

# Adapting and integrating existing open source projects

## Ethical Visualization in the Age of Big Data

Ludovic Moncla  
INSA Lyon, LIRIS CNRS UMR 5205



# 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/`

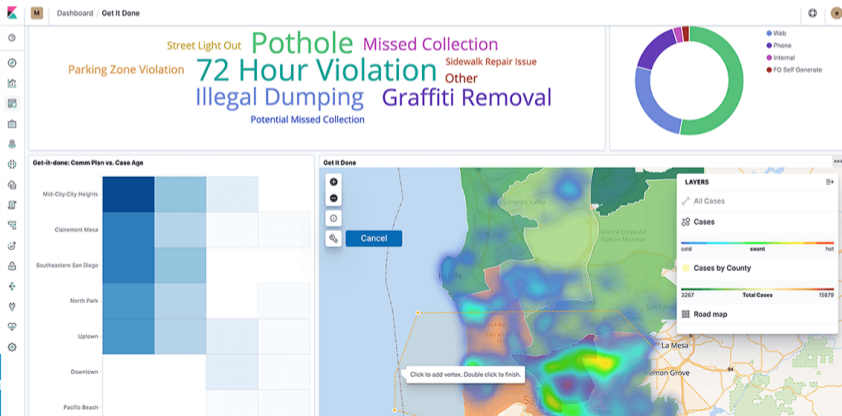
- Historic American newspapers browsable and searchable on the web
- Originally used to make Chronicling America
- It provides a set of functionality for loading, modeling and indexing data

# 1. Browse and share documents

## Elastic search & Kibana

<https://www.elastic.co/fr/products/maps>

- Indexing, searching and visualizing information

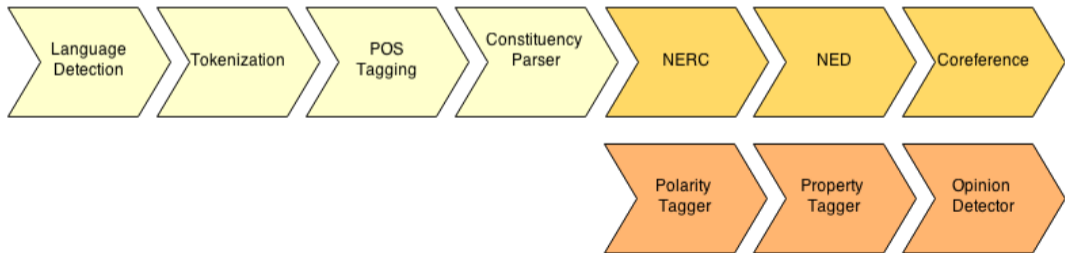


## 2. NLP tools

### The OpenNER 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 Mursi 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 Mursi'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 Mursi 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 , unravelling 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 "let us exit the difference " " prominent opposition leader **ElBaradei** said . " I am waiting to

## 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

**Stanford Named Entity Tagger**

Classifier:

Output Format:

Preserve Spacing:

Please enter your text here:

THE GLENCOE ROAD  
There are signs that the Ministry of Transport is in for a hot time when Parliament meets over its plans for constructing a great motor highway through the Moor of Rannoch and Glencoe.  
There is, so far, complete unanimity among those who have written and spoken on the subject that the plan is in outrage on sentiment and good taste. It will ruin the scenery and character of a route which is hardly ever traversed save for the sake of its scenery and character.  
But the case against the Ministry's plans on the ground of common sense is at least equally strong.  
Sir John Stirling Maxwell, who probably knows the West Highlands as well as anybody in Scotland, declares in a letter to the "Times" that the present road could have been made good for the traffic that uses it for the sum spent by the Ministry in surveying and designing the new road.  
That on such a road the Ministry should propose to spend half a million at the time when it has in effect stopped the work on the three-quarters completed West Highland road from Glasgow to Inverness because of shortage of funds in the crowning wonder.  
A feeling is taking hold that the Ministry is not to be trusted where, as is frequently the case in roadmaking, good taste is involved. It has irretrievably ruined the beauty of many parts of the Great North road, and its concrete bridges are everywhere blots on the landscape and eyesores.

THE GLENCOE ROAD There are signs that the Ministry of Transport is in for a hot time when Parliament meets over its plans for constructing a great motor highway through the Moor of Rannoch and Glencoe. There is, so far, complete unanimity among those who have written and spoken on the subject that the plan is in outrage on sentiment and good taste. It will ruin the scenery and character of a route which is hardly ever traversed save for the sake of its scenery and character. But the case against the Ministry's plans on the ground of common sense is at least equally strong. Sir John Stirling Maxwell, who probably knows the West Highlands as well as anybody in Scotland, declares in a letter to the "Times" that the present road could have been made good for the traffic that uses it for the sum spent by the Ministry in surveying and designing the new road. That on such a road the Ministry should propose to spend half a million at the time when it has in effect stopped the work on the three-quarters completed West Highland road from Glasgow to Inverness because of shortage of funds in the crowning wonder. A feeling is taking hold that the Ministry is not to be trusted where, as is frequently the case in roadmaking, good taste is involved. It has irretrievably ruined the beauty of many parts of the Great North road, and its concrete bridges are everywhere blots on the landscape and eyesores.


Potential tags:  
**ORGANIZATION**  
**LOCATION**  
**PERSON**

Copyright © 2011,2017 Stanford University, All Rights Reserved.  
Code for this demo can be found at: [src/edu/stanford/nlp/ie/ner/webapp/](http://src/edu/stanford/nlp/ie/ner/webapp/)



## 2. NLP tools

### The Edinburgh Geoparser



The Edinburgh Geoparser Demo

Claire Grover, Richard Tobin, Kate Hynes and Beatrice Alia

Download


Web Analytics

Upload a plain TXT or XML file:  Aucun fichier sélectionné.

Choose a gazetteer:  ?

type is text/plain (plain), gazetteer is geonames

THE GLENCOE ROAD There are signs that the Ministry of Transport is in for a hot time when Parliament meets over its plans for constructing a great motor highway through the Moor of Ramoch and Glencoe. There is, so far, complete unanimity among those who have written and spoken on the subject that the plan is in outrage on sentiment and good taste. It will ruin the scenery and character of a route which is hardly ever traversed save for the sake of its scenery and character. But the case against the Ministry's plans on the ground of common sense is at least equally strong. Sir John Stirling Maxwell, who probably knows the West Highlands as well as anybody in Scotland, declares in a letter to the "Times" that the present road could have been made good for the traffic that uses it for the sum spent by the Ministry in surveying and designing the new road. That on such a road the Ministry should propose to spend half a million at the time when it has in effect stopped the work on the three-quarters completed West Highland road from Glasgow to Inverness because of shortage of funds in the crowning wonder. A feeling is taking hold that the Ministry is not to be trusted where, as is frequently the case in roadmaking, good taste is involved. It has irrevocably ruined the beauty of many parts of the Great North road, and its concrete bridges are everywhere blot on the landscape and eyesores.



Click on a lat/long to centre the map there.

GLENCOE	56.682,-5.102	39.549,-76.635	36.139,-79.428	42.135,-87.758	29.026,-80.972	38.715,-84.823	36.294
Ramoch	56.633,-4.367	43.255,-81.191	-36.846,140.683	-45.958,167.840			
Glencoe	56.682,-5.102	39.549,-76.635	36.139,-79.428	42.135,-87.758	29.026,-80.972	38.715,-84.823	36.294
West Highlands	36.098,-40.273	27.965,-82.466	33.408,-86.997	34.139,-117.231	46.213,-119.181		
Scotland	56.000,-4.000	39.969,-77.587	38.089,-76.362	41.523,-81.362	37.184,-76.786	30.577,-84.430	38.913
West Highland	42.636,-43.662	31.572,-84.174					
Glasgow	55.865,-4.258	55.869,-4.354	55.872,-4.434	39.605,-75.745	37.634,-79.450	40.658,-80.690	40.646
Inverness	57.479,-4.224	39.263,-76.492	28.836,-82.330	42.118,-88.096	41.701,-85.094	42.893,-77.869	32.015
Great North	-24.133,29.300	48.250,-46.500					

Adapting and integrating existing open source projects

10 January 2020 – 9/17

## 2. NLP tools

### PERDIDO Geoparser

THE **GLENCOE ROAD** There are signs that the **Ministry of Transport** is in for a hot time when **Parliament** meets over its plans for constructing a great motor **highway through the Moor of Rannoch and Glencoe**.

There is, so far, complete unanimity among those who have written and spoken on the subject that the plan is in outrage on sentiment and good taste.

It will ruin the scenery and character of a **route** which is hardly ever traversed save for the sake of its scenery and character. But the case against the **Ministry**'s plans on the ground of common sense is at least equally strong.

**Sir John Stirling Maxwell**, who probably knows **the West Highlands** as well as anybody in **Scotland**, declares in a letter to the **Times** that the present **road** could have been made good for the traffic that uses it for the sum spent by the **Ministry** in surveying and designing the new **road**.

That on such a **road** the **Ministry** should propose to spend half a million at the time when it has in effect stopped the work on the three-quarters completed **West Highland road from Glasgow to Inverness** because of shortage of funds in the crowning wonder.

A feeling is taking hold that the **Ministry** is not to be trusted where as is frequently the case in roadmaking good taste is involved.

It has irretrievably ruined the beauty of many parts of the **Great North road** and its concrete **bridges** are everywhere blots on the landscape and eyesores.

geographical term

spatial relation

temporal relation or entity

motion verb

perception verb

uncategorised extended named entity

extended spatial named entity

uncategorised named entity

person named entity

spatial named entity

uncategorised verbal phrase

motion verbal phrase

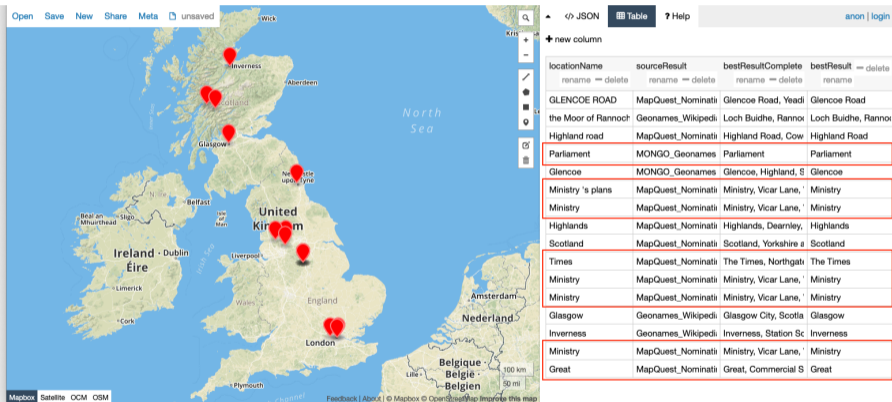
perception verbal phrase

spatial phrase

phrase containing at least a spatial relation, a geographical term, or an extended spatial entity (without verb)

## 2. NLP tools

### PERDIDO Geoparser



The screenshot displays the PERDIDO Geoparser interface. On the left is a map of the British Isles with red location pins. On the right is a data table with columns: locationName, sourceResult, bestResultComplete, and bestResult. The table contains 15 rows of data, with several rows highlighted in red. The interface also includes a search bar, a 'JSON' view toggle, and a 'Table' view toggle.

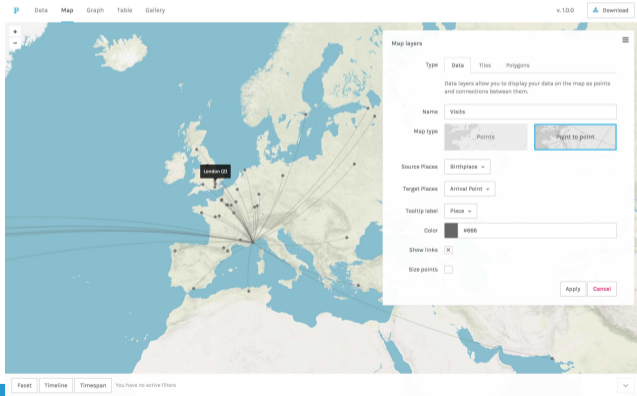
locationName	sourceResult	bestResultComplete	bestResult
GLENCEOE ROAD	MapQuest_Nominati	Glenceoe Road, Yeadi	Glenceoe Road
the Moor of Rannoch	Geonames_Wikipedi	Loch Buidhe, Rannoch	Loch Buidhe, Rannoch
Highland road	MapQuest_Nominati	Highland Road, Cow	Highland Road
Parliament	MONGO_Geonames	Parliament	Parliament
Glenceoe	MONGO_Geonames	Glenceoe, Highland, S	Glenceoe
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 s	Scotland
Times	MapQuest_Nominati	The Times, Northgat	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

Palladio : visualize complex historical data

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

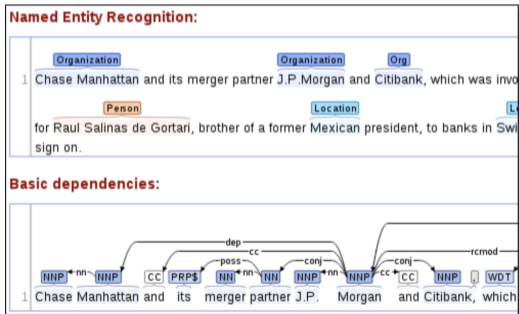


## 4. Manual annotation

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)

The screenshot shows a text editor interface for manual annotation. The text being annotated is: "Calypso, who had got him into a large cave and wanted to marry him. But as year go back to Ithaca; even then, however, when he was among his own people, his tro gun to piti". The word "Ithaca" is highlighted in blue. A dropdown menu is open over the highlighted word, showing three options: "Place" (with a location pin icon), "Person" (with a person icon), and "Event" (with a star icon). The "Place" option is selected, and a map is displayed below it. The map shows the location of Ithaca in the Aegean Sea, with a blue location pin. The map information panel shows: "Ithaca Ins.", "pleiades:530906", "An ancient place, cited: BAtlas 54 C5 Ithaca Ins.", "Ithaca", and "750 BC - 640". Below the map, there are three buttons: "Automatic Match" (with a warning icon), "Confirm", and "Change" (with a trash icon). At the bottom of the interface, there is a text input field labeled "Add a comment...".

# 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](mailto:ludovic.moncla@insa-lyon.fr)

**INSA** | INSTITUT NATIONAL  
DES SCIENCES  
APPLIQUÉES  
LYON

**LIRIS**