

# Open-Corpus Adaptive Hypermedia

Peter Brusilovsky  
School of Information Sciences  
University of Pittsburgh, USA  
<http://www.sis.pitt.edu/~peterb>

## Adaptive Hypermedia

- Hypermedia systems = Pages + Links
- Adaptive presentation
  - content adaptation
- Adaptive navigation support
  - link adaptation

## Adaptive Navigation Support

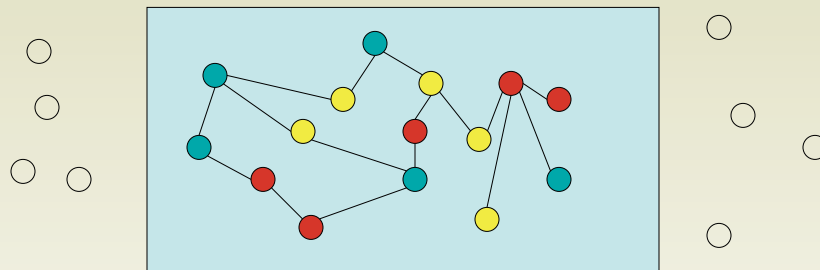
- Direct guidance
- Hiding, restricting, disabling
- Generation
- Ordering
- Annotation
- Map adaptation

## The Value of ANS

- Lower navigation overhead
  - Access the content at the right time
  - Find relevant information faster
- Encourages non-sequential navigation
  - Better use of *true hypertext* links
- Better learning outcomes
  - Achieve the same level of knowledge faster
  - Better results with fixed time

## The Problem

- Nearly all popular and efficient adaptive hypermedia technologies were built to operate with a relatively small set of documents that were structured and enhanced by metadata annotations at design time

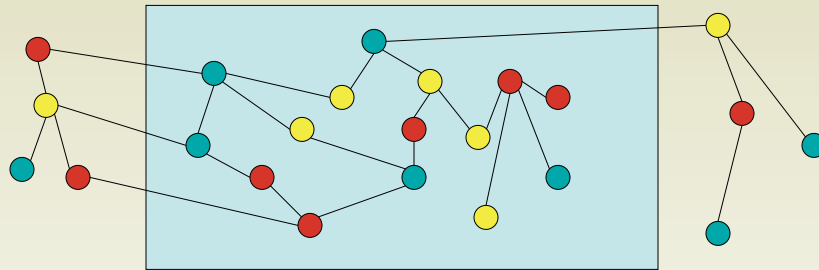


## Closed and Open Corpus AH

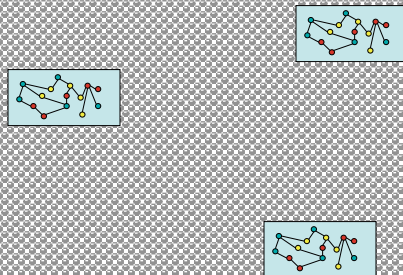
- Definition 1 (Closed Corpus Adaptive Hypermedia System)**
  - A closed corpus adaptive hypermedia system is an adaptive hypermedia system which operates on a closed corpus of documents, where documents and relationships between the documents are known to the system at design time.
- Definition 2 (Open Corpus Adaptive Hypermedia System)**
  - An open corpus adaptive hypermedia system is an adaptive hypermedia system which operates on an open corpus of documents, e.g., a set of documents that is not known at design time and, moreover, can constantly change and expand.

## The Open Corpus Problem

- Provide adaptation within a set of documents that is not known at design time and, moreover, can constantly change and expand



## The Open Corpus Problem in the Web Age

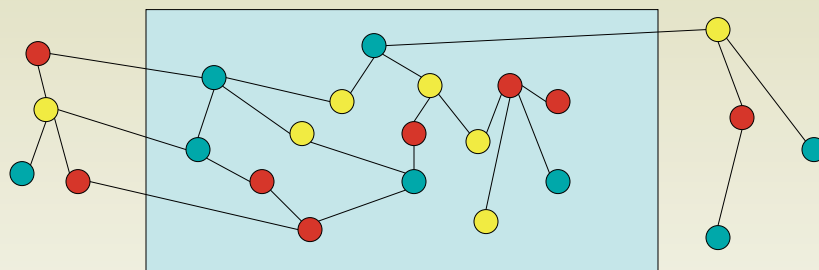


## Why it is a problem?

- Adaptive E-learning
  - Adaptive Java Tutorial vs. hundreds of Java books and Web pages
- Adaptive Tourist Guide
  - Guide pages vs. information about the same city from other sources
- Adaptive News System
  - Google News vs. news from other news providers and blogs

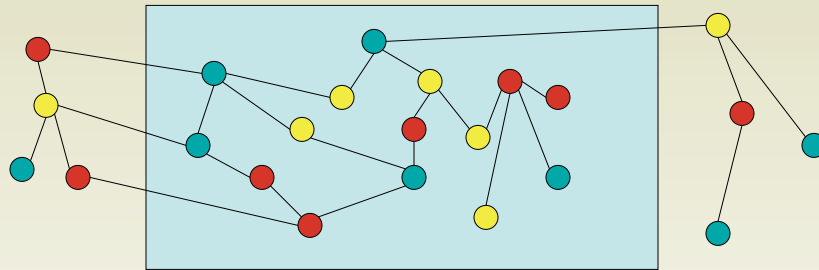
## The Anatomy of the Open Corpus Problem in AH

- Provide browsing-based access to open corpus (linking)
- Guide the user to the most appropriate content (adaptive navigation support)
- Present the open corpus content adaptively (adaptive presentation)



## The Anatomy of the Open Corpus Framework

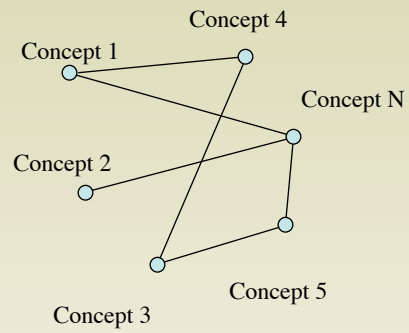
- Provide browsing-based access to open corpus (linking)
- Guide the user to the most appropriate content (adaptive navigation support)
- Present the open corpus content adaptively (adaptive presentation)



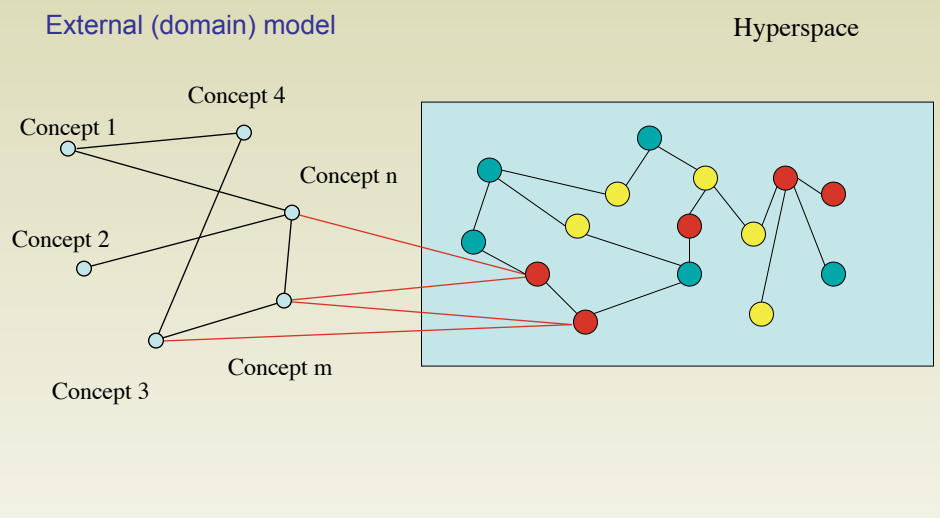
## How it is Done in Classic AH

- Classic AH use external models
  - Domain models, pedagogical modes, stereotype hierarchy, etc.
- Users are modeled in relation to these models
  - User is *field-independent*
  - User knowledge of *loops* is *high*
  - User is interested in *19th century architecture styles*
- Resources are connected (indexed) with elements of these models (aka *knowledge behind pages*)
  - This section presents *while loop* and *increment*
  - This page is for *field-independent* learners
  - This church is built in 1876

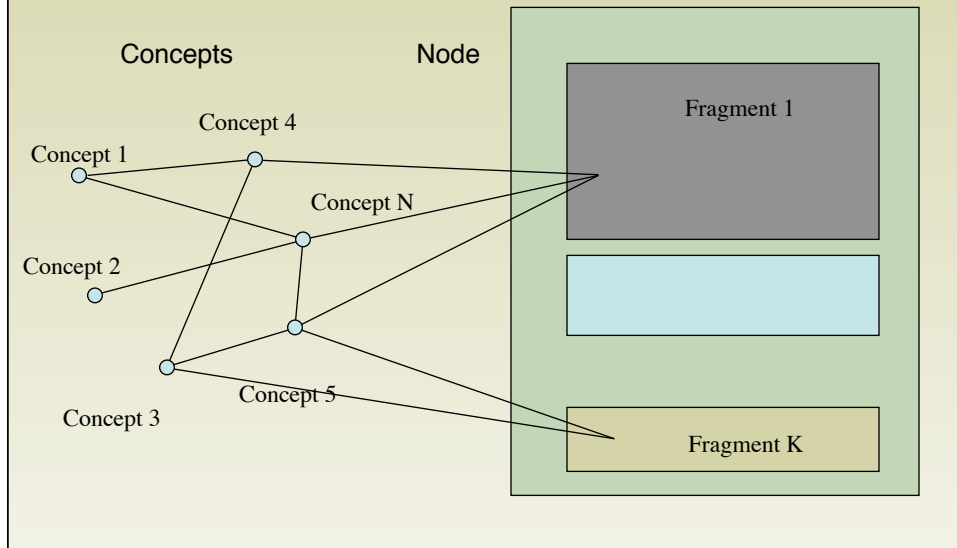
## An External Model



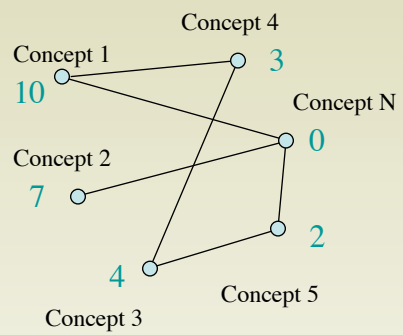
## Indexing of Nodes



## Indexing of Fragments



## Concept-Level User Model





## How to do it for the OC?

- How to build hyperspace?
- Where we can get external models?
- How we can index the hypertext nodes to accumulate “knowledge behind pages”?
- How we can build and maintain user models?

## Personalized Information Access 2000

- Adaptive IR systems (IR, from 1980)
  - Use word-level profile of interests and remedial feedback to adapt search and result presentation
- Adaptive hypermedia (HT, ITS, from 1990)
  - Use explicit domain models and manual indexing to deliver a range of adaptation effects to different aspects of user models
- Web recommenders (AI, ML, from 1995)
  - Use explicit and implicit interest indicators, apply clickstream analysis/log mining to recommend best resources for detected use interests
  - Content-based recommenders
  - Collaborative recommenders

# Personalized Information Access 2000

**HCI / HT**

Adaptive Hypermedia

- Concept-level domain models
- Concept-level user model
- Manual indexing at design time
- Use many adaptation techniques
- Adapt to many user factors
- Expressive, reliable adaptation

**AI / IR**

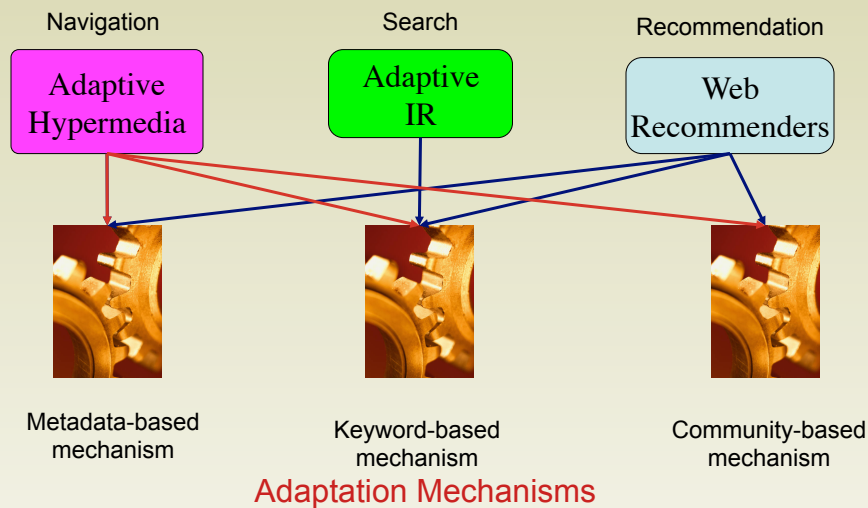
Adaptive IR

Web Recommenders

- No domain model
- Keyword-level user model
- No manual indexing
- Adapt to user interests
- Use ranked list of links/docs

## A Look under the Hood

Types of information access



## Building Open Corpus Adaptive Hypermedia with:

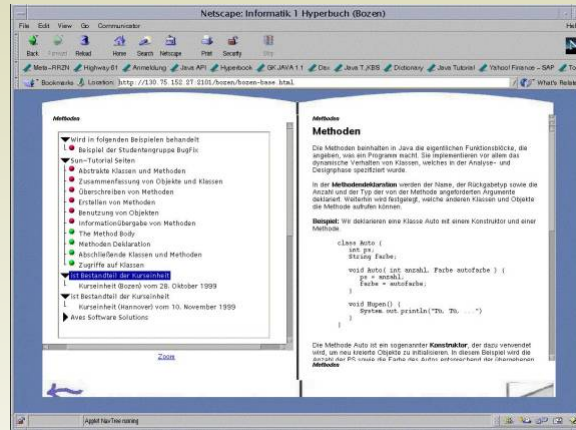
- Classic metadata-based (concept-based) mechanisms
  - Why not? If indexing can be done *after* the system design time
- Community-based mechanisms
  - Indexing done by users
- Keyword-based mechanisms
  - Classic IR text processing and indexing approaches

## Metadata-based OCAH

- Full-blown concept-level manual indexing
  - KBS-Hyperbook, SIGUE
- Simplified concept-level manual indexing: categorization
  - Topic-based adaptation in Quiz-GUIDE
- Automatic concept-level indexing
  - ELDIT, NavEx, concept-based Quiz-GUIDE
- Community-driven indexing
- Using metadata-enriched content
  - *Standard metadata*: Proactive
  - *Semantic Web*: Personal Reader

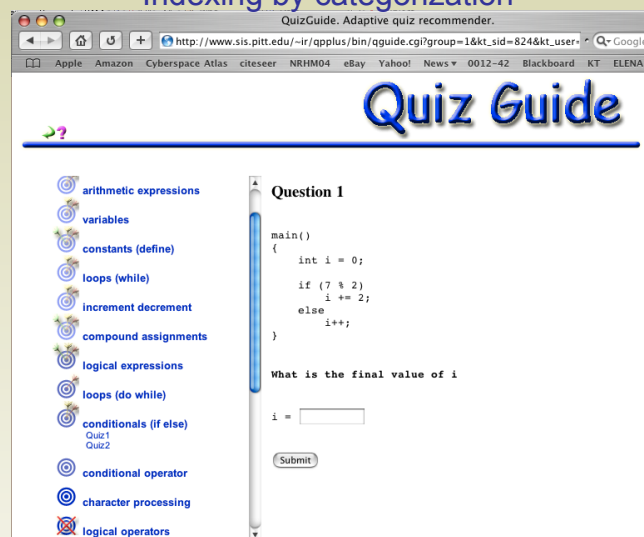
# KBS-HyperBook: Expandable AH

## Integrating new resources by indexing



# QuizGuide: Topic-Based AH

## Indexing by categorization



## NavEx: Automatic Indexing

Classic “traffic light” prerequisite-based mechanism based on automatic indexing

The screenshot shows a web browser window titled "L10: capital.c - Microsoft Internet Explorer". The address bar shows a URL from "http://kt2.exp.sis.pitt.edu:8080/". On the left, a navigation menu lists topics: L08: printing2.c, L09: exchange\_kiosk.c, L09: variables.c, L10: better\_kiosk.c, L10: capital.c (selected), L10: fahr\_input.c, L11: capitalyears.c, L11: countdown1.c, L11: countdown2.c, L11: factorial.c, L12: accumulator.c, L12: apples\_kids.c, L12: capitatable.c, and L12: f2c table.c. The main content area is titled "L10: capital.c" and "Code Dissection". It contains "Example 3.3" with the instruction "Click on a green bullet to see the annotation." Below this is a code snippet:
 

```
#include <stdio.h>
main() {
    float
    interest_rate; /*
    interest rate in
    percents */
    int capital; /*
```

 A yellow highlight covers the code from "#include <stdio.h>" to "during". To the right of the code, an annotation reads: "This is a preprocessor command. System header file stdio.h will be inserted there during".

## Concept-Based QuizGuide

The screenshot shows a web browser window titled "QuizGuide: Topic-based Adaptive Navigation for Quizzes - Mozilla Firefox". The interface includes a sidebar on the left with a list of SQL concepts: SELECT-FROM, Arithmetic Expressions, SELECT-FROM-WHERE, Pattern Matching, Multiple Table Queries, ORDER-BY, Set Operations, Aggregate Functions, GROUP-BY and HAVING, and Sub-Queries. The main content area is titled "Question:" and contains the text: "Based on the tables below, write the required SQL expression." and "Task: Show all the information contained in table 'store'." Below the text is a text input field with the placeholder "Enter your answer here." and two buttons: "Submit Answer" and "Go to SQL-Lab". At the bottom, there is a table with the following data:
 

actor_id	first_name	last_name	last_update
1	PENELOPE	GUINNESS	2006-02-15 04:34:33.0
2	NICK	WAHLBERG	2006-02-15 04:34:33.0

 Below the table, there are two more table headers: "address(+)" with columns address\_id, address, address2, district, city\_id, postal\_code, phone, last\_update; and "category(+)" with columns category\_id, name, last\_update.

# Proactive: Metadata for ANS

Recommendation and navigation support based on available metadata indexing

**Proactive** [Preference] [Feedback] [Help] [Logout]

Welcome to Proactive, danielle

[\[Most Recent Jobs\]](#)
[\[Matched to Preference\]](#)
[\[Recommended Jobs\]](#)
[\[My Saved Jobs\]](#)
[\[Advanced Search\]](#)

Job recommendation based on your saved jobs ::

Total: 50(1/3) Matched to Preference Visualize this list 1 2 3

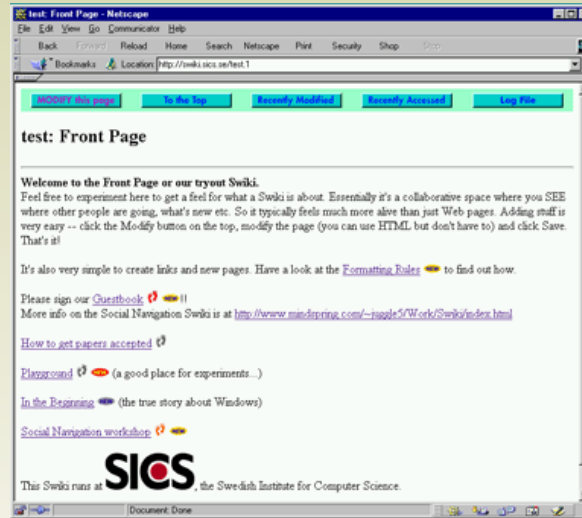
Job Category	Title	Company	City	State	Position Type	Salary	Experience	Education	Post Date	Relevance
Help Desk Analyst	Help Desk Support Analyst <a href="#">[SAVE]</a>	Cost Plus World Market	Oakland	CA	Full-Time, Employee	Unspecified	1-2 Years Experience	Associates	12-15-07	*****
E-Commerce Director	Director/VP of Applications-Wireless <a href="#">[SAVE]</a>	Terran Systems	San Jose	CA	Full-Time, Employee	Unspecified	10-15 Years Experience	Master of Science	12-15-07	*****
Java Server Engineer	Java Server Engineer <a href="#">[SAVE]</a>	Cross Creek Systems	San Jose	CA	Full-Time, Employee	Unspecified	5-10 Years Experience	Bachelor of Science	12-15-07	****
Information Services Consultant	Financial Services Consultant <a href="#">[SAVE]</a>	IBM	Boston	MA	Full-Time, Employee	Unspecified	0-1 Years Experience	unspecified	12-15-07	****
Information Services Consultant	Oracle Service Consultant <a href="#">[SAVE]</a>	IBM	Boston	MA	Full-Time, Employee	Unspecified	Unspecified	unspecified	12-15-07	****
Java Server Engineer	QA Engineer - Web UI Automation (Selenium/Java) <a href="#">[SAVE]</a>	Yahoo! Inc	Santa Clara	CA	Full-Time, Employee	Unspecified	2-5 Years Experience	unspecified	12-15-07	****
Help Desk Analyst	HELP DESK ANALYST - BOSTON <a href="#">[SAVE]</a>	SuccessFactors.com	Boston	MA	Full-Time, Employee	Unspecified	2-5 Years Experience	Bachelor of Science	12-15-07	****
Java Server Engineer	Principal SW Engineer - Ad Server <a href="#">[SAVE]</a>	Yahoo! Inc	Santa Clara	CA	Full-Time, Employee	Unspecified	5-10 Years Experience	unspecified	12-15-07	****
Java Server Engineer	Java Developer Server Side <a href="#">[SAVE]</a>	Informative People Inc	San Francisco	CA	Full-Time, Employee	Unspecified	1-2 Years Experience	unspecified	12-15-07	***
E-Commerce Director	Director of Software <a href="#">[SAVE]</a>	Terran Systems	San Francisco	CA	Full-Time, Employee	\$140K - \$190K	10-15 Years Experience	unspecified	12-15-07	***
E-Commerce Director	Director of Categorization <a href="#">[SAVE]</a>	Yahoo! Inc	Sunnyvale	CA	Full-Time, Employee	Unspecified	Unspecified	unspecified	12-15-07	***
Information Services Consultant	Financial Services Consultant <a href="#">[SAVE]</a>	IBM	San Francisco	CA	Full-Time, Employee	Unspecified	0-1 Years Experience	unspecified	12-15-07	***
Java Server Engineer	Senior Java Server Engineer <a href="#">[SAVE]</a>	ArcSight	Cupertino	CA	Full-Time, Employee	Unspecified	5-10 Years Experience	Bachelor of Science	12-15-07	***
Information Services	Oracle Service Consultant <a href="#">[SAVE]</a>	IBM	San Francisco	CA	Full-Time, Employee	Unspecified	Unspecified	unspecified	12-15-07	***

Internet | Protected Mode: Off 100%

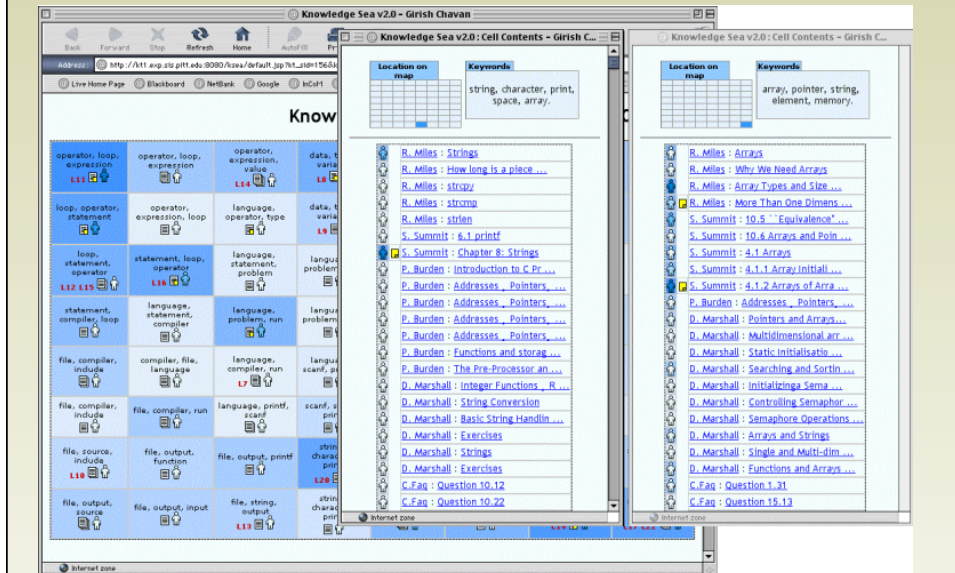
## Community-based OCAH

- Footprint-based social navigation
  - Footprints, CoWeb, Knowledge Sea II, ASSIST
- Action-based social navigation (annotation, scheduling...)
  - Knowledge Sea II, Conference Navigator
- Direct feedback for navigation support
  - CourseAgent, PittCult
- Tag-based social navigation

# CoWeb: Footprint Social NS



# Knowledge Sea II



# Conference Navigator

**E-Learn 2006**  
World Conference on E-Learning in Corporate, Government, Healthcare, & Higher Education

**Considers user visits, scheduling, annotation**

**E-Learn 2006 Presentation Schedule/Search**

Search:

Searches for United States or United Kingdom, use "USA" or "UK".

Date:  All Days  Saturday, Oct. 14, 2006  Monday, Oct. 16, 2006  
 Sunday, Oct. 15, 2006  Tuesday, Oct. 17, 2006

Topics:

**Next Page**  
Presentations Matching Search: 77  
Sat, Oct. 14

Time	Room	Title	Type	Topic	Proceedings Starting Page #
10:00 AM	3	Annotation: A Social Navigation and Annotation Service for Web-based Educational Resources Rosta Farzan, University of Pittsburgh, USA; Peter Bruckovskiy, University of Pittsburgh, USA <a href="#">View Details</a> <a href="#">View PowerPoints Presentation</a>	Full Paper	Tools & Systems	2794
10:30 AM	9	Using iPods to Support Content Area Learning in a Japanese College Lecture Course Chiyoaki Sooki, Waseda University, Japan; Shoji Hirohima, Waseda University, Japan; Shogo Kato, Waseda University, Japan <a href="#">View Details</a>	Full Paper	Tools & Systems	3014
11:15 AM	3	Enabling the collective to assist the individual: CoDEAD, a self-organising reading environment Katherine Charles, MCGill University, Canada; Suzanne Laporte, MCGill University, Canada <a href="#">View Details</a> <a href="#">View PowerPoints Presentation</a>	Brief Paper	Tools & Systems	2753
11:35 AM	3	Drag and Drop Streaming: an Inexpensive Method for Recording and Delivering Lectures is Becoming the Next Revolution in E-Learning Salvador Baez-Franceschi, University of Houston, USA; Alfhaid Dinslaw, University of Houston, USA; Ian Evans, University of Houston, USA; Donald McKeown, University of Houston, USA <a href="#">View Details</a>	Brief Paper	Tools & Systems	2715
11:55 AM	3	Supporting Diagnostic Problem Solving in Medical Education Using an Integrated Classroom - E-Learning Model Sonia Fiermo, MCGill University, Canada; Suzanne Laporte, MCGill University, Canada; Genevieve Cloutier, MCGill University, Canada; Jeffrey Wolman, MCGill University, Canada <a href="#">View Details</a>	Brief Paper	Tools & Systems	2788
1:30 PM	8	Finding Out Chunks of Understanding with Learner's Contexts in C-Language Programming	Full Paper	Tools & Systems	3040

# CourseAgent

Spring 2006 List

**CourseAgent**  
Adaptive Online Course Recommendation System

Control Panel | Schedules | Career Scope | Course Catalog | Faculty | Register

Rosta's CourseAgent Help Log off

**Schedule of spring 2006**

Taken Courses, Planned Courses, Currently Taken Courses, Recommend by Advisor, Degree of Relevance to Career Goals

CRN	Course No	Title	Day	Time	Location	Instructor	Workload	Relevance	Action
2692	ISECOM 2992	PRACTICUM	sept			Richard Thompson			Plan It
16084	INESC1 2120	INFORMATION AND COOKING THEORY	tue	6:00-8:50 P	302 CL	Paul Munro	YY 00	○○○	Plan It
16077	INESC1 2130	DECISION ANALYSIS AND DECISION SUPPORT SYSTEMS	wed	6:00-8:50	411 IS	Marek Coucatal	YY 00	○○○	Plan It
16080	LIS 2128	ETHICS IN THE INFORMATION SOCIETY	mon	3:00-5:50 P	403 IS	Ironi Carbo		○○○	Plan It
16099	INESC1 2130	HUMAN FACTORS IN SYSTEMS	thu	6:00-8:50 P	411 IS	Michael Lewis	YY 00	○○○	Register It
16056	INESC1 2470	INTERACTIVE SYSTEM DESIGN	wed	6:00-8:50 P	405 IS	Peter Rozalozzski	YY 00	○○○	Evaluate It
16079	INESC1 2511	INFORMATION SYSTEM ANALYSIS, DESIGN, AND EVALUATION	tue	6:00-8:50 P	411 IS	Glen Bax	Y 0	○○○	Plan It
16011	INESC1 2610	DATA STRUCTURES	thu	3:00-5:50 P	501 IS	Roger Pharo	YY 00	○○○	Plan It
16118	INESC1 2611	ALGORITHM DESIGN	tue	3:00-5:50 P	406 IS	Hassan Karami	YY 00	○○○	Plan It
16065	INESC1 2720	GEOGRAPHIC INFORMATION SYSTEMS	thu	6:00-8:50 P	405 IS	Hassan Karami	YY 00	○○○	Plan It

**Planned to take (can be registered)**

**Already taken (can be evaluated)**

**Difficulty level of the course**  
Low , Medium , High

**Degree of relevance to students' career goal**  
Marginally relevant   
Relevant   
Very Relevant



## Keyword-based OCAH

- Siskill and Webert
  - Link ordering and annotation
- ML-Tutor
  - Link ordering and generation
- ScentTrails
  - Link annotation
- YourNews/TaskSieve
  - Link ordering and generation

## ML Tutor: Keyword-based link ordering and generation

The screenshot displays the ML Tutor web application. On the left, a 'Suggested Topics' list is visible, containing items like 'Acid Rain', 'Acid Rain Facts', 'Longroad', 'Pollution', 'ScentTrails', 'ScentTrails', 'ScentTrails', and 'ScentTrails'. A label 'ML Tutor's suggestions' points to this list. The main content area features a cartoon character and text discussing acid rain's effects on people. Below the text is a 'Web Site Bookmarks' section with links to 'Environmental Problems', 'Acid Rain 2000', 'Clear Air Action', and 'Cleaner Drinking Water'. The interface is presented in a browser window with a taskbar at the bottom showing 'Start', 'Microsoft Word - Document', 'ML Tutor - Microsoft Internet Explorer', 'ML Tutor', and 'Java Console'.

# ScentTrails: Keyword-based Adaptive Link Annotation

## Departmental and Production Copiers

(60 & up Copies per Minute, Volume above 75,000 Copies per Month)

**5665 Copier:** 60 copies/min. Space efficient design, highlight color, versatile and feature rich with extensive sorter finishing options.

**5065 Copier:** 62 copies/min. Zoom R/E, up to 171"x22" originals & 11"x17" copies, feeder, duplex, other high end features.

**5365 Copier:** 62 copies/min. 100 sheet feeder, zoom R/E, up to 171"x22" originals & 11"x17" copies, duplex, other high end features.

**Document Centre 265 Digital Copier:** 65 copies/min. Scans your originals only once, and then prints as many copies as you need. Duplex, zoom reduce/enlarge.

**5385 Copier:** 80 copies/min. Up to 171"x22" originals & copies, 100 sheet feeder, highlight color, image editing, many features & options.

**5680 Copier:** 80 copies/min. Space efficient design, 100 sheet feeder, auto insertion of covers & transparency sheets, collating, stapling.

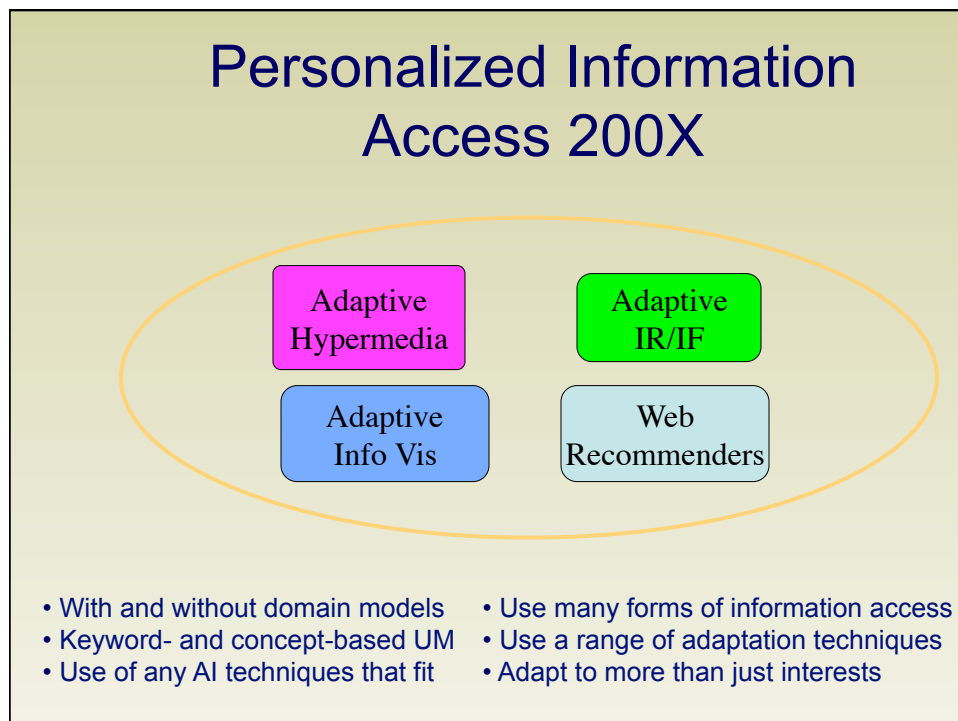
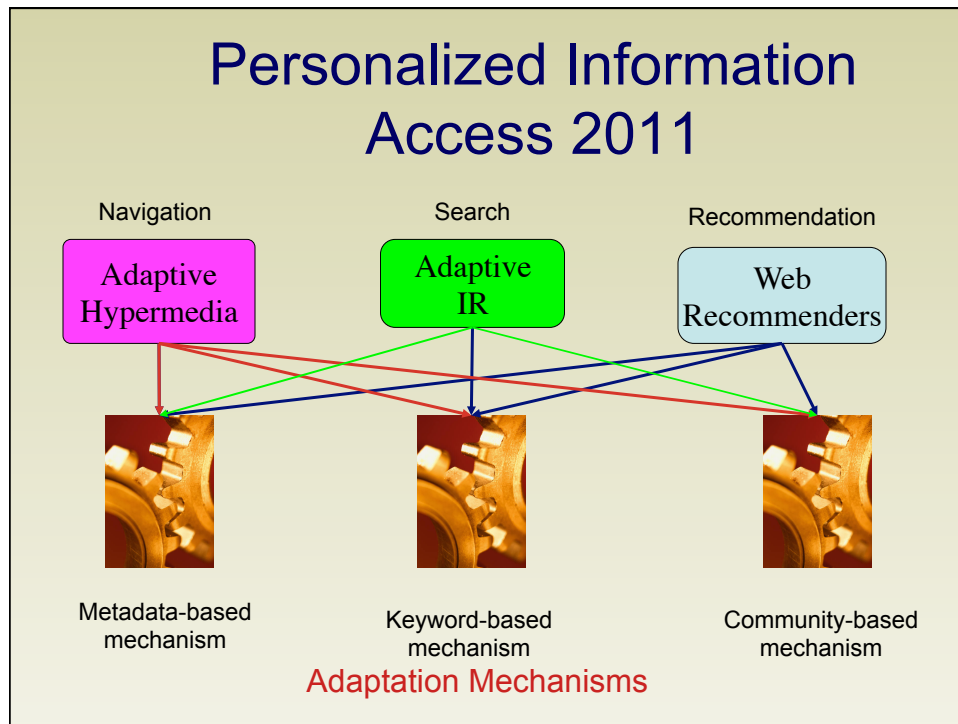
**5388 Copier:** 92 pages/min. Updated and enhanced design of the popular 1090 copier. Wide range of capabilities and capacities.

**5892 Copier:** 92 pages/min. Compact size, photo mode, background suppression, and 100-sheet universal document feeder. Easy-to-use control panel with message display and color graphics.

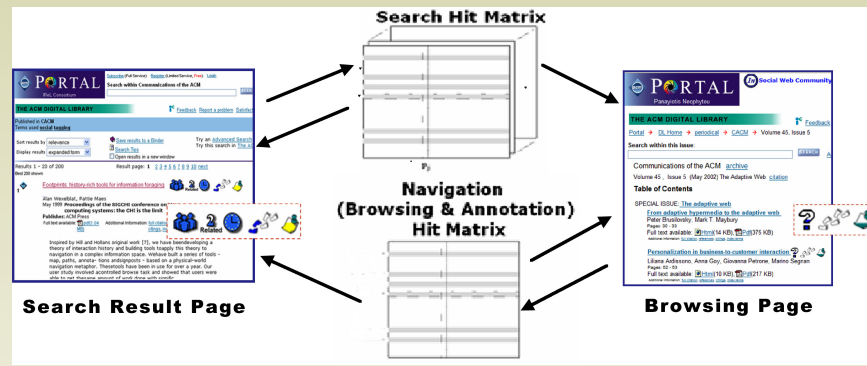
# YourNews: Open Keyword-Level User Models

Keyword-level user model is visible and editable

The screenshot shows a web browser window titled "YourNews - Mozilla Firefox". The address bar shows the URL "http://tr.exp.sis.pitt.edu/gale/news-study/personalize". The page content includes a navigation menu with categories like "All", "Headlines", "National", "World", "Business", "Technology", "Sports", "Entertainment", and "Health". Below the menu, there are several news articles. A sidebar on the right, titled "YourNews", displays a list of keywords for "Health News". The keywords include "COLI SPINACH", "BOY", "IDAHO", "PLANT", "BOISE", "PROBE", "KIDNEY", "INFECT", "DEATH", "BACTERIA", "TODDLER", "NATIONWIDE", "STRAIN", "OFFICIAL", "FAILURE", "OUTBREAK", "CRIMINAL", "FRESH", "PROMPT", "BLAME", "WARNING", "CONSUMER", "LINK", "HEALTH", "LAUNCH", "DIE", "OLD", "MONTH", "THURSDAY", "KILL". Below the keywords, there is a section for "CONTAMINATION" with an input field containing the word "contamination" and an "OK" button.



# ASSIST-ACM



Re-ranking result-list based on search and browsing history information

Augmenting the links based on search and browsing history information

## More Information

- Read
  - Brusilovsky, P. and Henze, N. (2007) Open corpus adaptive educational hypermedia. *The Adaptive Web: Methods and Strategies of Web Personalization*. Lecture Notes in Computer Science, Vol. 4321, Springer-Verlag, pp. 671-696.
- Explore
  - Try our systems at PAWS Community portal: <http://www.sis.pitt.edu/~paws>