Digitizing physical presence across spatiotemporal dimensions

I had planned to digitize my physical presence on my bike in regards to spatiotemporal dimensions (Wright et al., 2003, p. 48) of data, in order to probe the epistemic implications for such data. In order to accomplish this, my plan had been to decide on a data-collecting platform fit for the task; I chose Strava because of its approachability, supposed reliability, and – critically – its promises to uphold my rights to export any acquired data relating to me in a machine-readable format, as a EU citizen. I then strapped my phone to my road-bike via a cheaply acquired silicone phone-holder designed for bikes, and routinely and systematically recorded any and all bike rides conducted on my road-bike – meaning regular day-to-day commutes have been excluded from collection. The data collection was in effect from August 27th until September 24th of 2020.

A part of what I wanted to experience and examine was indeed "going in blind" from that point on; not knowing exactly in what format the data would be delivered to me, and, critically, with which taxonomies and curational efforts (in terms of categories, metadata and structuring) *already* applied to it, as described by Weinstock (2020) as a

> "[…] curatorial exploration of personal cosmologies placed a special emphasis on catalogues, collections, and taxonomies - formats that seem to have proven most qualified to suggest idiosyncratic reconnections of "the order of things"" (Weinstock, 2020, p. 231).

The entire dataset has been appended as the ZIP-file named "export_66439101.zip". I've been careful in the preservation of the original structure and naming-schemes, as a measure of portraying the above-mentioned efforts. I have, though, made sure to anonymize pieces of data which can be considered personally identifiable throughout the dataset. But I haven't anonymized the start- and end-coordinates from the rides, neither in the 'raw' (Gitelman, 2013) activity-data, nor as they have been converted from GPX- into CSV-data via the website "https://www.gpsvisualizer.com", as I would consider that 'tampering' with the data itself. I consider the aggregated dataset consisting of timestamped coordinates as the end result of my inquiry. Removing singular data entries on a subjective basis would endanger those exact epistemic considerations which I seek

to question. That dataset has been appended as 20200924080218-22138-data.csv. Some ethical considerations here have been weighed between the ability of the data to expose personally identifiable information (relating to me and myself only, critically), and the benefits of research (Markham & Buchanan, 2017).

The data handed over to me, about me, from Strava consist of mainly quantitatively generated metadata (such as total lengths, durations spent, elevations and speed), but there are a few critical 'sprinkles' of qualitative data, consisting of my perceived exertion (as Strava prompts the user to enter when an activity finished), and naming of the activity – which coincidentally is algorithmically determined if not manually filled in, being attributed various quantitative metadata acquired from the activity itself and the chosen privacy settings (most of my bike-rides are simply named "Afternoon Ride"; what does that tell me about me, collectively, and the rides in themselves?). I found it interesting how these metadata are algorithmically derived, although at some point determined by a human actor. More on this consideration from an ANT-perspective below.

I was reacting to the data as the process of converting and aggregating GPX to CSV in a meaningful way. Getting the data off of the collection platform (Strava, in this case) was at no time throughout the process considered as difficult, as it is my right, defined in the GDPR (European Parliament and Council of European Union, 2016, article 15). It posed a challenge actually remembering and differentiating my memories of the bike rides, and in order to name the individual rides in the final dataset (as a way of distancing myself from the algorithmically derived names) I had to consult my calendar. The questionable nature of the naming process is reflected in the names, as they appear in the dataset.

Throughput this part of the process, I was questioning the epistemic implications of the digitization of my "being-in-the-world" (Zahavi, 2019), and whether I was building a dataset which corresponds to my own being-in-the-world, or rather just collecting a bunch of timestamped coordinates? A phenomenological perspective, perhaps with a hermeneutical angle, would most likely prove highly interesting.

Considering an ANT-perspective

Following the same train of thought going forward as Nanna Thylstrup with "[t]hinking about mass digitization as an "assemblage" [which] allows us to abandon the image of a circumscribed entity in favor of approaching it as an aggregate of many highly varied

components and their contingent connections" (Thylstrup, 2018, p. 23) seems, to me, quite prolific in terms of approaching a such complex phenomenon as mass digitization. The way of perceiving and approaching a process of mass digitization, as Thylstrup describes with a quote from Bruno Latour; "Groups are not silent things, but rather the provisional product of a constant uproar made by the millions of contradictory voices about what is a group and who pertains to what." (Latour, 2005, p. 35; as quoted in Thylstrup, 2018, pp. 23–24), regarding a founding principle of the Actor Network Theory (ANT) approach of the dissecting and disassembling of seemingly tightly knitted networks which can tend to appear as a single unit – in this case the entire concept and phenomenon of mass digitization.

Considering the collected GPS-data as mass digitization

With a departure of the 'mundane' data (Pink et al., 2017) being termed as 'social', the considerations of Langlois (2015) makes sense, in that

> "social data is not only a product of computer science; it is also linked to processes of archiving, of remembering and forgetting. From this perspective, social data becomes thick and multidimensional. It does not simply exist to be classified, measured, and correlated, but as a trace of practices that establish cultural and social continuities and discontinuities" (Langlois et al., 2015, p. 11).

I, as an individual, has physically been where the data is pointing and experienced that which has been digitized – that's the only reason the data's there in the first place. This digitized "external storing" of my physical presence is being called into existence purely for archival and statistical reasons, as it – in its own – does not enhance what I knew, or add anything inherently new, referring to what Richard Rogers (2017) describes as 'Natively Digital' (Rogers, 2017). But in the 'mass'-part of the term 'mass digitization', something is happening with the data; the aggregation of data as a whole – a seemingly mundane transformation of data, from a bunch of timestamped coordinates into a 'set' of data, upon which metadata can be extracted, and mathematical equations can be performed.

Bibliography

European Parliament and Council of European Union. (2016, April 27). *Regulation (EU) 2016/679* [Official Journal of the European Union]. Official Journal of the European Union. https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:32016R0679&rid=3

Gitelman, L. (2013). *"Raw data" is an oxymoron*. The MIT Press.

Langlois, G., Redden, J., & Elmer, G. (Eds.). (2015). *Compromised data: From social media to big data*. Bloomsbury Academic, an imprint of Bloomsbury Publishing, Inc.

Latour, Bruno. (2005). *Reassembling the social an introduction to actor-network-theory*. OUP.

Markham, A., & Buchanan, E. (2017). Research Ethics in Context. In M. T. Schäfer & K. van Es (Eds.), *The Datafied Society* (pp. 201–210). Amsterdam University Press; JSTOR. https://doi.org/10.2307/j.ctt1v2xsqn.19

Pink, S., Sumartojo, S., Lupton, D., & Heyes La Bond, C. (2017). Mundane data: The routines, contingencies and accomplishments of digital living. *Big Data & Society*, *4*(1), 12. https://doi.org/10.1177/2053951717700924

Rogers, R. (2017). Foundations of Digital Methods. In M. T. Schäfer & K. van Es (Eds.), *The Datafied Society* (pp. 75–94). Amsterdam University Press. https://doi.org/10.1515/9789048531011-008

Thylstrup, N. B. (2018). *The politics of mass digitization*. The MIT Press.

Weinstock, K. (2020). Rearranging the World: Found Objects and the Collection, Pre- and Post-Internet. In B. von Bismarck & B. Meyer-Krahmer (Eds.), *Curatorial Things: Cultures of the Curatorial 4* (pp. 229–253). STERNBER PR.

Wright, P., McCarthy, J., & Meekison, L. (2003). Making Sense of Experience. In M. A. Blythe, K. Overbeeke, A. F. Monk, & P. C. Wright (Eds.), *Funology* (Vol. 3, pp. 43–53). Springer Netherlands. https://doi.org/10.1007/1-4020-2967-5_5

Zahavi, D. (2019). *Phenomenology: The basics* (Original edition). Routledge, Taylor & Francis Group.