Automatic Keyphrase Extraction for News Media

"Want information, not documents"

--User

Supervised By :

Dr. Dhaval Patel

Computer Science Department

Presented By :

Nikita Jain

14535031

News Observatory System for Evolving World Events

News Headlines for World Events

| World | Africa | Australia | Europe | Latin America | Middle East | US & Canada |

Paris attacks: Key questions after Abaaoud killed

By Laurence Peter
BBC News

20 November 2015 | Europe

The Gulf: Palm fronds and shifting sands

- Sheena Bora murder case: Peter Mukerjea arrested; CBI files chargesheet

Nitish still the popular leader

Bihar polls: 10 big India Today-Cicero survey takeaways

Onward robotic soldiers: IIT students pioneer cutting-edge research

icture this: Robots braving bullets while ferrying weapons and ammunition to soldiers on the battle front. Or, a robotic arm resembling the human variety that can work in hazardous areas like blast furnaces. Students at IIT-Roorkee are swotting to turn these ideas into reality.
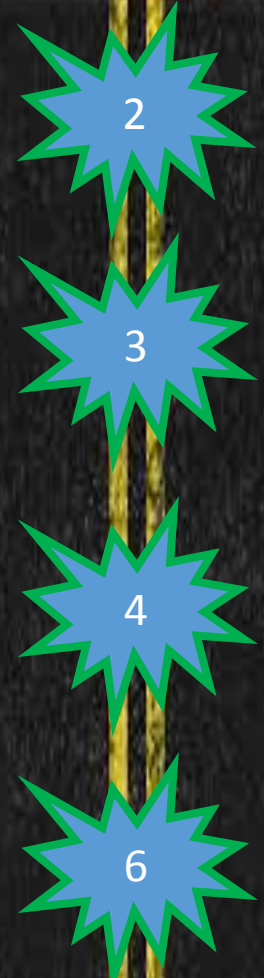
- 1984 anti-Sikh riots: Withdraw Rajiv Gandhi's Bharat Ratna, demands HS Phoolka

Keyphrases for World Events

- **Paris attacks**
- **The Gulf**
- **Bihar Polls**
- **Sheena Bora murder case**
- **Onward Robotic Soldiers**
- **1984 Anti-sikh Riots**
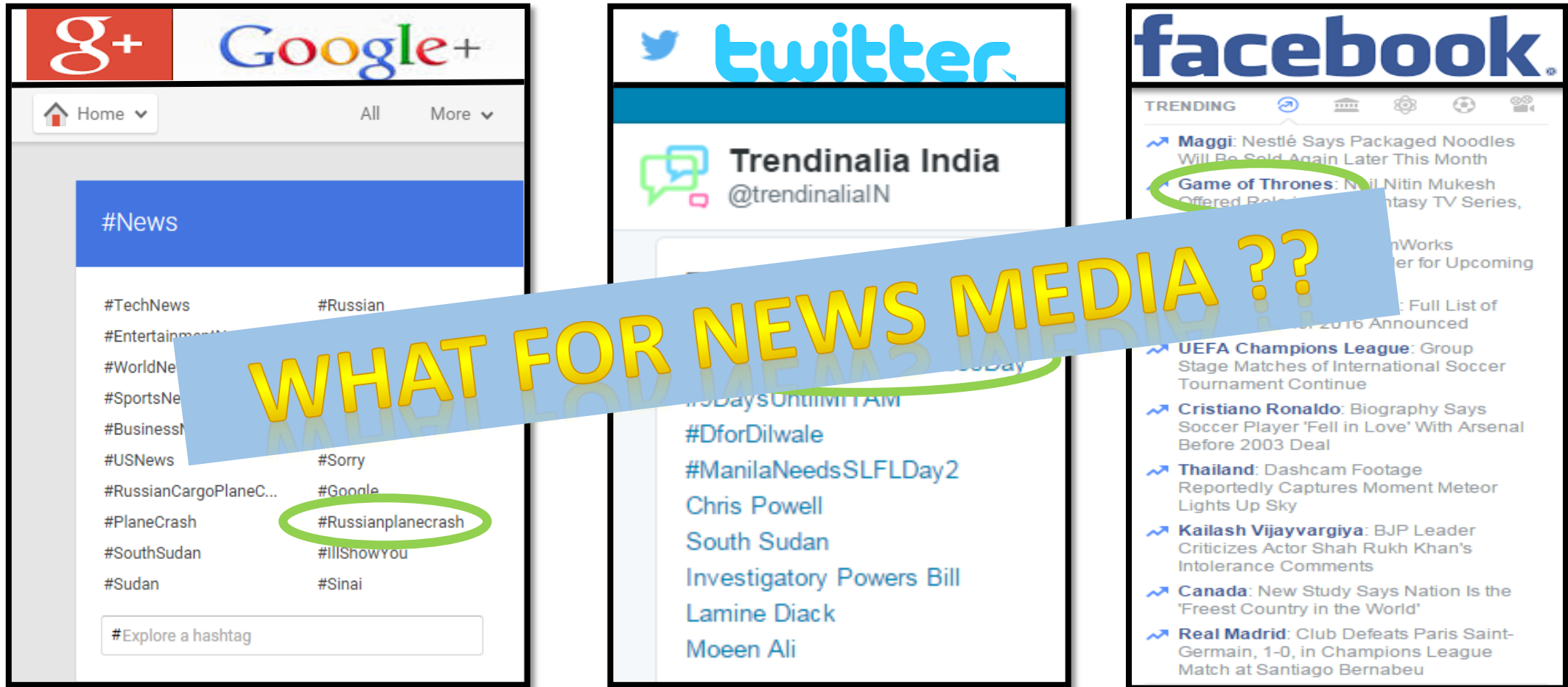
The Roadmap

Automatic Keyphrase Extraction for News Media

- *Group of **WORDS**, i.e., Short in length*
- *Expresses an Important **CONCEPT***
- **NO** structure or grammatical **RULES**

Keyphrases in other social media

Keyphrases used by various social media to summarize their **trending content**

An Example of Keyphrases in News

Keyphrase uses

- Better summarisation
- Better indexing
- Better browsing
- Better metadata construction
- Classification and clustering
- Topic/content identification
- Back of book/document

Some more Applications

Keyphrases extracted from News corpus can be used News channels,

- Make it easy to skim news articles by Highlighting keyphrases.
- Search temporal events by using keyphrases as index terms.
- Refine news queries on search engine by using keyphrases as suggestions.
- Find similar news content by using keyphrases as a similarity metric.

Problem Statement

For a given concept*, mine *Interesting Keyphrases* from *News media information*, where a keyphrase is considered interesting if when it is shown to N persons, and more than N/2 persons find it interesting on the basis of *frequency*, *collocation* and *completeness* of the topic covered by the output set of keyphrases.

Concept can be any name, place, event, organization, etc.