I. General Information
Time/location: MW 3:30-4:45pm, CI-117
Professor: Dr. Longzhuang Li
Office: CI-323
Office hours: MW, 1:00m - 2:00pm, or by appointment
Phone: (361) 825-2406
Email: Longzhuang.Li@tamucc.edu
http://www.sci.tamucc.edu/~lli/teaching/Fall2010/cosc5336/cosc5336.htm
Prerequisite: COSC 5335 and COSC 5321

II. Catalog description
Data integration has been recognized as a research topic of big practical importance. The availability of integrated data from multiple independent, heterogeneous data sources is crucial for many applications. Data integration requires combining and matching information in different sources, and resolving a variety of conflicts. With the number of data sources growing very quickly, data integration is bound to become even more important in the future. This course will survey selected theoretical and practical issues arising in data integration. Also includes the study of contemporary database management system concepts, terminology, and methodology for design, use, and implementation. Emphasis is on the relational model.

III. Student Learning Outcomes
Upon successful completion of this course, a student will
• comprehend the essential principles of the design, analysis and use of contemporary DBMS systems;
• be able to implement Web database applications that interact with a back end DBMS;
• be familiar with techniques associated with data integration;

IV. Course Evaluation
The course grade will be determined as follows:
- Project and term research paper 50%
- research paper presentation 15%
- exam 25%
- class attendance and participation 10%

V. Late Submissions
Late submissions are allowed but penalties apply. Late submissions will receive a 10% penalty each day.

VI. Class Attendance
Class attendance is strongly encouraged. Good attendance and class participation will count up to 10% in the final grade. Five or more unexcused absences will make the student ineligible for this 10%. If the student is absent from class, he or she is responsible for any materials covered, handouts and any announcements made in class, regarding (but not limited to) class schedule, homework, project and exams.
VII. Tentative Class Schedule

<table>
<thead>
<tr>
<th>Week</th>
<th>Topic</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Review the basics of DBMS</td>
</tr>
<tr>
<td>2</td>
<td>Overview of data integration</td>
</tr>
<tr>
<td>3 - 6</td>
<td>Meditor-based integration techniques</td>
</tr>
<tr>
<td>7</td>
<td>Schema matching methods</td>
</tr>
<tr>
<td>8 - 10</td>
<td>Peer-to-peer based data integration</td>
</tr>
<tr>
<td>11</td>
<td>Data archive techniques</td>
</tr>
<tr>
<td>12 – 13</td>
<td>Dataspase</td>
</tr>
</tbody>
</table>
and/or guidance in the grade appeal process, students may contact the Office of Student Affairs.

*Note: Any changes to the syllabus will be announced in class. It is the student's responsibility to obtain this information.