

Hot Topics

Hot Topics in Web Search

Pierre Senellart
(`pierre.senellart@telecom-paristech.fr`)



20 March 2009

New Aspects of Web Search

- **Web 2.0**: interaction-rich interfaces, social networks, mash-ups, user-provided content
- **Deep Web**: data accessible only through HTML forms
- **Information Extraction**: search at the level of objects inside pages
- **Semantic Web**: machine-readable content, complex search queries

Web 2.0

Definition

Web 2.0: **buzzword** about:

- **rich dynamic interfaces**, especially with the help of **AJAX** (Asynchronous JavaScript and XML) technologies: GMail, Google Suggest
- **user-editable** content, **collaborative** work and **social networks**: blogs, Wikipedia, MySpace, Facebook
- **aggregation** of content from multiple sources and **personalization**: Netvibes, Yahoo! Pipes

Web 2.0 Impact on Web Search

- How to crawl AJAX Web sites?
- Previous lecture about social networks and search in social networks
- How to exploit user feedback in search engines: **relevance feedback**, cf. slides by C. Manning and P. Raghavan, Stanford:
<http://www.stanford.edu/class/cs276/handouts/lecture9-queryexpansion-1up.pdf>
- Mash-ups, integration, cf <http://pipes.yahoo.com/>,
<http://www.google.com/ig>

Deep Web

cf. slides on the course website:

<http://pierre.senellart.com/enseignement/2008-2009/inf396/5-hot-topics/deep-web.pdf>

Information Extraction

cf. course by William Cohen, CMU:

<http://www.cs.cmu.edu/~wcohen/10-707/01-16-introduction.ppt>

Semantic Web

cf. online tutorial by Ivan Herman:

<http://www.w3.org/2006/Talks/0428-Beijing-IH/>