
How IT Supports Knowledge Discovery and Learning Processes on the Web

Tracy Jenkin
Queen's School of Business

*The Monieson Centre Seminar
February 20th, 2008*

Agenda

- Introduction
 - Motivation
 - Key concepts
- Theoretical Framework
 - Framework Development
 - Study 1: Experiment
- Contributions
 - Theory
 - Practice

Introduction

Motivation

- Learning about the **environment** is important for survival, competitiveness and growth (Daft & Weick 1984, Hambrick 1983, Penrose 1959, Summers 1980).
- Knowledge discovery on the Web
 - Used extensively by organizations
 - Challenges due to information overload
- Theoretical framework linking tools, knowledge discovery and learning.

Key Concepts

- Knowledge:
 - As object, state of mind, and capability (Alavi & Leidner 2001).
 - Information with direction, which enables action and decisions and increases the capacity for effective action (Beccerra-Fernandez et al. 2004, Huber 1991, Kim 1993).
- Mental models:
 - Frameworks to simplify, organize, and make sense of information (e.g. Crossan et al. 1999).
- External knowledge-discovery:
 - Development or acquisition of new knowledge (Beccerra-Fernandez et al. 2004) outside of the organization.

Theoretical Framework

Knowledge Discovery

1. **Question:** specific, record-oriented (e.g. Dworman et al. 2000, Vandenbosch & Higgins 1996).
Answer: focused, specific knowledge.

What types of questions are being asked?

What types of answers are being found?

2. **Question:** general, pattern-oriented questions (e.g. Dworman et al. 2000)
Answer: patterns, relationships between related concepts.

3. **Question:** vague or unarticulable.
Answer: novel knowledge or “surprising associations” (Kimbrough 2001).

Learning
Mode

Learning Mode

1. **Mental Model Maintenance:** confirm and validate existing mental models (Vandenbosch & Higgins 1996).

To what degree are mental models changed?

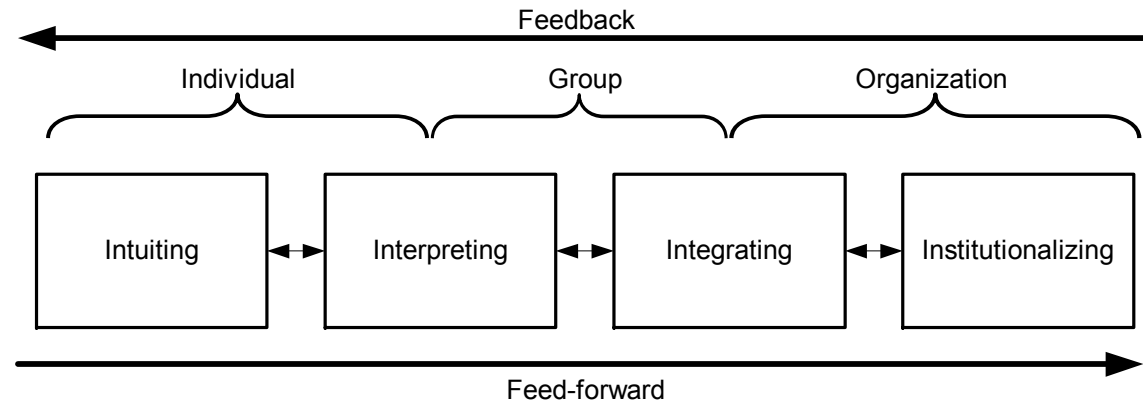
2. **Mental Model Building:** generates new mental models, existing knowledge is defined, changed and more efficient, specialized (Norman 1982, Vandenbosch & Higgins 1996).

- 3.

Learning
Process

How are mental
models changed?

Learning Process



(adapted from Crossan et al. 1999)

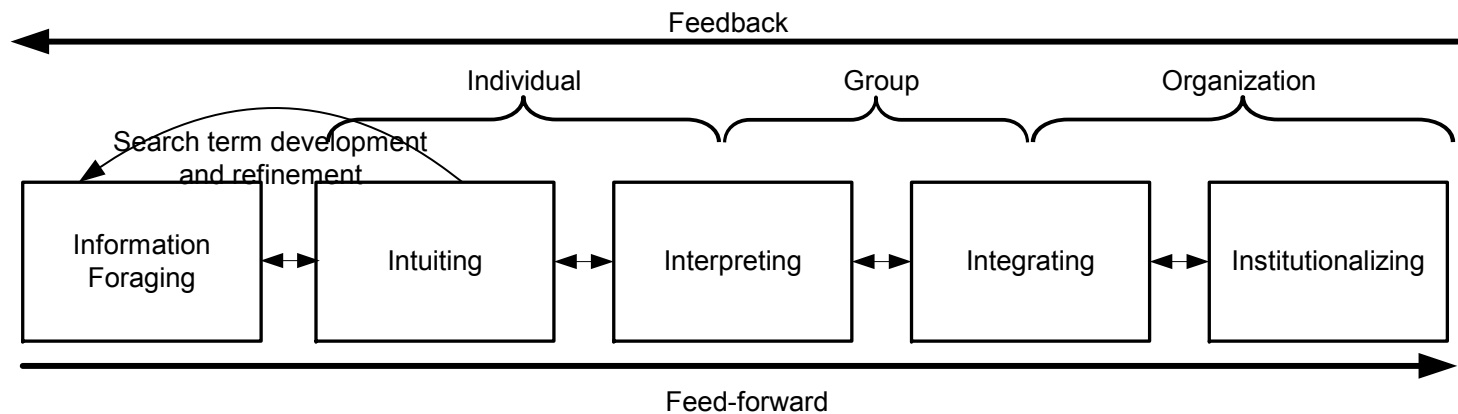
Other learning process models include:

- Scanning (Daft & Weick 1984)
- Reconceptualizing the problem
- Deciding to search
- Search and evaluate (Majchrzak et al. 2004)

Learning Process

Learning Process

How are mental models changed?



(adapted from Crossan et al. 1999)

Other learning process models include:

- Scanning (Daft & Weick 1984)
- Reconceptualizing the problem
- Deciding to search
- Search and evaluate (Majchrzak et al. 2004)

How do tool characteristics support different types of questions?

Tools

- 1. Question:** focused, specific knowledge.
Tools: search engines.
Example: Google.
- 2. Question:** general, pattern-oriented questions
Tools: clustering and categorization tools, semantic web technologies.
Example: Vivisimo.
- 3. Question:** vague or unarticulable (novel knowledge and surprising associations).
Tools: text mining tools
Example: **Athens**

**Novel-Knowledge Discovery
(NKD) Design Theory**

Tools

Tools

General Tool Characteristics

LEVEL 1

- Searches across a domain of content using keywords entered by the user. (*, #)
- Produces list of results matching search terms. (*)
- Includes links to relevant content and meta-knowledge of that content. (*, #)
- Sorts results by relevance / similarity to search terms (for example, weighted keyword scores, and page ranking). (*)

LEVEL 2

- Includes characteristics of level 1 tools (where noted *).
- Groups results into clusters of similar content, thus identifying various dimensions of the search term and related concepts. (#)
- Allows user to drill down into clusters for more detailed results. (#)

LEVEL 3

- ~~Novel Knowledge Discovery (NKD) Tools (Where noted #).~~ Novel Knowledge Discovery (NKD) Tools (Where noted #).
- Distills and summarizes results to a limited set.
- Directs user's attention to most interesting, relevant and novel results that are indirectly connected original search term.
- Provides a measurement of interestingness to enable evaluation of results.
- Reframes existing knowledge about the search term entered.
- Supports users in challenging existing mental models regarding search term entered.



Clustered Results

Top 139 results of at least 180 retrieved for the query "raynaud's syndrome" + "blood viscosity" (Details)

- ▶ "raynaud's syndrome" + "blood viscosity" (139)
- ▶ Raynaud's Disease (26)
- ▶ Fish Oil (22)
- ▶ Ginkgo, Biloba (20)
- ▶ Patients with Raynaud's syndrome (13)
- ▶ Red blood cell (11)
- ▶ Blood Viscosity Abnormalities (7)
- ▶ Medical (9)
- ▶ Disorder (8)
- ▶ Enhanced Blood Viscosity (7)
- ▶ Articles, Med (5)
- More

Find in clusters:
Enter Keywords Go

- [Raynaud's Syndrome - Abstracts : Online Reference For Health Concerns](#) [new window] [frame] [cache] [preview] [clusters]
Raynaud's Syndrome - Abstracts : Online Reference for Health Concerns ... inhibiting spontaneous platelet and red cell aggregation and reducing **blood viscosity**.
www.lef.org/protocols/abstracts/abstr-128.html - Gigablast 1, MSN 2, Ask 3
- [Raynaud's Syndrome, Raynaud's Disease treatment](#) [new window] [frame] [cache] [preview] [clusters]
... Platelet and **blood viscosity** abnormalities are also implicated. ... **Raynaud's syndrome** may involve increased sensitivity of alpha-2-adrenergic ... **Raynaud's Syndrome, Raynaud's Disease treatment** with naturopathic ...
Description: Raynaud's Syndrome, Raynaud's Disease treatment with naturopathic medicine, vitamins, herbs, nutrition
www.naturdoctor.com/Chapters/Diseases/Raynauds.html - Ask 1, Gigablast 3
- [Raynaud's Syndrome](#) [new window] [frame] [cache] [preview] [clusters]
Raynaud's Syndrome. A Clinical Update on **Raynaud's Syndrome.** **Raynaud's** ... important for **Raynaud's syndrome** with its vasospasms.. Ginkgo biloba ... **Raynaud's syndrome:** Using a range. of therapies to help patients.
...
www.dr-bylenok.com/raynaud's_syndrome.htm - Gigablast 2, Ask 6
- [eMedicine - Raynaud Phenomenon : Article by Jeffrey R Lisse, MD, FACP](#) [new window] [frame] [cache] [preview] [clusters]
The decreased **blood flow** may be a result of increased **blood viscosity**, abnormalities of the ... Intermittent epoprostenol (prostacyclin) infusion in patients with **Raynaud's syndrome**.
www.emedicine.com/med/topic1993.htm - MSN 1, Ask 22, Gigablast 41
- [Ginkgo biloba - RxList Alternatives](#) [new window] [frame] [cache] [preview] [clusters]
... Peripheral and cerebral circulatory stimulation, Peripheral vascular disease, **Raynaud's syndrome** ... It has also been reported as reducing **blood viscosity**. It's ability to increase vascular ...
www.rxlist.com/cgi/alt/ginkgo.htm - Gigablast 11, Ask 12, MSN 20
- [kiwi.uchicago.edu](#) [new window] [frame] [cache] [preview] [clusters]
TI - Fish oil, **Raynaud's syndrome**, and undiscovered public knowledge. PG - 7-18 ... MH - **Blood Viscosity**/*drug effects. MH - Fish Oils/pharmacology/*therapeutic use
kiwi.uchicago.edu/webwork/PMME/INFformat1 - MSN 4, Gigablast 17, Ask 73



1. Athens Summary Page

address H:\Thesis\Experiment\temp\index.html

A T H E N S
Novel Information Discovery

This page generated Thu Feb 08 14:59:16
Original Query: "raynaud's syndrome" + "blood"

Root

[Cluster 1](#): (Protein, Gene, Genome, Books, Domain, RefSeq, alpha, PopSet, Probe, BioAssay)

[Cluster 2](#): (Marshall, Clin, Protein, Gene, Genome, Auton, Books, Domain, RefSeq, Mohan)

[Cluster 3](#): (Marshall, Protein, Genome, Gene, Clin, Auton, Books, Domain, RefSeq, Mohan)

[Cluster 4](#): (REMARK, GeneRIF, Bbeta, Haemost, Brennan, Fibrinogen, Nottingham, CommentFeaturesSequence, DBSOURCE, JavaScript)

[Cluster 5](#): (Equine, Assoc, EIPH, Protein, Gene, Tobin, Pascoe, Smith, UniGene, thoroughbred)

Root | [Cluster 1](#)

Novel Cluster 1.1 Books RefSeq Entrez BioAssay Domain Probe HomoloGene PopSet UniGene Journals GeneRIF G UniSTS Matcher DataSet

2. Athens Novel Clusters Page

A T H E N S
Novel Information Discovery

This page generated Thu Feb 08
Original Query: "raynaud's syndrome"

3 clusters generated from 29 links.

Protein Genome Gene Books Domain Valberg Mickelson

[Cluster 1](#): (Protein, Genome, Gene, Books, Domain, RefSeq, HomoloGene, PopSet, Probe, BioAssay)

- http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&list_uids=7490716&dopt=Abstract
- http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&list_uids=15318347&dopt=Abstract
- http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&list_uids=3676281&dopt=Abstract
- http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=pubmed&list_uids=16151699&cmd=Retrieve&indexed=google
- http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&list_uids=16151699&dopt=Abstract
- http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&list_uids=10619752&dopt=Abstract
- http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&list_uids=14970703&dopt=Abstract
- http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&list_uids=16820721&dopt=Abstract
- http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&list_uids=11113112&dopt=Abstract
- http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&list_uids=10677289&dopt=Citation
- http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Abstract&list_uids=11867233
- http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&list_uids=15350187&dopt=Abstract
- http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&list_uids=12857913&dopt=Abstract
- http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&list_uids=3119826&dopt=Abstract
- http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&list_uids=9593490&dopt=Abstract
- http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&list_uids=8275703&dopt=Abstract
- http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&list_uids=9475402&dopt=Abstract

Theoretical Framework

Tools	Knowledge Discovery	Learning Process Support	Learning Mode
Level 1 Tools	Specific, focused knowledge	Information Foraging	Mental Model Maintenance
Level 2 Tools	Patterns and relationships	Information Foraging; Interpreting (partial)	Mental Model Tuning
Level 3 Tools	Novel knowledge and surprising associations	Information Foraging; Intuiting (partial); Interpreting	Mental Model Building

Caveats

- Type of question:
 - Context specific.
 - ≠ type of answer found.
- Type of knowledge discovered:
 - Depends on individual's prior knowledge.
- Learning process support:
 - Related to learning mode.
- Other theoretical lenses possible.

Competing Theory

- Technological Frames
 - Technology-specific mental models (Orlikowski & Gash 1994).
 - Expectations, assumptions and meanings surrounding use and outcomes of specific technologies.
 - Dominant design contributes to technological frames.
 - Impose technological frame of the “familiar” onto “unfamiliar” technology (e.g. Hargadon & Douglas 2001, Orlikowski & Gash 1994).
 - Tool congruence with technological frame will increase performance.

Study 1 - Experiment

- Goals:
 - Test theoretical framework and evaluate instantiation of NKD design theory – Athens.
- Subjects:
 - 99 students.
- Task:
 - Discover novel knowledge on the Web.
*“...find **novel** ways of applying their MP3 expertise – for example, to develop new services, new products, or new markets”*
- Design & Analysis:
 - Multiple equivalent control groups, with pretest.
 - Multivariate analysis of covariance (MANCOVA).

Dependent Variables

Type of Knowledge
Discovered

Learning
Process Supported

Learning
Mode Enacted

Specific, focused,
“**deep**” knowledge.

Patterns,
relationships,
“**broad**” knowledge.

Surprising
associations,
“**novel**” knowledge.

Information Foraging

Information Foraging;
Interpreting (partial)

Information Foraging;
Intuiting (partial);
Interpreting

Mental Model
Maintenance (**MMM**)

Mental Model
Tuning (**TUN**)

Mental Model
Building (**MMB**)

Independent Variable

Treatment

Type of Knowledge
Discovered

Learning
Process Supported

Learning
Mode Enacted

<p>Google (Level 1 Tool)</p>	<p>Specific, focused, “deep” knowledge.</p>	<p>Information Foraging</p>	<p>Mental Model Maintenance (MMM)</p>
<p>Vivisimo (Level 2 Tool)</p>	<p>Patterns, relationships, “broad” knowledge.</p>	<p>Information Foraging; Interpreting (partial)</p>	<p>Mental Model Tuning (TUN)</p>
<p>Athens (Level 3 Tool)</p>	<p>Surprising associations, “novel” knowledge.</p>	<p>Information Foraging; Intuiting (partial); Interpreting</p>	<p>Mental Model Building (MMB)</p>

Covariates

Original Level of
Knowledge

General Computer
Self-efficacy (GCSE)

Personal
Innovativeness
with IT (PIIT)

Web Self-efficacy
(WSE)

Demographics

Other Variables

Satisfaction with Tool

Training

Clarity of Instructions

Manipulation Check:
Tool Level

Manipulation Check:
Task

Tool Features

Tool Challenges

Tool Improvements

Hypotheses – Theory 1

Knowledge Type		
H1a	Novel	Athens > Vivisimo , Google
H2a	Broad	Vivisimo > Google , Athens
H3a	Deep	Google > Vivisimo , Athens
Learning Mode		
H4a	MMB	Athens > Vivisimo , Google
H5a	TUN	Vivisimo > Google , Athens
H6a	MMM	Google > Vivisimo , Athens
Learning Process		
H7a	Intuit	Athens > Vivisimo , Google
H8a	Interp	Athens > Vivisimo , Google
H9a	Forage	Athens > Vivisimo , Google
H10a	Interp	Vivisimo > Google

Competing Hypotheses – Theory 2

Knowledge Type		
H1b	Novel	Google > Vivisimo, Athens
H2b	Broad	Google > Vivisimo, Athens
H3b*	Deep	Google > Vivisimo, Athens
Learning Mode		
H4b	MMB	Google > Vivisimo, Athens
H5b	TUN	Google > Vivisimo, Athens
H6b*	MMM	Google > Vivisimo, Athens
Learning Process		
H7b	Intuit	Google > Vivisimo, Athens
H8b	Interp	Google > Vivisimo, Athens
H9b	Forage	Google > Vivisimo, Athens
H10b	Interp	Google > Vivisimo, Athens

Results

- Omnibus test:
 - Significant differences between tools (p 0 .001).
- Theory 1:
 - H3a (**Deep Knowledge**) supported (**G>V,A**).
- Theory 2:
 - H3b (**Deep Knowledge**) supported (**G>V,A**).
 - H8b/H10b (**Interpretation**) supported (**G>V,A**).
 - H9b (**Foraging**) supported (**G>V,A**).
 - **Google** highest mean in ALL outcome measures.

Theory 2 dominates

Qualitative Results

- Q: What were you trying to accomplish in your search?
 - Google
 - “I was trying to find ideas about novel applications to mp3 players. I was also trying to validate an idea that I already had, which was to create playlists on the go.”
 - Vivisimo
 - Any new idea that would be clear from what was seen on the results screen. This way I would not have to open each search result and actually read the page in depth. I wanted obvious answers in front of me, which I did not often get, hence why I only had 3 ideas. I am impatient and dislike reading webpages in more detail unless I know they will be 100% useful.
 - I was trying to find additional information and background facts about ideas that I already had a basic idea about. For 3 or 4 out of 5 of my ideas, I came up with a brief outline of what I was looking for in my head before I began searching. The search tool gave me the initial idea to search for hearing aid possibilities in MP3 players.
 - I was trying to generate new ideas for the company, but this was hard without specific knowledge of the direction I already wanted to go in, which I did not have.

Qualitative Results

- Athens comments:
 - “Being new to Athens, I required a few minutes to play around on it to better familiarize myself with its setup. Also, I was a little overwhelmed with all of the information at the beginning, while I simply found it different from typical search engines.”
 - “Did not give descriptions of website links like traditional search websites (aka - Google).”
 - “I found it almost no better than a regular search tool, such as Google.”
 - “It is difficult to say, because the tool uses a different approach. I would appreciate, if it provided some of the Google-type search too, because it would be more helpful with the search words I chose.”
 - “My search terms were not relevant enough to what I was looking for so I came up with many clusters that were not relevant to what I wanted to do.”
 - “A lot of the results had little to do with mp3's or anything that I would have wanted to find.”
 - “The results were irrelevant to what I was looking for.”
 - “The searches were not what I expected.”

Further Analysis

- Reviewed results - item level:
 - Theory 2 and **Google** still dominate.
 - New evidence supporting **Athens**, **Vivisimo** and Theory 1.
 - **Athens** and **Vivisimo** highest mean on several items.
- Reviewed results assuming Google outperforms:
 - H_0 – **Google** will outperform other tools 5 – 15%.

Further Analysis

H_0 – Google will outperform other tools 5 – 15%

Items	5%	10%	15%
Novel1	Accept	Accept	Accept
Novel2	Accept → New	Accept	Reject (A)/Accept (V)
Novel3	Accept → Surprisingly new	Accept	Accept
Novel4	Accept	Accept	Reject
Broad3	Accept	Accept	Accept
Broad4	Accept (A)/Reject (V) → Relationships	Accept (A)/Reject (V)	Accept (A)/Reject (V)
Broad5	Accept	Accept	Accept
Deep2	Accept	Accept	Accept
Deep3	Accept	Accept	Accept (A)/Reject (V)
Deep4	Accept	Accept	Accept

Further Analysis

- Qualitative Results:
 - Athens Comments:
 - “Got me thinking outside of the box.”
 - “Some of the findings the tool provided were the least I would expect to draw my attention to otherwise.”
 - “It got information from web pages that connected MP3 technology to ideas I would never have thought of on my own.”
 - “The ability to find concepts not directly connected to the search terms exposed me to potential applications for mp3 players I wouldn't have otherwise thought of.”
 - “The tool helped me brainstorm what types of services might be useful for my company, which do not come to mind right away.”
 - “The clusters assisted me with today's task. The clusters made it easy to see where similar words were grouped and allowed me to quickly narrow down which cluster was more interesting.”

Further Analysis

- Key Findings:
 - **Athens** outperforms for:
 - “surprisingly new” knowledge (NOVEL).
 - “reframing” mental models (MMB).
 - **Vivisimo** outperforms for:
 - Finding “relationships” (BROAD).
 - “Interpreting” raw results (INTERP).
 - Developing “new insights” (INTUIT).

Study 1 - Conclusions

- Theory 1
 - Promising findings regarding Athens and “**surprisingly new**” knowledge and “**reframing**”.
 - **Clustering** is useful as a interpretation tool.
 - Athens prototype needs additional work to support **intuiting and interpreting**.

Study 1 - Conclusions

- Theory 2
 - Difficult to break out of “**focused search**” and expectations.
 - **Changed task** from level 3 to level 1.
 - Individuals have strong “**technological frames**” oriented towards Google.
 - Clear tool differences regarding **information foraging**.

Contributions

Contributions

- Theory:
 - Develops theoretical framework.
 - Extends past research linking tools to learning.
 - Extends 4I organizational learning process model.
 - Test competing theories.

Contributions

- Practice:
 - Help practitioners “manage” their learning.
- Addresses calls for additional research:
 - How different tool characteristics support different aspects of learning.
 - Finer distinctions of knowledge categories.
 - IT support for mental model modification.

Questions?



Thank you!

