Engaging with sensor-based methods for social sciences research is necessary, overdue and potentially rewarding

blogs.lse.ac.uk/impactofsocialsciences/2017/04/17/engaging-with-sensor-based-methods-for-social-sciences-research-is-necessary-overdue-and-potentially-rewarding/

*Sensors are an important source of big data. Developments at the heart of "smart cities" or the exploding "quantified self" movement are all reliant on sensors. However, attempts by social scientists to engage with sensors from a methodological perspective have been rare. **Jörg Müller** argues that such engagement is not only necessary and overdue, but also potentially rewarding. It's important to address concerns over the reliability of the data obtained, and also to remain cautious about certain interpretations of it, but sensors can open up new ways of doing social research.*

This post is part of our digital methodologies series, arising from a stream at the ISA Logic and Methodologies conference, September 2016, organised by Mark Carrigan, Christian Bokhove, Sarah Lewthwaite and Richard Wiggins.

"Big data" has become a major buzzword across the business sector and many scientific disciplines. The ease of access to data goes some way to explaining the popularity of social media-based research such as analysis of Twitter or other social networks, or the measurement of co-author collaboration patterns among scientists. Indeed, there are many examples of how social scientists engage with the abundant informational traces we leave behind. Surprisingly, however, data streams from sensors remain largely ignored by social scientists.

Sensors are an important source of big data. They may be situated at the heart of "smart cities" to monitor pollution and manage flows of traffic or people, for example. Assisted living, remote health monitoring for ageing populations, or "intelligent" buildings all rely on sensors. The exploding "quantified self" movement would be unthinkable without wearable sensors tracking heart rate, blood oxygen levels, sexual activity or sleep patterns. While sensors quietly accompany more and more of our actions, hands-on engagement from the social sciences community is scarce. We see the well-rehearsed critical reflex regarding the shortcomings and pitfalls of "smart cities", "intelligent houses", "quantified self", or "assisted living", but there are fewer attempts to engage with sensors from a methodological perspective, despite the potential to expand the social science toolkit beyond sample surveys, in-depth interviews, and participant observations among others.

Image credit: smart city by Rene Adamos. This work is licensed under a CC BY-SA 2.0 license.

In part, this reluctance to engage with sensors from a methodological vantage point is arguably due to the suspect revival of a radical behaviourism implicit in many sensor-based applications. Sensors deliver data that undercut the subject's consciousness and agency; the very purpose of health apps is to monitor our actions beyond our good intentions, for example. As Steven Pool recently wrote in *The New Statesman*, "by using such devices, we voluntarily offer ourselves up to a denial of our voluntary selves, becoming atomised lab rats, to be manipulated electronically through the corporate cloud". Alongside a healthy skepticism of a bleak, Orwellian future of total surveillance through sensors, it is perhaps not surprising that practical engagements from social scientists have been sparse at best, specifically relating to methodological innovations. However, it might be that engagement is not only necessary and overdue, but also potentially rewarding.

Addressing the reliability of sensor data

The first argument for exploring sensors as new methodological tools builds upon a basic concern over the reliability of the data obtained. In cases where research uses this type of data, sensor data is often taken as a self-evident, objective window of reality. Since sensors deliver data without any obvious involvement on the part of a human researcher, one could assume that the obtained data is a reliable representation. However, this is certainly not the case. Sensors break down, wear out and need to be calibrated. Currently, there are no reporting standards for the reliability of sensor-based data measurements. Does the same sensor deliver the same data at different points in time? And under varying atmospheric conditions? How about similar sensor types but different manufacturers? Sensor readings, even under ideal conditions, are never 100% precise but can fluctuate within a certain range. As long as this random error is normally distributed, the "true" value should emerge over large number of observations. But this only raises the question of how long sensors should be deployed in the field; what's the minimum timeframe for getting a reliable mapping of face-to-face interaction patterns within a group of people, for example? Is one day enough? A week? A month? At what point do readings become redundant? There is an urgent need for basic methodological considerations, including quality reporting standards, to be developed.

Addressing the interpretations of sensor data

The second argument to take sensor-based methodologies more seriously has to do with the very jurisdiction of the

social sciences. Machines continuously approximate human capabilities and supersede them. Computers are now better than humans on face recognition tasks. The AI community is thus moving on, addressing new challenges such as "social signal processing". The aim is to enable machines to "understand" social situations through real-world locomotion and interaction patterns among people. What merits scrutiny here, especially from the social sciences, is the often black-boxed leap from the available raw sensor data to established higher-level social and psychological constructs.

Take proximity data as an example; i.e. sensors that detect physical co-presence between persons. This has been used to study the influence of "shared paths to the lab" on scientific collaboration. Here, physical co-location is interpreted as an opportunity for social interaction that has effects on research output. But the question is: under what circumstances is "proximity", as detected by sensors, really significant for a socially meaningful exchange? Is it that people are actually (emotionally) close or is it just that they share the same desk in a small office while quite possibly also despising each other?

In another example, what assumptions are built into algorithms that identify " the most dominant person in a group"? Up to now there has been little research on the contingencies of sensor data; i.e. the discriminant validity of these digital footprints of social interactions. The question of how to move from sensor data to sociological or psychological constructs should not be left to the AI community exclusively.

Opening up new ways of doing (social) research using sensor data

Additionally, sensor data may open up or, at least, facilitate access to hard-to-reach new empirical fields. Take non-verbal behavior, for example. Only part of our communication in face-to-face interactions is intentional and explicit. Non-verbal body language provides a battery of contextual cues starting with the appearance of people but also involving tone of voice, gestures, gaze, and even include spatiotemporal codes such as the perception of (private) space around one's body. Sensors offer the possibility to expand upon previous work on how stereotypes and bias shape these (non)verbal behaviors and interaction patterns. Staying with the example of non-verbal behavior, expectation states theory explains how interpersonal status hierarchies emerge when groups of people interact. This has been particularly demonstrated in relation to gendered competency expectations. The use of new, sensor-based, sociometric data raises a number of interesting (methodological) questions that allow us to reexamine and expand upon these findings in light of what Alex Pentland calls "honest signals". Honest signals can reveal patterns of dominance between interaction partners but also predict the outcomes of speed dating or business pitches by measuring largely unconscious, nonverbal signals. Leaving aside the self-optimisation reflex that underlies many of these applications, these sociometric data open up new ways of doing (social) research.

In short, if the above concerns over sensor-based data are addressed, exciting research opportunities can be developed; research that not only scrutinises hard-to-observe nonverbal behavior but also offers the possibility of better engagement with the temporal dynamics of social interactions. Finally, there is, of course, the need to embed these technological possibilities within a debate on the wider ethical implications and self-optimisation imperatives available through sensor-based data. Ignoring it for whatever reason won't make it go away.

Note: This article gives the views of the author, and not the position of the LSE Impact Blog, nor of the London School of Economics. Please review our comments policy if you have any concerns on posting a comment below.

About the author

__Jörg Müller__ is a senior researcher at the Internet Interdisciplinary Institute (IN3) at the Universitat Oberta de Catalunya, Spain. He is currently coordinating a H2020 project GEDII – Gender-Diverstiy-Impact – which looks at the influence of gender diversity in research teams. He is particularly interested in practice-based approaches to gender (in)equality in science, besides exploring the implications of (sensor-based) data developments from a gender perspective. His ORCID iD is 0000-0001-7727-2117.