

# Web Search and Mining

## Course description

The World Wide Web is evolving into the largest and most diverse information resource ever created. Web search and mining are two most frequently used technologies for retrieving information and discovering knowledge from the WWW. This course provides a comprehensive introduction to the state-of-art web search and mining techniques as well as their critical applications in the fields of business, bioinformatics, and healthcare informatics.

## Learning outcomes

- Understanding basic web search and mining techniques.
- Understanding basic working mechanisms of those popular web search engines like Google, AskJeeve, Overture (Yahoo), and Dogpile.
- Understanding some critical applications of web search and mining technologies in the fields of business, bioinformatics, and healthcare informatics.

## Instructor

Dr. Ying Xie (<http://science.kennesaw.edu/~yxie2/>)

## Textbook

Soumen Chakrabarti, *Mining the Web: Discovering Knowledge from Hypertext Data*, Morgan-Kaufmann Publishers, ISBN 81-8147-886-X.

## Prerequisite

Program admission

## Course Topics

1. Web crawling
2. Web indexing
3. Query languages
5. Retrieval models
6. Meta-search engines
7. Web page clustering
8. Web page classification
9. Business, bioinformatics, and healthcare informatics applications of web search and mining