Demos

- BibSonomy - Understanding and Shaping Social Bookmarking
- FireMe! - Shaping User Awareness of Social Media Dynamics
- H-Probe - Observing the Structure of Network Traffic
- Linked Data Observatory - Observing and Understanding the Web of Data
- Web History Repository - Understanding User Interaction with the Web

Datasets published as part of the Web Observatory can also be found on the research data platform of L3S.

Descriptions

**BibSonomy**

Understanding and Shaping Social Bookmarking

BibSonomy is a social web system which allows users to store, organize, and share bibliographic data and annotate it with freely chosen keywords – so-called tags. One of the core motivations to develop and run BibSonomy was to have a system to study social bookmarking. The effort paid off: the regularly published datasets have attracted the interest of more than 600 scientists, Google Scholar now finds more than 100 publications that mention the datasets, funding for several projects could be successfully acquired, and many of our research results are based on BibSonomy and have been integrated into it: from tag recommendations over tag similarity measures and the social ranking algorithm FolkRank to spam classification methods.

**FireMe!**

Shaping User Awareness of Social Media Dynamics

Messages posted via Twitter are public and anyone can instantly see your updates. This may lead to negative consequences. FireMe! collects and displays tweets from people who supposedly hate their jobs. In addition, the FireMe! leaderboard features those users with the highest scores on bad language and who are most likely to get fired. In a user study, FireMe! sent out more than 4000 tongue-in-cheek warnings. The wording of the warning had a direct impact on whether people deleted their questionable tweet. The results show that young or inexperienced users would certainly benefit from post-hoc privacy alerts and warnings like FireMe!
**H-Probe**

Observing the Structure of Network Traffic

H-probe is an online active-probing tool for estimating the traffic correlation structure in Internet paths. H-probe enables users to perform lightweight, end-to-end measurements to any destination supporting ICMP. To extract the correlations of the traffic traversing a path, the software sends randomly spaced probes and analyzes the associated latencies. Aggregated Internet traffic is known to exhibit long range dependency, a characteristic which has a significant impact on network performance. H-probe enables researchers to obtain a better understanding of the network behavior and to improve the quality of web services.

**Linked Data Observatory**

Observing and Understanding the Web of Data

The Web consists of both unstructured resources such as Web pages and documents and increasing amounts of (semi-)structured open Web data. Linked (Open) Data in particular has provided large amounts of datasets across the Web, where datasets differ fundamentally with respect to their addressed topics, resource types or the provided quality. To facilitate data reuse and take-up, the Web Observatory has to observe and understand the nature of existing Web datasets and their evolution over time. The **Linked Data Observatory** provides an explorative way to browse and search through existing datasets in the entire **Linked Open Data (LOD) Cloud** according to the topics which are covered. By deploying entity recognition, sampling and ranking techniques, the demo allows to find datasets providing data for a given set of topics or to discover datasets covering similar fields. Capturing the evolution of such topic profiles is part of our ongoing work.

**Web History Repository**

Understanding User Interaction with the Web

Web usage data is essential for the development and evaluation for new search algorithms, recommendation algorithms and web browser functionality. In contrast to companies like Google, Microsoft and Facebook, open source software developers and researchers normally have no access to this data. With the **Web History Repository** we aim to create a public repository of web usage data. For privacy considerations, all data submitted to the project is encrypted and anonymized. The WHR initiative and the analysis of the data have been awarded with the Engelbart Best Paper Award at the 22nd ACM Conference on Hypertext and Hypermedia.