COMP90049 Knowledge Technologies

Credit Points:	12.5
Level:	9 (Graduate/Postgraduate)
Dates & Locations:	2015, Parkville This subject commences in the following study period/s: Semester 1, Parkville - Taught on campus. Semester 2, Parkville - Taught on campus.
Time Commitment:	Contact Hours: 36 hours, comprising of two 1-hour lectures and one 1-hour workshop per week Total Time Commitment: 200 hours
Prerequisites:	**One of the following:** **Subject** / **Study Period Commencement:** / **Credit Points:** COMP20003 Algorithms and Data Structures — Semester 2 — 12.50 COMP90038 Algorithms and Complexity — Semester 1, Semester 2 — 12.50 COMP20007 Design of Algorithms — Semester 1 — 12.50 **OR** 433-253 Algorithms and Data Structures
Corequisites:	None
Recommended Background Knowledge:	None
Non Allowed Subjects:	**Subject** / **Study Period Commencement:** / **Credit Points:** COMP30018 Knowledge Technologies — Semester 1, Semester 2 — 12.50 **OR** 433-352 Data on the Web
Core Participation Requirements:	For the purposes of considering request for Reasonable Adjustments under the Disability Standards for Education (Cwth 2005), and Student Support and Engagement Policy, academic requirements for this subject are articulated in the Subject Overview, Learning Outcomes, Assessment and Generic Skills sections of this entry. It is University policy to take all reasonable steps to minimise the impact of disability upon academic study, and reasonable adjustments will be made to enhance a student's participation in the University's programs. Students who feel their disability may impact on meeting the requirements of this subject are encouraged to discuss this matter with a Faculty Student Adviser and Student Equity and Disability Support: http://services.unimelb.edu.au/disability
Coordinator:	Assoc Prof Karin Verspoor, Prof Rao Kotagiri
Contact:	email: **karin.verspoor@unimelb.edu.au (mailto:karin.verspoor@unimelb.edu.au)**
Subject Overview:	**AIMS** Much of the world's knowledge is stored in the form of unstructured data (e.g. text) or implicitly in structured data (e.g. databases). In this subject students will learn algorithms and data

	structures for extracting, retrieving and analysing explicit knowledge from various data sources, with a focus on the web. Topics include: data encoding and markup, web crawling, regular expressions, document indexing, text retrieval, clustering, classification and prediction, pattern mining, and approaches to evaluation of knowledge technologies. **INDICATIVE CONTENT** Introduction to Knowledge Technologies; String search; Genomics; Text processing and search; Web search and retrieval; Introduction to Data Mining; Introduction to basic Probability; Classification; Association Rules; Clustering; Evaluation measures. Examples of projects that students may completed are: # A method for automatically predicting the geo-location of a Twitter user on the basis of their posts # An automatic method for tagging multilingual Wikipedia documents with Wikipedia categories # A search engine for Twitter data, which takes into account the time stamp of the query and documents # A search engine for web user forum data # A search engine servicing mixed monolingual queries (as in monolingual queries from a range of languages) over a large-scale document collection # Classification and prediction of some real world problems using machine learning techniques.
Learning Outcomes:	**INTENDED LEARNING OUTCOMES (ILO)** Having completed this unit the student is expected to describe and apply the fundamentals of knowledge systems, including data acquisition and aggregation, knowledge extraction, text retrieval, machine learning and data mining. On completion of this subject the student is expected to: 1 Gain an understanding of a representative selection of knowledge technology techniques in both theoretical and applied contexts 2 Develop familiarity with component technologies used in commonly-deployed knowledge technology systems 3 Get a feel for what research is all about, especially relating to knowledge technology-related projects underway at The University of Melbourne
Assessment:	Project work during semester, requiring approximately 50 - 60 hours of work; one project due approximately mid-semester, and a second due in Week 11 or 12 (40%) One mid-semester test (10%) One 2-hour examination held during the examination period (50%). Hurdle requirement: To pass the subject, students must obtain at least: 50% overall 20/40 in project work, and 30/60 in the mid-semester test and end-of-semester written examination combined. ILO 1 is addressed in the projects (applied) and the mid-semester test and final exam (theoretical). ILO 2 is addressed in the projects (through using a range of systems that are provided to students or that students experiment with themselves). ILO 3 is also addressed in the projects (which are generally themed around projects underway at the University, to give them a more applied feel).
Prescribed Texts:	None
Breadth Options:	This subject is not available as a breadth subject.
Fees Information:	Subject EFTSL, Level, Discipline & Census Date, http://enrolment.unimelb.edu.au/fees
Generic Skills:	On completion of this subject, students should have the following generic skills: # General skills include the ability to undertake problem identification, formulation, and developing solutions especially exploiting acquired data # In addition this subject exposes students to use various data processing tools and make them learn integration of these tools to build more complex software systems # As a result the student will develop skills to utilise a systems approach to complex problems.
Notes:	**LEARNING AND TEACHING METHODS**

	This course is taught over 12 weeks, each week with two one hour formal lectures and a one hour workshop. During the workshops the students are given problems to solve to reinforce the previous week's lecturing material. The problem solving nature of the workshops is geared for the students to learn and understand the concepts of the subject material. **INDICATIVE KEY LEARNING RESOURCES** Christopher D. Manning, Prabhakar Raghavan, and Hinrich Schutze (2008), Information Retrieval, Cambridge University Press. Freely available at informationretrieval.org Pang-Ning Tan, Michael Steinbach and Vipin Kumar (2005) Introduction to Data Mining, Addison-Wesley. **CAREERS / INDUSTRY LINKS** This subject is relevant to many fields including Engineering, Commerce, Government Organizations, Research Institutes and Institutions in Medicine where data analysis can play a significant improvement in delivering services or improving profits.
Related Course(s):	Master of Information Systems Master of Information Systems Master of Information Systems Master of Information Technology Master of Information Technology Master of Philosophy - Engineering Master of Science (Computer Science) Master of Software Systems Engineering Ph.D.- Engineering
Related Majors/Minors/ Specialisations:	Computer Science Computer Science MIS Professional Specialisation MIS Research Specialisation MIT Computing Specialisation MIT Distributed Computing Specialisation MIT Health Specialisation MIT Spatial Specialisation Master of Engineering (Software with Business) Master of Engineering (Software)