Project Plan

Team Name: Politician 360

1. Vision

Our goal is to provide a platform with trustful information about the politicians to people. People can have all the information at one place, so that they can know more about the individual politician and compare with others.
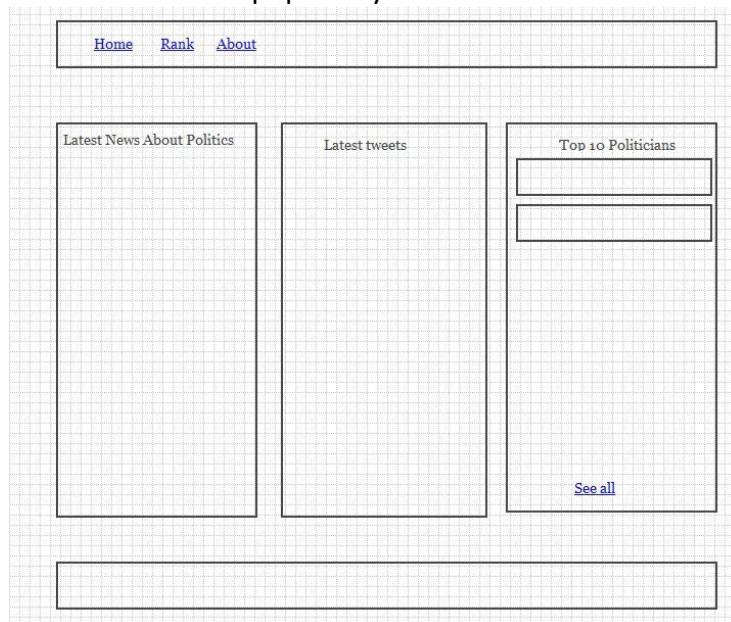
- What we are doing?
  - We are building a web app that will give trustful information about Irish politicians through a ranking system in which it ranks them by popularity, sentiment analysis or corruption mentions. The system will collect and mine the information from irish newspapers and twitter.
- Why we are doing it?
  - We think that people would like to know better what the news and social networks are saying about Irish politicians. We believe that a rank could be a good indicator of which politician is most popular and what the general opinion thinks about them, specially with the sentiment indicator. For instance, it is essential to know if a politician has any relation with corruption. It could be that he is fighting corruption or doing corruption. Just the fact that he is mentioned in corruption news is a starting point to know what the politician is doing regarding that.
- Who will use it?
  - The system is designed for anyone interested on Irish Politics and willing to have a centralized place where they can easily have access to news related to the Irish politicians and extra information such as a rank of the politicians. Basically, the users could be voters, news agencies, politics students and researchers.
- How will they use it?
  - Users could use the Irish politician latest news and tweets section for having a easy way to get this specific information instead of looking at the web trying to find the related news by themselves
  - Users could use the TOP 10 popularity rank to easily know what are the most popular politicians.
  - User could use the rank by popularity to compare politicians and know which politician is being talked about more or less on the news
  - Users could use the rank by sentiment analysis to compare politicians and know which of them has news and tweets talking about him negatively, positively or neutrally. That could lead users to understand if a politician is doing good or bad.
  - Users could use the politician information page to know more about a specific politician.
    - What news and tweets have a relation with him?
    - How does he look like(Photo)?
    - What is his party and constituency?
    - Is he a senator or deputy?

2. Minimum Viable Product

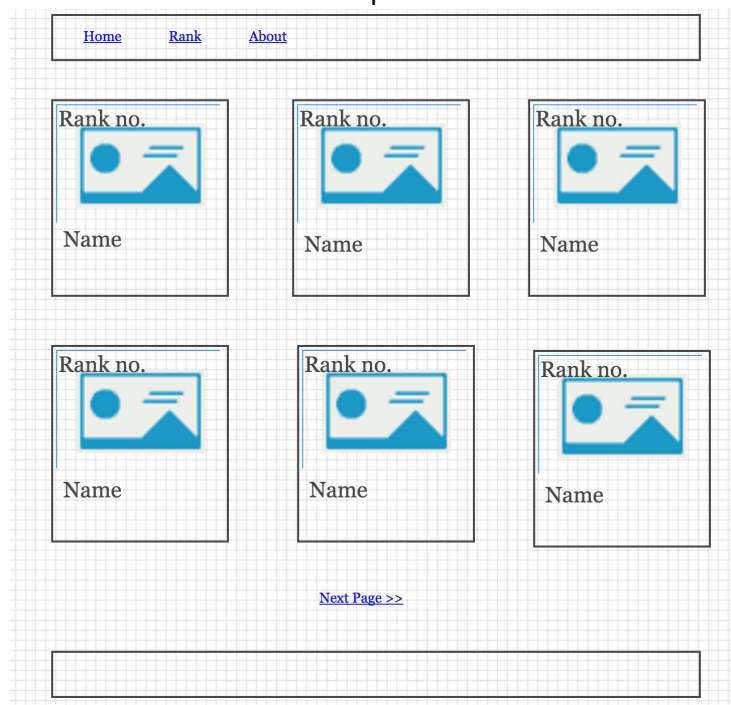Our MVP is a web app which will have the following sections
- ● Home page
  - ○ It will show the Latest news and tweets about the Irish politicians and the TOP 10 popularity rank

- ● Rank page
  - ○ It will show a rank of the politicians by popularity, sentiment analysis and number of corruption mentions

- ● About page
  - ○ It will show a description of our project.
- ● Politician individual page

○ It will show details about the politician such as, photo, party, constituency, related news and tweets

	Full Name	
[photo]	Party Constituency	
	Gender, Age	
	Website:	
	Email:	
	Short Description	

Tweets by @name	Articles related
	Link 1
	Link 2
	Link 3
	Link 4
	More ...

User Stories

● As a user I want to see the latest news and tweets about politicians so I get it in one place for easy access
  ○ Given a user goes to the politician360 website
  ○ When he goes to the homepage
  ○ Then he will see the latest news and tweets about the politicians
● As a user I want to see the TOP 10 most popular politicians so I can know who they are
  ○ Given a user goes to the politician360 website
  ○ When he goes to the homepage
  ○ Then he will see a list with the TOP 10 most popular politicians
● As a user I want to see a rank of all politicians by popularity(number of articles and tweets), sentiment(negative, positive and neutral) and corruption(number of articles and tweets talking about corruption)
  ○ Given a user goes to the politicians360 website
  ○ When they click on the Rank link
  ○ They will see a page with a rank of all politicians with default sort by popularity
  ○ And an option to rank them by sentiment or corruption
● As a user I want to see details about politicians so I know more about each of them
  ○ Given a user goes to the politician360 rank page
  ○ When he selects a politician
  ○ Then he will see more details about him
● As a user I want to see the news and tweets related to a politician so I can trust on the rank info
  ○ Given a user goes to the politicians360 rank page

      ○ When he selects a politician
      ○ Then he will see the news and tweets related to the politician

System Design

We are building our system using Docker which allow us to separate it into containers. For instance, we will have a container for the front end, another for the backend and another for the database. All the different containers will be managed by docker compose. In this matter we will leverage an open source project seed which already has a working system with django, postgres and angular.

Front-end
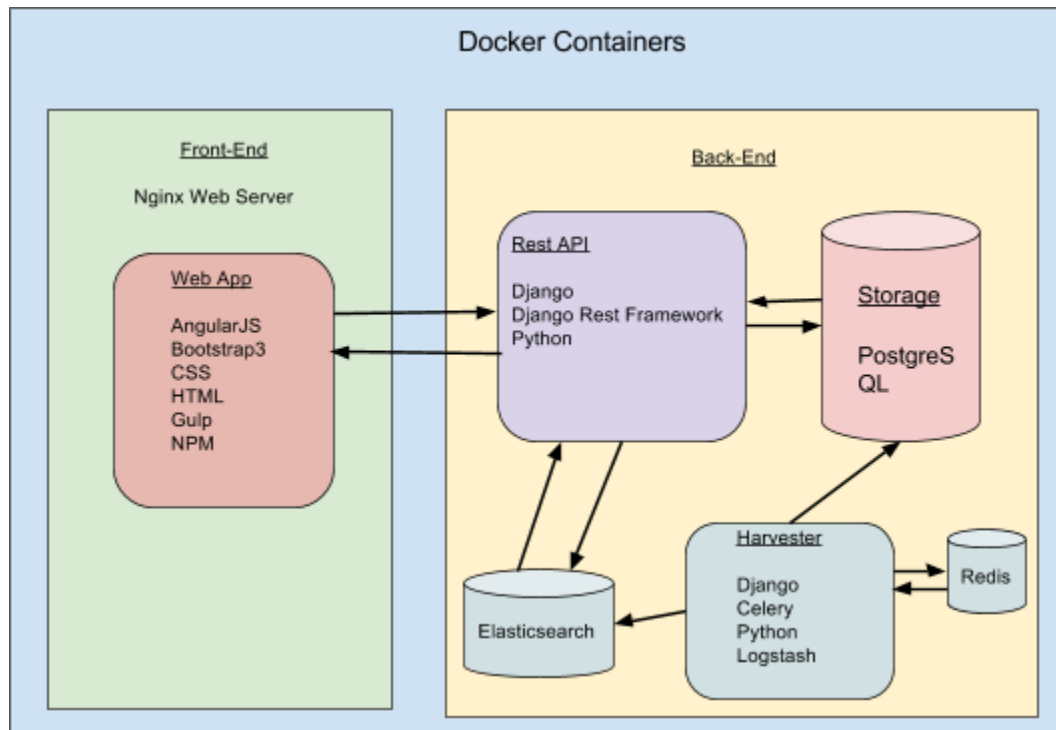The front-end will be implemented using the following technologies
- Javascript, HTML and CSS
- Bootstrap - UI front-end framework
- AngularJs - a javascript MVW framework
- NPM - javascript package manager
- Bower - web app dependency manager
- Gulp - build task manager

Back-end
The back-end will be implemented using the following technologies
- Python
- Django with Restful framework
- PostgreSQL
- Celery
- Redis
- Logstash
- Elasticsearch

The following diagram show a high level design of the entire system with the communication between the parts.

Basically, our front-end app will communicate with our back-end using JSON through a rest api which will hold the business logic of the system. This rest api will also consume the data from our database and index system. Both will be populated by a harvester system which will be comprised of 2 modules, one for reading and mining news articles and another for reading and mining tweets. The harvester jobs will run as often as possible through scheduled tasks so we keep the database and index up to date. We will make sure we respect the twitter api limits for instance.

Data sources

As a start point we have a list of elected politicians from 2016 elections. This was used on the Insight4News application. The system will then use that list to get other info about the politician from the IEOD (Irish Election Open Database) candidates API.
For each politicians we will need information from news articles and twitter. We have considered the main Irish newspapers which cover politics. The list was taking from FUJO (The Institute for Future Media & Journalism)
- RTE NEWS
- IRISH EXAMINER
- THE IRISH TIMES
- IRISH DEPENDENT
- THE JOURNAL.IE
- NEWSTALK

First of all the system will read and store their RSS for future reference. Then it will go through each article talking about a politician. Here the system will use context techniques for entity disambiguation. Once the article is found to be related to a politician it will then check what is its sentiment and if there is anything talking about

corruption on it. It will also increase the count of articles for the politician so it can be used on the popularity rank.

Moreover, we will also use Twitter data through the streaming and search APIs. As the search API only accepts 150 requests every 15 minutes. We will need at least two twitter accounts for searching tweets for all the elected Irish Politicians (218). The idea is to make 218 search requests every 15 minutes. All tweets are going to be indexed so we can easily mine them.

Evaluation

We will measure how useful the application is for the users. The idea is to assess people's knowledge about Irish Politicians when using the open web or our system. Our system should be able to give more info about a politician in a certain period of time, such as 10 minutes, than looking on the web. We can also add social network sharable links and count the number of shares.

We will also validate if our system is identifying and disambiguating entities accurately.

3. Project Management

We are developing the system on Scrum agile methodology. We are running on one week sprints with daily standups meetings on hangout, sprint planning and sprint retrospection where every member is assigned a task for the week and he/she is required to give update on his/her task assigned and discuss with other team members the problems he/she is facing and helping others to solve the issue.

For task management we are using Trello and for communication we are using slack where everyone doing the work updates on trello and the same is communicated to everyone in the group and is finally validated at the completion of the task.

We are using google docs for collaboratively writing the blogs, reports and presentations where the person in turn creates his blog using google docs. It helps sharing with other team members so that everyone has the visibility of what the individual is presenting and if any modification is required it can be passed on.

During all this Project management process,every task is managed by every team members and everyone is responsible for all the task. The task is finally placed under done category only when it is validated thoroughly by each and every members of the team.