

# BigX: A geographical dataset visualisation tool

Stefanie Wiegand<sup>1</sup>

<sup>1</sup> Bioengineering Sciences Research Group, School of Engineering, Faculty of Engineering and Physical Sciences, University of Southampton

DOI: [10.21105/joss.02537](https://doi.org/10.21105/joss.02537)

## Software

- [Review](#) ↗
- [Repository](#) ↗
- [Archive](#) ↗

---

Editor: [Hugo Ledoux](#) ↗

## Reviewers:

- [@liberostelios](#)
- [@jvdkwast](#)

Submitted: 18 June 2020

Published: 18 November 2020

## License

Authors of papers retain copyright and release the work under a Creative Commons Attribution 4.0 International License ([CC BY 4.0](#)).

## Statement of Need

Geographic open data is widely available today not only to researchers but to the general public. Globally operating organisations such as [OpenStreetMap](#), and more localised portals such as the [Open Geography Portal](#) and [Open Development Cambodia](#) provide geographic data in a multitude of formats to download and analyse.

There are a variety of tools available to suit users with different objectives, requirements and technical abilities, all of which have different advantages and disadvantages regarding their ease of use, interoperability, extensibility, licence, cost, and other factors.

Technically inclined users can harness programming languages such as [Python](#) or [R](#) to process and analyse their data and plot it using tools and libraries such as [MatLab](#), [Matplotlib](#) or [geoplot](#) ([Bilogur, 2016](#)) and the Javascript-based library [Leaflet](#) ([Agafonkin, 2010](#)) has made the visualisation of geographic data more accessible to software developers outside the scientific community.

Closed-source commercial desktop-based tools such as [ArcGIS](#) are popular, but open-source alternatives such as [QGIS](#) and [uDig](#) exist. For paying subscribers, online services such as [ArcGIS Online](#), and [Carto](#) are available. While free services such as [QGIS Cloud](#) exist, private maps are typically premium features, which raises questions of data privacy. In order to securely share interactive maps with a small number of trusted users online while retaining control over the physical location of the data in accordance with data protection laws, users can publish their own maps using tools such as [qgis2web](#) ([Chadwin et.al., 2015](#)) to convert their maps to static web-based maps although “many elements cannot be reproduced” ([Chadwin et.al., 2015](#)) and a new static version has to be produced after each change to the original map.

Statistical data can readily be obtained from places, such as [data.gov.uk](#), but it often relates to geographic locations using identifiers. While these are typically defined in publicly available datasets, they often come in formats such as CSV and need to be linked to the geographic data before the data can be visualised.

Without the technical skills to perform preliminary data processing, cleaning, and mapping, or the money to commission such work, certain groups are left out:

- Researchers from non-technical fields such as social sciences miss out on the benefits data visualisation can give them or lose time on manually processing large datasets.
- Users and organisations with an interest in analysing their own data but without the budget to pay for professional software or services often have no other option than handing over their data to online services, but
- in places where internet connectivity is bad (such as Lower- and Middle Income Countries), transferring large datasets is not always possible and
- sensitive data, as used in medical-related fields, often comes with restrictions on how and where it can be stored, leaving users at the mercy of providers of such services and their data protection policies.

## Summary

BigX addresses these issues as follows:

- By using an open-source license, it is freely available to anyone.
- All data processing is done client-side, which makes it easily configurable to work offline or semi-offline (using an online base-map but local data). This means that:
  - no data ever leaves the machine, so it cannot be intercepted or otherwise leaked,
  - because data does not need to be transferred, response times are much faster (subject to the machine it runs on).
- BigX can be deployed using simple access control measures such as .htaccess files and allows making updates to maps in a simple text editor, thus supporting environments where GUI access may not be present.
- The markup-like definition of layers does not require the user to have programming skills — a basic understanding of the data is sufficient to generate a data layer. [Figure 1](#) shows an example for a layer markup and the result. There are various layer types and options available to suit different types of data, all of which are explained in the [manual](#).



**Figure 1:** The layer markup (left) and the generated layer on the map (right). Contains Ordnance Survey, Office of National Statistics, National Records Scotland and LPS Intellectual Property data © Crown copyright and database right 2016. Data licensed under the terms of the Open Government Licence. Ordnance Survey data covered by OS OpenData Licence. Contains HM Land Registry data © Crown copyright and database right 2018. This data is licensed under the Open Government Licence v3.0.

So far, BigX has been used to map population movements in China during the COVID-19 pandemic ([Southampton, 2020](#)) and is being used in current research to map travel distances for prosthetics patients in Cambodia as part of a wider project ([Alexander Dickinson, 2019](#)).

## Acknowledgements

While BigX was not funded directly, the author would like to thank Dr. Alex Dickinson and Prof. Andy Tatem for the opportunity to apply it to real-world research.

## References

- Agafonkin, Vladimir. (2010). *Leaflet*. GitHub. <https://github.com/Leaflet/Leaflet>
- Alexander Dickinson, et. al. (2019). *Technologies to enhance quality and access to prosthetics & orthotics: the importance of a multidisciplinary, user-centred approach*. <https://doi.org/10.5258/SOTON/P0014>
- Bilogur, A. (2016). *Geoplot*. GitHub. <https://github.com/ResidentMario/geoplot>
- Chadwin et.al., Tom. (2015). *qgis2web*. GitHub. <https://github.com/tomchadwin/qgis2web>
- Southampton, U. of. (2020). *Population mapping covid-19*. YouTube. <https://www.youtube.com/watch?v=Z8AxazmthAw>