



Supporting Complex Search Tasks

Nick Belkin Charlie Clarke Ning Gao Jaap Kamps Jussi Karlgren

SIGIR 2011 Workshop, July 28, 2011

Search support has massively improved



PLACE
STAMP
HERE

GOOGLE
CLASSIC

QUERY:

IMAGES NEWS VIDEO MAPS OTHER

SEND YOUR QUERY TO: GOOGLE INC., 1600 AMPHITHEATRE PARKWAY, MOUNTAIN VIEW, CA 94043, UNITED STATES

PLEASE ALLOW 30 DAYS FOR SEARCH RESULTS



New technology to solve the old search problem



A person with tattoos and a green t-shirt is shown in a state of intense frustration, screaming at a computer monitor. The person's mouth is wide open, and their hands are raised in a gesture of exasperation. The background is a cluttered office space with various items on the desk and shelves. The overall scene conveys a sense of overwhelming complexity and frustration.

Complex tasks are still painstaking!

Many queries, tabs, notes, cut-and-paste, ...

entertain me - Google Search

Web Images Videos Maps News Shopping Gmail more ▾

Web History | Search settings | Sign in

Google

entertain me × Search Instant is on ▾

About 7,490,000 results (0.09 seconds) [Advanced search](#)

Everything
Images

▶ [Belief-O-Matic- Beliefnet.com](#)

Even if YOU don't know what faith you are, Belief-O-Matic TM knows. Answer 20 questions

entertain me - Bing

Web Images Videos Shopping News Maps More | MSN Hotmail Sign in ▾ Amsterdam, No

bing™

entertain me

Web News Videos

RELATED SEARCHES
[I Am Bored](#)
[I'm Bored](#)

ALL RESULTS 1-10 of 30,100,000 results · [Advanced](#)

[2011 Entertainment® Book](#) Sponsored sites
[www.entertainment.com](#) · \$19.99 for the First Book. Extra Books Are \$15. All Books

Can we deal with whole search task?

YAHOO!

Web Images Video Local Shopping News More ▾

entertain me Search

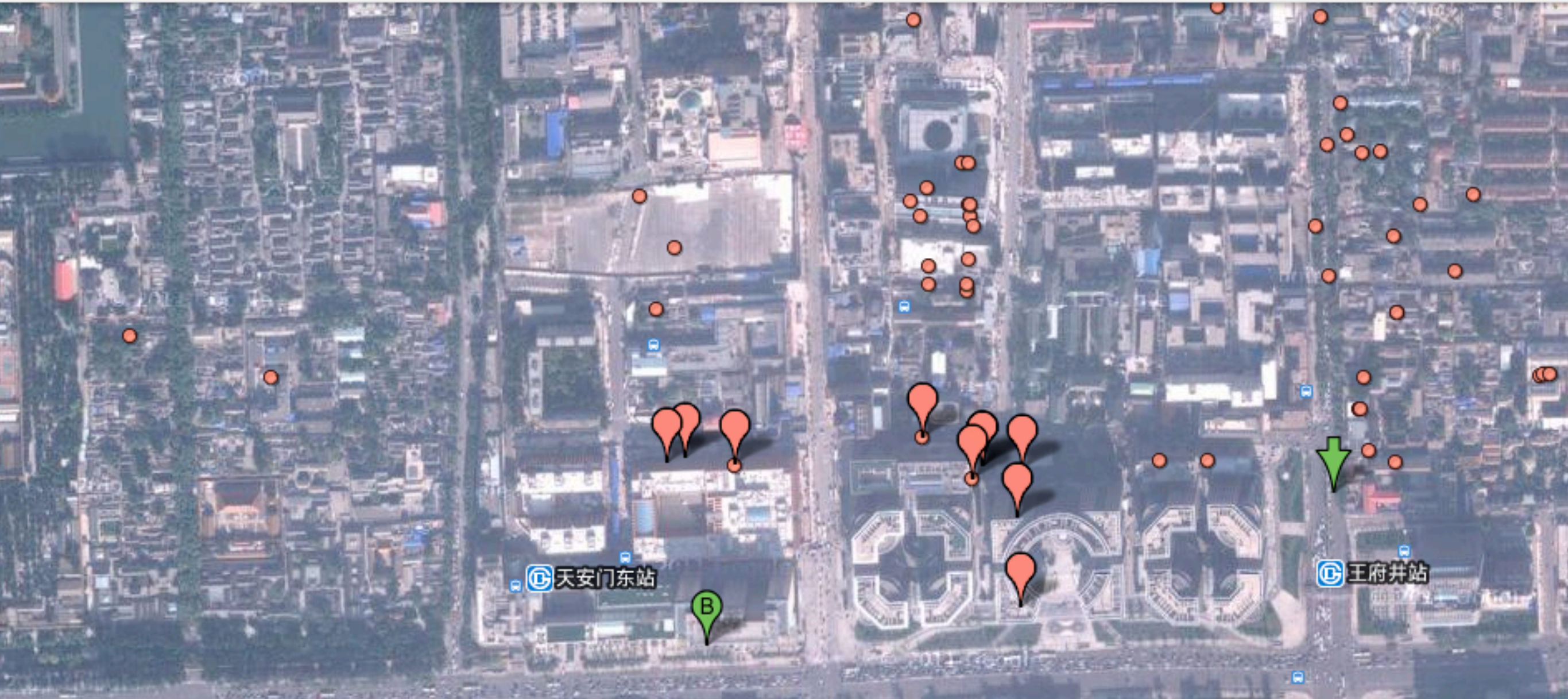
QuickApps

Search Pad

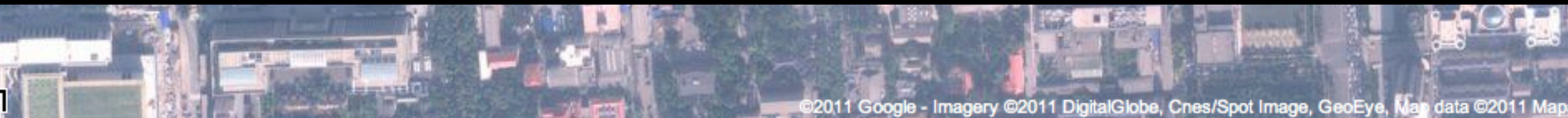
SearchScan - On

Also try: [im bored entertain me](#), [boppy entertain me play gym](#), [more...](#)

[2011 Entertainment® Book](#) Sponsored Results
\$19.99 for the First Book. Extra Books Are \$15. All Books Ship Free!
[www.entertainment.com](#)



Rich context is extremely powerful




```

<?xml version="1.0" encoding="UTF-16"?>
<queryplan end-time="2010-05-04T15:55:35.038Z" start-time="2010-05-04T15:55:35.026Z"
  xDB-version="xDB main@621581">
  <XQueryQuery accumulatedTime="10" calls="1" pagesRead="5" values="1">
    <querytext>declare option xhive:index-debug 'true';
declare option xhive:queryplan-debug 'true';
declare option xhive:pathexpr-debug 'true';
(: declare option xhive:ignore-indexes 'mp1';:)
let $othello docs := /feed/doc[. contains text 'Othello'],
    $books := for $book in /bib/book[author/last = 'Stevens'] return $book
return <res>{ $othello docs, $books }</res></querytext>
    <functions/>
    <variables/>
    <modules/>
    <let accumulatedTime="10" calls="1" location="query:5:1" pagesRead="5"
      type="item()*" values="1" variable="othello docs@0">
      <path accumulatedTime="6" calls="1" location="query:5:21" numExpr="2"
        onlyChildren="true" pagesRead="3" returnBlobs="false"
        usesNotOrOr="false" values="0">
        <indexplans>
          <indexplan context="/" node="primary">
            <lookup accumulatedTime="2" calls="1" conditions="1"
              index="mp1" lookup="server-side" pagesRead="2" type="11"
              values="0"/>
          </indexplan>
        </indexplans>
      <path path=".../child::feed/child::doc[. contains text Othello]">
        <root accumulatedTime="0" calls="1" location="query:5:21"
          pagesRead="0" values="0"/>
      </path>
    </let>
  </XQueryQuery>
</queryplan>

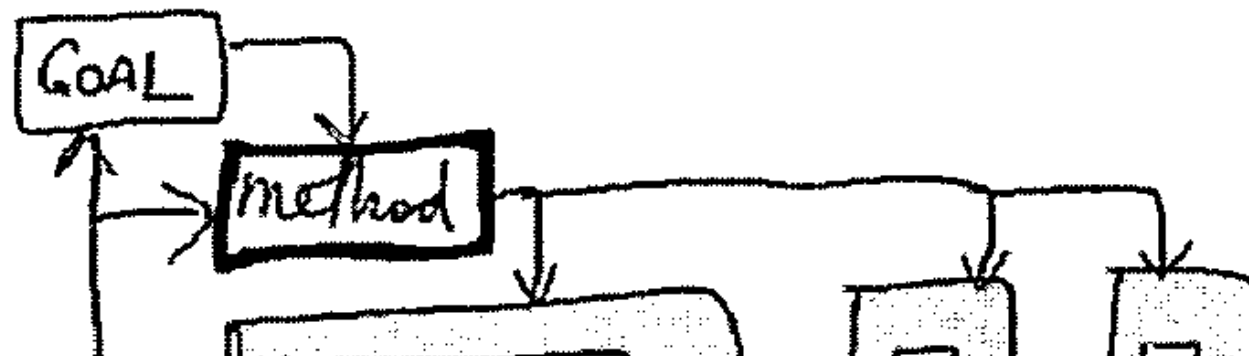
```

Interactively construct a (hidden) query

```

type="item()*" values="2" variable="book@1">
<path accumulatedTime="1" calls="1" location="query:6:28"
  numExpr="2" onlyChildren="true" pagesRead="2"
  returnBlobs="false" usesNotOrOr="false" values="2">
  <indexplans>
    <indexplan context="/" node="primary"/>
    <indexplan context="/dewiki20m.xml" node="primary"/>
    <indexplan context="/bib.xml" node="primary"/>
  </indexplans>
  <path path=".../child::bib/child::book
    [child::author/child::last[. = &quot;Stevens&quot;]]">
    <root accumulatedTime="0" calls="1" location="query:6:28"
      pagesRead="0" values="0"/>
  </path>

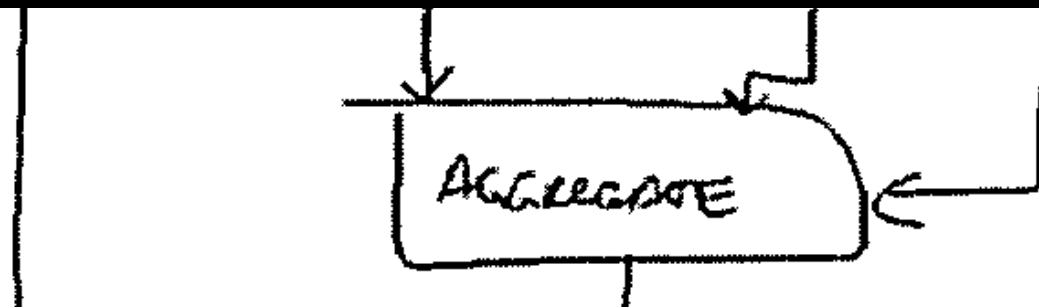
```



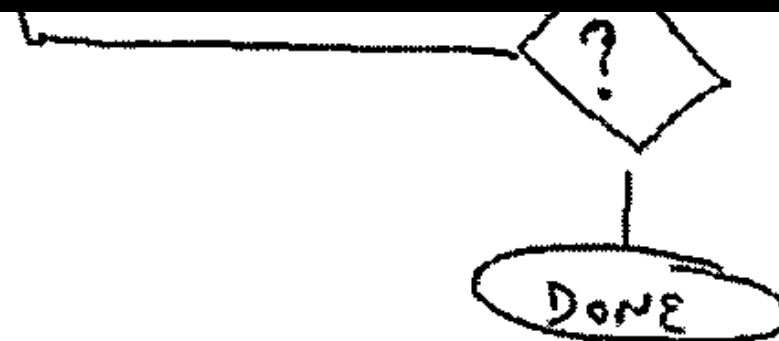
What are these tasks?



What does the system need to know?



What are the key system components?



Keynotes by Qi Lu and ChengXai Zhai

- What is a "task" and how can we formally model a task? (task vs. intent vs. information needs)
- How to design a task specification language?
- How do we design a set of general analysis operators to accommodate many different tasks?
- What does ranking mean in an analysis engine (ranking terms, documents, topics, operators)?
- What should the user interface look like?
- How can we seamlessly integrate search and analysis?
- How should we evaluate an analysis engine?
- ...



The banner for the ACM SIGIR 2011 conference features a central image of a speaker at a podium. The background of the banner includes logos for Microsoft, Baidu, Emap, Google, Sogou, and Yahoo!, along with the text 'Microsoft', 'Baidu', 'EMAP', 'Google', 'SOGOU', 'Yahoo!', 'eBay', 'HP', and 'Search Index'. The main text on the banner reads 'ACM SIGIR 2011', 'July 24-28, 2011, Beijing, China', and 'The 34th Annual International ACM SIGIR Conference'. A small logo is visible in the top right corner of the banner.

- What is a "task" and how can we formally model a task? (task vs. intent vs. information needs)
- How to design a task specification language?
- How do we design a set of general analysis operators to accommodate many different tasks?
- What does ranking mean in an analysis engine (ranking terms, documents, topics, operators)?
- What should the user interface look like?
- How can we seamlessly integrate search and analysis?
- How should we evaluate an analysis engine?
- ...

Program

- 09:00-09:15 Opening
- 09:15-09:45 Jussi Karlgren (Invited)
- 09:45-10:30 Papers "Complex Search Needs and Use-Cases"
- 10:30-11:00 Coffee break
- 11:00-12:00 Papers "Eliciting Complex Needs and Queries"
- 12:00-13:00 Papers "Task Context and Success"
- 13:00-14:00 Lunch break
- 14:00-15:30 Breakout groups on the three themes
- 15:30-16:00 Coffee break
- 16:00-17:00 Reports & Wrap-up
- 17:30-late Social evening program

Outcome?

1. Clear grasp on the problem:

- where are we now?
- where do we want to go?
- how to get there?

2. Plan to take this further:

- Report , white-paper, roadmap
- Proposal for next year's workshop
- Something we can run at trec/ntcir/clef/inex/...

Discussion is the Objective

- This is a *workshop* (!)
- but before we start
- let's introduce ourselves...

ENTERTAINMENT

