

Faculty Showcase '05: Partners in Learning

Sponsored by Academic Technology at Simmons College



Gerald Benoit
Graduate School of
Library and
Information Science

Goals

- Create re-usable library of Java tools to parse full-text documents and XML records;
- Use the parsed records for novel resource discovery in biomedical, mixed-language, and full-text documents;
- Create 2D and 3D interfaces with control panels so end-users can manipulate the data in ways that make sense to them.

Overview

“Information Retrieval” (IR) is a computer-based solution to locating machine-readable resources in response to an end-user’s information need. Internet search engines such as Google are one example of IR.

Unlike the traditional database approach, IR applies statistical probabilities of word occurrence in documents to the end-user’s expression, called the query. Improving the match between query and document collection is one of the main research goals within the community of people working on IR. In addition, IR research focuses on developing interactive interfaces designed to help end-users understand retrieval results.

Gerald Benoit has applied the parser and interface tools library to several projects: medical records, molecular biology literature, and collections of mixed-language documents. As a student at UCLA he worked on a NSF-funded visualization project and as an assistant professor at the University of Kentucky he applied this work to an IMLS-funded project to discover unanticipated links in the biomedical literature.

These libraries are being converted from Java servlets to signed applets and tested in new environments: an XML-Maker utility for web-based digital library creation and a program to integrate locally-created records with records distributed over the Internet.

Applets for Information Retrieval

Zazu - Curriculum Planning Applet
http://web.simmons.edu/~benoit/zazuApplet/Applet1.html

Area: INFOTECH Electives: ARCHIVES # of classes/term: 4 All Campuses Update

Term 1: LIS405, LIS426, LIS433, LIS438
Term 2: LIS439, LIS440, LIS455, LIS455
Term 3: LIS456, LIS458, LIS461, LIS461
Term 4: LIS462, LIS471, LIS486, LIS498
LIS446, LIS440, LIS461, LIS458, LIS465, LIS461

Classes to skip: Core Archives IT School Media
 LIS403 LIS404 LIS407 LIS405 LIS488 LIS406
 LIS408 LIS409 LIS413 LIS414 LIS416 LIS417
 LIS418 LIS419 LIS420 LIS421 LIS424 LIS426
 LIS427 LIS428 LIS430 LIS431 LIS433 LIS434
 LIS435 LIS437 LIS438 LIS439 LIS440 LIS446
 LIS450 LIS451 LIS452 LIS453 LIS454 LIS455
 LIS456 LIS457 LIS458 LIS460 LIS461 LIS462
 LIS465 LIS470 LIS471 LIS475 LIS479 LIS481
 LIS482 LIS483 LIS484 LIS485 LIS486 LIS489
 LIS490 LIS492 LIS495 LIS498

Classes to include:
 LIS403+ LIS404+ LIS407+ LIS405+ LIS488+ LIS406+
 LIS408+ LIS409+ LIS413+ LIS414+ LIS416+ LIS417+
 LIS418+ LIS419+ LIS420+ LIS421+ LIS424+ LIS426+
 LIS427+ LIS428+ LIS430+ LIS431+ LIS433+ LIS434+
 LIS435+ LIS437+ LIS438+ LIS439+ LIS440+ LIS446+
 LIS450+ LIS451+ LIS452+ LIS453+ LIS454+ LIS455+
 LIS456+ LIS457+ LIS458+ LIS460+ LIS461+ LIS462+
 LIS465+ LIS470+ LIS471+ LIS475+ LIS479+ LIS481+
 LIS482+ LIS483+ LIS484+ LIS485+ LIS486+ LIS489+
 LIS490+ LIS492+ LIS495+ LIS498+

Class numbers in *italics* indicates the class is a prerequisite.