Cognitive Research for Exploratory Search (CRES)

Yuka Egusa (National Institute for Educational Policy Research, Japan),
Noriko Kando (National Institute of Informatics/The Graduate University for Advanced Studies, Japan),
Hitomi Saito (Aichi University of Education, Japan),
Masao Takaku (Tsukuba University, Japan),
Hitoshi Terai (Nagoya University, Japan),
Makiko Miwa (Open University of Japan, Japan)

URL: http://cres.jpn.org/   Email: cres@nii.ac.jp

CRES project

- Aims to investigate user's behavior and cognitive processes during various information seeking tasks on the Web.
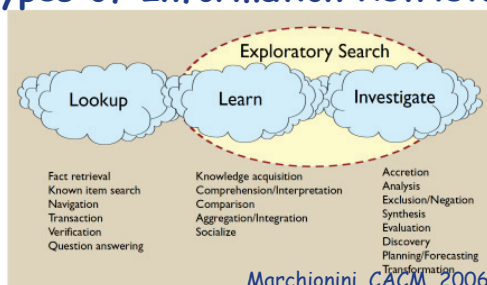
Background

In daily life, users have a lot of things that they want to know, but don't know how to look for them.

- What should I do this coming weekend?
- What is the trial committee system?
- Where can I find a good kindergarten?
- Where can I buy a good car?
- How to write a good project proposal?
- How to plan an enjoyable trip?
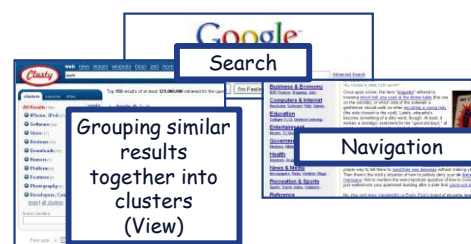- What should be seen at a museum?
- How can I write a good report?
etc.

Types of Information Retrieval

Exploratory Search

Lookup — Learn — Investigate

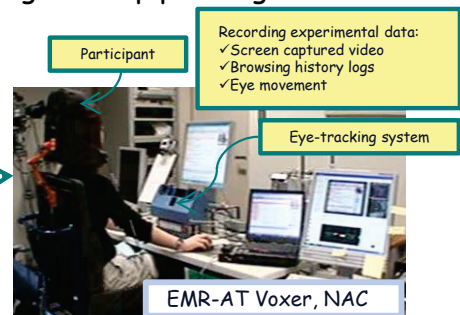|		
Fact retrieval	Knowledge acquisition	Accretion
Known item search	Comprehension/Interpretation	Analysis
Navigation	Comparison	Exclusion/Negation
Transaction	Aggregation/Integration	Synthesis
Verification	Socialize	Evaluation
Question answering		Discovery
		Planning/Forecasting
		Transformation

Marchionini, CACM, 2006

Research questions:
- How users explore the Web for results from search engines
- What users think when they explore the Web.
    -> Analysis of server-side logs is insufficient.

Current Search Engines

Search

Grouping similar results together into clusters (View)

Navigation

In depth analysis of user information seeking behavior is indispensable

● Differences between tasks and information needs:
  - Info. gathering for report writing vs. Info. gathering for trip planning
● Type of users (differences of experience):
  - Undergraduates (11) vs. Graduates (5)

Collected data:
- Browsing history logs
- Screen captured video
- Eye movement
- Thinking aloud
- Interviews, etc…

Procedure

Pre-questionnaires
◆ How much do you use Web weekly?
◆ What browsers do you use?
◆ What search engines do you use?

Repeat twice

Search Task
Report writing   Trip planning
15 min.

Post-questionnaire
◆ How difficult was task?
◆ Satisfied with search results?

Interview
◆ Information-seeking processes
◆ Show screen-captured video
≒30 min.

Instructions
✓ You have 15 minutes to collect related information through the Web.
✓ Add pages to browser's book marks if useful.

We proposed analytical frameworks:

"Lookzone": Set of categories indicating which part of resulting pages participants looked at.

"Web action categories": We defined 10 action categories for analyzing a user's behavior on the Web

"Link Depth": How far searchers browse into search engine result pages (SERP).

"Concept map": Measuring the change in the user's knowledge due to the search by comparing these before and after maps.

"Taxonomy of Knowledge Modification and Knowledge Utilization Patterns": Frameworks for content-analyzing the think-aloud and interview data.

We developed tools for collecting and analyzing data

"COPATT": Platform for integrating browsing history logs, screen captured video, eye movement, and annotating users' actions.

"QT-Honey": New client side logging tool based on Lemur Query Log Toolbar (Lemur project).

"VizCMaps": Visualization tool for pre- and post-concept maps

"Scanpath2SVG": Visualization tool for eye movement data.

Recording experimental data:
✓ Screen captured video
✓ Browsing history logs
✓ Eye movement

Participant

Eye-tracking system

EMR-AT Voxer, NAC

T60, Tobii

(Updated 2013-06-22)

CRES: Cognitive Research for Exploratory Search (2)

"Web action categories"

> ⬜ (red) > ⬜ (blue) Group
> ⬜ (red) > ⬜ (orange) Task

Web action categories		Undergraduates (n=9)		Graduates (n=5)	
		Report Mean(SD)	Trip Mean(SD)	Report Mean(SD)	Trip Mean(SD)
Search	searching using search engine	8.00 (4.37)	6.27 (4.92)	9.20 (2.99)	7.80 (5.27)
Link	clicking on page links	19.36 (6.26)	35.64 (8.65)	28.80 (7.28)	33.20 (8.57)
Next	going to next page	0.45 (0.78)	0.91 (1.08)	0.80 (0.75)	0.20 (0.40)
Back	going back to previous page	17.45 (7.51)	22.27 (13.80)	10.40 (8.11)	10.80 (7.19)
Jump	going to bookmarked or history page	2.64 (1.61)	2.64 (1.92)	2.20 (1.72)	3.40 (2.25)
Browse	browsing new search results	1.82 (2.25)	0.18 (0.57)	0.80 (1.17)	0.60 (1.20)
Submit	clicking submit button	1.27 (2.60)	3.00 (2.80)	7.60 (11.29)	4.60 (4.84)
Bookmark	adding bookmarks	4.55 (2.06)	4.55 (2.31)	8.00 (1.26)	8.00 (5.76)
Change	changing from one tab to another	2.45 (5.37)	3.55 (3.23)	43.60 (23.59)	28.40 (17.85)
Close	closing tabs or windows	0.36 (0.64)	2.36 (1.77)	4.20 (3.54)	6.00 (8.79)

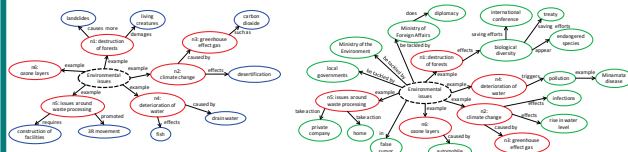Undergraduates were more likely to click links during the trip task than during the repot task.
Graduates tend to use multiple tabs and windows.
Graduates bookmarked more pages than the undergraduates.

"Lookzone"

	Lookzone	Undergraduates (n=9)		Graduates (n=5)	
		Report Mean(SD)	Trip Mean(SD)	Report Mean(SD)	Trip Mean(SD)
Operation for browser	1 Title bar	3.78 (6.81)	1.00 (1.56)	0.40 (0.80)	0.80 (0.98)
	2 Menu	0.22 (0.42)	0.11 (0.31)	1.80 (3.12)	0.00 (0.00)
	3 Bookmark	4.22 (5.90)	0.00 (0.00)	0.00 (0.00)	0.20 (0.40)
	4 Toolbar	1.33 (1.63)	1.22 (1.40)	0.40 (0.80)	0.40 (0.80)
	5 URL bar	0.56 (1.07)	0.22 (0.42)	0.40 (0.49)	0.00 (0.00)
	6 Search bar	0.00 (0.00)	0.00 (0.00)	6.40 (7.50)	4.00 (7.04)
	7 Search bar button	0.00 (0.00)	0.00 (0.00)	0.40 (0.49)	0.20 (0.40)
	8 Tab	8.11 (9.81)	9.22 (17.94)	12.00 (14.13)	6.00 (6.63)
	12 Scroll bar	0.11 (0.31)	0.00 (0.00)	0.60 (0.80)	0.00 (0.00)
	21 Find in a page	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)
	22 Status bar	1.78 (3.39)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)
Contents area of SERP	9 Link for services	17.67 (23.44)	5.11 (9.33)	2.40 (2.06)	2.20 (2.14)
	10 Query box	36.89 (36.71)	12.56 (11.93)	5.60 (4.36)	3.00 (4.65)
	11 Search button	0.89 (1.10)	0.67 (0.82)	0.40 (0.49)	0.20 (0.40)
	13 Number of hits	0.44 (0.96)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)
	14 Sponsor link	6.67 (7.85)	12.44 (9.93)	0.00 (0.00)	11.40 (13.99)
	15 Spell check	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.20 (0.40)
	16 Title	59.67 (38.92)	42.11 (34.19)	41.20 (26.80)	39.20 (40.82)
	17 Snippet	91.11 (55.59)	37.00 (32.84)	74.80 (42.56)	28.40 (28.00)
	18 URL	40.89 (34.27)	15.56 (11.35)	18.00 (9.21)	12.40 (11.83)
	19 Related search	3.00 (4.03)	2.56 (4.11)	1.20 (1.94)	1.20 (1.17)
	20 Link for next page	1.00 (1.89)	0.78 (1.03)	1.00 (1.10)	1.00 (2.00)
Misc	Out of lookzone	52.89 (53.43)	18.89 (14.51)	21.60 (14.47)	17.00 (8.60)
	Lack of eye position	83.44 (73.10)	70.78 101.06)	15.00 (10.94)	7.20 (4.71)

> ⬜ (red) > ⬜ (blue) Group
> ⬜ (red) > ⬜ (orange) Task

Graduates tend to use more advanced web browser features. e.g. search bars and tabs. Snippet areas more attractive in report tasks than trip tasks.

Eye movement and click ranking
View and click ranking

Undergraduates

Graduates

- Report task: Gazing
- Trip task: Gazing
- Report task: Clicking
- Trip task: Clicking

Users viewed more low-ranking pages for report tasks than for trip tasks. Users' viewing behavior was heavily influenced by type of query, i.e. navigational or informational.

Link Depth Overview
How far users explore the Web

SE — (a) Search — SR (b) — Blank (c) Link — 1 (d) Link — Return — (e) — 2 (f) Link — (g) — Link New tab — (h) Addtab — 3 (i) — Jump

Examples of Link Depth
Report   vs.   Trip

In the trip task, most of the participants explored more deeply on the Web than in the report task. This result reflects the relationship between tasks and the physical characteristics of the Web.

"Concept map"

Graphical representation (Novak & Gowin, 1984)
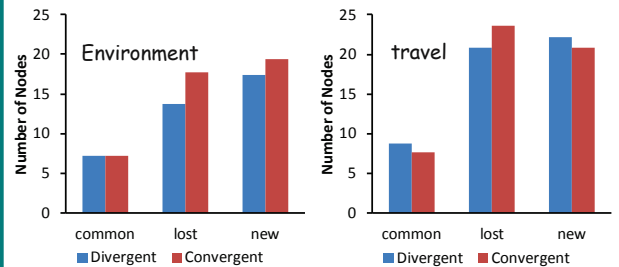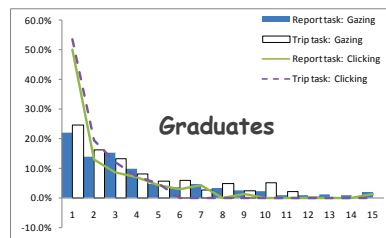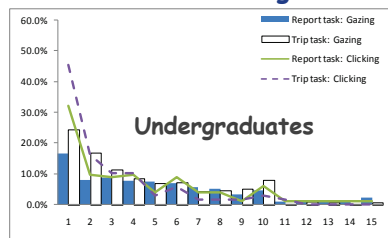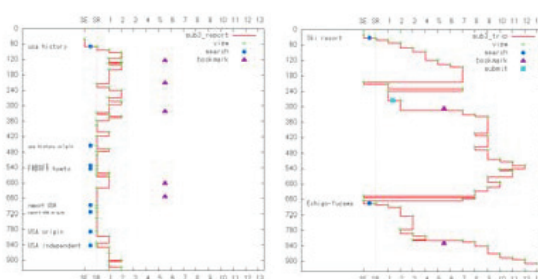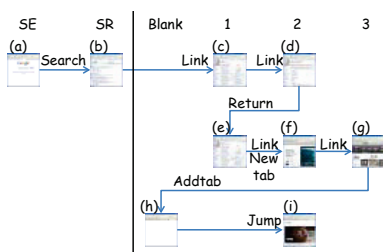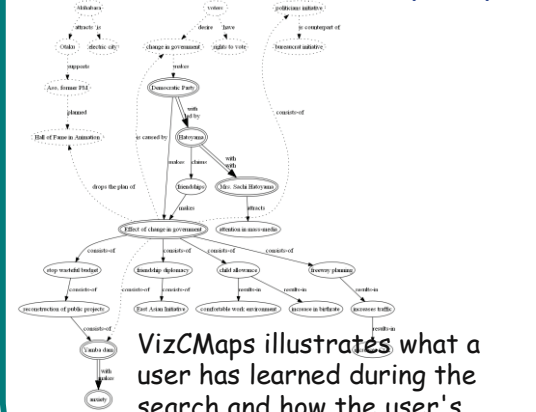→ To allow people to represent their knowledge

Pre-search concept map

Pre-search concept map

Common Node : Nodes exist in both pre and post cmaps
Lost Node : Nodes exist only in a pre cmap
New Node : Nodes exist only in a post cmap

Environment

travel

Number of Nodes

common   lost   new

- Divergent
- Convergent

Common Nodes < Lost and New nodes
→Cmaps dynamically changed before and after searching.
Number of Lost and New nodes: Environment < Travel

VizCMaps: Combined graph of pre- and post-search concept maps

VizCMaps illustrates what a user has learned during the search and how the user's knowledge representations have changed.

Outcomes

(Detailed info is available at **http://cres.jpn.org/**)
- Terai et al., IIix2008 (Oct. '08)
- Kando et al., Dagstuhl Seminar (Mar. '09)
- Miwa, Organizing a Panel at ASIS&T (Nov. '09)
- Egusa et al., HICSS 2009, (Jan. '10)
- Takaku et al., JSIK Journal (Sep. '10; online: May '10)
- Egusa et al., IIiX2010(Aug. '10)
- Saito et al., AIRS2010 (Dec. '10)
- Miwa et al., Information Research (Sep. '11)
- Saito et al., CogSci2012 (Aug. '12)