# A Framework and Practical Guidelines for Sharing Open Benchmark Datasets in Cartographic User Research Utilizing Neuroscientific Methods

Merve Keskin*, Tong Qin**, Bing Liu***

* The Department of Geoinformatics (Z_GIS), Paris Lodron University of Salzburg (PLUS), Austria
** Geography Department, Ghent University, Belgium
*** BYD Co. Ltd, China

**Abstract.** This paper presents a structured approach to creating benchmark datasets and proposes guidelines for open-data sharing in the field of cartographic user research using neuroscientific methods including but not limited to eye tracking, EEG, and fMRI. The unique complexities introduced by geospatial data and maps make geospatial tasks fundamentally different from those encountered in the experimental psychology or cognitive visualization domains. We argue that datasets capable of addressing specific cartographic problems possess significant value and hold the potential to become benchmarks. For instance, studying the cognitive load and strategies employed by map users during various map tasks can provide valuable insights for map design and serve as benchmarks in developing complexity algorithms for cartography. We emphasize that benchmarks should be tailored to specific scientific issues rather than solely focusing on data standards. Such benchmarks not only contribute to map usability research but also play a pivotal role in developing predictive models that consider the visual attention and map use capabilities of users. Researchers across domains bear the responsibility of actively seeking concrete methods to encourage the open sharing of experimental data, complemented by high-quality metadata. By fostering the creation of benchmark datasets and promoting open-data sharing, collaboration is enhanced, cartographic research advances, and the scientific community is empowered to effectively address cartographic challenges.

**Keywords.** Eye tracking, EEG, fMRI, cartographic user research, open data benchmark dataset guidelines

Published in "Proceedings of the 18th International Conference on Location Based Services (LBS 2023)", edited by Haosheng Huang, Nico Van de Weghe and Georg Gartner, LBS 2023, 20-22 November 2023 Ghent, Belgium.

This contribution underwent single-blind peer review based on the paper.
https://doi.org/10.34726/5701 

## 1. Introduction

Research methods in cognition and neuroscience, such as eye tracking, EEG (electroencephalogram), and fMRI (Functional Magnetic Resonance Imaging), have greatly enhanced our comprehension of visual attention, cognitive processes, and problem-solving strategies. Through their direct and objective assessment of eye movements and brain activities, these methods have become increasingly prevalent in spatial cognition and map usability research. They have demonstrated their efficacy as powerful instruments in evaluating the visual attention capabilities of map users. For instance, De Cock *et al.* (2022) performed an eye tracking usability study in VR to investigate the interaction between route instruction types and building configuration on cognitive load during indoor route guidance. Keskin *et al.* (2020) studied the memorability of map landmarks by assessing the influence of a set of visual variables (i.e., location, size, shape, color), task difficulty, and expertise on recognition using eye tracking and EEG. Qin & Huang (2022) investigated the cognitive processes of map users during different map tasks (i.e., global search; distance comparison; route following and route planning), integrating both overt (visual attention through eye tracking) and covert (neural activities through EEG) perspectives, with the hypothesis that distinct eye movements and brain activities are associated with different map tasks. Dong *et al.* (2022) used fMRI to investigate the relationship between human brain activities and the spatial and temporal characteristics of public transport travel whereas Liu *et al.* (2019) investigated the influence of regular and irregular road network patterns on spatial cognition using fMRI.

Compared to standard behavioral measurements such as response time and accuracy, the added value of neuroscientific methods (*e.g.*, eye tracking, EEG, fMRI) lies in providing a deeper and more nuanced understanding of how our eyes and brains interact with maps, allowing for more targeted and effective interventions to gain insights into map users' behaviors and possibly improve map usability and map usage related capabilities. When combined with other physiological measures (*e.g.*, galvanic skin response (EDA/GSR), skin temperature, heart rate, *etc.*) or user-generated feedback such as (digital) sketch maps, they enhance the understanding of spatial cognition processes, limitations and capabilities of map users (Keskin *et al.*, 2018; Xu *et al.*, 2022). These methods can also be complemented by qualitative user feedback such as structured interviews, pre- and post-test questionnaires. In this context, it is important to consult the existing review guidelines and suggestions for mixed methods research in map usability and cartography (Roth, 2015; Štěrba *et al.*, 2014).

It is not straightforward to design a neuroscientific user experiment. Preparing all the equipment, experimental tasks and stimuli, instructing participants, collecting, preprocessing, analyzing, and making sense of the processed data requires a lot of time, expertise and effort. Despite the great endeavor, such data seems to be disposable only after one or two publications and left somewhere in the lab where it becomes another data island with limited external connectivity. In the meantime, some other researchers might be going through the same process for similar data and map reading tasks. Therefore, we must prioritize open data and open science efforts for cartographic user studies. This involves publishing raw experiment data (*e.g.,* eye tracking, EEG, fMRI, *etc.*), defining possible causes and solutions for a lack of user data benchmarks, and repurposing collected datasets as benchmarks. One possible reason for the lack of such data benchmarks for cartography is that existing benchmarks are typically used to serve computational models and they fail to replicate the similar results when it comes to complex stimuli as maps and unique tasks involved in map reading. However, if the dataset is collected to address a specific cartographic problem, and recorded during map reading, it can have significant value and the potential to become a benchmark. For instance, cognitive load and cognitive strategies of map users during different map tasks can be used as inputs in map design or can be used as benchmarks for developing complexity algorithms in cartography. Overall, benchmarks should be aimed at specific scientific issues rather than just data standards. Such open benchmark datasets for map usability research can then provide several benefits in terms of:

- Accessibility: Open benchmark datasets can increase accessibility to research findings, particularly for researchers who may not have the resources to collect their own data.
- Reproducibility & Transparency: By sharing data openly, other researchers can attempt to replicate the findings of a study and build upon them, as well as save time and resources by avoiding the repetition of similar experiments.
- Collaboration: Encouraging collaboration between researchers working on similar problems, open datasets can foster collaboration and improve the credibility, quality, and accuracy of the research, and with multi-purpose datasets, this can also happen in cross-disciplinary.
- Innovation: By providing a shared resource for testing and improving new methods and algorithms, open datasets can lead to innovative solutions, *e.g.*, *further analysis as well as exploring new methodological approaches (e.g., by including new indices and/or aggregated visualization methods, and predictive AI algorithms) can help the process of computational modeling of visual attention for this specific type of visual stimuli.*

## 1.1. Existing datasets in general *vs.* datasets for geospatial tasks

There have been several efforts to provide open datasets, such as MIT/Tuebingen's Saliency Benchmark (Kümmerer *et al.*, 2018) which is one of the largest repositories that provides a collection of eye-tracking data and saliency maps for various types of stimuli. They plan to include other eye movement benchmarking tasks with an initial focus on scanpath prediction in free-viewing and visual search settings. However, many open eye movement datasets including MIT/Tuebingen's Saliency Benchmark typically lack information for task-driven cognitive assessments. They are largely tailored for saliency analyses of general images (*e.g.*, Tliba *et al.*, 2022), and are not diverse enough to address complex cartographic visualizations and their associated tasks. Recently, there has been increasing recognition of the value of open datasets in the cartographic usability domain. For instance, **EyeTrackUAV2** was created to study how the participants' visual attention was influenced by the UAV videos under free viewing and surveillance viewing (i.e., object detection) tasks. It includes a large-scale eye tracking data (*i.e.*, gaze position, fixation duration, and saccade amplitude) from 30 participants watching 43 videos of UAVs recorded from different perspectives and at different speeds. **CartoGaze** (Keskin *et al.*, 2023) contains a comprehensive and reproducible set of large eye movement data from a controlled memorability experiment with 38 participants balanced in age and gender, along with 37 corresponding map stimuli, AOI files, task descriptions, and full procedural details of data collection and analysis framework. **GeoEye** (He *et al.*, 2023), on the other hand, constitutes 110 college-aged participants' eye movement data when free viewing 500 geospatial images, including thematic maps, remote sensing images, and street view images, which demonstrate the scientific benefits and applications in saliency prediction and map user identification. The development and provision of more varied datasets remain essential to expand the scope of cartographic research and make it possible to study new and emerging topics.

Despite EEG and fMRI research producing a wealth of data that can be used to study brain functions and activities, such data can be difficult to access and use. In recent years, there has been a growing movement to make brain imaging data more accessible through publicly available datasets.
Here we list some of the useful open datasets and repositories available for eye tracking, EEG, and fMRI research and/or co-registration of those:

- A list of all public **EEG datasets (github)**: This repository includes EEG data from a variety of tasks, including visual perception, memory, and motor control (Agarwal, 2023).
- **Donders Data Repository**: is designed to accommodate brain imaging (*e.g.*, fMRI, MEG, EEG, *etc.*) research data management workflows throughout the research life cycle. It ensures the long-

term preservation of large datasets from a variety of tasks, including language, attention, and emotion, and helps researchers adhere to the FAIR (Findable, Accessible, Interoperable, and Reusable) principles and Radboud University's research data management policy (URL 1).

- **The SJTU Emotion EEG Dataset (SEED)**: is a collection of EEG datasets provided by the BCMI laboratory for emotion recognition (URL 2).
- **Radboud Coregistration Corpus of Narrative Sentences (RaCCooNS)**: the first freely available corpus of eye-tracking-with-EEG data collected while participants read narrative sentences in Dutch. The collection is intended for studying human sentence comprehension and for evaluating the cognitive validity of computational language models (Frank & Aumeistere, 2022).

There is still a growing body of research into brain imaging research and to the best of our knowledge, no open brain imagining datasets have been published for cartographic tasks. Therefore, future research should focus on the data sharing standards that are needed for benchmarking.

## 2. Proposed structure

Standardization plays a crucial role in facilitating reliable and reproducible research outcomes. Data standards/principles such as FAIR exist but how applicable is it to address the research questions in neuroscientific user studies, yet in the cartographic domain? A key aspect of achieving standardization in neuroscientific user research is the implementation of detailed documentation for shared data, along with the development of empirically derived guidelines. Geospatial data and maps present unique complexities that differentiate them from traditional stimuli subjected to experimental psychology or cognitive visualization. Therefore, it becomes imperative to address these differences and adapt existing standardization practices to suit the specific requirements of geospatial tasks. In this context, we propose a comprehensive structure for benchmark datasets and present guidelines specifically tailored to open-data sharing in neuroscientific user research related to cartography to foster greater consistency and generalizability of research outcomes.

### 2.1. The characteristics of benchmark datasets

For the reusability of the shared datasets, information about the participants' characteristics, recording device, experimental tasks, stimuli and conditions are bare minimum, in order other words, "must-have". Here we list and detail all the dimensions that are "nice-to-have" and to be taken into consideration while reporting metadata related to datasets:

1. Controlled conditions
    - Medium/display: mobile (smartphone, tablet), laptop and/or desktop with integrated or standalone webcams
    - Performance for data collection and system specification: data quality, claimed error and calibration accuracy of the recording system and device
    - Input modality: the means by which a user interacts with a computer or other electronic device. It can include various input methods such as a keyboard, mouse, touchscreen, joystick, voice recognition, or gaze control
    - Recording devices: remote/screen-based or mobile eye trackers, webcam eye trackers, standalone EEG recording modules (*e.g.*, EEG caps & electrodes), or headsets with integrated eye tracking, VR, AR capabilities, EEG, fMRI, or other sensors
    - Extraneous variables such as lighting conditions, noise, shielded room, impedance, curiosity about the experimental procedure and equipment
2. Well-defined tasks and research questions
    - The purpose of the experiment with keywords defining the study as this is useful for others to access the datasets
    - Full procedural details of the experiment (*e.g., perhaps a standard flowchart can be prepared if not released with a research paper*)
    - Free viewing vs. task-specific
    - In labs or in real-world environments
    - Visuospatial or perceptual tasks
    - Trial tasks, orientation, and task instructions
    - Task design (*e.g.*, randomized block design, event-related design, *etc.*)
    - Task duration and total recording length as it is important due to fatigue, performance, and focus. Typically experiments should not take longer than 45 min for eye tracking and no longer than an hour for EEG or fMRI
3. Well-defined data
    - Artifact-free (if so, preprocessing steps) or raw data
    - The data quality
    - Sufficiently large data samples to ensure the generalizability of the results
    - The data format and compatibility
    - Detailed documentation including data collection, preprocessing and analysis protocols, and open codes for such analysis

- Attaching relevant scientific research if applicable and/or other relevant references
- Data specific descriptions:
  *Eye tracking specific*: dominant eye, resolution (60Hz, 120Hz, *etc*.), fixation recognition algorithm/parameters
  *EEG specific*: resolution, the number of electrodes, the type of electrodes, their spatial distribution (*e.g*., 10-20 system)
  *fMRI specific*: the number of channel head coils, repetition time (for functional/structural images), echo time (for functional/structural images), layer scan

4. Well-defined stimulus properties:
   - Screen map (mobile, laptop, desktop), animation, web-service
   - 2D, 3D or XR
   - Static, dynamic, interactive
   - Size, position, and format of images or other media used in the experiments
   - Visual or task-related manipulations if applicable
   - Experimental stimuli preparation details (*e.g.,* source, authorship, existing or new)
5. Well-defined participant characteristics
   - Sample size: we often need a large sample size for EEG and fMRI due to noise but optimization is important when using mixed methods
   - Individual characteristics of the participants (age, gender, education), additional tests to classify participants based on spatial abilities (*e.g.,* NASA TLX), if needed
   - Special concerns: Color blindness, users with other disabilities
   - Self-reports, pre- or post-test questionnaires, and structured verbal interviews
6. Well-defined metrics
   - Behavioral metrics: response time, response accuracy
   - Eye movement metrics: fixation- or saccade-related, AOI- (area of interest) specific metrics, scanpaths, heatmaps
   - EEG: time-domain: Event-Related Potentials (ERP) (*e.g.,* P300); frequency-domain: Power Spectral Density (PSD) (*e.g.,* hemispheric differences: Frontal Alpha Asymmetry (FAA)); time frequency-domain: Event-Related Synchronization& Desynchronization (ERS/ERD)
   - fMRI: Blood Oxygen Level Dependent (BOLD) Signal (*e.g*., changes in whole-brain or AOI), Functional Connectivity
7. Ethics
   - Asking local ethics committees for permission if needed
   - Adhering to ethical standards, including obtaining informed consent from participants and protecting their privacy.

- Anonymization of participants' data

With above metadata being listed, we would like to emphasize that it is important to make pre-processing and analysis steps clear, available and somewhat transferrable to other use-cases. Hence, linking shared datasets with published scientific work is the ideal approach. For instance, the open EEG dataset collection published by Popov *et al.* (2018) portrays a good example of sharing the metadata, as well as all raw data, metrics, and analysis scripts necessary to reproduce the results of the original study.

### 2.2. The guidelines for sharing data openly

- Accessibility: The data should be stored somewhere accessible to a wide audience.
- Stability: The data provider should make sure the data is taken care of and always accessible (at least within certain years).
- Safety: The data provider should make sure there is no malware that might attack its users.

## 3. Conclusion

Open data and open science for cartographic user studies are essential to improving the quality and accessibility of cartographic research. We must prioritize these initiatives in order to ensure that our research is as rigorous and impactful as possible. Benchmarks that are useful for addressing specific cartographic issues not only create value for map usability research but also are essential parts of the development of predictive models considering the visual attention and map use capabilities of map users. The biggest responsibility for us researchers in all domains is to seek concrete ways to encourage ourselves and the community to share experimental data openly with high-quality metadata.

## 4. Acknowledgement

We would like to express our sincere gratitude to the cartographic community (i.e., members of the Commissions of International Cartographic Association (ICA) on Geovisuzaliation, User Experience, and Cognitive Visualization) for various conversations and especially to Dr. Amy Griffin who initiated the Remote Eye tracking Benchmarking exercise.

## References


Agarwal, M. (2023, January 1). EEG-Datasets. GitHub. Retrieved from https://github.com/meagmohit/EEG-Datasets. Last accessed on 12.7.2023

De Cock, L., Van de Weghe, N., Ooms, K., Saenen, I., Van Kets, N., Van Wallendael, G., ... & De Maeyer, P. (2022). Linking the cognitive load induced by route instruction types and building configuration during indoor route guidance, a usability study in VR. International Journal of Geographical Information Science, 36(10), 1978-2008.

Dong, W., Qin, T., Yang, T., Liao, H., Liu, B., Meng, L., & Liu, Y. (2022). Wayfinding behavior and spatial knowledge acquisition: Are they the same in virtual reality and in real-world environments?. Annals of the American Association of Geographers, 112(1), 226-246.

Frank, S. L., & Aumeistere, A. (2022, June 1). An eye-tracking-with-EEG Coregistration Corpus of Narrative Sentences. https://doi.org/10.31234/osf.io/j5fgd Last accessed on 12.7.2023

He, B., Dong, W., Liao, H., Ying, Q., Shi, B., Liu, J., & Wang, Y. (2023). A geospatial image based eye movement dataset for cartography and GIS. Cartography and Geographic Information Science, 50(1), 96-111, https://doi.org/10.1080/15230406.2022.2153172

Keskin, M., Ooms, K., Dogru, A. O., & De Maeyer, P. (2018). Digital sketch maps and eye tracking statistics as instruments to obtain insights into spatial cognition. Journal of Eye Movement Research, 11(3). https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7733313/

Keskin, M., Ooms, K., Dogru, A. O., & De Maeyer, P. (2020). Exploring the cognitive load of expert and novice map users using EEG and eye tracking. ISPRS International Journal of Geo-Information, 9(7), 429. https://www.mdpi.com/2220-9964/9/7/429

Keskin, M., Krassanakis, V., & Çöltekin, A. (2023). Visual Attention and Recognition Differences Based on Expertise in a Map Reading and Memorability Study. ISPRS International Journal of Geo-Information, 12(1), 21. https://doi.org/10.3390/ijgi12010021

Kümmerer, M., Wallis, T. S. A., & Bethge, M. (2018). Saliency Benchmarking Made Easy: Separating Models, Maps and Metrics. In V. Ferrari, M. Hebert, C. Sminchisescu, & Y. Weiss (Eds.), Computer Vision – ECCV 2018 (pp. 798-814). Lecture Notes in Computer Science. Springer International Publishing. https://saliency.tuebingen.ai/datasets.html

Liu, B., Dong, W., Zhu, L., Liu, H., & Meng, L. (2019). Using fMRI to explore the influence of road network patterns on geospatial cognition. Proc. Int. Cartogr. Assoc, 2, 75. https://pdfs.semanticscholar.org/c297/bb1631f9ff86477fe1ca7678cbd879beb85d.pdf

Popov, T., Oostenveld, R., & Schoffelen, J. M. (2018). Characterization of auditory steady stated brain responses in time, frequency and space.Version 2. Radboud University. (dataset). https://doi.org/10.34973/fkgz-8d22

Qin, T., & Huang, H. (2022). Visual attention and neuro-cognitive processes in map use. Abstracts of the ICA, 5, 1-2. https://ica-abs.copernicus.org/articles/5/114/2022/ica-abs-5-114-2022.pdf

Qin, T., Dong, W., & Huang, H. (2023). Perceptions of space and time of public transport travel associated with human brain activities: A case study of bus travel in Beijing. Computers, Environment and Urban Systems, 99, 101919. https://www.sciencedirect.com/science/article/pii/S0198971522001636?via%3Dihub

Roth, R. E. (2015, August). Challenges for Human Subjects Research in Cartography. In Proceedings of the ICA Workshop on Envisioning the Future of Cartographic Research, Curitiba, Brazil (Vol. 21). https://cogvis.icaci.org/pdf/icc2015/Roth_2015_ICCprecon_Final.pdf

Štěrba, Z., Šašinka, Č., Stachoň, Z., Kubíček, P., & Tamm, S. (2014). Mixed research design in cartography: a combination of qualitative and quantitative approaches. Kartographische Nachrichten, 64(5), 262-269.

Tliba, M., Kerkouri, M. A., Ghariba, B., Chetouani, A., Çöltekin, A., Shehata, M. S., & Bruno, A. (2022). Satsal: A multi-level self-attention based architecture for visual saliency prediction. IEEE Access, 10, 20701-20713.

URL 1. Donders Repository (n.d.) Retrieved from https://data.donders.ru.nl/collections/published?1. Last accessed on 12.7.2023

URL 2. SEED Dataset Retrieved from https://bcmi.sjtu.edu.cn/home/seed. Last accessed on 12.7.2023