

# COMP90042 Web Search and Text Analysis

<b>Credit Points:</b>	12.50									
<b>Level:</b>	9 (Graduate/Postgraduate)									
<b>Dates &amp; Locations:</b>	2012, Parkville This subject commences in the following study period/s: Semester 1, Parkville - Taught on campus.									
<b>Time Commitment:</b>	Contact Hours: 24 one-hour lectures (two-per week), 12 one-hour workshops (one-per week). Total Time Commitment: 120 hours.									
<b>Prerequisites:</b>	One of the following: <table border="1" data-bbox="387 573 1485 779"> <thead> <tr> <th>Subject</th> <th>Study Period Commencement:</th> <th>Credit Points:</th> </tr> </thead> <tbody> <tr> <td>COMP30018 Knowledge Technologies</td> <td>Semester 2</td> <td>12.50</td> </tr> <tr> <td>COMP90049 Knowledge Technologies</td> <td>Semester 1</td> <td>12.50</td> </tr> </tbody> </table>	Subject	Study Period Commencement:	Credit Points:	COMP30018 Knowledge Technologies	Semester 2	12.50	COMP90049 Knowledge Technologies	Semester 1	12.50
Subject	Study Period Commencement:	Credit Points:								
COMP30018 Knowledge Technologies	Semester 2	12.50								
COMP90049 Knowledge Technologies	Semester 1	12.50								
<b>Corequisites:</b>	None									
<b>Recommended Background Knowledge:</b>	None									
<b>Non Allowed Subjects:</b>	433-460 Human Language Technology 433-467 Text and Document Management 433-660 Human Language Technology 433-667 Text and Document Management 433-476 Text and Document Management									
<b>Core Participation Requirements:</b>	For the purposes of considering request for Reasonable Adjustments under the Disability Standards for Education (Cwth 2005), and Students Experiencing Academic Disadvantage Policy, academic requirements for this subject are articulated in the Subject Description, Subject Objectives, Generic Skills and Assessment Requirements of this entry. The University is dedicated to provide support to those with special requirements. Further details on the Disability support scheme can be found at the Disability Liaison Unit Website: <a href="http://www.services.unimelb.edu.au/disability/">http://www.services.unimelb.edu.au/disability/</a>									
<b>Coordinator:</b>	Assoc Prof Steven Bird									
<b>Contact:</b>	Associate Professor Tim Baldwin email: <a href="mailto:tbaldwin@unimelb.edu.au">tbaldwin@unimelb.edu.au</a> ( <a href="mailto:tbaldwin@unimelb.edu.au">mailto:tbaldwin@unimelb.edu.au</a> )									
<b>Subject Overview:</b>	The web is a vast and expanding storehouse of semi-structured textual information. Accessing and processing this information is one of the major challenges of the information age. In this subject, students study the technologies behind search engines, spam filtering, plagiarism detection, information extraction, question answering and newly emerging fields of information engineering. Topics include: web indexing, query evaluation, probabilistic language modelling, document classification and filtering, grammar and spelling correction, topic detection, cross-language information retrieval, machine translation and summarisation.									
<b>Objectives:</b>	On completion of this subject students should be able to: <ul style="list-style-type: none"> <li># Articulate issues relevant to the efficient implementation of web search systems and information retrieval systems</li> <li># Apply information retrieval methodologies as they relate to textual data</li> <li># Apply symbolic and statistical natural language processing techniques in textual analysis tasks</li> </ul>									

	<ul style="list-style-type: none"> <li># Develop and evaluate computational models of language</li> <li># Apply core information engineering technologies in the management and exploitation of online information</li> </ul>
<b>Assessment:</b>	Two collaborative and/or individual projects due around weeks 6 and 11 of semester expected to take about 36 hours (20% each) A research-oriented workshop presentation (10%) And an end-of-semester written examination not exceeding 3 hours (50%).
<b>Prescribed Texts:</b>	None
<b>Breadth Options:</b>	This subject is not available as a breadth subject.
<b>Fees Information:</b>	Subject EFTSL, Level, Discipline & Census Date, <a href="http://enrolment.unimelb.edu.au/fees">http://enrolment.unimelb.edu.au/fees</a>
<b>Generic Skills:</b>	<p>On completion of this subject students should have the:</p> <ul style="list-style-type: none"> <li># Ability to undertake problem identification, formulation, and solution</li> <li># Ability to utilise a systems approach to complex problems and to design for operational performance</li> <li># Ability to manage information and documentation</li> <li># Capacity for creativity and innovation</li> <li># Ability to communicate effectively, with the engineering team and with the community at large</li> </ul>
<b>Related Course(s):</b>	Bachelor of Computer Science (Honours) Master of Engineering in Distributed Computing Master of Science (Computer Science) Master of Software Systems Engineering
<b>Related Majors/Minors/Specialisations:</b>	B-ENG Software Engineering stream Computer Science Master of Engineering (Software)