Bringing (Web) Databases to the Masses

Alon Halevy

Google Inc., USA
`halevy@google.com`

The World-Wide Web contains vast quantities of structured data on a variety of domains, such as hobbies, products and reference data. Moreover, the Web provides a platform that can encourage publishing more data sets from governments and other public organizations and support new data management opportunities, such as effective crisis response, data journalism and crowd-sourcing data sets. To enable such wide-spread dissemination and use of structured data on the Web, we need to create a ecosystem that makes it easier for users to discover, manage, visualize and publish structured data on the Web.

I will describe some of the efforts we are conducting at Google towards this goal and the technical challenges they raise. In particular, I will describe Google Fusion Tables, a service that makes it easy for users to contribute data and visualizations to the Web, and the WebTables Project that attempts to discover high-quality tables on the Web and provide effective search over the resulting collection of 200 million tables.