Outline

- Action Items
- Component Design
- Test Evaluation
- Project Evaluation
- Future Work
- Lessons Learned
- Demonstration
- Questions / Comments
Action Items

- Future Project Use and Extension
  - Created a FileLoader class for easy project extension
  - Created Technical Instructions for adding onto the project
Action Items (cont.)

- Update Project Plan to include new cost estimate and details
  - Project Plan was revised to include the new information
Action Items (cont.)

- Investigate derivation of Webpage and KrestEntity in UML Model
  - Updated WebObject and KrestEntity objects to show that they are abstract classes
  - Updated the links to show the derivation relationship

- Investigate mutual exclusion of types for WebObject in formal specification
  - Updated OCL to include invariants for WebObject
Component Design

- Developed using Fujaba
- Contains class variables and functions
- Added separate table to describe the roles of functions
Component Design

- Overall Package View
Component Design

- Controller Package:
Component Design

- View Package:
Component Design

- Model Package:
Component Design (cont.)

### EntityObserver Class

<table>
<thead>
<tr>
<th>Class</th>
<th>Visibility</th>
<th>Extends</th>
<th>Implements</th>
</tr>
</thead>
<tbody>
<tr>
<td>EntityObserver</td>
<td>public</td>
<td>none</td>
<td>none</td>
</tr>
<tr>
<td>matchingEntitiesTable</td>
<td>public</td>
<td>JTable</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Function</th>
<th>Visibility</th>
<th>Parameters</th>
<th>Returns</th>
<th>Actions</th>
</tr>
</thead>
<tbody>
<tr>
<td>EntityObserver</td>
<td>public</td>
<td>JTable</td>
<td>void</td>
<td>Constructor to set up the class</td>
</tr>
<tr>
<td>sortListByMatchSize</td>
<td>private</td>
<td>ArrayList</td>
<td>ArrayList</td>
<td>Sorts the entities found by the number of pages they were found on</td>
</tr>
<tr>
<td>updateEntitySearchResults</td>
<td>public</td>
<td>KrestObjectLibrary, String</td>
<td>void</td>
<td>Updates the matching entities found</td>
</tr>
</tbody>
</table>
## Test Evaluation

<table>
<thead>
<tr>
<th>Test Case</th>
<th>Main Functionality Tested</th>
<th>Pass/Fail</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test Case 1</td>
<td>Application Functionality</td>
<td>PASS</td>
</tr>
<tr>
<td>Test Case 2</td>
<td>Web Crawling Functionality</td>
<td>PASS</td>
</tr>
<tr>
<td>Test Case 3</td>
<td>Web Searching</td>
<td>PASS</td>
</tr>
<tr>
<td>Test Case 4</td>
<td>Entity Searching</td>
<td>PASS</td>
</tr>
<tr>
<td>Test Case 5</td>
<td>Reproducing the results from Tao Cheng’s work</td>
<td>PASS</td>
</tr>
</tbody>
</table>
Test Evaluation (cont.)

- Reproducing Tao Cheng’s work
  - Downloaded 12 separate data sets from WebBase repository
    - Data sets ranged from 300 – 1200 web pages each
    - The data sets each contained one item of contact info from Tao’s work
    - Helped ensure that the entity searcher could find specific contact info
Problems encountered:

- Two problems found
  - Web crawl stopping at random time
    - Due to internet connection issues at home
  - Case sensitivity in web and entity search
Project Evaluation

- Problems encountered
  - Web Crawler Thread Control
  - Java Class Size Limitations
  - Jigloo GUI Builder
Project Evaluation (cont).

- **SLOC Evaluation**
  - Phase One Estimate: 2K
  - Phase Two Estimate: 2.3K
  - Final SLOC Developed: 2.9K
    - 1.4K Comment Lines

- **Analysis**
  - Estimates fairly close
  - Most of the extra code was generated by Jigloo
    - get( ) methods for widgets
Project Evaluation (cont).

- Project Schedule:

<table>
<thead>
<tr>
<th>Phase</th>
<th>Expected Completion Date</th>
<th>Actual Completion Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>November 13, 2007</td>
<td>November 13, 2007</td>
</tr>
<tr>
<td>2</td>
<td>February 15, 2008</td>
<td>February 13, 2008</td>
</tr>
<tr>
<td>3</td>
<td>April 25, 2008</td>
<td>April 21, 2008</td>
</tr>
</tbody>
</table>
Project Evaluation (cont).

- Time Spent During in Each Phase:

![Pie chart showing time spent in each phase]

- Phase 1: 55.92 hours, 37%
- Phase 2: 57.83 hours, 38%
- Phase 3: 37.67 hours, 25%
Project Evaluation (cont).

- Overall Task Breakdown:

![Pie chart showing time spent on various project activities]
Project Evaluation (cont).

- Phase 1 Task Breakdown:

![Pie chart showing time spent per project activity during Phase 1 (in Hours)]

- Discussion: 4.33, 7.75%
- Research: 1.92, 3.43%
- Reading: 9.50, 16.99%
- Webpage: 1.00, 1.79%
- Presentation: 3.67, 6.56%
- Documentation: 17.50, 31.30%
- Environment: 4.58, 8.20%
- Coding: 12.58, 22.50%
- Timelog: 0.83, 1.49%
- Integration: 0.00, 0.00%
Phase 2 Task Breakdown:

- Coding, 30.33 hours, 52.45%
- Documentation, 18.67 hours, 32.28%
- Webpage, 0.25 hour, 0.43%
- Presentation, 7.92 hours, 13.69%
- Timelog, 0.67 hour, 1.15%

![Pie chart showing time spent per project activity during Phase 2](chart.png)
Project Evaluation (cont).

- Phase 3 Task Breakdown:

![Pie chart showing time spent per project activity during Phase 3 in hours: Coding 15.00, 40%; Integration 0.75, 2%; Documentation 21.00, 56%; Presentation 0.00, 0%; Timelog 0.00, 0%; Discussion 0.17, 0%; Webpage 0.75, 2%.]
Future Work

- Described in more detail in Software Technical Instruction for Reuse and Extension Document
  - Add Open Source Web Crawler
  - Implement Database for Crawled/Loaded Web Pages
  - Add support for more loading more types (e.g. XML)
Lessons Learned

- **Design Development**
  - MS Visio
  - Lots of design tools exist, many aren’t very robust

- **Java Tool Usage**
  - Eclipse IDE
  - GUI builders – should do more in depth search before choosing one
Demonstration
Final Steps

- Final Examination Ballot
- Make any necessary changes to the MSE Portfolio
- Deliver the Portfolio
- Put ISO of all MSE files on KDD Tiki page
Questions / Comments