Extraction of places of interest from VGI

by Junru Lin

In many application scenarios, such as urban planning, traffic guidance, travel planning, POI (place of insterest) plays a vital role in supporting decision making. Large amounts of VGI data are generated daily on OpenStreetMap and social media, including Flickr, Facebook, Instagram, etc., which provides new possibilities for POI extraction and description methods. Therefore, this research aims to develop a workflow to visualize and summarize POIs or AOIs for travel planning purposes, on a multi-scale and national-range interactive web map, based on three VGI datasets within Germany.

Datasets

Three VGI (Volunteered Geographic Information) [1] datasets are applied in this research: aggregated place name dataset (from Instagram, Twitter, and Facebook), aggregated post dataset (from Flickr, Instagram, Facebook, and Twitter), and Flickr CCBy post dataset (from Flickr).

based on NUTS (Nomenclature of Territorial Units for Statistics) boundaries and postcode. This aggregation produces five tables containing NUTS code or postcode, area name, postcount, usercount, post_per_capita, user_per_capita, and geometry of the areas on five separate levels for visualising on multi-scale maps.

Grid-based aggregation

Aggregated post dataset is aggregated based on grids in PostgreSQL. After importing spatial data from a flat file, a a geohash_reduce function is applied to aggregate latitude and longitude coordinates on 4 km-grids and 20 km-grids. Meanwhile, user_hll and post_hll as two HyperLogLog (HLL) [2] type columns are also unioned together with the geometry column aggregation using "hll_union_agg" function. The distinct number of posts and users for the new grid can be derived by calculating the cardinalities.

Density-based data clustering

For extracting POIs on a local level such as famous buildings, parks, or bridges, DBSCAN (density-based spatial clustering of applications with noise) is applied to aggregate Flickr CCBy post dataset. For the study case in Dresden, two required parameters (eps and minPts) by DBSCAN are selected as 10 meters and 30 with a good understanding of the data. Some tourist attractions within Innere Altstadt, such as Frauenkirche, Altmarkt, and Zwinger, can identified when the parameters are adequately picked.

Administrative boundaries-based aggregation

Aggregated post dataset is aggregated

Interactive visualisation

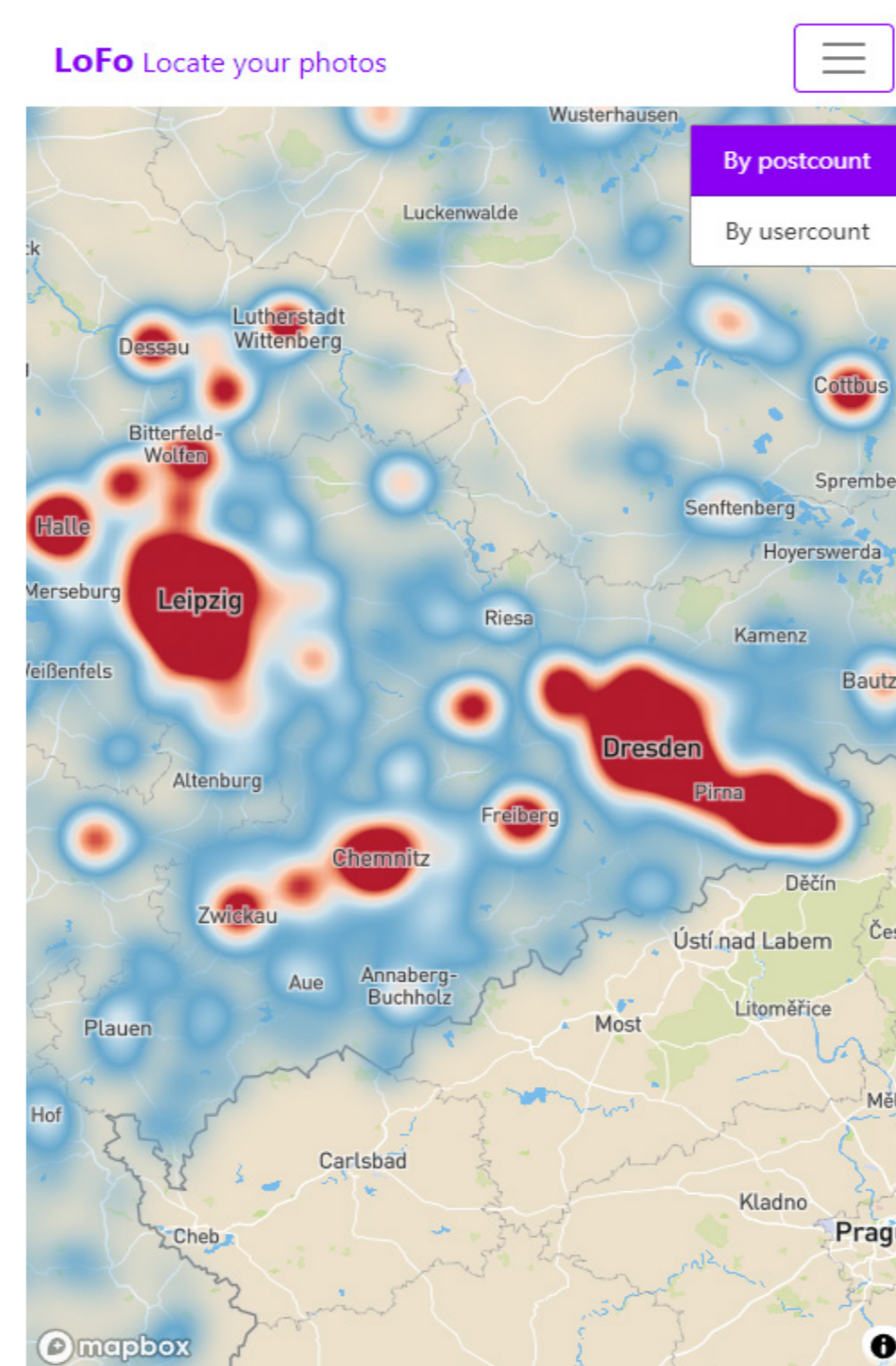A web application based on HTML5, CSS3,

Fig. 2. A Heatmap of LoFo on a mobile device displaying the region around Saxony (at zoom level 8)

and JavaScript is developed to help visualize the VGI data that has been aggregated and achieve interactions with targeted users. The application (Fig. 1) named LoFo is avaiable at: https://lofo.netlify.app. The results from grid-based aggregation are visualised as heatmaps (Fig. 2), while the results from administrative boundaries-based aggregation are illustrated as choropleth maps. The local POIs extracted in Berlin and Dresden (Fig. 1) study cases are marked on the based map as circles with variable colors and radii decided by Post Count. Possible interactions include zoom in and out, pan, mouse hover, mouse click, and type search.

Conclusion

This research developed a workflow to visualize and summarize POIs for tourist guiding purposes, on a multi-scale and national-range interactive web map application, based on three national VGI datasets using grid-based aggregation, administrative boundaries-based aggregation, and density-based clustering. By combining three different types of multi-scale maps, the application gives users a good overview of POIs within Germany. Furthermore, it allows users to explore and compare POI extraction results using different metrics for visualization at different scales.
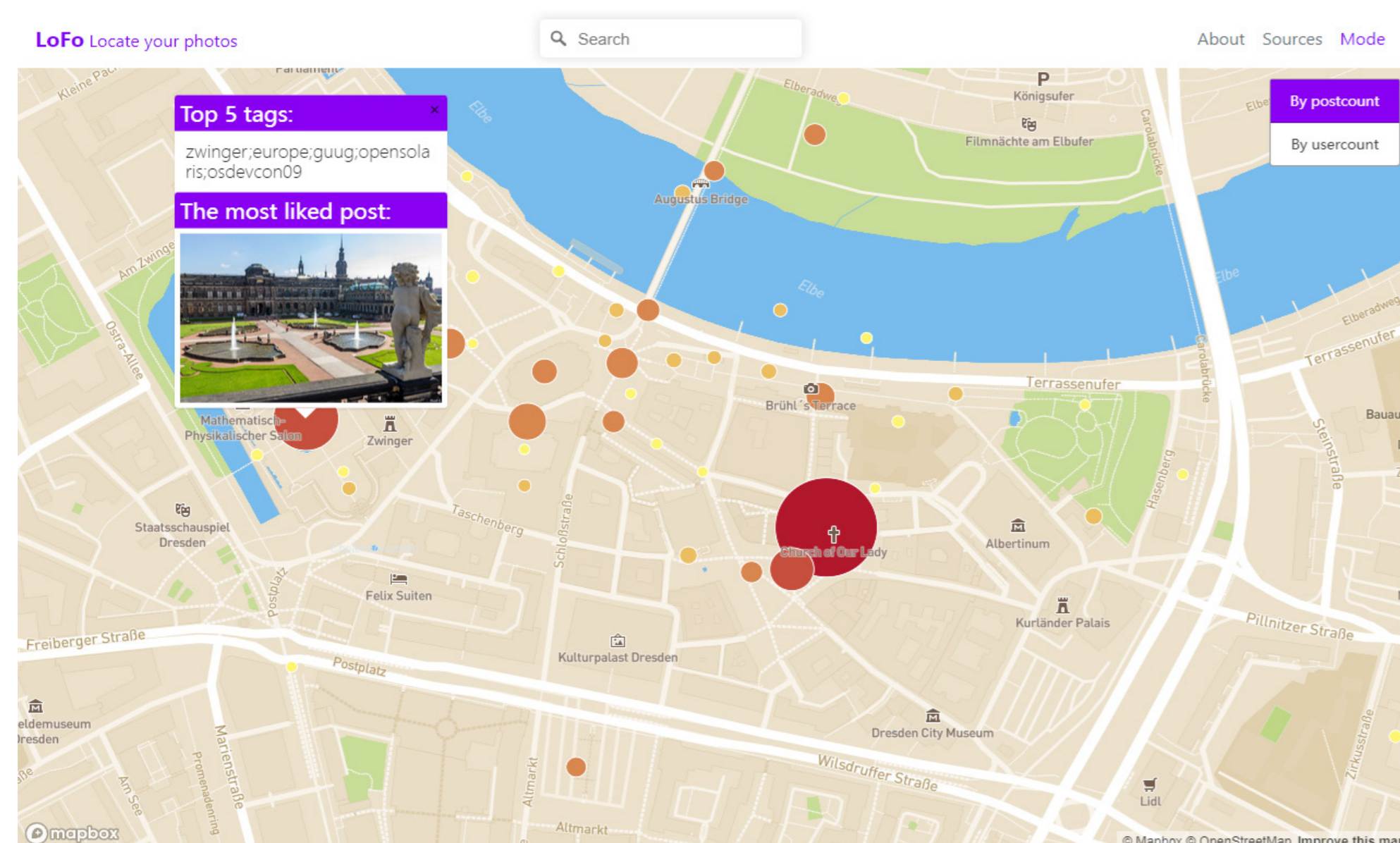

Fig. 1. The user interface of LoFo showing POIs in Dresden city center in the map display area

References

[1] Goodchild, M. F. (2007). Citizens as sensors: The world of volunteered geography. GeoJournal, 69(4), 211–221. https://doi.org/10.1007/s10708-007-9111-y

[2] Dunkel, A., Löchner, M., & Burghardt, D. (2020). Privacy-Aware Visualization of Volunteered Geographic Information (VGI) to Analyze Spatial Activity: A Benchmark Implementation. ISPRS International Journal of Geo-Information, 9(10), 607. https://doi.org/10.3390/ijgi9100607

[3] Schubert, E., Sander, J., Ester, M., Kriegel, H. P., & Xu, X. (2017). DBSCAN Revisited, Revisited: Why and How You Should (Still) Use DBSCAN. ACM Transactions on Database Systems, 42(3), 19:1-19:21. https://doi.org/10.1145/3068335

cartographymaster.eu