Data and Web Science Seminar
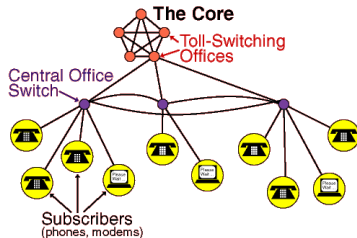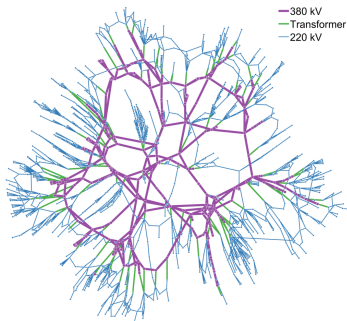
Graph Mining and Learning from Graphs

Kiril Gashteovski, Prof. Dr. Rainer Gemulla
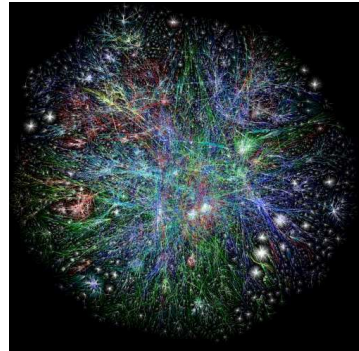
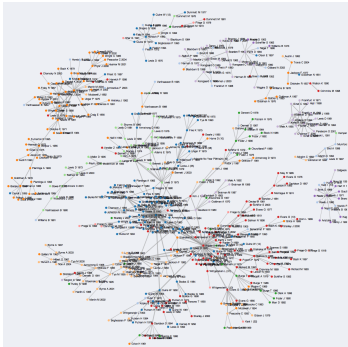# Why are graphs important?

- Graph data is everywhere
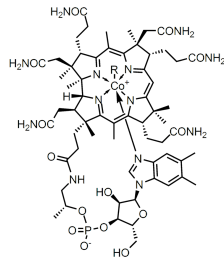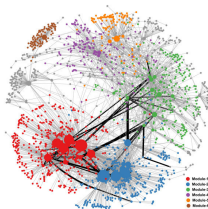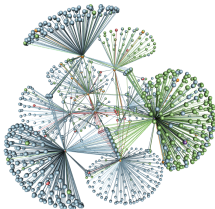 - **Technology:** power grids, telephone networks, …

Why are graphs important?

- Graph data is everywhere
  - **Knowledge and information:** citation networks, the Web, …

Why are graphs important?

- Graph data is everywhere
  - **Biology and chemistry:** protein-protein interactions, gene-regulation networks, chemical compounds, ...

Goals

Discuss recent techniques in graph mining and learning from graphs

– Models, training/inference, applications, pros/cons

You will

– **explore and read** scientific literature
– **summarize** your own topic in a concise report
– **present** your own topic in two presentations
– **discuss** other topics
– **review** your peers' output

Grading

- Report (50 %)
- 3-minutes flash presentation (10%)
- 15-minutes final presentation + 5 minutes discussion (30%)
- Peer review for slides and reports (10 %)

Note!

- Do not plagiarize
- Plagiarize $\Rightarrow$ lose 50 % immediately

Guidelines: Report

- Assigned papers are a **starting point**
  - also relevant: background material, prior work & subsequent work
  - a good starting point are papers that are cited by or cite the article (e.g., try Google scholar)
- Report should
  - Summarize the paper as well as relevant context (see above)
  - Discuss strength and weaknesses of the proposed methods
- 8-10 pages
- Use your own words
- Prefer **understandability** over technical details
- **Cite** related work properly

Guidelines: Presentation

- Flash presentation
    - ~ 3 minutes
    - think: elevator pitch
    - "what" and "why": attract your audience and convey excitement
    - don't be technical
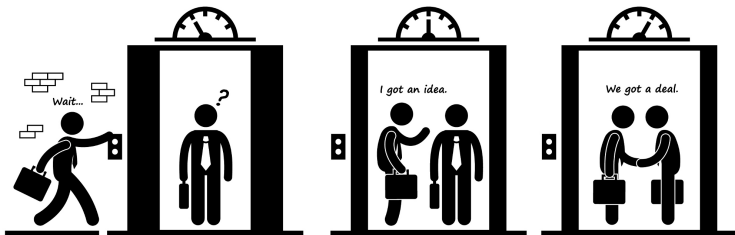
Guidelines: Presentation

- Flash presentation
    - ~ 3 minutes
    - think: elevator pitch
    - "what" and "why": attract your audience and convey excitement
    - don't be technical
- Final presentation
    - ~ 15 minutes
    - "what", "why" and "how"
    - add key technical points + your own thoughts
    - ~ 5 minutes moderated discussion:
        - audience: ask questions, comments, ideas
        - presenter: address questions and comments

Guidelines: Peer review

- Give feedback for other's report and slides using template
  - well-structured?
  - easy to understand?
  - strong points?
  - weak points?
  - ...
- Be honest, be constructive
- Saying everything is great is not a useful peer review (and neither is saying everything is bad)

Tips

- – Start early
- – Review your knowledge in optimization, linear algebra, and/or statistics
- – Google Scholar and bibtex are your friends
- – Various online tutorials exist for this topic
- – *"Giving Conference Talks"* by Prof.Dr. Rainer Gemulla
- – *"Writing for Computer Science"* by Justin Zobel