Querying into Data in Asylum Decision-making

Author: Trine Rask Nielsen, E-mail: trn@di.ku.dk

03-06-2021

In this blog post, I will give some examples from the DATA4ALL project on how we have set up weekly 'data crunching' sessions as a first important step towards querying into data from asylum cases. These sessions were set up as a collective space for querying into data, taking different topics (e.g., fairness in AI) as a lens. In these sessions, we have explored, across the team of data scientists, legal scholars, and computer scientists, data-driven technologies in the context of asylum decision-making, what it means to produce, exchange, and make sense of data, and what a responsible approach to our research will look like, coming from disciplines across the sciences. These 'crunching data' sessions became an important driver for the initial understanding of asylum cases that represent a critical part of our data in the DATA4ALL project and for shaping my own perspective (as a new PhD student) on asylum data.

What do we mean by data?

The 'crunching data' sessions have, for me at least, led to some very interesting, insightful, and important discussions and reflections such as the difficulty in prototyping counter data or what is sometimes termed missing data: data that are currently not being collected because of e.g., bias, lack of social and political will, or structural oppression (D'Ignazio et al. 2020). What once exist in the formal asylum cases is thus difficult to imagine differently. One of the things that I am continuously reminded of from our sessions is, that it is critical that we have these ongoing conversations about how data is produced and the ways in which they harbour interpretations within our team: speculating is one technique to get closer to what could be important counter data in an asylum context but requires translation to other disciplines where the concept of what counts as data and data science techniques is different. In this regard, questions have come up such as: what do we even mean, when we, for example, speak about 'data', 'raw data', and 'cleaned data' within our interdisciplinary research team? What kind of interpretation do we cook into these concepts? Do we interpret these concepts equally across our different research disciplines?

Data and numbers never speak for themselves

When working with data, we as researchers, are responsible for actively preventing numbers from speaking for themselves, especially when the numbers have to do with people and in this case, vulnerable people. Data most often entails a reduction in complexity and context, so in order to get one step closer to conducting ethical data analysis responsibly, situating data in context, is therefore necessary for making any real sense of the data (Møller et al. 2019; Randall et al. 2007). In other

words, it is essential to unpack and understand the complexity of what the data actually represents and contextualize and evaluate the findings of our research to ensure that its "situatedness" (Haraway 1998) is taken into account.

Context as a necessity for making sense of data

This means, that just like the concept of data, it is important that we also discuss the notion of 'context' as a slippery concept (Dourish 2004). When context is not appropriately considered, "data lose meaning and value" (Neff et al. 2017 following boyd et al. 2012). Similarly, we must ask critical questions, within our team, about the kind of work 'context' is doing. It is important that we bring attention to the ways in which context may be conceived differently and understand the perhaps different approaches to the concept of context across disciplines.

> "On the one hand, it is a technical notion, one that offers system developers new ways to conceptualise human action and the relationship between that action and computational systems to support it. On the other hand, it is also a notion drawn from social science, drawing analytic attention to certain aspects of social settings" (Dourish 2004 pp. 20).

Thus, in order to make sense of data, it is therefore vital that we are aware of the fact that the notion of context can be understood as one thing in social science and as another in data science. What are we really talking about when we talk about context within our team? A stable description of the world? Or details of lived experience?

Interdisciplinary research takes effort and care

As the 'crunching data' sessions have changed into more focused sessions of data analysis, we reflect on our process across the interdisciplinary team of researchers. The establishing of a shared vocabulary is difficult across disciplines, we learned, and as a PhD student, wanting to work both independently and contribute, I must admit, that I sometimes have found these sessions a bit overwhelming. However, the time it takes to become familiar with the asylum domain and get a clearer view of my role in the project – and other peoples' roles – is a necessary part of conducting the kind of interdisciplinary research that the DATA4ALL project set out to do. The sensemaking of our own disciplines and epistemologies is a task that takes effort and care, as also identified by Neff

(2017) and others: Sensemaking is a collective process and thus, in order to take a responsible approach to our research, it is imperative that we keep on prioritizing conversations within our interdisciplinary team around the production, interpretation, and use of data.

My perspective on querying into data in asylum decision-making

One important learning for me, at this point of the project is, that there are no such thing as "raw" data collection. Rather, data collection practices are shaped by values and judgement. This means that data production and interpretations of data make who the 'asylum seeker' is – perhaps even through data the individual applying for asylum has no idea is being produced, shared, and used in the decision-making process. Whether the applicant recognize the data or not, these data will affect the asylum decision-making. Thus, from my perspective on data in asylum decision-making, important questions concerning data justice arises such as: Who determines where the data comes from and what data is available and what data is not? Who is not in the data? How do legal asylum authorities ensure that all individuals applying for asylum are treated properly and equally – that is, that data points actually represent the individual applying for asylum and that all individuals applying for asylum are part of the data production, processing, and sharing on equal terms? What and whose values are encoded and reproduced in the data collection informing asylum decision-making?

References

Boyd, D. and Crawford, K. (2012): CRITICAL QUESTIONS FOR BIG DATA. Information, Communication & Society, 15, 662 - 679.

D'Ignazio, C., and Klein, L. F. (2020): Data feminism, MIT Press.

Dourish. P. (2004): What we talk about when we talk about context. Personal Ubiquitous Comput. 8, 1 (February 2004), 19–30.

Haraway, D. (1988): Situated knowledges: The science question in feminism and the privilege of partial perspective. Feminist studies 14, 3 (1988), 575–599.

Møller, N. H., Fitzpatrick, G., and Le Dantec, C. A. (2019): Assembling the Case: Citizens' Strategies for Exercising Authority and Personal Autonomy in Social Welfare. Proceedings of the ACM on Human-Computer Interaction 3, GROUP (2019), 1–21.

Neff, G., Tanweer, A., Fiore-Gartland, B., and Osburn, L. (2017): Critique and contribute: A practice-based framework for improving critical data studies and data science. Big Data, 5(2), 85–97.

Randall, d. Harper, R., and Rouncefield, M. (2007): Fieldwork for design: theory and practice. Springer Science & Business Media.