Introduction

Multimedia Information Systems 2 VU (707.025)
(“Web-based Visual Data Analysis” in the future)

SS 2016

Vedran Sabol

Know-Center

March 8th 2016
Overview

- Organisational information
- Goals of the course
- Course topics
- Practical part: projects
  - Topics, Deadlines
  - Tasks: will be given in early April
- Course structure and calendar
- Presentations and grading
Course

- Multimedia Information Systems 2 VU 707.025 (3.0 SSt, 5 ECTS credits)
- Elective (optional) course for
  - Computer Science
  - Software Development and Business Management
  - Doctoral Studies
- Catalogues: Multimedia Information Systems, Knowledge Technologies
Name: Vedran Sabol
Affiliation: Know-Center, KTI
Office: Inffeldgasse 13, 6th floor, room 082
Office hours: by appointment
Phone: +43 316 873 30850
Email: vsabol@know-center.at
Language

- Master course: lectures in English
- Communication in German/English
- If in German: please informally (Du)!
- Project: German/English
- Presentation: German/English
Organization of the Course

• Lectures
  ▪ When: Tuesday, 10:15 – 12:45
  ▪ Where: HS i9
• Registration for the course in TUGOnline until **09.03.2015**
• Presence at lectures is not obligatory, but recommended(!)
• **Presentations ARE obligatory**
Organization of the Course

• Course Homepage:
  http://kti.tugraz.at/staff/vsabol/courses/mmis2
    ▪ Lecture slides, links to external resources
• Newsgroup: tu-graz.lv.mmis2
  ▪ News server: news.tu-graz.ac.at
  ▪ Newsgroup is the preferred way of communication for this course
  ▪ The study assistant and the lecturer will actively participate in the newsgroup
Goals of the course (VU 707.025)

• 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
  ▪ Unstructured data (e.g. text, images)
  ▪ Semi-structured data (e.g. resources described by rich metadata)
  ▪ Network data (e.g. interlinked documents, social networks)
  ▪ Multi-dimensional data sets
  ▪ Semantically described data (ontologies)
  ▪ Sensor and time-oriented data
Goals of the course (VU 707.025)

• **Goal**: learn about the structure of complex data in the Web
  - Social networks and processes
  - Semantic knowledge bases: ontologies, linked open data cloud, RDF Data Cubes
  - Multimedia documents described by rich metadata
  - Sensor and event data collected by mobile devices

• **Goal**: learn about presenting Web content with visual means
  - In an suitable, easy to understand way
  - Using Web technologies (primarily HTML5)

• **Goal**: comprehend the Web data as an object of analysis
  - Knowledge Discovery in the Web (also known as Web Mining)
  - Visual Analytics for the Web
  - Apply algorithmic and visual methods for analysis of Web data
Goals of the course (VU 707.025)

- **Automated analysis: Knowledge Discovery Process**
  - Processing chain involving: selection, preprocessing, transformation, mining and interpretation of data
  - Mainly an automatic process

- **Involve humans in the analytical process: Visual Analytics**
  - Use visualisation to support analysis of complex data
  - Combining visual and automatic analysis methods

- **Goal:** learn how to apply Visual Analytics methods in the Web
  - on Web data
  - using Web technologies
  - in selected Web-based scenarios
Non-Goals (VU 707.025)

- MMIS2 is not about Web programming, Web frameworks, Service-oriented or Enterprise Architectures
  - MMIS1 dealt with some of those issues
- An advanced course on the above topics: 706.052 AK Informationssysteme (WS)
  - also deals with J2EE, architecture of Web applications, Data Warehousing etc.
Non-Goals (VU 707.025)

• MMIS2 is not about providing a comprehensive overview of Knowledge Discovery and Visual Analytics methods

• Advanced courses on the above topics
  ▪ 707.003 Knowledge Discovery and Data Mining 1 (VO, winter semester)
  ▪ 707.004 Knowledge Discovery and Data Mining 2 (VU, summer semester)
  ▪ 710.220 Visual Analytics (VU, summer semester)
Topics of the course (VU 707.025)

- **Automatic Web data analysis**
  - The Knowledge Discovery (KDD) process
  - Data selection and cleaning, feature engineering, data mining algorithms...
  - Discussion of selected data mining algorithms (e.g. clustering)
  - Applications on text, graph and sensor data

- **Recommendation User Interfaces**
  - Recommenders as ahead of time information retrieval engines
  - Adaptive visualisation interfaces for metadata-rich recommendations
  - Examples using a browser plug-in
Topics of the course (VU 707.025)

• Visual Analytics for Web Data
  ▪ Combined automatic and visual analysis – human in the loop
  ▪ Information landscapes
  ▪ Social network visualization
  ▪ Ordination and layout algorithms

• Visualisation of Semantic Data (RDF)
  ▪ Introduction to RDF
  ▪ Geo-spatial and temporal data
  ▪ Using semantics to automate visualisation
  ▪ Visual ontology alignment
Topics of the course (VU 707.025)

• High-dimensional data visualisation
  ▪ Multi-visualisation interfaces
  ▪ View coordination
  ▪ RDF Data Cube Visualisation
  ▪ Visual metaphors for multidimensional data

• Visual exploration of sensor and time-oriented data
  ▪ Scalable sensor-data visualization
  ▪ Visualisation of multiple sensor channels
  ▪ Interactive exploration techniques for sensor data
Example - Geovisualisation

• Which is the happiest city in the USA?

• Sentiment detection to extract “happiness” from geo-tagged tweets

• Geo-visualisation with colour coding to convey “happiness”
Example – EEXCESS uRank

- Content-based exploration of recommendations
- Significantly easier to use than list scanning

1. pick keywords
2. change weights
3. Re-ranking of documents
4. Evaluation shows: significantly easier than list scanning!
5. Inspection: highlight keywords in content
Example – EEXCESS Recommendation Dashboard

• Multiple visualisations
  - Timeline
  - GeoView
  - BarChart
• Filtering of recommendations
• Organising recommendations in collections
Practical Part – Project (VU 707.025)

• Implement a Web-based system for visual data analysis
  ▪ Team work: **groups of 2-3 students**

• Topical areas
  1. Visual exploration of network data (AFEL EU Project)
     • Social network data
  2. Automated visualisation of semantic data (AFEL and CODE EU projects)
     • Ontologies, multi-dimensional data sets (RDF-cubes)
  3. Visualisation of recommender results (EEXCESS EU project)
     • Recommendations incl. content and metadata (time stamps, geo-references...)
  4. Visualisation of sensor data (MoreGrasp EU project)
     • Sensor data from mobile devices, industrial sensors, bio-med sensors etc.

→ **Project tasks will be given in the lecture on 12.04.2016**

• **Attendance highly recommended!**
TeachCenter: for all matters concerning practicals
- [https://tugtc.tugraz.at/wbtmaster/courseMain.htm?707025](https://tugtc.tugraz.at/wbtmaster/courseMain.htm?707025)
- Detailed information on the practicals, development environment etc.
- Registration for projects, presentation slots etc.
- Will be set up over the following days
  - Announcement in a newsgroup posting
Practical Part – Tasks (VU 707.025)

- Team building: group member names, chosen project
- Project plan: goals, time estimate, group member responsibilities
- Implementation: working, well-documented code
- Project report: scientific paper-like document
  - Title + Abstract
  - Motivation and goals (which problem you are solving for the chosen data)
  - Description of your solution: methodology, algorithms, design, use case
  - Discussion and outlook: what worked well, what could be improved
  - References: software libraries, data sets, papers...
  - Length: 6 pages for groups of three students, 4 pages for groups of two
  - Format: Springer LNCS
    - [http://www.springer.com/computer/lncs?SGWID=0-164-6-793341-0](http://www.springer.com/computer/lncs?SGWID=0-164-6-793341-0)
Structure of the Course (VU 707.025)

- Lectures: theoretical part
  - Topics directly related to the projects
- Practical part: project (in groups)
- Plan presentations – mandatory participation
  - Present and discuss your project idea, receive feedback
- Project presentations – mandatory participation
  - Presentation of your project results: slides/demo
  - Oral exam in the form of question answering and a discussion of results (immediately after the presentation)
Timetable (VU 707.025)

  - Group registration in TeachCenter
- Project plan: 29.04.2016
- **Plan presentations: 03.05.2016**
  - 10:00 to 12:00 in HS i9
  - 13:00 to 15:00 in Know-Center meeting room, Inffeldgasse 13, 6th floor
  ➡️ **Project submission deadline: 17.06.2016**!
    - Implementation, code documentation, project report
    - Submissions through SVN
- **Project presentations: 21.06. and 28.06.2016**
  - 10:00 to 12:00 in HS i9
  - 13:00 to 17:00 in Know-Center meeting room, Inffeldgasse 13, 6th floor
- Study assistant: Benham Taraghi (b.taraghi@tugraz.at)
• **08.03.2016**: Course Introduction (Vedran)
  - Course organisation and schedule, examination mode, topics overview
• **15.03.2016**: Automatic Data Analysis in Visual Analytics (Vedran)

Easter holydays (21.03. – 09.04.2016)

• **12.04.2016**: Practicals Tutorial (Beni) and Projects (Vedran)
  - Coding standards, documentation, reports, organisation
  - *Project tasks and data sets*
• **19.04.2016**: Recommendation User Interfaces, Sensor Data Visualisation (Cecilia, Gerwald)
• **26.04.2016**: Personalised, Automated Visualisation of High-dimensional Data (Belgin)
Course Calendar (VU 707.025)

• 03.05.2016: Development plan presentations - students
• 10.05.2016: Visual Analytics for Unstructured and Network Data (Vedran)
• 24.05.2015: TBD - guest professor lecture
• 31.05.2015: TBD - guest professor lecture
• 07.06.2016: no lecture
• 14.06.2016: no lecture
• 21.06.2016: Project presentations - students
• 28.06.2016: Project presentations - students
Development Plan Presentations (VU 707.025)

- Present what you plan to implement
- Receive feedback from the lecturer and colleagues
- Short presentation: **max. 5min**, with slides (and initial results, if any)
- Time slot reservation: in the TeachCenter
- All groups must present and all team members must attend
Project Presentations (VU 707.025)

- Show us what you did, explain how it works
- Presentation/demo: **max. 10 minutes** (sharp)
- Time slot reservation: in the TeachCenter
- Exam: briefly answer 1-2 questions (2-3 minutes)
- Important for a good mark:
  - Argue why you did something (the way you did it)
  - Discuss advantages/disadvantages and possible improvements
- Similar to e.g. Master’s Thesis exam
- All groups must present and all team members must attend
Grading

• Plan and plan presentation: 10%
• Implementation: 30%
• Technical documentation: 10%
• Project report: 20%
• Project presentation and answering questions: 30%
• Grading
  ▪ 0 – 50: 5
  ▪ 51 – 62: 4
  ▪ 63 – 74: 3
  ▪ 75 – 87: 2
  ▪ 88 – 100: 1
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

• You don’t need everything, but some of these will be helpful
Exploit your Project Results

- 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
  - Possibility for scientific publication (if results adequate)
Thank you

Questions?