Goals, Tasks, and Information Retrieval

Nicholas J. Belkin
School of Communication & Information
Rutgers University
New Brunswick, NJ, USA
belkin@rutgers.edu

Overview

- The traditional concerns of information retrieval (IR) research, and their problems
- An alternative view of the concerns of IR, and their consequences (session-oriented IR)
- Overview, with some examples, of current "session-oriented" research
- An example of such research specific to patent search
- Possible future directions for session-oriented IR systems

“Traditional” IR Research

- Concerned with providing “relevant” documents (or parts of documents) in response to a query, which represents an “information need”
- Topical relevance is understood as the minimal condition for relevant documents
- Additional conditions can be imposed, such as “authority” (e.g. PageRank)
- The response to a query is a ranked list of items
- Performance is evaluated with respect to “goodness” of the response, with relevance as the basic criterion.

Some Fundamental Assumptions of Traditional IR Research

- The person enters the IR system with a *static* "information need"
- The goal of the system is to satisfy that need
- A query is a (relatively) good representation of the need
- A system satisfies the need through response to a query
- Evaluation of the performance of the IR system should be according to how well it responds to a "best" query

Some Problems with the Traditional IR Research Assumptions

- People engage in information seeking because of a goal to achieve, a task to accomplish; information is a means for doing that, not a "need"
  - Wilson (1980) made this point very clearly
- People *change* during information seeking interactions
  - Goal/task may remain constant, but knowledge does not
- Most serious encounters with IR systems are more than single queries, but rather "search sessions"

Alternative Assumptions about IR

- People engage in information seeking because they want to achieve a goal, accomplish a task
- The goal of IR is to support people in achieving their motivating goal/task
- People engage in information seeking sessions (ISSs)
- People change during the course of ISSs
- People try to accomplish many things during ISSs
- IR systems should be evaluated according to how well they support the motivating task/goal

Implications of Alternative Assumptions for IR Research

- Study, and understand, motivating tasks/goals and their relationships to information seeking behaviors
- Study, support and evaluate search sessions as a whole
- Study, understand, and support people's various *information seeking intentions* in IR interaction
- Study, understand, and account for contextual and situational variables in IR interaction
- *Personalize* support for information seeking

Motivating Goals/Tasks

- Classifying or otherwise characterizing goals or tasks
- Research initially concerned with work tasks/goals, but extended to everyday life information seeking
- Most often based on task *complexity*, but other characteristics also considered
- Usually explicitly related to information search goals, e.g. factual, exploratory.

Li & Belkin (2008) Facet Analysis of Task

- Source of Task
  - Self, Group, Assigned
- Task Doer
  - Individual, Group
- Time
  - Frequency
  - Length
  - Stage
- Product
  - Physical, Intellectual, Decision, Factual
- Process
  - One-time, Multiple
- Items
  - Named or Not
  - Whole or Part
- Goal
  - Quality
    - Specific, Amorphous, Mixed
  - Quantity
    - Single or multiple goals
- Common attributes of task, e.g.
  - Objective/Subjective task complexity, Urgency, Salience, Difficulty, …

Example Task and Classification

Assignment 1. Copy Editing (CPE)
Your Assignment: You are a copy editor at a newspaper and you have only 20 minutes to check the accuracy of six italicized statements in the excerpt of a piece of news story below.
Your Task: Please find and save an authoritative page that either confirms or disconfirms each statement.

Product: Fact; **Items**: Named/Part; **Goal**: Specific

Example Task and Classification

Assignment 2. Story Pitch (STP)
Your Assignment: You are planning to pitch a science story to your editor and need to identify interesting facts about the coelacanth ("see-la-kanth"), a fish that dates from the time of dinosaurs and was thought to be extinct.
Your Task: Find and save web pages that contain the six most interesting facts about coelacanths and/or research about their preservation.

Product: Fact; **Items**: Not Named/Part; **Goal**: Specific

Example Task and Classification

Assignment 3. Relationships (REL)
Your Assignment: You are writing an article about coelacanths and conservation efforts. You have found an interesting article about coelacanths but in order to develop your article you need to be able to explain the relationship between key facts you have learned.
Your Task: In the following there are five italicized passages, find an authoritative web page that explains the relationship between two of the italicized facts.

Product: Intellectual; **Items**: Named/Part; **Goal**: Mixed (Specific + Amorphous)

Example Task and Classification

Assignment 4. Interview Preparation (INT)
Your Assignment: You are writing an article that profiles a scientist and their research work. You are preparing to interview Mark Erdmann, a marine biologist, about coelacanths and conservation programs.
Your Task: Identify and save authoritative web pages for the following: Identify two (living) people who likely can provide some personal stories about Dr. Erdmann and his work. Find the three most interesting facts about Dr. Erdmann's research. Find an interesting potential impact of Dr. Erdmann's work.

Product: Intellectual; **Items**: Not-Named/Whole; **Goal**: Amorphous

Kelly, et al. (2015) Tasks According to Cognitive Complexity

- Based on Anderson & Krathwohl (2001)

PROCESS	TARGET OUTCOME	MENTAL ACTIVITY
Remember(**R**)	Fact	Identify
Understand(**U**)	List (set)	Identify, Compile
Analyze(**A**)	List (prioritize), Description	Identify, Compile, Describe
Evaluate(**E**)	Recommendation	Identify, Compile, Describe, Compare, Decide, Justify
Create(**C**)	Plan	Identify, Compile, Describe, Compare, Decide, Make

Example Tasks

- Remember

You recently watched a documentary about people living with HIV in the United States. You thought the disease was nearly eradicated, and are now curious to know more about the prevalence of the disease. Specifically, how many people in the US are currently living with HIV?

- Understand

Your nephew is considering trying out for a football team. Most of your relatives are supportive of the idea, but you think the sport is dangerous and are worried about the potential health risks. Specifically, what are some long-term health risks faced by football players?

- Analyze

Having heard some of the recent reports on risks of natural tanning, it seems like a better idea to sport an artificial tan this summer. What are some of the different types of artificial tanning methods? What are the health risks associated with each method?

Example Tasks

- Evaluate

One of your siblings got a spur of the moment tattoo and now regrets it. What are the current available methods for tattoo removal, and how effective are they? Which method do you think is best? Why?

- Create

After the NASCAR season opened this year, your niece became really interested in soapbox derby racing. Since her parents are both really busy, you've agreed to help her build a car so that she can enter a local race. The first step is to figure out how to build a car. Identify some basic designs that you might use and create a basic plan for constructing the car.

Personalization of IR

- Can knowledge of characteristics of the task that motivated a person to engage in information seeking, knowledge about that person's domain/ topic knowledge, other characteristics, be derived from observation of search behaviors?
- Can such knowledge then be used to interpret those behaviors in order to personalize the search experience
- These questions led to the IMLS-funded project: Personalization of the Digital Library Experience

General Structure of PoODLE

- Journalism tasks given to journalism students to perform
- Tasks structured by Li & Belkin (2008)
- Three major user studies with goals of:
  - Interpret behaviors in order to predict document usefulness for relevance feedback; identify searcher characteristics through behaviors; characterize tasks through behaviors
- Participants' search behaviors, including eye tracking, logged; post-hoc search replay evaluated document usefulness

Tasks and Behaviors

- Characterizing tasks according to behaviors: what kind of a task led this person to engage in information seeking behavior?
- Cole, M.J., Gwizdka, J., Liu, C., Bierig, R., Belkin, N.J. & Zhang, X. (2011) Task and user effects on reading patterns in information search, *Interacting with Computers*, v.23 (4): 346-362.
- Differences in eye fixation patterns in tasks with different values of the Task Facets : Product; and Items.

Tasks and Behaviors

- Cole, M.J., Hendahewa, C., Belkin, N.J. & Shah, C. (2015) User activity patterns during information search. *ACM Transactions on Information Systems*, 33 (1): Article No. 1 (39 p.)
- Classification of sequences of behaviors (eye fixations; page type transitions) lead to a 8-11 clusters, whose frequency of occurrence differ in tasks that vary in values of Li & Belkin task classification facets.

Search Behaviors and Knowledge

- Zhang, X., Liu, J., Cole, M. & Belkin, N.J. (2015) Predicting users' domain knowledge in information retrieval using multiple regression analysis of search behaviors. *Journal of the Association for Information Science and Technology*, 66 (5): 980-1000.
- Liu, J., Liu, C. & Belkin, N.J. (2016) Predicting information searchers' topic knowledge at different search stages. *Journal of the Association for Information Science and Technology,* v. 67, no. 11: 2652-2666.

What do Searchers Try to do During a Search Session?

- Substantial evidence that searchers *want* to do more than get "relevant" documents during the course of a search session
- If this is the case, then
  - What are the things that they want to do?
  - How can these be supported by IR systems?
  - How do they relate to the task/goal that motivated the search session?

Xie's (2002) Interactive [Search] Intentions

- Identify search information (Something to start; Something more to search)
- Learn (Domain knowledge; Database content)
- Find (Known item; Specific information; Sharing named characteristic; Without predefined criteria)
- Keep record
- Access item or set of items
- Evaluate (Correctness; Usefulness; Best; Specificity; Duplication)
- Obtain (Specific information; Part of item; Whole item)

Behaviors and Search Intentions

- Cole, Hendahewa, Belkin & Shah (2015), and Cole, (2016) showed that it is possible to identify a (relatively) small number of sequences of behaviors, that vary in frequency and in sequence according to task facets.
- Could these sequences of behaviors be associated with "cognitive-level" search intentions?
- This question led to current NSF-funded research project: Characterizing and Evaluating Whole-Session Interactive Information Retrieval.

General CHEWS-IIR Procedure

- Journalism tasks yet again
- Journalism students yet again
- Logging as usual
- Post-hoc annotation of entire search sessions, eliciting, from participants:
  - Search intentions for each query segment, and success
  - Saved document usefulness
  - Reasons for query reformulation
- Participant knowledge of task and topic, and evaluation of success in task

Early Results

- Mitsui, M., Shah, C. & Belkin, N.J. (2016) Extracting information seeking intentions for web search sessions. In: *Proceedings of the 39th International ACM SIGIR Conference on Research & Development in Information Retrieval, SIGIR 2016* (pp. 841-844). New York: ACM.
- Early results indicating that the set of intentions actually was necessary, and perhaps sufficient, for participants to characterize what they wanted to accomplish during a query segment

Early Results

- Rha, E.Y., Belkin, N.J., Mitsui, M. & Shah, C. (2016) Exploring the relationships between search intentions and query reformulations. In: *Proceedings of the 79h Annual Meeting of the Association for Information Science and Technology,* (9 pp.). Silver Spring, MD: Association for Information Science and Technology
- Preliminary results relating type of query reformulation to search intention, indicating some regularities

Early Results

- Mitsui, M., Liu, J., Belkin, N.J. & Shah, C. (2017) Predicting information seeking intentions from search behaviors. In *Proceedings of the 40th International ACM SIGIR Conference on Research & Development in Information Retrieval, SIGIR 2017* (in press). New York: ACM.
- More substantive results, which demonstrate reasonable prediction of search intention on the basis of search behaviors (no eye tracking yet), using a linear classifier.

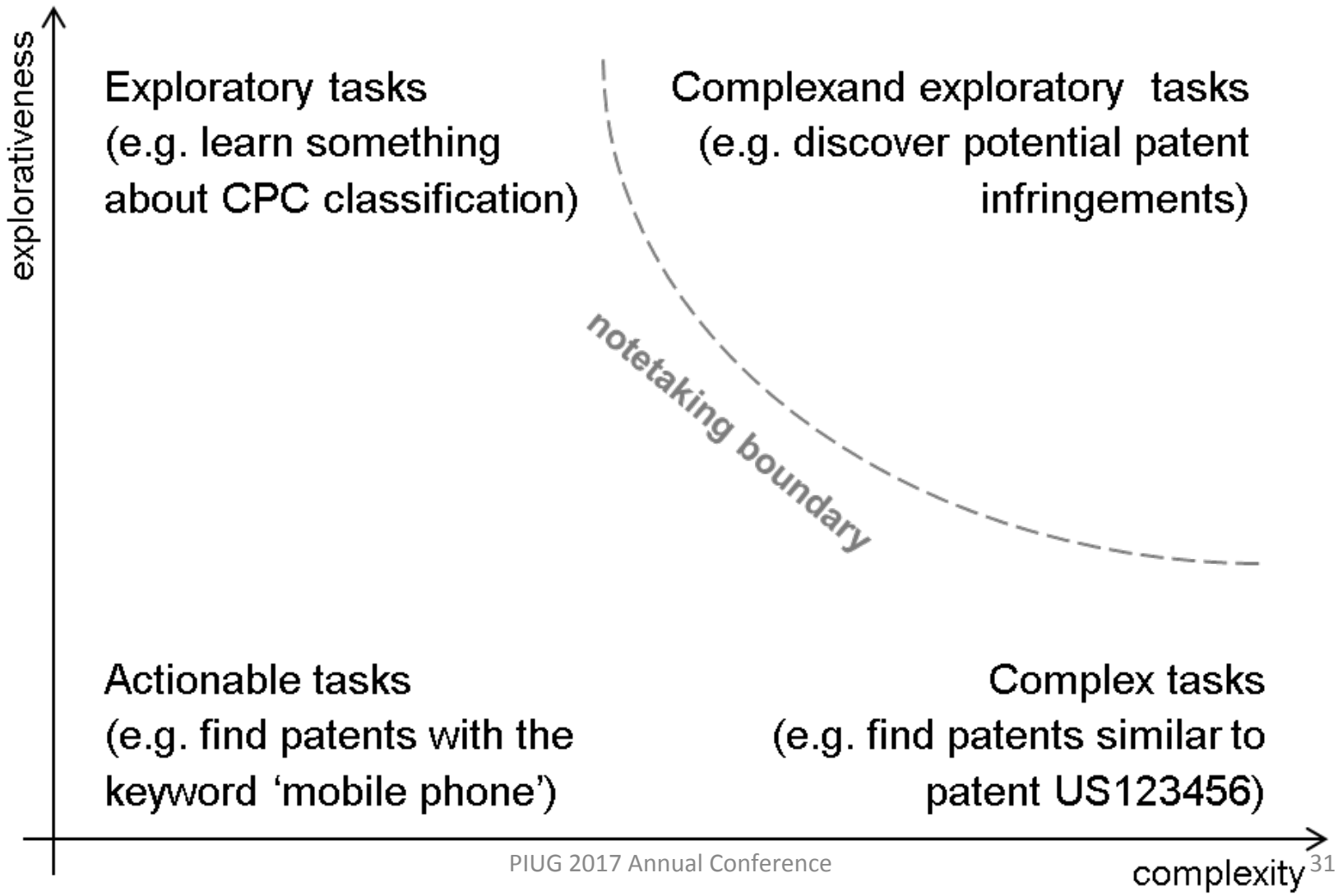
Example Task-Oriented Research in Patent Domain

- Kuhlahcioglu, T., Fradkin, D. & Palanivelu, S. (2017) Incorporating task analysis in the design of a tool for a complex and exploratory search task. In Proceedings of the Second ACM SIGIR Conference on Human Information Interaction and Retrieval (CHIIR '17) (pp. 373-376). New York: ACM.
- Done at Siemens and Rutgers University, with respect to supporting Siemens activities

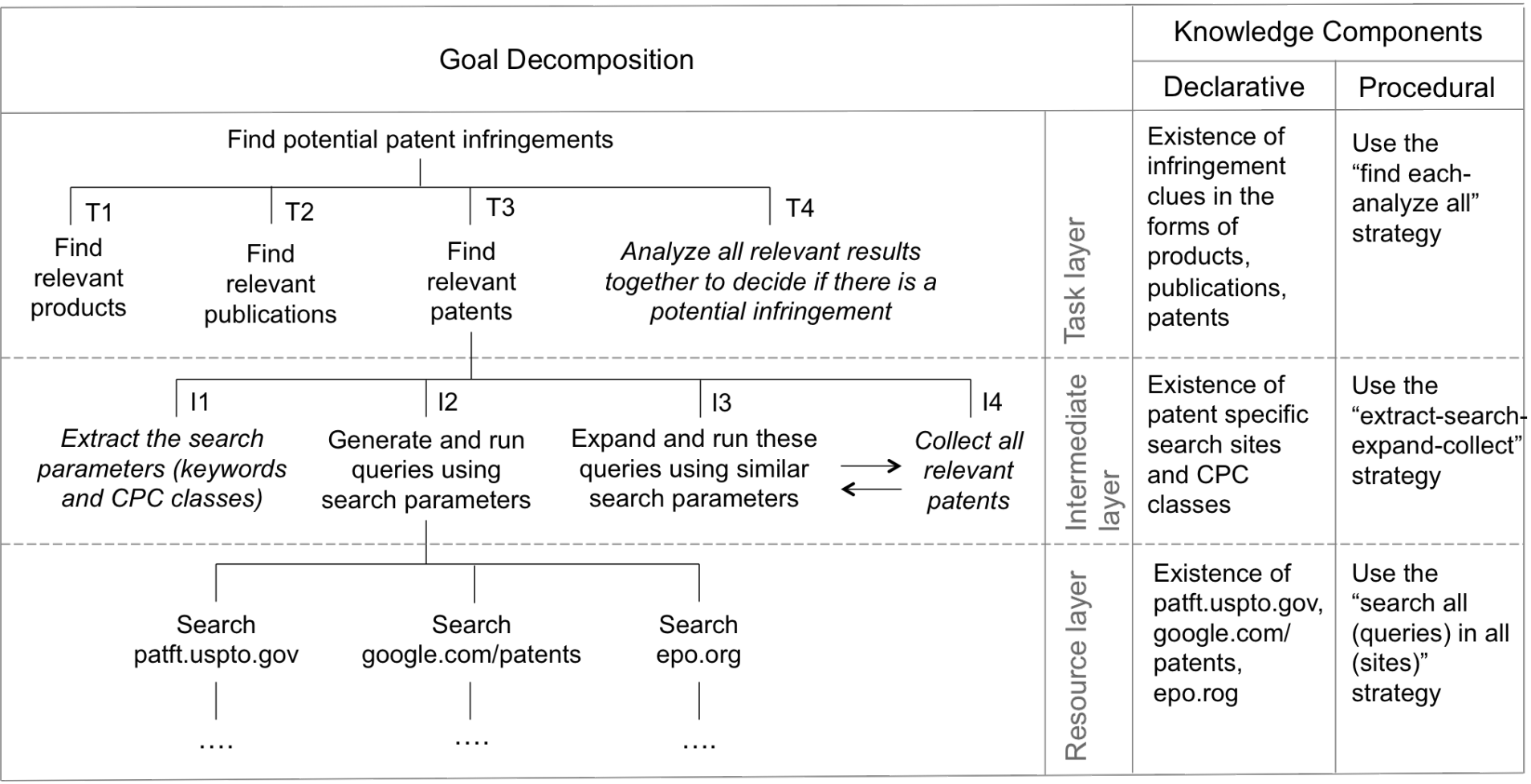
Designing to Support a Complex Task

- Applied “cognitive task analysis” to the task of patent infringement discovery
- First, hierarchical task analysis, to identify tasks
- Then, cognitive task analysis, to identify knowledge for tasks
- Then, tool design to reduce workload

Patent Infringement Tasks Example



Patent Infringement Goal Decomposition



Possible Tool Support for PI Task

- Automatically suggest search parameters relevant to input patent (I1)
- Tool for saving interim results from different verticals (I4, T4)
- Clustering mechanisms (T4)

Tasks and Intentions

- The PI example indicates that persons engaged in this task will have different intentions and different points or stages of the task
- Different intentions require different support techniques
- Search behaviors can (we believe) be indicative of search intentions
- Predicting search intentions can lead to offering "just in time" support

Search Session Support Evaluation

- The TREC "Session Track" has addressed this issue, without much success
- My colleagues and I have proposed that the appropriate criterion for interactive IR evaluation is *usefulness*, NOT relevance
- Belkin, N.J. (2015) Salton Award Lecture: People, interacting with information. *SIGIR Forum*, v. 49(2): 13-27. http://sigir.org/files/forum/2015D/p013.pdf
- Belkin, N.J. (2010) On the evaluation of interactive information retrieval systems. In: B. Larsen, J.W. Schneider & F. Åström (Eds.) *The Janus Faced Scholar. A Festschrift in Honour of Peter Ingwersen* (pp. 13-21). Copenhagen: Royal School of Library and Information Science. Retrieved on 7 January 2014 at http://www.issi-society.org/peteringwersen/

Concluding Comments

- An overview of a strand of current IR research, that has focused primarily on just one research group's work
- Substantial other research is taking place at many other groups, including industrial research labs (e.g. MSR, Google, Amazon, Walmart!, …)
- Other, related research directions, include "conversational search", design for specific tasks, rather than inferring tasks, and "searcher simulation" for design and evaluation purposes
- First steps only, and we still don't have good ways to evaluate support for whole search sessions

Concluding Comments

- The basis for this line of research is the realization that understanding people (i.e. users), their goals, tasks, intentions, differences, is central to the problem of information retrieval
- That's why I'm so happy to be able to talk about this research at this conference of USERS

THANKS FOR YOUR ATTENTION

- Acknowledgements due to all of the members of the PoODLE and CHEWS-IIR projects, and to our funders
- Some work reported here was supported through the National Science Foundation, grant #IIS-1423239.
- Some work reported here was supported by IMLS grant LG#06-07-0105-07
- Some work reported here was supported by a Google Faculty Research Award to N.J. Belkin & C. Shah