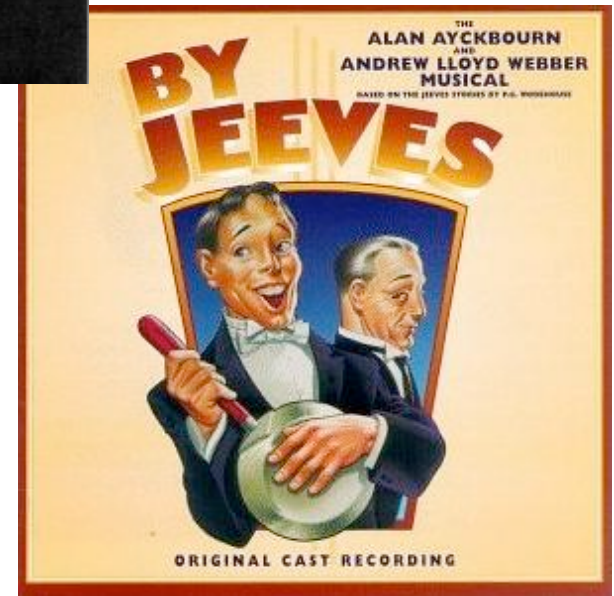
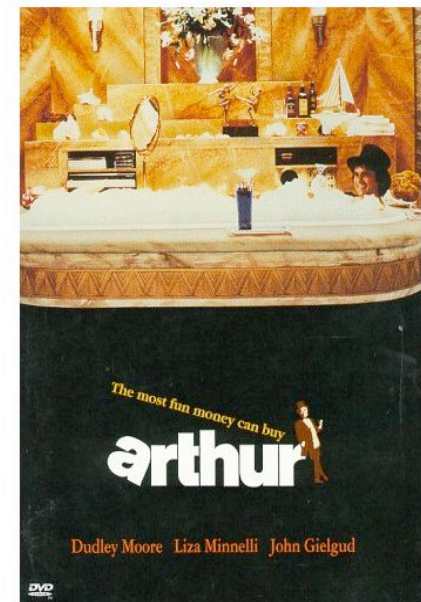




INLS 490-154: Information Retrieval Systems
Design and Implementation
Spring 2009

IR with MySQL and Text Files
Moving from structured to unstructured IR
January 22, 2009

Chirag Shah
School of Information & Library Science (SILS)
UNC Chapel Hill



INLS 490-154. Spring 2009

Source: Wikipedia

© 2009 Chirag Shah. UNC Chapel Hill



Ask Jeeves in 1996

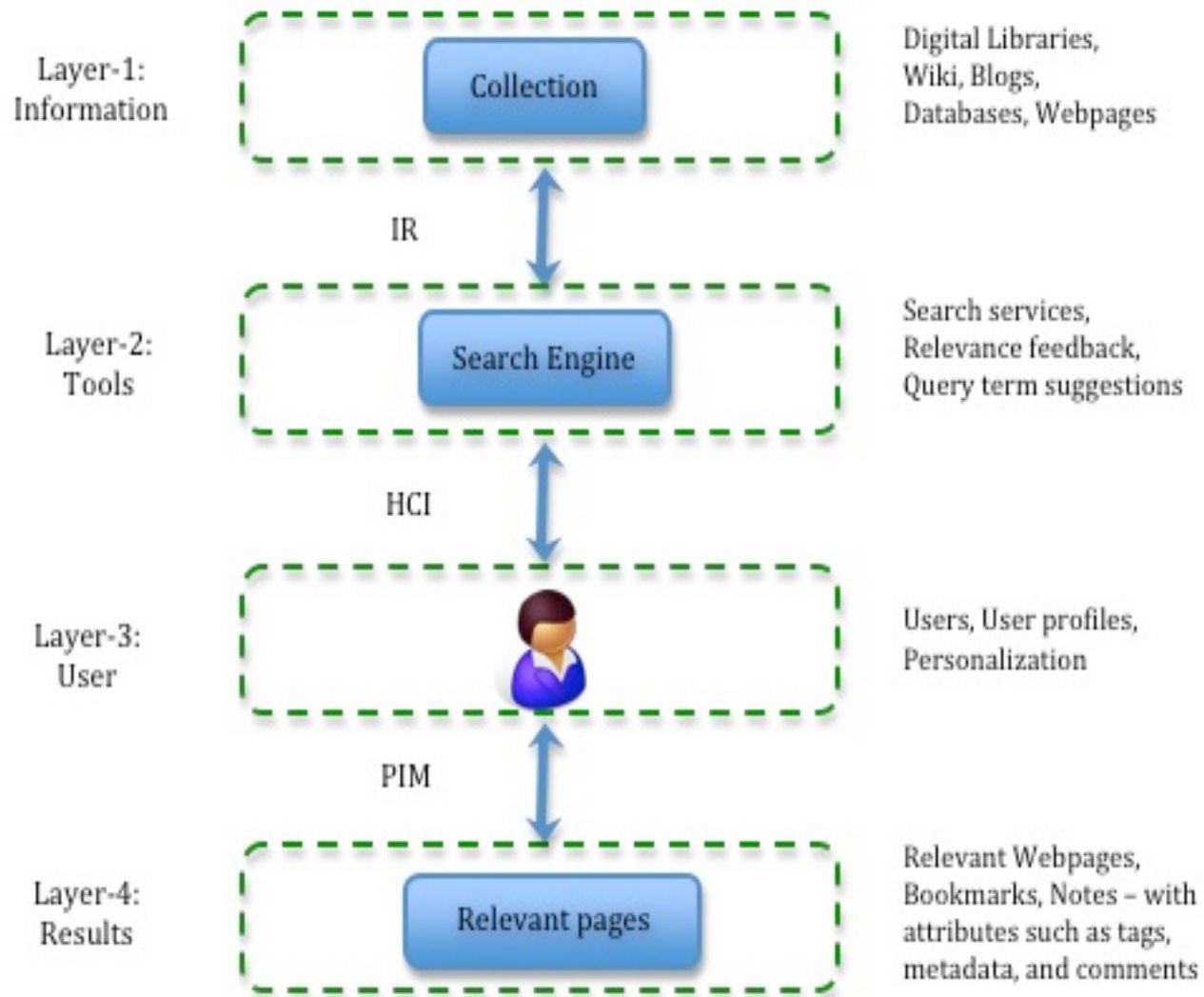
Source: Wikipedia



Today's lab

- Perform text searches on MySQL
- Build full-text search index for MySQL
- Play with PHP
- Develop a front-end for accessing MySQL data
- Work with text-files for indexing and retrieving

A model for information seeking



Summary

- Full-text searching in MySQL
- Front-end of a MySQL database for IR
- Searching through text files

- Concepts:
 - Field search vs. Full-text search
 - Serial scanning
 - Bag-of-words
 - Indexing
 - Stopwords
 - Stemming

Next time

- Extend our understanding of IR “model” by taking it out in the “open field”
- Setup Lemur Toolkit for language modeling and information retrieval
- Index a collection of documents and get them ready for search and retrieval

