# Context Disambiguation in Social Media

Akshaya Iyengar

akshaya1@umbc.edu

CSEE, UMBC

## The Idea

- Bring together different forms of user-generated content pertaining to a real-world entity or event, and returning it to the user in a unified form.
- Exploit context for more relevant results
- User generated content as a formal source of information

## What is Different?

- Identify entities in queries and online content
- Use context to accurately disambiguate entities
- When in doubt - in Wikipedia we trust
- Faceted browser – browse features of data

## Related Work

- Usage of document level metadata information
- Twitter for disaster and crisis situations
- Category-wise display of search result

## Why do we need Context?

- Identify links between the different forms of UGC, pertaining to the same entity.
- Analyze tags, description or comments - short but limited information gain
- Different sources collectively contain more data.
- Information gained from different source used for better searches

## Hurdles

- Manage and store data from multiple sources
- Data collection  - real-time vs offline corpus
- Understanding un-moderated user content
- Accurate entity recognition, context identification

## Applications

- More relevant search results for each individual network.
- Detecting activity across multiple social networks.
- Rich dataset for opinion sentiment analysis.
- Identify gravity of events as they happen – help plan relief work

## Future Work

- Extend to more data sources.
- Retrieve data on the fly.
- Add an inbuilt module for sentiment analysis
- Validate the data using the dataset or online sources

## References

[1] Amit Sheth and Meenakshi Nagarajan. Semantics-empowered social computing. IEEE Internet Computing, 13(1):76{80, 2009. ",
[2] Susan Dumais and Edward Cutrell. Optimizing search by showing results in context. pages 277{284. ACM Press, 2001.