

Reproducibility Review of: Spatio-Temporal Knowledge Graph from Unstructured Texts: A Multi-Scale Approach for Food Security Monitoring



Item	Value
Title	Spatio-Temporal Knowledge Graph from Unstructured Texts: A Multi-Scale Approach for Food Security Monitoring
Authors	Charles Abdoulaye Ngom, Landy Rajaonarivo, Maguelonne Teisseire, Sarah Valentin
Ref. paper	Ngom, C. A., Rajaonarivo, L., Teisseire, M., and Valentin, S.: Spatio-Temporal Knowledge Graph from Unstructured Texts: A Multi-Scale Approach for Food Security Monitoring, AGILE GIScience Ser., 7, 16, https://doi.org/10.5194/agile-giss-7-16-2026 , 2026.
Codechecker(s)	Sophie Teichmann
Date of Check	2026-04-10
Summary	Partial reproduction
Repository	https://github.com/reproducible-agile/STKG-FS
Ref. certificate	10.17605/OSF.IO/BV4MK

Table 1: Reproduction metadata

Output

Category	Count	Percentage
Location and Date correct	928	81.91%
Location corrected only	166	14.65%
Date corrected only	27	2.38%
Both correct	12	1.06%
Total	1133	100.00%

Comment

Recreation of Table 2 with the script
Triplet_Formation/Triplet_Algo.ipynb

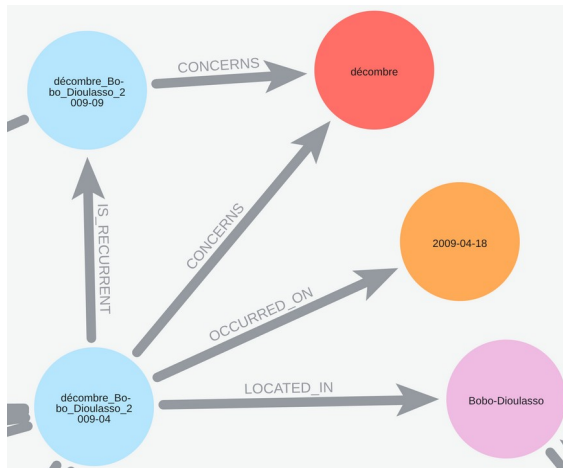


Figure 3 was generated with Query 1 in
neo4j Desktop.

Figure 5 was generated using Query 4.

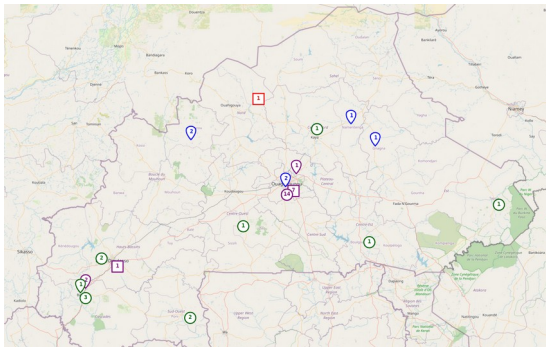


Table 2: Summary of output files generated

Summary

The authors provided a well documented Git Hub repository with detailed instructions in the README. Due to the high computational demand, for the training of the model, the reproductions in only partial. The individual steps are very well documented and I was able to recreate the results from step 3 onward. Overall, I would consider the reproduction a success.

Reproducibility reviewer notes

I performed the reproduction on an Ubuntu 22.04 machine with python3.10.20. The work requires quite some computational resources. The Step 2 in the repository the fine tuning was skipped, due to the required of computational resources. I appreciate them provide intermediate results for each step, once you get stuck.

Installation prerequisites and computational environment

The installation steps are documented are documented in the README.md. For newer python versions the installation of pyrobuf fails. This is as well handled in the README.md, with the annotation that pyrobuf is not essential for the pipeline. With some minor adjustments to the requirements, the installation via pip worked flawless. I did not need to do the HeideTime installation as described, since I was able to install it via pip.

Data preparation

The authors provided the file 'raw_articles.csv' upon request. With that and some minor adjustments to the paths, the step 1, the data preprocessing works nicely.

Running the code

I did not test the scripts needed in 'Step 2: Fine-tune CamemBERT for Spatial NER' due to intensive use of computational resources. The model and the corresponding dataset have the DOIs [10.57967/hf/7766](https://doi.org/10.57967/hf/7766) and [10.57967/hf/7767](https://doi.org/10.57967/hf/7767), respectively. All the follow-up steps, I was able to execute. The instructions in the README.md were very instructive and detailed. Step 2 includes the model

Outputs and results

All the expected output files were generated in each of the steps 3: Evaluate Spatial Entity Detection to step 8:Generate Interactive Risk Map. I will show selected figures and tables from the manuscript.

Table 2

The results of the reproduced table above and the original below align perfectly. The table is generated 'Triplet_Formation/Triplet_Algo.ipynb'.

Category	Count	Percentage
Location and Date correct	928	81.91%
Location corrected only	166	14.65%
Date corrected only	27	2.38%
Both correct	12	1.06%
Total	1133	100.00%

Category	Count	Percentage
Location and Date correct	928	81.91%
Location corrected only	166	14.65%
Date corrected only	27	2.38%
Both corrected	12	1.06%
Total	1,133	100%

Figure 3

Left is the recreated version of Figure 3 a-c and right the original version of the Figure 3 a-c. Since we have a reduced dataset, it is expected to get reduced results. Qualitatively, all three relationship to types are available. The recreation was done with in the explore mode with Neo4j Desktop.

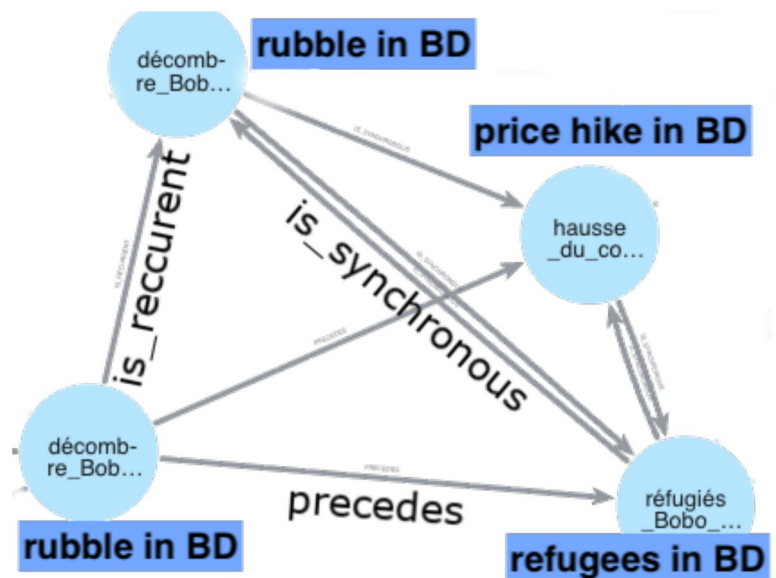
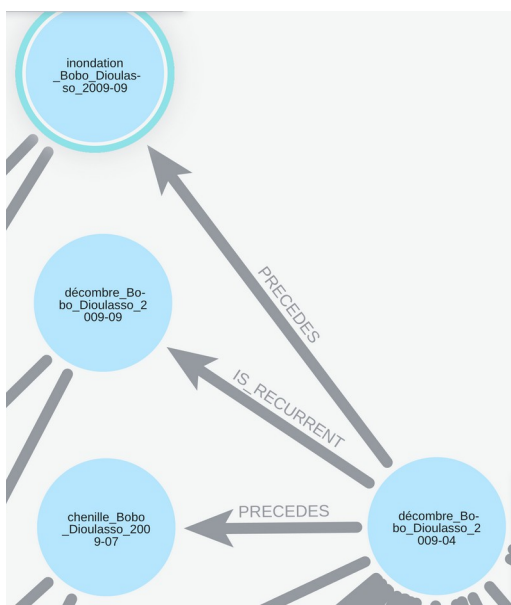
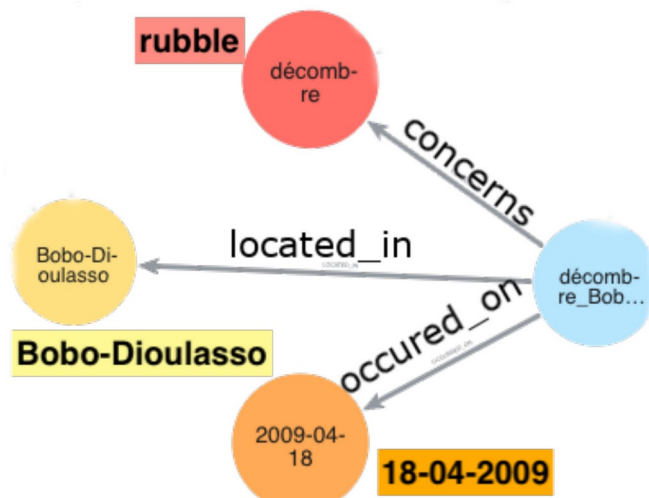
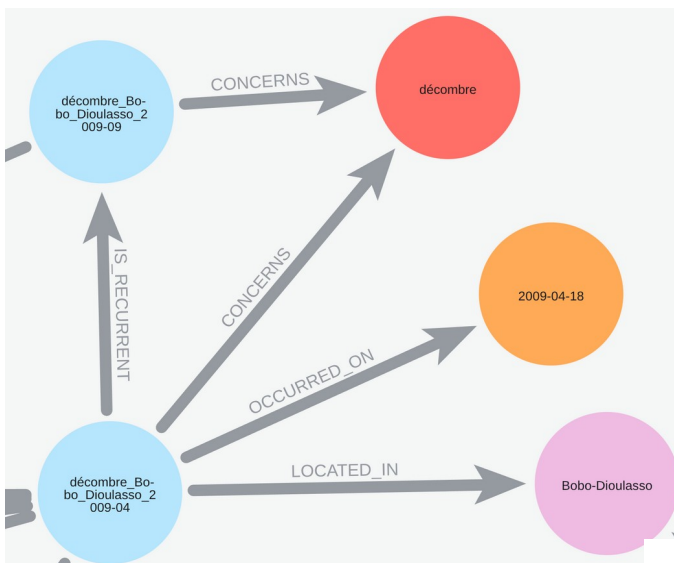
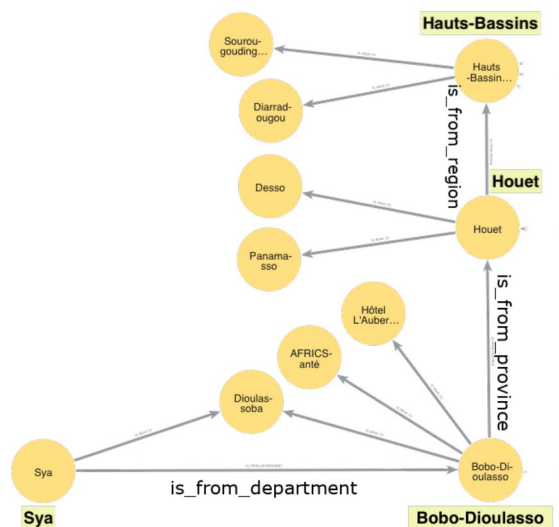
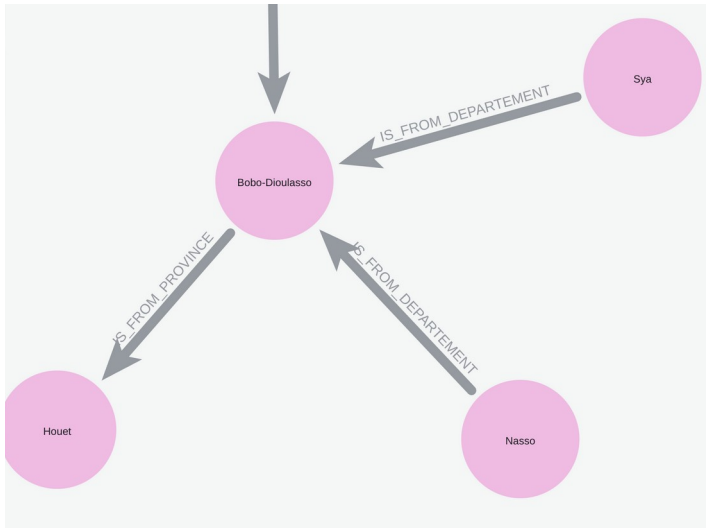
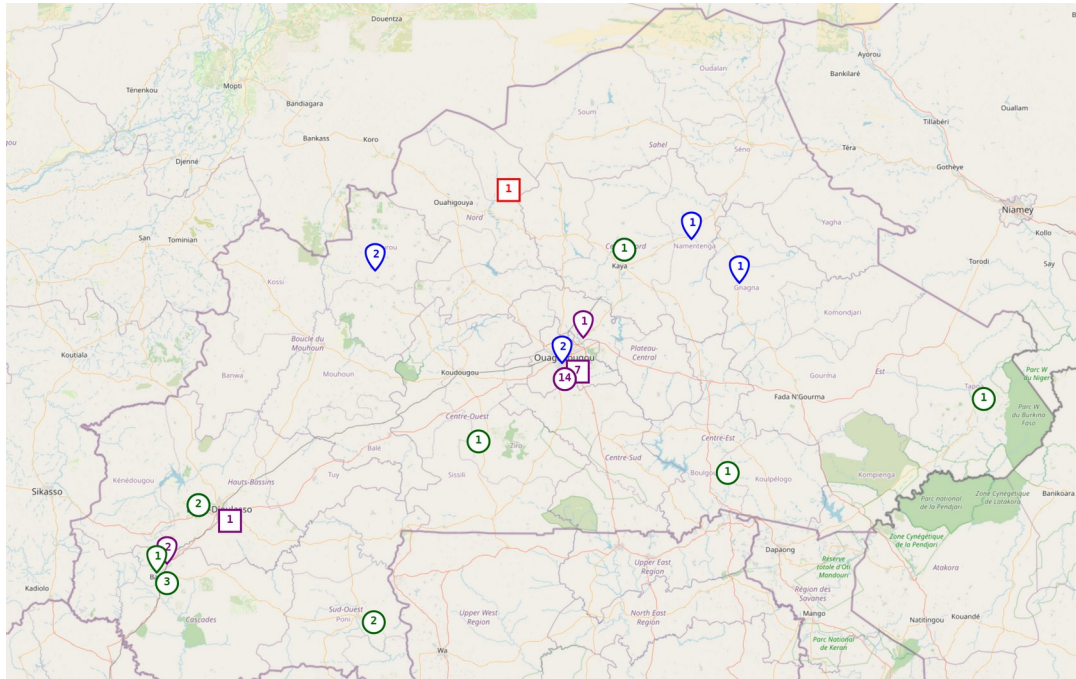
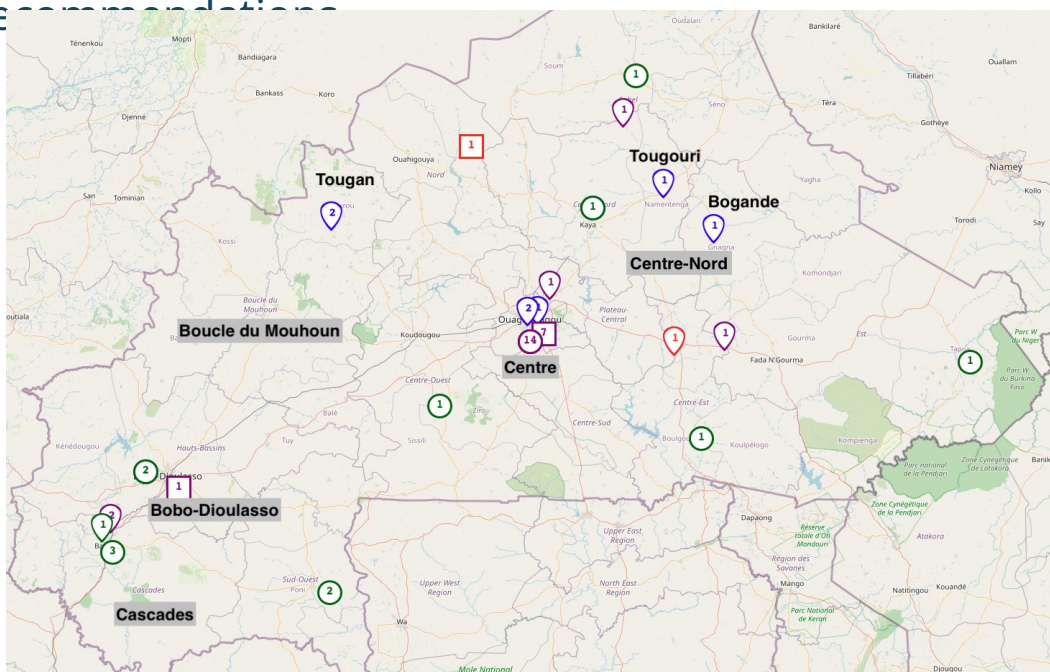


Figure 5

On top is the recreated version of the map generated in the steps 7 and 8. It differs from the original map in the label of maps. Some of the locations coordinates were added manually. This was done with Query 4 from the README.



Recommendations



The authors

adjusted the file requirements.txt (added folium and resolved issues with conflicting requirements after fairseq2n v0.3.0rc1 graduated to fairseq2n v0.3.01)

They also adjusted paths in the scripts. Additionally, they added a comment in the README that :source command is not supported on Neo4j Desktop. All my suggestions were already perfectly implemented by the authors.

Citing this document

This report is part of the reproducibility review at the AGILE conference. For more information see <https://reproducible-agile.github.io/>. This document is published on OSF/RE at OSF/RE *10.17605/OSF.IO/BV4MK*.

To cite the report use

10.17605/OSF.IO/BV4MK