



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

**Introduction**

*Get ready, get started, get going*

January 15, 2009

**Chirag Shah**

School of Information & Library Science (SILS)

UNC Chapel Hill

---

## Guess..!

- It started in a campus trailer
- It was titled “Jerry and David's Guide to the World Wide Web”
- They called it “Yet Another Hierarchical Officious Oracle”

# The story of Yahoo!



David Filo



Jerry Yang

**YAHOO!**

Started: February, 1994

First million-hit day in Fall, 1994 with almost 100,000 users

\$2 million invested in Yahoo! Inc. in Spring, 1995



## We're not going to build Yahoo!, but...

- We'll see how such search services can be created.
- Learn about search engine “secrets”.
- Deal with unstructured data (mostly).
- Get our hands (and knees) dirty by crawling on the Web.
- Build cool interfaces.



## What we're NOT going to do

- Work with structured data (with small exceptions)
- Learn how to program.
- Build websites that “can fly”.
- Use proprietary software.

## Requirements

- Basic experience with some programming environment.
- Laptop and a bunch of software (available for free).
- Open mind and a sense of adventure!

```
begin
  $query = <STDIN>;
  $document = "Hello world";
  if ((strcmp($query, $document) == 0)
      print "Match\n";
  else
      print "Sorry, no match found.\n";
end
```

# The gory details

- Quizzes
  - 5 or 6.
  - Lowest score dropped.
  - 25% of the grade.
- Assignments
  - Almost every week. Total 10 to 12.
  - Carries 45% of the grade.
  - Late policy.
- Project
  - Proposal due March 27<sup>th</sup>.
  - Final project due May 1<sup>st</sup>.
  - 25% of the grade.
- Grades
- Office hours
  - Manning 303. Monday 1:30-3:30pm, or by appointment.
- Website, class slides, and class notes
- Blackboard



# Quiz #0



## Today's lab

- Get your machines ready
- Learn a little bit about UNIX
- Access the database
- Create a table
- Access the information in the table



## Assignment- I

- Assigned: 01/15/2009 (today)
- Due: 01/20/2009 (Tuesday)
- Email the answers with “INLS 490: Assignment- I” in the subject.



## A word of advice to make your life easy!

- Review the basics.
- Create a “good” working environment.
- Try one thing at a time.
- Start a problem/assignment as soon as you can.
- Keep backups.



## Summary

- Why we're here, what we'll do, what we'll stay away from.
- This is not a “traditional” class.
- Try it to learn it.
- Basic UNIX
- Basic MySQL



## Next time

- Create webpage on the fly using data in MySQL.
- Unstructured data access from (1) MySQL and (2) text files.