



SEBA EXERCISE 2

- Ozel Christo
- Lisandro Pablo
- Hafsa Asif
- Urmimala Chakraborty
- Sangkyun Kim

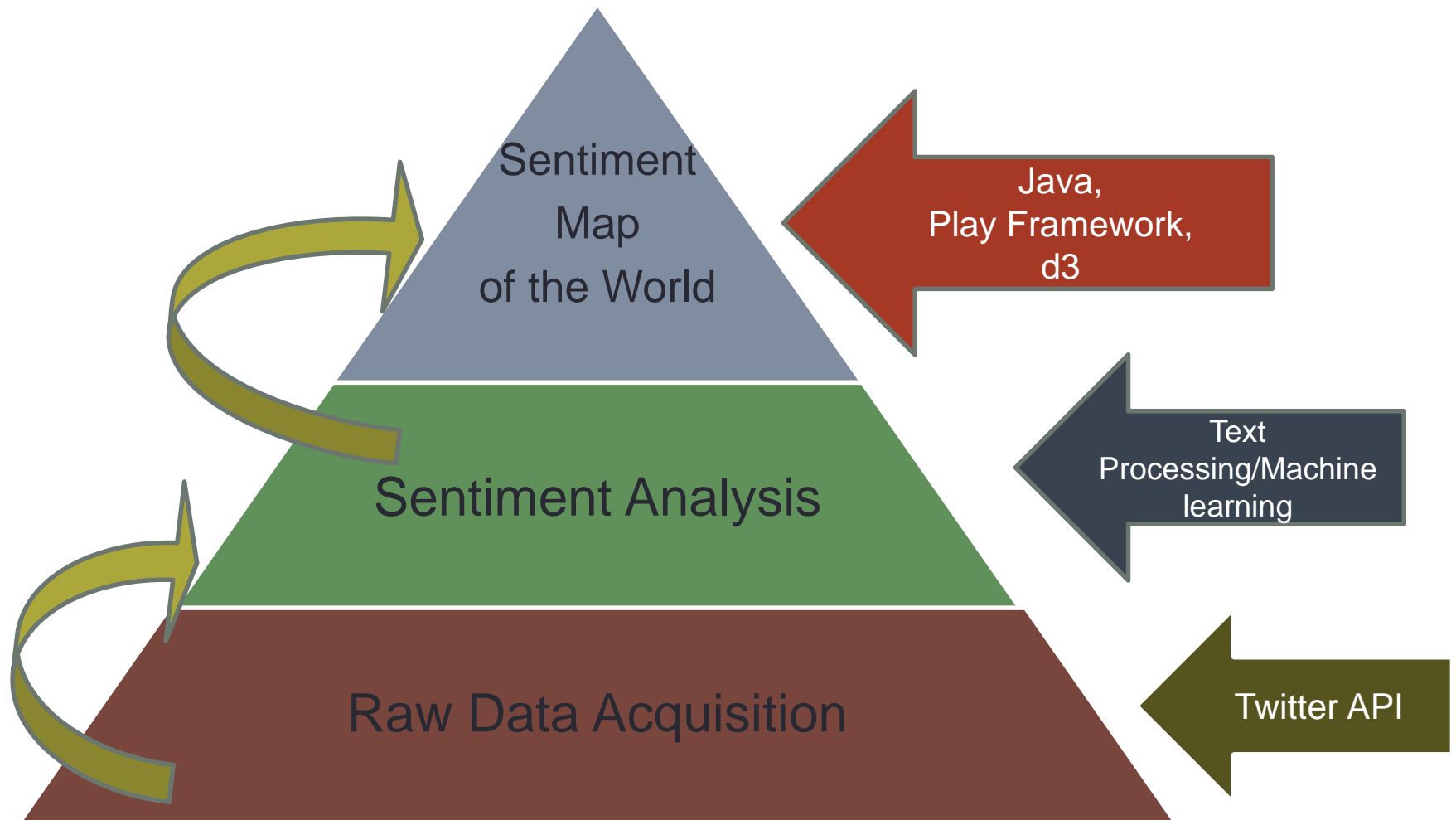
A Quick Review:

- Our application binds sentiment data to a world map
- **Specifically :**
 - our application takes geo-tagged tweets from around the globe,
 - batches them together based on geography
 - uses sentiment analysis algorithms to detect the polarity of the content
 - then color-maps this sentiment to the region of the world it came from.
 - The end result would be a sentiment heat map of the entire world.

What we have so far...

- ✓ Acquired text based twitter data
- ✓ Applied our machine learning algorithm ~80% accuracy
- ✓ Used the play framework and the d3 libraries to produce a basic heat map based on static data

Our Framework:



Raw Data Acquisition

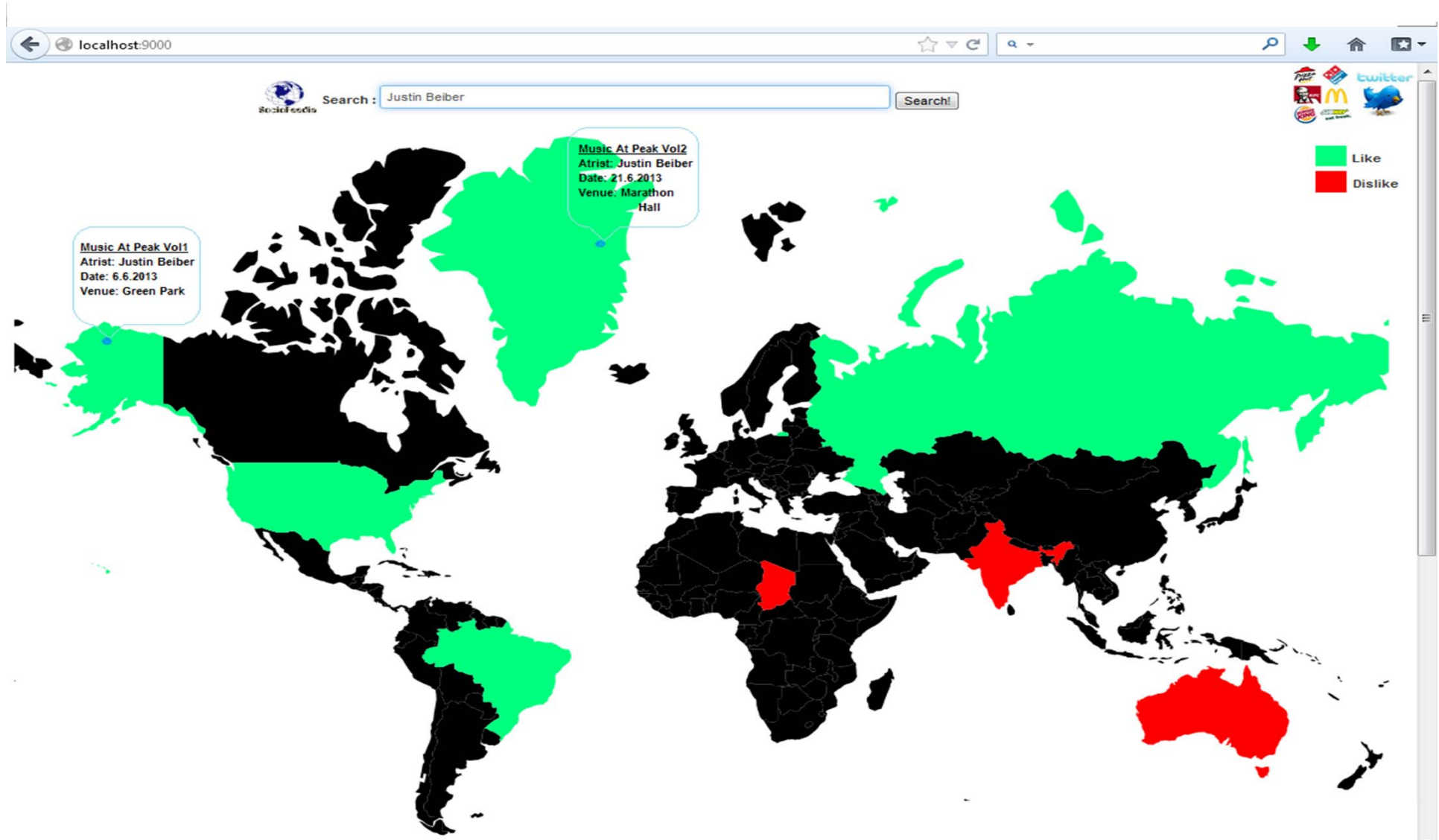
```
{"created_at":"Tue May 21 21:38:35 +0000 2013",...,
```

```
"text":"Bayern munich is the best football team in the history. I've  
never seen this huge success before",
```

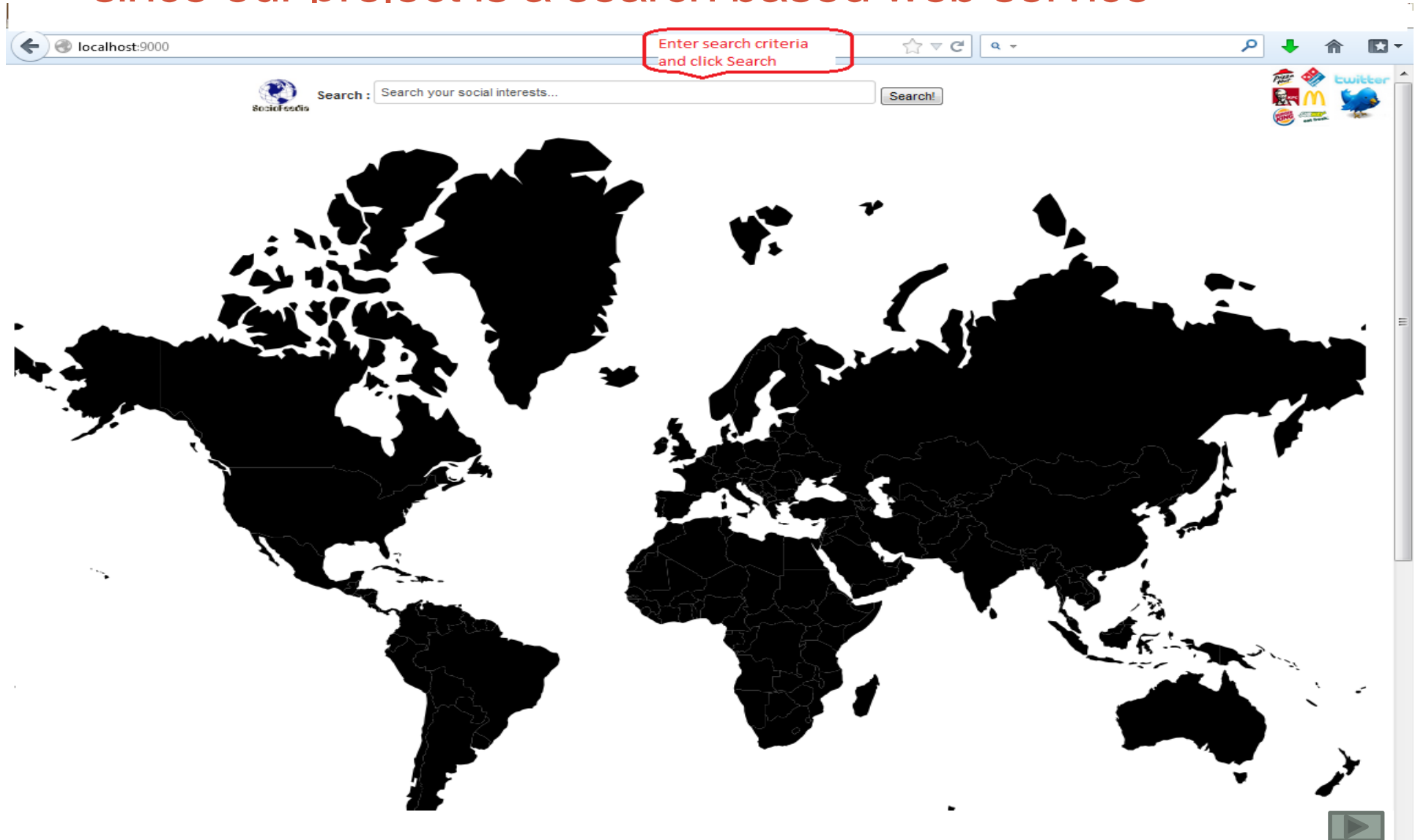
← The data, which we retrieve for the sentiment analysis.

```
..., "possibly_sensitive":false, "filter_level":"medium", "lang":"en"}
```


Geo-Mapping of Sentiment

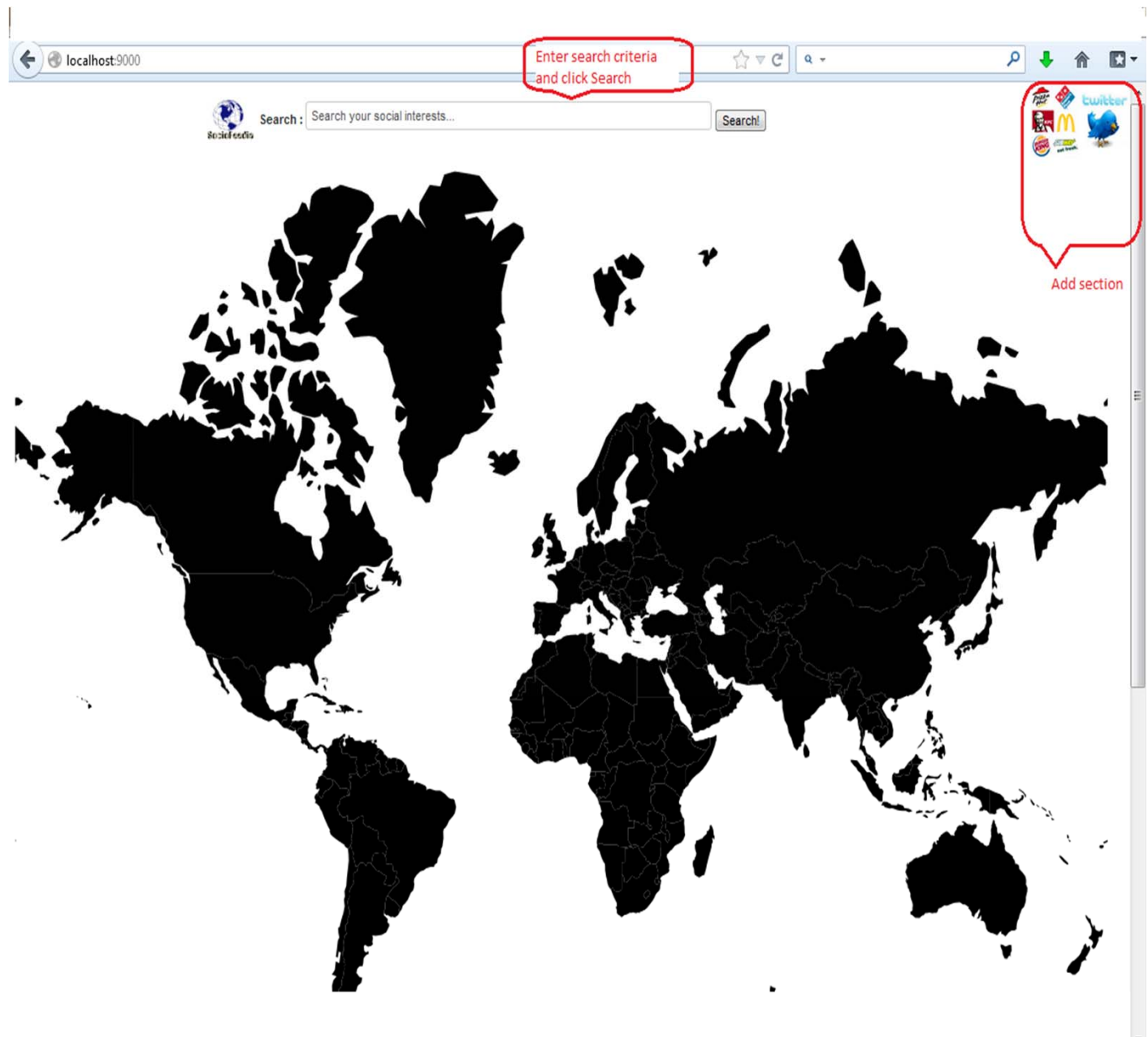


Navigation: We chose **J2** as our web design pattern since our project is a search based web service



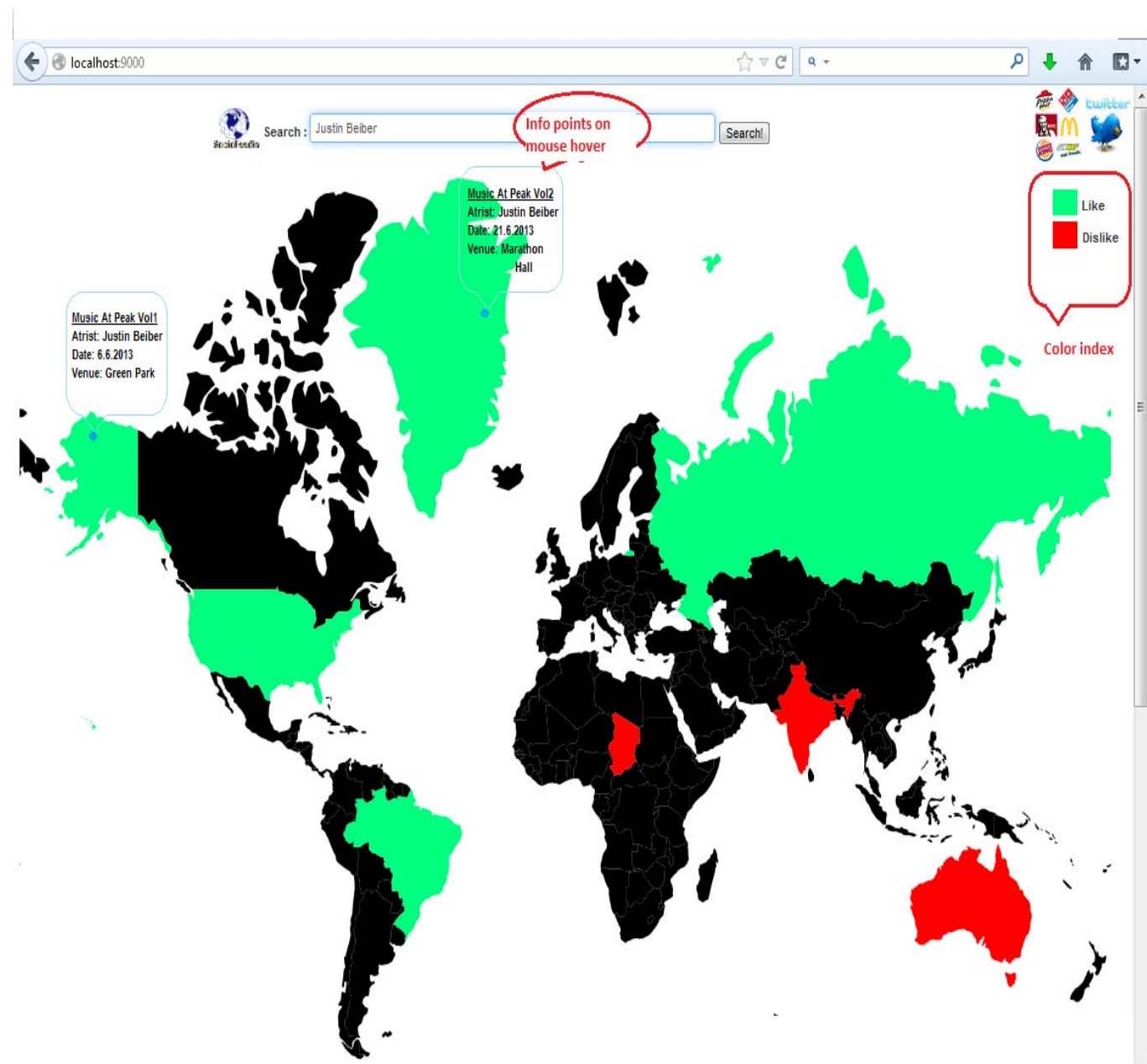
Managing Content

- Search option is always in the top centre position in each page as it is the primary action to be performed by user
- The map occupies the central position
- The advertisements possibly on the upper right corner



Contd..

- On Search result page Info points are viewable on mouse hover over highlighted areas
- The page has a color index on the upper right below ad section



Building Trust and Credibility

- Transparency on what data we use and how we use that data
- Experiential credibility as our application leads users to new knowledge discovery
- Once our application has a rich enough sentiment tapestry, we will introduce “sign in” for a more personal experience
- Perpetual improvement of algorithm and computational methods
- Localization

Current Challenges

- Making full use of the twitter streaming api in collecting and streaming raw data
- Improving the sentiment analysis algorithm (currently at ~80% accuracy)
- Overcoming memory limitations encountered when using large datasets
- Creating efficient back-end to front-end infrastructure so that raw data is converted to sentiment in a seamlessly in real-time
- Finalizing the name of our application!

Next Steps

- Creating a fine-mesh grid over the map for a heat-map visual effect
- Full back-end to front-end integration
- Making the map interactive and dynamic (zoom, click, search)
- Extending to sentiment algorithm to other English-speaking countries



Questions

