Opportunity

These special topic courses are intended as a forum within which to explore the development of a variety of specialized systems to support an accessible and controlled public information system.

The opportunities include:

- Efficient manipulation of XML
- DBMS controlled data/documents
- Location based services
- Development of spiders and RSS systems
- Applets and Java applications
- Streaming video

RAPAI is an embedded information portal supporting decentralized role based access and differentiation.
Research Possibilities

- Distributed systems
- Web services
- Directory services
- Agents
- Structured document systems
  - RSS feeds
  - Database documents
  - XPath, XSLT and XSLT-FO
- Interactive systems
  - Navigation & Visualization
  - Location Based Services
  - Embedded systems

Course Expectations

- Learning what you need to know
- Facility with mail, ftp, and utilities
- Knowledge of operating systems, particularly unix
- Ability to use development tools
- Ability to write code, any language
- Preparation
  - Reading and thinking
  - Thinking and “doing” – beyond what I have “done”
  - Doing a complete job
- Engagement
  - Communication
  - Experimentation
  - Documentation

Course Expectations (2)

Ignorance is OK…
Stupidity is not

Mistakes are acceptable…
Carelessness is not
Course Expectations (3)

Don’t use tools/code
You don’t understand

Don’t just do something
Stand there

Don’t just stand there
Do something

Course Resources

• The textbook – read and digested
• The lectures – they will go beyond and around the readings and assignments
• The web and manuals – there is more around than we will ever digest
  • Web standards
  • New ideas – e.g. digital libraries
• Experimentation and discussion with other students

Course Overview

• Understand system
• Document system
• Clean up system
• Augment system
• Develop new subsystems
  • Directory access
• XML Basics
• Java
• Java for XML
  • Production
  • Access
  • Manipulation
• XSLT
Imperatives for the Seminar

- Speak up and ask for help
- Plan for the term – understand what needs to be done
- Get the big picture and understand where the “cracks” are
- Learn thoroughly what you need to succeed:
  - XML
  - Unix
  - String handling
  - Protocols

First Steps

- Learn about Unix
  - Where are things located
  - How do web servers work
  - What are file protections
  - How to comment files
  - How to name files
- Learn about RAPAI
  - How do we access Oracle
  - Where is tomcat
  - Where is apache

Web Technology Overview

- BROWSER Javascript Vbscript HTML CSS XML Applet Plug-in
- SERVER Request mgt. Security Logging Program mgt. Output
- Pages for delivery Programs that produce pages
- Java Application
- Java Application
Unix OS

- Unix is still an important environment for web servers
- Under Unix how do you…
  - Add and modify users
  - Change file protections and access rights
  - Run programs that are X window system based
  - Compile and run java programs
  - Establish cron files
  - Start and stop servers

Java Programming

- Programming languages use core components and extensions
- The core components of a language include:
  - Declarations, assignments, and operators
  - Flow control
    - Sequencing
    - Selection
    - Repetition
  - I/O operations
- The extensions involve libraries, packages, modules, etc., that can be reused by others:
  - Network interfaces
  - DBMS interfaces
  - String handling

Standards

- Standards are critical to understanding the web.
- Reading standards is a little intimidating at first
- The basic standards are:
  - HTML4.0/XHTML2.0
  - HTTP1.1
- The XML suite includes
  - XML
  - XPath
  - XSLT
- WWW standards are available online at w3c.org
Strings

- Strings are represented on different machines in different ways
  - ftp file transfers – images and perl scripts
  - Unix defines a string as any set of characters terminated a null – the character whose value is 0
  - A newline can be a part of a string
  - A non-terminated string can lead to segmentation faults
  - There are a whole series of functions to process strings in different languages
  - Perl has extensive string processing capabilities

Protocols

- A protocol, in the client server world, is the rules by which two programs communicate
  - There are many different kinds of protocols
  - The http protocol is a very simple one
    - The client makes a request
    - The server responds to the request
  - In reality the http protocol is a little more complex in that it sits on top of TCP/IP
    - The host location is resolved
    - A connection is made
    - A request is made and responded to
    - The connection is closed

Projects

- Spiders – URL normalization, page parsing, history management, data analysis
- RSS feeds
- Servlets
- Voice XML/IM interfaces
- RMI applications
- XML processing
- Directory server for nomadic systems
- Location based services for students
A dream list for the system

- Anyone at any time can enter simple data to create fancy posters and announcements
- People walk up to displays that change for them
- Streams of information relevant to the profession constantly get updated on walls
- All panels are interactive based on need
- Students get tickled by announcements that are important for them to be aware of

Other Projects

- Faculty create syllabi and they are automatically moved to multiple formats and made accessible
  - Webpage
  - Courseinfo
  - Department listings
- Staff have single source pages that impact all needed actions
  - Mail to registrar
  - RAPAI updates
  - Website updates

First Steps

- Figure out how the system operates
- Grab and clean a piece of code
- Improve the performance of a piece of code