Selected Papers of #AoIR2024:
The 25th Annual Conference of the
Association of Internet Researchers
Sheffield, UK / 30 Oct - 2 Nov 2024

VISUALIZING 10 THOUSAND CITIES? UBER'S DATA STORIES ON KNOWING URBAN SPACE

Abel Guerra
London School of Economics and Political Science

In 2020, Uber celebrated its global scaling by highlighting its granularity. More than simply entering 70 countries, the platform announced: "we're excited as ever to deliver on our vision in the over 10,000 communities [cities] we are in" (Uber, 2020). To "deliver a vision" based on the algorithmic management of labour and territories, the platform creates visions of these diverse cities and communities - it makes them visible and intelligible by investing on a unified vision, ready to be algorithmically processed, anticipated and managed. Such is the role of H3, a now open-sourced hierarchical hexagonal grid system for indexing geospatial data, cleaning up the messiness of cities by datafying them.

Representation of H3's hexagonal grid layered over a city map. The original caption reads "cars in a city; cars in hexagons; and hexagons shaded by number of cars" (UEB, 2018)

While workers on the ground precariously move across territories, Uber's engineering teams invest in abstract and disembodied views of the city. Taking Uber's knowledge production about cities as an object of inquiry, this paper explores the epistemic dimensions of platformisation, delving into platforms technical and narrative reliance on datafication as they position themselves as relevant epistemic actors.

Besides its own team of experts, Uber enrols a wider community of practitioners through open-sourced software, APIs, and databases, as exemplified by H3 and Deck.gl, a

Suggested Citation (APA): Guerra, A. (2024, October). *Visualizing 10 thousand cities? Uber's Data Stories on Knowing Urban Space*. Paper presented at AoIR2024: The 25th Annual Conference of the Association of Internet Researchers. Sheffield, UK: AoIR. Retrieved from http://spir.aoir.org.

framework for geospatial data visualisation. In this paper, I delve into Uber Engineering Blog[1] (UEB) to explore Uber's positioning within a wider spatial knowledge infrastructure (Edwards, 2010; Plantin, 2018) in articulation with techno-corporate discourses, focusing on the platform's data narratives about knowing urban space.

Following Dourish and Gómez Cruz (2018) propositions on narrating data and narrating *with* data, I use data narratives to refer to processes "by which data is found to be meaningful" (p.5), attending to metaphors, tropes and conventions at work, as well as patterns of interpretation that may be indicative of epistemic orders imposed into urban space and those inhabiting it. Considering this, I ask: what stories about urban space are told by Uber and its engineers with and through data? How do these stories make data itself meaningful? What counts as knowledge and who counts as a knowing subject?

Methods and Data

Similarly to other tech companies' engineering blogs, UEB compiles a series of publications authored by Uber's data and engineering experts and interviews with said experts. UEB combines the community appeal and the chronological organisation of early practitioners' software engineering blogs (Rainer & Williams, 2019; Pagano & Maleej, 2013) to a PR move towards ideals of transparency and public interest. As such, it works as a "performed backstage" (Gillespie 2014; Hilgartner, 2000) that allows the public to peak behind the curtains of its technological developments and research culture.

It is also a venue for the dissemination of APIs and open-sourced initiatives with a collection of external links to Github repositories and use cases. The driving community *ethos*, often summarised in a gesture of "giving back" to the community, is central to an understanding as to who is addressed as a knowing subjects with shared interests, meanings, practices – industry, researchers, independent developers, and amateur enthusiasts.

This combination of a technical and corporate discourse makes UEB a rich source to explore some of the narratives and imaginaries that permeate Uber's knowledge production vis-à-vis its business strategies, as well as the epistemological assumptions inscribed in it. For this study, I manually collected 105 blog posts from UEB, screened them in detail according to their stance towards the production of knowledge about urban space. 41 selected posts, from August 2015 to October 2022, were analysed through inductive coding and situational analysis (Clarke, 2005).

Discussion

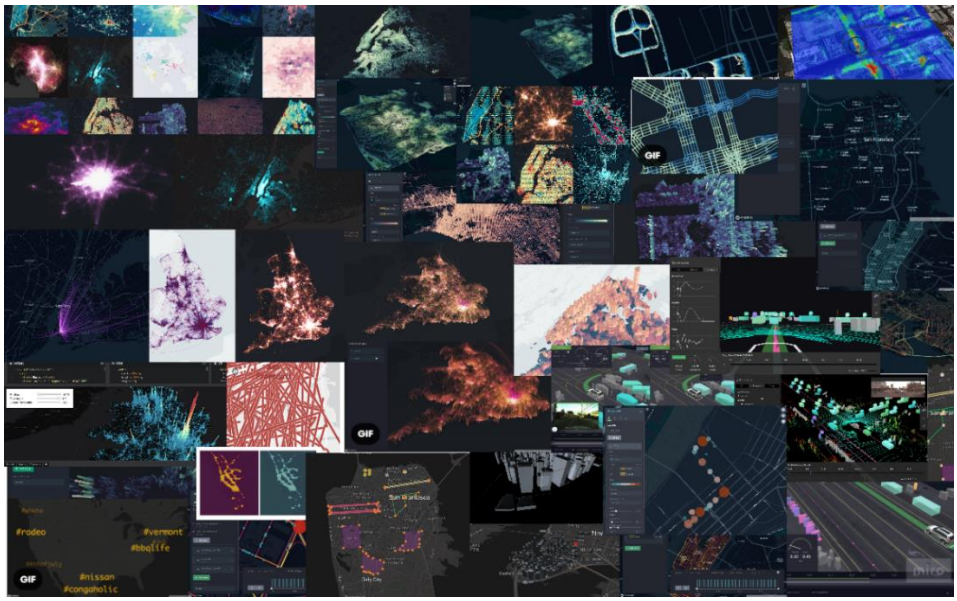
While UEB brings together a series of stories told with and through data about urban space and knowing urban space, these are also stories of and about data. They narrate how data comes to be and for what purposes, what data can do and what they make possible.

[1] Previously eng.uber.com, now accessible via uber.com/en-GB/blog/engineering/

The overarching narratives of data and urban space found in UEB echo well-known promises and assumptions of datafication and dataism (van Djick, 2014), deriving authority from the transformation of urban spaces into simplified and decontextualized metrics (Bowker & Star, 1999; Espeland & Sauder, 2007). Through large volumes of fine-grained data and machine learning models, Uber would be able to know, manage and anticipate a vastitude of territories and practices. Imposing an epistemic order into "the real world" is a pre-condition to make space algorithmically measurable, calculable, predictable. H3 is central in these regards: by partitioning the globe into multi-resolution hexagons (UEB, 2018), it allows for the transposition of transposing a "messy geographic space" into a "neat and clean mathematic space" (UEB,2019b). These reflect the ideal of objective, disembodied knowledge, a view from nowhere Haraway (1988) calls god-trick.

In a more focused analysis, UEB publications also provide an entry point to more circumscribed narratives and imagined interlocutors. The Blog speaks to and of different types of data-publics. In its timeline, there seems to be a move from internal interoperability (how different teams interact and work on similar data structures) to external publics such as other tech companies, external developers, and urban planners. As different external publics, or users, are enrolled, a discursive drive towards accessibility and usability for people with different levels of experience and technical knowledge becomes more evident.

Data visualization plays a central role as both rhetorical and epistemic devices (Mackenzie, 2015), being central to the interpretation of large amounts of multidimensional data and in the validation of knowledge produced. In the case of Uber, there is an explicit aesthetic concern with public-facing data and data-storytelling (UEB,2016). Deck.gl and Keplr.gl are central characters: not only are they foregrounded as key innovations and disseminated for external use, but they are also instruments to "create *beautiful visualization*", which do the work of providing a sense of objectivity and the promise of exciting insights (Kennedy et al., 2016).



Collage of examples data-visualizations and machine-oriented representations of the city retrieved from UEB

Uber's mission to make the "physical" world readable and intelligible is also very clearly articulated in relation to various algorithmic systems, aimed at, for example, anticipating patterns of supply and demand and triggering surge pricing (Brodsky, 2018). Among different human and non-human gazes addressed in these data narratives, self-driving vehicles (SDVs) often take the centre-stage being portrayed as complex, cognisant entities, knowing subjects in their own right. Their machinic visions are critical both because data and data visualisations are catered to them and because they themselves produce data. They also introduce a shift from the predominantly bird-eye visualisations of urban space to a ground-level view into "a city full of data" (UEB,2019c).

Interface of Uber's Autonomous Visualisation System (AVS) for visualising self-driving vehicles performance. Retrieved from Uber Engineering Blog (UEB, 2019a)

Notably absent in this community of algorithms, SDVs, engineers, developers and enthusiasts are those whose highly precarised labour is in the basis of Uber's business model and data production: platform workers moving around diverse territories in cars, motorcycles and bikes. Whereas SDVs are assigned the capacities of "seeing", "perceiving", "understanding" and "predicting", workers are abstracted and instrumentalised into the least-significant half of variables – "driver-location", driver-behaviour" – functioning as data points or sensors feeding inputs into Uber's data pipeline. In a way, UEB's data stories echo Uber own patterns of exploitation and failure to respond to workers demands or attend to their needs - workers are sites of silence in a map made by privileged knowers.

References

Dourish, P., & Gómez Cruz, E. (2018). Datafication and data fiction: Narrating data and narrating with data. Big Data & Society, 5(2). doi.org/10.1177/2053951718784083

Bowker, G.C. & Star, S.L. (1999) Sorting Things Out: Classification and Its Consequences. Cambridge, MA, USA: MIT Press (Inside Technology).

Clarke, A.. (2005). Situational Analysis: Grounded Theory After the Postmodern Turn. Thousand Oaks, CA: Sage.

Espeland, W.N. & Sauder, M. (2007) 'Rankings and Reactivity: How Public Measures Recreate Social Worlds', American Journal of Sociology, 113(1), pp. 1–40. doi:10.1086/517897.

Edwards, P.N. (2010) A vast machine: computer models, climate data, and the politics of global warming. Cambridge, Mass: MIT Press

Kennedy H., Hill R.L., Aiello G., & Allen, W. (2016) The work that visualisation conventions do. Information, Communication & Society 19(6): 715–735.

Gillespie, T. (2014) The relevance of algorithms, In Gillespie,T; Boczkowski, PT & Foot, KA (Eds.) Media technologies: Essays on communication, materiality, and society (pp. 167–193) Cambridge: MIT Press.

Haraway, D. (1988) 'Situated Knowledges: The Science Question in Feminism and the Privilege of Partial Perspective', Feminist Studies, 14(3), pp. 575–599. doi:10.2307/3178066.

Hilgartner, S. (2000) Science on Stage: Expert Advice as Public Drama, Stanford, CA: Stanford University Press

Mackenzie, A. (2015) 'The production of prediction: What does machine learning want?', European Journal of Cultural Studies, 18(4–5), pp. 429–445. doi:10.1177/1367549415577384.

Mejias, U.A. & Couldry, N. (2019) 'Datafication', Internet Policy Review, 8(4). doi:10.14763/2019.4.1428.

Pagano, D., Maalej, W. How do open source communities blog?. *Empir Software Eng* 18, 1090–1124 (2013). https://doi.org/10.1007/s10664-012-9211-2

Plantin, J.-C. (2018) 'Digital Traces in Context| Google Maps as Cartographic Infrastructure: From Participatory Mapmaking to Database Maintenance', *International Journal of Communicatio*n, 12(0), p. 18.

Poell, T., Nieborg, D. & van Dijck, J. (2019) 'Platformisation', *Internet Policy Review*, 8(4). DOI: 10.14763/2019.4.1425

Rainer, A. & Williams, A (2019). Using blog-like documents to investigate software practice: benefits, challenges, and research directions. *Journal of Software: evolution and process.* doi.org/10.1002/smr.2197

Van Dijck, J. (2014). Datafication, dataism and dataveillance: Big Data between scientific paradigm and ideology. *Surveillance & society*, *12*(2), 197-208.

UEB (2016) From Beautiful Maps to Actionable Insights: Introducing kepler.gl, Uber's Open Source Geospatial Toolbox. *Uber Engineering Blog*. Available at https://www.uber.com/en-GB/blog/keplergl/

UEB (2018) H3: Uber's Hexagonal Hierarchical Spatial Index. *Uber Engineering Blog*. Available at: https://eng.uber.com/h3/

UEB (2019a) Introducing AVS, an Open Standard for Autonomous Vehicle Visualization from Uber. Available at: https://www.uber.com/en-GB/blog/avs-autonomous-vehicle-visualization/

UEB (2019b) [Uber Open Source] Engineering Sub-City Geos for a Hyper-Local Marketplace with Uber. *Uber Engineering Blog YouTube Channel* Available at: https://www.youtube.com/watch?v=wDuKeUkNLkQ

UEB (2019c) Power On: Accelerating Uber's Self-Driving Vehicle Development with Data. *Uber Engineering Blog*. Available at: https://www.uber.com/en-GB/blog/accelerating-self-driving-vehicle-development-with-data/