

CIT 496: Technology Support Admin "Search Engines/Information Retrieval"

[Spring 2012: Dr. Leyla Zhuhadar]

Instructor Information

Instructor: Leyla Zhuhadar, PhD.

Computer Engineering and Computer Science



Office Hours: By appointment: (Mon through Thu) 9:00 am-4:00pm

Address: Western Kentucky University

1906 College Heights Blvd. Bowling Green, KY 42101

Office: Garrett Conference Center, Office #104

Telephone #: Office: 270-745-8774

Cell phone: 270-535-8132 (only in case of emergency)

Email: <u>Leyla.zhuhadar@wku.edu</u>

Course Website: http://ecourses.wku.edu

Instructor Website: http://people.wku.edu/leyla.zhuhadar

Prerequisite: CIT 300

Technical problems with Blackboard or in your WKU Email account:

Call Help Desk @ 270-745-7000 for any problem you have related to your online course. All information regarding technical issues can be found at: http://www.wku.edu/infotech

Technical Problems

with Tegrity

multimedia Server

Email: <u>leyla.zhuhadar@wku.edu</u> and describe your technical problem in details, also provide your telephone number.



Course Description

Google Search Engine became one of the most important resources that everyone uses on daily basis! The effective integration and indexing of content, retrieving, and querying is expected to lead to the next generation of more useful and more intelligent Web applications. This sets the stage for the future of the semantic web and intelligent information retrieval. This course provides a fundamental working knowledge of the varied aspects of Technology Support Admin and prepares students for future studies in more specialized topics within subject area such as, Information Retrieval, Search Engines, and Semantic Web. Students will understand how technology can support administrator in building strong, reliable E-commerce platform using effective, smart Search Engine that can attract customers interests.

Students will increase their awareness of the overall environment and function of the WWW and how to be able to discover knowledge from the Web by indexing and retrieving information from the web as well as observe its contributions to society. This course also covers practical (hands-on) examples using open source software as a main resource for practicing, like "Galago Search Toolkit". In addition, students will be introduced to the usage of up-to-date Search Engine software, such as Google "Google-fastcat". As summary, this course prepares CIT students with knowledge of the real world environment. It provides insight about the architectural design and the management behind the implementation of the leader's companies on WWW, such as Google search engine, and Yahoo.

Course Objectives

By the end of this course, you should be able to:

- Understand how technology can support administrator in building strong, reliable E-commerce platform using effective, smart search engine that can attract customers' interests.
- Differentiate between different types of basic Indexing and Retrieving data techniques
- Describe how search engines, such as Google works and have the capacity to build your own search engine (link analysis, Web Crawling)



Course Time Management

It is your responsibility to manage your time which is the most difficult issue in online course. Most face-to-face classes meet three hours a week, in addition to that students are expected to spend up to nine hours per week in reading and homework assignments. Therefore, you can expect to spend up to twelve hours per week in any settings face-to-face or online. In this course, you will be provided with different technological components that might facilitate your class time. Therefore, start early with your preparations and do not procrastinate till the last minute.

Very Important: Technical Requirement

You need to have a PC (not an Apple computer). Also, you need to have a wired connection (not wireless connection!). Your PC will be used as a server. You can test your PC before taking this course by: (1) turning the remote access on from your system properties, then, go to command line (start>cmd) and write (ipconfig). Write down your IP-address, send this IP to a friend and see if he/she can access your computer from his/her computer (along with your IP-address, provide your friend your User Name and Password).

If you don't have a PC that fulfils this requirement, you will not be able to do any assignment in this course!!!

Emails to Instructor

- Emailing your Instructor: the email must have the following format: CIT-496, Last name, Topic (emails not following this format will be gently returned to you to revise the subject, I urge you to follow the protocol, so I don't lose your emails).
- Replying to your Instructor's email: When you reply to your Instructor's email, you need to keep the original email included in your replay, don't change the subject (your email will be Re: "Previous Subject", your Instructor next respond, would include all previous conversations and the subject remain the same, as long the email is intended to the same topic.
- Please use formal language in your emails.

Submission of Assignments

• Watch the following video that teaches you how to submit an Assignment:



http://www.wku.edu/infotech/atech/videotutorials/Bb9SubmittinganAssignment.htm

■ To Check your grades regarding an Assignment watch the following video tutorials: http://www.wku.edu/infotech/atech/videotutorials/Bb9CheckingGrades.htm

Late Assignments

Assignments turned in after due dates during the semester will result in a 10% reduction per day unless prior arrangements were made with the instructor. Any assignments turned in after the last due date (see course calendar) will result in a 20% reduction per day unless prior arrangements were made with the instructor. Technical problems are NOT an excuse unless reported to the instructor prior to 24 hours before due date.

Grade Book																	
Week#	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	Pts
WKU Calendar Events	ents will acce																
Watch Video Lectures								60									
Welcome Blog	20														20		
Career Blog	100							100									
Discussion Board																	
Quizzes			20														20



Course Syllabus p. 5

Test		100		100		200	
Final Project		(Mid-term	•	(Final-Project) 300	600		
Total	1000						
	Final Opportunity: If you did not make 3.0 in this course! You may got chance to revise your Final Project (300 points), but only by given extra 3 days to resubmit your analysis.						
Final Grades	On Monday 5/8/2012 Final Grades due are by 4:00 pm						

Grading Scale							
Total Points	Letter Grade	Percentage					
900-1000	A	90-100%					
800-899	В	80-89%					
700-799	С	70-79%					
600-699	D	60-69%					
000-599	F	0-59%					

Course Contents

This course does not use a traditional textbook. Instead, you will be reading from different websites and from different articles, and watching Video Guides recorded by your Instructor.

You will have access to:

- 1. Download **videos** from **Tegrity** at any time by clicking the Tegrity Campus tab at the very top-right of Blackboard page (graded).
 - ✓ First, have your Internet Explorer as your default browser, watch the following video tutorial to set your IE as default Browser:

 http://www.wku.edu/infotech/atech/videotutorials/Setting_the_default_browser.htm



- ✓ Second: You have different ways to watch Tegrity video, in case you are connected to Internet, all the time, you can watch the Tegrity video streamed over the Internet (see the following video tutorial: view) or by subscribing to Tegirty video as podcast/iTunes/MP3/RSS feed (see the following video tutorial: view). Which means you can watch lectures on your ipod, iPhone, iPad or listening as podcast on MP3 while commuting, working, or even while exercising at the gym!
- ✓ Third: If you have limited connection to Internet, you can visit an Internet Café or while you are at WKU university and download the video on your laptop while you are connected to internet, then, watching those videos, offline, on your laptop at any place you are at (no need to Internet connection in that case (see the following video tutorial: view)
 Note: Any problem you may have using Tegrity, you will have to contact me personally via email (It happens that in my second job in the Office of Distance Learning. Lam the Tegrity
 - in the Office of Distance Learning, I am the Tegrity
 Administrator and who does the Technical support for Tegrity
 on both sides User & Server).
- ✓ Fortunately, you will have no problem Tegrity videos! I track your login to those videos in two ways: how many times you logged-in and how much time you spent watching them. Trust me that you will benefit, the most, from watching those videos to get good grades; since it will help you understand the course materials.
- ✓ All the Materials used in the Tegrity videos are available to download in two formats (pdf, ppt) under course documents.

2. **Blog**:

- ✓ Welcome Blog (graded): You will use this Blog to introduce yourself to your colleagues, upload your picture and to share ideas and even to post questions to your colleagues.
- ✓ Career Blog OR Creating Futuristic Social Network (graded):

 This Blog will be graded, but who provides creative ideas might get extra credits. The purpose of this Blog is to share knowledge related to the course topics by suggesting Social Network connections, such as *LinkedIn* (of course, not *My Space* or *Facebook*!). At the end of the course, each one of you should create his own Career Network in



the area of interest that might help his/her future job's hunting. My job is to help you by teaching applications that are known as #1 most wanted applications for jobs in the area of Search Engine. If you suggest ideas that your colleagues approve, you might get extra credit at the end of the semester!

3. Discussion Boards:

✓ Q & A Forum: This is your "S.O.S" forum! Look at this forum as your S.O.S rescue place when you are puzzled and need help from your colleagues.

Below is an outline of the content and activities in each unit of this course. Course Documents

Unit 1 (from 1/24 to 2/7): Introduction to CIT-496

- **Reading 1**: Syllabus (in details)
- Watch Video Tutorial 1
- Welcome Blog: Post your first Blog entry: Introduce yourself in details, post a picture and share knowledge with your Colleagues-- Due dates: from 1/24 to 2/13
- Google Analytics (20 points): Assignment #1 (20 points).

Unit 2 (from 2/7 to 2/14): Introduction to Search & IR

- Watch Video Tutorial 2
- **Reading 2**: Intro to Information Retrieval in Practice
- Continue: Welcome Blog (20 points)-- until 2/13

Unit 3 (from 2/14 to 2/21): Search Engine Architecture

- Watch Video Tutorials
- Reading 3: Build your first Search Engine
- Start: Q & A Forum "S.O.S" Discussion Board-thread #1



 LinkedIn (Career Blog Posting) (100 points) -- Due dates: from 2/14 to 5/8

Unit 4 (from 2/21to 2/28): Prepare for Test #1

- Watch Video Tutorials about your Test #1
- **Reading 4:** Notes on Test #1
- Test 1 (100 points) -- Due date: <u>2/28/2011</u>
- Continue LinkedIn (Career Blog Posting) -- until 5/8/2011

Unit 5 (from 3/7 to 3/14): Web Crawler

- Watch Video Tutorials
- Reading 5
- Continue LinkedIn (Career Blog Posting)-- until 5/8/2011
- Start: Q & A Forum "S.O.S" Discussion Board-thread #2
- Test 2 (100 points)--Due date at: 3/28/2011

Unit 6 (from 3/14 to 3/21): Processing Text

- Watch Video Tutorials
- Start: Q & A Forum "S.O.S" Discussion Board- thread #3
- Reading 6
- Continue LinkedIn (Career Blog Posting) -- until
 5/8/2011
- Mid-term Project (300 points) -- Due dates: from 3/14 to 4/11/2011

Unit 7 (from 3/21 to 3/28): Indexes and Ranking

- Watch Video Tutorials
- Reading7



- Continue LinkedIn (Career Blog Posting) -- until
 5/8/2011
- Start: Q & A Forum "S.O.S" Discussion Board- thread #4
- Continue working on Mid-term Project

Unit 8 (from 3/28 to 4/4): Information Needs

- Watch Video Tutorials
- Reading 8
- Continue LinkedIn (Career Blog Posting) -- until
 5/8/2011
- Start: Q & A Forum "S.O.S" Discussion Board- thread
 #5
- Continue working Mid-term Project

Unit 9 (from 4/4 to 4/11): Retrieval Model Overview

- Submit Mid-term Project—Due Date 4/11/2011
- Watch Video Tutorials
- Reading 9
- Continue LinkedIn (Career Blog Posting) -- until 5/8/2011
- Start: Q & A Forum "S.O.S" Discussion Board- thread #6

Unit 10 (from 4/11 to 4/18): Evaluation of Search Engines

- Watch Video Tutorials
- Reading 10
- Continue LinkedIn (Career Blog Posting) -- until 5/8/2011
- Start: Q & A Forum "S.O.S" Discussion Board- thread #7

Unit 11 (from 4/18 to 4/25): Classification and Clustering in IR

- Watch Video Tutorials
- Reading 11
- Continue LinkedIn (Career Blog Posting) -- until 5/8/2011



Start: Q & A Forum "S.O.S" Discussion Board- thread #8

Unit 12 (from 4/25 to 5/1): Social Search

- Watch Video Tutorials
- Reading 12
- Continue LinkedIn (Career Blog Posting) -- until 5/8/2011
- Start: Q & A Forum "S.O.S" Discussion Board- thread #9

Unit 13 (from 5/1 to 5/8): (300 points)--Final Project Submission Due date on 5/8/2011

Submitting your Final Project

from 5/8 to 5/10: Final opportunity

• Final Opportunity for you to revise your Final Project in case your grade is less than 3.0 in this course

Final Grades (5/13 at 4 pm)

- Project Feedback (individual)
- Final Grade Feedback (individual)

Plagiarism

WKU is subscribed to **TurnItIn**, a plagiarism service which gives professors and originality report for each student paper turned in. Your assignments may be checked through this service.

Disability Accommodations Statement



Course Syllabus p. 11

Disability Accommodations Statement: "Students with disabilities who require accommodations (academic adjustment and/or auxiliary aids or services) for this course must contact the Office for Student Disability Services: http://www.wku.edu/Dept/Support/AcadAffairs/SDS/sds.htm. The Office for Student Disability Services (OFSDS) telephone number is 270-745-5004. Please DO NOT request accommodations directly from the professor or instructor without a letter of accommodation from the OFSDS."