

DM79 Algorithms for Web Indexing and Searching

Fall 2004

Exam Questions

1. Crawling.
2. Indexing.
3. Ranking.
4. Querying.
5. The anatomy of Google.
6. Link based ranking schemes.
7. The I/O-model and external sorting.
8. The vector space model and latent semantic indexing.
9. Advanced lexicon structures.
10. The Web as a graph - statistics and models.
11. Document similarity.

The exam takes place on January 27. At IMADA's departmental office, you can find a sheet with enumerated slots in which you must enter your name. This determines which number student you are at the exam day. We expect to take a new student in approximately every half an hour (except for a half hour lunch break). Since some students may decide not to appear, and since the external examiner does not receive any payment for students who do not show up, you must be ready *at least 1 hour before the exam time you signed up for*. If not, you may lose your exam try.

The exam itself is without preparation. Thus, after you draw a topic, the examination starts immediately. You have the opportunity of making a 10–15 minute presentation using the blackboard (you cannot bring your own transparencies). We may ask brief questions during your presentation. Afterwards, we ask questions related to the topic and to the rest of the curriculum. The total examination time will normally be 25 minutes.