

Web Dynamics

<http://www.mpi-inf.mpg.de/departments/d5/teaching/ss09/dyn/>

Lecturers



Dr.-Ing. Ralf Schenkel
Group Leader, MMCI
Associated Senior Researcher, MPI-INF



Dr. Marc Spaniol
Postdoc, MPI-INF



Dimitar Denev (Excercises)
PhD Student, MPI-INF

Organization

- **Lecture:** Thursday 14-16 in E1.4/024
first on April 23rd, last probably on July 23rd
12 lectures
- **Assignments:**
 - both theoretical and practical (details in May)
 - date/time/location for exercise group needs to be determined
 - starting approx. May 13th
- **Requirements for obtaining 6 credit points:**
 - You **must** successfully participate in an oral exam at the end of the semester (probably July 30/31)
 - You **must** successfully solve the practical assignments
 - You **should** actively participate in the exercise group

Planned Outline

- (1) Introduction (today) [RS]
 - Dimensions of dynamics on the Web
 - Application examples
- (2) Modeling static and evolving graphs (2 lectures) [RS]
 - The Web graph and its static properties
 - Properties of dynamic graphs
 - Generative models for random graphs
- (3) Searching the dynamic Web (2 lectures) [RS/MS]
 - Crawling and recrawling policies
 - Search engine coverage
 - Accessing the Hidden Web

Planned Outline (ctd.)

(4) Preserving the Past: Web Archiving (2 lectures) [MS]

- Architecture of a large-scale archive
- Archive consistency, coherence, coverage
- Challenges of long-term preservation

(5) Searching the Past (2 lectures) [RS]

- Time-travel queries
- Temporal inverted indexes and query processing
- Temporal measures of page importance

(6) Human Behaviour on the Web (2 lectures) [RS]

- Modeling browsing behaviour
- Access Log Mining
- Personalization

Planned Outline (ctd.)

(7) Web Spam (1 lecture) [MS]

- Definitions and Examples
- Countermeasures

(8) Recent Results (1 lecture) [RS/MS]

- Interesting papers from SIGMOD'09 and SIGIR'09

(9) Summary and Outlook [RS/MS]

Selected Literature

- Pierre Baldi, Paolo Frasconi, Padhraic Smyth: **Modeling the Internet and the Web**, Wiley, 2003
- Julien Masanes (Ed.): **Web Archiving**, Springer, 2006
- Mark Levene, Alexandra Poulouvassilis (Eds.): **Web Dynamics**, Springer, 2004

The lecture mostly relies on scientific papers
(a list of references will be provided in the slides)