EUPEG: Towards an Extensible and Unified Platform for Evaluating Geoparsers

Yingjie Hu
University at Buffalo, NY, USA
yhu42@buffalo.edu

ABSTRACT

Geoparsing, namely recognizing and geo-locating place mentions from unstructured texts, is a critical task in geographic information retrieval (GIR). While a number of geoparsers have been developed, they were often tested on different datasets using different performance metrics. Consequently, it is difficult to compare multiple geoparsers directly. In recent years, open corpora with human annotations have been developed and shared by researchers. However, much effort is still needed for implementing and applying previous geoparsers to these annotated corpora or to *rehydrate* certain datasets (e.g., tweets) due to data sharing restrictions. This short paper presents the early work of EUPEG: an Extensible and Unified Platform for Evaluating Geoparsers. EUPEG is an open-source and Web-based platform which incorporates existing geoparsers and hosts a set of annotated corpora. A newly developed geoparser can be connected to EUPEG and compared with other geoparsers based on the hosted datasets. The objectives of EUPEG are to enable systematic comparison of geoparsers across datasets and to reduce the amount of time and effort that researchers have to spend in implementing previous geoparsers as baselines.

CCS CONCEPTS

• **Information systems → Retrieval effectiveness**; • **Computing methodologies → Information extraction**;

KEYWORDS

Geoparsing, benchmarking framework, evaluation, GIR

ACM Reference Format:
Yingjie Hu. 2018. EUPEG: Towards an Extensible and Unified Platform for Evaluating Geoparsers. In *12th Workshop on Geographic Information Retrieval (GIR'18), November 6, 2018, Seattle, WA, USA.* ACM, New York, NY, USA, 2 pages. https://doi.org/10.1145/3281354.3281357

1 INTRODUCTION

Geoparsing is a critical task in geographic information retrieval (GIR) [7]. Unstructured texts, such as Web pages, news articles, social media posts, and historical archives, contain large amounts of valuable geographic information. A developed geoparsing system, called a *geoparser*, can take unstructured texts as the input, and output the recognized place names and their corresponding spatial footprints [2, 5, 9].

A number of geoparsers have already been developed, such as MetaCarta [4], GeoTxt [8], Edinburgh Geoparser [1], and TopoCluster [3]. While successfully handling many geoparsing tasks, existing geoparsers were often tested on problem-specific datasets, and it is difficult to compare the performances of different geoparsers on the same dataset or to compare the performances of the same geoparser across multiple datasets. It is worth noting that using a problem-specific dataset for testing is often necessary in order to distinguish a new geoparser from existing ones. For example, most geoparsers can have high performances when given a corpus containing only unambiguous names of major countries, while only a few can still perform well when the corpus contains many highly ambiguous place names. However, these problem-specific datasets were often not shared openly. In recent years, researchers have developed and shared open datasets, such as Local-Global Lexicon (LGL) corpus [10], WikToR [6], and GeoCorpora [14]. These datasets are extremely valuable for testing new geoparsers. Meanwhile, much effort is still needed for researchers to implement or apply existing geoparsers (and their own parsers) to these datasets. In addition, some annotated datasets (e.g., tweets) cannot be completely shared openly due to their policy restrictions, and researchers have to *rehydrate* these datasets (i.e., writing a program to retrieve tweets based on their IDs) before experiments.

This short paper presents the early work of EUPEG: an Extensible and Unified Platform for Evaluating Geoparsers. A main goal of this project is to reduce the time that researchers have to spend in implementing existing baselines and preparing datasets for evaluation experiments, so that researchers can focus on developing their own geoparsing methods.

2 SYSTEM DESCRIPTION

EUPEG is designed as an open-source and Web-based platform. It hosts a set of human-annotated corpora and provides connections to a number of existing geoparsers. A newly developed geoparser can be connected to EUPEG and compared with other geoparsers, and additional datasets can also be added into this platform. EUPEG was inspired by GERBIL, a benchmarking framework for evaluating entity annotators [13], and the great efforts in the GIR community, especially the works of [6, 11, 12, 14].

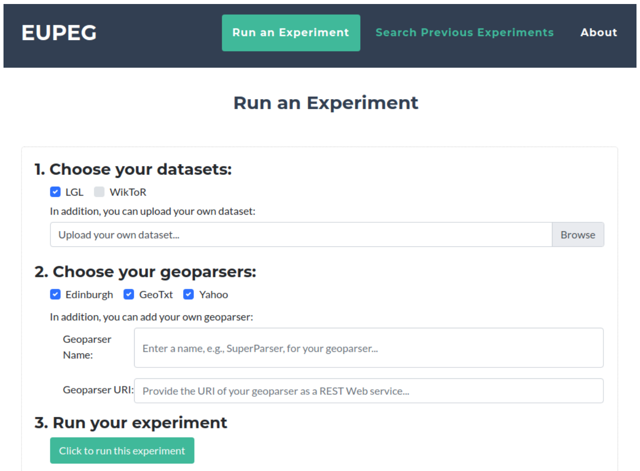
Two annotated datasets are currently hosted on EUPEG: LGL and WikToR. LGL is a news article corpus which was developed by Lieberman et al. [10]. It contains 588 articles published by 78 local newspapers from highly ambiguous places, such as *Paris News* (Texas) and *Paris Beacon-News* (Illinois). WikToR is a Wikipedia article corpus which was developed by Gritta et al. [6]. It contains 5,000 Wikipedia entries with ambiguous names, such as *Lima, Peru,*

GIR'18, November 6, 2018, Seattle, WA, USA

ACM ISBN 978-1-4503-6034-0/18/11…$15.00
https://doi.org/10.1145/3281354.3281357

Lima, Ohio, and *Lima, Oklahoma*. We are currently in the process of incorporating GeoCopora [14], a dataset of annotated tweets, into EUPEG. A new dataset created based on a pre-defined format can also be uploaded to the platform for experiments.

Three geoparsers are currently available on EUPEG, which are Edinburgh geoparser [1], GeoTxt [8], and Yahoo! PlaceSpotter. For the geoparsers that provide a REST API, such as GeoTxt and Yahoo!, EUPEG directly sends texts to their API and retrieves the annotated results. For the geoparsers without a REST API, such as Edinburgh geoparser, they are deployed on the local server of EUPEG. A user can also connect a new geoparser to EUPEG by providing the REST API of the geoparser (the geoparser has to be published as a REST service following a pre-defined format first). Figure 1 provides a screenshot of EUPEG. A user can (1) select the existing datasets or add a new dataset, (2) choose the existing geoparsers and add his/her own geoparser, and (3) click the run experiment button.



Run an Experiment

1. Choose your datasets:

☑ LGL ☐ WikToR

In addition, you can upload your own dataset:

Upload your own dataset...

2. Choose your geoparsers:

☑ Edinburgh ☑ GeoTxt ☑ Yahoo

In addition, you can add your own geoparser:

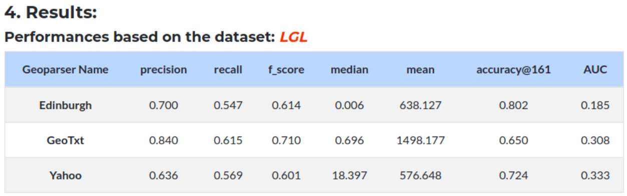
Geoparser Name: Enter a name, e.g., SuperParser, for your geoparser...

Geoparser URI: Provide the URI of your geoparser as a REST Web service...

3. Run your experiment

Figure 1: A screenshot of the current EUPEG interface.

Seven evaluation metrics are currently provided on EUPEG. These metrics include *precision*, *recall*, and *F-score*, which evaluate the capability of the geoparsers in identifying the correct place instances without considering the offsets of the identified locations. The metrics, *median* and *mean*, evaluate how the locations identified by a geoparser deviate from the ground-truth locations (in kilometers). The metric, *accuracy@161 km*, measures the percentages of the geoparsed locations that are within 161 km (100 miles) of the ground truths. The metric, *Area Under the Curve* (AUC), quantifies the area under the distance error curve [6]. Figure 2 shows an example of the obtained evaluation results. Depending on specific needs, one can select a subset of the metrics that are more suitable for the evaluation experiment.



4. Results:

Performances based on the dataset: *LGL*

Geoparser Name	precision	recall	f_score	median	mean	accuracy@161	AUC
Edinburgh	0.700	0.547	0.614	0.006	638.127	0.802	0.185
GeoTxt	0.840	0.615	0.710	0.696	1498.177	0.650	0.308
Yahoo	0.636	0.569	0.601	18.397	576.648	0.724	0.333

Figure 2: A screenshot of EUPEG showing the evaluating results of three geoparsers based on the LGL dataset.

3 DISCUSSION AND FUTURE WORK

This short paper reports the early work of EUPEG, an open-source and Web-based platform for facilitating the evaluation of geoparsers. EUPEG provides a unified platform that combines annotated corpora, developed geoparsers, and evaluation metrics from existing literature. Much work still needs to be done to make EUPEG a full-fledged platform. First, more datasets (e.g., GeoCorpora [14]) and geoparsers (e.g., Topocluster[3]) in the existing literature should be added. Meanwhile, the data formats for EUPEG to automatically accept newly developed corpora and geoparsers need to be specified. This part could use the feedback from the GIR community, and the GIR'18 workshop provides a great opportunity for obtaining such feedback. In addition, current evaluation metrics are mainly based on point locations which may not accurately capture the spatial footprints of certain geographic features, such as rivers and states. Metrics based on lines and polygons can be introduced in the near future. Finally, EUPEG needs to be published as a free Web service and its source code will be shared.

REFERENCES

[1] Beatrice Alex, Kate Byrne, Claire Grover, and Richard Tobin. 2015. Adapting the Edinburgh geoparser for historical georeferencing. *International Journal of Humanities and Arts Computing* 9, 1 (2015), 15–35.

[2] Adrien Barbaresi. 2017. Towards a toolbox to map historical text collections. In *Proceedings of the 11th Workshop on Geographic Information Retrieval*. ACM, New York, NY, USA, 5.

[3] Grant DeLozier, Jason Baldridge, and Loretta London. 2015. Gazetteer-Independent Toponym Resolution Using Geographic Word Profiles. In *Proceedings of the AAAI Conference on Artificial Intelligence*. AAAI Press, Palo Alto, CA, USA, 2382–2388.

[4] John R Frank, Erik M Rauch, and Karen Donoghue. 2006. Spatially coding and displaying information. US Patent 7,117,199.

[5] Nuno Freire, José Borbinha, Pável Calado, and Bruno Martins. 2011. A metadata geoparsing system for place name recognition and resolution in metadata records. In *Proceedings of the 11th annual international ACM/IEEE joint conference on Digital libraries*. ACM, New York, NY, USA, 339–348.

[6] Milan Gritta, Mohammad Taher Pilehvar, Nut Limsopatham, and Nigel Collier. 2018. What's missing in geographical parsing? *Language Resources and Evaluation* 52, 2 (2018), 603–623.

[7] Christopher B. Jones and Ross S. Purves. 2008. Geographical information retrieval. *International Journal of Geographical Information Science* 22, 3 (2008), 219–228.

[8] Morteza Karimzadeh, Wenyi Huang, Siddhartha Banerjee, Jan Oliver Wallgrün, Frank Hardisty, Scott Pezanowski, Prasenjit Mitra, and Alan M MacEachren. 2013. GeoTxt: a web API to leverage place references in text. In *Proceedings of the 7th workshop on geographic information retrieval*. ACM, New York, NY, USA, 72–73.

[9] Jochen L Leidner. 2008. *Toponym resolution in text: Annotation, evaluation and applications of spatial grounding of place names*. Universal-Publishers, Irvine, CA, USA.

[10] Michael D Lieberman, Hanan Samet, and Jagan Sankaranarayanan. 2010. Geotagging with local lexicons to build indexes for textually-specified spatial data. In *2010 IEEE 26th International Conference on Data Engineering (ICDE)*. IEEE, Long Beach, CA, USA, 201–212.

[11] Ross S Purves, Paul Clough, Christopher B Jones, Mark H Hall, Vanessa Murdock, et al. 2018. Geographic Information Retrieval: Progress and Challenges in Spatial Search of Text. *Foundations and Trends® in Information Retrieval* 12, 2-3 (2018), 164–318.

[12] Ludwig Richter, Johanna Geiß, Andreas Spitz, and Michael Gertz. 2017. HeidelPlace: An Extensible Framework for Geoparsing. In *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing: System Demonstrations*. ACL, Stroudsburg, PA, USA, 85–90.

[13] Ricardo Usbeck, Michael Röder, Axel-Cyrille Ngonga Ngomo, Ciro Baron, Andreas Both, Martin Brümmer, Diego Ceccarelli, Marco Cornolti, Didier Cherix, Bernd Eickmann, et al. 2015. GERBIL: general entity annotator benchmarking framework. In *Proceedings of the 24th International Conference on World Wide Web*. International World Wide Web Conferences Steering Committee, Republic and Canton of Geneva, Switzerland, 1133–1143.

[14] Jan Oliver Wallgrün, Morteza Karimzadeh, Alan M. MacEachren, and Scott Pezanowski. 2018. GeoCorpora: building a corpus to test and train microblog geoparsers. *International Journal of Geographical Information Science* 32, 1 (2018), 1–29.