Automated Analysis of e-Participation Data

Peter Teuff
Peter Parycek
Udo Payer
Outline

- e-Participation data and automated analysis
- Analysis framework
- Live demonstration of implemented framework
e-Participation Analysis

- Costs
- Time
- Domain experts
- Wide range of data sources
Automated Analysis

- Natural Language Processing (NLP)
- Machine Learning
- Artificial Intelligence
- Different areas: security, event correlation, text analysis
The Model

- Use semantic network to store arbitrary information and relations.
- Activate nodes (terms, concepts) and follow relations according to the task at hand, extract Activation Pattern.
- Use Activation Patterns for a wide range of analysis processes.
Framework

- Analysis Module (implements L1-L5)
- Webservice for providing the analyzed data
- Browser application in Flash
Live Demonstration

- Sources:
  - Twitter news headlines extracted from 35 news sources covering 08/14 to 09/01
  - Mitmachen: Austrian e-Participation project
- Why Twitter?
- Temporal analysis: time Line, related terms, clustering
- Semantic relations, search queries
The Core - Activation Patterns

- Take NLP processed data and store information in a semantic network
- Activate nodes and spread activation to related nodes (terms)
- Extract activation values of all nodes and store them in a vector (the Activation Pattern)
- Apply arbitrary analysis techniques (clustering, semantic search queries, ...)

Future

- Use sophisticated lexical parsers
- Concentrate on evolution over time (e.g. how does the relation between terms change)
- Improve by applying to e-Participation projects, security related areas, event correlation
Thanks!

- Peter Teufl, Peter Parycek, Udo Payer
- http://www.iaik.tugraz.at/content/about_iaik/people/teufl_peter/
- http://www.donau-uni.ac.at/egov