



CNLP

Center for Natural Language Processing



MetaTest

**Automatic Generation Of Metadata And Preliminary
Evaluation Of Its Utility In Information Seeking**

**Liz Liddy
Tina Finneran**

**Center for Natural Language Processing
School of Information Studies
Syracuse University**



MetaTest Overview

CNLP

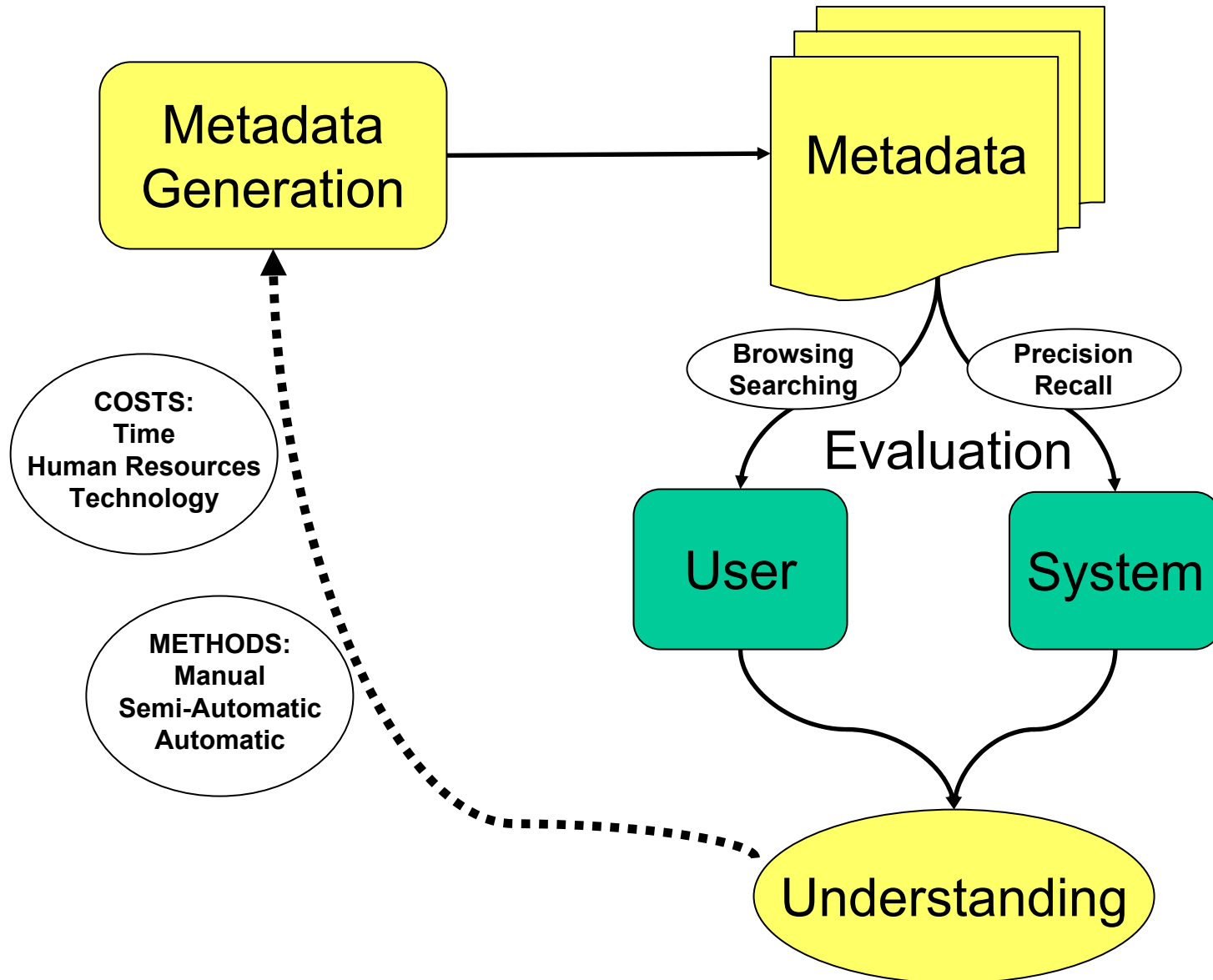
Center for Natural Language Processing



Testing of Assumptions

- Do we need metadata?
 - Why?
- How much metadata do we need?
 - For what purposes?
- Which elements do we need?
 - For which digital library tasks?
- How do information-seekers utilize the metadata?
 - When browsing / searching / previewing?
- Can automatically generated metadata perform as well as manually assigned metadata?

GOAL: Measure Quality & Usefulness of Metadata



Evaluation Methodology

- ✓ Automatically metatag a Digital Library collection that has already been manually meta-tagged
- ✓ Solicit range of appropriate Digital Library users
- ✓ Have users qualitatively evaluate metadata tags
 - Conduct searching & browsing user studies with eye-tracking protocols
 - Conduct information retrieval experiments
- Develop metrics of relative utility of each metadata element (manual & automatic) for searching and browsing

- ✓ Completed
- Underway
- Future

Desired Achievements

- Provide experimental results to guide Digital Library development
- Develop metrics of metadata quality & utility
- Inform HCI design
- Reduce the metadata generation bottleneck
 - Refine metadata standards to include only useful elements
 - Determine if automatic metadata generation can perform comparably to manually generated metadata

Target Metadata Schema

Dublin Core Metadata Elements

- Contributor
- Coverage
- Creator
- Date
- Description
- Format
- Identifier
- Language
- Publisher
- Relation
- Rights
- Source
- Subject
- Title
- Type

GEM Metadata Elements

- Audience
- Cataloging
- Duration
- Essential Resources
- Grade
- Pedagogy
- Quality
- Standards



Qualitative Study

Comparing the Quality of Manual & Automatic Metadata

CNLP

Center for Natural Language Processing



Evaluation Instructions - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Back Forward Stop Home Search Favorites Media Refresh Print Mail New Tab

Address http://katzer.syr.edu/~tina/lesson-plan-bkup/instructions.php Go Links

Instructions for Evaluating Summaries

On the next page, you will see a Web page similar to the one below.

The screenshot shows a web browser window with a page titled "Questionnaire". The page is divided into three main sections. At the top is the title "Questionnaire" with a small blue link to the right. Below this, the page is split into two vertical columns. The left column is labeled "Lesson Plan" and the right column is labeled "Summary of Lesson Plan". The browser's address bar shows the URL: http://katzer.syr.edu/~tina/lesson-plan-bkup/instructions.php.

You will be given 7 lesson plans, each with its own Summary. Expect that the parts of the Summary might change slightly with each Lesson Plan.

Please evaluate how well the Summary reflects the Lesson Plan, regardless of how good the Lesson Plan itself is. In other words, imagine that a search engine returns the Summary on the right -- how well does it represent the Lesson Plan on the left. Below are step-by-step instructions:

1. Read the question in the Questionnaire
2. Skim the Lesson Plan in the left bottom frame
3. Compare the Summary on the right to the Lesson Plan on the left to answer each question
4. The final question for each Summary will give you space to write comments that you think are important but are not

Done Internet

Lesson Plan Survey - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Back Forward Stop Home Search Favorites Media Print Mail New Tab Person

Address <http://katzer.syr.edu/~tina/lesson-plan-bkup/evalFrameset.php> Go Links


QUESTIONNAIRE (remember to click the **Submit** button)

[Return to Instructions](#)

How well does the **TITLE** on the right match the title of the lesson plan on the left?

Very Poorly Poorly Well Very Well Unsure


If Unsure/Poorly/Very Poorly: What is wrong with the **TITLE** on the right?



Lesson Plan

Home Next

Wildflowers



Grade Level:

Fifth or sixth grade

LESSON PLAN SUMMARY (4 of 7)

Title: Lesson Plan

Description: Given the proper instructions, the students will be able to dry/preserve wildflowers with 100% participation. The teacher will use the overhead projector to project the instructions for the students to view while the teacher goes over the steps with the students. Next, the teacher will demonstrate the procedure as he/she gives the directions that are projected. The students will then prepare to press their flowers ...

Grade: Fifth or sixth grade;

Keywords: flower, wildflower, place, fold, newspaper, heavy book, telephone_book, step, construction_paper, top, dry

Materials: 2 pages of newspapers for each student; 2 or more wildflowers for each student; 1 heavy book per student; 1 telephone book for each student; overhead projector; prepared transparency;

Pedagogical view-write

Internet

How well does the **GRADE** on the right match the grade level of the lesson plan?

- Very Poorly Poorly Well Very Well Unsure

If Unsure/Poorly/Very Poorly: What is wrong with the **GRADE** on the right?

How well does the **SUBJECT** on the right represent the content of the lesson plan?

- Very Poorly Poorly Well Very Well Unsure

If Unsure/Poorly/Very Poorly: What is wrong with the **SUBJECT** on the right?

How well do the **KEYWORDS** on the right represent the content of the lesson plan?

- Very Poorly Poorly Well Very Well Unsure

If Unsure/Poorly/Very Poorly: What is wrong with the **KEYWORDS** on the right?

How well does the **DURATION** on the right represent the content of the lesson plan?

- Very Poorly Poorly Well Very Well Unsure

If Unsure/Poorly/Very Poorly: What is wrong with the **DURATION** on the right?

Any final comments about the summary on the right?

By submitting my responses I know that I am agreeing to participate in this research project. My participation is voluntary and any personal identifying information will be held in confidence. For questions, contact [Tina Finneran](#) or call 315-443-5484.

Who Were the Respondents?

Type of Educator	
Elementary Teacher	6%
Middle School Teacher	6%
High School Teacher	66%
Higher Education Teacher	6%
Instructional Designer	3%
School Media	3%
Other	11%

Subject Taught	
Science	69%
Math	6%
Engineering	3%
Combination	11%
Other	11%

Experience with Lesson Plans	
<1 Year	6%
1-3 Years	29%
3-9 Years	29%
10+ Years	37%

Statistical Analysis

- Ordinal data used to measure metadata quality
 - Unsure, Very Poorly, Poorly, Well, Very Well
- Used the Mann-Whitney Test on Independent Pairs (Non-parametric test)
 - Accepts Ordinal data
 - Does not require normal distribution, homogeneity of variance, or same sample size
 - Ranks the scores from each group and

Medians of Metadata Element Quality

Median Score
Inter-Quartile Range
 Mean Rank

	Title*	Description	Grade	Keyword*	Duration	Material	Pedagogy Method	Pedagogy Process	Pedagogy Assessment	Pedagogy Group
Manual Quality	3 2-4 132	3 3-4 122	3 2-4 73	3 3-4 127	3 2.75-4 29	3.5 3-4 49	3 0.5-3 30	--	--	3 1.5-3 14
Automatic Quality	3 1-4 105	3 2-4 113	3 3-4 80	3 2-4 99	3 2-3.25 25	3 2-4 39	3 2-4 33	3 1-3 53	3 2-3.5 9	3 2.5-4 18

* Indicates statistically significantly difference in the medians

Of the metadata elements above which had data to compare, we conducted a Mann-Whitney U-Test to test if the quality scores for each metadata element were different between the manually and automatically generated methods. The only statistically significant findings were:

- The manually generated Title element was shown to have significantly higher scores than those that were automatically generated (U = 5238, p < .001).
- The manually generated Keyword element was shown to have significantly higher scores than those that were automatically generated (U = 4756, p < .001).

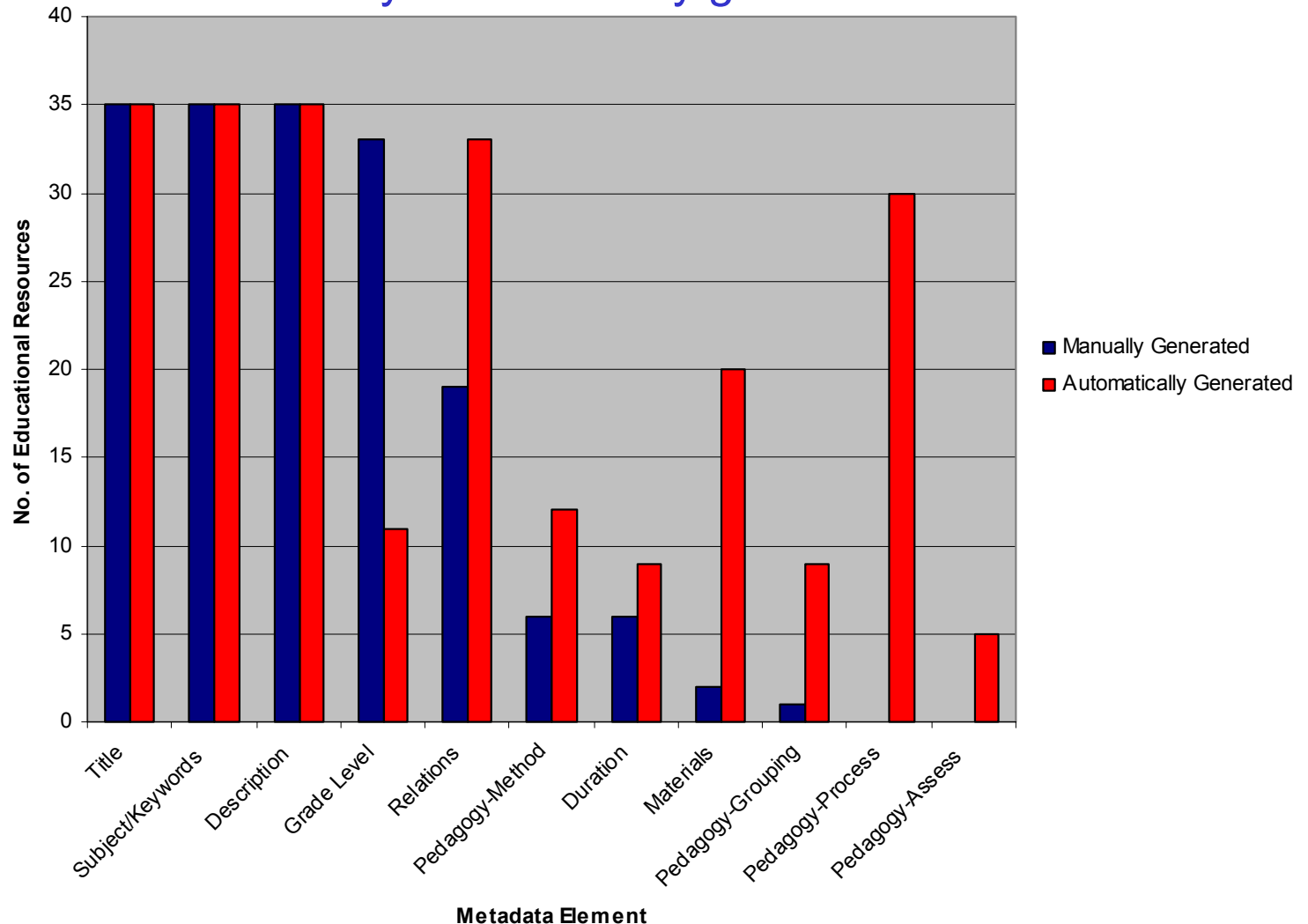
Teachers' Judgments on Title

Some respondents appeared to be judging the actual title, not automatic extraction of the title

- “Title should indicate it is a student/ home survey based lesson”
 - Scored Poorly on Automatic
- “It is actually a good match, but it would be much better if it also had the unit title”
 - Scored Poorly on Automatic
 - Scored Well/Very Well on Manual which was enhanced Title
- “It matches, but doesn’t address what the lesson is about”
 - Scored Poorly on Automatic
 - Scored Well/Very Well on Manual which was enhanced Title

Comparison of Elements with Values

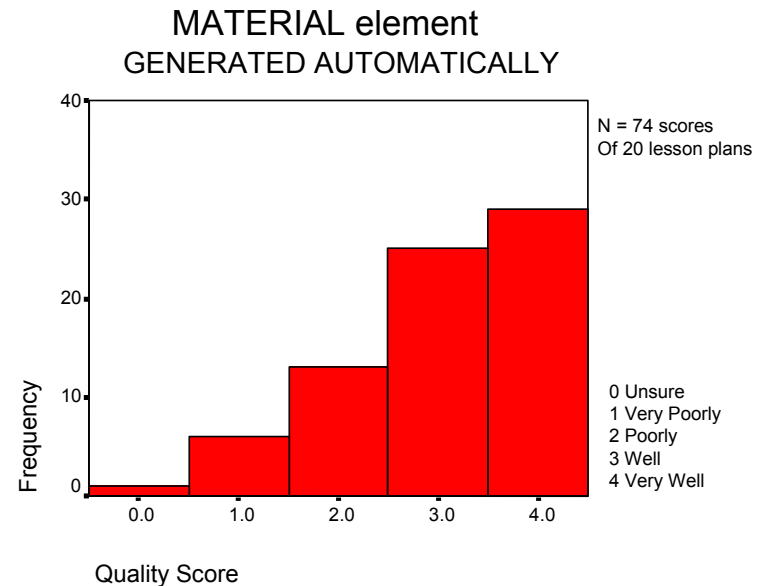
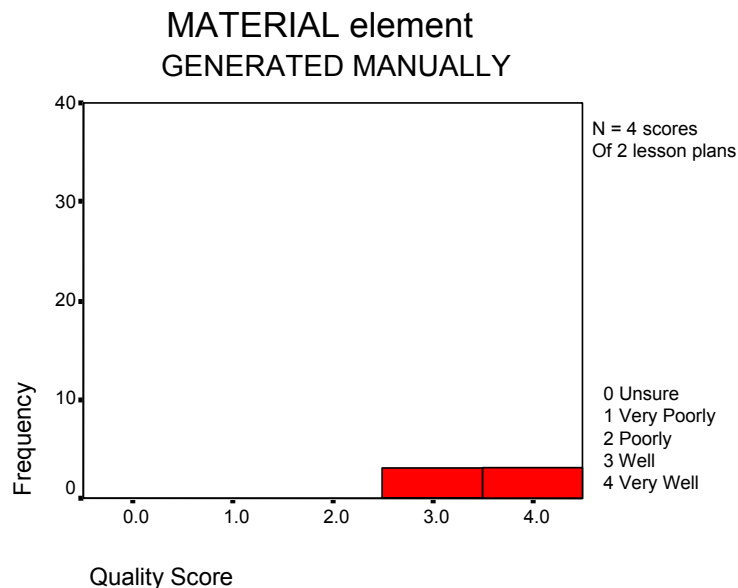
Of the 35 systematically selected lesson plans and activities from the GEM Gateway the following graph shows the presence of the metadata elements for automatically and manually generated metadata.



Quality v. Quantity

Consider quality versus amount of effort to generate manually

- Manually generated values for the Materials element were always rated well or very well, however only 2 of the 35 lesson plans had values for this element.
- Of the 20 lesson plans which had automatically generated Material elements, their scores were more dispersed.





IR Experiment

Comparing Information Retrieval with
Manual Metadata
Automatic Metadata
Free-text of Document

CNLP

Center for Natural Language Processing



IR Experiment Design

We will run three distinct trials as well as combinations of Trials I & II with Trial III.

Trial I: Fielded retrieval on metadata (automatic & manual)

Automatic
Metadata

Title: Black Hole Lesson
Grade: 9-11
Keywords: galaxy, black_hole, set, model, object, dark matter

Manual
Metadata

Title: GalaxSee: Black Hole Lesson
Grade: 9, 10, 11, 12
Keywords: Interactive teaching, simulations, Universe, GalaxSee software

Trial II: Free-text retrieval on metadata (automatic & manual)

Black Hole Lesson, Grade: 9-11, galaxy, black_hole, set, model, object, dark matter, ...

GalaxSee: Black Hole Lesson, Grade: 9, 10, 11, 12, Interactive teaching, simulations, Universe, GalaxSee software, ...

Trial III: Free-text retrieval on lesson plan





User Studies

Pilot Study with Eye-Tracking

CNLP

Center for Natural Language Processing



What the Eyes Can Tell Us

- Indices of ocular behavior are used to infer cognitive processing, e.g.,
 - Attention
 - Decision making
 - Knowledge organization
 - Encoding and access
- For example, the longer an individual fixates on an area, the more difficult or complex that information is to process. Similarly, the first few fixations indicate areas of particular importance or informativeness.



User Study: Data Collection

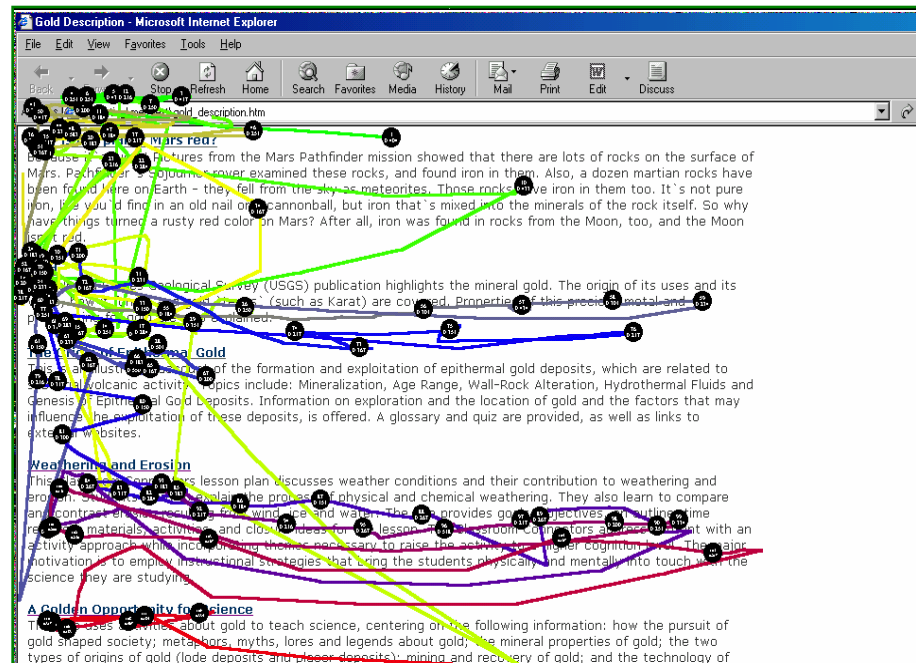


User wears an eye-tracking device while browsing or searching Science, Technology, Engineering or Math educational resources

The eye fixations (stops) and saccades (gaze paths) are recorded.

Fixations, typically 150-300 milliseconds, enable a person to gather information. No information can be acquired during saccades.

The colors represent different intervals of time (from green through red).



Eye Tracking in Digital Libraries

- How users of Digital Libraries use metadata and process metadata?
 - Test on three conditions
 - Records with descriptions
 - Records with Metadata
 - Records with both descriptions and metadata



Method

- Pre-exposure search attempt
 - 3 trials to enter search terms using free text, modifiers, boolean expressions etc.
- Exposure to test stimuli – Information in 1 of 3 formats. Eye track during exposure.
 - Metadata only
 - Description only
 - Metadata and Description
- Post- exposure search attempt and follow-up interview

Metadata Display



Title	Why is the planet Mars red?	
Subject:	Oxidation	Creator: Steven W. Squyres
		Publisher: Cornell University
		Format: text/html 3802 bytes
		Type: Text
		Date: 1999-07-16
		Language: en
Identifier:	http://www.ccmr.cornell.edu/education/ask/index.html?	

Title	Gold	
Subject	Mineralogy or petrology Geology	Creator Kirkemo, Harold Newman, William
		Contributor USGS Information Services
		Type Image Text
		Format: text/html
		Language: en
Identifier	http://pubs.usgs.gov/gip/gold/index.html	
Rights	Information presented on this website is considered public information and may be distributed or copied. Use of appropriate byline/photo/image credit is requested.	

Title	The Origin of Epithermal Gold	
Subject	Geochemistry Geology	Creator Clarke, Beresford
		Contributor Suthren, Roger
		Format text/html
		Type InteractiveResource
		vcDate 1998-07
		Language en
Identifier	http://www.brookes.ac.uk/geology/8361/1998/berry/gold1.html	
Rights	Many of the photographs are copyright of the South African Chamber of Mines	

Title	Weathering and Erosion	
Subject	Physical geography	Contributor Field, Maurice
		Publisher University of Tennessee at Martin, Center for Environmental and Conservation Education
		Format text/html
		Type Interactive Resource

Why is the planet Mars red?

Because it's rusty! Pictures from the Mars Pathfinder mission showed of Mars. Pathfinder's Sojourner rover examined these rocks, and found have been found here on Earth - they fell from the sky as meteorites. pure iron, like you'd find in an old nail or a cannonball, but iron that's why have things turned a rusty red color on Mars? After all, iron was f Moon isn't red.

Gold

This United States Geological Survey (USGS) publication highlights the value, how it forms, and gold terms (such as Karat) are covered. Pro prospecting for gold are also explained.

The Origin of Epithermal Gold

This is an illustrated account of the formation and exploitation of epith subaerial volcanic activity. Topics include: Mineralization, Age Range, and Genesis of Epithermal Gold Deposits. Information on exploration an may influence the exploitation of these deposits, is offered. A glossary external websites.

Weathering and Erosion

This Classroom Connectors lesson plan discusses weather conditions a erosion. Students learn to explain the process of physical and chemical and contrast erosion resulting from wind, ice and water. The site provi required, materials, activities, and closure ideas for the lesson. The Classroom Connectors address content with an activity approach while incorporating themes necessary to raise the activity to a higher cognition level. The major motivation is to employ instructional strategies that bring the students physically and mentally into touch with the science they are studying.

A Golden Opportunity for Science

This site uses activities about gold to teach science, centering on the following information: how the pursuit of gold shaped society; metaphors, myths, lores and legends about gold; the mineral properties of gold; the two types of origins of gold (lode deposits and placer deposits); mining and recovery of gold; and the technology of heap leaching. There is also a debate between different interest groups about a small gold-mining town for students to consider, mention of a few historic mining sites and ghost towns for field trips, a map of U.S. regions that have produced gold, and a poster to illustrate the gold mining process, such as exploration, extraction,

Description Display



Scanpath of Metadata Only Condition

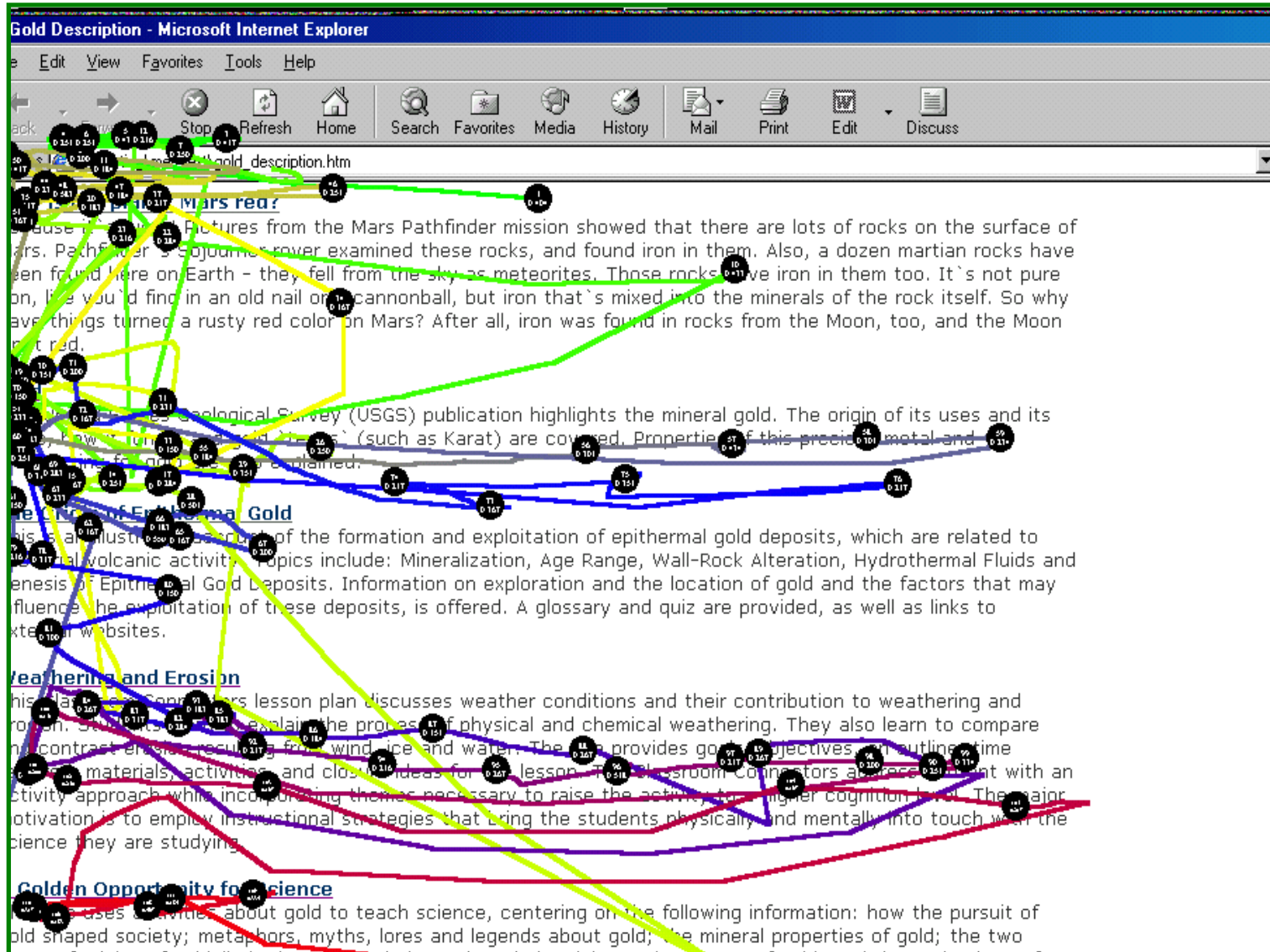
File Edit View Favorites Tools Help

Back Forward Stop Refresh Home Search Favorites Media History Mail Print Edit Discuss

Address \\Testing\metatest\gold_metadata_ready.htm

	vcDate	1998-07
	Language	en
Identifier	http://www.brookes.ac.uk/geology/8361/1998/berry/gold1.html	
Rights	Many of the photographs are copyright of the South African Chamber of Mines	
Subject	Erosion Geography Gold, Maurice Tennessee at Martin, Center for Environmental and Conservation Education	
Identifier	http://www.utm.edu/departments/ed/geose/seventh/7J2.shtml	
Rights:	Copyright 2001, The University of Tennessee at Martin	
Title:	Opportunities for Science	
Subject	Geology Mineralogy Petrology	
Creator	Barna, Carl; Bly, Richard S. O'Sullivan, Shelley Tisdale, Mary Fishman, Shelly	
Contributor	Bureau of Land Management, Office of	
Publisher	Public Affairs	
Format	text/html	
Type	text/html	
Language	en	
Identifier	http://www.blm.gov/education/going_4_the_gold/gold1/er.html	
Rights	Copyright and Other Restrictions Information is Unknown	

Scanpath of Description Only Condition



Scanpath of Metadata & Description Condition

File Edit View Favorites Tools Help

Back Forward Stop Refresh Home Search Favorites Media History Mail Print Edit Discuss

Address http://testing\metatest\silicon_MDEDESC2.htm

Format	text/html
Language	en
Identifier	http://chemistry.beloit.edu/Chip/index.html
Rights	Copyright 2000 by the Trustees of Beloit College and the Regents of the University of California. Published through exclusive license with John Wiley Sons

The lattice structure of the Chip Firing Game and Related Models

In this paper, we use a system, the Chip Firing Game, as a model in physics, economics, computer science. We show that the set of configurations (i.e. the configuration space) of each system started in any configuration has a structural properties. The lattice structure of the configuration space of a Chip Firing Game is of great interest since it implies convergence (at least) if the configuration space is finite. We show that another kind of convergence: all the configurations reachable from two given configurations are reachable from their infimum. In other words, there is a unique first configuration which is reachable from two given configurations. Moreover, the Chip Firing Game is a very general model, and we show how known models can be encoded as Chip Firing Games, and how some results about them can be deduced from this paper. Finally, we define a new model, which is a generalization of the Chip Firing Game, and about which many interesting questions arise.

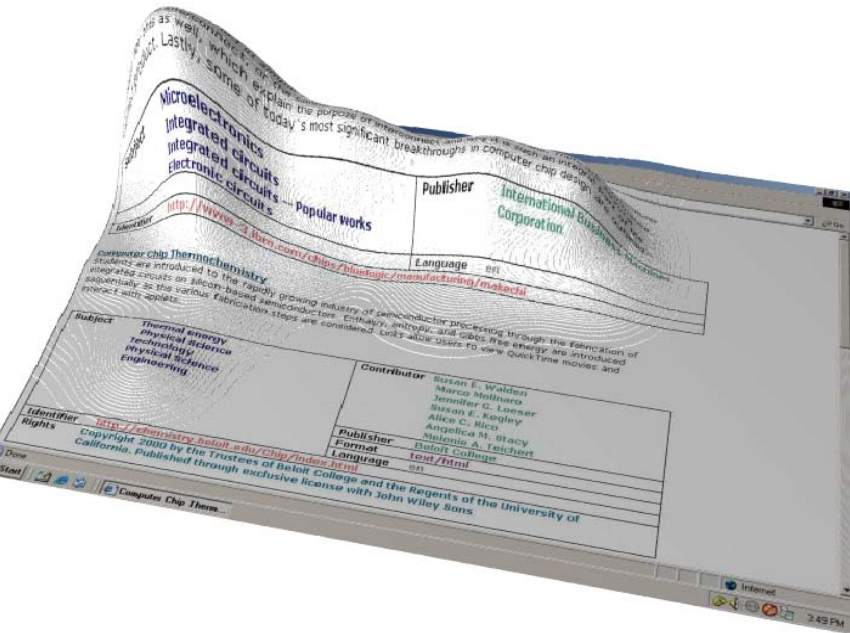
Subject	Pattern Formation and Solitons	Creator	Phan, H. D.
	Adaptation and Self-Organizing Systems		Latapy, M.
	Cellular Automata and Lattice Gases	Format	text
		Type	Text
		Date	2000-08-02

Identifier	http://arXiv.org/abs/nlin/0008003
------------	---

HyperTransport Technology: Simplifying System Design

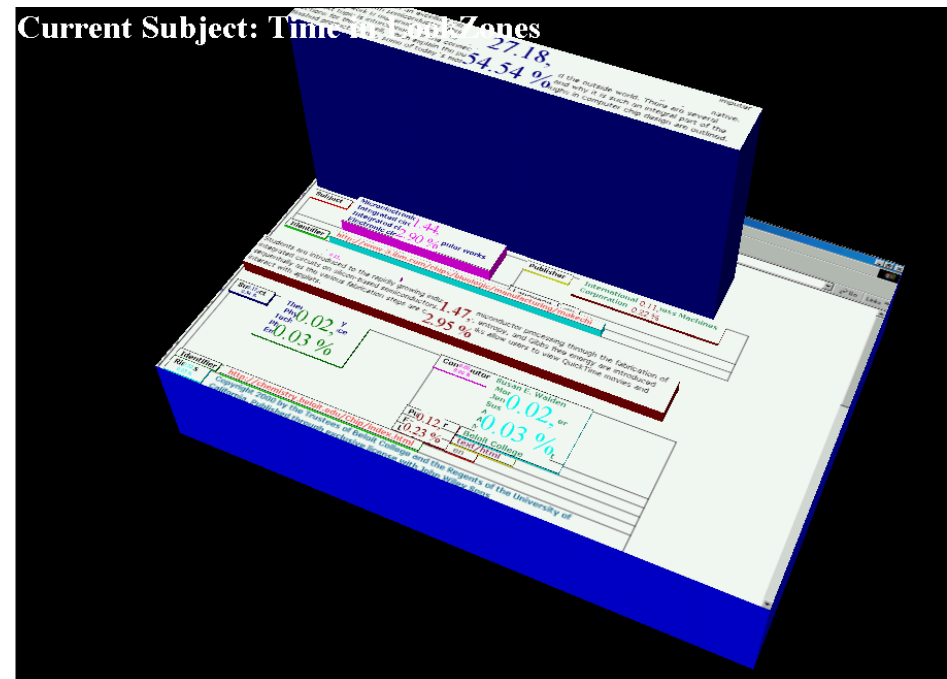
HyperTransport technology is an architecture that allows data transfer between chips in excess of ten gigabytes per second. This white paper from Advanced Micro Devices (AMD), released in October 2002, considers the potential of HyperTransport in designing an input/output bus for a microprocessor. The technology, while accommodating high speed demands, requires relatively little power, making it ideal for a wide range of

Graphically Understanding the Data



Above image: The contour map shows the aggregate of eye fixations (stops). Peak fixation areas are on the Description element, with some interest in the identifier (URL) and subject elements. Note the dominance of the left side.

Below image: The LookZone shows the amount of time spent in each zone (paragraphs in this case). The user spent 27 seconds or 54% of her time looking at the description metadata element. Very little time was spent with the other elements.



Preliminary Findings: Eye Tracking

- Descriptions are viewed in a linear order, but metadata is not
- Titles and sources are the mostly viewed metadata
- The first few sentences in the description are read more carefully; the rest of them are skimmed
- Before selection, a re-visit of the records for confirmation
- Subjects focus on descriptions when both descriptions and metadata are on the same page.

Preliminary Findings: Interview Data

- 100% indicated they would continue to search for more information beyond the document they retrieved.
- 65% changed their initial search terms after exposure to test stimuli; proportionately more so in the combined condition.
- 20% indicated that they would use their chosen document for the intended purpose; all of these were women.
- 60% said they learned something from the retrieved document that helped them restructure their next search. Proportionately more of these were woman than men.
- 100% indicated that they use Google when searching for lecture/lesson information.
- Less than half of the participants knew what metadata was.

Preliminary Findings: Search Attempts

- On post-exposure search attempts the mean number of search terms increased by 25% for those in the combined condition. The number of search terms decreased for both of the other conditions.
- Men used more search terms on their first 3 query attempts, while women used more on their 2nd (post-exposure) query attempts. Men were also more likely to use modifiers and full text queries. Women tended to use more Boolean expressions.

Next Steps for User Study

- **Finer grain analysis of eye tracking metrics**
 - What metadata elements do searchers fixate?
 - What kinds of words are attended to in descriptive information?
 - What do the scan paths of efficient searchers look like?
 - How does ocular behavior map onto actual search behavior and decision making?

