

# Structured Data on the Web (or, a Personal Journey Away From and Back To Ontologies)

Alon Halevy

Google Research, USA

For the first time since the emergence of the Web, structured data is playing a key role in search engines and is therefore being collected via a concerted effort. Much of this data is being extracted from the Web, which 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 encourages publishing more data sets from governments and other public organizations. The Web also supports new data management opportunities, such as effective crisis response, data journalism and crowd-sourcing data sets.

I will describe some of the efforts we are conducting at Google to collect structured data, filter the high-quality content, and serve it to our users. These efforts include providing Google Fusion Tables, a service for easily ingesting, visualizing and integrating data, mining the Web for high-quality HTML tables, and contributing these data assets to Google's other services. The talk will touch at various points on the role of ontologies in managing all this structured data and the new kind of ontological tools we may require.