

CS3245

# Information Retrieval

Lecture 0: Course Organization





Why should you care about

# INFORMATION RETRIEVAL?




PORTFOLIO &gt; CAREERS

## The 10 hottest jobs of 2016


15 DEC 2015 846 SHARES



### LATEST NEWS


HTC One A9 review: phone, if a bit expensive  
52 MINUTES AGO


Limerick TechTown: create 1,000 jobs for 3 years  
56 MINUTES AGO


Fare play as Uber has spent €4m in Limerick  
3 HOURS AGO



After gro America premises  
3 HOURS AGO


### Companies and organizations will struggle to find data talent


There are expected to be 4.4 million people employed worldwide in positions directly involved with big data analysis by next year. But this won't be enough. By next year, 70% of US businesses will either have a data strategy in place or will be planning one for the near future, according to one survey. The number of colleges offering courses related to big data analysis continues to grow rapidly, but there will continue to be a shortage of workers trained in the necessary skills for the foreseeable future.


### Textual analysis will become more widely used


Increasingly, much of the data we are storing for analysis is in an unstructured form. Textual analysis has become increasingly sophisticated in the last few years and that trend will continue. Computers will become more proficient at "reading" a piece of text (or voice converted to text) and spotting themes and sentiments – meaning it can be classified and analyzed in the same way as structured data.



Big Data: The Predictions For 2015  
Bernard Marr



The 2 Most Powerful Words A Manager Can Use  
Bernard Marr


Why You Need Emotional Intelligence To Succeed  
Dr. Travis Bradberry


How To Avoid Big Company Disease How Toyota Got Their  
Colin Baird


Give Yourself Space to Think  
Rita J. King


My Best Job Search Tip: How To Create An 'Interview Bucket'  
J.T. O'Donnell


3 Scripts that Keep Women Stuck at Work  
Sandra Ford Walston



**Bernard Marr**  **fluencer**  
Best-Selling Author, Keynote Speaker and  
Consultant in Strategy, Performance Management,  
[Follow](#)

# Information retrieval requests put a major strain on IT departments



By **Darren Allan** | Published 1 year ago

No Comments

Like 5

Share 14

G+ 0

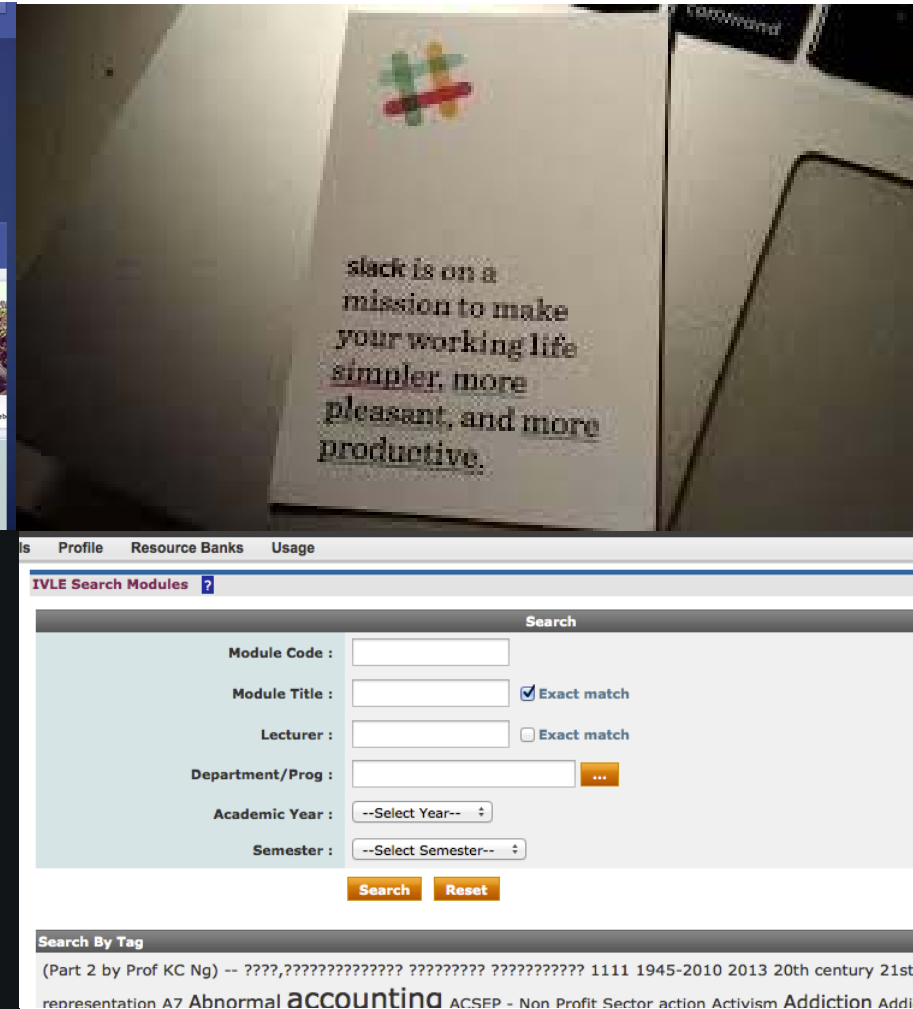
Tweet



Time consuming data retrieval requests are putting a big strain on IT teams, at least according to a new piece of research.



# A keystone of today's tech





# Jobs everywhere



## Sr Software Engineer – Information Retrieval (IR), Machine Learning (ML), Natural Language

quick apply

LinkedIn

San Francisco

📅 Posted 6 days ago    👤 20 views



Your connections **Dan and Yiming** work at LinkedIn



## SVP, Head of Search Engineering

Lazada Group

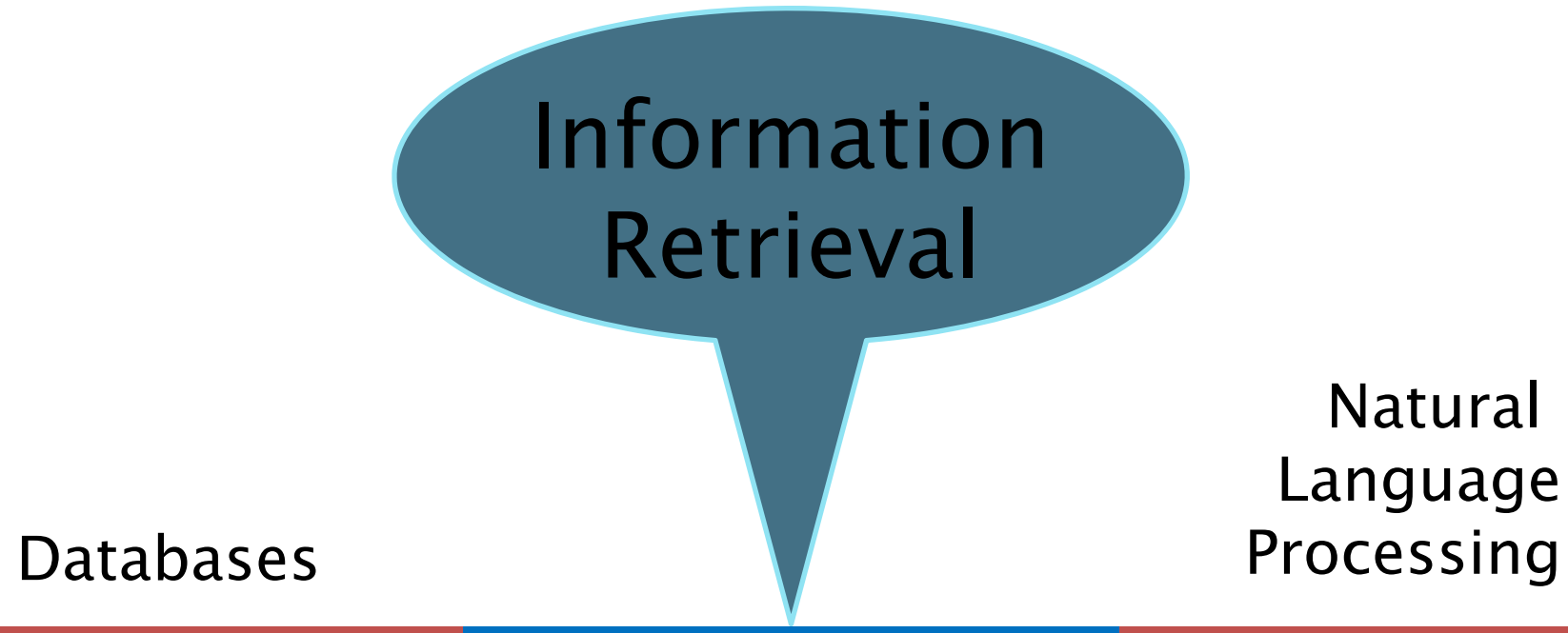
Singapore

📅 Posted 26 days ago    👤 152 views



1 person from your school works at Lazada Group

# Right at the center of attention



# How is it structured?

---



- 
- 
- 
- 
- 
- 

- 
- 
- 
-





And now for the

# **COURSE ORGANIZATION**

# Lecturer

ZHAO Jin

(Just call me “Jin” for short)

[zhaojin@comp.nus.edu.sg](mailto:zhaojin@comp.nus.edu.sg)

<http://www.comp.nus.edu.sg/~zhaojin>

Office: COM2-02-10    Tel: 66011083

Hours: 13:00-14:00 Fri,  
or by appointment

Hobbies:

Music, exercise and games



# Teaching Assistant

Kishaloy HALDER

[kishaloy@comp.nus.edu.sg](mailto:kishaloy@comp.nus.edu.sg)

<http://www.comp.nus.edu.sg/~kishaloy/>

Office: AS6-05-21 (Media Lab 6)





# Course web sites

<http://www.comp.nus.edu.sg/~cs3245>

- Course materials
- Homework
- Other supplementary content

When using a search engine to find our course materials, make sure you find our site for 2017/18 Sem II.

<http://ivle.nus.edu.sg/>

- Announcements
- Discussion forum (participation counts!)
  - Any questions related to the course should be raised on this forum
  - Emails to me are considered public unless otherwise specified
- Homework submissions (via workbin)

# Freedom of information rule



- Collaboration is acceptable and encouraged.
- To assure that all collaboration is on the level, **you must always fill in the name(s) of your collaborators on your assignment.**
- You will be assessed for the parts for which you claim is your own contribution.



# Facebook rule

---

- You are free to meet with fellow students(s) and discuss assignments with them.
- Writing on a board or shared piece of paper is acceptable during the meeting; however, you **may not take any written (electronic or otherwise) record away from the meeting.**
- After the meeting, do something else for at least a half-hour (Facebook, or doing an assignment for a different class), before working on the assignment.
  - This will assure that you are able to reconstruct what you learned from the meeting, **by yourself, using your own brain.**
  - You will be asked to certify that you meet this requirement per assignment.

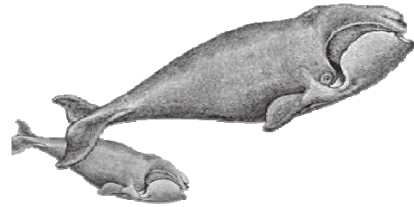


# Python



- We'll be using the Python programming language for our entire class homework assignments.
- You are expected to learn Python and use it for the assignments.
- Don't worry, Python is an easy language to transition to for most programmers and it is very useful for string manipulation (a key characteristic for IR).

# NLTK



- We will also be using the Natural Language Tool Kit (NLTK) for helping us deal with some parts of the course.
- It's coded in Python (surprise!)
  - Yes, we know that this is a course on **information retrieval** and not **natural language processing**, but the two are forever intertwined.
- Please install both Python (2.7 or 3.6) and NLTK on your own personal computer.