

Bedeutung und Struktur der Nutzung digitaler Methoden und Verfahren in der empirischen Sozialwissenschaft

Andreas Jungherr

Juniorprofessor

Social Science Data Collection and Analysis

Center for Data and Methods

Fachbereich Politik- und Verwaltungswissenschaft

Universität Konstanz

Nutzung

- Computational Social Science (Lazer et al. 2009; Macy and Willer 2002) & Big Data (Golder and Macy 2014; Lazer and Radford 2017)
- Predictions (Lazer et al. 2014)
- Sensors (Pentland 2008; Weidmann and Schutte 2017)
- Text as Data (Grimmer and Stewart 2013)

Methoden

- Agent-Based Modeling (Waldherr 2014)
- Computer Assisted Text Analysis (Trilling 2015)
- Data Collection & Processing (Jürgens and Jungherr 2016; King 2011)
- Digital Trace Data (Howison, Wiggins, and Crowston 2011; Jungherr 2015; Jungherr et al. 2017)
- Experiments (Bail et al. 2018; Salganik, Dodds, and Watts 2006)
- Machine Learning (Theocharis et al. 2016)
- Network Analysis (Goel et al. 2016)
- Time Series Analysis (Jungherr 2014; Jungherr and Jürgens 2014; Wells et al. 2016)

Zur Zeit findet die Nutzung digitaler Methoden und Verfahren an den Rändern der empirischen Sozialwissenschaft statt. Gründe dafür sind:

- Methoden zur Datensammlung und -analyse bleiben oft esoterisch;
- Studien bleiben oft nur illustrativ für eine Datenquelle oder Methode;
- Niedriges Interesse der Sozialwissenschaft an Phänomenen und Prozessen der gesellschaftlichen Digitalisierung;
- Niedriges Bewusstsein der CSS für Konzeptualisierung und Messtheorie.

Strukturelle Fragen

Methoden- versus Grundlagenforschung

Es besteht eine Kluft zwischen dem selbsterklärten Anspruch der CSS und ihrem bisherigen Beitrag.

Die meisten vorliegenden Studien sind überwiegend technischer Natur und tragen wenig zu Debatten im Kern der empirischen Sozialwissenschaft bei.

Dies beginnt sich zu ändern, da der Neuigkeitswert von Daten und Methoden sinkt. Forschungsdesign (Salganik 2017) und bewusste Konzeptualisierung (Jungherr 2018; Jungherr and Theocharis 2017) beginnen wichtiger zu werden.

Aktuelle Rekrutierungsmuster spiegeln diese Entwicklung jedoch noch nicht. Hier werden überwiegend Kandidaten auf Basis ihrer Methodenkenntnis auf Positionen in der Sozialwissenschaft berufen nicht auf Basis von konzeptionellen oder inhaltlichen Beiträgen.

Erfahrungen aus dem interdisziplinären Master-Programm Social and Economic Data Science in Konstanz zeigen:

- Methodenkurse werden von Studierenden gut aufgenommen;
- Konzeptionelle und inhaltliche Kurse sind schwerer zu vermitteln;
- Berufliche Chancen der Absolventen sind außerakademisch sehr hoch, dies ist hauptsächlich in ihrer Methodenkenntnis begründet;
- Akademische Ausbildungspipeline bleibt problematisch: Chancen für attraktive Promotions- und Post-Doc Stellen sind sehr hoch; es bleibt aber fraglich ob sich dies in volle Professuren übersetzen lässt.

Strukturelle Verankerung

Es stellen sich die folgenden strukturelle Fragen:

- Wird die CSS Teil der normalen sozialwissenschaftlichen Methodenausbildung oder werden gezielt Professuren geschaffen?
- Bleibt die CSS Hilfswissenschaft der etablierten sozialwissenschaftlichen Schwerpunkte oder werden Professuren mit neuen inhaltlichen Schwerpunkten geschaffen, die die CSS konzeptionell begleiten können?
- Entwickelt sich Lehre und Forschung mit Methoden und Konzepten der CSS in die Breite oder bleibt sie auf einige wenige Institutionen beschränkt?
- Werden entsprechende Professuren mit Sozialwissenschaftlern besetzt oder mit Informatikern?

Die sozialwissenschaftliche Arbeit mit großen und heterogenen Datensätzen erfordert immer vielfältigere Kenntnisse und Ressourcen:

- Datensammlung;
- Datenaufarbeitung & -verwaltung;
- Große und stetig wachsende Methodenvielfalt.

Der in der Informatik erfolgreich etablierte Lab-Ansatz ist vielversprechend aber in der Sozialwissenschaft schwierig umsetzbar (King 2011).

Interdisziplinarität bietet vielversprechende Lösungen der oben genannten Herausforderungen. Aber:

- Projekte müssen so strukturiert werden, dass alle Beteiligten in ihren Fachbereichen interessante Probleme lösen können;
- Interdisziplinarität rhetorisch erwünscht, praktisch bleiben attraktive Publikationsformate weiterhin stark disziplinär gebunden;
- Dies gilt auch für akademische Berufsperspektiven.

Herausforderung, gleichberechtigte Partnerschaft zu entwickeln:

- Informatik darf nicht als Hilfswissenschaft, Datenlieferant oder Methodenentwickler behandelt werden;
- Sozialwissenschaftler dürfen sich nicht darauf beschränken Methodenanwender zu bleiben;
- Sozialwissenschaftler müssen sich jedoch auch auf ihre Stärken in Konzeptentwicklung und Interpretation von Analysen konzentrieren und nicht Energien verschwenden, um schlechte Informatiker zu werden.

Projektstruktur und Begutachtungsprozess

Daten- und methodenintensive Projekte in der Sozialwissenschaft teilen mehr mit Projekten der Informatik als der Sozialwissenschaft:

- Personalkosten;
- Bedeutung von Post-Docs in Koordinierung des Teams;
- Kosten von Infrastruktur und Datenbeschaffung.

Aber:

- Finanzierung über externe Geldgeber (z.B. Wirtschaft) deutlich schwerer;
- Gutachtungsprozess ist Lotterie ob Reviewer Methoden oder Inhalte sehen wollen;
- Unterschiedlicher Begutachtungsstil in der Sozialwissenschaft und Informatik.

Fragen?




Literatur


References



Bail, Christopher A., Lisa P. Argyle, Taylor W. Brown, John P. Bumpus, Haohan Chen, M.B. Fallin Hunzaker, Jaemin Lee, Marcus Mann, Friedolin Merhout, and Alexander Volfovsky. 2018. "Exposure to opposing views on social media can increase political polarization". *Proceedings of the National Academy of Sciences*. doi:10.1073/pnas.1804840115.

References ii

-  Goel, Sharad, Ashton Anderson, Jake Hofman, and Duncan J. Watts. 2016. “The Structural Virality of Online Diffusion”. *Management Science* 62 (1): 180–196. doi:10.1287/mnsc.2015.2158.
-  Golder, Scott A., 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.
-  Grimmer, Justin, 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.

-  Howison, James, 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.
-  Jungherr, Andreas. 2015. *Analyzing Political Communication with Digital Trace Data: The Role of Twitter Messages in Social Science Research*. Cham, CH: Springer.
doi:10.1007/978-3-319-20319-5.
-  – . 2018. "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.

References iv

-  – . 2014. “The Logic of Political Coverage on Twitter: Temporal Dynamics and Content”. *Journal of Communication* 64 (2): 239–259. doi:10.1111/jcom.12087.
-  Jungherr, Andreas, and Pascal Jürgens. 2014. “Through a glass, darkly: Tactical support and symbolic association in Twitter messages commenting on Stuttgart 21”. *Social Science Computer Review* 32 (1): 74–89. doi:10.1177/0894439313500022.
-  Jungherr, Andreas, Harald Schoen, Oliver Posegga, and Pascal Jürgens. 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.

References v

-  Jungherr, Andreas, and Yannis Theocharis. 2017. "The Empiricist's Challenge: Asking Meaningful Questions in Political Science in the Age of Big Data". *Journal of Information Technology & Politics* 14 (1): 97–109.
doi:10.1080/19331681.2017.1312187.
-  Jürgens, Pascal, and Andreas Jungherr. 2016. *A Tutorial for Using Twitter-Data in the Social Sciences: Data Collection, Preparation, and Analysis*. Social Science Research Network (SSRN). doi:10.2139/ssrn.2710146.
<http://ssrn.com/abstract=2710146>.
-  King, Gary. 2011. "Ensuring the data-rich future of the social sciences". *Science* 331 (6018): 719–721.
doi:10.1126/science.1197872.

References vi

-  Lazer, David, Ryan Kennedy, Gary King, and Alessandro Vespignani. 2014. “The Parable of Google Flu: Traps in Big Data Analysis”. *Science* 343 (6176): 1203–1205. doi:10.1126/science.1248506.
-  Lazer, David, Alex Pentland, Lada Adamic, Sinan Aral, Albert-László Barabási, Devon Brewer, Nicholas Christakis, Noshir Contractor, James Fowler, Myron Gutmann, Tony Jebara, Gary King, Michael W. Macy, Deb Roy, and Marshall Van Alstyne. 2009. “Computational Social Science”. *Science* 323 (5915): 721–723. doi:10.1126/science.1167742.
-  Lazer, David, 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.

-  Macy, Michael W., 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.
-  Pentland, Alex. 2008. *Honest Signals: How They Shape Our World*. Cambridge, MA: The MIT Press.
-  Salganik, Matthew J. 2017. *Bit By Bit: Social Research in the Digital Age*. Princeton, NJ: Princeton University Press.
-  Salganik, Matthew J., Peter S. Dodds, and Duncan J. Watts. 2006. "Experimental study of inequality and unpredictability in an artificial cultural market". *Science* 311 (5762): 854–856.
doi:10.1126/science.1121066.

References viii

-  Theocharis, Yannis, Pablo Barberá, Zoltán Fazekas, Sebastian Adrian Popa, and Olivier Parnet. 2016. “A Bad Workman Blames His Tweets: The Consequences of Citizens’ Uncivil Twitter Use When Interacting With Party Candidates”. *Journal of Communication* 66 (6): 1007–1031. doi:10.1111/jcom.12259.
-  Trilling, Damian. 2015. “Two Different Debates? Investigating the Relationship Between a Political Debate on TV and Simultaneous Comments on Twitter”. *Social Science Computer Review* 33 (3): 259–276. doi:10.1177/0894439314537886.
-  Waldherr, Annie. 2014. “Emergence of News Waves: A Social Simulation Approach”. *Journal of Communication* 64 (5): 852–873. doi:10.1111/jcom.12117.



Weidmann, Nils B., and Sebastian Schutte. 2017. "Using night light emissions for the prediction of local wealth". *Journal of Peace Research* 54 (2): 125–140.
doi:10.1177/0022343316630359.



Wells, Chris, Dhavan V. Shah Jon C. Pevehouse, JungHwan Yang, Ayellet Pelled, Frederick Boehm, Josephine Lukito, Shreenita Ghosh, and Jessica L. Schmidt. 2016. "How Trump Drove Coverage to the Nomination: Hybrid Media Campaigning". *Political Communicaton* 33 (4): 669–676.
doi:10.1080/10584609.2016.1224416.