

NSF WORKSHOP ON TASK-BASED SEARCH

TASK-BASED SEARCH: A SEARCH ENGINE PERSPECTIVE

#nsftasks

Susan Dumais, Microsoft Research

Outline



- Why tasks are important
- A taxonomy of Web browsing/searching tasks
- Detecting tasks
 - ▣ Within and across sessions
 - ▣ Implicit detection using behavioral interaction
- Supporting tasks
 - ▣ Ranking/recommending
 - ▣ Interaction
- Evaluating task support

Search and Context

User Context



Query Words

Query Words

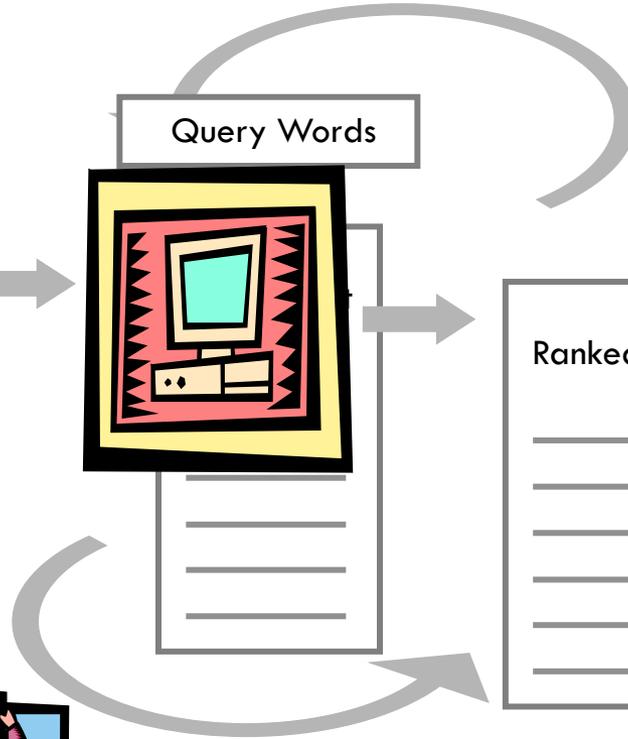
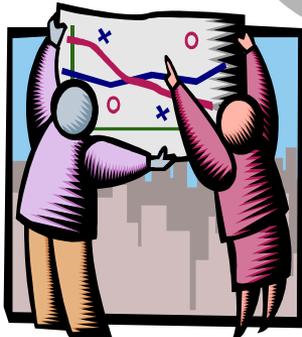


Ranked List

Document Context

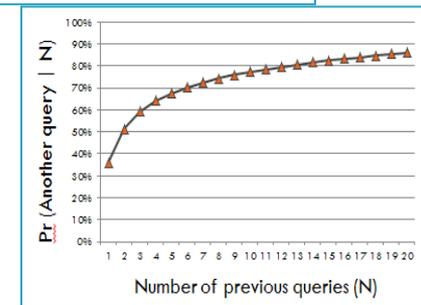
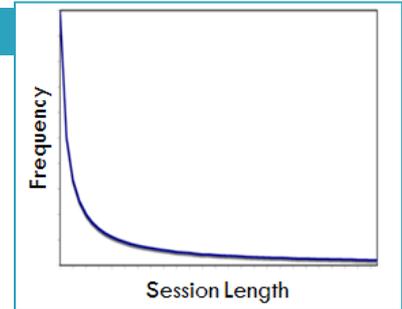


Task Context



Why “Tasks” Are Important

- Long sessions are common
 - ▣ 40% of sessions contain multiple queries
 - ▣ 40% of sessions take 3+ minutes
 - 5% of these take 30+ minutes
- Long sessions are more likely to continue
- Long sessions account for most of the search time
 - ▣ ~50% of time spent in sessions of 30+ mins
- What are people doing in these long sessions?
- How can we support them?



Tasks != Sessions

- Are interleaved with other tasks
- Extend across sessions and across devices
- Can extend over long periods of time
 - ▣ E.g., Queries related to “mortgage” over time

Time Period				
0–30 min	1–7 days	7–30d	30–90d	90–365d
mortgage	realtors	llc	kohls	patio
mortgage	owner	associates	bath	harbor
mortgage	homes	insurance	overstock	outdoor
calculator	mls	lowes	barn	replacement
mortgages	remax	notary	sears	pools
lenders	property	depot	linens	hampton
calculators	financial	savings	beyond	lawn
countrywide	appraisers	construction	kmart	enterprise
gmac	builders	condo	pottery	ymca
refinance	prudential	business	walmart	vehicle
rates	zillow	secretary	outlet	supply
interest	bankruptcy	furniture	costco	resorts
broker	real	allstate	target	lake
lending	keller	companies	pier	rv
lender	properties	contractors	bed	walgreens
payment	agreement	cost	grill	newport
loan	appraisals	reverse	kitchen	lumber

Which Tasks To Focus On

- Several factors considered when deciding which tasks to support
 - ▣ Coverage
 - What proportion of tasks are impacted?
 - ▣ Accuracy
 - How accurately can we detect a task?
 - ▣ Impact
 - What can we do about it?
 - How does it changes behavior or outcome?

A Taxonomy of Browsing Behavior

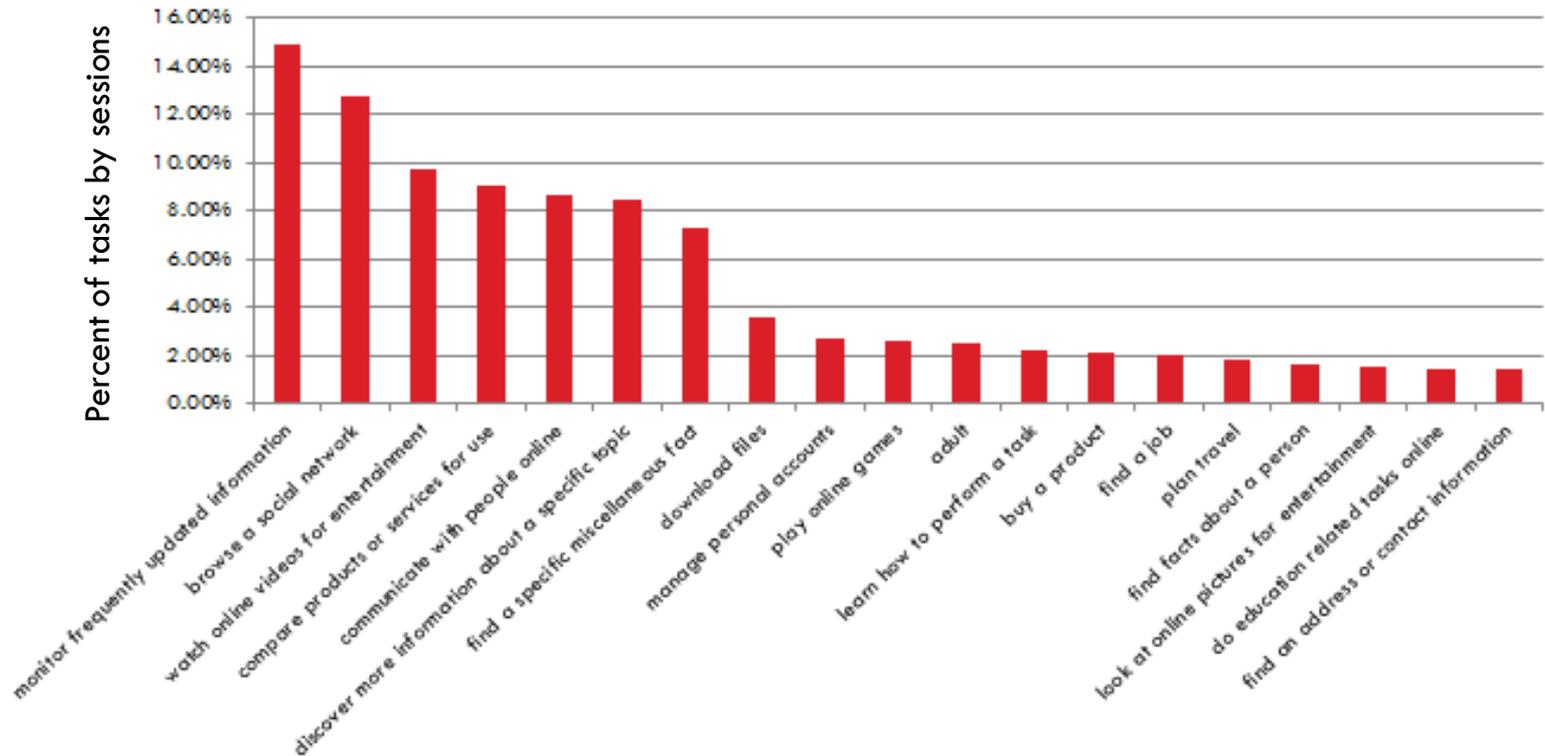
- Developed a new task taxonomy for Web browsing and query-focused browsing behavior
- Log data from: 187 users, 453 sessions, 1913 tasks
- Used iterative taxonomy development, 5 judges
- **Verb**-based, not domain- or search-activity based
 - ▣ Action-topic pairs
 - E.g., find-address, plan-trip, download-song, compare-product
- NOTE: many other methods for identifying tasks of interest including ethnography, surveys, critical incidents

Compare to Broder & Russell et al.

[Broder 2002]	[Google 2009]	Action	Topic
<i>Informational</i>	<i>Find-Simple</i>	Find: a specific miscellaneous fact Find: an address or contact information	
	<i>Find-Complex</i>	Find: a date online Find: a job Find: facts about a person Find: real estate	
	<i>Explore/Learn</i>	Browse: a social network Compare: products or services for use Discover: more information about a specific topic Discover: leisure activities Learn: how to perform a task Plan: event Plan: travel	
<i>Transactional</i>	<i>Locate/Acquire</i>	Buy: a product Download: files Plan: event Plan: travel Sell: a product	
	<i>Play</i>	Look at online pictures for entertainment Watch: online videos for entertainment Listen: to online music Play: online games	
	<i>Meta</i>		
<i>Navigational</i>	<i>Navigate</i>		
<i>Other</i>	<i>Other</i>	Communicate: with people online Do: education related tasks (online homework, etc.) Manage: personal accounts Read or write: on blogs or forums	

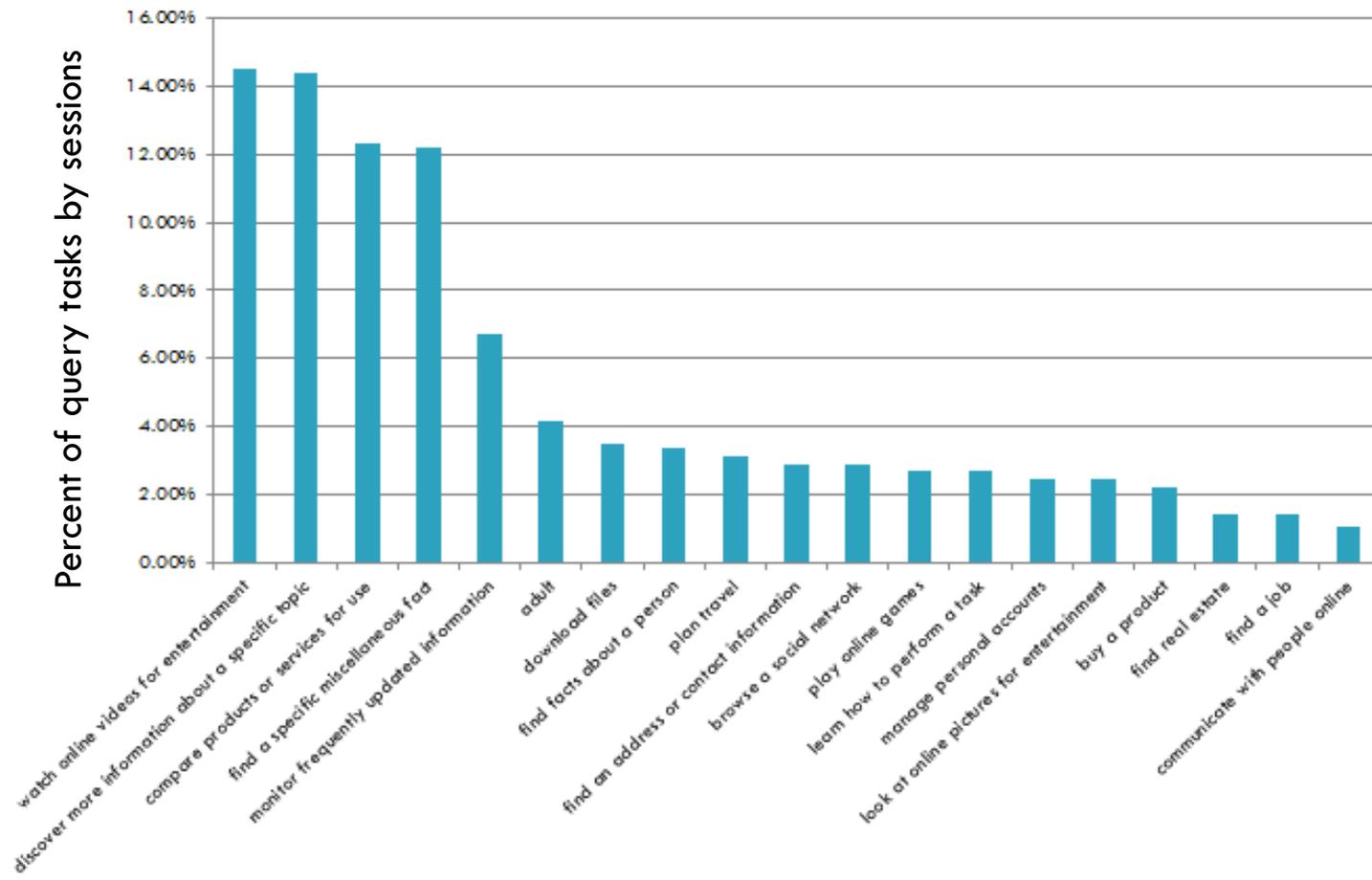
Top Web Tasks by Session

- Web Tasks: All web activities including browsing behavior and search behavior



Top Query Tasks by Session

- Query Tasks: Contiguously labeled tasks within a session which contain a query issued to Google, Yahoo! or Bing



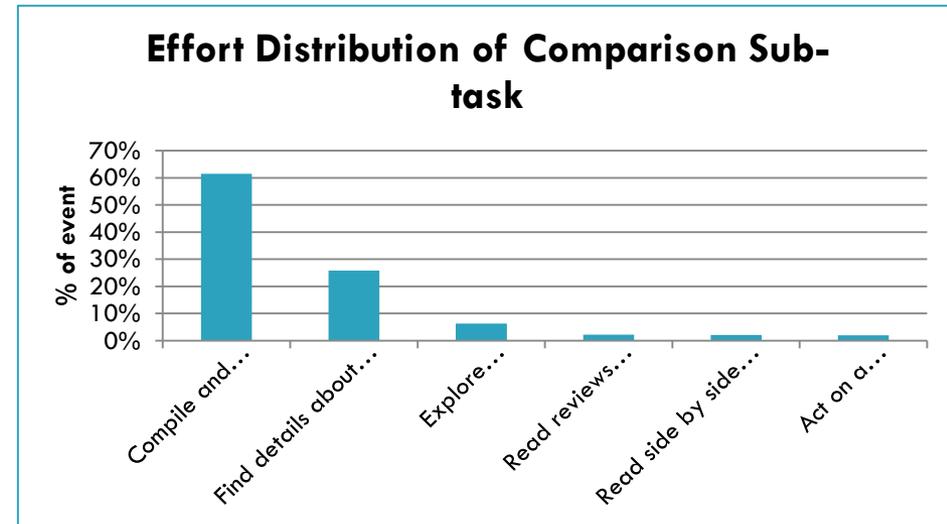
Bold = good engagement, opportunity for enhanced support

Other Task Properties

Task	Queries per task	Avg. events per task	Avg. length (mins)
<i>adult</i>	18.7	18.3	48.6
<i>look at online pictures for entertainment</i>	16.2	50.7	15.7
learn how to perform a task	13.0	11.9	8.5
<i>download files</i>	11.7	31.2	15.0
<i>watch online videos for entertainment</i>	7.5	19.5	19.0
find facts about a person	6.9	18.9	4.8
discover more information about a specific topic	6.8	24.8	13.5
compare products or services for use	6.8	22.3	24.8
find real estate	5.1	11.7	14.9
plan travel	4.7	5.1	12.0
find an address or contact information	4.2	48.5	7.5
not a task	3.9	10.3	21.6
monitor frequently updated information	3.6	24.1	20.6
find a specific miscellaneous fact	3.2	40.0	7.9
buy a product	3.1	15.4	8.9
play online games	2.0	21.1	16.2
manage personal accounts	1.8	42.0	7.5
find a job	1.8	29.4	18.0
communicate with people online	1.8	9.0	5.6
browse a social network	1.5	7.1	24.7

E.g., Compare Task

- Average compare task has lots of engagement
 - ▣ Contains 7 queries, 22 actions
 - ▣ Lasts for 25 minutes
- Compare sub-tasks
 - ▣ Explore dimensions for comparison (e.g., size, color, capacity)
 - ▣ **Compile and refine list of choices**
 - ▣ **Find details**
 - ▣ Read reviews
 - ▣ Read side-by-side
 - ▣ Act on a comparison decision



Automatically Detecting Tasks

- Approach
 - ▣ Encode search interaction features
 - ▣ Label some tasks
 - ▣ Learn a model that links implicit behaviors to explicit task labels
 - ▣ Run learned model in open-loop
- Used to model
 - ▣ Tasks within sessions
 - ▣ Tasks that extend across session and/or device

Cross-Session Tasks

- Many tasks continue across sessions
 - ▣ E.g., Medical diagnosis and treatment, event planning, how-to advice, shopping research, academic research, etc.

Time	Query	NClicks
1/22/2011 1:11pm	ken more	
1/22/2011 1:13pm	kenmore elite parts	
1/22/2011 1:14pm	kenmore elite washer	
1/22/2011 1:15pm	kenmore elite washer troubleshooting	3
1/22/2011 3:11pm	hotmail	
1/22/2011 4:15pm	weather nyc	
1/22/2011 4:19pm	hotmail	
1/22/2011 4:46pm	ebay	10
1/22/2011 6:40pm	how to fix kenmore elite washer	2
1/22/2011 6:44pm	kenmore elite error codes	
1/23/2011 6:20pm	kenmore elite error codes	
1/23/2011 6:22pm	kenmore elite error codes ce	1
1/24/2011 6:30pm	rangers scores	
1/24/2011 6:40pm	kenmore elite washer manual	
1/24/2011 6:55pm	kenmore elite calypso washer manual	1
1/24/2011 7:01pm	elite calypso wash motion	3
1/24/2011 8:31pm	hotmail	
1/25/2011 6:29pm	sears washer/dryer reviews	3
1/25/2011 7:31pm	kenmore elite parts	
1/25/2011 7:35pm	kenmore elite parts sears	1
1/25/2011 7:37pm	sears nyc hours	1
1/25/2011 8:00pm	nytimes.com	5
1/25/2011 8:35pm	regal cinema schedule	1

fixing a kenmore washing machine

- 16 queries
- 25 clicks
- several hours
- over 3 days

Cross-Session Tasks

- Many tasks continue across sessions
 - ▣ E.g., Medical diagnosis and treatment, event planning, how-to advice, shopping research, academic research, etc.
- Can we *identify same task, and predict task resumption?*
- Data and labeling
 - ▣ 270k people w/ 5+ search sessions and 10+ queries
 - ▣ Identify an early-dominant task
 - Task occurs during first 2 days
 - 2 or more unique queries on same task
 - ▣ Automatically cluster 10k queries into tasks
 - ▣ Manually label 1.2k of these tasks

Time	Query	NClicks
1/22/2011 1:11pm	ken more	
1/22/2011 1:13pm	kenmore elite parts	
1/22/2011 1:14pm	kenmore elite washer	
1/22/2011 1:15pm	kenmore elite washer troubleshooting	3
1/22/2011 3:11pm	hotmail	
1/22/2011 4:15pm	weather nyc	
1/22/2011 4:19pm	hotmail	
1/22/2011 4:49pm	ebay	10
1/22/2011 6:40pm	how to fix kenmore elite washer	2
1/22/2011 6:44pm	kenmore elite error codes	
1/23/2011 6:20pm	kenmore elite error codes	
1/23/2011 6:22pm	kenmore elite error codes ce	1
1/24/2011 6:30pm	rangers scores	
1/24/2011 6:40pm	kenmore elite washer manual	
1/24/2011 6:55pm	kenmore elite calypso washer manual	1
1/24/2011 7:01pm	elite calypso wash mclon	3
1/24/2011 8:31pm	hotmail	
1/25/2011 6:29pm	sears washer/dryer reviews	3
1/25/2011 7:31pm	kenmore elite parts	
1/25/2011 7:35pm	kenmore elite parts sears	1
1/25/2011 7:37pm	sears nyc hours	1
1/25/2011 8:00pm	nytimes.com	5
1/25/2011 8:35pm	regal cinema schedule	1

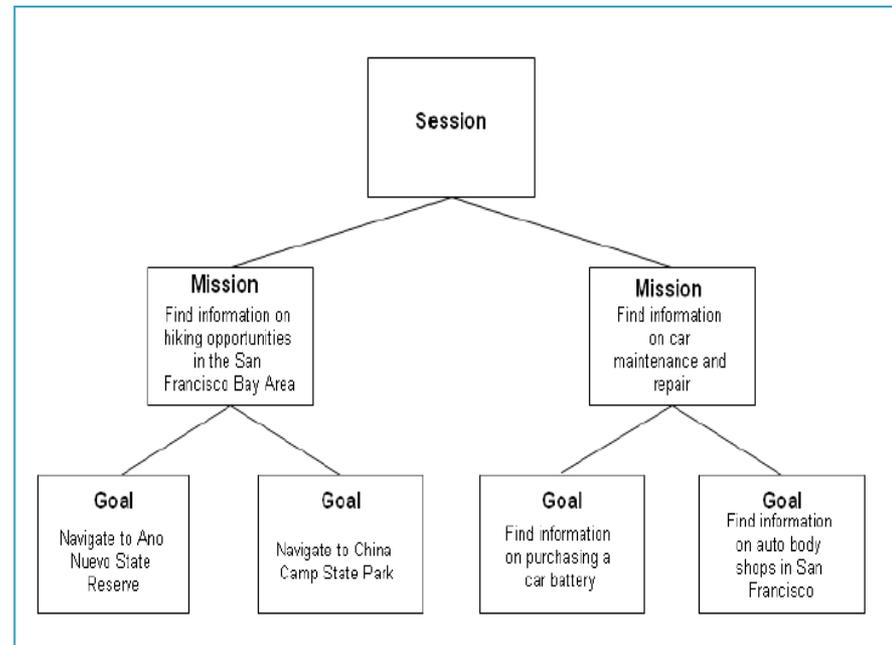
Cross-Session Tasks (cont'd)

- 15% of early-dominant tasks continue across sessions
 - ▣ 25% of queries are from multi-session tasks
- Two prediction tasks
 - ▣ Identify Same Task: *Given current query, find (previous) queries on this task*
 - ▣ Predict Task Resumption: *Will searcher resume the task (within the week)*
- Behavioral interaction features
 - ▣ Query, Session, History, Pair-wise
- Labels (automatic, human)
- Learned model (logistic regression, MART)
- Prediction accuracy
 - ▣ Same Task: $F1 = 0.67$
 - ▣ Task Continuation: $F1 = 0.75$

Feature	Weight	Feature Type
SameQueryHist	1.11	History-based
NumSessHist	0.60	History-based
NumDomQueriesHist	0.39	History-based (Table 6)
AvgInterQTimeHist	0.24	History-based (Table 6)
FreqDomQueriesHist	0.24	History-based (Table 6)
NumDwell30Hist	0.22	History-based (Table 6)
NumQueryHist	0.21	History-based
NumTop10Clicks	-0.16	Query-based
AvgInterQTimeSess	-0.17	Session-based (Table 6)
NumClicksHist	-0.18	History-based
NumQueryChars	-0.21	Query-based
SubQueryHist	-0.23	History-based
SupQuerySess	-0.40	Session-based
SupQueryHist	-0.40	History-based
SubQuerySess	-0.49	Session-based

Search Missions and Goals

- Hierarchical segmentation of search logs into missions and goals
- Data and labeling
 - ▣ 312 user sessions, 3 days
 - ▣ 1820 missions, 2922 goals, 8266 queries
 - ▣ Label all goals and missions within session
- Supervised learning
 - ▣ Task boundary detection
 - ▣ Same task identification



Research Missions and Search Pad

- Search missions and goals (from Jones & Klinkner)
 - ▣ 10% of search sessions contain “research” missions
- Can research missions be identified on-the-fly during a search session?
 - ▣ Input features representing tasks and engagement
 - Textual – similarity $q1, q1$
 - Session-based – queries, clicks, queries since last click, etc.
 - Time-based – time between $q1, q2$, total session time, etc.
 - ▣ Three general signals compute probability of research mission
 - Research_mission ($q1, q2$) boosted dt classifier
 - Same_mission ($q1, q2$) boosted dt classifier
 - Similarity (topics ($q1$), topics ($q2$))
- If research mission detected, show Yahoo! Search Pad

Research Missions and Search Pad

- If research mission detected, show Yahoo! Search Pad
- Deployed broadly, for a while

Scenario:

Searcher is planning a trip to Barcelona

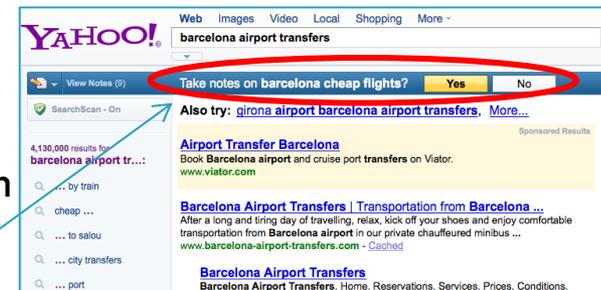
Issues a few related queries:

“*Barcelona cheap flights*”
(several clicks)

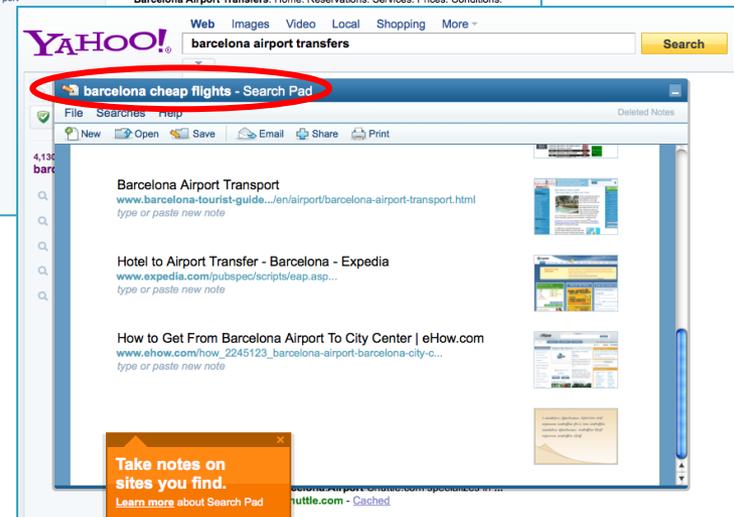
“*Barcelona airport*” (several clicks)

Then, when issuing the query
“*Barcelona airport transfer*”,
she might see ...

Research Mission
detected



Search Pad
invoked



How to Support Tasks

□ (Better Ranking, Query Suggestion, etc.)

□ History of queries and/or URLs

□ Richer snippets

□ Verticals



Welcome to [SIGIR Conference 2013 / Home](#)

www.sigir2013.ie

Welcome We are delighted to welcome SIGIR 2013 to Dublin, Ireland. SIGIR was last held in Dublin almost 20 years ago in 1994. The intervening years have seen huge ...

[About Dublin](#)

[Contact Us](#)

[Programme](#)

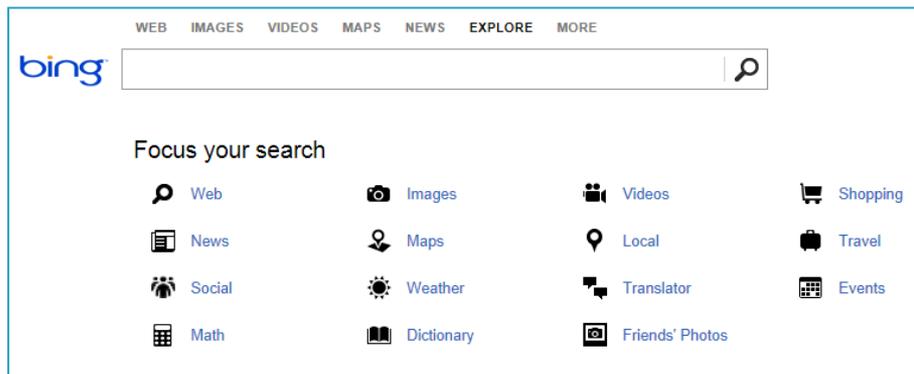
[Social Programme](#)

[Accommodation](#)

[Sponsorship](#)

[Organisation](#)

[Conference Venue](#)



How to Support Tasks

- (Better Query Suggestion, Ranking, etc.)
- History of queries and/or URLs
- Richer snippets
- Verticals
- Inline “answers”

Flight status for United 1637
flightstats.com · 1 minute ago

 **Arriving early at 3:07 PM in EWR**
Track this flight in real time

FROM	SEA Seattle	6:56 AM (was 7:00 AM) 04/09/2013	Terminal N, Gate 15 Airport map
TO	EWR Newark	3:07 PM (was 3:25 PM) 04/09/2013	Terminal C, Gate 80 Airport map

Other flight segments: [EWR-TPA](#)

bing weather chapel hill nc

15,300,000 RESULTS

Weather in Chapel Hill, North Carolina
bing.com/weather · Data from AccuWeather
[Hourly](#) · [Radar & Maps](#)

Now	Wed	Thu	Fri	Sat	Sun	Next 5 days
49°						
Clear						
°F °C	57° / 30°	53° / 32°	60° / 42°	71° / 39°	59° / 44°	

Change providers: [iMap Weather](#) · [Foreca](#) · [Compare all](#)

27514 Weather Forecast and Conditions - weather.com
[www.weather.com/weather/today/Chapel+Hill+NC+27514](#)
Mar 12, 2013 · 27514 weather forecast and weather conditions. Today's and tonight's 27514 weather forecast plus Doppler radar from weather.com.

dollars to euros

18,600,000 RESULTS

Currency Converter
[investing.money.msn.com/investments/currency-converter-calculator](#)

10/5/2012 1/1/2013 4/9/2013

1	US Dollar	To	Euro	Convert
---	-----------	----	------	-------------------------


0.75
0.70

1 US Dollar = 0.763802 Euro
1 EUR = 1.30923982 USD

How to Support Tasks

- (Better Query
- History of que
- Richer snippet
- Verticals
- Inline “answers
- Custom experiences
e.g., entities, exploration

The screenshot shows a Bing search for "trellis kirkland wa". The search results include several entries for "Trellis Restaurant" and "The Heathman Hotel". A map of Kirkland, WA, is displayed, showing the location of Trellis. Social results from Facebook and other networks are also visible.

Search Results:

- The Heathman Kirkland | heathmankirkland.com**
www.heathmankirkland.com
Downtown Kirkland Boutique Hotel near Lake Washington. Book Now!
Location · Accommodations · Why Us · Special & Packages
- Kirkland Washington Restaurant - Trellis Restaurant**
https://www.heathmankirkland.com/trellis-trellis-restaurant.aspx
More Information Accepting Reservations Now VinTimate Video Kirkland, Washington Restaurant - Trellis Restaurant Nestled in the heart of charming downtown Kirkland...
- Kirkland Washington Hotels - The Heathman Hotel**
https://www.heathmankirkland.com
The Heathman - Luxury Historic Hotel in Kirkland, WA Luxury Hotel - ... A celebration of our land. Trellis creates a true "farm to table" dining experience.
- Trellis Restaurant - Kirkland, WA**
www.yelp.com · Restaurants · American (New) ·
220 Kirkland Ave (between 3rd St & Main St) Kirkland, WA 98033
Dec 28, 2012 · (425) 288-3996 · "I'm a fan of Trellis. They work hard to bring some pretty solid farm-to-table food in Kirkland. ..."
- Kirkland Washington Hotel Restaurant**
www.heathmankirkland.com/trellis
Dine at Trellis, an award winning Kirkland restaurant located at The Heathman Hotel which also offers private dining menus for special events, weddings and meetings.
220 Kirkland Ave, Kirkland, WA · (425) 288-3996 · Directions · Menu · \$\$\$
- Trellis Restaurant - Kirkland, WA**
www.yelp.com · Restaurants · American (New) ·
220 Kirkland Ave (between 3rd St & Main St) Kirkland, WA 98033
(425) 288-3996 · "The big negative was the service. We waited for 55 minutes to be seated for a 7-45 reservation, which means we ..."
- Trellis, Kirkland - Restaurant Reviews - TripAdvisor**
www.tripadvisor.com · ... Kirkland · Kirkland Restaurants ·
220 Kirkland Avenue, Kirkland, WA
Feb 26, 2013 · "This past Sunday, my wife and I were looking

Map: Trellis, 220 Kirkland Ave, Kirkland, WA. Directions · Streetside

Open today*

Mon - Thu	6:00 AM - 10:30 AM, 11:00 AM - 2:00 PM, 5:00 PM - 9:00 PM
Fri	6:00 AM - 10:30 AM, 11:00 AM - 2:00 PM, 5:00 PM - 10:00 PM
Sat	6:00 AM - 2:00 PM, 5:00 PM - 10:00 PM
Sun	6:00 AM - 2:00 PM, 4:00 PM - 9:00 PM

*subject to change

Recent reviews

- ★★★★ Feb 15, 2013
We sat out in the warmed patio for our dinner. It was nice and toasty, so I wasn't uncomfortable. Read more
- ★★★★ Feb 02, 2013
Only visit this establishment for happy hour. The prices are reasonable at that ... Read more

See all 151 Yelp reviews

Social Results

Search is better with friends! See posts and photos from your friends when you search on Bing.

Connect with Facebook

From social networks (4)

- Melissa O.
- 07 Aug 2012
Trellis - Kirkland
foursquare.com
Love the simple value included as their guest, makes client meetings a breeze!
- Andru E.
- 04 Aug 2012
Trellis - Kirkland
foursquare.com
Try the Mopan Chocolate Cake and spot it on Facebook!
- Eric R.
- 19 Aug 2012
Trellis - Kirkland

sigir 2013

51,300 RESULTS

[Welcome to SIGIR Conference 2013 / Home](#)
www.sigir2013.ie

Welcome We are delighted to welcome SIGIR 2013 to Dublin, Ireland. SIGIR was last held in Dublin almost 20 years ago in 1994. The intervening years have seen huge ...

About Dublin
SIGIR is the Association for Computing Machinery's Special ...

Organisation
SIGIR 2013 / Committee

Programme
SIGIR is the Association for Computing Machinery's Special ...

Contact Us
SIGIR 2013 / Contact Us

Accommodation
SIGIR is the Association for Computing Machinery's Special ...

Social Programme
SIGIR 2013 / Conference

See results only from sigir2013.ie

SIGIR 2013 / Conference
sigir2013.ie/posterdemonstrations.html

SIGIR is the Association for Computing Machinery's Special Interest Group on Information Retrieval. Today, SIGIR alternates in a 3 year cycle between a North ...

Sigir 2013 | Facebook
www.facebook.com/pages/Sigir-2013/148345471890928

Sigir 2013, Dublin. 80 likes · 2 talking about this. ... Historic Trinity College Dublin, where the Main SIGIR Conference will be held...

SIGIR 2013

ACM SIGIR is the Association for Computing Machinery's Special Interest Group on Information Retrieval ...

www.cng.ie

Dates: Jul 28 - Aug 01, 2013
Location: Dublin, Ireland
Website: Official site

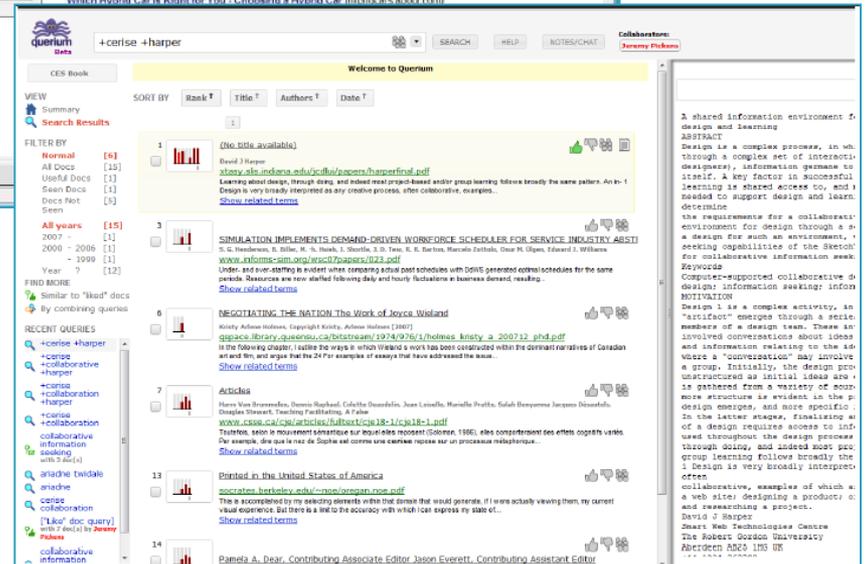
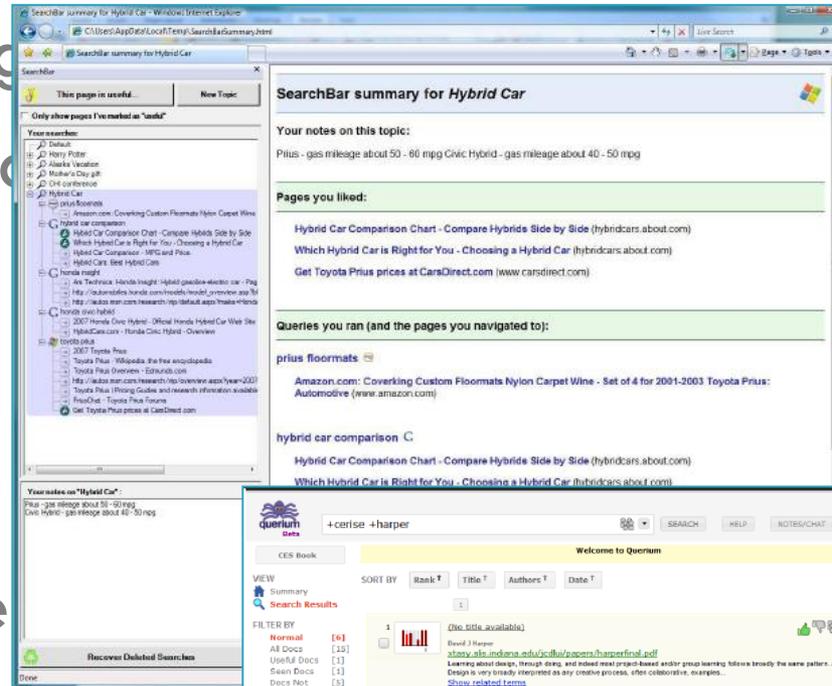
People also search for

- IJCAI-13
- WWW 2013
- WSDM 2013

Data from: sigir2013
Report a problem

How to Support Tasks

- (Better Query Sugg
- History of queries of
- Richer snippets
- Verticals
- Inline “answers”
- Custom experience
- Richer Sensemaking



How to Support Tasks

- (Better Query Suggestion, Ranking, etc.)
- History of queries and/or URLs
- Richer snippets
- Verticals
- Inline “answers”
- Custom experiences – e.g., entities
- Richer Sensemaking
- Apps, apps and more apps



How to Evaluate Task Support

□ How to measure success

□ Explicit

- Retrospective (By expert judges, crowd workers or individual)
- *In situ* (By individual e.g., Curious Browser)

□ Implicit

- Off-line using learned models
- On-line operational system (using controlled experiments)

□ What to measure

□ Many metrics: “scorecards”

- E.g., clicks, dwell, engagement, time on task, sustained use, etc.

□ System as whole (and system components)

- *“there is a quantum of barriers in task integration and work task contexts”*

Summary

- Tasks are important
 - ▣ Significant time is devote to some tasks which require multiple queries, sessions and devices to complete
- Some progress in modeling and support
 - ▣ Automatic detection and prediction of tasks, using behavioral modeling, implicit feedback and machine learning
 - ▣ Support for tasks increasing within Web search engines and in stand-alone apps
- Many challenges/opportunities remain
 - ▣ Task selection: where are biggest opportunities
 - ▣ Support: broadly available vs. specific task
 - Are there reusable components that generalize across tasks?



□ Thanks !

□ Questions / Comments ???

(Some) References

- Bailey, *Under task understanding*. NII Shonan 2012
- Bennett et al., *Modeling the impact of short- and long-term behavior on search personalization*. SIGIR 2012
- Broder, *A taxonomy of web search*. SIGIR Forum 2002
- Donato et al., *Do you want to take notes? Identifying research missions in Yahoo! Search Pad*. WWW 2011
- Fox et al., *Evaluating implicit measures to improve the search experience*. TOIS 2005
- Jones & Klinker, *Beyond the session timeout: Automatic hierarchical segmentation of search topics in a query log*. CIKM 2008
- Kotov et al., *Modeling and analysis of cross-session search tasks*. SIGIR 2011
- Richardson, *Learning about the world through long-term query logs*. TWEB 2008
- Russell et al., *Retrospective cued recall: A method for accurately recalling previous user behaviors*. HICCS 2009