



Issue #6 (closed): [REVIEW] Open Materials Review

@chatchavan on
Sep 01, 2024 10:26: [opened]

@chatchavan on
Sep 01, 2024 10:26:

Instructions for Open Practices Chair

- If there is anything in the list below that you are unsure of, you can reply to this issue (after it has been created) and tag the responsible organizers. Organizers' Github IDs are listed on [People page](#).
- If anything in the below list is missing (e.g. required paper sections, supplemental materials, etc), [mention](#) the authors on this issue and ask them to add those materials. Authors' Github IDs are listed at the top of README.md.
- Once authors have provided any missing material (or if no material is missing) and all requirements are met, close this issue.

Open Materials Checklist

- ☐ Paper license is CC-BY
- ☐ The "Research materials" section in the paper exists.
- ☐ Links to supplements and materials are listed in at least the following locations:
 - ☐ Abstract (according to the [structured abstract requirement](#))
 - ☐ "Research materials" section of the paper.
- ☐ All links to repositories and materials in the paper can be reached.
- ☐ All materials are in repositories that meet the [FAIR principles](#) (e.g. OSF, Zenodo, etc). Non-long-term hosting solutions, such as Github or self-hosted websites, are acceptable only if they are also archived on a long term repository like OSF or Zenodo. One easy way to archive + DOI an existing Github repo is with Zenodo; see [here](#)
- ☐ Check that all other requirements under the [transparency section of the author guide](#) are fulfilled.

@chatchavan on
Sep 01, 2024 10:27:

[referenced from [#Ensure all appropriate supplemental materials are included: e.g. experimental data, experimental analysis code, source code for artifacts, etc. For a full list of requirements see the transparency section of the author guide](#)]

@abhsarma on
Sep 06, 2024 14:22:

@vuluongj20

Currently, the OSF repository that is included is not anonymously accessible by the public. To address this, please create an OSF registration (see [instructions](#)). Once this is complete, please update the URL in the paper and post it here as well. I can then complete the Open Materials Review.

@vuluongj20 on
Sep 12, 2024 06:13:

Hi @abhsarma! I created a new registration on OSF but couldn't get past the final approval step "there seems to be a server issue. I'll try this again tomorrow and contact them if it fails again.



OSFHOME ▼

My Projects

Search

Support

Donate

Vu Luong ▼

Unable to resolve

OSF was unable to resolve your request. If this issue persists, please report it to support@osf.io.

@vuluongj20 on
Sep 12, 2024 18:34:

@abhsarma The server issue has been resolved! The public registration is now available at <https://osf.io/5zwe9>. I've updated the paper with the new URL.

@abhsarma on
Sep 18, 2024 16:58:

thanks!

@abhsarma on
Sep 28, 2024 01:54:

@vuluongj20 please find the open practices review below:

License: Please add a CC-BY or CC-BY-SA license to your [paper](#). Please also add a license to your OSF repository (Creative Commons or an [OSI-approved license](#)).

Abstract: Paper lacks an abstract (thereby no links to supplements and materials are present in the abstract. If you do not wish the abstract to be at the top, you may add it as a section at the end of the article. (When we publish your article, we will need an abstract as a part of the metadata for registering a DOI.)

Missing DOIs: Please add DOIs for the following: [2](#), [8](#), [10](#), [20](#), [26](#); [20] is linked but not to the actual publication link. I have provided the DOIs for these here (see the links above).

Dataset: Authors are using a public dataset, including citation. However, the link within the citation (<http://lcg-www.uia.ac.be/conll2003/ner/>) is broken. Please include a link to the dataset which is accessible. Moreover, this dataset consists of annotations (publicly available) on the original text (which has limited access due to copyright). The homepage of the dataset itself is also not persistent. I'd recommend the following suggestions: If the authors did not use the copyrighted part in their work, they can include the public part in their OSF repository together with a file that provides appropriate credit to the data source. If the authors used the copyrighted part, they can quote the build instructions from the dataset homepage in one of their README files either in their paper or on OSF.

@vuluongj20 on

Oct 09, 2024 03:15: Hi @abhsarma, thank you for the detailed feedback and instructions!

License: I’ve added a CC-BY license to the OSF [project](#) and [registration](#). The paper already contains a License section at the end with a CC-BY hyperlink. Is that sufficient or is there something else I need to do?

License

This work is licensed under a [Creative Commons Attribution 4.0 International License](#).

Abstract: The introduction (everything before “Probabilistic Graphical Models”) was meant to double as an abstract. I’ve added a subheading to make that more clear.

Abstract

“The UK” is a country, but “The UK Department of Transport” is an organization within that country. In a named entity recognition (NER) task, where we want to label each word with a name tag (organization/person/location/other/not a name), how can a computer model know one from the other?

In this case, contextual information is key. The fact that “The UK” is followed by “Department” is compelling evidence that when taken together the phrase refers to an

DOIs: I’ve added the DOIs to the References section. Thank you for providing them!

Dataset: I’ve added a functioning download link to the Research Materials Statement section at the bottom of the article. It's the same

source used by [Tensorflow Data](#). We were not aware of the copyright issues. We used the annotations for both training and validation. I’ve removed the data files from the OSF project and added instructions for how to download the data in a root-level [README file](#) and the project’s [Wiki](#). Since OSF doesn’t allow modifying files in a registration, I’ve withdrawn [the outdated registration](#), created [a new one](#), and updated the corresponding link in the article.

Data: all three models were trained and evaluated on the [CoNLL-2003 dataset](#), available for download [here](#).

@abhsarma on Oct 16, 2024 18:33: Hi @vuluongj20 Thanks for addressing the issues! Everything looks good to me.

@abhsarma on Nov 03, 2024 13:14: [closed]