# Subject Description Form

<table>
<thead>
<tr>
<th>Subject Code</th>
<th>COMP5324</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subject Title</td>
<td>Internet Information Retrieval</td>
</tr>
<tr>
<td>Credit Value</td>
<td>3</td>
</tr>
<tr>
<td>Level</td>
<td>5</td>
</tr>
<tr>
<td>Pre-requisite/ Exclusion</td>
<td>Nil</td>
</tr>
</tbody>
</table>

## Objectives
The objectives of this subject are to:

1. Study the fundamental knowledge and techniques in information retrieval (IR) and information extraction (IE);
2. apply the IR and IE fundamentals to various internet applications; and
3. explore the use of information retrieval technology in advanced IR internet applications, like information filtering.

## Intended Learning Outcomes
Upon completion of the subject, students will be able to:

a) be aware of various classical information retrieval models;
b) comprehend the main difference between classical information retrieval and Internet information retrieval, and handle the problems particularly associated to Internal information retrieval;
c) design and implement effective retrieval systems;
d) apply retrieval evaluation techniques to improve retrieval system; and
e) extract important pieces of information from the retrieved text and convert from unstructured text to structured database.

## Subject Synopsis/Indicative Syllabus
- **Presentation of Information in the Internet**: Tagging and Processing: HTML, XML and SGML.
- **Classical Information Retrieval**: Architecture, IR models, Term selection and weighting, Ranking, Query processing, Evaluation techniques, indexing and search engine fundamentals.
- **Information Extraction**: Extraction: Keyword identification, NP extraction, String pattern extraction
- **Distributed Information Retrieval**: Web-graph analysis, Server ranking, Meta search engines.
- **Applications**: Digital library, Wireless information access
- **Advanced Information Retrieval**: Relevance feedback, Advanced indexing techniques, issues in Multilingual/Multimedia information retrieval, Information filtering and text categorization.

## Teaching/Learning Methodology
Class activities including - lecture, tutorial, lab, workshop seminar where applicable

## Assessment Methods in Alignment with Intended Learning Outcomes

<table>
<thead>
<tr>
<th>Specific Assessment Methods/Tasks</th>
<th>% weighting</th>
<th>Intended subject learning outcomes to be assessed</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>b</td>
<td>c</td>
</tr>
<tr>
<td>d</td>
<td>e</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Assignments, Tests &amp; Projects</td>
<td></td>
</tr>
<tr>
<td>--------------------------------</td>
<td>-------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>Final Examination</td>
<td>45</td>
<td>✔</td>
</tr>
<tr>
<td>Total</td>
<td>100</td>
<td></td>
</tr>
</tbody>
</table>

**Student study effort expected**

**Class Contact:**
- Class activities (lecture, tutorial, lab) 39 hours

**Other student study effort:**
- Assignments, Quizzes, Projects, Exams 65 hours
- **Total student study effort 104 hours**

**Reading list and references**