When Semi-structured data meets the web

Alon Halevy
RESearch

Cheaper

Better

Bigger

Faster
Structured Data is Here, Big Time

• GoogleBase
  – Provide a feed of structured data
  – A few recommended attributes, but anything goes.
  – Still in beta.

• Deep Web
  – To crawl or not to crawl?

• Vertical search
Challenges

• Users don’t understand our fine distinctions:
  – Give me everything from one search box.

• Schema:
  – Lots of them. “Mass schema design”
  – Will incentives finally get people to converge?
  – Does it matter?
The Semantic What??

• This IS the Semantic Web!

• Finally something for us DB folks to do.