**Crowd Pulse**

Crowd Pulse is a platform for the semantic analysys of social contents. It allows to extract, filter, manipulate and analyze data retrieved from social (Twitter in our case). We can perform all these operations by means of a configuration file written in JSON inside the Crowd Pulse platform. Each configuration file is called **project**. Each project is composed by:

* Process: it contains information about the name of the project and logs’ path;
* Nodes: it contains plugins and the configurations of them;
* Edges: describe the sequence in which plugins activate;

The core of Crowd Pulse are the available plugins.

# Plugins

## I.I twitter-extractor

Different options available:

* Based on content (terms, hashtag, etc.);
* Based on user (sent and received);
* Based on geo-coordinates(**point\_radius** approach);
* Based on filters (language, date, etc.);
* Combination of filters;

## I.II message-persist

It permits to store data in a database given the name.

## I.III tokenizer-opennlp

It executes the tokenization of the retrieved tweets. In its configuration we can specify the minimum number of characters a word must contain to be considered useful and if we want to consider mentions, urls or hashtags.

## I.IV message-fetch

It permits to retrieve data given the name of the database.

## I.V simple

It executes the stopword removal for the tweets. In its configuration we can specify to which parts it must execute like tokens, tags or categories and the dictionary to be used to remove unrelevant words. We can also specify custom words to remove.

## I.VI lemmatizer-stanford

Plugin for lemmatization.

## I.VII tagme

Plugin responsible for entity-linking. It uses TAG.ME. Before this plugin, lemmatizer-stanford must be applied. You must specify a threshold for relevance (~0.15).

## I.VIII sentiment-sentit

It is a plugin which performs a Machine Learning algorithm called SentIt that returns the polarity of tweets (1: positive, 0:neutral, -1:negative).

At the end of the execution, we can see the results through the View panel. We can view pie charts, histograms and word clouds based on the viewpoint we focus on.