Journal of Visualization and Interaction
JoVI review snapshot created at Jan 27, 2026 08:55.

Issue #12 (open): [REVIEW] GLADOS: Graph Layout Algorithm Benchmark Datasets for Open Science â€“ GD Benchmark Sets

@gdrawing-anon-reviewer on Mar 20, 2025 13:49: [opened]

@gdrawing-anon-reviewer on Mar 20, 2025 13:49:

Conflicts of interest

- ☒ I declare that I have no known conflicts of interest with the authors.

Reviewed version

f41e831

Review

This paper presents GLADOS, an extensive webpage and archive for a large collection of popular graph drawing datasets. Additionally, the authors provide more information in the form of e.g. documentation, code and unified graph data format. The paper itself describes the problem of reproducibility in science and vanishing datasets, and offers a solution of an overarching taxonomy of datasets and their features. GLADOS is easy enough to explore so that any experienced an inexperienced user should be able to view existing datasets, where they have been used, their classifications, and their attributes.

The paper documents a massive collaborative effort to not just collect but also categorize the most popular graph drawing datasets used in recent years. Moreover, the authors put in a lot of effort to acquire datasets that were lost by either reproducing them (with the instructions given in their respective papers) or by contacting the author for each lost dataset. There are a few limitations, which have also been identified by the authors themselves, such as the limited number of venues and the large bias towards datasets used in recent years. However, the paper is strong enough to hold substantial merit irrespective of these limitations. The paper therefore does not need major revising but could provide a few more arguments on their limitations.

Strengths

(S1) Data collection By far, the data collection is the strongest aspect of the paper. Not only is the accumulated collection of datasets incredibly vast, it has also been categorized and described to help researchers with making choices on what datasets to use for their future works. This collection will make future experiments with newly proposed graph drawing algorithms much more comparable as it allows researchers to have access to standard benchmark datasets.

(S2) Extensive Data Processing and Gathering Next to collecting the publicly available datasets from various sources, the authors also attempted to recreate datasets of papers according to the papers' specifications. Furthermore, the authors have made all collected graphs in the datasets available in three commonly-used formats.

(S3) Archive As the authors themselves noted, direct access to datasets and even papers occasionally vanish, making it impossible to reproduce research and compare newer techniques with existing ones. To counter this problem, the authors establish a long-term archive on the Open Science Framework, ensuring the preservation and accessibility of the collected datasets.

Weakness

(W1) Maintenance and Future Datasets While the creation of a long-term archive ensures its existence for a long period of time, the maintenance of the webpage and the archive has not been described in the paper. I am particularly interested in whether the authors considered a roadmap or some sort of protocol for including newer datasets that will be released in future works, i.e. how can these datasets be added to the collection?

(W2) The information on the popularity of the datasets is not integrated in the tag-filtered navigation filtering system. In Section 4.1 the authors do mention for each dataset the original paper and example papers in which the dataset is used; on top of that, Figure 8 displays the frequency of a dataset being used in a paper over time. However, it would be useful to have information on the popularity readily available in the filtering system to quickly observe which datasets are more commonly used.

(W3) As mentioned by the authors themselves, the selection of venues limited the total number of papers and datasets included in the collection. I know for a fact that there are a handful of GD papers posted in other venues in which other datasets were reused and created for the purpose of GD algorithms e.g. see CoRe-GD of ICLR (https://openreview.net/forum?id=vtyasLn4RM). While the aforementioned example paper could (probably) still be found by the collection process described in Section 3, I fear that other papers that focus more on the algorithmic part and were therefore published in other venues might have been overlooked. While the authors mention that the collection is not comprehensive but a starting point, they do not mention how they or others might extend this work or could continue the upkeep of collecting datasets.

(W4) The collection process simply mentions that only papers were included that were cited more than 4 times in the last 6 years of graph drawing. While I do understand the need for setting some inclusion criterion, and I might be nit-picky here, why was this criterion set to 4 citations? I would imagine that this, for instance, would include less relevant papers that were only cited 5 times in the last 20 years and would exclude newly released papers that might be more relevant.

Minor notes: Section 3, 2nd paragraph, 2nd to last sentence: 'redistirbution' Section 3.1, 1st paragraph, last sentence: 'replicablility'

Reviewer note: I have seen similar issues arise in the field of Dimensionality Reduction (DR). To my knowledge, a work similar to GLADOS for DR does not exist. The authors could raise this as a potential future work for the DR community to take inspiration from GLADOS.

Final Recommendation I believe this paper is a strong and especially useful contribution to the field of Graph Drawing. With some minor modifications to the webpage and some additional clarity on the upkeep, I believe this paper will be a valuable contribution to the Journal of Visualization.

Openness/Transparency

Openness/Transparency As described in the Strengths section of my review, the paper at hand is incredibly open and transparent. The data processing has been explained in detail and all of the datasets have been made publicly available in multiple formats.

Submission categories

- ☐ Registered Report
- ☐ Replication Study
- ☐ Empirical Research - Quantitative
- ☐ Empirical Research - Qualitative
- ☐ Systems or design research
- ☐ Commentary
- ☒ Systematic Literature Review

Suggested outcome

Minor revisions: this paper requires some smaller changes, after which I am confident I would be

able to endorse it.

Requested changes

- Additional Information on the Upkeep: As mentioned in (W1), I believe that it would be worthwhile describing how other researchers could contribute to the collection in the future, as well as how GLADOS will be maintained.
- Improve the Functionality of the Filtering System: As mentioned in (W2), the inclusion of popularity of a dataset will help researchers in finding the benchmark dataset(s) that are more commonly used for the classes described in the filtering system.

ORCID

No response

@gdrawing-anon-reviewer on Sep 09, 2025 09:04:

Round 2 Decision â€¢ Endorse:* I am willing to endorse this paper, with at most minor copyediting.

The authors addressed the issue on maintenance by providing a template for contributors to request datasets to be added to the collection. The authors also provided a suitable explanation on the in/exclusion criteria and fixed all the small. The only thing that has not been addressed was the improvement of the filtering system in the GLADOS interface, which was raised as an improvement point by multiple reviewers, though I admit that it is a low priority issue.
