

**Title of Session:** Search and Retrieval Tools

**Moderator:** Steve Schatz

**Title of File:** 20040212sandr

**Date:** February 12, 2004

Room: After School Online

**BJ:** let's start with introductions, Steve...

**BJ:** in case we have any latecomers

**SteveS:** OK..

**SteveS:** I'll start .. I'm Steve Schatz.. among other things, I'm working on my PHD in instructional systems at Indiana University

**BJ:** Welcome to the New Search and Retrieval Tools discussion. Our discussion leader is Steve Schatz

**BJ:** I'm an art teacher in Pennsylvania and a helpdesk volunteer for Tapped In

**IbrahimK:** I am pre-service teacher enrolled in professional urban multiculture action

**SteveS:** interesting

**BJ:** very interesting, Ibrahim!

**SteveS:** ok.. so.. I'll talk a little bit about my work and new methods of tagging

**SteveS:** search/retrieval

**SteveS:** as we are a few folk.. feel free to ask questions at any point.

**IbrahimK:** Thank you and I hope we learn from each other

**SteveS:** it is semi technical.. so if I get lost in jargon.. stop me and ask

**SteveS:** Ok..

**BJ:** will do, Steve.

**SteveS:** so.. I came to playing with this from an interest in learning objects

**SteveS:** I don't know if you have heard of them.. it's more a deal in training than in P-12 ed

**LynneW** joined the room.

**LynneW:** Hello

**SteveS:** but the idea (which has some fundamental flaws) is

**SteveS:** hello Lynne

**SteveS:** I'm just getting started on background stuff

**LynneW:** Ok

**SteveS:** OK.. so learning objects..

**SteveS:** the idea is .. if you have a lot of web based pieces of instruction

**SteveS:** wouldn't it be swell if those could be tagged and retrieved

**SteveS:** so that a teacher in one country.. might be able to find and use a piece made by a teacher in another country

**BJ:** can you give us an example, Steve?

**SteveS:** now. there are some theoretical problems.. like can learning be objectified.. flies in the face of constructivism

**SteveS:** sure.. an object can be anything.. but a good idea is an animation of the water cycle

**SteveS:** say you have taken all the time to build a very cool little animation..

**SteveS:** you could post it on the web.. and others can find it and use it

**BJ** nods

**SteveS:** so.. what I've realized over time

**SteveS:** is that the important thing about learning objects

**SteveS:** is that they have a standard method of tagging the objects

**SteveS:** so it is easier to find them

**SteveS:** ok.. now.. bear with me here

**SteveS:** there are a few tech things you have to get

**SteveS:** ok.. if you have a web page

**BJ** listens carefully

**SteveS:** you can put tags.. META tags

**SteveS:** in the header

**SteveS:** this was an idea from early in the web

**SteveS:** that would help with search and retrieval

**SteveS:** if we could put information for finding web pages in the header

**SteveS:** then searching would be easier

**SteveS:** it's like a card catalog at a library

**IbrahimK:** good idea Steve

**SteveS:** that has meta data.. data about the books

**SteveS:** well.. it IS a good idea.. but the problem was.. that there is no real standardization on the web

**SteveS:** if you look at a card catalog.. it is going to have a set amount of information

**SteveS:** and the labeling.. for finding the book.. will either be Dewey decimal.. or MARC.. which is library of congress

**SteveS:** in the world o the web.. there is NO standardization

**SteveS:** and marketing folks took to putting information in meta tags that would drive search engines to their sites

**SteveS:** so.. that idea never really firmed up

**SteveS:** that make sense so far?

**LynneW:** absolutely

**BJ** nods.

**SteveS:** good

**SteveS:** ok.. now a few years ago

**SteveS:** some folks started thinking about having a standard set of tags

**SteveS:** specifically for learning

**SteveS:** ok.. here's another conceptual leap

**SteveS:** get ready

**SteveS:** there are LOTS of different tagging sets.. or schemas

**SteveS:** they go from very simple (the most simple is Dublin core.. with 17 elements)

**SteveS:** to very complex

**SteveS:** whenever you go to a search system.. like lexis, or ERIC...

**SteveS:** they are using a tagging schema

**SteveS:** ok. so groups set up a tagging schema for learning

**SteveS:** OK

**SteveS:** now that brings us to a very interesting place

**SteveS:** IF we can think about a standardized schema for learning

**BJ . o O ( you have to get educators to agree on something?! )**

**SteveS:** oh.. much harder than that - grin

**BJ** chuckles

**SteveS:** it is educators, and software makers, and content makers

**SteveS:** and you hit on a major problem

**SteveS:** in order to be all things to as many people as possible.. for educators.. and industrial trainers.. and simulation makers.. and things like blackboard and webct

**SteveS:** the standards groups had to make some pretty major compromises

**SteveS:** trying to please all folks

**SteveS:** now.. there is a problem with having general tags

**SteveS:** imagine a tag - author

**SteveS:** ok.. now image you have a web site

**SteveS:** and on that web site is a photo of a statue in the Getty museum

**SteveS:** who is the author?

**SteveS:** the web site builder?

**SteveS:** the web site owner?

**SteveS:** the sculptor?

**SteveS:** the photographer?

**SteveS:** the Getty?

**BJ:** yikes...that is a problem!

**SteveS:** and even if you can decide on an answer

**SteveS:** how do you know that you will have uniformity?

**SteveS:** I was in a class where we tried this..

**SteveS:** and almost NEVER agreed

**SteveS:** so.. given THAT as a problem

**SteveS:** I started thinking..

**SteveS:** what about having a schema focused on P-12 educators

**SteveS:** and develop that schema

**SteveS:** not trying to describe objects..

**SteveS:** but by talking to teachers

**SteveS:** and asking what problems they would turn to the web to solve

**SteveS:** what kinds of problems..

**SteveS:** and what kinds of information would they need

**SteveS:** if one uses a PROBLEM focus

**SteveS:** instead of an object focus

**SteveS:** a very interesting thing happens

**SteveS:** you only need a very small tag set to describe the information

**SteveS:** and you have a very different kind of searching

**SteveS:** ok.. you have all been very patient

**SteveS:** now.. let me show you some examples

**SteveS:** <http://www.powerstart.com/cgi-bin/tagger/tag.pl>

**SteveS:** it will probably pop up with a username and password

**SteveS:** use e139

**SteveS:** for both.. lower case

**BJ:** I got right in, Steve

**SteveS:** what this is .. is a database of over 500 objects that have been tagged

**SteveS:** is everybody in?

**LynneW:** Yes

**SteveS:** ok.. now go to 503..

**SteveS:** and click on edit

**SteveS:** that shows the tags

**SteveS:** for that particular object

**SteveS:** if you click on the link of 503.. you will be able to see the object

**SteveS:** if you notice.. the tagging set is actually pretty simple

**SteveS:** it breaks down to this.. location.. title and description

**SteveS:** then the other things have controlled vocabulary

**SteveS:** so that everyone will call a rose a rose

**SteveS:** so.. once you have a bunch of these objects

**SteveS:** 2 important things happen

**SteveS:** 1 is that it is fast to search (I'll show you that in a minute)

**SteveS:** and the other.. which I think is more significant

**SteveS:** is that we can have a site

**SteveS:** where we can have teachers adding objects

**SteveS:** and they can tag them for easy search and retrieval

**BJ** smiles...very cool, Steve

**SteveS:** without taking a lot of time

**SteveS:** thanks

**SteveS:** ok.. so close those windows

**BJ:** done

**SteveS:** and go here [http://www.powerstart.com/cgi-bin/survey/search.pl?which\\_search=tag](http://www.powerstart.com/cgi-bin/survey/search.pl?which_search=tag)

**SteveS:** if you need to sign in.. again use e139

**SteveS:** for both

**SteveS:** now.. this is a hard jump for folks

**SteveS:** it is a different kind of searching

**SteveS:** if you are searching using a standard web search engine

**SteveS:** this is how it works

**SteveS:** the engine goes out and reads web pages

**SteveS:** and it creates an index

**SteveS:** if the page has moon on it

**SteveS:** then the search engine will go to its index

**SteveS:** under moon.. and say.. ok.. this page has moon

**SteveS:** so.. when you type in moon..

**SteveS:** it goes to its index

**SteveS:** and returns all the references in its index

**SteveS:** it isn't searching the web..it's searching its index

**SteveS:** now..

**SteveS:** the problem is.. it will return moon.. and moon pie.. and I moon for you

**SteveS:** and reverand moon

**SteveS:** and EVERYTHING that has moon

**SteveS:** that's why we get millions of hits

**BJ** nods

**SteveS:** NOW.. this is important

**SteveS:** if you don't type in anything to a text search  
**SteveS:** you will not get anything back  
**SteveS:** right?  
**SteveS:** but a tag search is different  
**BJ:** right  
**SteveS:** a tag search is NOT searching the document  
**SteveS:** it has NOT read the document  
**SteveS:** it is looking at the TAGS  
**SteveS:** and when you select tags.. it is focusing the search  
**SteveS:** if you don't enter anything in a tag search.. you get EVERYTHING back  
**SteveS:** so.. you use very few things to trim down the search  
**SteveS:** let's give it a try on the tag search window  
**SteveS:** try clicking teacher  
**SteveS:** in what role  
**SteveS:** and video in format  
**SteveS:** and nothing else  
**SteveS:** you should get quite a few things  
**SteveS:** go to the bottom of the page.. and click search again  
**SteveS:** and this time.. click some more.. to try to cut down the number.. maybe..  
teacher.. and Activity under what kind of information.. and Video.. and Using technology  
**SteveS:** that returns nothing  
**SteveS:** try some other combinations  
**SteveS:** what you will see is that you get a very powerful way of cutting down.. and  
getting exactly what you want.. IF THE DOCUMENT EXISTS  
**SteveS:** is this making sense?  
**SteveS:** it's pretty tricky  
**SteveS:** if I am searching for moon..  
**BJ:** yes, indeed. It is tricky, but part of it is that you are familiar with your page and we're  
still trying to find all the categories to click  
**SteveS:** a lot of times.. I will start with the text box  
**SteveS:** and type in moon..and see what comes up  
**SteveS:** then I will start to cut it back..  
**SteveS:** just a tag at a time  
**SteveS:** ok.. while you are playing with that..  
**SteveS:** I will talk about the final piece that I'm working on  
**SteveS:** and then open it up to questions  
**SteveS:** have you ever wondered why there weren't many search engines just for  
teachers.. or for specific groups?  
**SteveS:** the reason is that there is no real way now to evaluate if search engines work  
**SteveS:** there is a field.. called information retrieval evaluation  
**SteveS:** and it is the way that all search engines are tested  
**SteveS:** when someone says that they like google more than yahoo.. that is evaluation  
**SteveS:** the theory behind evaluation..however is very thin  
**SteveS:** it has a very faulty assumption.. a main assumption is that a GOOD system  
**IbrahimK:** sorry guys I have to attend a scheduled course meet you in the next  
discussion and I am sorry to miss the conclusion Bye

**SteveS:** thanks for coming  
**SteveS:** if you have any questions.. lemme know  
**SteveS:** [schatz@powerstart.com](mailto:schatz@powerstart.com)  
**SteveS:** ok.. so..a GOOD system.. is one that returns as many as possible documents  
**SteveS:** that are relevant to your question  
**SteveS:** in the early days of the web.. that might have been good  
**SteveS:** but now.. with millions being returned.. that assumption is pretty bad  
**SteveS:** the way that systems are evaluated  
**SteveS:** is with 2 measures..  
**SteveS:** precision and recall  
**SteveS:** the way they are evaluated is this.. you get a set of documents  
**SteveS:** and a set of queries  
**SteveS:** and a set of judgments.. which documents  
**SteveS:** are RELEVANT  
**SteveS:** to which query  
**SteveS:** then you use the system  
**SteveS:** and see how it does  
**SteveS:** the problem with that is..  
**SteveS:** WHAT is relevant?  
**SteveS:** who decides?  
**SteveS:** there has been a lot of research.. and it has shown that having relevance based measures is really pretty bad  
**SteveS:** but they are still the ones used  
**SteveS:** the reason.. well.. why do we still have standardized tests?  
**SteveS:** they give a number  
**SteveS:** even if it is a flawed number  
**BJ** nods sadly  
**SteveS:** so.. my latest work is looking at other ways of evaluating systems  
**SteveS:** using measures that are NOT based on relevance  
**SteveS:** and that.. is 5 years of conceptual jumps in 50 minutes  
**SteveS:** so.. I'm going to throw it open to questions  
**BJ** applauds. Fascinating, Steve!  
**SteveS:** thankee  
**LynneW:** You broke this information down very well. Thank you for clarity.  
**SteveS:** I will be making a site based on tagging  
**BJ:** you are assuming that teachers will voluntarily enter tags on the database?  
**SteveS:** and letting teachers add their own objects  
**SteveS:** yes  
**SteveS:** it is a question if they will  
**LynneW:** Steve, this might be a great learning unit for pre-service teachers.  
**SteveS:** but I think it is a possibility.. since it is a simple set  
**SteveS:** I'd love to show it to pre service teachers  
**SteveS:** the idea is that if it is fast and easy (it takes me about 30 seconds to tag an object  
**LynneW:** I would like to invite you to TI2 as a guest speaker for my classes.  
**SteveS:** then teachers will add to the set of documents  
**SteveS:** and we can teach each other.. providing cool objects

**SteveS:** sure..how would that work?

**LynneW:** I am just starting my information literacy unit and did some experiential stuff with search engines yesterday. This is all very timely

**SteveS:** is TI2 another discussion group on Tapped In?

**SteveS:** very cool

**LynneW:** Sorry, TI2 is Tapped In new and improved

**SteveS:** I will be launching my site.. teachershelpdesk.com within the next few months

**SteveS:** very cool

**BJ** smiles. We're in TI2 now, Steve

**SteveS:** I would be happy to.. just give me a time and place

**SteveS:** oh..

**LynneW:** I just need to know what type of information they should have to scaffold them in preparation for your interaction and demonstration with them.

**LynneW:** I teach undergrads this semester and MATs in the summer.

**BJ:** Steve, why don't you give Lynne your email and you can hash out the details?

**SteveS:** well.. they can come at it pretty cold

**SteveS:** great idea

**SteveS:** [schatz@powerstart.com](mailto:schatz@powerstart.com)

**SteveS:** where physically are you Lynne?

**LynneW:** I have an office here and you can reach me at [lwolters@cu-portland.edu](mailto:lwolters@cu-portland.edu)

Lynne Wolters

**LynneW:** Portland, OR

**BJ:** I've got to run to another f2f meeting. Thanks, Steve, for doing this...excellent job

**LynneW:** Thank you BJ

**SteveS:** thanks BJ

**BJ** . o O ( you both can stay here and talk...nothing scheduled for a while )

**BJ** waves bye

**LynneW:** Cool