DATA INQUIRY: METHODOLOGICAL CONSIDERATIONS ON DATAFICATION IN SOCIAL RESEARCH

Irina Zakharova
University of Bremen

Background and objective

Datafication is widely acknowledged as a process "transforming all things under the sun into a data format" (van Dijck, 2017, p. 11). Data emerging out of such a transformation are often used for social research, although they are produced for purposes other than academic inquiry. For example, social media scholars analyse data produced by people using platforms designed by big tech companies to extract value; governance researchers make sense of data identified as relevant by administrative and political actors. This is also true for instruments used to retrieve and analyse data for academic purposes such as platform APIs or visualisations. As data, therefore, become both objects and instruments of academic inquiry, many scholars call for attention to the ways datafication reconfigures scholarly knowledge production, its methodological opportunities, and challenges (Lomborg et al., 2020). In this hindsight, it becomes ever more important to investigate the methodological approaches taken to study datafication and the concepts about it they provide. Exploring how the futures of datafication are produced by academic inquiry today allows for a better understanding of interdependencies between technological advances and social realities, required to continue important academic work towards challenging existing power relations.

To contribute to explorations of that interdependence within the domains of critical data studies and media studies, I extend on the notions of methods performativity (Barad, 2007) and methods assemblages (Law, 2004) developed within science and technology studies. A methods assemblage can be conceived of as "a continuing process of crafting and enacting necessary boundaries [and relations]" (Law, 2004, p. 144) associating researchers with other human and non-human actors, spaces, times, materialities, and discourses relevant to the research situations. Within the methods assemblage, all its elements (re)make each other and the world dynamically. I argue that defining datafication should be seen in concert with the methods assemblages applied to study these.

Methods

The presented data inquiry develops three kinds of methods assemblages based on a review of empirical social research on datafication. The key question is what kinds of methods assemblages are being applied in current datafication research and what concepts of datafication they produce. The contribution is based on a study of 32 expert interviews with leading scholars in the domains of critical data studies and media studies. The interview sample follows a literature review (Scopus, Web of Science) of original research articles on datafication published between 2015 and 2020 that reported findings from empirical projects. According to the purpose of study in identifying methods assemblages, conceptual work was excluded. For the interviews, authors—academics at various career stages—of the sampled literature were contacted. Additionally, interview partners were asked to recommend further experts. Informed consent was obtained, and all interviews and gathered materials were anonymised. Subsequently, qualitative content analysis (Saldaña, 2016) including both deductively and inductively developed categories was conducted. Deductive categories focused on the formal aspects of empirical research such as philosophical assumptions, particular research procedures (e.g. sampling), and "sociopolitical commitments" of the researchers (Onwuegbuzie & Frels, 2016, p. 51). Inductive categories were central for analysis and highlighted concepts about datafication produced in reported methods assemblages from a practice-based viewpoint. The interviewed experts discussed research on datafication in different societal domains, including the use of data and media in the everyday, data-driven governance, migration, education, activism. Most scholars reported applying a mix of qualitative methods (interviews, ethnographic methods, document analyses), while some applied software analyses, quantitative, computational methods, experiential, and theory-driven methodological approaches.

Preliminary results and implications

Empirically, methods assemblages include: researchers (and their positionings, subjectivities), the researched individuals, communities, organisations and /or things (e.g. documents, technologies), research situations in their spatial and temporal contexts. These entities are related to each other through sets of practices; some of these practices can be conceived of as research procedures (e.g. interviewing), other—organising research (e.g. securing funding), and communicating results. Central to distinguishing between methods assemblages are the *ways of associating of the involved actors and things*. In my analysis the questions of (1) what we are talking about when talking about datafication and (2) kinds of knowledges that researchers were interested in producing can be understood as such ways of associating (Table 1).

Three kinds of methods assemblages were developed:
- An encounter with data representations,
- Tracing dynamics of data infrastructures,
- Reconstructing algorithmic regimes.

Table 1. Three methods assemblages.

Methods assemblages	What are we talking about	What kinds of knowledges are sought?	Distinctive characteristics

	talking about datafication?		
Encounter with data representations	Data representations (visualisations or numbers— rankings, ratings)	Lived experiences of people in everyday and work settings	Learning *with* and not „only" about participants
Tracing dynamics of data infrastructures	Data infrastructures	Historical, social, cultural, political, economical aspects	Long-term focus on research situations and spaces
Reconstructing algorithmic regimes	Algorithms & software	Categorisations of the world inscribed in algorithms & software	Research on data production & processing

An encounter with data representations foregrounds people's lived experiences and can be configured to learn together with the study participants, allowing to explore why they (do not) value certain datafication processes (for an example of research illustrating some aspects of this methods assemblage see e.g. Kennedy & Hill, 2018). Tracing dynamics of data infrastructures situates these in their historical, socio-political contexts as scholars gather tacit knowledge about studied regions and matters over a long period of time (e.g. Taylor & Richter, 2017). Reconstructing algorithmic regimes addresses how certain categorisations of the world are inscribed in software and algorithms acknowledging the challenges of access to such information (e.g. Redden, 2018).

The methods assemblages contribute to critical data studies by producing accounts about datafication processes that are in concert with the methods assemblages applied to study these. They synthesise how current scholarship addresses datafication processes based on either data representations, data infrastructures, or algorithms based on the study of lived experiences, broader contextual aspects, or categorisations inscribed in software. The developed methods assemblages also provide critical data and media scholars with additional means to develop methodological sensitivity directed towards multiple aspects of datafication processes for their future inquiries.

References
Barad, K. M. (2007). *Meeting the universe halfway: Quantum physics and the entanglement of matter and meaning*. Duke University Press.

Kennedy, H., & Hill, R. L. (2018). The Feeling of Numbers: Emotions in Everyday Engagements with Data and Their Visualisation. *Sociology*, *52*(4), 830–848.

Law, J. (2004). *After method mess in social science research*. Routledge.

Lomborg, S., Dencik, L., & Moe, H. (2020). Methods for datafication, datafication of methods: Introduction to the Special Issue. *European Journal of Communication.*

Onwuegbuzie, A. J., & Frels, R. (2016). *Seven steps to a comprehensive literature review: A multimodal and cultural approach*. Sage.

Redden, J. (2018). Democratic governance in an age of datafication: Lessons from mapping government discourses and practices. *Big Data & Society*, *5*(2).

Saldaña, J. (2016). *The coding manual for qualitative researchers* (3E [Third edition]). SAGE.

Taylor, L., & Richter, C. (2017). The Power of Smart Solutions: Knowledge, Citizenship, and the Datafication of Bangalore's Water Supply. *Television & New Media*, *18*(8), 721–733.

van Dijck, J. (2017). Foreword. In M. T. Schäfer & K. van Es (Eds.), *The Datafied Society. Studying Culture through Data*. Amsterdam University Press.