

Dr. Peter Farago, Präsident CPS

Big Data in den Sozialwissenschaften – Herausforderungen und Chancen

Skizze für eine SAGW-Tagung im November 2018

1 Ausgangslage

Wie in der breiten Öffentlichkeit so ist auch in der sozialwissenschaftlichen Profession „Big Data“ zu einem Schlagwort geworden, an dem nicht vorbeizukommen ist. An der Tagung der European Survey Research Association ESRA im Juli 2017 mit rund 850 Teilnehmenden waren die von Mario Callegaro (Google Research UK) und Kollegen organisierten Sessions zu diesem Thema buchstäblich überlaufen. Der Artikel von Prof. Mick Couper (University of Michigan, Vorsitzender des Wissenschaftlichen Beirats von FORS) „Is the Sky Falling? New Technology, Changing Media, and the Future of Surveys“, publiziert in der führenden Methodenzeitschrift „Survey Research Methods“ ist mit über 1000 Downloads der mit Abstand am meisten beachtete Beitrag in diesem Journal (<https://ojs.ub.uni-konstanz.de/srm/article/view/5751/5289>). Und der nächste grosse wissenschaftliche Event zum Thema ist bereits angekündigt: Ende Oktober 2018 findet der Kongress „BigSurv18 – Big Data meets Survey Science“ (www.bigsurv18.org) statt.

In der Schweiz wird die Diskussion bislang vorwiegend unter informatikbezogenen und juristischen Aspekten geführt, so etwa beim unter der Ägide der beiden ETH eingerichteten Swiss Data Science Center SDSC (<https://datascience.ch/>). Auch das vom SNF durchgeführte NFP 75 „Big Data“ folgt weitgehend diesen Prioritäten (<http://www.nfp75.ch/de>). Lediglich eines der zur Zeit 36 Projekte ist direkt mit den Sozialwissenschaften befasst. Es trägt den weitgefassten Titel „Der Umgang mit Big Data: Methoden für eine Soziologie des 21. Jahrhunderts“, konzentriert sich jedoch auf die Methodenlehre in der deutschsprachigen Soziologie sowie auf die Berufsbilder „data science“ und „data journalism“. Es wird von Prof. Sophie Mützel (Universität Luzern) geleitet (<https://sites.google.com/view/facingbigdata>).

Bei all dem ist jedoch durchaus unklar, was mit „Big Data“ genau gemeint ist. Oft werden unrealistische Erwartungen an die Nutzung solcher Daten für die sozialwissenschaftliche Forschung, aber auch für evidence based politics geknüpft. Kleiner et al. 2015 (download: http://forscenter.ch/wp-content/uploads/2014/06/FORS_WPS_2015-02_Kleiner.pdf) haben in einer mit Fallstudien angereicherten Gesamtschau vier Typen von Big Data im Hinblick auf ihre forschungsrelevanten Stärken und Schwächen analysiert. Diese Studie ist ein geeigneter Ausgangspunkt für die Konzeption der Tagung, die der CPS an seiner letzten Sitzung in Aussicht genommen hat.

2 Zielsetzung, Zielpublikum

Die Tagung soll eher **policy-orientiert** als rein wissenschaftlich sein. Die Fragen nach Möglichkeiten und Grenzen von Big Data für die sozialwissenschaftliche Forschung und die daraus zu erwartenden Ergebnisse/Erkenntnisse für die gesellschaftliche und politische Praxis sollen in erster Linie diskutiert werden: z.B. die oft kolportierte, jedoch verkürzte Ansicht, Big

Data könnten Surveys ersetzen oder die mit der Verknüpfung unterschiedlicher Datenquellen einhergehenden daten- und personenschutzrechtlichen Probleme und deren Lösungsansätze.

Ein wichtiges Ziel der Tagung ist es, der Legende von den Big Data-ignoranten Sozialwissenschaften Fakten entgegenzusetzen, die zeigen, dass die Beschäftigung mit Big Data nicht nur schon längst am Laufen ist, sondern dass sie auch wissenschaftlich und praktisch relevante Resultate gezeitigt hat.

Als **Zielpublikum** sind dementsprechend Akteure aus dem öffentlichen Bereich (Bund, Kantone, swissuniversities etc.) sowie Multiplikatoren (Medien, Hochschulen) anzupeilen. Forschende sollen eher aktiv referieren als passiv teilnehmen.

3 Aufbau

Kleiner et al. unterscheiden vier Typen von Big Data: administrative Daten (z.B. aus den Sozialversicherungen), Social Media Daten (z.B. Twitter, Facebook), Transaktionsdaten (z.B. Kunden- oder Kreditkarten) und Textdaten (z.B. Inhalte von persönlichen Nachrichten). Und sie orten drei Arten von Herausforderungen: die analytische Nützlichkeit, die Datenqualität und Fragen des Zugangs zu den Daten.

In einem **ersten Teil** der Tagung könnten in einem Grundsatz- und in 2-3 beispielhaft vertiefenden Ergänzungsreferaten diese Herausforderungen illustriert und Wege zu ihrer Bewältigung aufgezeigt werden.

In einem **zweiten Teil** könnten im Rahmen eines moderierten Podiumsgesprächs unter Stakeholdern (s.u.) Folgerungen und Nutzenwendungen thematisiert werden.

Wünschbar ist ein abschliessendes Referat oder Statement zu Perspektiven und absehbaren Entwicklungen.

4 Stakeholders

Die Stakeholders werden in alphabetischer Reihenfolge aufgelistet, die keinerlei Wertung oder Prioritätensetzung impliziert. Die Aufzählung ist nicht abschliessend.

- Akademien und wissenschaftliche Gesellschaften
- Forschungseinrichtungen (akademische und nicht-akademische Stellen und Institute)
- Geldgeber (SNF, SBFJ, andere öffentliche und private Auftraggeber)
- Medien (Publikums- und Fachmedien)
- Öffentliche Statistik (BFS, andere statistikproduzierende Stellen bei Bund, Kantonen und Städten)
- Zivilgesellschaftliche Akteure (opendata.ch, Swiss Data Alliance und ähnliche auf dem Gebiet tätige Vereinigungen)

5 Sprache(n), Ort, Termin

Grundsätzlich wünschbar wäre eine **mehrsprachige** Durchführung (D,F,E). Das würde aber möglicherweise Simultanübersetzungen bedingen. Je nach Zusammensetzung der Referierenden müsste womöglich auf Englisch ausgewichen werden.

Wie bei nationalen Veranstaltungen üblich, sollte die Tagung in **Bern** stattfinden können.

Aus Gründen der Auslastung des SAGW-Tagungskalenders, aber auch im Hinblick auf den BigSurv18-Kongress Ende Oktober ist eine Durchführung im **November 2018** angezeigt. Wichtige Erkenntnisse aus diesem Kongress könnten so noch kurzfristig einbezogen werden.

6 Nächste Schritte

Die hier vorliegende, erst summarische Skizze ist in einem nächsten Schritt sowohl inhaltlich wie vorgehensmässig zu detaillieren: inhaltlich in Bezug auf die Zielsetzung und den Aufbau (und ggf. den Titel), vorgehensmässig in Bezug auf die Stakeholders und die Akquisition von Beiträgen (Ausschreibung oder Einladung). Die notwendigen Planungsschritte und deren Terminierung sind mit der Geschäftsstelle der SAGW zu koordinieren.

Für diese Arbeiten soll eine kleine Gruppe eingesetzt werden, bestehend aus dem Autor, einer Vertretung der SAGW-Geschäftsstelle sowie 1-2 weiteren sachkundigen Personen. Zusätzliches Know-how kann bei Bedarf gezielt beigezogen werden.