Project: "Social Media Sentiment Monitor" [about SM Caen]

Description:

The project consists of two (sub-)tasks. In a first step, **social Web contents** about SM Caen need to **collected/crawled**. These contents (min 1.000 documents per match day) should cover at least three matches and should be, e.g., acquired from tweets containing an associated hashtag (#). For further analytics, the contents should be parsed and the "meaningful" text extracted in order to "Identify" the overall sentiment (resp. sentiments in different sub-communities).

In a second step, a **Web-based search interface** needs to be developed. This interface should at least support **temporal search and analytics** for three match days (before and after). The analytics module should to this end be capable of visualizing the overall sentiment and/or retrieving the relevant documents, terms and/or entities associated with positive/negative sentiments. Enhancements include but are not limited to the identification of the most important entities, linking and/or incorporating match statistics as well trend analytics.

Specification:

The software you are supposed to develop should be deployed on the Web as a "Web search engine". The interface needs to be productive (fast, usable and "appealing"). The minimum functionality specified above needs to be available, while additional "features" are a surplus. The choice of software is up to you, but if third party software/libraries is/are used, it must be open-source.