Contextual Authority Tagging : Expertise Location via Social Labeling

Dissertation Proposal

Terrell Russell

University of North Carolina at Chapel Hill
School of Information and Library Science

May 4, 2010

UNC
SCHOOL OF INFORMATION
AND LIBRARY SCIENCE

Contextual Authority Tagging

- Problem
- Background
- Proposed Research
- Methods
- Evaluation
- Pilot Test Data: Example Evaluation

Problem

Locating the topical expert requires topical knowledge.

The one who seeks the expert does not have the knowledge.

- Knowledge management within organizations has focused on expertise location, among other things, usually through the tracking and mining of created documents and artifacts.
- People within organizations use their networks.
- Source selection is critical.
- Need to tap into the *cognitive authority* of those around us (Wilson1983).

Background

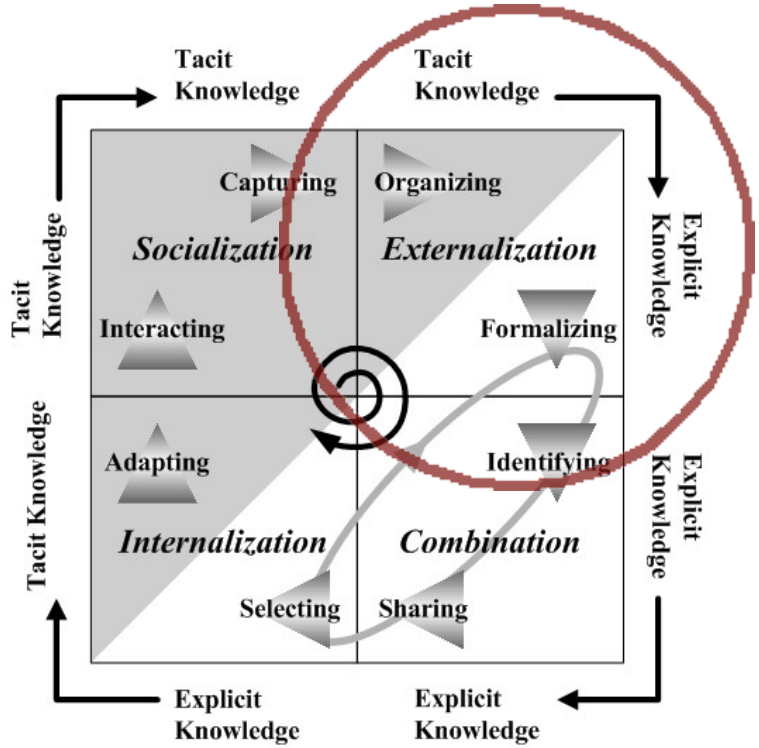
Identity & Reputation

Tagging

Collective Intelligence

Social Epistemology

Expertise



Tacit and Explicit Knowledge

- **Nonaka1991** – Feedback loop, Bringing the tacit into the open is part of the cycle

Background

Knowledge Management

- **Stein1995** – Organizational memory
  - Knowledge acquisition, retention, maintenance, retrieval

- **Dieng1999** – Corporate memory management
  - Detection of needs, knowledge construction, distribution, use, evolution

Background

Source Selection

- **O'Reilly1982** – Quality sources are relevant, timely, specific, and accurate
- **Nilan1988**, **Halpern1988** – Authority, expertise, and trust are the most cited criteria for acceptance/rejection of an information source
- **Rieh2002** – Source credibility depends largely on reliability (reputation, prior work, apparent authenticity)

Background

Expertise Location

- **Abecker1997** – requires multiple sources, integration, little overhead, clear presentation, remaining up-to-date

- **Ehrlich2003** – successful systems must be fast, easy to use, engender trust, scale to entire enterprise, and used by management

Proposed Research

This research will explore the ability of a group to identify the areas of expertise of its members.

- Collective intelligence and distributed cognition of humans
- Visibility of relevant information
- Generation of a positive feedback loop

Proposed Research

Inquiry based on the Delphi Method

- Panel of experts
- Anonymous
- Iterated

Tagging using free-text keywords / labels

Asking a focused, direct question:

- “What are this person's areas of expertise?”
- “What does this person know about?”

Validation through convergence and confidence assessment

Research Question 1

Does CAT work?

- **Similarity** – How similar are a group member's opinion of his/her own areas of expertise and the group's opinion of his/her areas of expertise?

- **Convergence** – How does the similarity behave over time? Do the two opinions converge? If so, how long does it take? If not, is there a persistent gap?

Research Question 2

How acceptable is CAT?

- **Comfort –** How comfortable are group members in participating? What are the main factors influencing their comfort level?
- **Confidence –** How confident are group members in a system like this? What is the quality of the output of this system? Does this system provide a valid credential? Does this system increase users' trust in one another?
- **Usefulness –** What is useful about a system like this? What did participants learn? How would using this system affect participants' decision making?

Proposed Research

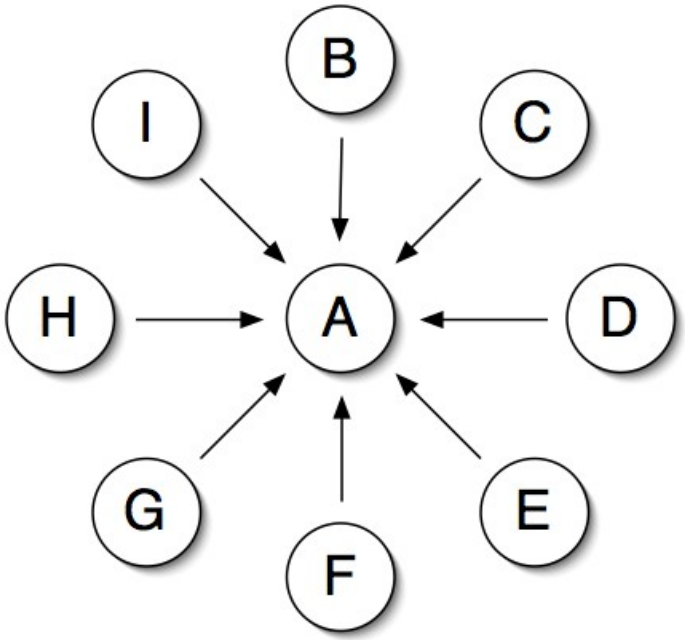
Recruitment → **Experimentation** → **Evaluation**

- 8-10 groups
- 8-10 people each

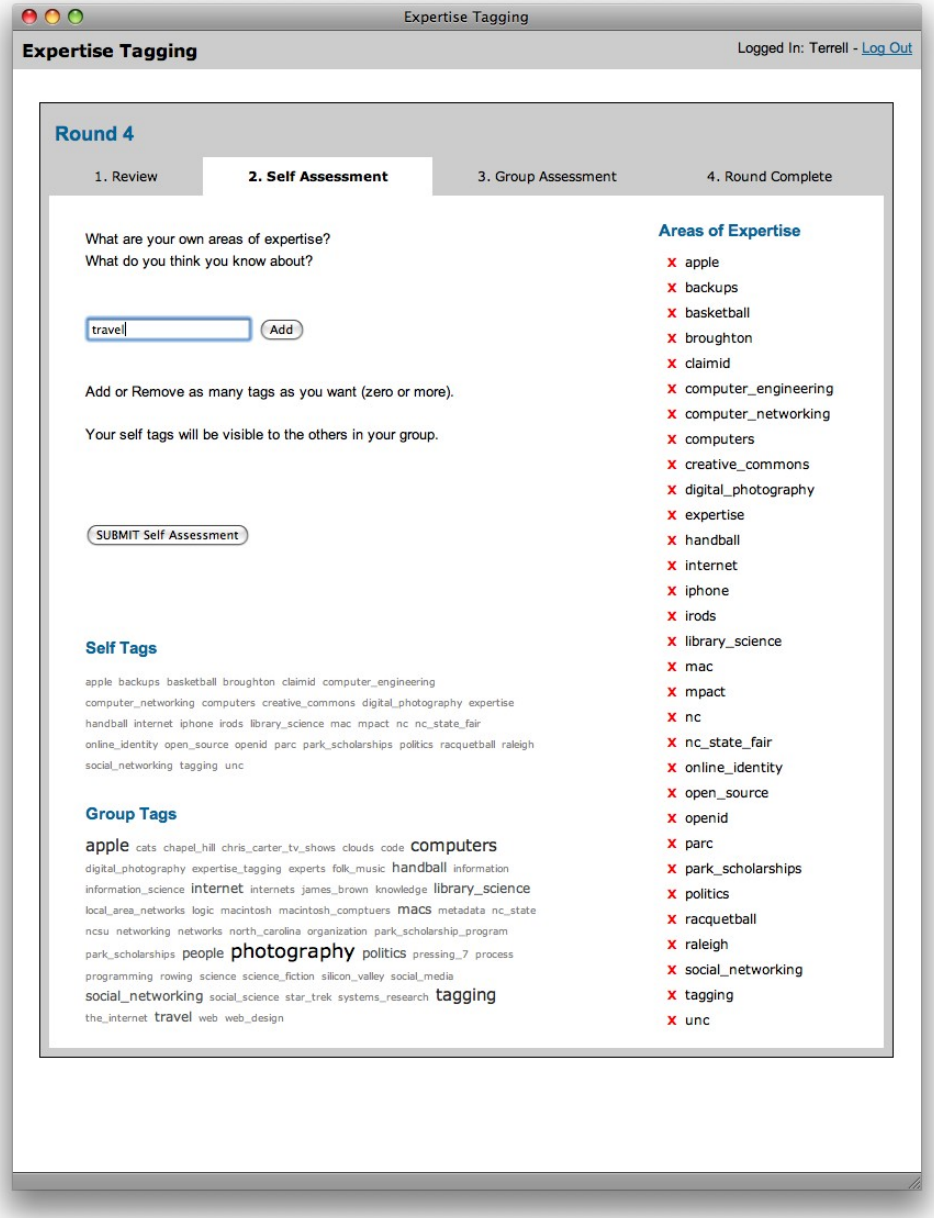
- Survey Pre-Test
- CAT – 5 rounds
- Survey Post-Test
- Follow-up Interview

- Similarity
  - Amazon's Mechanical Turk
  - WordNet Algorithm
- Acceptability
  - Survey
  - Interview

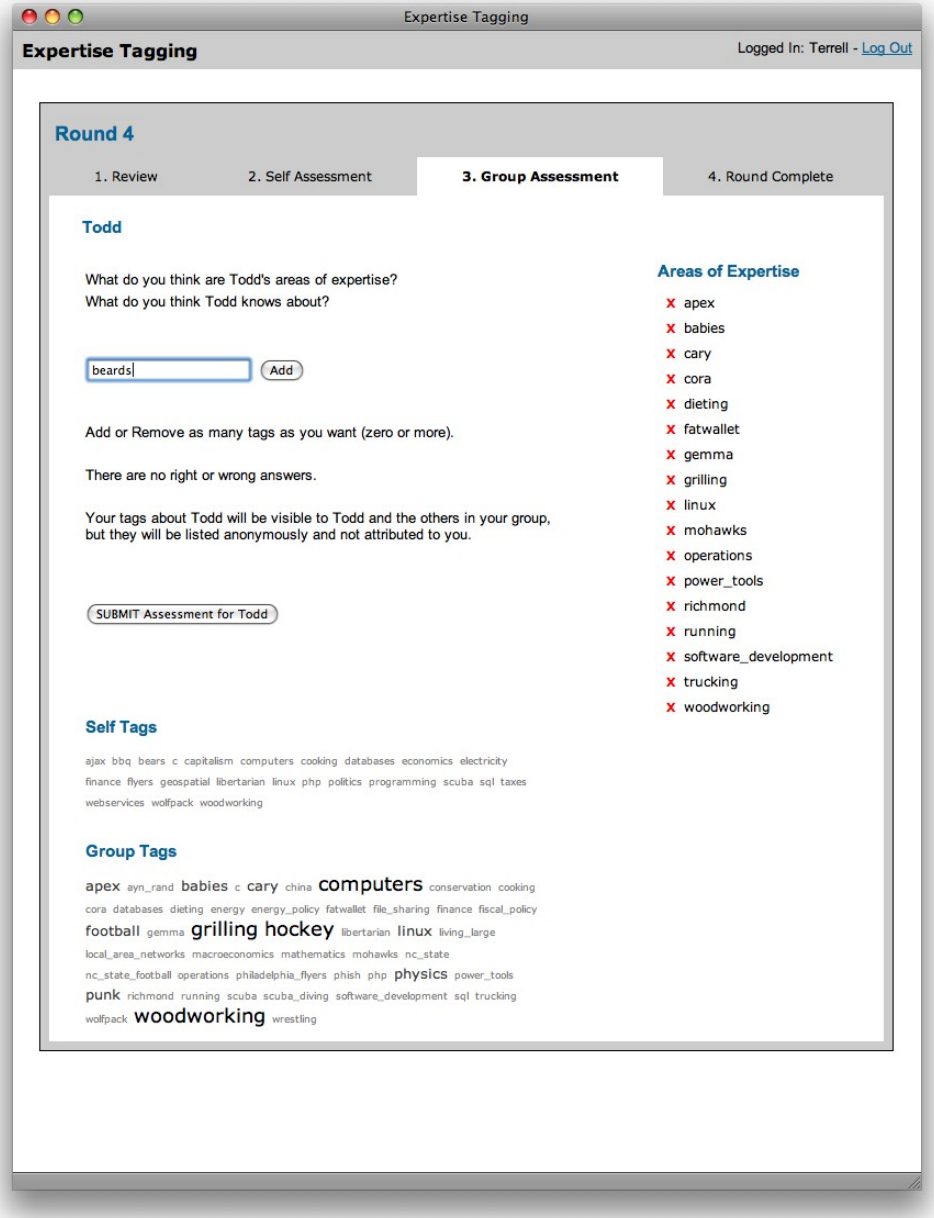
Methods: Contextual Authority Tagging



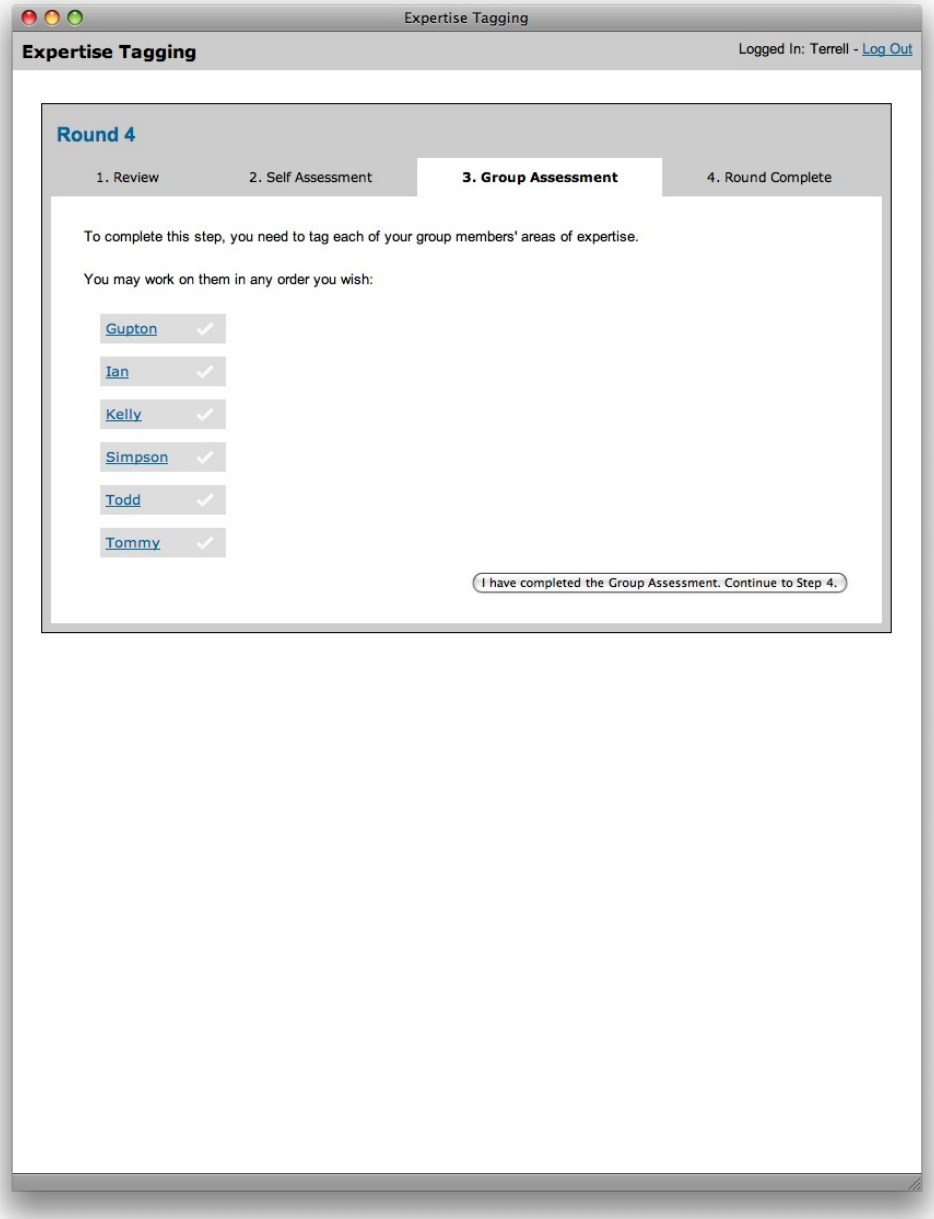
Methods: Contextual Authority Tagging



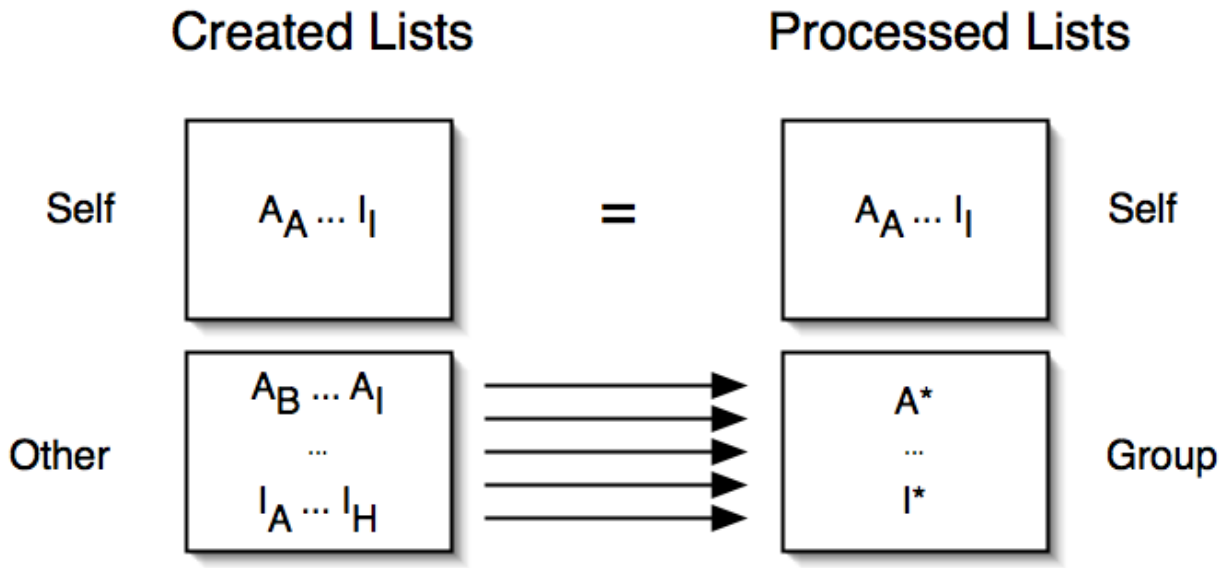
Methods: Contextual Authority Tagging



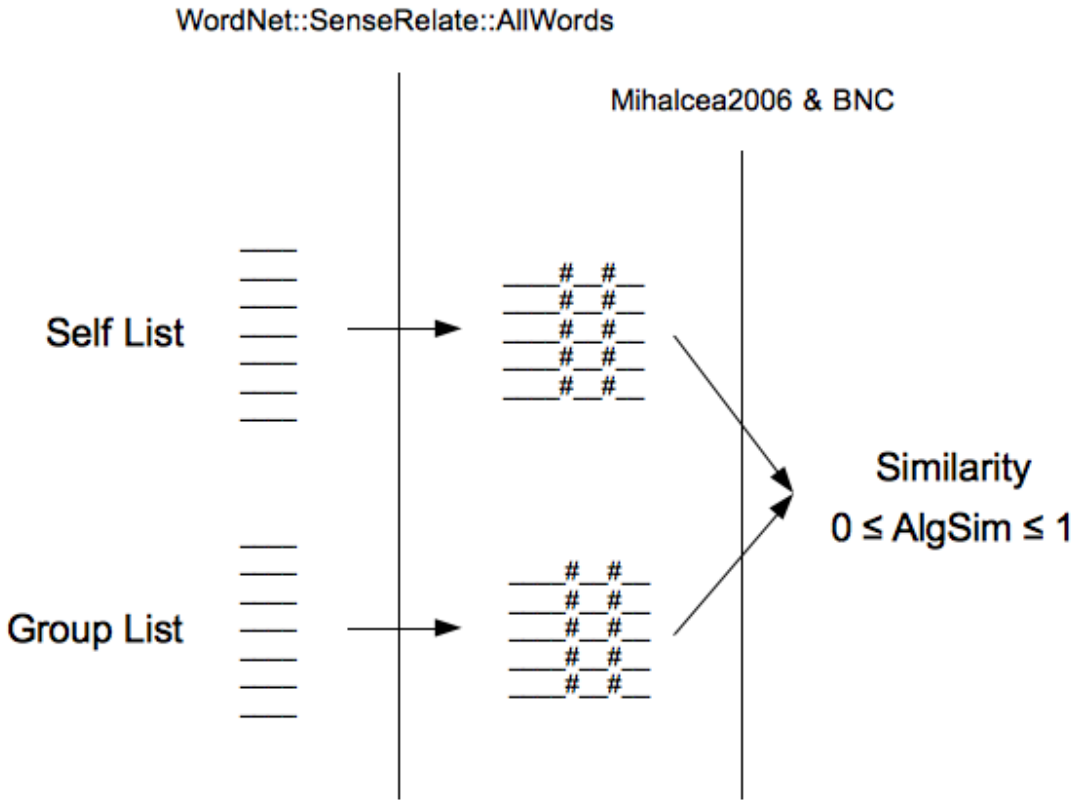
Methods: Contextual Authority Tagging



Methods: Contextual Authority Tagging



Evaluation: Similarity – WordNet



Evaluation: Similarity – WordNet

$$AlgSim(A,B) = \frac{1}{2}\left(\frac{\sum_{w \in \{A\}} (maxSim(w,B) * idf(w))}{\sum_{w \in \{A\}} idf(w)} + \frac{\sum_{w \in \{B\}} (maxSim(w,A) * idf(w))}{\sum_{w \in \{B\}} idf(w)}\right)$$

Mihalcea2006

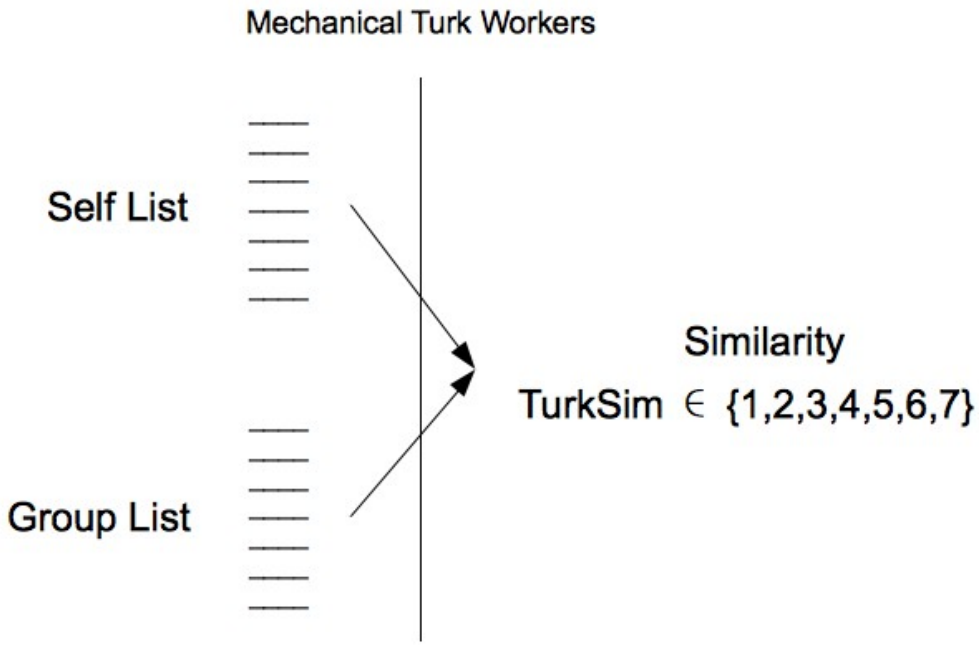
Evaluation: Similarity – WordNet

Comparisons

- self vs group-all
- self vs group-common (2+ occurances)

Each word exists only once in each list.

Evaluation: Similarity – MTurk



"I think these two lists describe similar concepts and ideas."

Evaluation: Similarity – MTurk

7-point Likert

- Extremely Disagree … Extremely Agree

Current Comparisons

- Self-raw vs Group-all-raw
- Self-raw vs Group-common-raw (2+ occurances)
- Self-matching vs Group-all-matching (WordNet identified)
- Self-matching vs Group-common-matching

Possible Comparisons (involving the weighted terms from a Group listing)

- Self-raw vs Group-all-raw-weighted
- Self-raw vs Group-common-raw-weighted
- Self-matching vs Group-all-matching-weighted
- Self-matching vs Group-common-matching-weighted

Evaluation: Similarity & Convergence

XY Plots

- For each user – Similarity graph over time

Box Plots

- For each group – Aggregated similarity graph over time
- For experiment – Aggregated similarity graph over time

ANOVAs to show change between rounds

- Increasing similarity = Convergence

Can also compare and contrast Human vs. Algorithm

Evaluation: Acceptability

Survey

- Likerts are 7-point
- ANOVAs to show differences in Pre-Test/Post-Test
- Existing validated scales will address Research Question 2

Interviews

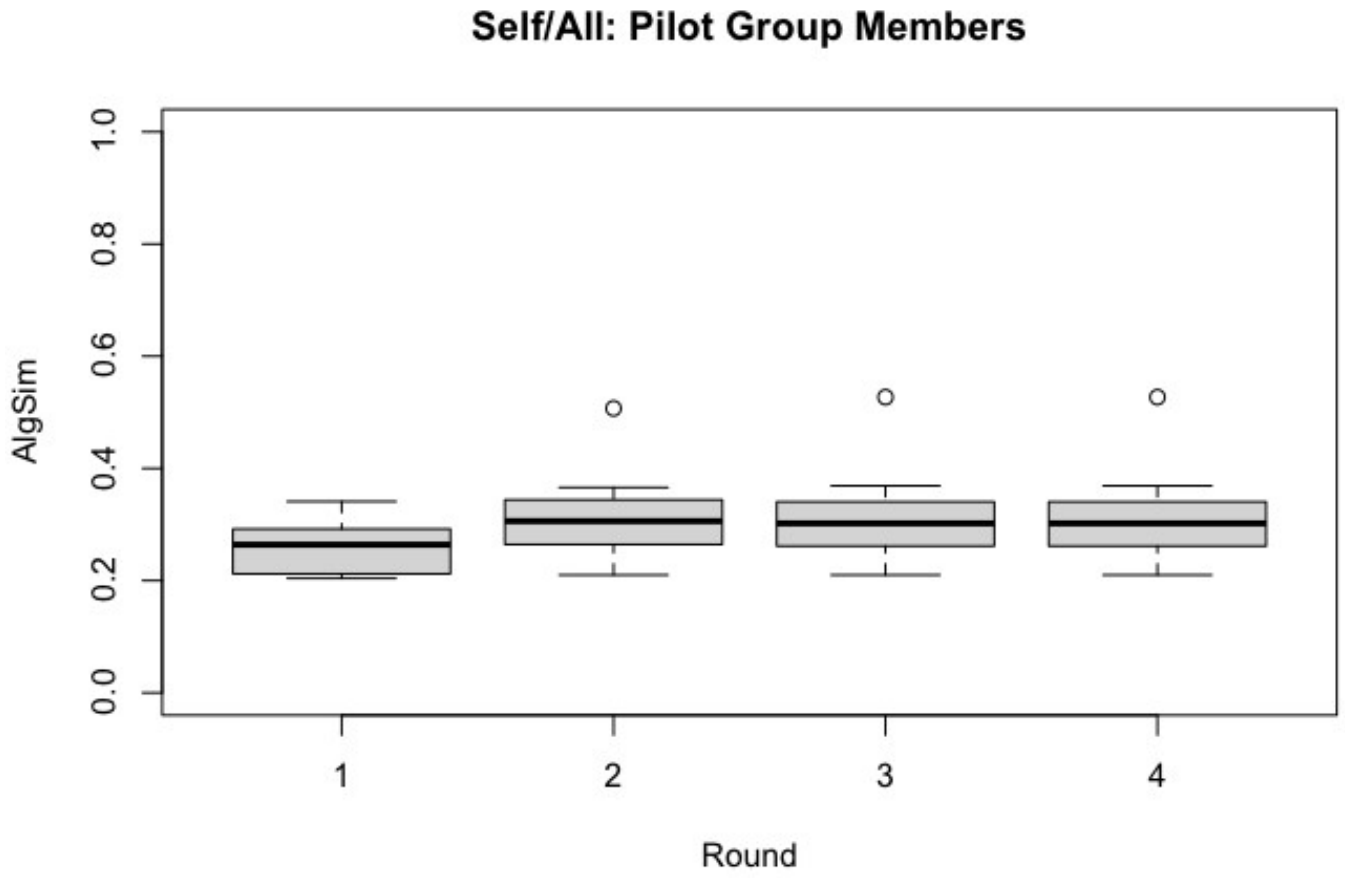
- Selected participants, probably liaisons and others
- Definitely any dropouts, if possible
- Grounded Theory, Open Coding, Inductive
- Will largely address Research Question 2a - Comfort

Pilot Test Data: Example Evaluation

- Friends Dataset
  - 7 friends, 4 rounds
  - Used CAT prototype software

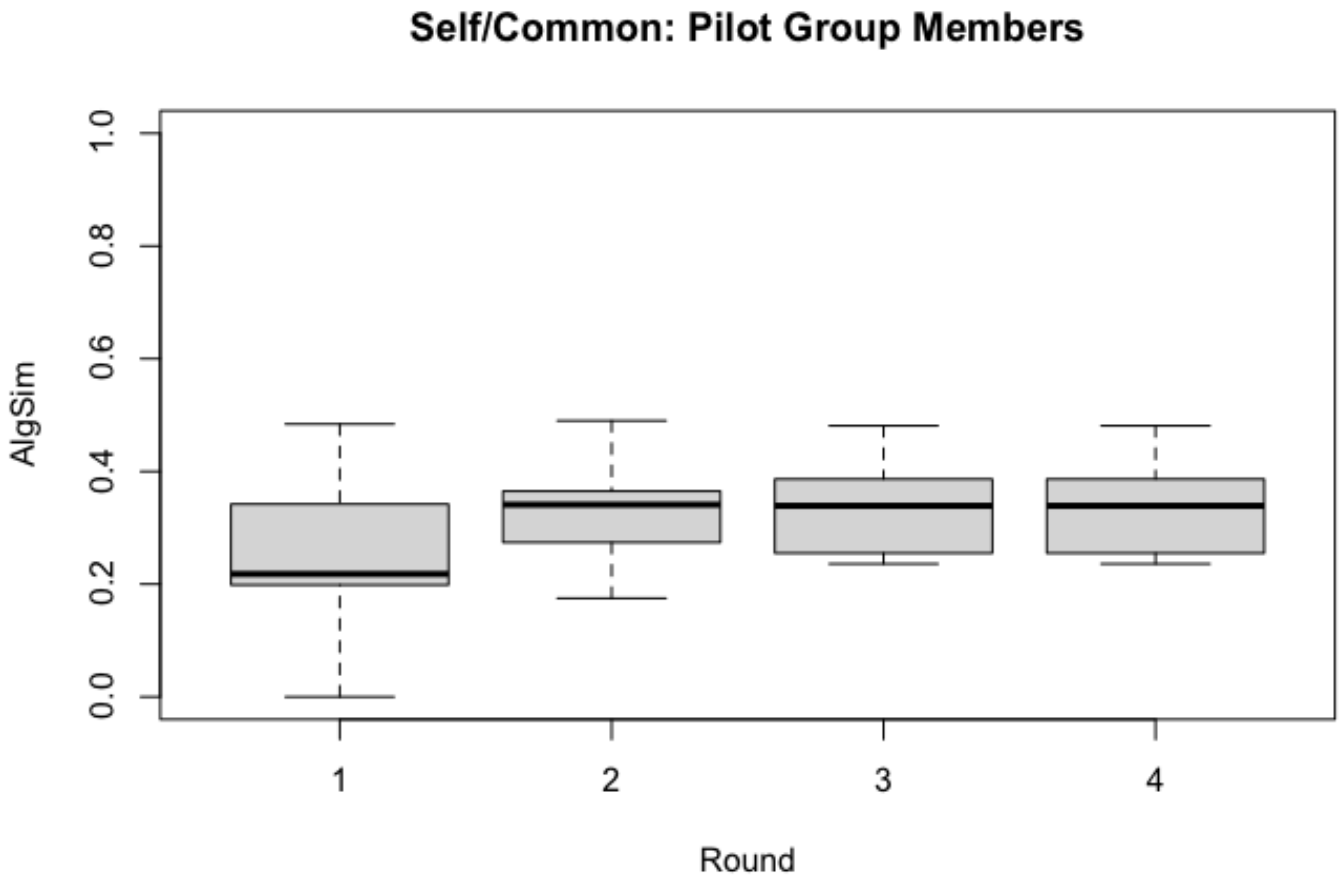
- Evaluation
  - Ran AlgSim
  - Ran TurkSim – Data not complete
  - No Survey or Interviews

Pilot Test Data: AlgSim



N = 7, p = 0.2379

Pilot Test Data: AlgSim



N = 7, p = 0.1971

Contextual Authority Tagging : Expertise Location via Social Labeling

New Technique

Loose Credentialing

Tacit Expertise Location

- Visible
- Up-to-date
- Trusted Collective Opinion

References

Abecker, A., Bernardi, A., Hinkelmann, K., Kühn, O., and Sintek, M. (1997). Towards a well-founded technology for organizational memories. American Association for Artificial Intelligence Technical Report, SS-97-01.

Dieng, R., Corby, O., Giboin, A., and Ribiére, M. (1999). Methods and tools for corporate knowledge management. International Journal of Human-Computer Studies, 51(3):567–598.

Ehrlich, K. (2003). Locating expertise: Design issues for an expertise locator system. In Sharing Expertise: Beyond Knowledge Management. MIT Press.

Halpern, D. and Nilan, M. S. (1988). A step toward shifting the research emphasis in information science from the system to the user: An empirical investigation of source-evaluation behavior, information seeking and use. Proceedings of the American Society of Information Science.

Mihalcea, R., Corley, C., and Strapparava, C. (2006). Corpus-based and knowledge-based measures of text semantic similarity. In Proceedings of the National Conference on Artificial Intelligence, volume 21, pages 775–780. AAAI.

Nilan, M. S., Peek, R. P., and Snyder, H. W. (1988). A methodology for tapping user evaluation behaviors: An exploration of users' strategy, source and information evaluating. Proceedings of the American Society of Information Science.

Nonaka, I. (1991). The knowledge-creating company. Harvard Business Review, 69(6):96–104.

O'Reilly, C. A. (1982). Variations in decision makers' use of information sources: The impact of quality and accessibility of information. The Academy of Management Journal, 25(4):756–771.

Rieh, S. Y. (2002). Judgement of information quality and cognitive authority in the web. Journal of the American Society for Information Science and Technology, 53(2):145–161.

Stein, E. W. (1995). Organization memory: Review of concepts and recommendations for management. International Journal of Information Management, 15(1):17–32.

Wilson, P. (1983). Second-hand knowledge: An inquiry into cognitive authority. Greenwood Press, Westport, CT.