. "The cultural environment: measuring culture with big data". *Theory and Society* 43 (3-4): 465–482. doi:`10.1007/s11186-014-9216-5`.

– Lisa Blaydes, Justin Grimmer, and Alison McQueen. 2018. "Mirrors for Princes and Sultans: Advice on the Art of Governance in the Medieval Christian and Islamic Worlds". *The Journal of Politics.* doi:`10.1086/699246`.

3.11   Week 11: Networks (November 28)

Required Reading:
– Gerald C. Kane et al. 2014. "What's different about social media networks? A framework and research agenda". *Management Information Systems Quarterly* 38 (1): 275–304.

Background Readings:
– Damon Centola. 2018. *How Behavior Spreads: The Science of Complex Contagions.* Princeton, NJ: Princeton University Press.

– David Easley and Jon Kleinberg. 2010. *Networks, Crowds, and Markets: Reasoning About a Highly Connected World.* Cambridge, UK: Cambridge University Press.

– Matthew O. Jackson. 2008. *Social and Economic Networks.* Princeton, NJ: Princeton University Press.

How To:
– Eric D. Kolaczyk and Gábor Csárdi. 2014. *Statistical Analysis of Network Data with R.* Cham, CH: Springer.

– Katherine Ognyanova. 2018. *Network visualization with R.* `http://kateto.net/network-visualization`.

Presentations:
– Sharad Goel et al. 2016. "The Structural Virality of Online Diffusion". *Management Science* 62 (1): 180–196. doi:`10.1287/mnsc.2015.2158`.

3.12 Week 12: Experiments (December 5)

Required Reading:
– Chapter 4: Running experiments (pp.147–220) in Matthew J. Salganik. 2017. *Bit By Bit: Social Research in the Digital Age*. Princeton, NJ: Princeton University Press.

Background Readings:
– Paul R. Rosenbaum. 2017. *Observation and Experiments: An Introduction to Causal Inference*. Cambridge, MA: Harvard University Press.

Presentations:
– Christopher A. Bail et al. 2018. "Exposure to opposing views on social media can increase political polarization". *Proceedings of the National Academy of Sciences*. doi:10.1073/pnas.1804840115.
– Matthew J. Salganik and Duncan J. Watts. 2009. "Web-based experiments for the study of collective social dynamics in cultural markets". *Topics in Cognitive Science* 1 (3): 439–468. doi:10.1111/j.1756-8765.2009.01030.x.

3.13 Week 13: Simulation (December 12)

Required Reading:
– Michael W. Macy and Robert Willer. 2002. "From Factors to Actors: Computational Sociology and Agent-Based Modeling". *Annual Review of Sociology* 28:143–166. doi:10.1146/annurev.soc.28.110601.141117.

Background Readings:
– Joshua M. Epstein. 2006. *Generative Social Science: Studies in Agent-Based Computational Modeling*. Princeton, NJ: Princeton University Press.
– Dirk Helbing. 2012. "Modeling of Socio-Economic Systems". In *Social Self-Organization: Agent-Based Simulations and Experiments to Study Emergent Social Behavior*, ed. by Dirk Helbing, 1–24. Heidelberg, DE: Springer. doi:10.1007/978-3-642-24004-1_1.
– Dirk Helbing and Stefano Balietti. 2012. "Agend Based Modeling". In *Social Self-Organization: Agent-Based Simulations and Experiments to Study Emergent Social Behavior*, ed. by Dirk Helbing, 25–70. Heidelberg, DE: Springer. doi:10.1007/978-3-642-24004-1_2.
– John H. Miller and Scott E. Page. 2007. *Complex Adaptive Systems: An Introduction to Computational Models of Social Life*. Princeton, NJ: Princeton University Press.
– Melanie Mitchell. 2009. *Complexity: A Guided Tour*. Oxford, NY: Oxford University Press.

How To:
– Uri Willensky and William Rand. 2015. *An Introduction to Agent-Based Modeling: Modeling Natural, Social, and Engineered Complex Systems with NetLogo*. Boston, MA: The MIT Press.

Presentations:
– Annie Waldherr. 2014. "Emergence of News Waves: A Social Simulation Approach".
Journal of Communication 64 (5): 852–873. doi:`10.1111/jcom.12117`.

3.14   Week 14: Discussion of Open Questions and Term Paper (December 19)