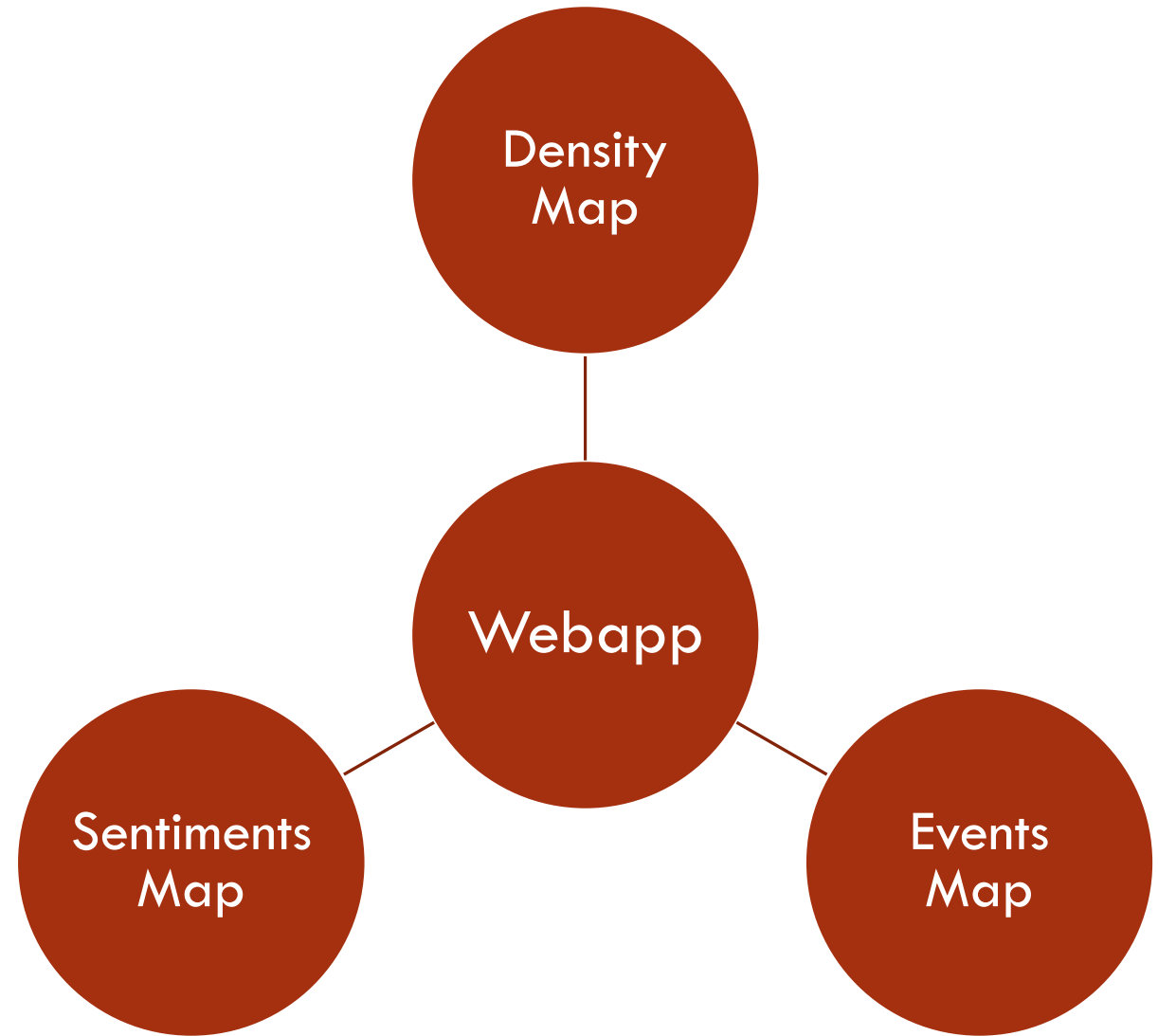


SwissTweets

VISUALIZATION TOOL FOR GEOLOCATED TWEETS

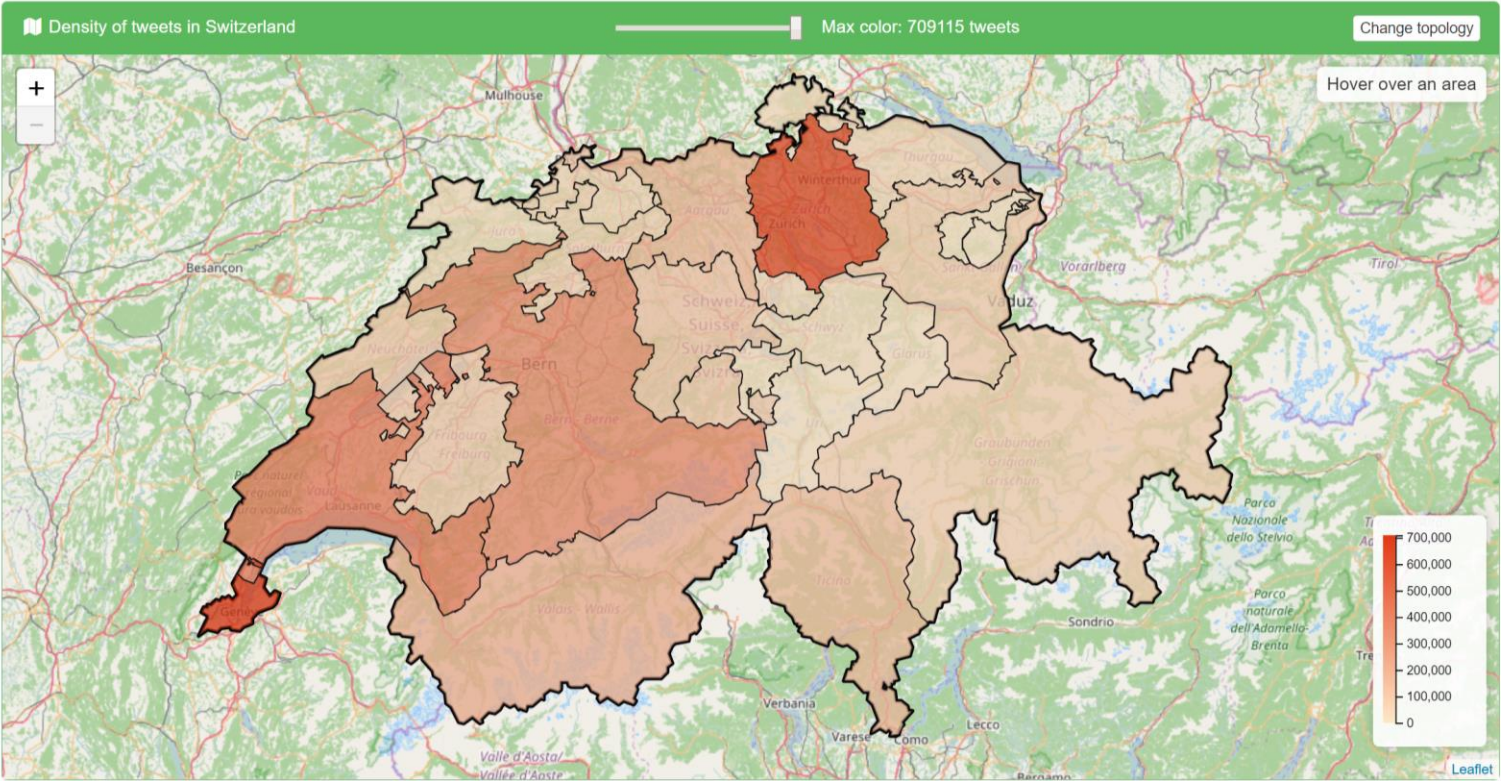
MAIN GOALS

- ↔ Interactive map
- ↔ Interactive timeline
- ↔ Intuitive interpretation

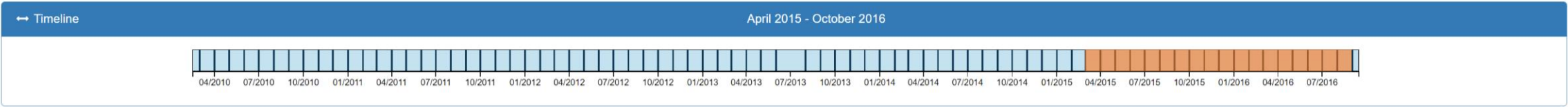


- Density map
- Sentiments map
- Events map
- About

Density map



Information

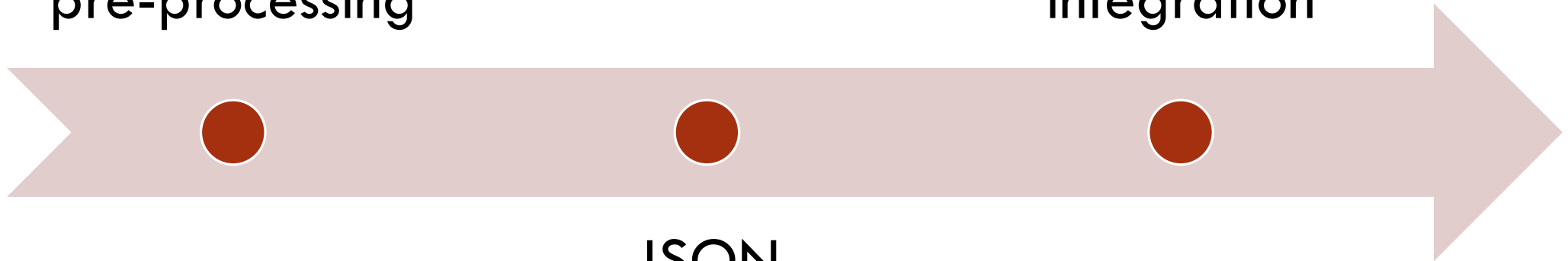


WORKFLOW

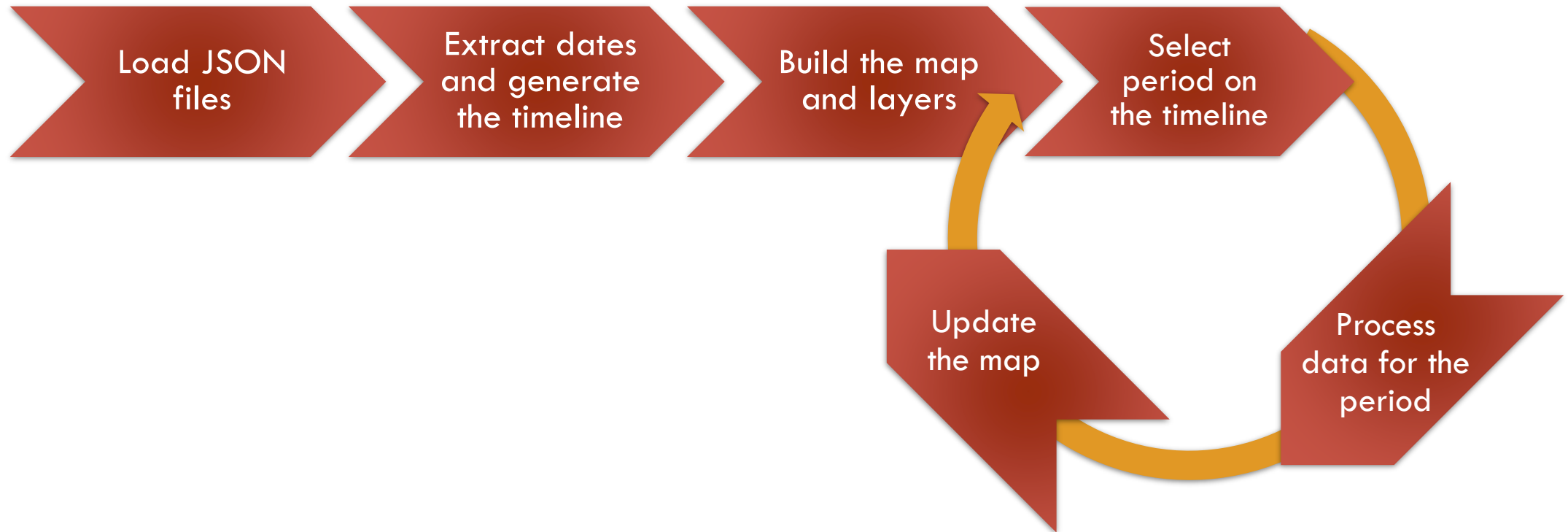
Data
pre-processing

Webapp
integration

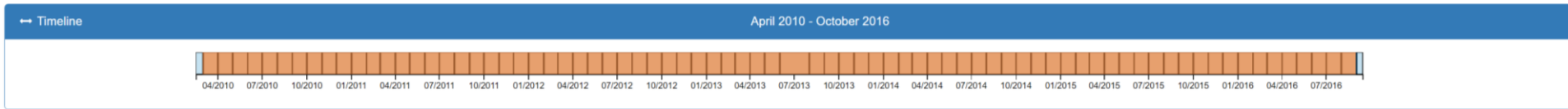
JSON
creation



WEBAPP — GENERAL PIPELINE

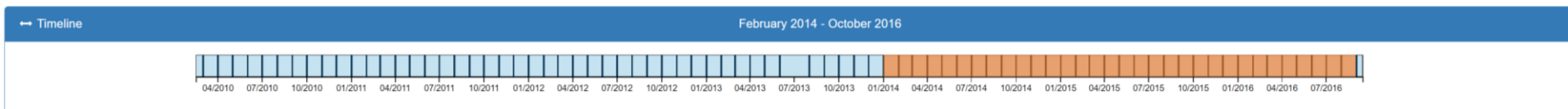


WEBAPP — INTERACTIVE TIMELINE

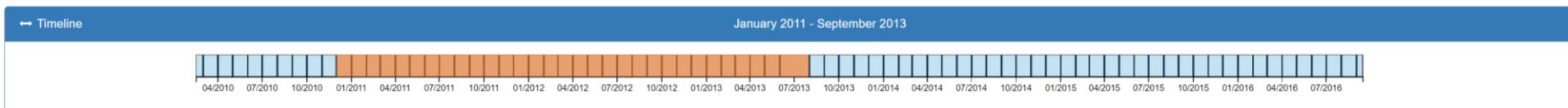


The user can change the **range of dates** concerning the displayed data by:

↔ **adjusting the width** of the orange bar



↔ **horizontally dragging** the orange bar



WEBAPP — TECHNOLOGIES

- ↔ JavaScript/HTML5
- ↔ GeoJSON/TopoJSON
- ↔ LeafletJS (map with layers)
- ↔ D3.js (timeline and legends)
- ↔ Bootstrap

DATA DESCRIPTION

↔ **id** – the integer representation of the unique identifier for the tweet

↔ **createdAt** – the UTC time when the tweet was created

(timestamp in the format 0000-00-00 00:00:00)

↔ **text** – the actual UTF-8 text of the status update

↔ **placeLongitude*** – the longitude of the place the tweet is associated to

↔ **placeLatitude*** – the latitude of the place the tweet is associated to

* we decided to use these fields instead of **latitude** and **longitude** given the fewer number of NaNs

DENSITY MAP

Swiss Tweets v1.0

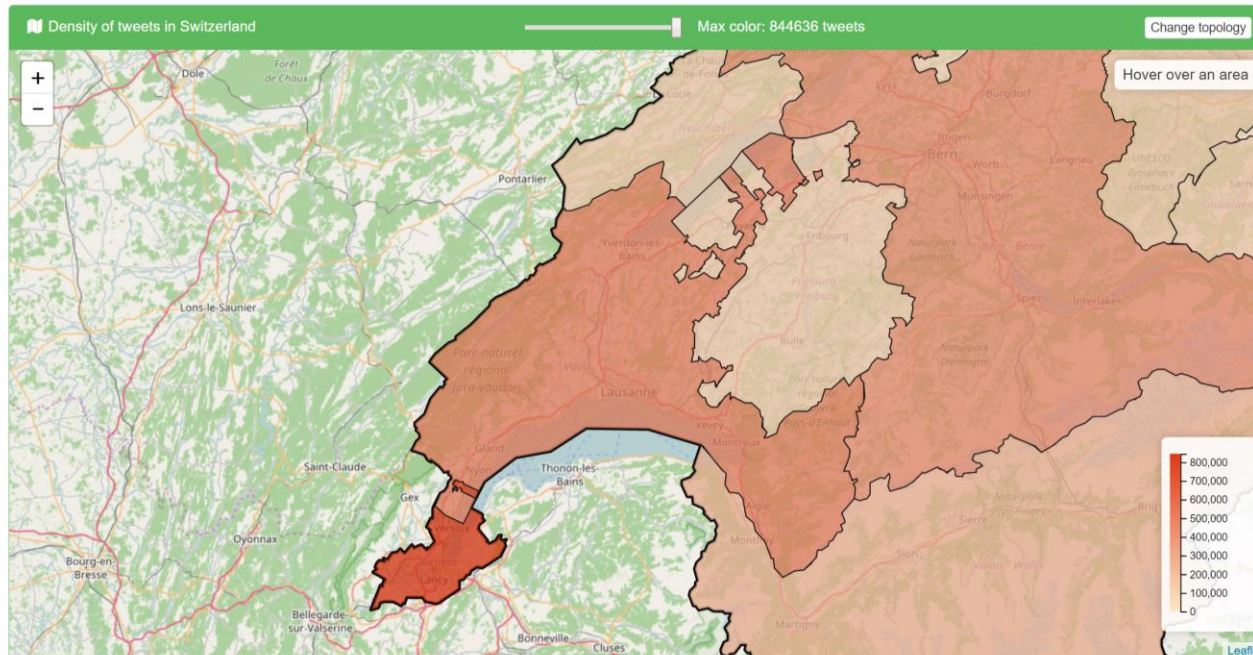
Density map

Sentiments map

Events map

About

Density map



Information

Name: Vaud
Number of tweets: 417397

Timeline

January 2015 - October 2016



DENSITY MAP — OBJECTIVES

- ↪ Distribution of tweets per canton/municipality
- ↪ Identification of regions where the users are more active on the platform
- ↪ Potential areas where events might take place

DENSITY MAP — PRE-PROCESSING

- ↔ Only consider tweets with **valid values** (**id**, **createdAt**, **placeLongitude** and **placeLatitude** fields)
- ↔ Restrict the **date range** (from 2010 to 2016) and convert the dates to **UNIX time**
- ↔ Get the **cantons IDs** and the **municipalities IDs** from the **GeoJSON** files
- ↔ Restrict the tweets to those which are associated with a **place in Switzerland**
- ↔ Backbone of the JSON files: number of **tweets grouped** by **date** (year + month in UNIX time) and by **canton/municipality**
- ↔ By the end of the pre-processing step we have **12.705.241 tweets***

* most of the tweets that are eliminated aren't associated to a place inside Switzerland

DENSITY MAP — VISUALIZATION

- ↔ Two **levels of granularity** (cantons and municipalities)
- ↔ The **color** (linear scale) is given by the **number of tweets** in the regions
- ↔ **Information** box gives the **name** of the **canton/municipality** selected and the **number of tweets** in the region
- ↔ Set a **maximum number of tweets** and **auto-adjust** the **color scale**
- ↔ Interactive **timeline** (the data that is shown is automatically updated)

SENTIMENTS MAP

Swiss Tweets v1.0

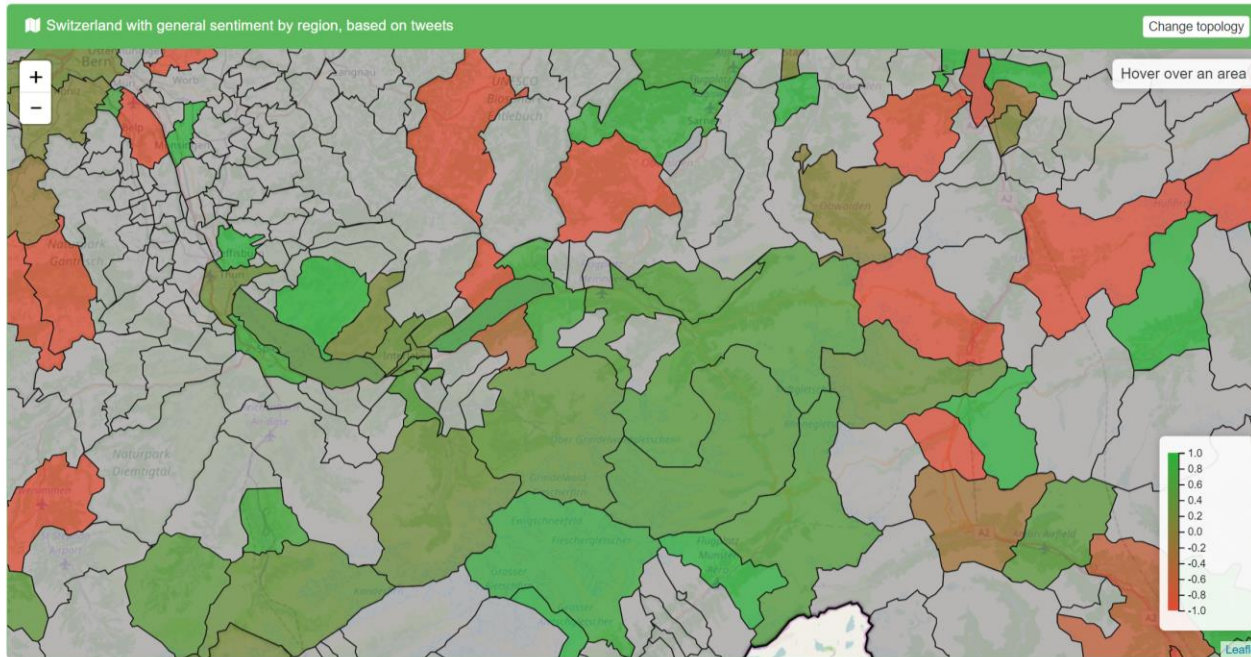
Density map

Sentiments map

Events map

About

Sentiment map

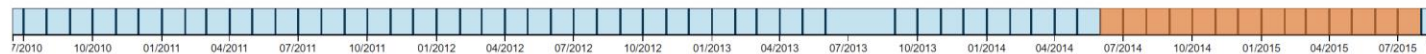


Information

Name: Grindelwald
Sentiment: 0.41 (-1 = bad, +1 = good)

Timeline

July 2014 - September 2015



SENTIMENTS MAP — OBJECTIVES

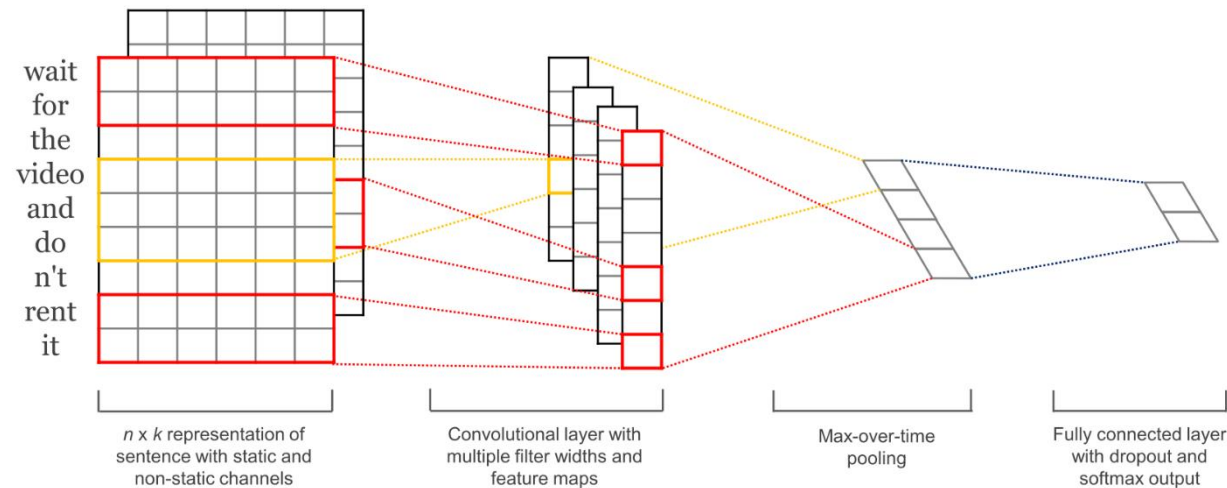
- ↔ Average sentiment associated with tweets per canton/municipality
- ↔ Identification of regions where the users tend to be more positive/negative
- ↔ Potential relationship between the sentiment of the tweets in a region and the kind of events found in the same area

SENTIMENTS MAP — PRE-PROCESSING (TRAINING)

- ↔ Load positive/negative sentences from raw training set
- ↔ Clean the text
- ↔ Pad each sentence to the maximum sentence length (padding sentences to the same length helps to batch the data efficiently)
- ↔ Build a vocabulary index and map each word to an integer between 0 and 18.765 (the vocabulary size). Each sentence becomes a vector of integers.

SENTIMENTS MAP — CNN

Sliding over 3, 4 or 5 words at a time + Dropout + Max Pooling Layer + Softmax Function



Training Score: 88%

SENTIMENTS MAP — PRE-PROCESSING (TEST SET)

- ➡ Split the dataset into 2MB files
- ➡ Filter the tweets and only take the ones written in English
- ➡ Clean the text from the filtered tweets

SENTIMENTS MAP — VISUALIZATION

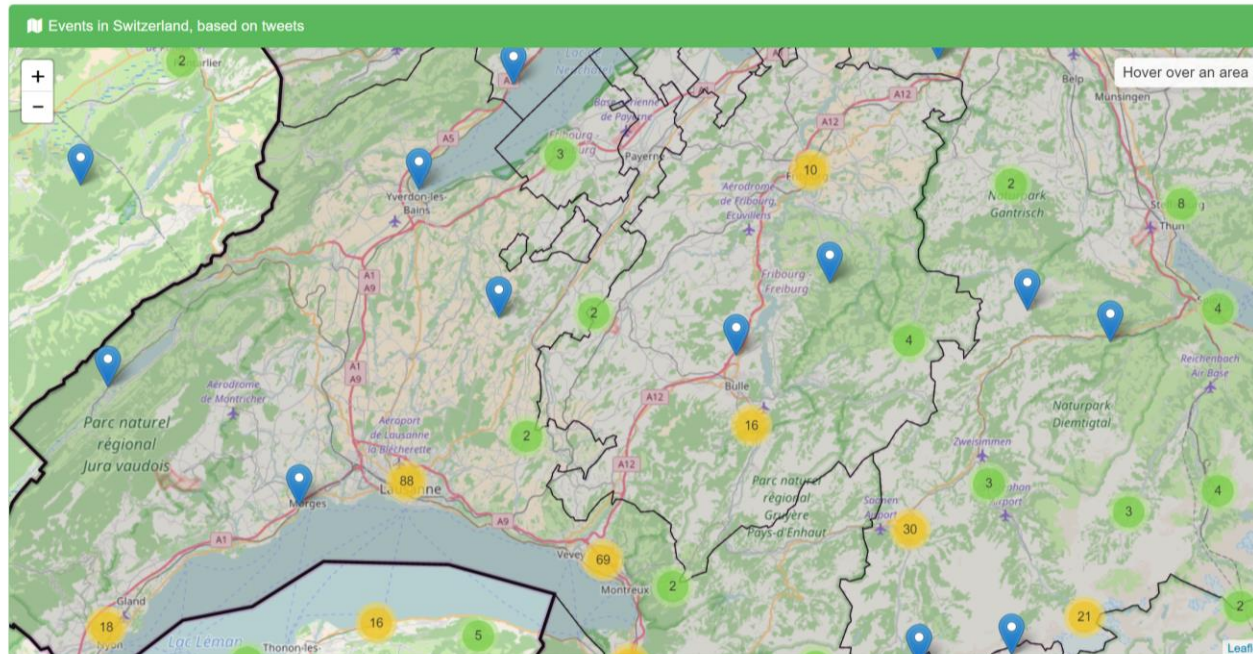
- ↔ Two **levels** of **granularity** (cantons and municipalities)
- ↔ The **color** (linear scale from red to green) is given by the **average sentiment** of the **tweets** in the regions
- ↔ **Information** box gives the **name** of the **canton/municipality** selected and the **average sentiment** of the **tweets** in the region
- ↔ Interactive **timeline** (the data that is shown is automatically updated)

EVENTS MAP

Swiss Tweets v1.0

- Density map
- Sentiments map
- Events map
- About

Events map



Timeline

January 2015 - October 2016



Information

Name: pmi
Date: 29 Jan 2015
Number of tweets: 12

List of tweets

- Ready for the @PMI_Switzerland event with at Ferring #PMI #PMICH #PMOT <http://t.co/TnT9sllWaq>
- Full house for @PMI_Switzerland event at Ferring tonite #PMI #PMICH <http://t.co/QWelab6YpU>
- Congratulations @celine_janvier and @PMI_Switzerland for excellent event! #PMI #PMICH
- Andrew Sadler, Portfolio Manager at Ferring, launching the @PMI_Switzerland #PMICH #PMI <http://t.co/VQ16Nriqy>
- Eli Fumoto presenting on #Creative #Leadership @PMI_Switzerland event at Ferring #PMICH #PMI <http://t.co/hS0o40J3DI>
- Eli Fumoto transforming #uncertainties into #opportunity @PMI_Switzerland event at Ferring Pharmaceutical #PMI #PMICH <http://t.co/3ESmiTCgKA>
- "#ProjectManagement is a #creative #profession" says Eli Fumoto @PMI_Switzerland YES!!! #PMICH #PMI <http://t.co/X3KskLz1P>
- @PMI_Switzerland It's both of course!!! #pmich #PMI
- Eli Fumoto presents 7 habits of creative #ProjectManager @PMI_Switzerland #PMI #PMICH <http://t.co/OXye3gl16x>
- 7 habits of creative #ProjectManager continued with Eli Fumoto @PMI_Switzerland #PMICH #PMI <http://t.co/SC6EVUGW2a>
- Distribution of Creative tools by Eli Fumoto @PMI_Switzerland #PMICH

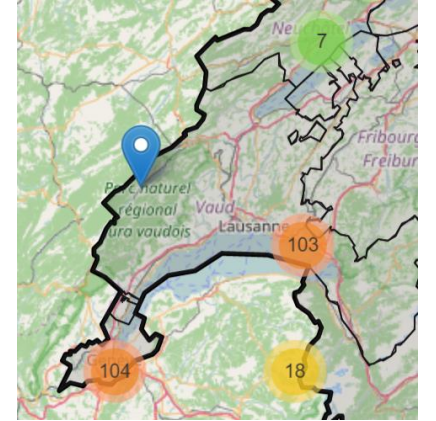
EVENTS MAP — OBJECTIVES

- ↔ Distribution of events across Switzerland
- ↔ Temporal evolution of the events' distribution in Switzerland
- ↔ Keywords related to the events (hashtags)
- ↔ All this information is shown at the same time, giving the user an interactive way of manipulating the data he/she wants to visualize

EVENTS MAP — PRE-PROCESSING/DETECTION

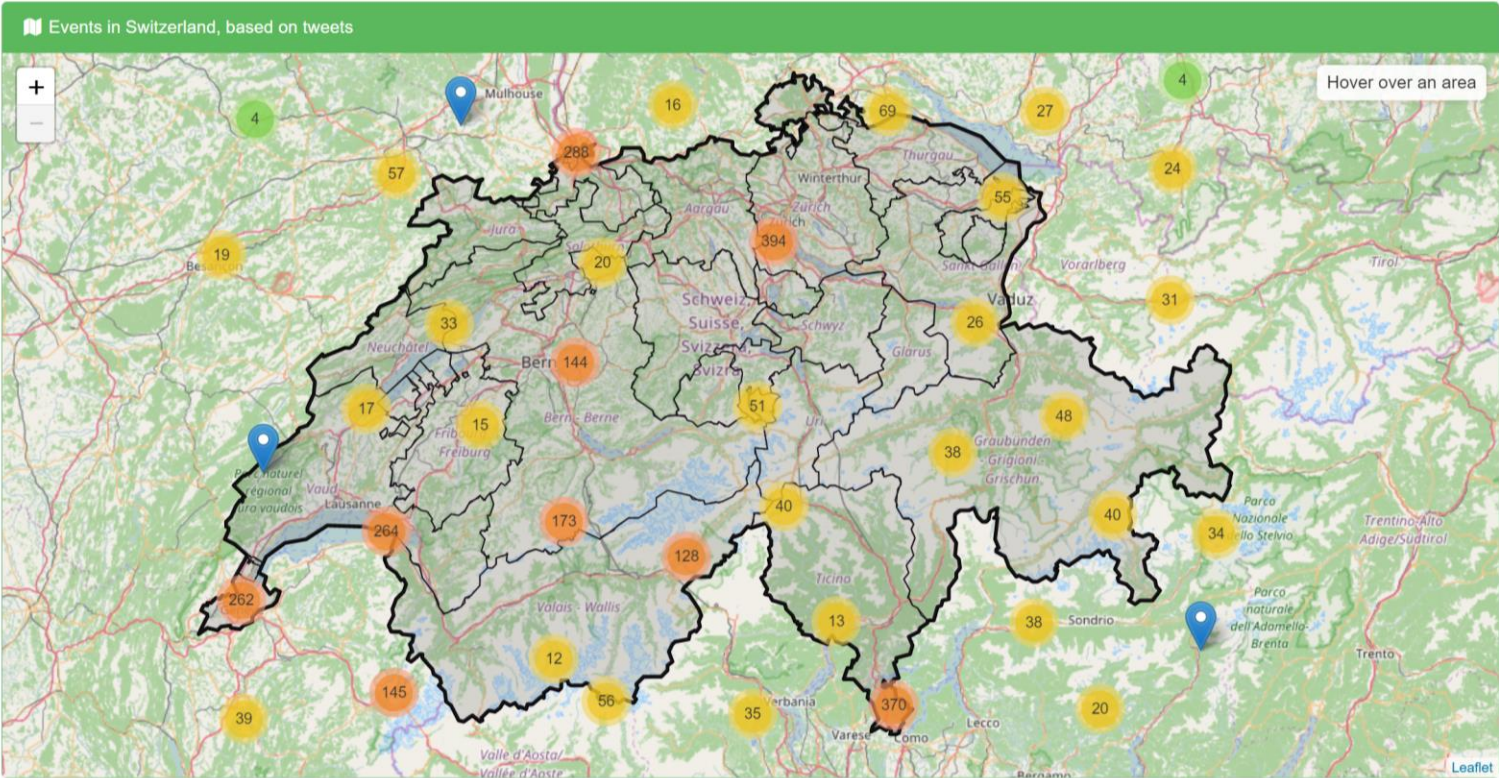
- ↪ **Teamed up** with another group (responsible for the pre-processing of the data and the event detection using a Machine Learning approach)
- ↪ JSON file following a pre-defined structure
- ↪ Backbone of the JSON file: **events grouped by date** (year + month in **UNIX time**)

EVENTS MAP — VISUALIZATION



- ↔ Two **levels of granularity** (groups of events and single events)
- ↔ **Single events** are depicted by **blue markers**
- ↔ For **grouped events**, the number and the **color** (green-yellow-red) of the **circle** indicates the **number of events** grouped together
- ↔ **Information** box gives the **name** of the **event** selected, the **day**, the **number of tweets** and the **tweets** themselves associated with the event
- ↔ Interactive **timeline** (the data that is shown is automatically updated)

Events map

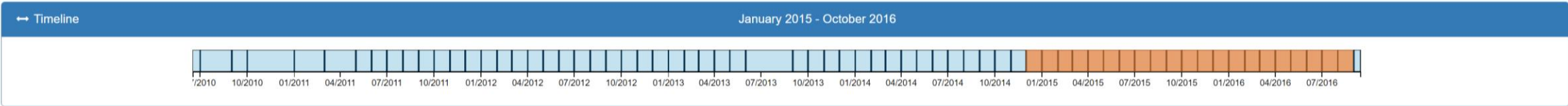


Information

Name: jaegerlecoultre
Date: 20 Apr 2016
Number of tweets: 7

List of tweets

- Scheletrato e inciso a mano... #jaegerlecoultre @ Jaeger-Lecoultre <https://t.co/CSLc6sG56Y>
- Ripetizione in stile van Gogh #jaegerlecoultre @ Jaeger-Lecoultre <https://t.co/ERbfaPdrMy>
- Incastonatore #jaegerlecoultre @ Jaeger-Lecoultre <https://t.co/DWY1vvLFaj>
- Incisore #jaegerlecoultre @ Jaeger-Lecoultre <https://t.co/JaVsKge7UG>
- Micropittura #jaegerlecoultre @ Jaeger-Lecoultre <https://t.co/li8deiLUll>
- Ripetizione a minuti #jaegerlecoultre @ Jaeger-Lecoultre <https://t.co/wgW6b3dcUp>
- Cronografo monopulsante con minuti digitali #jaegerlecoultre @ Jaeger-Lecoultre <https://t.co/O4ShfmPe3N>



Inês Valentim

`ines.rentevalentim@epfl.ch`

Pierre Colombo

`pierre.colombo@epfl.ch`

Syméon del Marmol

`symeon.delmarmol@epfl.ch`

Supervised by **Professor Michele Catasta**