GIR relevant papers / works / Data

End 2 End Geocoding

end to end geocoding model – This is what I was thinking of doing !

https://arxiv.org/abs/2107.00080

I was thinking of doing this !

Spatial Language Representation with Multi-Level Geocoding

geocoding model – classifay multi levels locations using convolution network

https://arxiv.org/pdf/2008.09236.pdf

A Coherent Unsupervised Model for Toponym Resolution

Relation ranking

Leverages the connections between nearby place names as evidence to resolve toponyms.
We explore the interactions between multiple interpretations of mentions and the relationships between different toponyms in a document to build a model that finds the most coherent resolution

https://dl.acm.org/doi/pdf/10.1145/3178876.3186027

Which Melbourne? Augmenting Geocoding with Maps

CamCoder convolution geo coder

https://aclanthology.org/P18-1119.pdf

Mordecai full text geoparsing

Python package – Extract the place names from a piece of English-language text

https://github.com/openeventdata/mordecai

GeoText 2019

GeoTxt is a scalable geoparsing system for the recognition and geolocation of place names in unstructured text. GeoTxt offers six named entity recognition (NER) algorithms for place name recognition, and utilizes Apache Solr for the indexing, ranking, and retrieval of toponyms, enabling scalable geoparsing for streaming text. GeoTxt offers a flexible application programming interface (API), generating a GeoJSON FeatureCollection as output 2019.
https://github.com/geovista/GeoTxt

Extracting and modeling geographic information from scientific articles 2021

extract and represent relevant locations from scientific articles 2021

https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0244918

Bi-directional Recurrent Neural Network Models for Geographic Location Extraction in Biomedical Literature 2018

extracting geographic locations and their effective contribution to the disambiguation task using population heuristics 2018.

https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0244918

UniMelb at SemEval-Task 12: Multi-model Combination for Toponym Resolution 2019

Toponym resolution in scientific paper task of SemEval 2019

https://aclanthology.org/S19-2231.pdf

I'm Eating a Sandwich in Hong Kong Modeling Tweets 2011

model on twitter for geo loacation 2011

https://docs.google.com/document/d/1QQ2NO1aViLDc7UYc9PnQLfMpmg5T3k04eoYfqGCqxNk/edit

Improving Geocoding for City-Level Locations

geocoding improvments

https://ieeexplore.ieee.org/document/8665524

Lost in space: Geolocation in event data

classified correct location words

https://arxiv.org/pdf/1611.04837.pdf

Data

GeoNames geographical database

contains over 27 million geographical names and consists of over 12 million unique features whereof 4.8 million populated places and 15 million alternate names. All features are categorized into one out of nine feature classes and further subcategorized into one out of 645 feature codes.

GeoNames is integrating geographical data such as names of places in various languages, elevation, population and others from various sources. All lat/long coordinates are in WGS84 (World Geodetic System 1984)
https://www.geonames.org/

Twitter spritzer:

A simple collection of JSON grabbed from the general twitter stream, for the purposes of research, history, testing and memory. This is the "Spritzer" version, the most light and shallow of Twitter grabs. Unfortunately, we do not currently have access to the Sprinkler or Garden Hose versions of the stream.
https://archive.org/details/twitterstream?&sort=-week&page=2