Exploring Ideas of Geoprivacy Through Social Media: Facebook as a Case Study

Maximilian Bakenhus, William Mackaness
m.bakenhus@outlook.de, William.mackaness@ed.ac.uk
The School of GeoSciences, The University of Edinburgh

Summary

The general public is blissfully unaware of how social media erodes their geo-privacy. We present research exploring people's awareness of how private data can be used to generate geographical profiles. We created a web-based tool for explorative data analysis (EDAMS), which was used to expose our participants to their visualized digital Facebook history. Though users were aware of privacy issues, we found that the aspects of location and geography related to privacy (geoprivacy) were for most of our interviewees unknown. Acknowledging the role of GDPR we believe there is a strong argument in favour of increased efforts towards education (making available open-source visualisation tools), in order to raise awareness of geoprivacy issues when generating and sharing social media content.

KEYWORDS: visualisation, geoprivacy, social media, Facebook

1. The Role of Place in Revealing your Digital Personae

Activities in the geo-sciences is a massive contributor to the gathering and interpretation of data. These activities are also of huge value to surveillance capitalist companies; it is the spatial dimensions that are so crucial to modelling movement, identifying hot spots of activities, creating geographical profiles - enabling linkages that lead to commercial gain (Zuboff 2019). Knowledge of our habits of movement have even been used in planning crime. The spatial and temporal dimensions to data are what give it its value. McKenzie et al. (2016) argue that geoprivacy needs to extend well beyond the mere protection of coordinates: '*[…] everyday digital footprints such as timestamps, geosocial check-ins, and short social media messages, (e.g. tweets) are indicative of the user's location*'. Geoprivacy arises from a conglomerate of semantic layers that contain a locational signature in both space and time (eg a time stamped coordinate, an IP address, a place name). The integration of contextualising information (a time stamped activity, knowledge of the function of a building or place, a shared location with others, etc) is what enables us to give meaning to a set of abstract timestamped coordinates; thus the value lies in creating place from space.

Methods such as the profiling approaches proposed by Matz and Kosinski (2017) can be cross-referenced with self-organising maps or other AI-based spatial clustering methods, thus enabling us to gain insight into an individual's motivations, and how their behaviour differs from 'the ordinary'. In the abstract, space has no value. It is in the context of place that we can unpick a person's activities, and so in turn, their thinking. The research reported here sought to play back to the user, information garnered from their historical social media profile, and assess whether this altered their sense of security and exposure. In other words, would their raised awareness of geoprivacy lead them to behave differently in response to all those pop up terms and conditions and cookies that we so readily agree to when first we sign up to a service! The research was also interested in the efficacy of visualisation methodologies. The research leads to wider questions concerning governance and whether the human right to privacy is sufficiently protected.

2. Methodology – Visualising User's Digital Footprints

Figure 1 is an overview of the schema of our web-based explorative data analysis, EDAMS (Explorative Data Analysis of Massive Social Network Footprints) which was created in order to visualise historical Facebook content, the intention being to expose users to their own pre-existing digital footprints thus enabling them to reflect on the degree of individual exposure and the potential implications of this. The data was sourced via a request from the personal Facebook page of the participant. Each participant was able to interact with the data, and in the interests of expediency was guided by the project supervisor (Figure 3). The half-structured interviews (Thielsch and Weltzin, 2009) were used to collect

quantitative measures and responses to a questionnaire measuring their general awareness on privacy issues.

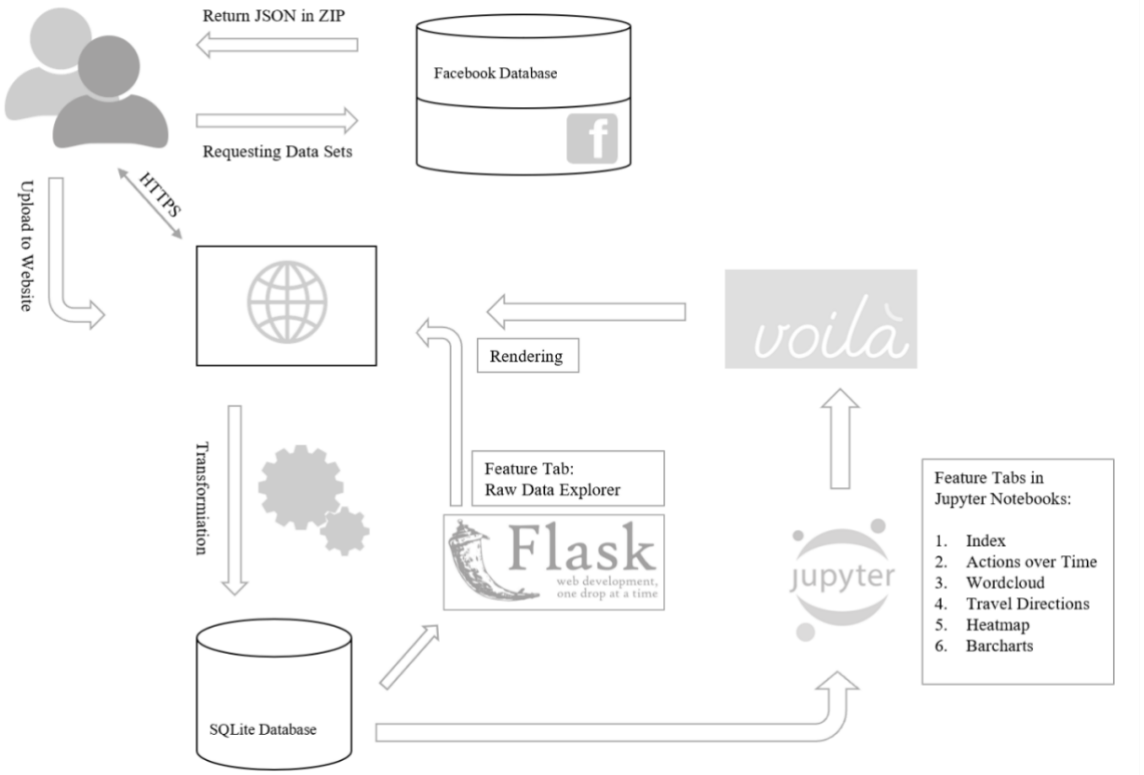

Figure 1: Schema of EDMAS

Beyond the need to have a Facebook account, users were split into one of two groups: those who had considerable exposure to GIS, and those that had none. Our participants were highly-educated young professionals (they are not seen as a representative sample of the public). The participants were sourced via convenience sampling methods (Boehnke et al., 2011) and care was taken to process the data locally, in a secure manner, encrypted and password protected. In total, eight participants were interviewed (age range of 20 - 35 years, median 28, with men and women equally represented). The participants mostly belonged to a mixed European ethnicity. The consultations were conducted remotely because of the pandemic. The interviewees were guided and audio-recorded using Microsoft Teams (Figure 2) and their conversations were fully manually transcribed. Pre-defined categories of responses were used to classify their responses in a thematic analysis with some dynamically added for unexpected responses. Examples of EDAMS visualizations are elaborated in the following sections in relation to the participants responses.

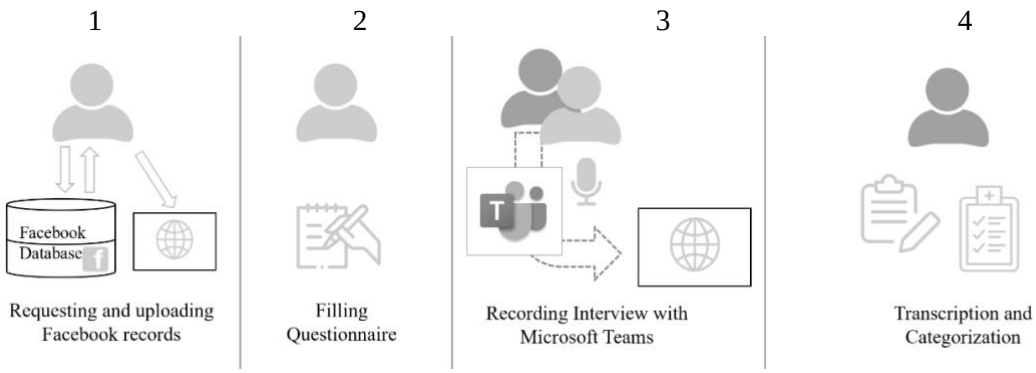

Figure 2: Schematic of the interview process

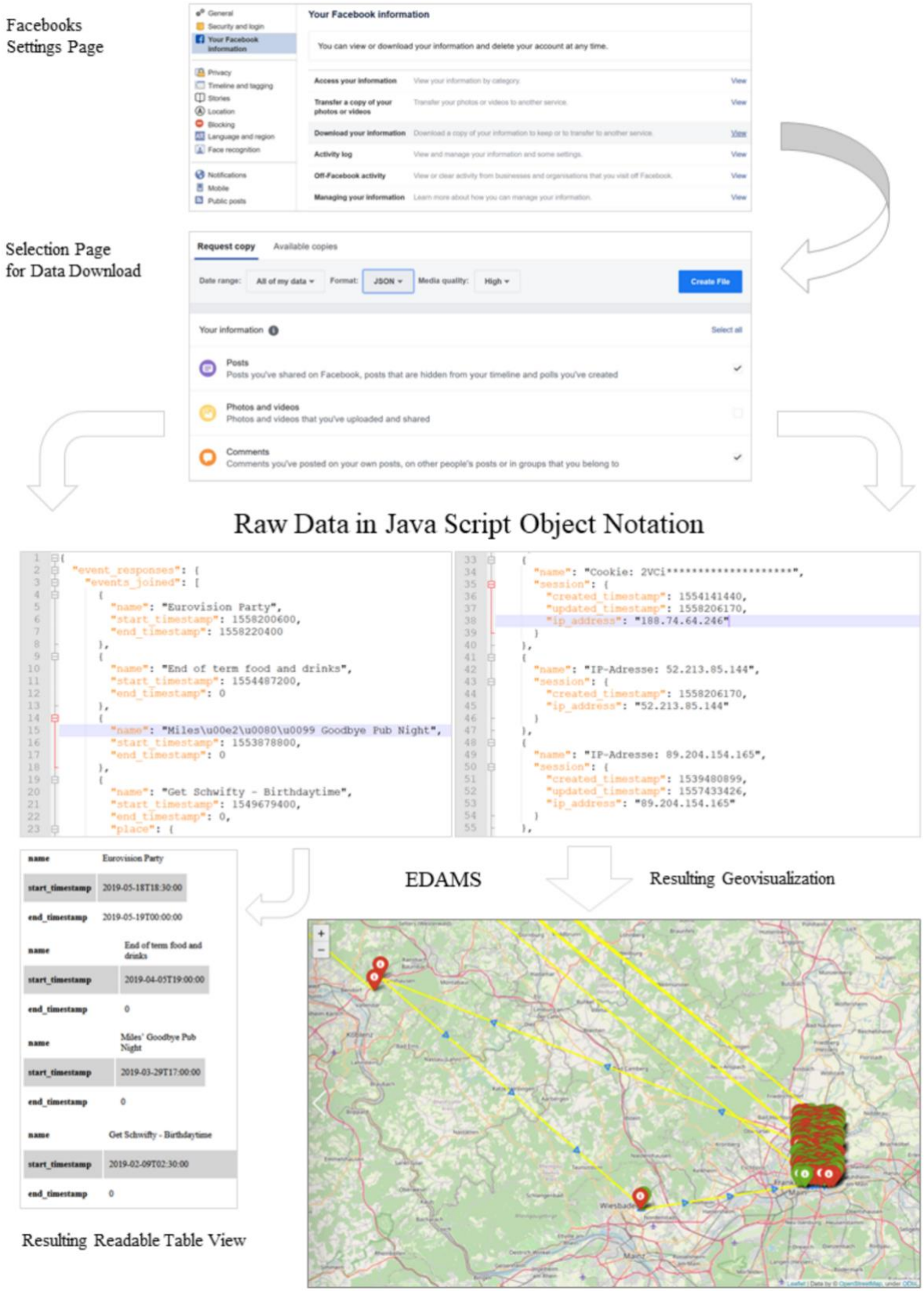

Figure 3: The process from the participant's perspective

3. Results

Various reasons were given as to why privacy was important, most of which focused around a desire to publish a controlled pictured of themselves - one that was not harmful: *'They could picture to me*

according to me in a way that I would never present myself or even see myself and perhaps […] I don't know what is worse, if they get the wrong picture or if they get the right picture' (Participant 8).

Some participants were shocked by what was revealed. In one instance, a participant responded by saying: '*I will delete Facebook again*' (Participant 5).
The tab 'Your Address Book' was a focus of concern. This revealed synchronised phone contacts, telephone numbers and mail addressees of contacts. Participants were surprised by 'Your Off Facebook Activity' which recorded those websites which had been visited outside of Facebook. Participant comments included: '*Super scary*'; '*That is bizarre*'; '*It's scary*'; '*S: Do you think this is too intrusive? P: Definitely yes yeah.*' (Participant 1,4,7).

It was interesting to observe participants assessing their tolerance to the gathering of such information: '*Storing your phones, your contact points, shadow profiles. Seems a bit weird, seems a bit a lot. Out of Facebook's range of control. […] It exceeds a point of privacy.* […] *Okay* (30 sec. quiet)*. That's a bit too much. That's something I feel a bit uncomfortable with.*' (Participant 2).

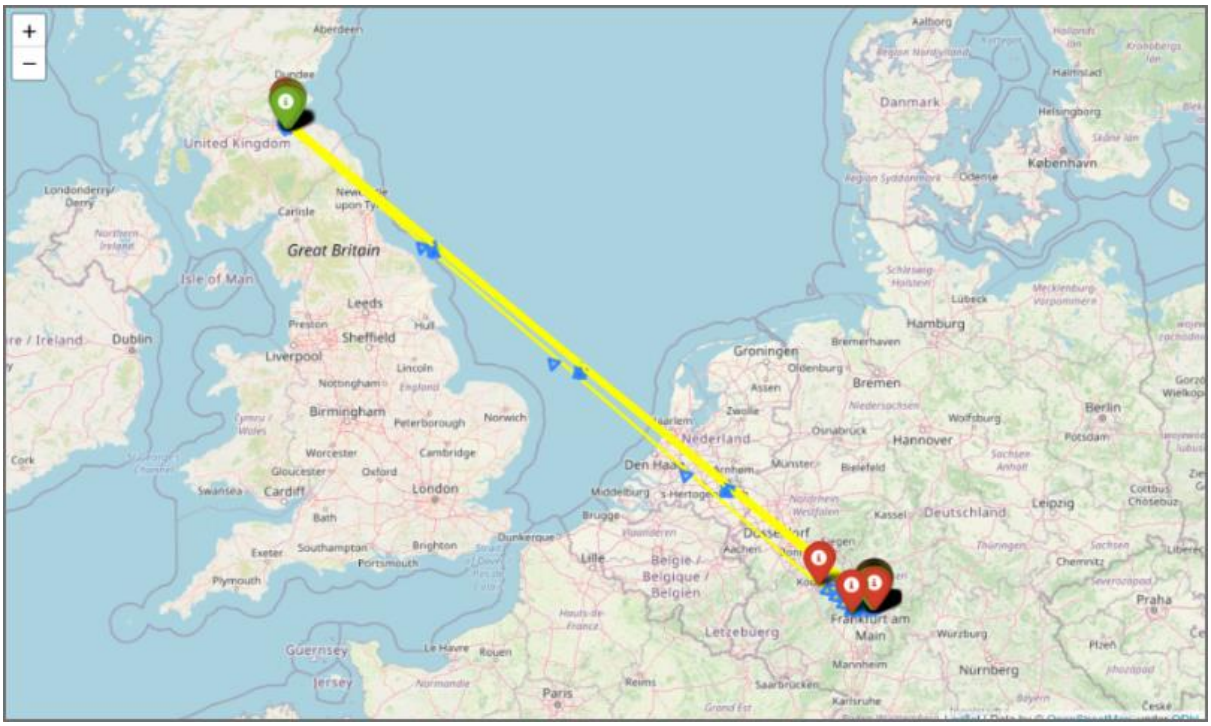

Figure 4: Identified travel directions

The feature 'Travel Directions' (Figure 4) was used to visualize the travel patterns of participants. The exploratory nature meant participants could explore the link between where they were, and the points at which friends had been added over time. Additionally, interactive linear plots were generated showing a variety of variables in relation to each other (Figure 5). Both were then combined to create clusters of activity (Figure 6).

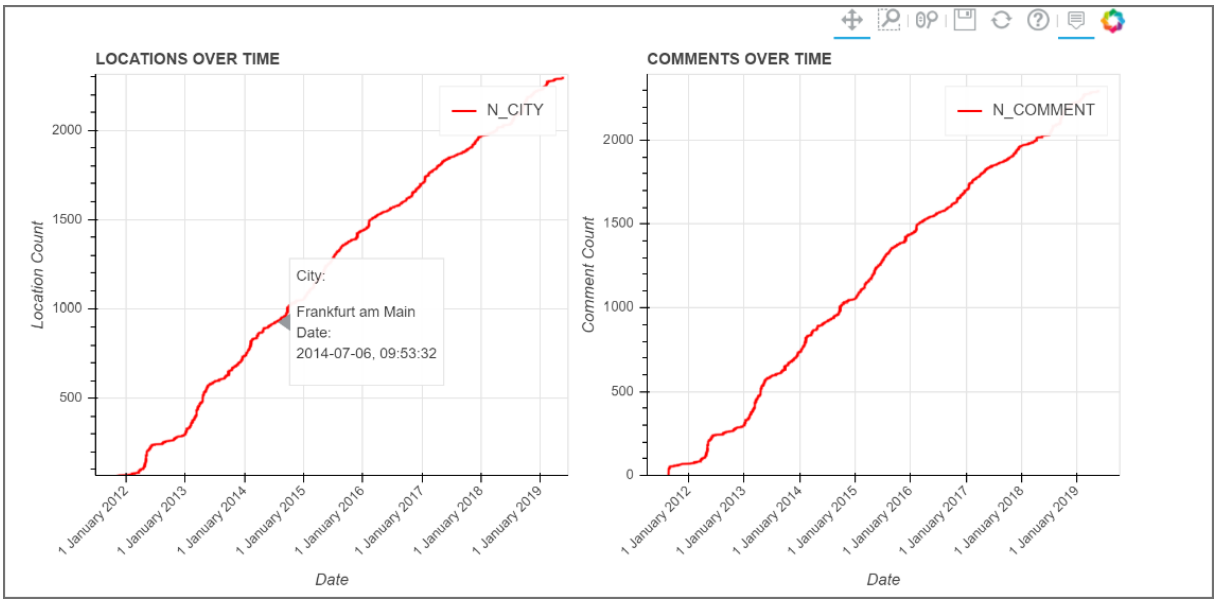

Figure 5: Sample of generated Interactive Linear Plots

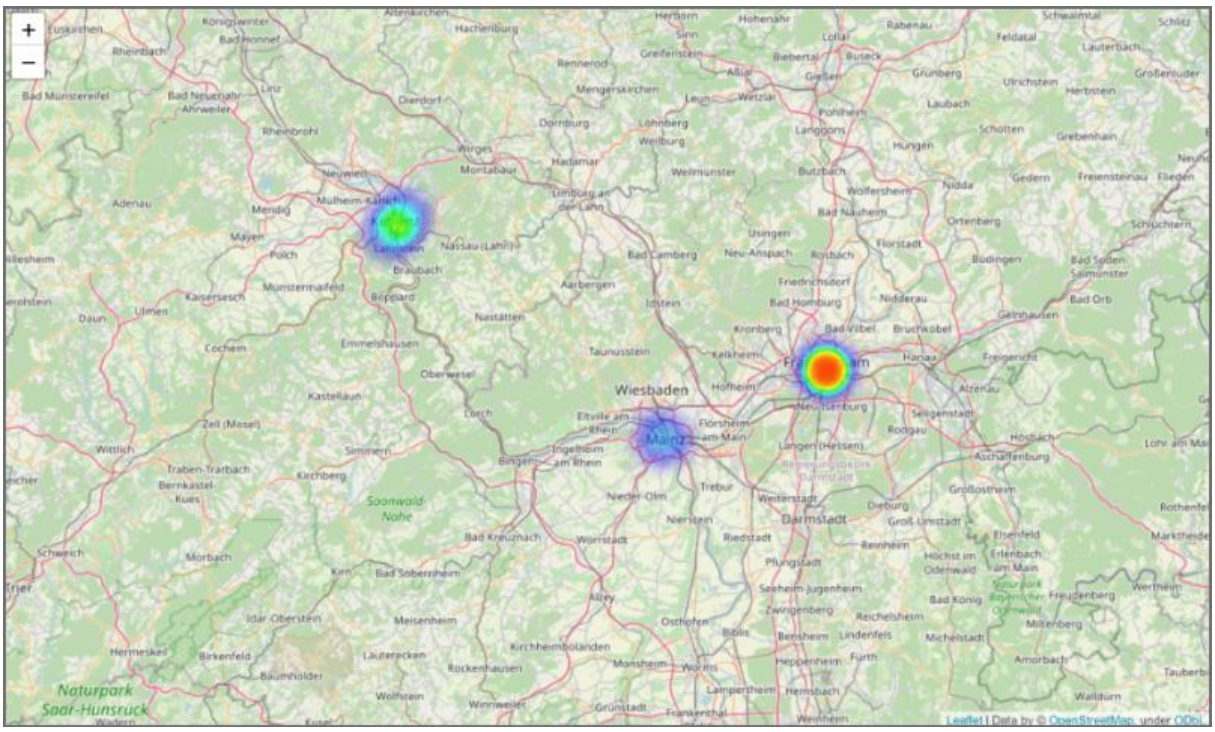

Figure 6: Clusters of Facebook activity

This tool also generated interesting responses of surprise and reflection: '*Every piece of information on the internet has a location. There is no non-locational action, therefore geoscience plays a part in every part of the discussion. Even not immediately obvious information are linked to a location in an indirect manner. The time pattern and content can be indicative or correlated to a recorded location.* […] *in other words the geoscientist has the ability to combine all of their recommendations into a space.*' (Participant 6).

Non-GI participants were surprised to see how much geography was contained within the data. Even with GPS deactivated, every action on the internet had a location and can therefore be tracked. '*S: Why do you think it is important to consider, where something has happened? P: Because you can be traced down very easily. For example, if we stay at the friend's thing, you can check who met whom, where in the world. You can see where a person has been. I mean if you take my vacation to*

Indonesia and the year after I went to Israel. They know about that. S: Why would it be tricky if you go to Israel? Well maybe if you go to Palästina afterwards. P: Yeah exactly. Also Indonesia, there is the Islam the religion. So Israel, the Israelis are very strict about that.' (Participant 8).

4. Conclusion

Our research reveals that most users are unaware of the geography or spatio-temporal attributes of their internet actions or even of the extent of their exposure. Even through these simple exploratory tools it was clear users could be made aware of their existing exposure and could better understand the potential impact on their personal lives. The experiments evoked reactions of surprise, discontent and anger from our participants, with some feeling extremely appalled by the discoveries they made. All of our participants were also able to identify a personal threshold of tolerance to their exposure, which some felt has been crossed. Some of our respondents did not feel sufficiently protected by current legal frameworks. As the scandal around Cambridge Analytica revealed, modern technologies of commercial surveillance could and will continue to pose a threat to privacy if it is not better constrained by the law.

GDPR (EU 2016) allows the use of profiling and targeted advertising provided there is user consent. But the EU is dominated by a few, predominantly US based digital companies and their dominance acts to constrain choice with users having to agree to the terms and conditions of monopolistic providers and thus to data acquisition and profiling. Frameworks for Open-Source and privacy-protecting alternatives exist, but they often lack professional support and funding. The public is somewhat straight jacketed given the expectations in both the workplace and people's personal lives.

Debates around access to the internet being 'a human right' raises questions as to the role of Governments in 1) securing digital infrastructure and 2) protecting the privacy of individuals within a digital society. The protection of an individual's whereabouts should be seen as a crucial aspect of protecting their privacy and we argue that tools such as these could be used to raise awareness among users as to how geospatial data is being used to create digital profiles of their activities and thus unpick their privacy. In the future it is hoped to include new spatial clustering techniques such as self-organising maps or other machine learning approaches. These have the potential to highlight even more of the connections that lie among these semantic layers.

5. References

Boehnke, K., Lietz, P., Schreier, M.,Wilhelm, A., 2011. Sampling The selection of cases for culturally comparative psychological research. Culture and psychology. Cross-cultural research methods in psychology, pp. 101–129 URL: https://psycnet.apa.org/record/2010-22491-005. (visited on 2020-07-23).

EU, 2016. General Data Protection Regulation (GDPR) of the European Union (EU) 2016/679. URL: https://gdpr-info.eu/. (visited on 2020-07-23).

Kerren, A., Ebert, A., Meyer, J., 2007. Human-Centered Visualization Environments: GI-Dagstuhl Research Seminar, Dagstuhl Castle, Germany, March 5-8, 2006, Revised Papers. volume 4417. Springer.

Matz, S., Kosinski, M., 2017. Psychological targeting as an effective approach to digital mass persuasion. National Acad Sciences URL: https://www.pnas.org/content/114/48/12714/. (visited on 2020-04-21).

McKenzie, G., Janowicz, K., Seidl, D., 2016. Geo-privacy beyond coordinates, in: Geospatial Data in a ChangingWorld. Springer, pp. 157–175

Thielsch, M.T.,Weltzin, S., 2009. Online-Befragungen in der Praxis. Praxis der Wirtschaftspsychologie: Themen und Fallbeispiele für Studium und Praxis, pp. 69–85.

Zuboff, S., 2019. The Age of Surveillance Capitalism

Biographies

Maximilian Bakenhus completed his MSc in GIS at The University of Edinburgh with a Distinction. He is a dedicated open-source software engineer with a deep interest in issues of geoprivacy.

William Mackaness is a senior lecturer in GIS at The University of Edinburgh in the School of GeoSciences.