

**Jerash University**

**Faculty of Computer Science and Information Technology**

**Computer Sciences Department**

**Semester**: Fall Semester 2018/2019

|  |  |
| --- | --- |
| **Course symbol and number: 1002241** | **Course Name:** Information Retrieval System |
| **Teaching Language:** English | **Prerequisites: 1002140**. |
| **Credits:** 3 hours**.** | **Course Level:** 400 |

|  |
| --- |
| **Course Description** |
| The course will focus on basic information retrieval (IR) terminology and concepts, and provide further a large spectrum of practical applications of IR within industrial settings. The introductory components of the course include representation of information needs and documents, retrieval models, clustering algorithms, text categorization. A variety of other topics on adaptive information retrieval, web-based IR and user interfaces and visualization for IR will be also studied in this course. The lectures will cover topics which will be further exercised and exemplified in practice within the final assignments. The assignments involve both analytical, design and implementation skills. |

|  |
| --- |
| **Course Objectives** |
| * To gain a detailed understanding of automated indexing and retrieval and to survey various subfields of information retrieval and closely related topics in language processing. * To learn the underlying technology of search engines. * To Gain practical experience building simple, but true-to-practice retrieval software * To appreciate topics in the broad area of information retrieval, including evaluation, classification, cross-language retrieval, and computational linguistics. |

|  |
| --- |
| **Learning Outcomes** |
| |  | | --- | | At the end of the class, students should be able to master latest techniques of text indexing, web search, text mining and system evaluation including building index, calculating term weights and ranking scores, etc. Students will form teams and apply these techniques on real-world web data using IR tools. | |

|  |  |
| --- | --- |
|  | **Text Book(s)** |
| **Title** | Introduction to Information Retrieval |
| **Author(s)** | Kindle Edition |
| **Publisher** | John Wiley & sons |
| **Year** |  |
| **Edition** |  |

|  |  |
| --- | --- |
|  | **References** |
| **Books** |  |
| **Internet links** | http://www.jpu.edu.jo/lms |
| **Course link** |  |

|  |  |
| --- | --- |
|  | **Instructors** |
| **Instructor** | Dr.Ali Malkawi |
| **Office Location** | الطابق السابع - 715 |
| **Office Phone** |  |
| **E-mail** | [ali.amalkawi@jpu.edu.jo](mailto:Shquier@jpu.edu.jo) |

|  |  |  |  |
| --- | --- | --- | --- |
| **Topics Covered** | | | |
| **Topics** | **Chapters in Text** | **Week number** | **Teaching hours** |
| **Introduction**   * Basic concepts   + Information vs. Data Retrieval   + User Task   + Logical View of the Documents * History of Information Retrieval * Retrieval Process * Formal definition of IR model   Classical IR models | *Chapter 1,2* | *1* | *3* |
| **Query Language and Operations**   * IR Systems evaluation   + IR Systems performance   + Results Relevance   + Recall & Precision * Query Languages   + Keyword-based Querying   + Single-word queries   + Context queries   + Boolean queries   + Natural Language queries   + Pattern Matching * Structural queries | *Chapter 2,3* | *2,3* | *12* |
| **Using Relevance feedback during IR**   * Vector Model * Probabilistic Model | *R2* | *1* | *3* |
| **Text and Multimedia Languages and Properties**   * MetaData * Text   + Formats   + Information theory   + Modeling natural languages   + Similarity models. * Markup languages * Multimedia: formats, graphics and virtual reality | *Chapter 6* | *6,7* | *12* |
| *Reference with Pointers* |  |  |  |
| **Multimedia IR**   * Access to multimedia information * Content-based retrieval * Language Modeling for Information Retrieval (recap) * Generative Models for Images and Video * Evaluation * Query articulation | *Chapter 11,12* | *8,9* | *12* |
| **Search Engines**   * WWW search engine architecture * Web crawling * HTML parsing * indexing * retrieval * relevance * SPAM * Future of Search | *Chapter 13* | *10,11* | *12* |
| **User Interface and Visualization**   * Introduction * Human Computer Interaction   + Design Principles   + The rule of visualization   + Evaluating interactive systems * Information access process   + Models of interaction   + Earlier interface studies * Starting points   + List of collections   + Overviews   + Examples, dialogues and wizards   + Automated source selection   Query specifications | *Chapter 10* | *12,13* | *12* |
| **Review**  Project Discussions and Presentations | *Chapter 9* | *14,15* | *12* |
| **File System**   * Interface * File Concept * Access Methods * Directory Structure * File Sharing * Protection * File System Structure * File System Implementation * Directory Implementation * Allocation Methods | *Chapter 10* | *14,15* | *12* |

|  |  |  |
| --- | --- | --- |
|  | ***Evaluation*** |  |
| **Assignment and Projects** |  | ***20%*** |
| Project |  | **15%** |
| Presentation & Discussion |  | **5%** |
| **Individual Work** |  | *80%* |
| Attendance, Participation, Home works and short report | Chapter Homework’s, Discussions, Short Presentations | **10%** |
| Quizzes | Unannounced Short quizzes |
| First Exam | Multiple Choice Questions worth 25% and Essay Questions worth 75% of exam grade. | **15%** |
| Second Exam | Multiple Choice Questions worth 25% of and Essay Questions worth 75% of exam grade. | **15%** |
| A Comprehensive Final examination | Multiple Choice Questions worth 25% and Essay Questions worth 75% of exam grade. | **40%** |
| *total* |  | *100%* |

|  |  |
| --- | --- |
|  | ***Policy*** |
| ***Attendance*** | *Attendance is very important for the course. In accordance with university policy, students missing more than the allowed absence rate of total classes are subject to failure. Penalties may be assessed without regard to the student's performance. Attendance will be recorded at the beginning or end of each class.* |
| ***Exams*** | *All exams will be CLOSE-BOOK; necessary algorithms/equations/relations will be supplied as convenient.* |

|  |
| --- |
| **Class Schedule & Room** |

|  |  |  |
| --- | --- | --- |
| **Office Hours** | | |
| Sun: 12 - 2  Mon: 8 - 11  Tues: 12 - 2  Wed: 8 - 11 | | |
|  | \* Or by an appointment through email |  |

|  |  |
| --- | --- |
|  | **Teaching Assistant** |
| To announced later on. |  |

|  |  |
| --- | --- |
|  | **Prerequisites** |
| **Prerequisites by course** | 1001108 |