

gesis

Leibniz-Institut
für Sozialwissenschaften



A short introduction to Computational Social Science and Digital Behavioral Data

Meet the Experts

*Best practice methods in Survey Methodology and
Computational Social Science*

Katrin Weller, 24.06.2021

Logistics

- This talk will be recorded. We are not recording the Q&A session after the talk.
- Participants are muted during the session.
- For questions, please only use the private chat function and send questions exclusively to the „meetexperts“ account.
- If you send a message to the general chat, this message (incl. your name) will be visible to all participants.
- Questions will be collected and answered after the talk.

Speaker



Dr. Katrin Weller

- Deputy head of CSS Department at GESIS
- Team lead Social Analytics and Services
- PhD in information science
- Social media research methods, altmetrics
- Contact: katrin.weller@gesis.org

gesis

Leibniz-Institut
für Sozialwissenschaften



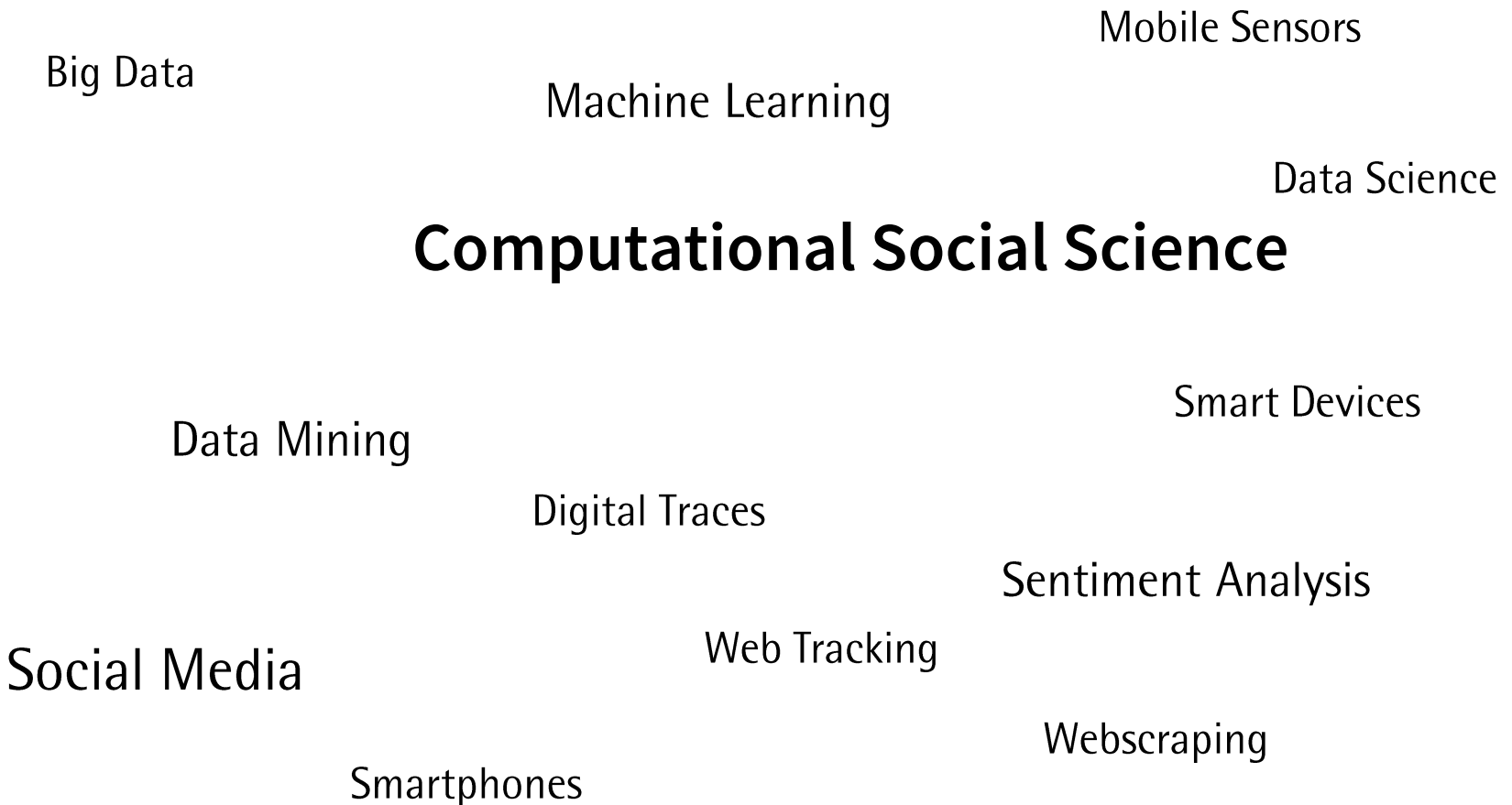
A short introduction to Computational Social Science and Digital Behavioral Data

Meet the Experts

*Best practice methods in Survey Methodology and
Computational Social Science*

Katrin Weller, 24.06.2021

Computational Social Science (CSS)



Computational Social Science (CSS)



Computational Social Science (CSS)

**What is Computational Social Science –
and what can it do for you?**

Let's take a first look:

How often do you talk
to your friends on the
phone?

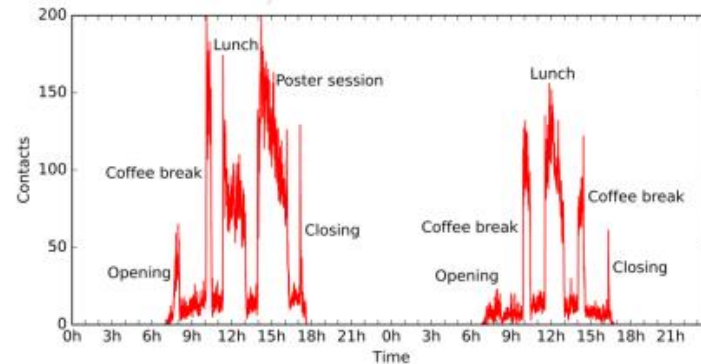
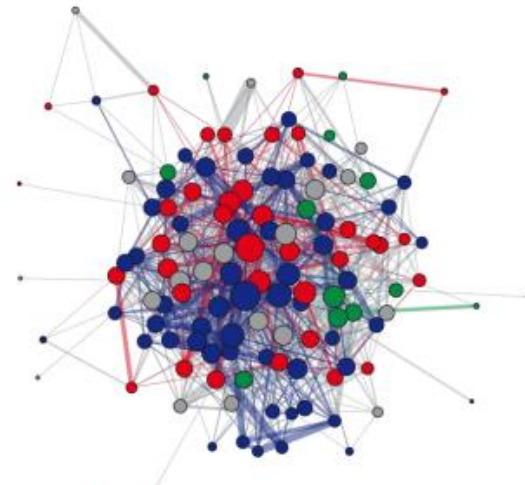
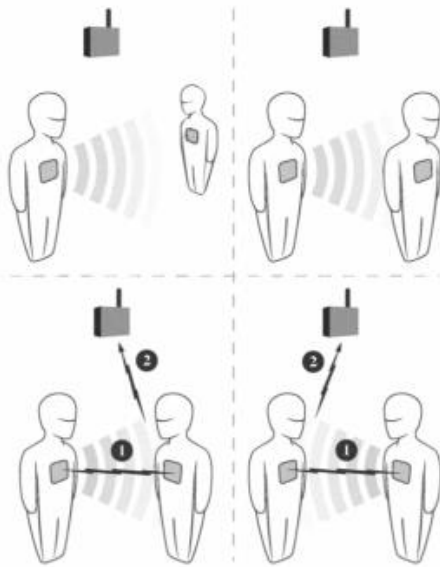
How many hours do
you spend outside
your home?

To how many people did
you talk during the
conference last week?

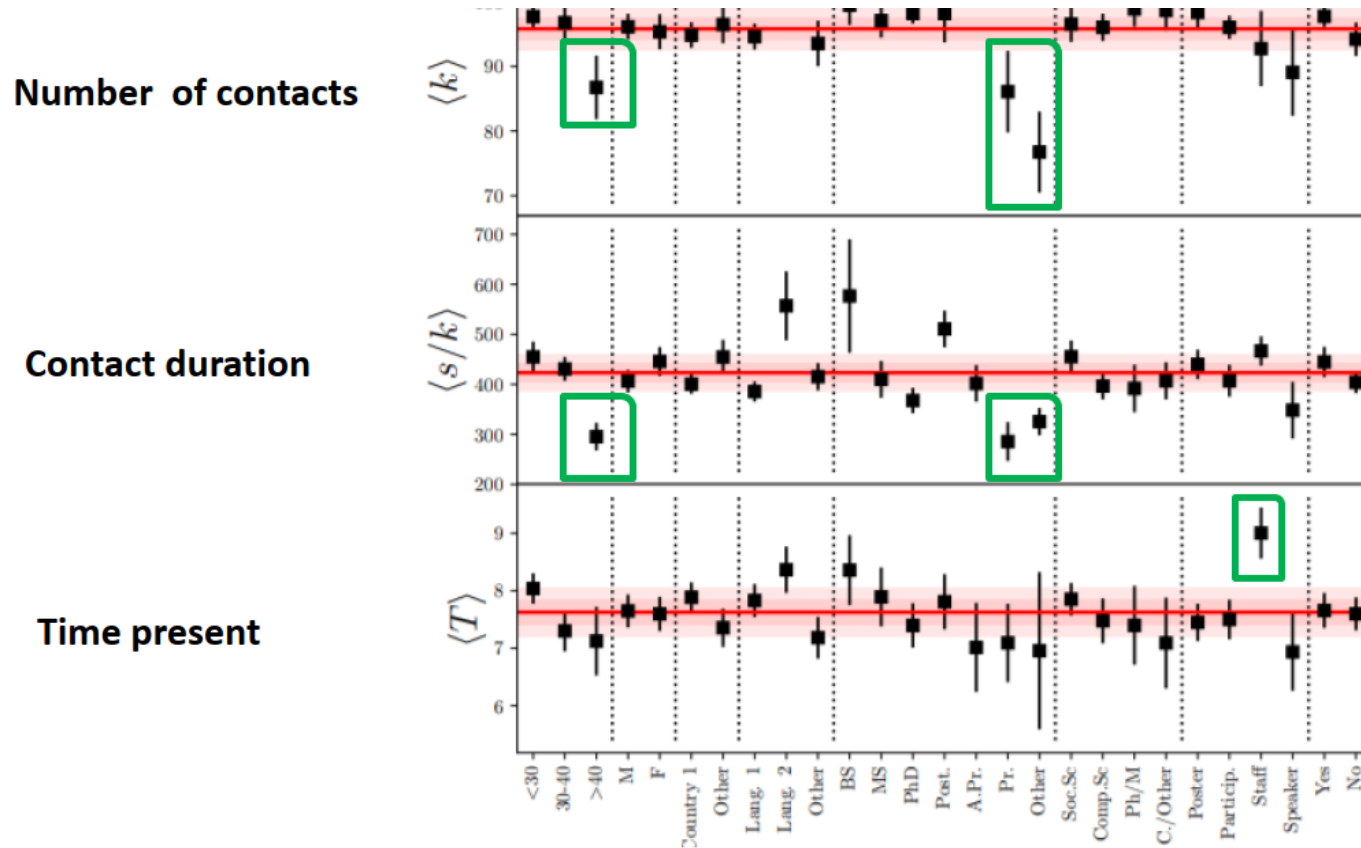
... hm, just guessing,
maybe... approximately ...

Digital devices or sensors may better recall certain facts than human memory.

Example: SocioPatterns – RFID sensors for measuring face to face interactions



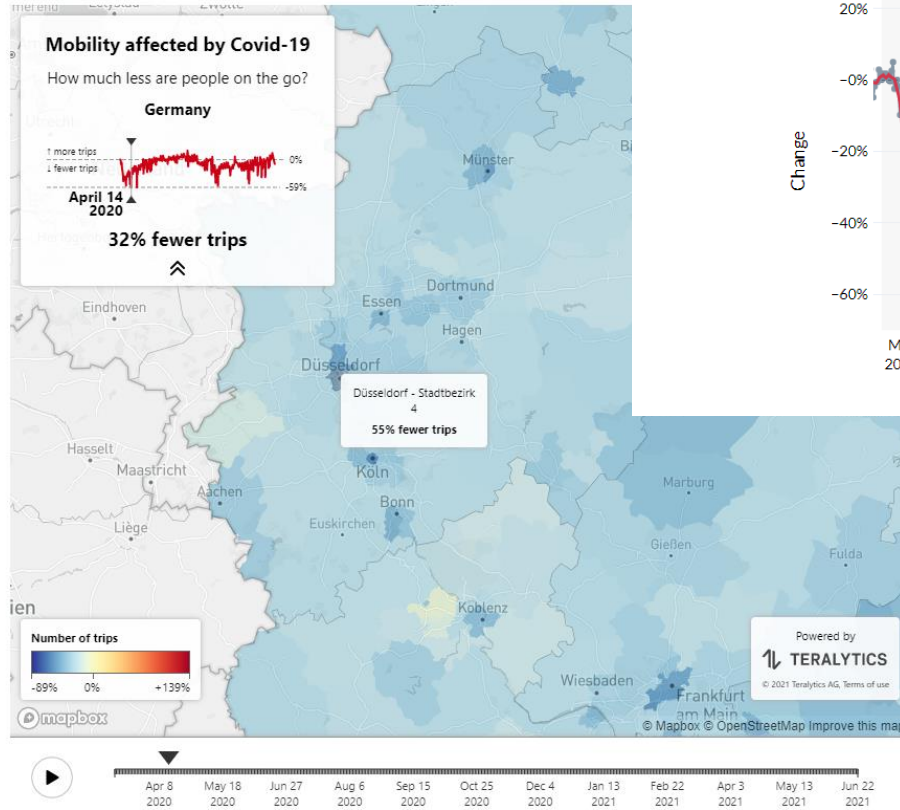
Example: SocioPatterns – RFID sensors for measuring face to face interactions



Génois, M., Zens, M., Lechner, C., Rammstedt, B., & Strohmaier, M. (2019). Building connections: How scientists meet each other during a conference. arXiv:1901.01182 [physics]. <http://arxiv.org/abs/1901.01182>

And often sensors are already built into everyday technology, producing **Digital Behavioral Data** as a "side product".

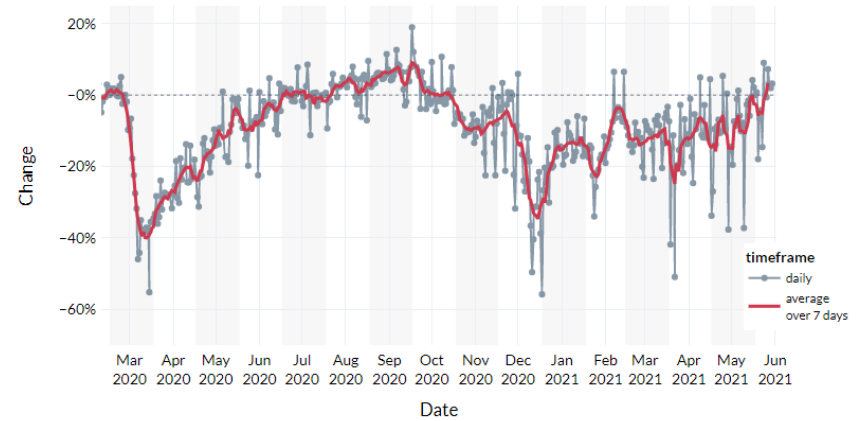
Mobility monitor



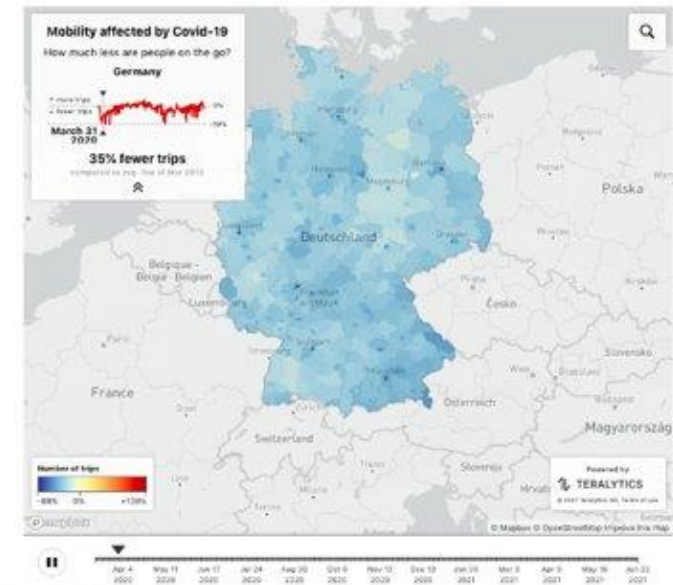
Change in mobility in Germany

compared to 2019

currently: 3% on 13. June



Mobility monitor



COVID-19 MOBILITY PROJECT

<https://www.covid-19-mobility.org/mobility-monitor/>

→ data from mobile phone companies.

Digital Behavioral Data

summarizes a variety of potential data sources, that record different types of activities.





The "ABC"

Digital behavioral data can help to identify

Attitudes and opinions,

Behavior,

Characteristics

of human users of digital technologies.

What did you think about
the preventive measures
during the first wave of
the covid-19 pandemic?

... well, let me
remember ...

Mining opinions from existing digital communication streams can be more timely than creating a survey. They are a valuable source, especially during unforeseeable events.



Sometimes digital behavioral data may enable looking into topics for which it would be difficult to recruit study participants otherwise.

They are often created without any stimulus from a researcher.

How do you feel?

Has your mood
changed over the past
6 weeks?

... sigh...

I feel good. As usual.

Measuring Subjective Wellbeing



Body functions
measured via
smartwatches

Usage patterns of
smartphone or social
media, writing style
in messages/posts



Photo content,
selfies, photo sharing



Luhmann, M. (2017). Using Big Data to study subjective well-being. *Current Opinion in Behavioral Sciences*, 18, 28–33. <https://doi.org/10.1016/j.cobeha.2017.07.006>

Studying online communities dedicated to intimate data

Confessional data selfies / life logging / data visualization



DataIsBeautiful

Beitreten

r/dataisbeautiful

Beiträge

Posting Rules

Top OC of the Week

Join our Discord

Robards, B., Lyall, B., & Moran, C. (2020). Confessional data selfies and intimate digital traces. *New Media & Society*, 146144482093403. <https://doi.org/10.1177/1461444820934032>

So, what is Computational Social Science?

Computational Social Science

is a research field that

- makes use of computational methods
- to collect and analyze data, especially large-scale and digital behavioral data,
- in order to study socially relevant phenomena.

“We define CSS as the development and application of computational methods to complex, typically large-scale, human (sometimes simulated) behavioral data.”

(Lazer et. al. 2020)

CSS as an evolving field

PERSPECTIVE | SOCIAL SCIENCE

Computational Social Science

David Lazer¹, Alex Pentland², Lada Adamic³, Sinan Aral^{2,4}, Albert-László Barabási⁵, Devon Brewer⁶, Nicholas C...

+ See all authors and affiliations

Science 06 Feb 2009:
Vol. 323, Issue 5915, pp. 721-723
DOI: 10.1126/science.1167742

SOCIAL SCIENCE

Computational social science: Obstacles and opportunities

Data sharing, research ethics, and incentives must improve

By **David M. J. Lazer^{1,2}, Alex Pentland³,
Duncan J. Watts⁴, Sinan Aral³, Susan
Athey⁵, Noshir Contractor⁶, Deen Freelon⁷,
Sandra Gonzalez-Bailon⁴, Gary King², Helen
Margetts^{8,9}, Alondra Nelson^{10,11}, Matthew
J. Salganik¹², Markus Strohmaier^{13,14},
Alessandro Vespignani¹, Claudia Wagner^{14,15}**

The field of computational social science (CSS) has exploded in prominence over the past decade, with thousands of papers published us-

dependencies within data. A loosely connected intellectual community of social scientists, computer scientists, statistical physicists, and others has coalesced under this umbrella phrase.

MISALIGNMENT OF UNIVERSITIES

Generally, incentives and structures at most universities are poorly aligned for this kind of multidisciplinary endeavor. Training tends to be siloed. Integrating computational training directly into social science (e.g.

els of administrative data rese
serving as platforms for analyzi
data while preserving privacy (c
important lessons for potential
with private companies, includ
opment of methodologies to l
data secure, yet accessible for
innovations in differential priv

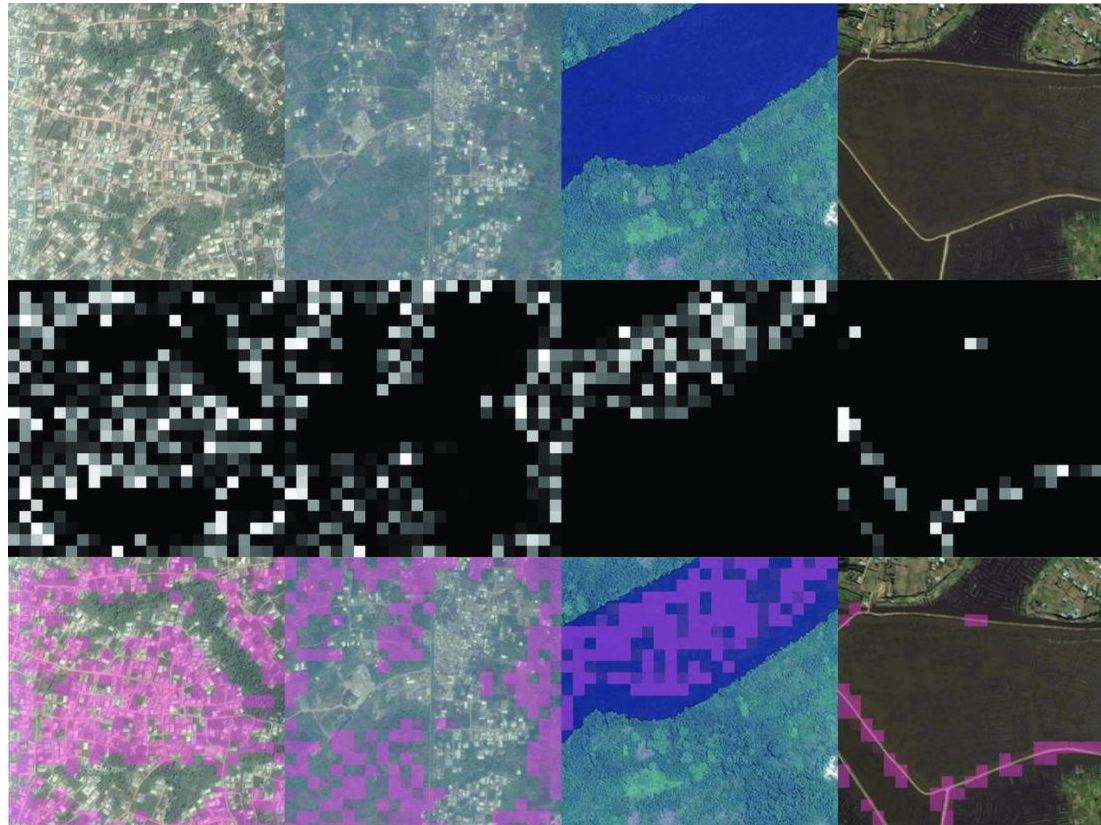
The value proposition for
panies is different and there l
dictably less progress. Data
government agencies are held
the public, whereas data held
are typically seen as a key pr
set. Public accountability inh
ing data is likely seen as a p
relevant stakeholders for gove
cies, but generally, far less so f
ers for private companies. A
from private companies is thu
able to academics, and when i
cally granted through a patc

Science 28 Aug 2020:
Vol. 369, Issue 6507, pp. 1060-1062
DOI: 10.1126/science.aaz8170

Computational methods may be a way to collect and analyze data at large scale, from various countries, over long periods of time.

Machine Learning methods are of particular interest to CSS researchers. Based on existing training data and the different data features, algorithms can „learn“ how to interpret or sort them into categories – and thus help to detect meaning in large datasets.

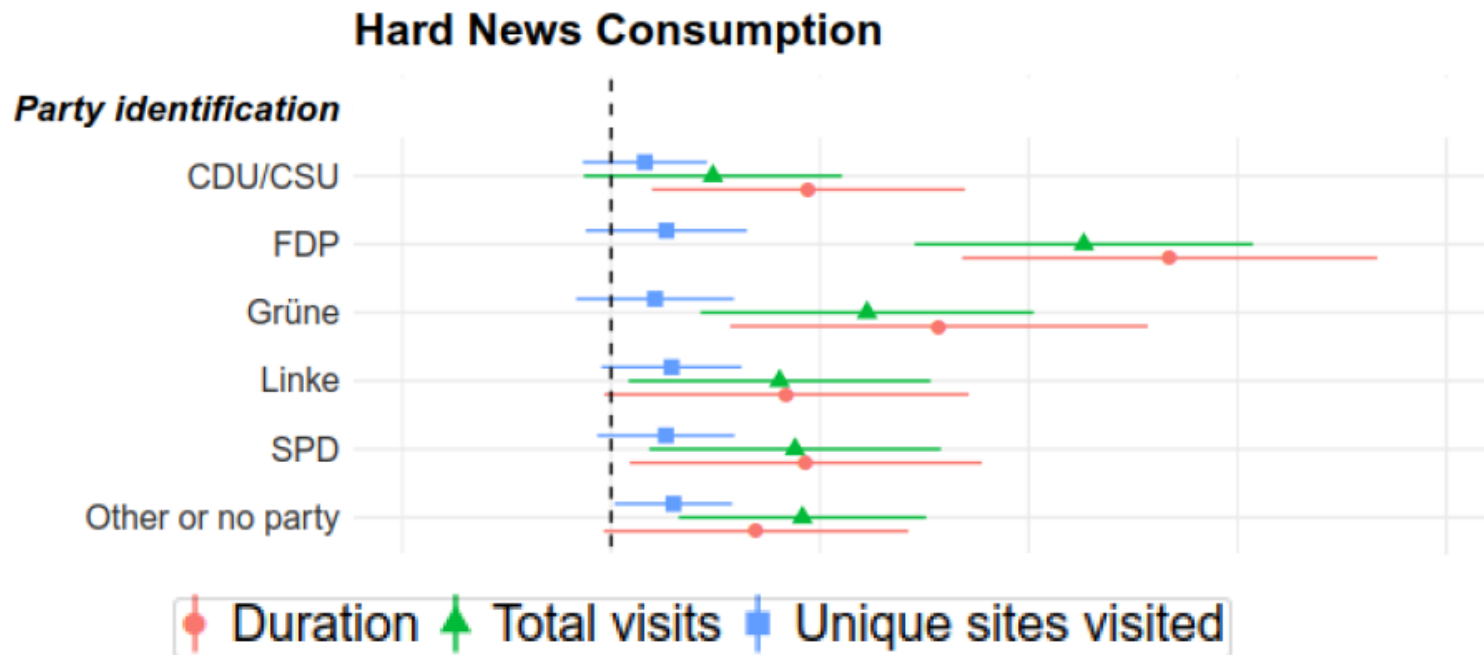
Visualizing Poverty



Jean, N., Burke, M., Xie, M., Davis, W. M., Lobell, D. B., & Ermon, S. (2016). Combining satellite imagery and machine learning to predict poverty. *Science*, 353(6301), 790–794. <https://doi.org/10.1126/science.aaf7894>

Digital Behavioral Data can also be combined
with survey data.

Combining Surveys and Web Tracking

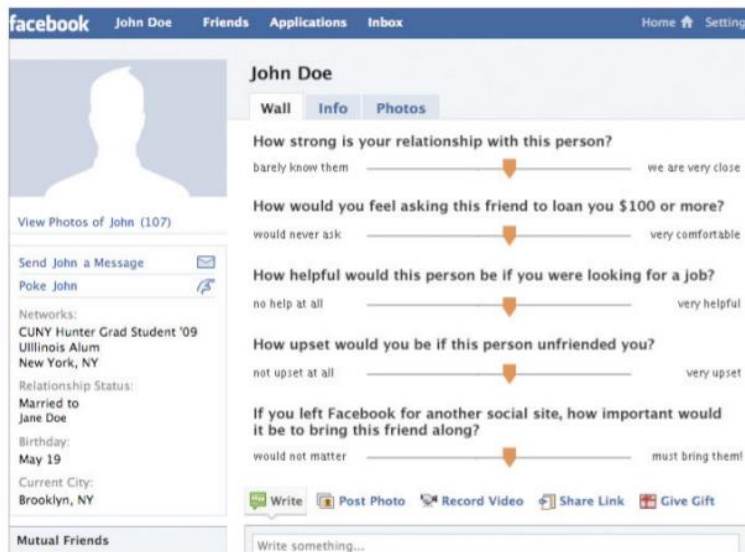


Stier, S., Breuer, J., Siegers, P., Gummer, T., Bleier, A., “Enemies of the People”: Party cues, populism and selective exposure to news – An investigation combining web tracking and survey data, 2019, working paper

Computational models can be developed to test social science concepts, e.g. with new types of data.

Predicting tie strengths

Can we predict loose acquaintances from close friends based on social media data?



How strong is your relationship with this person?

barely know them _____ we are very close

How would you feel asking this friend to loan you \$100 or more?

would never ask _____ very comfortable

How helpful would this person be if you were looking for a job?

no help at all _____ very helpful

How upset would you be if this person unfriended you?

not upset at all _____ very upset

If you left Facebook for another social site, how important would it be to bring this friend along?

would not matter _____ must bring them

Figure 1. The questions used to assess tie strength, embedded into a friend's profile as participants experienced them. An automated script guided participants through a random subset of their Facebook friends. As participants answered each question by dragging a slider, the script collected data describing the friendship. The questions reflect a diversity of views on tie strength.

Gilbert, E., & Karahalios, K. (2009). Predicting tie strength with social media. ACM Conference on Human Factors in Computing Systems - CHI '09, 211–220. <https://doi.org/10.1145/1518701.1518736>

Many approaches have been made for different types of predictions based on digital behavioral data – with varying success.

Predictions from Digital Behavioural Data

**Predicting elections with
twitter: What 140
characters reveal about
political sentiment,
Tumasjan et al., 2010**

**Detecting influenza
epidemics using search
engine query data.
Ginsberg et al., 2009**

Predictions from Digital Behavioral Data

**Predicting elections with
twitter: What 140**

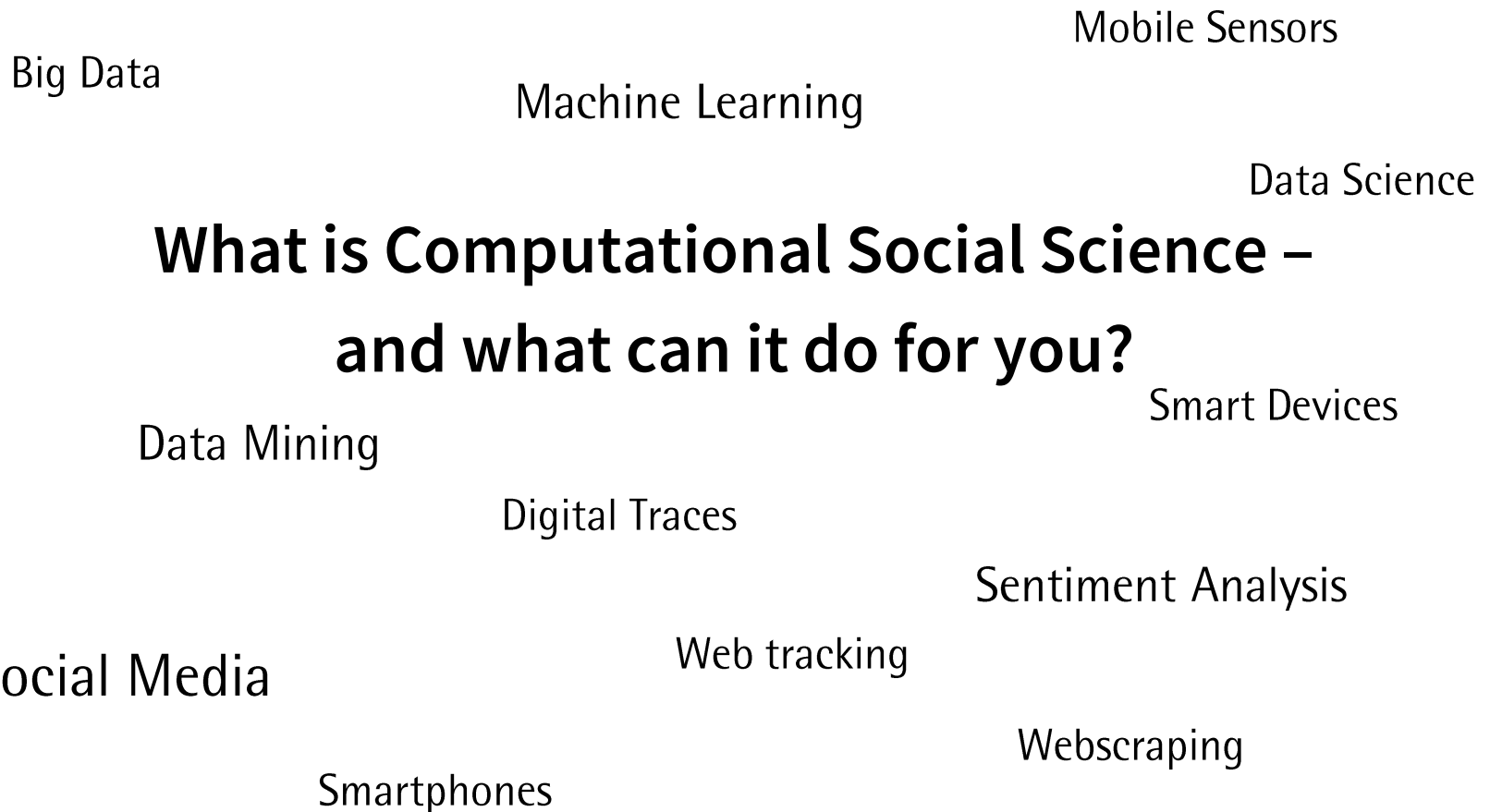
**characters reveal about
political sentiment,**
Tumasjan et al., 2010

**Why the pirate party won
the german election of
2009 or the trouble with
predictions: A response to
Tumasjan et al. [...]"**
Jungherr et al., 2012

**Detecting influenza
epidemics using search
engine query data**
Ginsberg et al., 2009

**The parable of Google Flu:
traps in big data analysis,**
Lazer et al., 2014

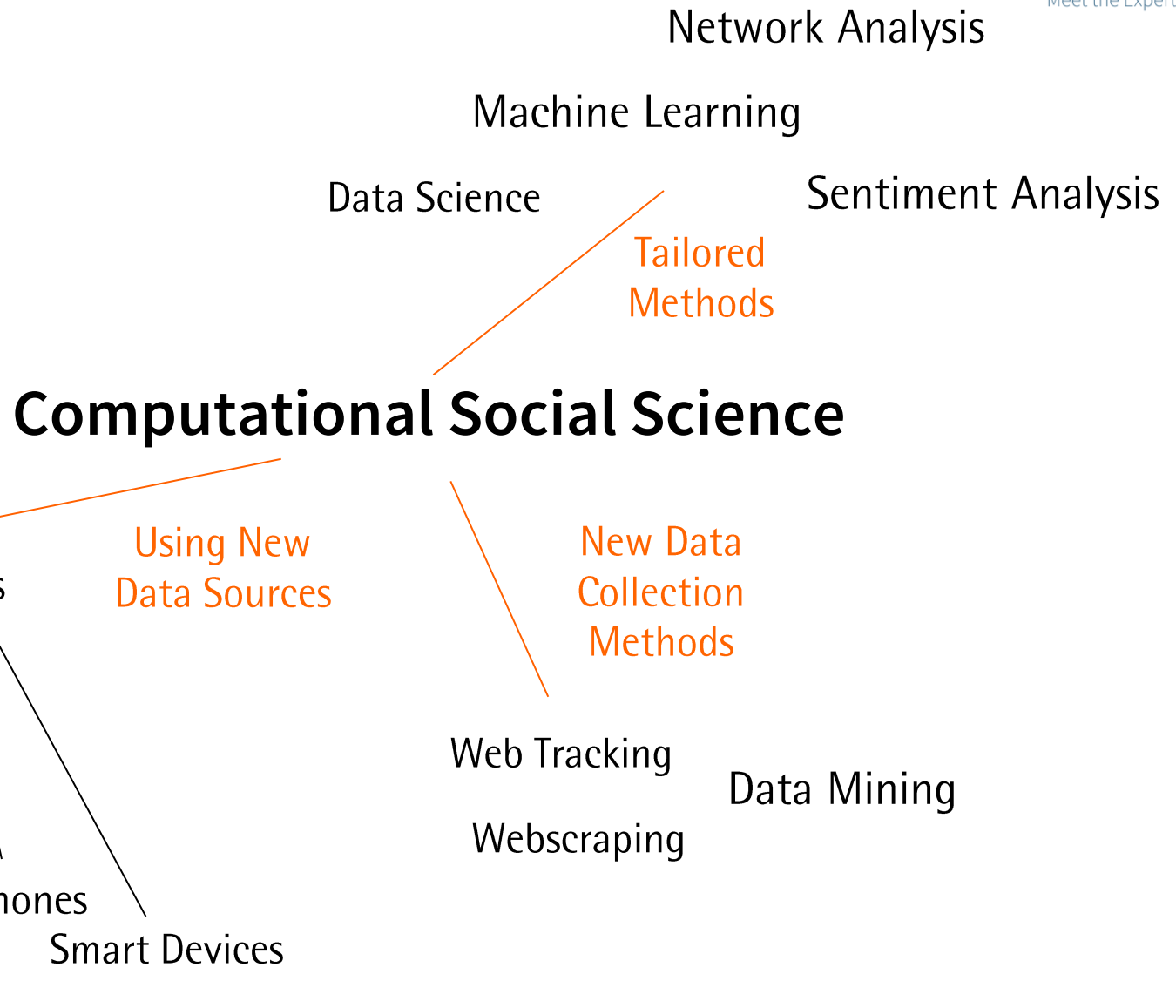
Computational Social Science (CSS)



Computational Social Science helps to

- study genuine digital phenomena
- collect and preprocess digital behavioral data
- apply new methods to analyze large scale datasets

CSS should be viewed as a new sub-area for social sciences, or as a new "tool box" to complement traditional social science approaches.





Data Quality?
Method Development?
Evaluation?
Collection Strategies?
Research Ethics?
Legal Constraints?

Computational Social Science

Network Analysis

Machine Learning

Data Science

Sentiment Analysis

Tailored
Methods

Using New
Data Sources

New Data
Collection
Methods

Big Data
Digital Traces

Social Media

Smartphones

Smart Devices

Mobile Sensors

Web Tracking

Webscraping

Data Mining

Outlook

Research with Digital Behavioral Data – more to come

01.07.2021

Digital traces of human behavior in online platforms – Research design and error sources

Fabian Flöck and Indira Sen

08.07.2021

Combining survey data and digital behavioral data

Sebastian Stier and Johannes Breuer

New *Meet the Experts* series with talks about CSS methods and data coming soon: September – December 2021

Other options to learn about CSS at GESIS

[GESIS Training](#) offers a wide range of seminars, workshops, and other courses, including:

- Sep 13- Oct. 1, 2021: [Fall Seminar in Computational Social Science](#)
- Nov 2-5, 2021: Workshop **Introduction to Social Media Research Data: Potentials and Pitfalls**, Katrin Weller and Indira Sen

GESIS Consulting

GESIS offers individual consulting in a number of areas – including survey design & methodology, data archiving, digital behavioral data & computational social science – and across the research data cycle. Please visit our website www.gesis.org for more detailed information.

GESIS consulting is *free of charge* for researchers who conduct

- scientific projects – financed institutionally or by third-party-funds – at universities or publicly funded research institutions, or
- scientific projects at institutions of the Federal Government or the *Länder* or other publicly funded institutions.


For other projects consulting is *subject to a charge* and to available resources.



Expert contact: katrin.weller@gesis.org

Please find on the GESIS website consulting contacts for:
[Planning Studies](#), [Accessing Data](#), [Analyzing Data](#), [Archiving Data](#)

More Services from GESIS

- [GESIS Survey Guidelines](#) provide short and hands-on explanations to frequent challenges in survey design and methodology.
- Use GESIS data services for [finding data](#) for secondary analysis and [sharing your own data](#).
- Get materials for [capacity building](#) in computational social science and take advantage of our expanding expertise and resources in [digital behavioral data](#).
- Check out the [GESIS blog](#) "Growing Knowledge in the Social Sciences" for topics, methods and discussions from the GESIS cosmos – and beyond.
- Keep up with GESIS activities and subscribe to our monthly [newsletter](#).
-  for publications, tools & services.

Thank you !

gesis

Leibniz-Institut
für Sozialwissenschaften

Leibniz
Leibniz
Gemeinschaft

Getting Started: Introductory Material for CSS

- Cioffi-Revilla, C. (2017). *Introduction to Computational Social Science: Principles and Applications* (2nd ed. 2017). Springer International Publishing : Imprint: Springer. <https://doi.org/10.1007/978-3-319-50131-4>
- Lazer, D. M. J., Pentland, A., Watts, D. J., Aral, S., Athey, S., Contractor, N., Freelon, D., Gonzalez-Bailon, S., King, G., Margetts, H., Nelson, A., Salganik, M. J., Strohmaier, M., Vespignani, A., & Wagner, C. (2020). Computational social science: Obstacles and opportunities. *Science*, 369(6507), 1060–1062. <https://doi.org/10.1126/science.aaz8170>
- Lazer, D., Pentland, A., Adamic, L., Aral, S., Barabasi, A.-L., Brewer, D., Christakis, N., Contractor, N., Fowler, J., Gutmann, M., Jebara, T., King, G., Macy, M., Roy, D., & Van Alstyne, M. (2009). SOCIAL SCIENCE: Computational Social Science. *Science*, 323(5915), 721–723. <https://doi.org/10.1126/science.1167742>
- Olteanu, A., Castillo, C., Diaz, F., & Kiciman, E. (2019). Social Data: Biases, Methodological Pitfalls, and Ethical Boundaries. *Frontiers in Big Data*, 2, 13. <https://doi.org/10.3389/fdata.2019.00013>

Getting Started: Introductory Material for CSS

- Ruths, D., & Pfeffer, J. (2014). Social media for large studies of behavior. *Science*, 346(6213), 1063–1064. <https://doi.org/10.1126/science.346.6213.1063>
- Salganik, M. J. (2018). *Bit by bit: Social research in the digital age*. Princeton University Press.
- Snee, H., Hine, C., Morey, Y., Roberts, S., & Watson, H. (2016). Digital Methods as Mainstream Methodology: An Introduction. In H. Snee, C. Hine, Y. Morey, S. Roberts, & H. Watson (Hrsg.), *Digital Methods for Social Science* (S. 1–11). Palgrave Macmillan UK. https://doi.org/10.1057/9781137453662_1

See also:

- SAGE Ocean's list of teaching materials: <https://ocean.sagepub.com/teaching-materials-for-computational-social-science>
- The Summer Institutes in Computational Social Sciences (SICSS): <https://sicss.io/curriculum>