



Connecting Searching with Learning

Kevyn Collins-Thompson

Associate Professor
University of Michigan

School of Information
College of Engineering, EECS



M | SCHOOL OF INFORMATION
UNIVERSITY OF MICHIGAN

University of Waterloo
Data Systems Seminar
May 16, 2016

Joint work with my collaborators:

Soo Young Rieh

Michigan

Rohail Syed

Michigan

Paul Bennett

Microsoft Research

Carl Haynes

Syracuse

Karthik Raman

Cornell / Google Research

Long-term goal:

Optimally connecting users & information for personalized learning

- Search technology is a key resource for learning
 - People use Web search as a primary resource for learning-related tasks
 - Learning applications are starting to rely on search as a backend service to find the right content

Learning-related Web search tasks are popular, but can take multiple queries + significant time

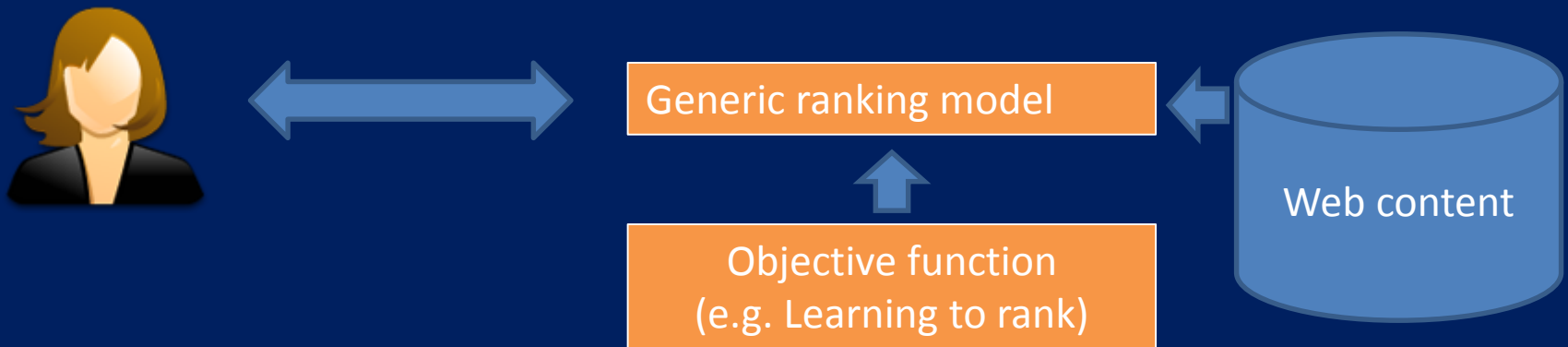
Search task	% of all sessions	Avg queries needed	Avg session len (minutes)
Discover more about specific topic	8.0%	6.8	13.5
Find specific fact	7.0%	3.2	7.9
Find facts about a person	1.8%	6.9	4.8
Learn how to perform a task	2.0%	13	8.5
Other education-related tasks	1.5%	n/a	n/a

Source: P. Bailey et al., User task understanding: a Web search engine perspective. NII Shonan Talk, 8 Oct 2012.

Long-term goal:

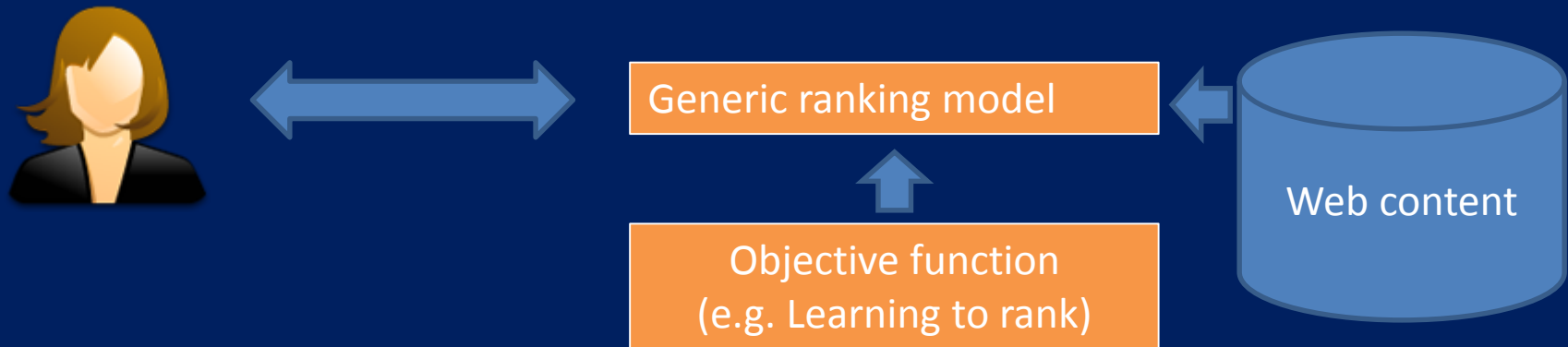
Optimally connecting users & information for personalized learning

- Current Web retrieval models are optimized for fast, relevant response to single queries
 - Not optimized for complex, multi-step tasks
 - Poor at personalization, especially for learning tasks
 - Ignore important cognitive aspects such as the difficulty of the material or concept dependencies

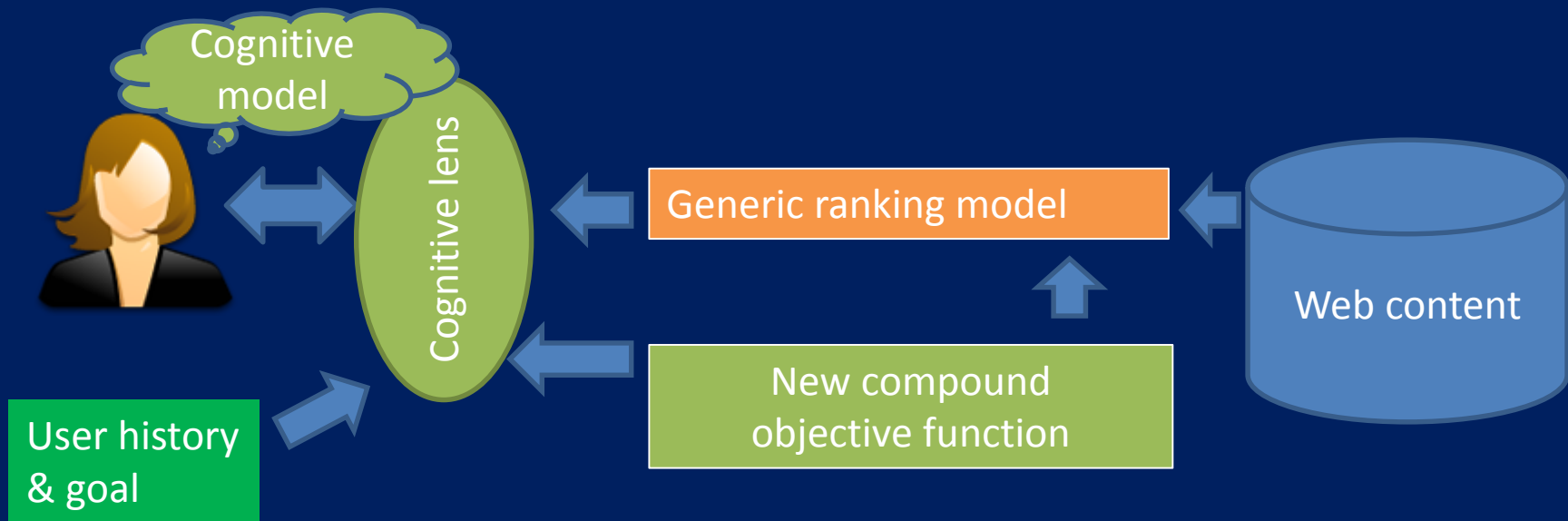


Long-term goal:

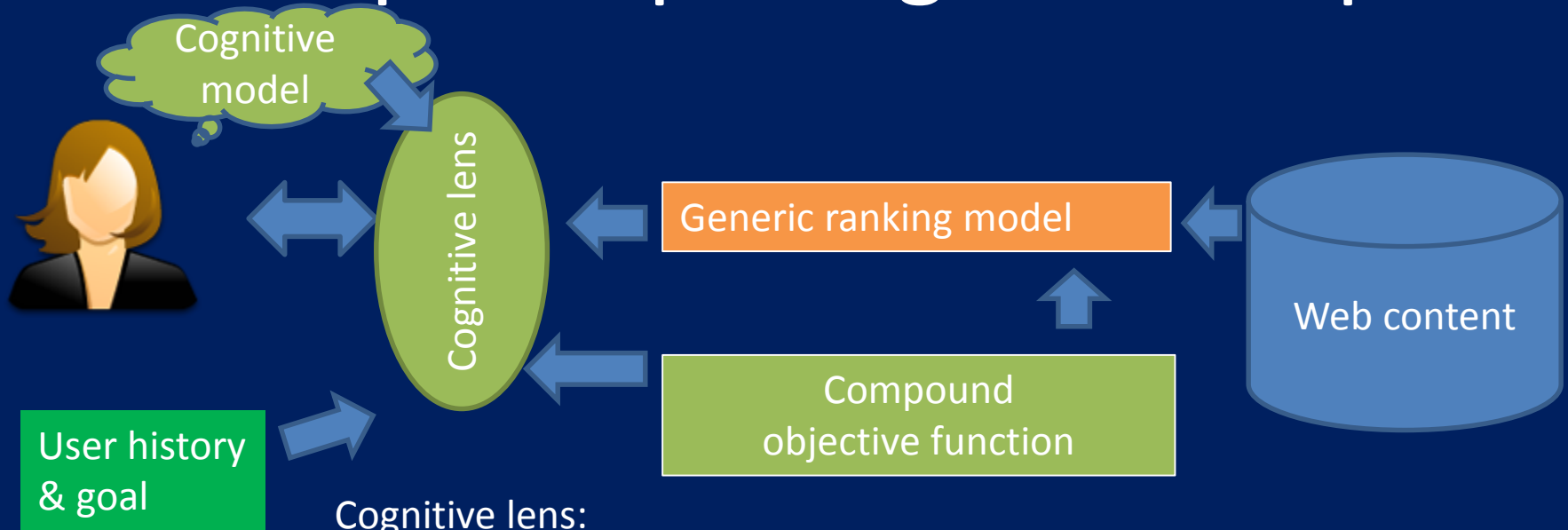
Optimally connecting users & information for personalized learning



Idea: Optimize the ranking as seen through the user's cognitive "lens"



Example: exploring a new topic



Cognitive lens:

- Find representative aspects of the topic: predict future queries
- Find the right level of effort
 - Repeated, supportive exposure to target concepts
 - Efficient coverage of target concepts
 - Consider user's prior knowledge

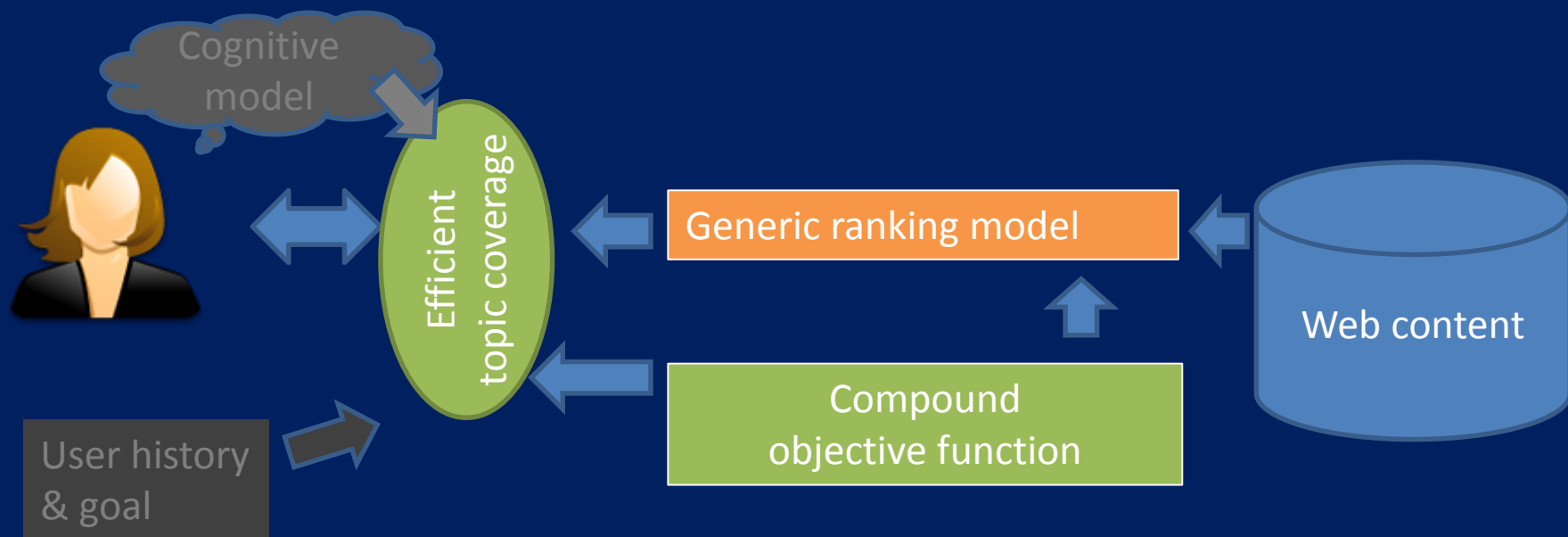
Multi-step objective function:

1. Identify candidate docs with efficient coverage of key vocab
2. Jointly select documents and aspects toward the vocab coverage goal

Implications of algorithms that account for user learning

- Web search (user: person)
 - Greatly improved support of learning tasks
 - More time well spent
 - Better retention, engagement
 - Support higher-level connections between ideas
- Learning systems (user: application)
 - Better learning outcomes from more effective content recommendation
 - Methods for adjudicating models of learning
 - New types of interactive applications
 - "What should I read to help understand this material"

Objective 1: Efficient exploration of a new topic



Intrinsic diversity: Helping users learn about multiple aspects of a single topic

[K. Raman, P. Bennett, Collins-Thompson SIGIR 2013]

- Diversify across aspects of a single intent
e.g.: *snow leopards*:
 - *Where do snow leopards live*
 - *Snow leopards lifespan*
 - *Snow leopard population*
 - *Snow leopards in captivity*
- Currently, users need multiple searches to find different aspects of their need.
- Can we help them find this information faster?



30,300,000 RESULTS Any time ▾

[Kelly Clarkson keeps national anthem Super Bowl performance ...](#)

[Entertainment Weekly Online](#) ▾

Feb 05, 2012 - There were no flubbed lyrics, and no screeching, when **Kelly Clarkson** sang the national anthem at **Super Bowl XLVI** on Sunday. In fact, **Clarkson** kept ...

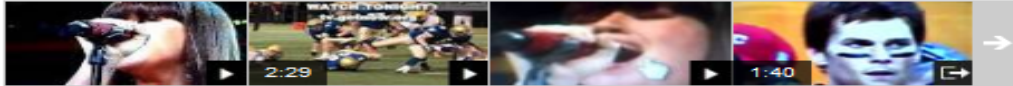
[Kelly Clarkson's Super Bowl Performance Truly Idol-Worship ...](#)

[YAHOO!](#) ▾

photo: Ezra Shaw/ Getty ImagesLast year at the **Super Bowl**, Christina Aguilera's national anthem **performance** was a disaster on an international scale. And this year ...

[Videos of kelly clarkson super bowl performance](#)

[bing.com/Videos](#)



[Kelly Clarkson National Anthem YouTube](#)

[Kelly Clarkson Super Bowl 201... Dailymotion](#)

[Kelly Clarkson's Super Bowl YouTube](#)

[Kelly Clarkson Canta el Himno metatube](#)

[Kelly Clarkson's Super Bowl Performance Truly Idol-Worship ...](#)

[omg! on Yahoo!](#) ▾

Last year at the **Super Bowl**, Christina Aguilera's national anthem **performance** was a disaster on an international scale. But this year, Xtina's new "The Voice ...

[Super Bowl | The Music Mix | EW.com](#)

[music-mix.ew.com/category/super-bowl](#) ▾

Kelly Clarkson keeps it quick and simple for her **Super Bowl performance** of the national anthem ... no screeching, when **Kelly Clarkson** sang the national anthem at **Super Bowl** ...

[Images of kelly clarkson super bowl performance](#)

[bing.com/images](#)



[News about kelly clarkson super bowl performance](#)

[bing.com/news](#)



[Kelly Clarkson Belts Out 'Dark Side' + 'Stronger' at 2012 MuchMusic Video Awards](#)

Kelly Clarkson may have lighter hair now, but she still has a 'Dark Side' — and she brought it, as well as 'Stronger,' to the 2012 MuchMusic Video Awards... [Popcrush](#) · 9 days ago

[The Super Bowl Is Now Everyone's Game](#) [Huffington Post](#)

[Duets' Goes to the Movies: TV Recap](#) [Wall Street Journal](#)

[Super Bowl 2012: Kelly Clarkson Nails National Anthem \(Video ...](#)

[www.hollywoodreporter.com/earshot/super-bowl-2012-kelly-clarkson...](#) ▾

Super Bowl 2012: Kelly Clarkson Nails National Anthem (Video) The Grammy-winning diva performs a ... Prior to her **Super Bowl performance**, country music power ...

RELATED SEARCHES

- [Kelly Clarkson Super Bowl 2012](#)
- [Kelly Clarkson Super Bowl Time](#)
- [Kelly Clarkson Super Bowl Length](#)
- [Kelly Clarkson Super Bowl Outfit](#)
- [Kelly Clarkson Singing Super Bowl](#)
- [Kelly Clarkson National Anthem Video](#)
- [Clarkson National Anthem](#)
- [Super Bowl Kelley Clarkson](#)



30,300,000 RESULTS Any time

Kelly Clarkson keeps national anthem Super Bowl performance ...

Entertainment Weekly Online

Feb 05, 2012 - There were no flubbed lyrics, and no screeching, when **Kelly Clarkson** sang the national anthem at **Super Bowl XLVI** on Sunday. In fact, **Clarkson** kept ...

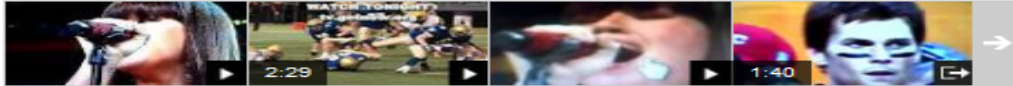
Kelly Clarkson's Super Bowl Performance Truly Idol-Worship ...

YAHOO!

photo: Ezra Shaw/ Getty ImagesLast year at the **Super Bowl**, Christina Aguilera's national anthem **performance** was a disaster on an international scale. And this year ...

Videos of kelly clarkson super bowl performance

bing.com/videos



Kelly Clarkson National Anthem YouTube

Kelly Clarkson Super Bowl 201... Dailymotion

Kelly Clarkson's Super Bowl YouTube

Kelly Clarkson Canta el Himno metatube

RELATED SEARCHES

Aspects

Kelly Clarkson Facts - I AM FAN! Alicia Keys · Jay-Z · Lil ...

www.iamfan.com/~kelly_clarkson/kelly-clarkson-facts.htm

Kelly Clarkson trivia - interesting facts about Kelly Clarkson

Facts About Kelly Clarkson

List of awards and nominations received by Kelly Clarkson ...

en.wikipedia.org/wiki/List_of_Kelly_Clarkson_awards

This is a list of **awards** that American singer-songwriter **Kelly Clarkson** has received throughout her career, which started following her coronation as the first ...

kelly clarkson awards

Images of kelly clarkson super bowl performance

bing.com/images



News about kelly clarkson super bowl performance

bing.com/news



Kelly Clarkson Belts Out 'Dark Side' + 'Stronger' at 2012 MuchMusic Video Awards

Kelly Clarkson may have lighter hair now, but she still has a 'Dark Side' — and she brought it, as well as 'Stronger,' to the 2012 MuchMusic Video Awards... Popcrush · 9 days ago

The Super Bowl Is Now Everyone's Game Huffington Post

Duets' Goes to the Movies: TV Recap Wall Street Journal

Kelly Clarkson makes her MMVA return with a flawless performance

blog.muchmusic.com/kelly-clarkson...return-with-a-flawless-performance

During last night's MMVAs, **Kelly Clarkson** blew the crowd away with her double **performance** of **Dark Side** and **Stronger**. Performing her motivational tracks from

kelly clarkson performance

kelly clarkson super bowl performance



30,300,000 RESULTS Any time

RELATED SEARCHES

Kelly Clarkson keeps national anthem Super Bowl performance ...

Entertainment Weekly Online

Feb 05, 2012 ... Kelly Clarkson sang

Kell

YAHOO

photo

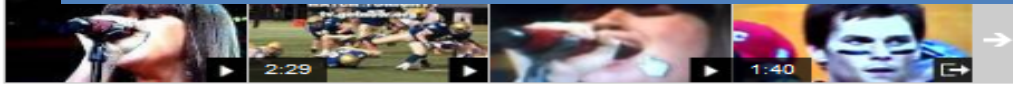
natio

...

Vide

bing.

1. Documents should be relevant to query q .
2. Document d_i should be relevant to associated aspect q_i .
3. Aspects should be relevant to ID task initiated by q .
4. Aspects should be diverse.



Kelly Clarkson National Anthem YouTube

Kelly Clarkson Super Bowl 201... Dailymotion

Kelly Clarkson's Super Bowl YouTube

Kelly Clarkson Canta el Himno metatube

Aspects

Kelly Clarkson Facts - I AM FAN! Alicia Keys · Jay-Z · Lil ...

www.iamfan.com/~kelly_clarkson/kelly-clarkson-facts.htm

Kelly Clarkson trivia - interesting facts about Kelly Clarkson

Facts About Kelly Clarkson

List of awards and nominations received by Kelly Clarkson ...

en.wikipedia.org/wiki/List_of_Kelly_Clarkson_awards

This is a list of **awards** that American singer-songwriter Kelly Clarkson has received throughout her career, which started following her coronation as the first...

kelly clarkson awards

Images of kelly clarkson super bowl performance

bing.com/images



News about kelly clarkson super bowl performance

bing.com/news



Kelly Clarkson Belts Out 'Dark Side' + 'Stronger' at 2012 MuchMusic Video Awards

Kelly Clarkson may have lighter hair now, but she still has a 'Dark Side' — and she brought it, as well as 'Stronger,' to the 2012 MuchMusic Video Awards... Popcrush · 9 days ago

The Super Bowl Is Now Everyone's Game Huffington Post

Duets' Goes to the Movies: TV Recap Wall Street Journal

4

Kelly Clarkson makes her MMVA return with a flawless performance

blog.muchmusic.com/kelly-clarkson...return-with-a-flawless-performance

During last night's MMVAs, **Kelly Clarkson** blew the crowd away with her double performance of Dark Side and Stronger. Performing her motivational tracks from

kelly clarkson performance

1

3

2

Ranking Algorithm

- Given query q :

Produce ranking $d_1, d_2..$ (with associated aspects $q_1, q_2..$)

1. Documents should be relevant to query q .
2. Document d_i should be relevant to associated aspect q_i .
3. Aspects should be relevant to ID task initiated by q .
4. Aspects should be diverse.

- Objective :

$$\operatorname{argmax}_{(d_1, q_1) \dots (d_n, q_n)} \sum_{i=1}^n \gamma_i \cdot R(d_i|q) \cdot R(d_i|q_i) \cdot e^{\beta \operatorname{Div}(q_i, Q)}$$

Breaking Down the Objective - 1

$$\operatorname{argmax}_{(d_1, q_1) \dots (d_n, q_n)} \sum_{i=1}^n \gamma_i \cdot R(d_i|q) \cdot R(d_i|q_i) \cdot e^{\beta \operatorname{Div}(q_i, \mathcal{Q})}$$

- Document relevance to query.
- Trained Relevance model (with 21 simple features) using Boosted Trees.

Breaking Down the Objective - 2

$$\operatorname{argmax}_{(d_1, q_1) \dots (d_n, q_n)} \sum_{i=1}^n \gamma_i \cdot R(d_i|q) \cdot R(d_i|q_i) \cdot e^{\beta \operatorname{Div}(q_i, \mathcal{Q})}$$

- Document relevance to aspect.
 - *Represents/Summarizes* the aspect.
- Can be estimated with same relevance model R

Breaking Down the Objective - 3

$\operatorname{argmax}_{(d_1, q_1)}$

Algorithm 1 Greedy-DynRR($\beta, \lambda, P(\cdot|\cdot), q$)

```
1:  $(y_D, y_Q) \leftarrow \phi$ 
2: for all  $q' \in \operatorname{Rel}Q(q)$  do
3:    $\operatorname{Next}(q') \leftarrow$  Document Ranking by  $R(\cdot|q) \cdot R(\cdot|q')$ .
4: for  $i = 1 \rightarrow n$  do
5:    $\operatorname{best}U \leftarrow -\infty$ 
6:   for all  $q' \in \operatorname{Rel}Q(q) / y_Q$  do
7:      $d' \leftarrow \operatorname{Top}(\operatorname{Next}(q') / y_D)$ 
8:      $v \leftarrow R(d'|q) \cdot R(d'|q') \cdot e^{\beta \cdot \operatorname{Div}(q', y_Q)}$ 
9:     if  $v > \operatorname{best}U$  then
10:       $\operatorname{best}U \leftarrow v$ 
11:       $\operatorname{best}Q \leftarrow q'$ 
12:       $\operatorname{best}D \leftarrow d'$ 
13:    $(y_D, y_Q) \leftarrow (y_D, y_Q) \oplus (\operatorname{best}D, \operatorname{best}Q)$ 
14: return  $y$ 
```

$e^{\beta \operatorname{Div}(q_i, Q)}$

- Aspect Divergence
- Maximize coherence

D

$\operatorname{ip}(q_j))$.

- Submodular Objective:
 - Can be optimized using an efficient greedy algorithm.

Sample re-ranking results

<u>REGULAR</u>	<u>RERANKED</u>
<p>Kelly Clarkson - Wikipedia, the free encyclopedia en.wikipedia.org/wiki/Kelly_Clarkson</p> <p>Kelly Brianne Blackstock (née Clarkson; born April 24, 1982) is an American singer-songwriter and occasional actress. She rose to fame after winning the first season of the reality singing competition American Idol in 2002, she rose to fame after winning the ... Life and career · Artistry · Personal life · Discography · Touring · Filmography</p> <p>Homepage Kelly Clarkson www.kellyclarkson.com/us/home</p> <p>Okay my baby is moving like crazy in my belly th...</p> <p>Kelly Clarkson (kelly clarkson) on Twitter twitter.com/kelly_clarkson</p> <p>9,488,708 Followers</p> <p>The latest from Kelly Clarkson</p> <p>Kelly Clarkson New Music www.mtv.com/artists/kelly-clarkson</p> <p>Kelly Clarkson new music, concerts, photos, videos, and more. Follow Kelly Clarkson's Twitter and Facebook.</p> <p>Music · Photos · News · Discography</p> <p>Kelly Clarkson - IMDb www.imdb.com/name/nm1225628</p> <p>Kelly Clarkson, Soundtrack: Love Actually. Kelly Clarkson was born on April 24, 1982 in Fort Worth, Texas, USA as Kelly Brianne Clarkson. She has been married to ... Born Apr 24, 1982 · News · 185 photos · Biography · Awards · Films</p> <p>Kelly Clarkson - Free listening, videos, concerts, stats and more at Last.fm www.last.fm/music/Kelly+Clarkson</p> <p>Kelly Brianne Clarkson (born April 24, 1982 in Fort Worth, Texas, USA) is a Grammy-winning American singer-songwriter and occasional actress. Clarkson recorded her debut album, Thank You, Next, in 2013. Clarkson is red hot while hitting the stage for a performance of her new single "Tie It Up" at the 2013 CMA Music Festival on Saturday (June 8 ...)</p> <p>Kelly Clarkson - People.com : Celebrity News, Celebrity Photos, and More www.people.com/people/kelly_clarkson</p> <p>Get everything Kelly Clarkson straight from America's #1 celebrity brand, PEOPLE. The latest Kelly Clarkson news, a full collection of photos, fun facts and her ... News · Biography · Photos</p> <p>Kelly Clarkson Biography - Facts, Birthday, Life Story ... www.biography.com › People</p> <p>Kelly Clarkson has never looked back after her show-stopping performance on American Idol . Learn more about the woman behind the hits, at Biography.com.</p>	<p>Kelly Clarkson - Wikipedia, the free encyclopedia en.wikipedia.org/wiki/Kelly_Clarkson</p> <p>Kelly Brianne Blackstock (née Clarkson; born April 24, 1982) is an American singer-songwriter and occasional actress. She rose to fame after winning the first season of the reality singing competition American Idol in 2002, she rose to fame after winning the ... Life and career · Artistry · Personal life · Discography · Touring · Filmography</p> <p>facts about kelly clarkson</p> <p>kelly clarkson awards</p> <p>kelly clarkson social media</p> <p>kelly clarkson performance</p> <p>kelly clarkson movies</p>

Our SIGIR 2013 paper showed that this algorithm does enable more efficient exploration... but what about learning outcomes?

Assessing learning via search

[Collins-Thompson, Rieh, Haynes, Syed. ACM CHIIR 2016]

RQ1: What kinds of measures and indicators can be developed to assess learning experiences and outcomes in interactive search systems?

RQ2: What query strategies best support human learning experiences and outcomes? (e.g. Intrinsic diversity helps?)

RQ3: To what extent is searchers' search behavior correlated with learning experiences and outcomes?

Study workflow

Assign user to one of three query conditions

Task completion (X 2)

Instructions



Read task description



Pre-search questions



Query condition instructions



Search

- Select an initial query
- Issue more queries (in condition)
- Browse results, save documents
- Completion button when done



Post-search questions



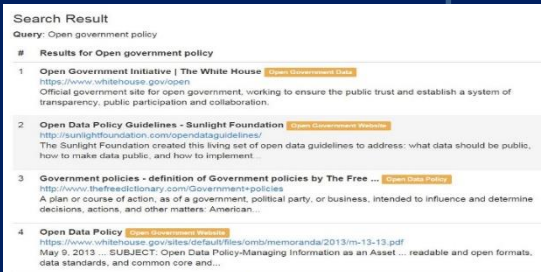
Background questionnaire

Search server

uFindIt-based
Amazon EC2



Google
Custom
Search API



Pre-search questionnaire

4 questions

P1: Subjects' prior knowledge level (1-5)

P2: Interest in the topic (1-5)

P3: Perceived difficulty of searching (1-5)

P4: "Please write what you know about this topic"



Post-search questionnaire

23 questions (Q1-Q23)

Question Group	Response Type	# of questions
Cognitive learning assessment (Q1-Q6)	Written	6
Search exploration	1-5	6
User experience with the system	1-5	4
Learner interest & motivation	1-5	5
Perceived learning and search success	0-100	2

See the paper for a complete inventory

Users were assigned to one of three "query strategy" conditions

- Single query (SQ, N=12)
 - Select initial query
 - Use initial results for the rest of the session
- Multiple query (MQ, N=15)
 - Same user interface as SQ condition
 - Select initial query
 - May issue new queries
- Multiple query + Intrinsic diversity (ID, N=15)
 - Uses ID presentation to MQ condition
 - Select initial query
 - May issue new queries OR use ID suggestions

Search Result

Query: Open government policy

Results for Open government policy

1	Open Government Policy The White House http://www.whitehouse.gov/open/about/policy The President's Memorandum on Transparency and Open Government The President's first executive action, the Open Government Memorandum calls for more transparent ...
2	Open Government Initiative The White House http://www.whitehouse.gov/Open/ Official government site for open government, working to ensure the public trust and establish a system of transparency, public participation and collaboration.
3	Open Government and Democracy Policy - Green Party of ... https://www.greens.org.nz/policy/open-government-and-democracy-policy Read the Open Government Policy Summary Download the Open Government Policy as PDF. Spokesperson James Shaw MP. Updated 16-Jan-2015. Vision. The Green Party ...
4	The Government of India's new open source policy ... http://opensesource.com/government/158/india-adopts-open-source-policy The Government of India (GOI) has adopted a comprehensive and supportive open source policy. It builds on their earlier efforts to adopt open standards for procurement.

Search Result

Query: Open government policy

Results for Open government policy

1	Open Government Policy The White House http://www.whitehouse.gov/open/about/policy The President's Memorandum on Transparency and Open Government The President's first executive action, the Open Government Memorandum calls for more transparent ...
2	Open Government Initiative The White House http://www.whitehouse.gov/Open/ Official government site for open government, working to ensure the public trust and establish a system of transparency, public participation and collaboration.
3	Open Government and Democracy Policy - Green Party of ... https://www.greens.org.nz/policy/open-government-and-democracy-policy Read the Open Government Policy Summary Download the Open Government Policy as PDF. Spokesperson James Shaw MP. Updated 16-Jan-2015. Vision. The Green Party ...
4	The Government of India's new open source policy ... http://opensesource.com/government/158/india-adopts-open-source-policy The Government of India (GOI) has adopted a comprehensive and supportive open source policy. It builds on their earlier efforts to adopt open standards for procurement.

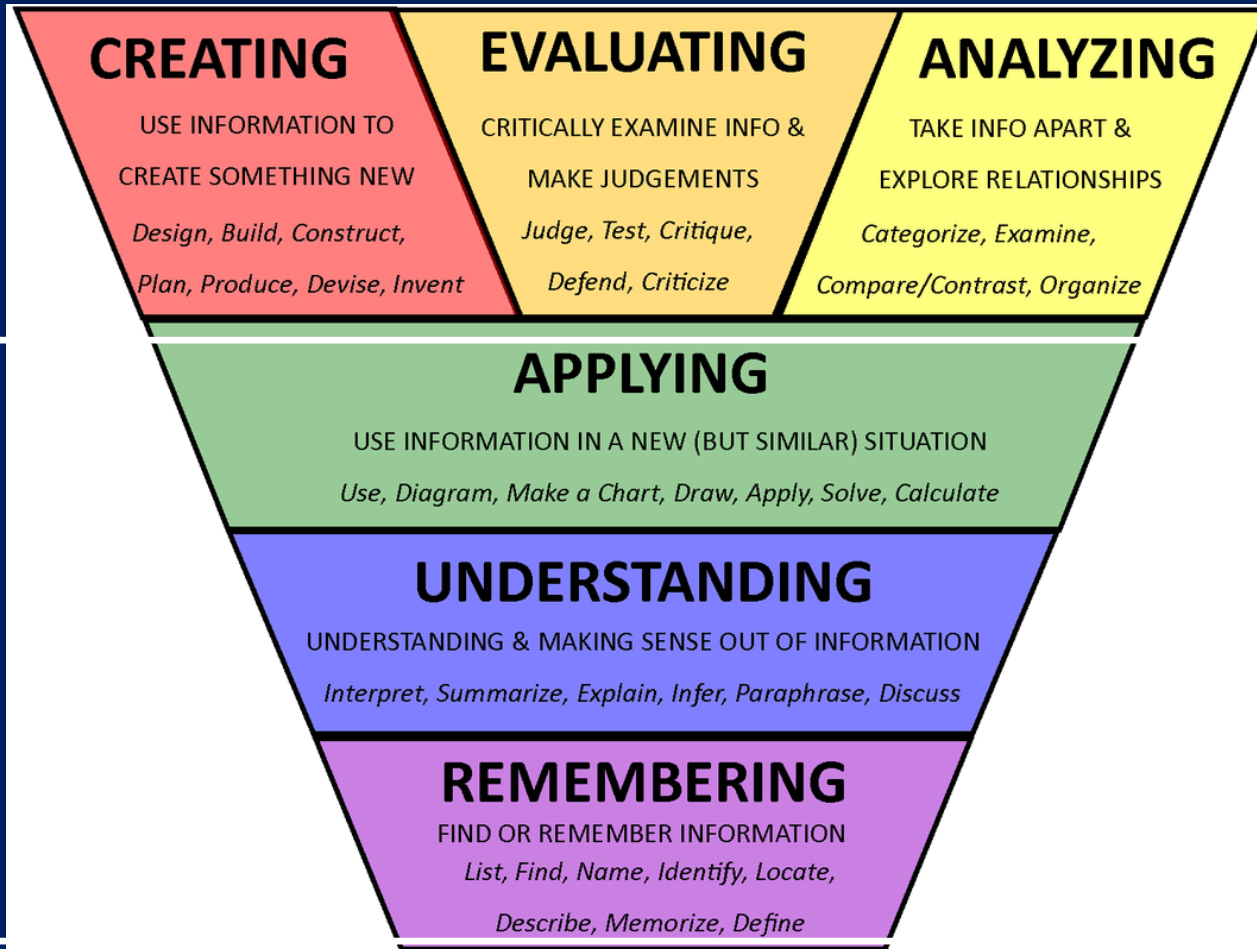
Search Result

Query: Open government policy

Results for Open government policy

1	Open Government Initiative The White House Open Government Website https://www.whitehouse.gov/open Official government site for open government, working to ensure the public trust and establish a system of transparency, public participation and collaboration.
2	Open Data Policy Guidelines - Sunlight Foundation Open Government Website http://sunlightfoundation.com/opendataguidelines/ The Sunlight Foundation created this living set of open data guidelines to address: what data should be public, how to make data public, and how to implement...
3	Government policies - definition of Government policies by The Free ... Open Data Policy http://www.thefreedictionary.com/Government+policies A plan or course of action, as of a government, political party, or business, intended to influence and determine decisions, actions, and other matters: American...
4	Open Data Policy Open Government Website https://www.whitehouse.gov/sites/default/files/omb/memoranda/2013/m-13-13.pdf May 9, 2013 . . . SUBJECT: Open Data Policy-Managing Information as an Asset ... readable and open formats, data standards, and common core and...

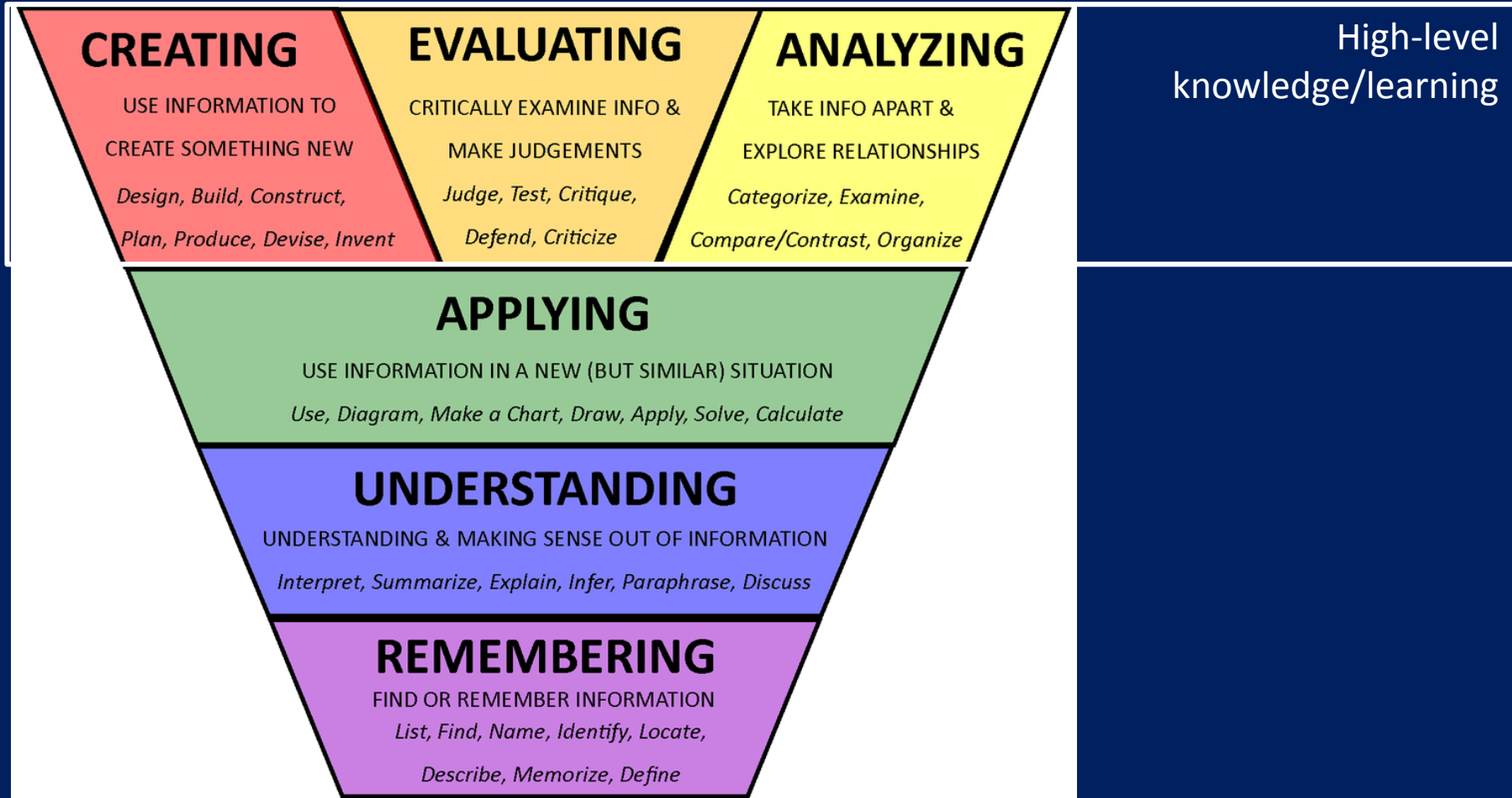
Our model of learning during search was based on Bloom's Taxonomy



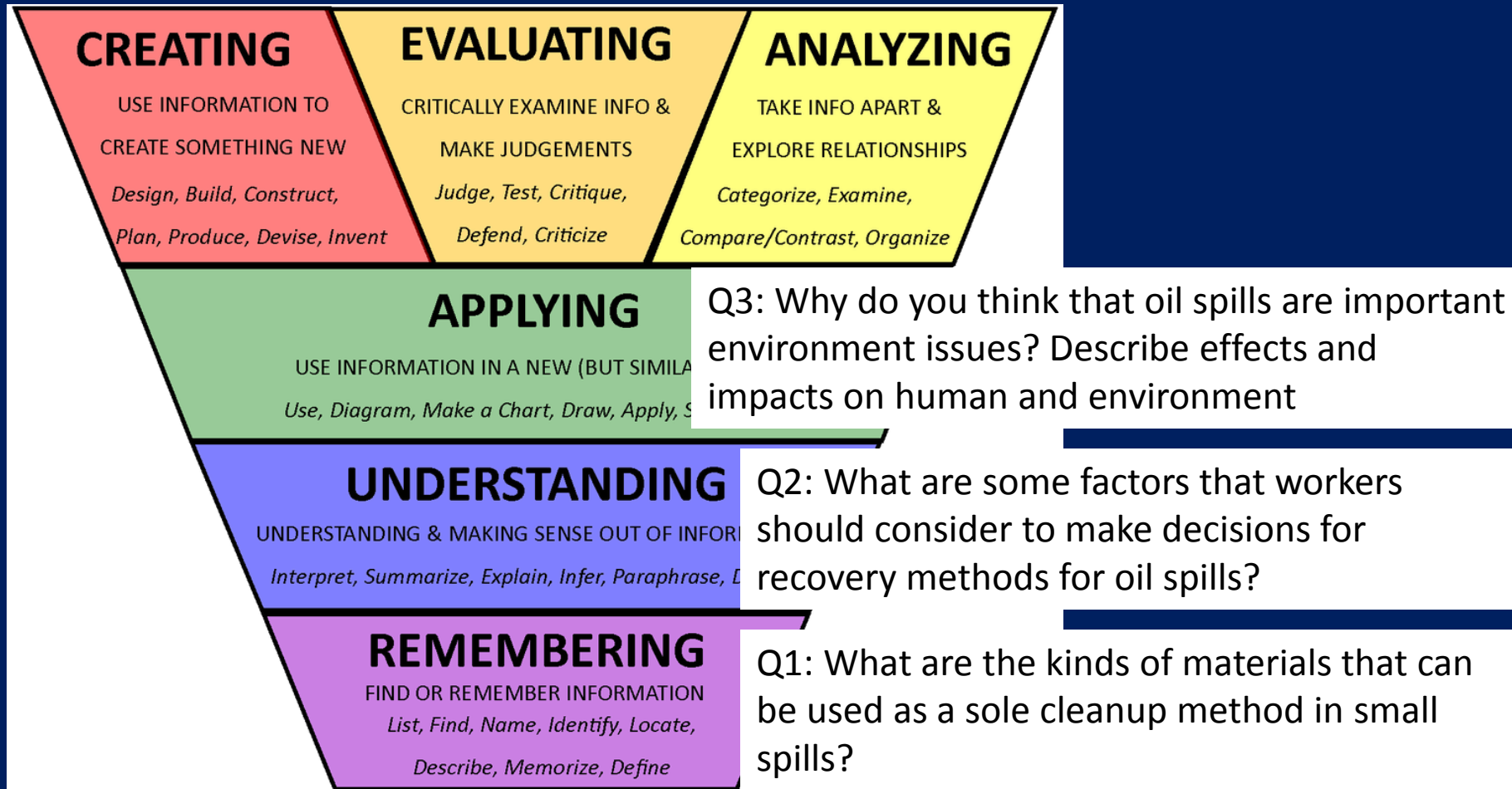
Low-level
knowledge/learning

Anderson, L. W. and Krathwohl, D. R. 2001. A Taxonomy for Learning, Teaching, and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives. Longman, New York.

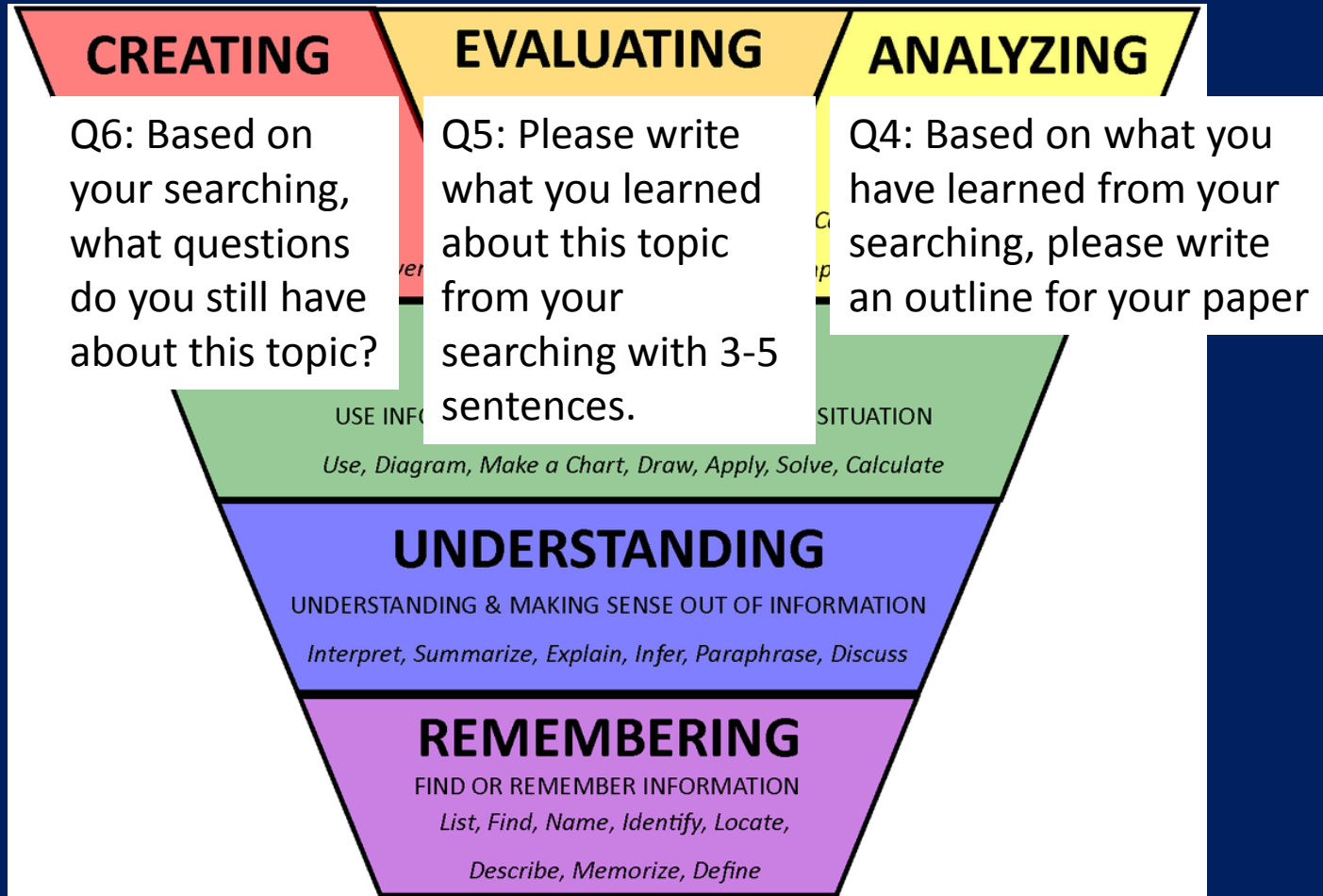
Our model of learning during search was based on Bloom's Taxonomy



Lower-level cognitive learning questions (Q1-Q3) were specific to each topic



Higher-level cognitive learning questions (Q4-Q6) were the same for both topics



Example coding for 'oil spill' Q1:

What are the kinds of materials that can be used as a sole cleanup method in small spills?

Cleanup from an oil spill is hard and depends upon many factors, including the type of oil, the temperature of the water, and the types of shoreline. Methods for cleanup include oil-eating bacteria, dispersants, and skimming.

	Oil Spill Coding Scheme	Score
Q1 Remember	1. List general terminology	1
	2. Recall elements/kinds of materials/processes	1
	3. Describe cleanup methods and techniques	1
	4. State appropriate procedures for particular circumstances or conditions	0
	5. Demonstrate conceptual knowledge of principles or concepts	1
	6. Awareness of knowledge of one's own cognition	0
	7. Show knowledge of classifications and categories	1

Raw assessment score for Q1: 5/7

For both search tasks, mean inter-coder reliability across Q1–Q6 (Holsti)

Oil Spill task: 0.914

Open Data task: 0.797

Writing-based Assessment Measure #1: Cognitive Score

Oil Spill Coding Scheme	Score
Q1. What are the kinds of materials that can be used as a sole cleanup method in small spills?	4
Q2. What are some factors that they should consider to make decisions for recovery methods?	2
Q3. Why do you think that oil spills are important environment issues? Describe its effects and impacts on human and environment	5
Q4. Based on what you have learned from your searching, please write an outline for your paper.	3
Q5. Please write what you learned about this topic from your searching with 3-5 sentences.	2
Q6. Based on your searching, what questions do you still have about this topic?	4

Lower level

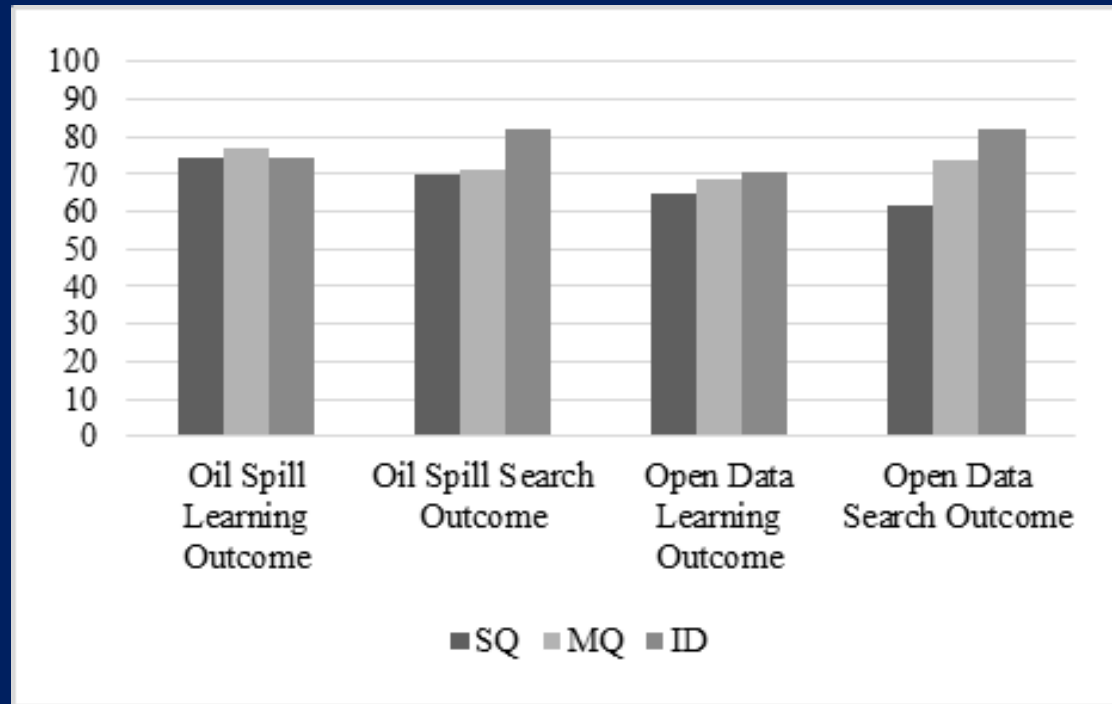
Higher level

Lower-level cognitive score: $Q1+Q2+Q3 = 4+2+5 = 11$

Higher-level cognitive score: $Q4+Q5+Q6 = 3+2+4 = 9$

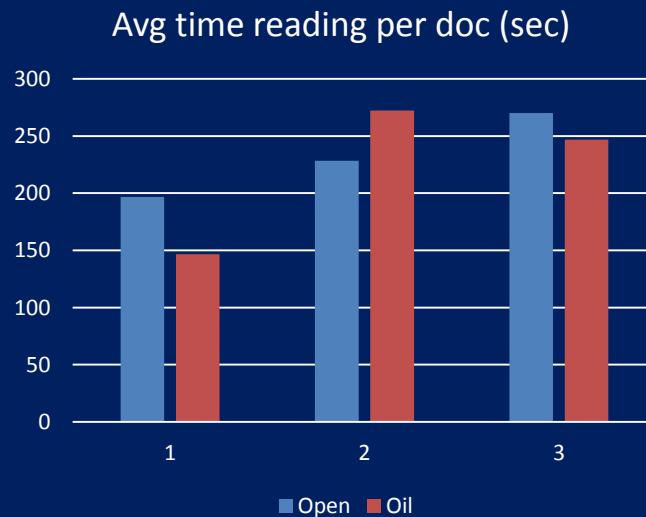
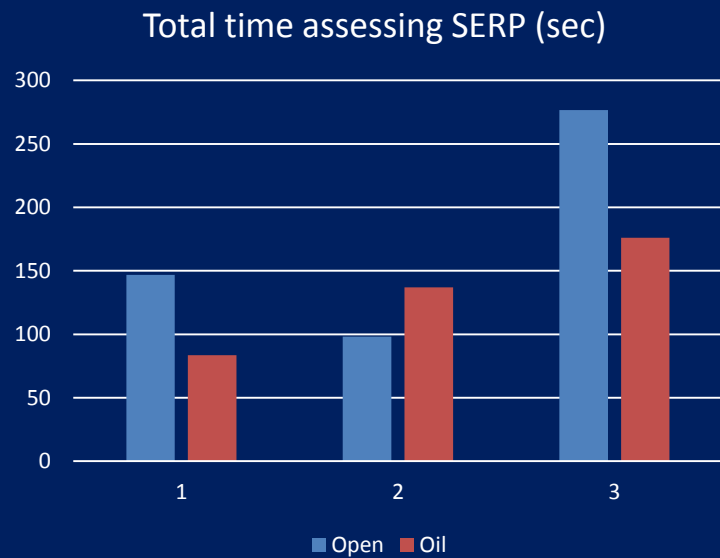
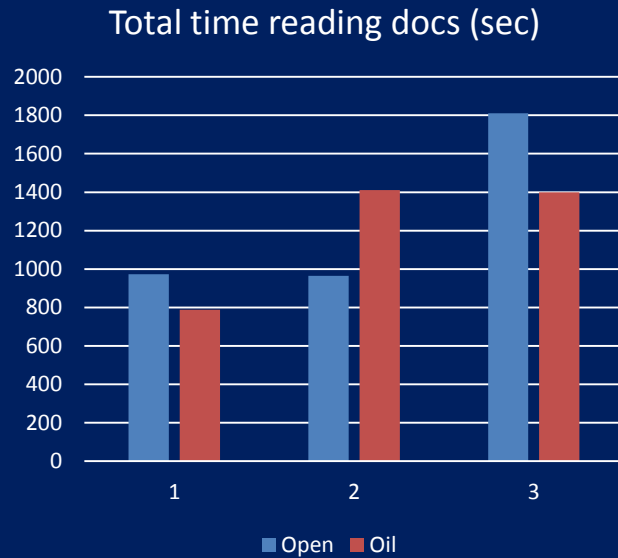
Overall cognitive score: $Q1 + \dots + Q6 = 20$

Users' self-reported perceived learning and search success by query condition and task (Q22, Q23)



- MQ > SQ across both tasks
- Intrinsic Diversity condition (Open Data task): higher search outcome ($p < 0.02$)

RQ2: ID users spent more time reading docs and assessing results



Writing-based assessment method # 2:

Knowledge level gain estimates the change in level of knowledge from prior (P4) to current (Q5)

Response to P4: Please write what you know on this topic...

I guess most oil spill cleanup is done using some kind of skimming device or vacuum cleaner. I don't know much more than that.

What is the highest level of knowledge exhibited in this writing?

0 = no knowledge 1 = factual knowledge 2 = conceptual knowledge

Response to Q5: Please write what you learned about this topic...

I discovered that cleanup from an oil spill is hard and depends upon many factors, including the type of oil, the temperature of the water, and the types of shoreline. This can determine which is the best cleanup method, which might include oil-eating bacteria, dispersants, and skimming.

What is the highest level of knowledge exhibited in this writing?

0 = no knowledge 1 = factual knowledge 2 = conceptual knowledge

Knowledge level gain: $2 - 1 = 1$

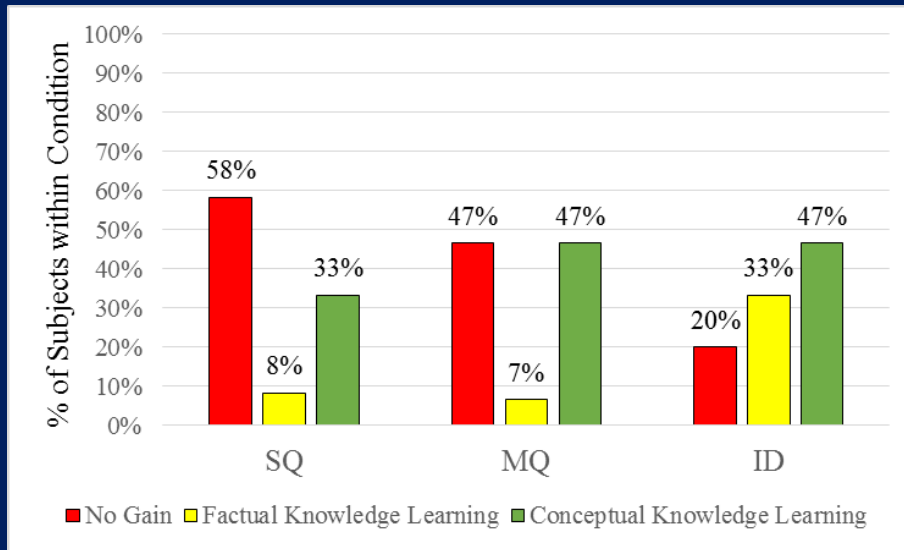
RQ1: Indicators of learning Knowledge Level Gain

Knowledge level gain	# Participants
+2 (none → conceptual)	10 (24%)
+1 (none → factual, factual → conceptual)	15 (35%)
+0	16 (38%)
-1 (conceptual → factual)	1 (3%)

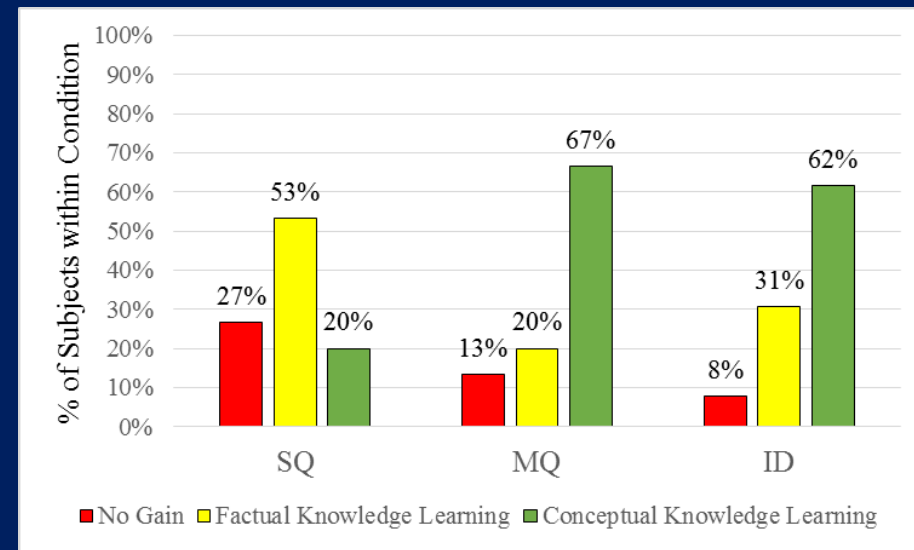
- We found a strong positive correlation between perceived learning outcomes and actual knowledge level gain in the ID condition, for both the Open Data ($r=0.69$) and Oil Spill ($r=0.64$) tasks.

RQ2: Knowledge Level Gain by Condition

Oil Spill task



Open Data task



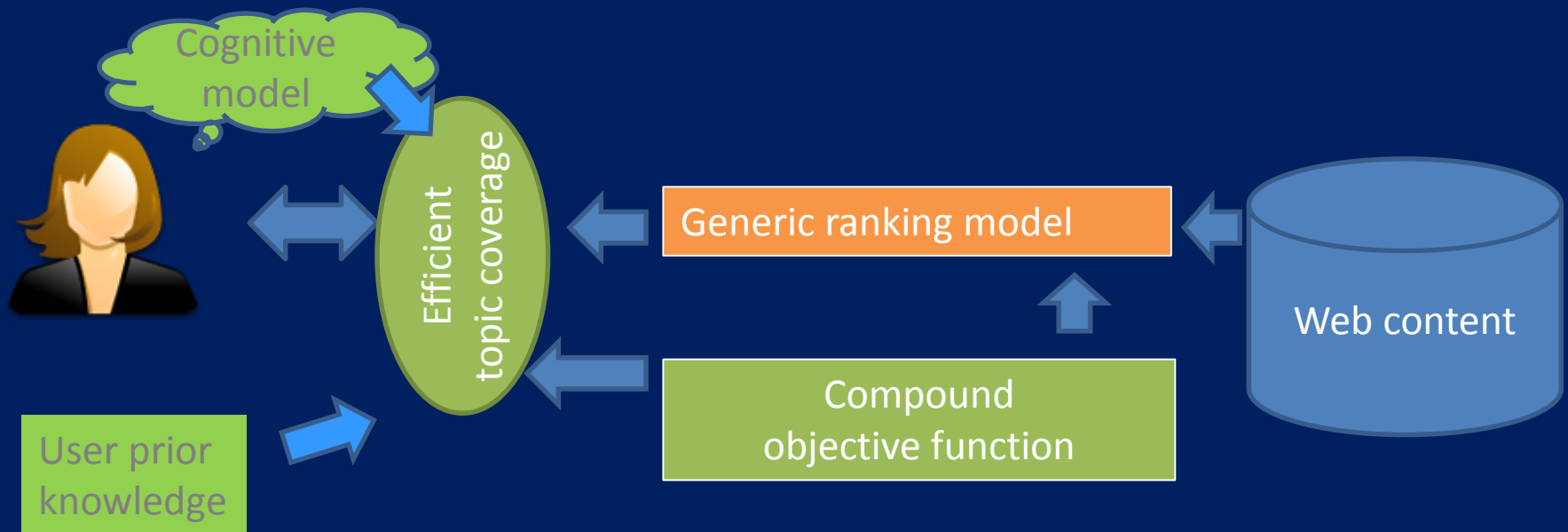
- Subjects exhibited different knowledge level gains depending on query condition.
- The ID condition gave the highest % users with combined factual + cognitive knowledge gain for both tasks
- Absolute gains in knowledge level in the ID condition were statistically significant for the Oil Spill task (SQ=0.41, MQ=0.73, ID=1.20: $p=.031$)

Key study contributions

- RQ1: Searchers' perceived learning outcomes closely matched actual learning outcomes.
- RQ2: The ID query condition may give an advantage for learning compared to SQ and MQ conditions:
e.g. ID was associated with the highest % of users achieving factual and conceptual knowledge gains.
- RQ3: Time spent per document while searching was highly and consistently correlated with higher-level cognitive learning scores.
- A detailed coding scheme for assessing learning from writing
- Exploration of implicit/explicit indicators for assessing learning in search
- Our study provides further insight into:
 - How learning may be assessed effectively during web search.
 - Which aspects of search interaction may be most effective for supporting superior learning outcomes.

Objective 2: Adding a vocabulary learning goal to topic exploration

[Syed & Collins-Thompson, in submission]



Topic	Keywords to learn
Igneous rock	magma, pyroclastic, extrusion, ...
Cytoplasm	cytoplasm, organelle, cell membrane, ...

Optimization: Compute an intrinsically diverse ranking that also achieves the desired keyword coverage

- We add a new term to the intrinsic diversity objective function

$$\arg \max_{\mathcal{D}} \sum_{i=1}^{|\mathcal{D}|} Rel(d_i|q) \cdot Rel(d_i|q_i) \cdot e^{\beta \eta_i} \cdot e^{\alpha \epsilon_i}$$

$$\epsilon_i = \frac{1}{|d_i|} \sum_{j=1}^N \begin{cases} C_{ij} & C_{ij} + C_{\mathcal{D}j} \leq s_j \\ \max(0, s_j - C_{\mathcal{D}j}) & \text{else} \end{cases}$$

- Parameter α effectively controls the keyword density of documents in the ranking, with the goal of covering all the 'target' keywords.

Crowdsourced pilot study

Task completion

Instructions



Assign user to a topic
& α condition

n = 175 per condition
n = 700 users total

Read task description



Vocabulary pre-test



Read ranked list of documents
- Browse results
- Completion button when done

Pre-computed
Intrinsically diverse
rankings

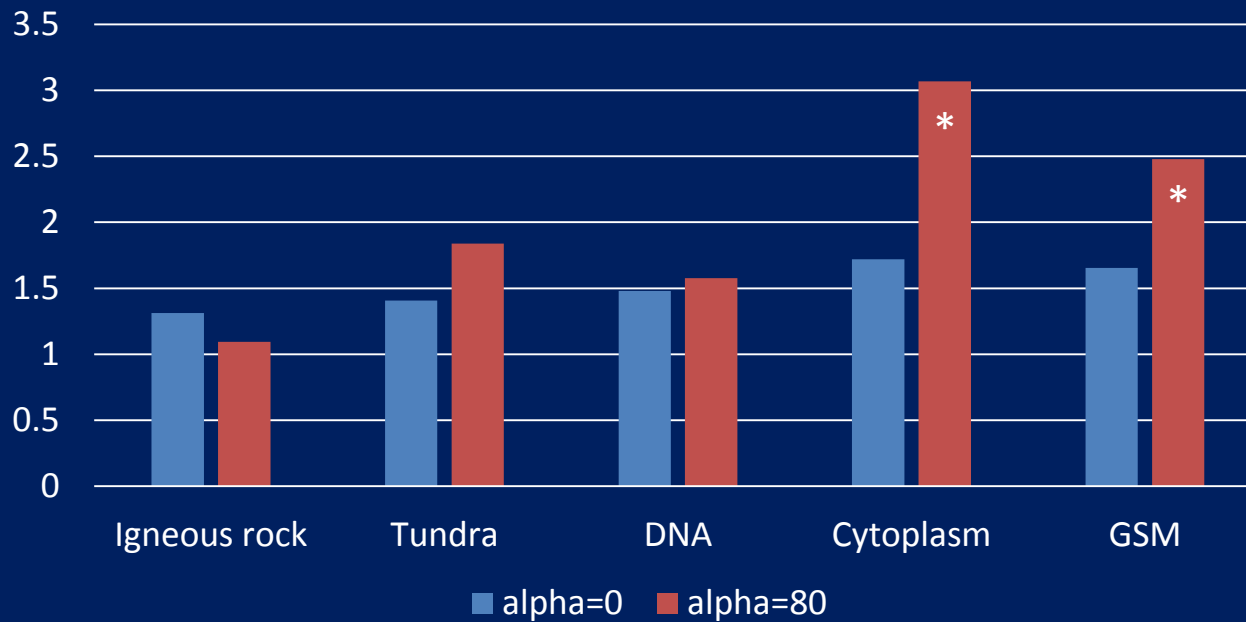


Vocabulary post-test

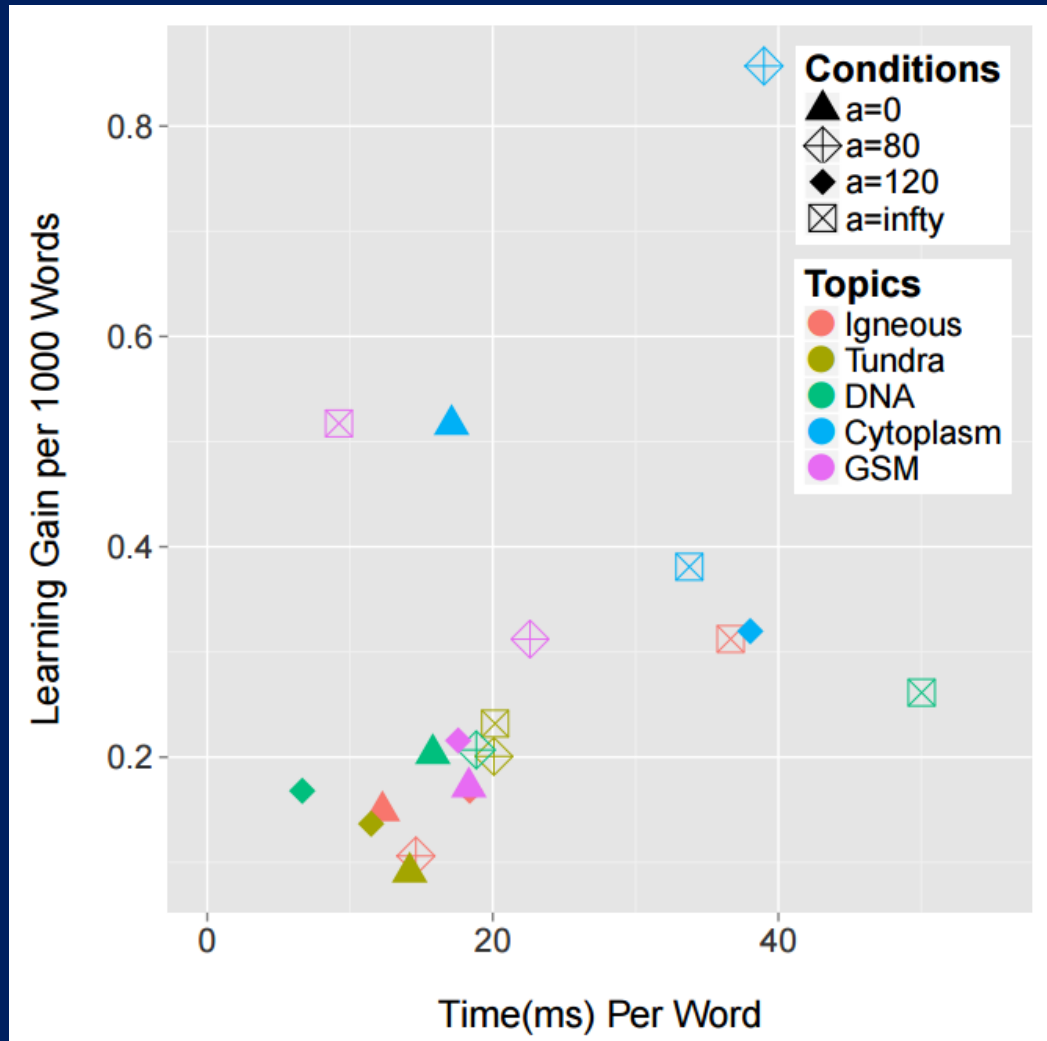
Learning objective: Maximize (Post – Pre) test score

Adding keyword coverage to the intrinsic diversity objective improved vocabulary learning outcomes for some topics

Absolute vocabulary learning gain by topic and alpha condition



Increasing α keyword coverage also led to rankings with more efficient learning gains



Reading time per word was correlated with learning gain per word ($r = 0.422$).

Documents with higher keyword density tended to be shorter with more supporting images.

Future directions: e.g. What resources should I use to help me understand this paper?

New Terms/Concepts:

- functional
- expected loss
- measure
- conditional distribution
- expected value
- independent
- identically distributed
- ...

Concept Dependencies:

measure → probability/topology
functional → linear algebra
IID → statistics

988 IEEE TRANSACTIONS ON NEURAL NETWORKS, VOL. 10, NO. 5, SEPTEMBER 1999

An Overview of Statistical Learning Theory

Vladimir N. Vapnik

Abstract—Statistical learning theory was introduced in the late 1960's. Until the 1990's it was a purely theoretical analysis of the problem of function estimation from a given collection of data. In the middle of the 1990's new types of learning algorithms (called support vector machines) based on the developed theory were proposed. This made statistical learning theory not only a tool for the theoretical analysis but also a tool for creating practical algorithms for estimating multidimensional functions. This article presents a very general overview of statistical learning theory including both theoretical and algorithmic aspects of the theory. The goal of this overview is to demonstrate how the abstract learning theory established conditions for generalization which are more general than those discussed in classical statistical paradigms and how the understanding of these conditions inspired new algorithmic approaches to function estimation problems. A more detailed overview of the theory (without proofs) can be found in Vapnik (1995). In Vapnik (1998) one can find detailed description of the theory (including proofs).

I. SETTING OF THE LEARNING PROBLEM

In this section we consider a model of the learning and show that analysis of this model can be conducted in the general statistical framework of minimizing expected loss using observed data. We show that practical problems such as pattern recognition, regression estimation, and density estimation are particular case of this general model.

A. Function Estimation Model

The model of learning from examples can be described using three components:

- 1) a generator of random vectors x , drawn independently from a fixed but unknown distribution $P(x)$;
- 2) a supervisor that returns an output vector y for every

B. Problem of Risk Minimization

In order to choose the best available approximation to the supervisor's response, one measures the loss or discrepancy $L(y, f(x, \alpha))$ between the response y of the supervisor to a given input x and the response $f(x, \alpha)$ provided by the learning machine. Consider the expected value of the loss, given by the risk functional

$$R(\alpha) = \int L(y, f(x, \alpha)) dP(x, y). \quad (2)$$

The goal is to find the function $f(x, \alpha_0)$ which minimizes the risk functional $R(\alpha)$ (over the class of functions $f(x, \alpha), \alpha \in \Lambda$) in the situation where the joint probability distribution $P(x, y)$ is unknown and the only available information is contained in the training set (1).

C. Three Main Learning Problems

This formulation of the learning problem is rather general. It encompasses many specific problems. Below we consider the main ones: the problems of pattern recognition, regression estimation, and density estimation.

The Problem of Pattern Recognition: Let the supervisor's output y take on only two values $y = \{0, 1\}$ and let $f(x, \alpha), \alpha \in \Lambda$, be a set of indicator functions (functions which take on only two values zero and one). Consider the following loss-function:

$$L(y, f(x, \alpha)) = \begin{cases} 0 & \text{if } y = f(x, \alpha) \\ 1 & \text{if } y \neq f(x, \alpha) \end{cases} \quad (3)$$

For this loss function, the functional (2) provides the probability of classification error (i.e., when the answers y given by supervisor and the answers given by indicator function

Future directions

- Optimizing long-term retention, not just immediate learning
- Modeling conceptual dependencies as constraints
- Integrate models of cognitive capacity, memory
- Applications for learning disabilities

Thanks!

Kevyn Collins-Thompson kevynct@umich.edu