Course Definition

Information Retrieval

Information

Technology

Engineering

Syrian Arab Republic	الجمهورية العربية السورية
Ministry of Higher Education and Scientific Research	وزارة التعليم العالـــي والبحث العلمي
Syrian Virtual University	الجامعة الافتراضية السورية

1. Basic Information:

Course Name	Information Retrieval
Course Code	SIR601
Number of Presentational Sessions*	20
Number of Synchronous Sessions**	10
Number of Shorter Tests***	2
Number of Exams***	1
Theoretical Sessions Work Load (hrs.)	60
Practical Sessions Work Load (hrs.)	30
Credit Hours	6

*Each presentational session comprises both recorded lecture (1.5 hrs.) and interactive learning content (1.5 hrs.).

**Each synchronous session comprises the interactive lecture carried out in real time in a virtual class (1.5 hrs.).

***Each shorter test is 0.5 hr. long. The final exam is 2 hrs. long.

N.B.

Generally, each chapter requires two presentational sessions: one for the recorded content and one for the interactive content (unless the chapter is too long, in which case it may require more sessions (. This note applies to synchronous sessions as well, where each chapter requires one synchronous session generally.

Syrian Arab Republic		الجمهورية العربية السورية
Ministry of Higher Education and Scientific Research		وزارة التعليم العالــي والبحث العلمي
Syrian Virtual University		الجامعة الافتراضية السورية

2. Prerequisites courses:

Course	Code
Algorithms and Data Structures (1)	BDA501
Discrete Mathematics	BDM501
Software Engineering II	SSE602

3. Course Objectives:

The main aims of this course "Information Retrieval" are:

- Familiarizing students with information retrieval algorithms
- View the mechanisms of building indexes
- Explanation of text processing technologies
- Using the Boolean model and the extended Boolean model
- Using the vector space model
- Evaluating information retrieval systems
- Adaptative systems
- Introducing of content-based image retrieval systems (CBIR)

4. Learning Outcomes (LO):

Upon completion of the course, the student must have:

- Finding results for Boolean queries
- Finding results for extended Boolean queries
- Text processing and index building
- Finding the results of the queries using vector space model
- Evaluating retrieval systems performance
- Adapting with users

Syrian Arab Republic		الجمهورية العربية السورية
Ministry of Higher Education and Scientific Research		وزارة التعليم العالــي والبحث العلمي
Syrian Virtual University		الجامعة الافتراضية السورية

- Learning retrieval basics depending on content

5. Assessment Results:

Chapter Number	Chapter Title	General Objectives	Assessment Type				
			Interactive Content & Recorded Sessions	Applied Activities (Synch. Sessions)	Final Exam* / Shorter Tests**	Presentations and Interviews***	Reports ***
CH1	Information Retrieval	Comprehension –Analytical Thinking –Tools and Application Hands– On	√	√	√	√	√
CH2	Boolean Model	Comprehension –Analytical Thinking –Tools and Application Hands– On	√	√	√	√	√
CH3	Extended Boolean Model	Comprehension –Analytical Thinking –Tools and Application Hands– On	√	√	√	√	√
CH4	Text Processing	Comprehension	√	√	√	√	√

		–Analytical Thinking –Tools and Application Hands– On					
CH5	Indexing	Comprehension –Analytical Thinking –Tools and Application Hands– On	√	√	√	√	√
CH6	Vector space model	Comprehension –Analytical Thinking –Tools and Application Hands– On	√	√	√	√	√
CH7	Evaluation	Comprehension –Analytical Thinking –Tools and Application Hands– On	√	√	√	√	√
CH8	Query Operation	Comprehension –Analytical Thinking –Tools and Application Hands– On	√	√	√	√	√
CH9	Links Analysis	Comprehension	√	√	√	√	√

		−Analytical Thinking −Tools and Application Hands− On					
CH10	Content− based image retrieval	Comprehension −Analytical Thinking −Tools and Application Hands− On	√	√	√	√	√

*The final exam is two hours long and is given at the end of the course.

**Shorter tests are about 30 minutes long and are given after three or four lectures throughout the semester during synchronous sessions.

***Presentations, interviews, and reports are submitted once after each three or four lectures throughout the semester during synchronous sessions.

6. Course Syllabus:

Chapter	Subject	Content	Number of Learning Objects	Number of synchronous Learning Objects
CH1	Information Retrieval	1. Introduction 2. Information Retrieval 3. Big data computing 4. Apache Hadoop 5. Hadoop Ecosystem 6. Relevance 7. Intelligent Retrieval Systems	7	3
CH2	Boolean Model	1. Introduction 2. Boolean Model 3. Boolean Model Disadvantages	3	1
CH3	Extended Boolean Model	1. Introduction 2. Extended Boolean Model 3. Extended Boolean Model Disadvantages	3	1
CH4	Text Processing	1. Introduction 2. Tokens Extraction 3. Hadoop MapReduce Framework 4. Basic Frame Recovery	8	4

		5. Stemming 6. Porter Stemmer 7. N–gram stemmers 8. Connected Phrases Recognition		
CH5	Indexing	1. Introduction 2. Inverted Index Construction 3. AND Query Processing	**3**	**1**
CH6	Vector space model	1. Introduction 2. Statistical Models 3. Vector Space Model 4. Analogy Measuring Tool 5. Inner Product 6. Angle Cosine as Analogy Measuring Tool 7. TF_IDF Weighing 8. Advantages and Disadvantages	**8**	**4**
CH7	Evaluation	1. Introduction 2. Evaluation Standards 3. Reverse Relation between Precision and Recall	**3**	**1**
CH8	Query Operation	1. Feedback 2. Query reforming 3. Thesaurus–based Query Expansion	**5**	**2**

		4. Statistical Thesaurus 5. Correlation matrix Query Expansion		
CH9	Links Analysis	1. Introduction 2. HITS Algorithm 3. PageRank Algorithm	**3**	**1**
CH10	Content−based image retrieval	1. DBIR 2. CBIR 3. Image Representation 4. Color features 5. Texture 6. GLCM 7. Shape	**7**	**3**

7. Practical Activity:

- **Tools and Labs:**

Tool Name	Description
C# or any suitable programming Language	Visual Studio or IDE

- **Practical Activities per Chapters:**

Chapter	Activities Type	Remarks
CH1	☑ Exercises ☑ Homework ☐ Webinars	

	Syrian Arab Republic		الجمهورية العربية السورية
	Ministry of Higher Education and Scientific Research	**SVU** الجامعـة الافتراضيـة السوريـة SYRIAN VIRTUAL UNIVERSITY	وزارة التعليـم العالـــي والبحث العلمي
	Syrian Virtual University		الجامعة الافتراضية السورية

	☐ Project ☐ Experiment ☐ Other	
CH2	☑ Exercises ☑ Homework ☐ Webinars ☐ Project ☐ Experiment ☐ Other	
CH3	☑ Exercises ☑ Homework ☑ Webinars ☐ Project ☐ Experiment ☐ Other	
CH4	☑ Exercises ☑ Homework ☐ Webinars ☐ Project ☐ Experiment ☐ Other	
CH5	☑ Exercises ☑ Homework ☐ Webinars ☐ Project ☐ Experiment	

	☐ Other	
CH6	☑ Exercises ☑ Homework ☐ Webinars ☐ Project ☐ Experiment ☑ Other	
CH7	☑ Exercises ☑ Homework ☐ Webinars ☐ Project ☐ Experiment ☑ Other	
CH8	☑ Exercises ☑ Homework ☐ Webinars ☐ Project ☐ Experiment ☑ Other	
CH9	☑ Exercises ☑ Homework ☐ Webinars ☐ Project ☐ Experiment ☑ Other	
CH10	☑ Exercises	

Syrian Arab Republic			الجمهورية العربية السورية
Ministry of Higher Education and Scientific Research			وزارة التعليم العالـــي والبحث العلمي
Syrian Virtual University			الجامعة الافتراضية السورية

	☑ Homework	
	☐ Webinars	
	☐ Project	
	☐ Experiment	
	☑ Other	

8. References:

Introduction to Information Retrieval, Cambridge University Press. 2008. Christopher D. Manning, Prabhakar Raghavan and Hinrich Schütz