Adapting and integrating existing open source projects

Ethical Visualization in the Age of Big Data

Ludovic Moncla
INSA Lyon, LIRIS CNRS UMR 5205

ludovic.moncla@insa-lyon.fr

10 January 2020

What tools for what needs ?

1. Browse and share a collection of documents
   - Open-ONI, PhiloLogic, etc.
2. NLP : NER, NED, Geoparsing, Sentiment analysis, etc.
   - OpeNER, NERC-fr, Edimburgh Geoparser, etc.
3. Visualization and exploration
   - Palladio, D3, etc.
4. Manual annotations and visualization
   - Brat, Recogito, GeoViz, etc.

1. Browse and share documents

Open-ONI : Open Online Newspaper Initiative

`https://open-oni.github.io/`

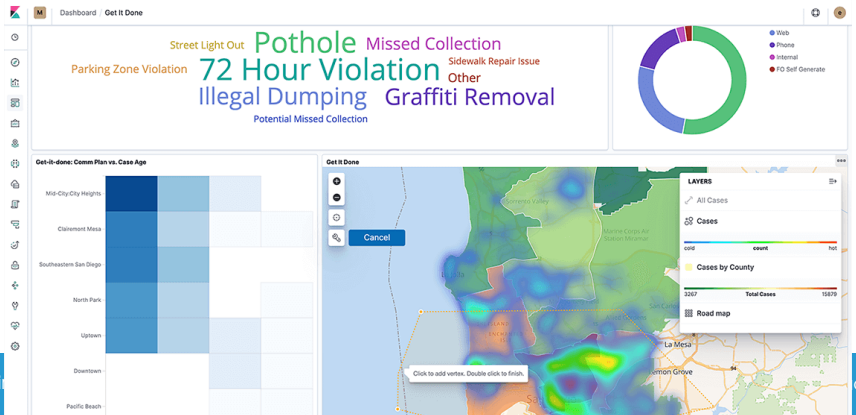
- Historic American newspapers browsable and searchable on the web
- Originally used to make Chronicling America
- It provides a set of functionality for loading, modeling and indexing data

1. Browse and share documents

Elastic search & Kibana

`https://www.elastic.co/fr/products/maps`

- Indexing, searching and visualizing information

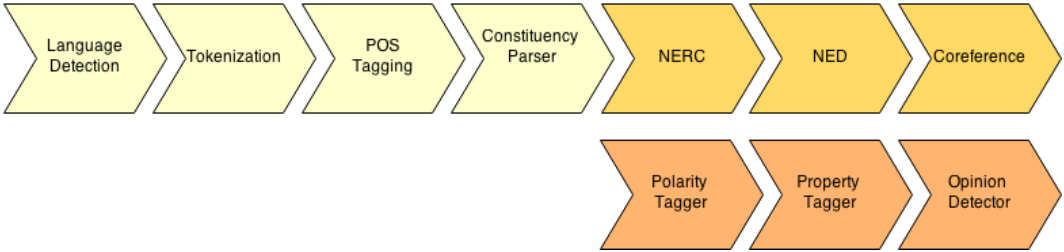


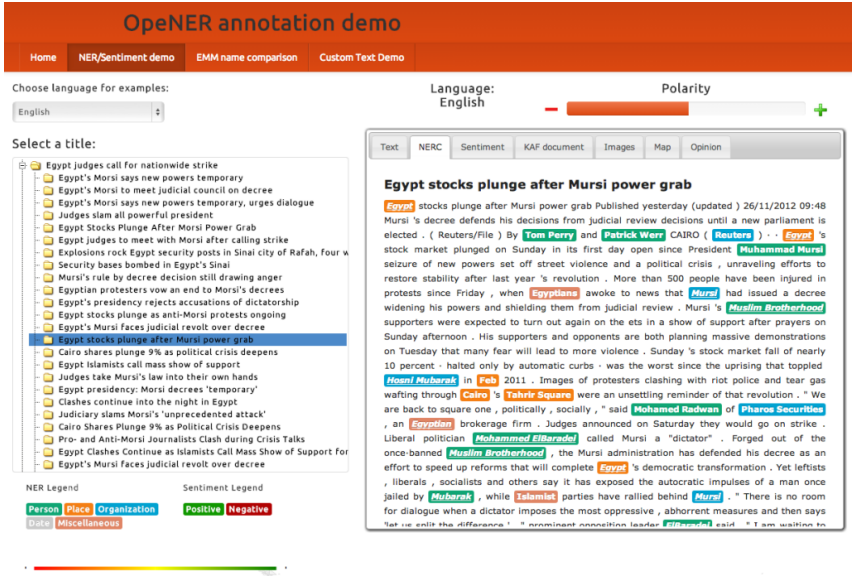
2. NLP tools

The OpeNER project (Open Polarity Enhanced Name Entity Recognition)

https://www.opener-project.eu/

- Detecting sentiments, opinions and named entities in texts
  - Open Source
  - Multilingual



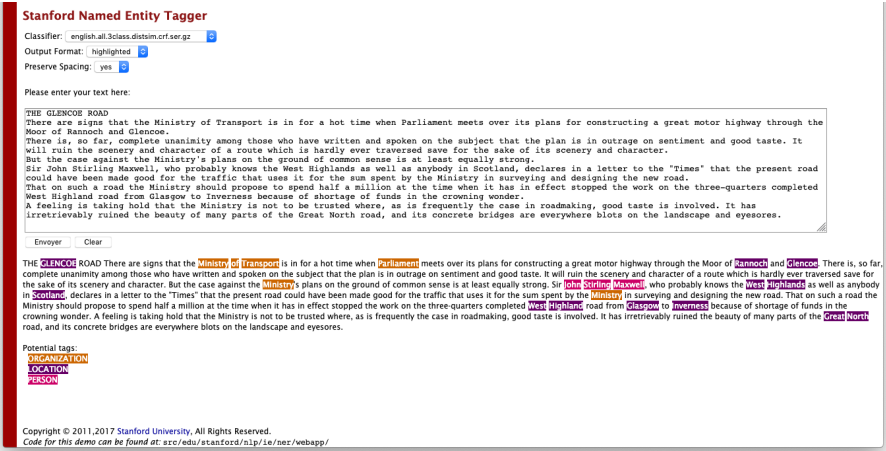

OpeNER annotation demo
Home
NER/Sentiment demo
EMM name comparison
Custom Text Demo
Choose language for examples:
English
Language:
English
Polarity
Select a title:
Egypt judges call for nationwide strike
Egypt's Morsi says new powers temporary
Egypt's Morsi to meet judicial council on decree
Egypt's Morsi says new powers temporary, urges dialogue
Judges slam all powerful president
Egypt Stocks Plunge After Morsi Power Grab
Egypt judges to meet with Morsi after calling strike
Explosions rock Egypt security posts in Sinai city of Rafah, four w
Security bases bombed in Egypt's Sinai
Mursi's rule by decree decision still drawing anger
Egyptian protesters vow an end to Morsi's decrees
Egypt's presidency rejects accusations of dictatorship
Egypt stocks plunge as anti-Morsi protests ongoing
Egypt's Morsi faces judicial revolt over decree
Egypt stocks plunge after Mursi power grab
Cairo shares plunge 9% as political crisis deepens
Egypt Islamists call mass show of support
Judges take Morsi's law into their own hands
Egypt presidency: Morsi decrees 'temporary'
Clashes continue into the night in Egypt
Judiciary slams Morsi's 'unprecedented attack'
Cairo Shares Plunge 9% as Political Crisis Deepens
Pro- and Anti-Morsi Journalists Clash during Crisis Talks
Egypt Clashes Continue as Islamists Call Mass Show of Support for
Egypt's Morsi faces judicial revolt over decree
NER Legend
Person
Place
Organization
Date
Miscellaneous
Sentiment Legend
Positive
Negative
Text
NERC
Sentiment
KAF document
Images
Map
Opinion
Egypt stocks plunge after Mursi power grab
Egypt stocks plunge after Mursi power grab Published yesterday (updated) 26/11/2012 09:48 Mursi 's decree defends his decisions from judicial review decisions until a new parliament is elected . (Reuters/File) By Tom Perry and Patrick Werr CAIRO (Reuters) - - Egypt 's stock market plunged on Sunday in its first day open since President Muhammad Mursi seizure of new powers set off street violence and a political crisis , unraveling efforts to restore stability after last year 's revolution . More than 500 people have been injured in protests since Friday , when Egyptians awoke to news that Mursi had issued a decree widening his powers and shielding them from judicial review . Mursi 's Muslim Brotherhood supporters were expected to turn out again on the ets in a show of support after prayers on Sunday afternoon . His supporters and opponents are both planning massive demonstrations on Tuesday that many fear will lead to more violence . Sunday 's stock market fall of nearly 10 percent · halted only by automatic curbs · was the worst since the uprising that toppled Hosni Mubarak in Feb 2011 . Images of protesters clashing with riot police and tear gas wafting through Cairo 's Tahrir Square were an unsettling reminder of that revolution . " We are back to square one , politically , socially , " said Mohamed Radwan of Pharos Securities , an Egyptian brokerage firm . Judges announced on Saturday they would go on strike . Liberal politician Mohammed ElBaradei called Mursi a "dictator" . Forged out of the once-banned Muslim Brotherhood , the Mursi administration has defended his decree as an effort to speed up reforms that will complete Egypt 's democratic transformation . Yet leftists , liberals , socialists and others say it has exposed the autocratic impulses of a man once jailed by Mubarak , while Islamist parties have rallied behind Mursi . " There is no room for dialogue when a dictator imposes the most oppressive , abhorrent measures and then says

2. NLP tools

	SPACY	NLTK	CORENLP
Programming language	Python	Python	Java / Python
Neural network models	✓	✗	✓
Integrated word vectors	✓	✗	✗
Multi-language support	✓	✓	✓
Tokenization	✓	✓	✓
Part-of-speech tagging	✓	✓	✓
Sentence segmentation	✓	✓	✓
Dependency parsing	✓	✗	✓
Entity recognition	✓	✓	✓
Entity linking	✓	✗	✗
Coreference resolution	✗	✗	✓

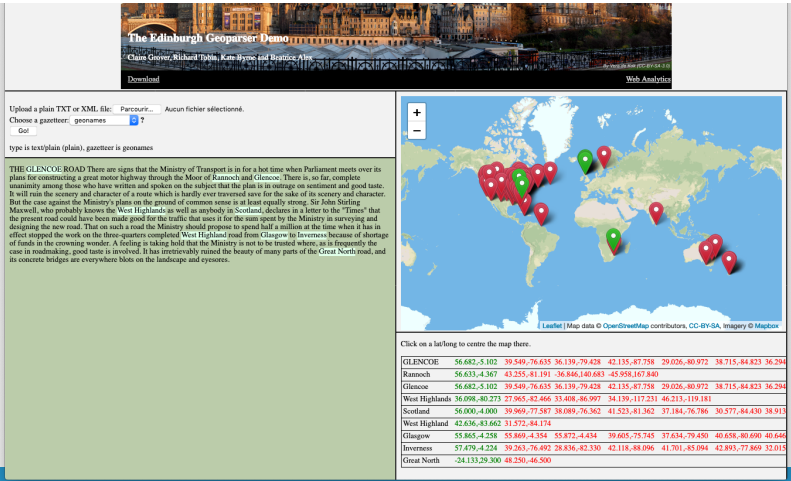
2. NLP tools

Stanford NER



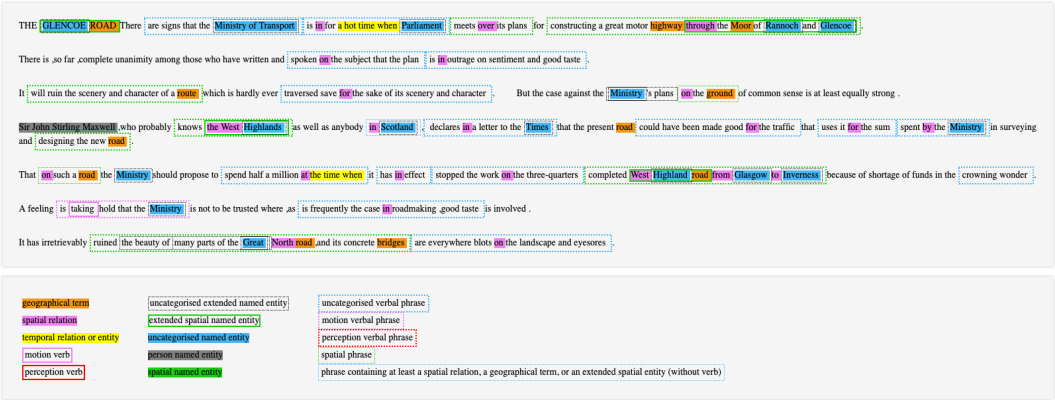
2. NLP tools

The Edimburgh Geoparser



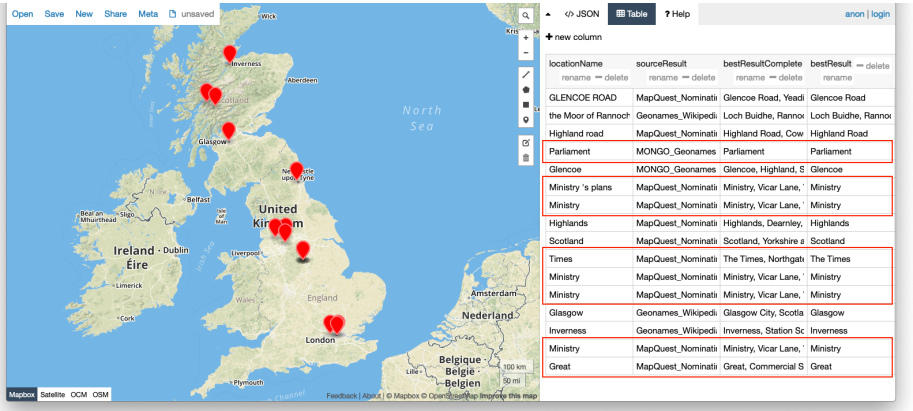
2. NLP tools

PERDIDO Geoparser



2. NLP tools

PERDIDO Geoparser



locationName	sourceResult	bestResultComplete	bestResult
rename delete	rename delete	rename delete	rename delete
GLENCOE ROAD	MapQuest_Nominati	Glencoe Road, Yeadi	Glencoe Road
the Moor of Rannoch	Geonames_Wikipedi	Loch Buidhe, Rannoc	Loch Buidhe, Rannoc
Highland road	MapQuest_Nominati	Highland Road, Cow	Highland Road
Parliament	MONGO_Geonames	Parliament	Parliament
Glencoe	MONGO_Geonames	Glencoe, Highland, S	Glencoe
Ministry 's plans	MapQuest_Nominati	Ministry, Vicar Lane, '	Ministry
Ministry	MapQuest_Nominati	Ministry, Vicar Lane, '	Ministry
Highlands	MapQuest_Nominati	Highlands, Dearnley,	Highlands
Scotland	MapQuest_Nominati	Scotland, Yorkshire a	Scotland
Times	MapQuest_Nominati	The Times, Northgate	The Times
Ministry	MapQuest_Nominati	Ministry, Vicar Lane, '	Ministry
Ministry	MapQuest_Nominati	Ministry, Vicar Lane, '	Ministry
Glasgow	Geonames_Wikipedi	Glasgow City, Scotla	Glasgow
Inverness	Geonames_Wikipedi	Inverness, Station Sc	Inverness
Ministry	MapQuest_Nominati	Ministry, Vicar Lane, '	Ministry
Great	MapQuest_Nominati	Great, Commercial S	Great

3. Visualization

Desktop software

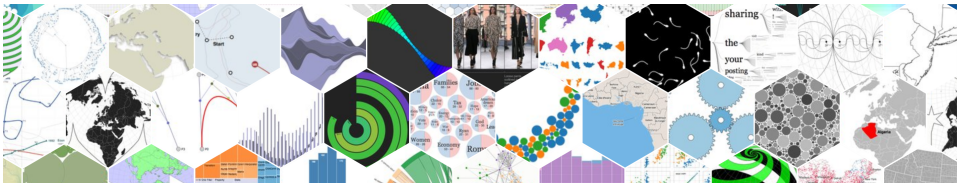
- Gephy (`https://gephi.org/`) for graphs and networks

Python library

- matplotlib, geoplotlib, folium, etc.

JavaScript library

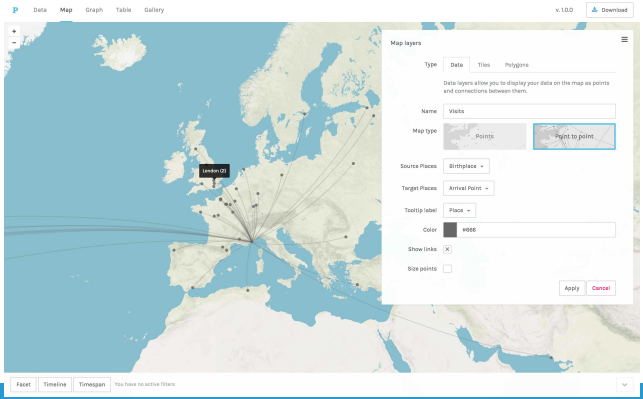
- D3 (Data-Driven Documents) : `https://d3js.org/`



3. Visualization

Palladio : visualize complex historical data

- `https://hdlab.stanford.edu/palladio`
- Load CSV or spreadsheet

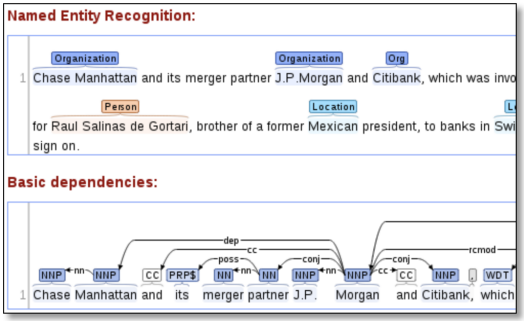


4. Manual annotation

Brat rapid annotation tool

`https://brat.nlplab.org/`

- Entity mention detection
- Visualization
- Evaluation system



4. Manual annotation

Pelagios & Recogito : History of places

`https://recogito.pelagios.org/`

- Semantic annotation tool for texts and images
- Collaborative document annotation
- Identify and map places (Automatic suggestions from gazetteers)

Questions

What are the needs ?

- Do we want to provide online tools for texts analysis ?
  - What kind of analysis ?
- Do we want to provide corpus of annotated or enriched documents ?
  - What kind of information do we want to annotate ? What format ? etc...
- Do we want to provide online tools for data vizualization ?
  - What kind of visualization ?

Thank you for your attention

CONTACT

Ludovic Moncla (INSA / LIRIS)
ludovic.moncla@insa-lyon.fr

INSA INSTITUT NATIONAL DES SCIENCES APPLIQUÉES LYON

LIRIS