

CS347

Lecture 6

April 25, 2001

©Prabhakar Raghavan

Today's topic

- Link-based ranking in web search engines

Web idiosyncrasies

- Distributed authorship
 - Millions of people creating pages with their own style, grammar, vocabulary, opinions, facts, falsehoods ...
 - Not all have the purest motives in providing high-quality information - commercial motives drive “spamming”.
 - The open web is largely a marketing tool.
 - IBM’s home page does not contain *computer*.