

SEARCH AND RETRIEVAL OF
LEARNING OBJECTS

University of
Nebraska
Lincoln

Ziyang Lin November 1, 2010

Overview

- Introduction
- Background
- Design of Learning Object Search Engine
- Future Work

2

Introduction

- What is Learning Object Search Engine
 - A tool that retrieves learning objects (LOs) indexed by metadata from online LO libraries based on users queries.

Introduction Background Design Future

3

Introduction

- What is a Learning Object(LO)?
 - provides reusable online instruction and tutorials (e.g. text, image and flash object). Some LOs contain assessment questions to evaluate users' performance.
 - Usually web-based (can be used in learning management system like Blackboard and Moodle)
 - Contains metadata to describe the structure and the characteristics of LO.

Introduction Background Design Future

Introduction

Advanced Recursion

Tutorial

The inorder traversal begins by looking at the root node. But does not immediately visit it, since because the first step of the traversal involves visiting the left child first. Since B also has a left child, the child gets visited and now we arrive at...

Introduction Background Design Future

Introduction

- What is Digital Library of Learning Object?
 - Online storage of learning objects with tools for search and retrieval.
 - Or indexing of the LO from other storage with tool for search and retrieval.

Introduction Background Design Future

Introduction

7

□ Search Tool in Moodle.org

Found records: 8/61 (Reset filters)

Name	Type
SIMS Word Plot - Flash sample	SCORM 2004
SIMS Rabbit Takeaway - Flash sample	SCORM 2004
SIMS Place Value Darts - Flash sample	SCORM 2004
SIMS Hidden Craft - Flash sample	SCORM 2004
SIMS Fruit Picker - Flash sample	SCORM 2004
SIMS Fraction Monkeys - Flash sample	SCORM 2004
SIMS Division Factory - Flash sample	SCORM 2004
SIMS Bracket Basics - Flash sample	SCORM 2004

Entries per page: 10 Search: math Sort by: Time added Descending Advanced search Save settings

Introduction Background Design Future

Introduction

8

□ Search Tool in the National Science Digital Library

NSDL
National Science Digital Library

General Search K-12 Higher Ed & Research Specialized Search

math

Found records: 7,243,913 (30,485) (Reset filters) Results Per Page

WebDescription

Math Practice
How far is it from the center of the Earth to the Sun? Use this interactive Math Practice to explore the number concept 'Number'.

Math
This is an interactive application to help students learn about the number concept 'Number' through a series of activities. It includes a number line, a number grid, and a number line.

Resources to use for 4th Grade Math
This resource includes a variety of activities and materials for 4th grade math. It includes a number line, a number grid, and a number line.

Introduction Background Design Future

Background

9

□ Searching LO and web page

- Common
 - Indexing, user query, scoring, ranking ...
- Difference
 - Web search is primarily text-occurrence-based. Reviewing the result can be fast.
 - LO search involves much range retrieval and domain knowledge from education which can explain the meaning of queries. Reviewing LOs take much time and effort (may need to download the LOs and deploy them in Learning Management System).
 - Others: scale, negative educational impact of using improper LOs ...
- General search tool is not suitable for LOs.

Introduction Background Design Future

Background

10

Common	Different
Indexing	Web search is primarily text-occurrence-based. Reviewing the result can be fast.
Query	LO search involves much range retrieval and domain knowledge from education which can explain the meaning of queries. Reviewing LOs take much time and effort.
Ranking	Others: scale, negative educational impact of using improper LOs ...

Introduction Background Design Future

Background

11

□ LO Metadata and Its Extension

- Metadata is an xml file with tags (e.g. tags of SCORM) that describe the structure (e.g. what page and image are contained?) and characteristics (e.g. author, difficulty level) of LO)
- New tags can be added as extensions to describe customized attributes (e.g. average score of the users who finish the assessment questions).

Introduction Background Design Future

Background

12

□ Search Form and Metadata Extension

- Many LO search engines only provide simple interface to users. Users can not specify values for different tags in the metadata in details.
- Some LO search engines have complicated interface designed for the variety of metadata and extensions but the LOs provided by third parties do not have the preferred metadata.

Introduction Background Design Future

Background

13

- User Preference Modeling and Ranking
 - User may prefer some LOs to other LOs in the result.
 - Preference may not be completely specified in the query.
 - Clicking behaviors (e.g. time of click, position of click) contains important information which can be used to model the preference.
 - Preferred LOs should be ranked higher in the result.

Introduction Background Design Future

Background

14

- User Preference Modeling and Ranking
 - "Crucial" technology which can improve the result greatly.
 - Widely used in major search engines for web page.
 - Few LO search engine uses it except NSDL.

Introduction Background Design Future

Background

15

- User Preference Modeling and Ranking
 - A Dynamic Bayesian Network is proposed (Chapelle, 2009) to predict the click through rates of URLs based on users' clicking history in different positions on the result page.
 - It introduces 3 hidden variables to describe whether a user examines, is attracted by or is satisfied by the URL approximated by using EM.
 - click through rates can be used as a parameter of other ranking function such as RankNet (next slide).

Introduction Background Design Future

Background

16

- User Preference Modeling and Ranking
 - RankNet (Burges, et al, 2005) will give a rank for each result in the result set using data of clicking behavior.
 - It is basically a neural network trained by pair-wised ranked examples with probability of the paired examples.
 - It ranks the results in the result set directly.

Introduction Background Design Future

Design of Learning Object Search Engine

17

- Two problems to be solved:
 - Third party LOs may not contain desired metadata.
 - Different standard metadata
 - Same standard metadata with different extensions
 - It costs much time and effort for the user to review the results. Reviewing LOs may include downloading the LOs or uploading to user's own learning management system.

Introduction Background Design Future

Design of Learning Object Search Engine

18

- How to populate metadata if the metadata of the LO is not designed in the desired way?
 - LO hosting
 - LOs developed by third parties can be hosted by the search engine.
 - LOs can be uploaded to the system and deployed to users for a period
 - Collect the usage data, analyze the data and add the result to metadata.

Introduction Background Design Future

Design of Learning Object Search Engine

19

- Two search forms for the metadata collected from LO hosting
 - Provide a general search form if users want to handle the extensions as text-occurrence-based search.
 - Provide an advanced search form and allow users to specify the requirement for each tag in the metadata in the query.

Introduction Background **Design** Future

Design of Learning Object Search Engine

20

- Retrieval of LO
 - Given the query, the search engine will go over the metadata of all the LOs to see whether the requirements of the query are met.
 - All the LOs in the result will be graded and ranked based on the fitness and user's implicit preference.
 - Fitness of a LO measures how well a LO matches user's requirements in the query.
 - Grading – computing the fitness for each LO.
 - User's implicit preference – derived from the clicking and browsing behaviors when a user is reviewing the results.
 - Ranking considers the fitness, the preference from the same user's previous search and the preference from other similar users.

Introduction Background **Design** Future

Design of Learning Object Search Engine

21

- How to improve the result and reduce the negative impact of reviewing LOs?
 - User Preference Modeling
 - When a user clicks and reviews a LO in the result, this interaction will be tracked for modeling the preference in the future.
 - Users are encouraged to rate the LOs as explicit feedback of the review of the result.
 - When the user comes back and searches again, the search engine will be able to predict what this user really wants besides the requirements in the query.

Introduction Background **Design** Future

Design of Learning Object Search Engine

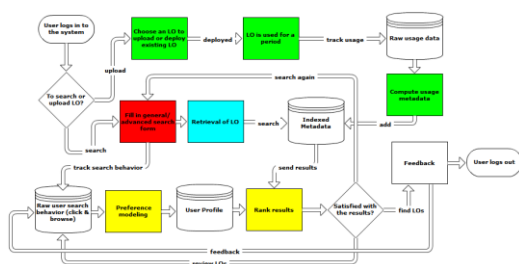
22

- The entire system contains 4 major components.
 - Search Form
 - General search form handles text-occurrence-based search.
 - Advanced search form handles the metadata with extensions.
 - Learning Object Hosting
 - Retrieval of LO
 - User Preference Modeling

Introduction Background **Design** Future

Design of Learning Object Search Engine

23



Introduction Background **Design** Future

Future Work

24

- Compare different algorithms of user preference modeling.
- Design a search form that handles arbitrary metadata coming along with the LOs when they are indexed.

Introduction Background **Design** Future

References

25

- Chapelle, O., Zhang, Y. (2009) A Dynamic Bayesian Network Click Model for Web Search Ranking. World Wide Web Conference 2009 Madrid, Spain.
- Burges, C., Shaked, T., Crenshaw, E., Lazier, A., Deeds, M., Hamilton, N., Hullender, G., (2005). Learning to Rank using Gradient Descent. In Proceedings of the 22nd International Conference on Machine Learning.
- S. Riley, L.D. Miller, L.-K. Soh, A.Samal, and G. Nugent (2009). Intelligent Learning Object Guide (iLOG): A Framework for Automatic Empirically-Based Metadata Generation, Proceedings of the International Conference on Artificial Intelligence in Education (AIED2009), Brighton, UK, July 6-10. pp. 515-522.
- Lagoze, C., Hoehn, W., Millman, D., Arms, W., Gan, S., Hillmann, D., Ingram, C., Krafft, D., Marisa, R., Phipps, J., Saylor, J., Terizzi, C., Allan, J., Guzman-Lara, S., Kati, T. (2002) Core services in the architecture of the national digital library for science education (NSDL). In: Proceedings of the second ACM/IEEE-CS joint conference on Digital libraries, Portland, Oregon, USA, pp. 201-209

Thank you

26

- Question?