Student Projects

Multimedia Information Systems 2 VU (707.025)

(“Visual Analytics”)

SS 2016

Vedran Sabol

Know-Center

April 12th 2016
Lecture Overview

- Motivation and Goals
- Four Project Topics
  - Overall project description
  - Implementation ideas
  - Data set suggestions
- Next Steps
Motivation

- Web is man made but it behaves as a natural phenomenon
  - Complex system: technological and social
- The Web is a technological infrastructure supporting processes of
  - Publishing, linking, connecting, communicating, collaborating etc.
- Result: creation of huge amounts of data
- Web data as object of analysis
  - Knowledge Discovery in the Web (Web Mining): automated analysis
  - Information and Data Visualisation: human visual pattern recognition
  - Visual Analytics: combine algorithmic and visual methods (human in the loop)
Goals

• Learn how to apply Visual Analytics methods in the Web
  ▪ on Web data
  ▪ using Web technologies
  ▪ in selected Web-based scenarios
• Learn about presenting Web data visually
  ▪ Using Web technologies (HTML5)
• to gain insights into
  ▪ Multidimensional data (tables)
  ▪ Recommended (multimedia) resources
  ▪ Sensor and event data
  ▪ Semantic knowledge bases (ontologies)
Projects

• Project topics
  1. Visual exploration and filtering of recommender data
  2. Sensor and time series visualisation
  3. Visualisation recommendation for tabular data sets
  4. Visualisation of semantic networks

• Each group picks one topic
  ▪ Number of groups per topic is limited
  ▪ First come, first served

• Topic registration: per Email to the tutor (b.taraghi@tugraz.at) and lecturer (vsabol@know-center.at)
  ▪ ➔ List your first and your second choice
  ▪ If your first choice is already booked out: you will be notified by the tutor and will have to live with your second choice
Projects

• The four project topics are fixed!
  ▪ Each team must pick one of them
  ▪ (or contact the lecturer directly if you think you have a much better idea)

• Presented implementation ideas are not binding
  ▪ But, they are aligned with the lecture topics

• The listed data sets are suggestions
  ▪ You are free to select any suitable data set for your demo

• You have the choice of
  ▪ implementing your own UI from scratch
  ▪ extending an existing UI (topics 1 and 3), such as the Recommendation Dashboard or VisWizard
Visualisation of Recommender Results
1. Recommender Interfaces

• Recommenders as ahead of time information retrieval engines
  ▪ Recommendations are automatically generated
  ▪ Depending on user’s context (and profile)
  ▪ E.g. what the user is reading in the browser

• Problem: recommendations may not be relevant
  ▪ It is hard to guess user’s needs

• Solution: visual tools for exploring, filtering and specifying interests
  ▪ Ideally Personalised and context-sensitive
1. Recommender Interfaces – Project Ideas

- **Recommendation Dashboard (RD) interface provides**
  - Filtering and bookmarking functionality
  - Views for temporal, geographical, topical and categorical data
- **Extend it with new views visualising e.g.**
  - keyword-relationships based on co-occurrence
  - Image similarity maps etc.
1. Recommender Interfaces – Project Ideas

• The RD **micro-visualisations** show the currently active filter set
  ▪ Temporal, spatial topical, categorical etc.

• Improvements
  ▪ Make micro-visualisations interactive
    • Add zooming, panning, selection etc.
    • Including touch interactivity for the mobile
  ▪ Add new/improved visual metaphors, e.g.
    • Hierarchies and graphs
    • Collection interfaces (topical overview, image browser etc.)
1. Recommender Interfaces – Project Ideas

• Improve the **uRank** topical exploration interface
  - New tag-cloud view for the keywords
  - Replace stacked bar with new document content visualisations
  - Implement a new re-ranking algorithm
1. Recommender Interfaces – Suggested Data Sets

- Scientific and cultural heritage data
  - Returned by the EEXCESS recommender and retrieved directly by the Recommendation Dashboard UI
  - Goodie: 2 integrated test data sets available for offline testing
  - Details to be introduced in the lecture on 19.04.2016
- Europeana data APIs: [http://labs.europeana.eu/api](http://labs.europeana.eu/api)
Visualisation of Sensor Data
2. Visualisation of Sensor Data

• Massive production of sensor data
  ▪ Mobile devices (quantify yourself)
  ▪ Industrial sensors (Industry 4.0): monitoring, prediction etc.
  ▪ Medicine: patient monitoring, brain-computer interfaces
  ▪ Transportation
  ▪ Climate, ...

• Problems to address:
  ▪ Scalability: visualize massive amounts of data (high-frequency, long time range)
  ▪ Handling many sensor channels at once
  ▪ Interactive exploration techniques for sensor data: annotation, brushing and filtering, searching etc.
2. Visualisation of Sensor Data – Project Ideas

- **Scalability**
  - methods to visualise massive signals: down-sampling techniques, LOD rendering, data transfer protocols etc.
  - Simultaneous visualisation of very many sensor channels: dense views

Downsampling can be problematic!
2. Visualisation of Sensor Data – Project Ideas

- Interactive exploration techniques for sensor data
- Annotation tools: users describe phenomena (collaboratively)
- Show a pattern overview grouped by annotations (on right)
2. Visualisation of Sensor Data – Project Ideas

• Brushing: multiple value-range filters, angle- (slope-) filter

• Searching interfaces: including similarity computation, ranking and result browsing
2. Visualisation of Sensor Data – Suggested Data Sets

- EEG Data:
  [http://sccn.ucsd.edu/~arno/fam2data/publicly_available_EEG_data.html](http://sccn.ucsd.edu/~arno/fam2data/publicly_available_EEG_data.html)
- Additional data sets will be introduced in a lecture on 19.04.2016
Visualisation of Tabular Data
3. Visualisation of Tabular Data

- **Data properties**
  - Multiple columns containing heterogeneous data types
  - A large number of rows
  - Potentially multiple values per cell

- **Data element is a row: described by multiple attributes**
  - Multi-dimensional data

- **Visualisation: specialised representations for different data types**
3. Visualisation of Tabular Data - Project Ideas

• Multi-visualisation UI
  ▪ Use data-type specific visualisations
  ▪ Choose meaningful representations for your data
  ▪ Implement view coordination for interactive analysis
    ▪ Interactions in one view are represented in all others
  ▪ Provide data aggregation and or filtering functions
• Extend the VisWizard or implement your own UI
3. Visualisation of Tabular Data - Project Ideas

- Algorithms for automated visualisation
  - Use knowledge about data, visualisations or even users to automate visualisation selection and configuration
  - Extract (or use available) data semantics to support the process
  - Consider the user profile
- Replace the current VisWizard algorithms
3. Visualisation of Tabular Data - Project Ideas

- Implement or extend metaphors for high-dimensional data
  - Extend parallel coordinates (e.g. with histograms or hierarchical information)
  - Implement a dimensionality reduction method to layout data in 2D
3. Visualisation of Tabular Data – Suggested Data Sets

- Open governmental data such as from
- Land Steiermark (CSV and Excel files):
- EU Open data Portal
  - Details to be introduced in the lecture on 26.04.2016
Visualisation Semantic Networks
4. Visualisation of Semantic Networks

- Display and navigate structure of large graphs, e.g.
- Ontologies: semantic networks
  - Consist of nodes and relations with precisely defined semantics
- Alignment of ontologies: map concepts from two semantic networks onto each other
- Extract graphs from unstructured text
  - Entities (persons, organisations, locations etc.): Natural Language Processing (NLP) methods
  - Relations between entities: e.g. co-occurrence in documents, sentences.
4. Visualisation of Semantic Networks – Project Ideas

- Graph layout algorithms, such as

Force-directed layout:

\[ d_1 \rightarrow d_2 \rightarrow d_3 \]

Edge bundling:

\[ F_s \rightarrow F_e \]

Edge Routing:
4. Visualisation of Semantic Networks – Project Ideas

• Graph visualisation showing semantic relationships
  ▪ Icon, shape and colour coding for relation and node types
  ▪ Interaction: expanding the network in a particular direction
  ▪ Visual methods for graph querying (e.g. “blossom node”: Chile)
4. Visualisation of Semantic Networks – Project Ideas

- **Ontology alignment**
  - Ontologies define concepts (vocabularies) and relationships between them
  - Problem: different domains and view-points - diversity of conceptualizations
  - Ontology alignment: map concepts from different ontologies
4. Visualisation of Semantic Networks – Suggested Data Sets

- DBPedia data sets: [http://wiki.dbpedia.org/Datasets](http://wiki.dbpedia.org/Datasets)
- Ontology Alignment Evaluation Initiative (OAEI): [http://oaei.ontologymatching.org/2012/](http://oaei.ontologymatching.org/2012/) (e.g. anatomy)
- More details to be given in the lecture on the 10.05.
Technical Prerequisites

• Client: HTML5/JavaScript (a must)
  ▪ With visualisation libraries such as D3.js, Sigma.js or Raphäel

• Server:
  ▪ Java (with Tomcat or Jetty)
    • Possibly using Apache Jena (Semantic Web framework)
  ▪ Python
    • Possibly with NumPy (large array/matrix), SciPy (scientific/technical computing)
    ▪ <your preferred Web development language/framework>

• Also see http://kti.tugraz.at/staff/vsabol/courses/mmis2/en/links.html

• You don’t need everything, but some of these will be helpful
Next Steps

• Attend the next lectures with **focus on the particular projects**
  - 19.04.2016: Recommendation User Interfaces, Sensor Data Visualisation (Cecilia, Gerwald)
  - 26.04.2016: Personalised, Automated Visualisation of Highdimensional Data (Belgin)
  - 10.05.2016: Visual Analytics for Unstructured and Network Data (Vedran)

• Lecture content
  - Introduction of visualisation and algorithm fundamentals
  - Technical information on software frameworks where you will integrate your results
  - Ask questions to the framework authors!
Next Steps

• Upcoming deadlines:
  ▪ Team building: 22.04.2016 (group registration in TeachCenter)
  ▪ Plan presentations: 03.05.2016
Thank you

Questions?
Exploit your Project Results

- Possibility to develop your MMIS2 projects further
  - as Bachelor or Master’s Thesis
  - Contribute to EU research projects (EEXCESS, AFEL, MoreGrasp)
  - Open-source code base
  - Perform usability evaluations
  - Scientific publication, if results are adequate