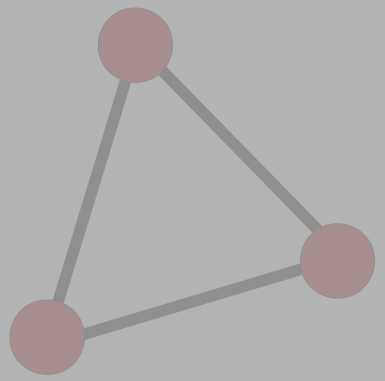
Contextual Authority Tagging: Expertise Location via Social Labeling



Terrell G. Russell
Dissertation Defense

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

Wednesday, April 27, 2011

Acknowledgements

Dr. Deborah Barreau, advisor

Dr. Gary Marchionini, advisor

Dr. Barbara M. Wildemuth, committee member

Dr. Sri Kalyanaraman, committee member

Dr. Phillip J. Windley, committee member

Stephen and Corinne Russell

Kelly Marks

Many Friends and Fellow Ph.D. Students

SILS Faculty and Staff

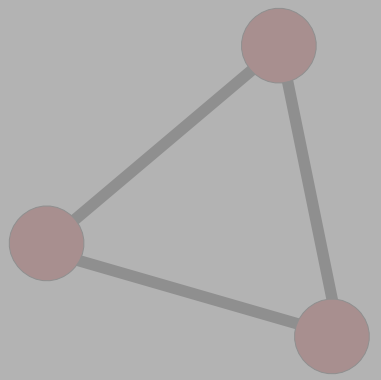
- **Motivation**

Proposal

Methodology

Findings

Summary



Flood of Information

- Internet has
  - democratized speech like never before
  - facilitated a flood of new information sources
  - created a need for better filtering

- We filter based on
  - history
  - credentials
  - reputation
  - ...

Flood of Information

- Similar, smaller scale problem exists within organizations
  - many people
  - many projects
  - changing details over time

We have a need for knowing

who knows what

Flood of Information

- Existing expertise location software systems are based on
  - self report
  - exhaust documents and/or activity

and do not necessarily

capture the opinions of others

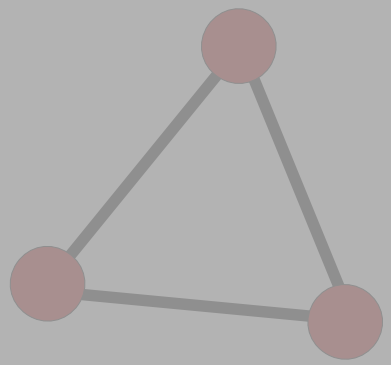
Motivation

- **Proposal**

Methodology

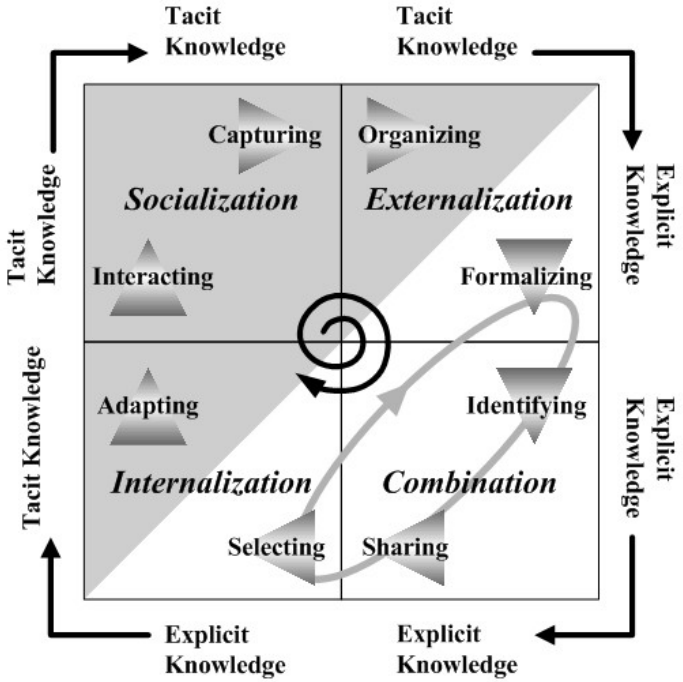
Findings

Summary



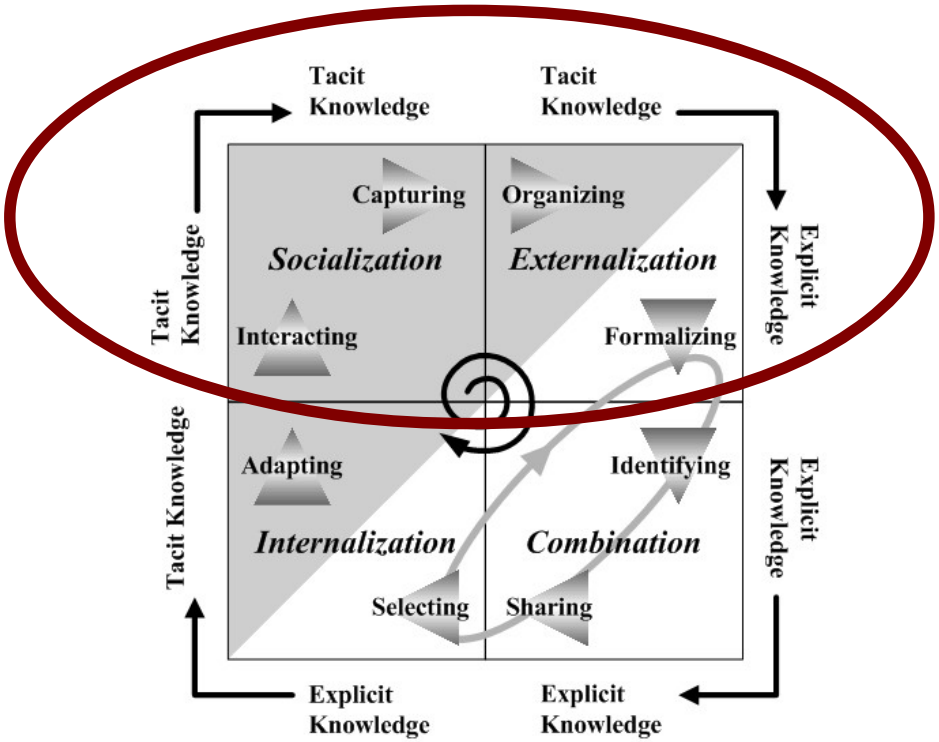
Contextual Authority Tagging

Nonaka, 1994



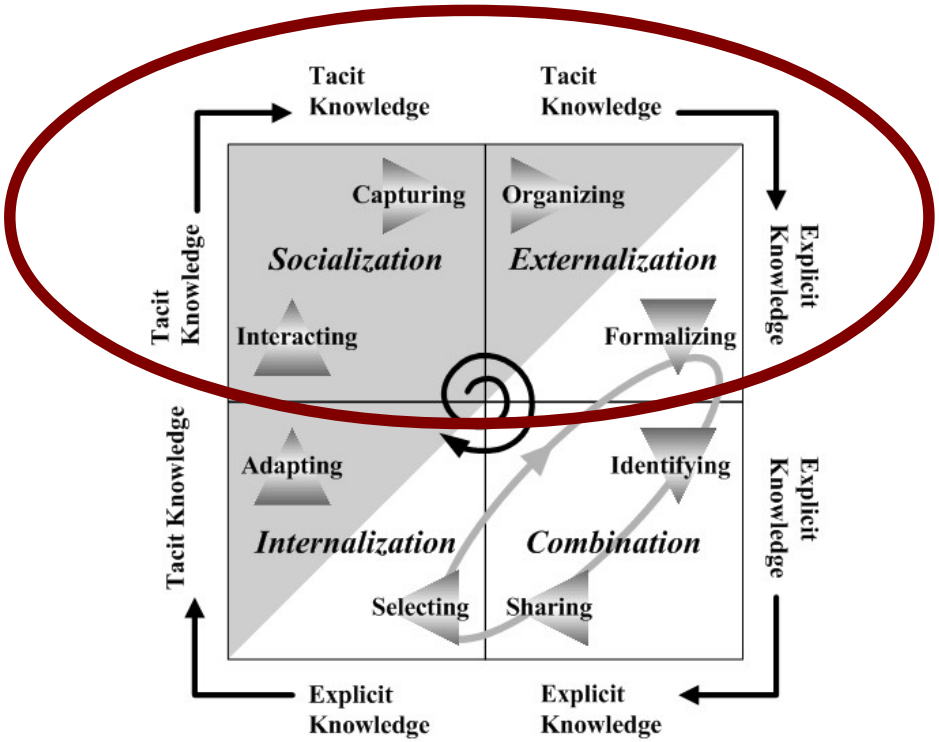
Contextual Authority Tagging

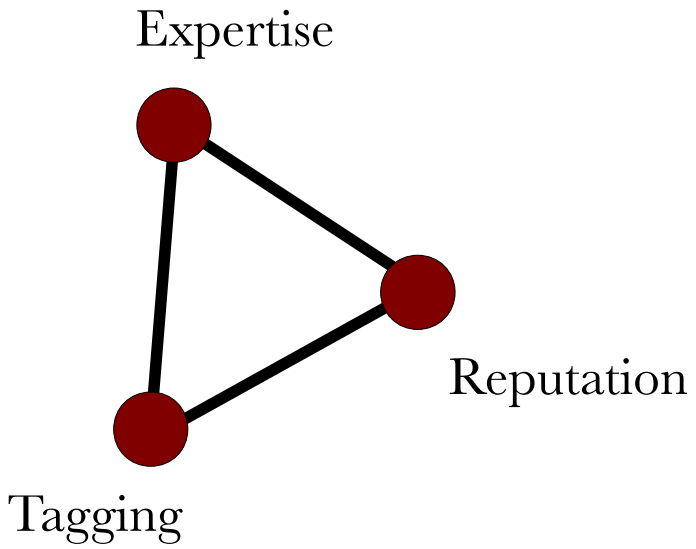
Nonaka, 1994



Contextual Authority Tagging

Nonaka, 1994





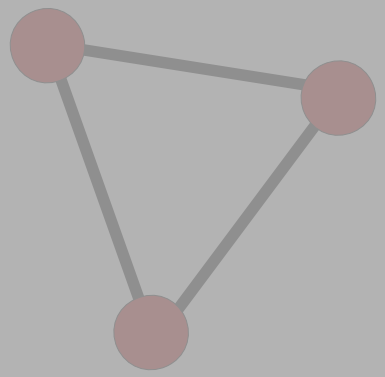
Motivation

Proposal

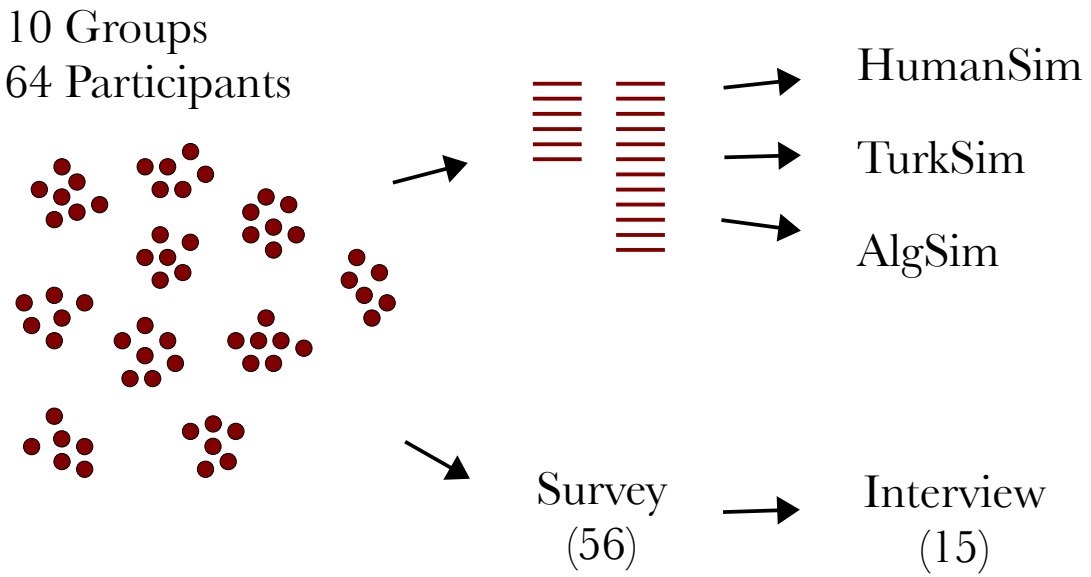
- **Methodology**

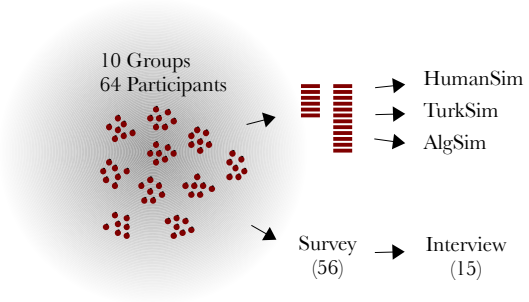
Findings

Summary



Study Design





Generating Tags

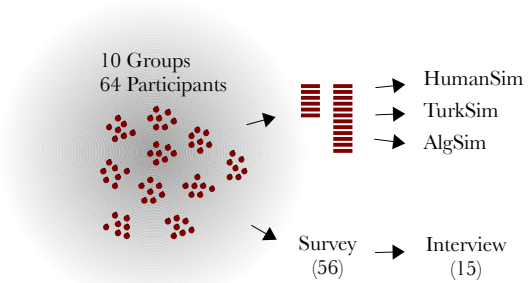
- Delphi
  - Experts
  - Anonymous
  - Iterated

Goal is to triangulate on a subjective truth.

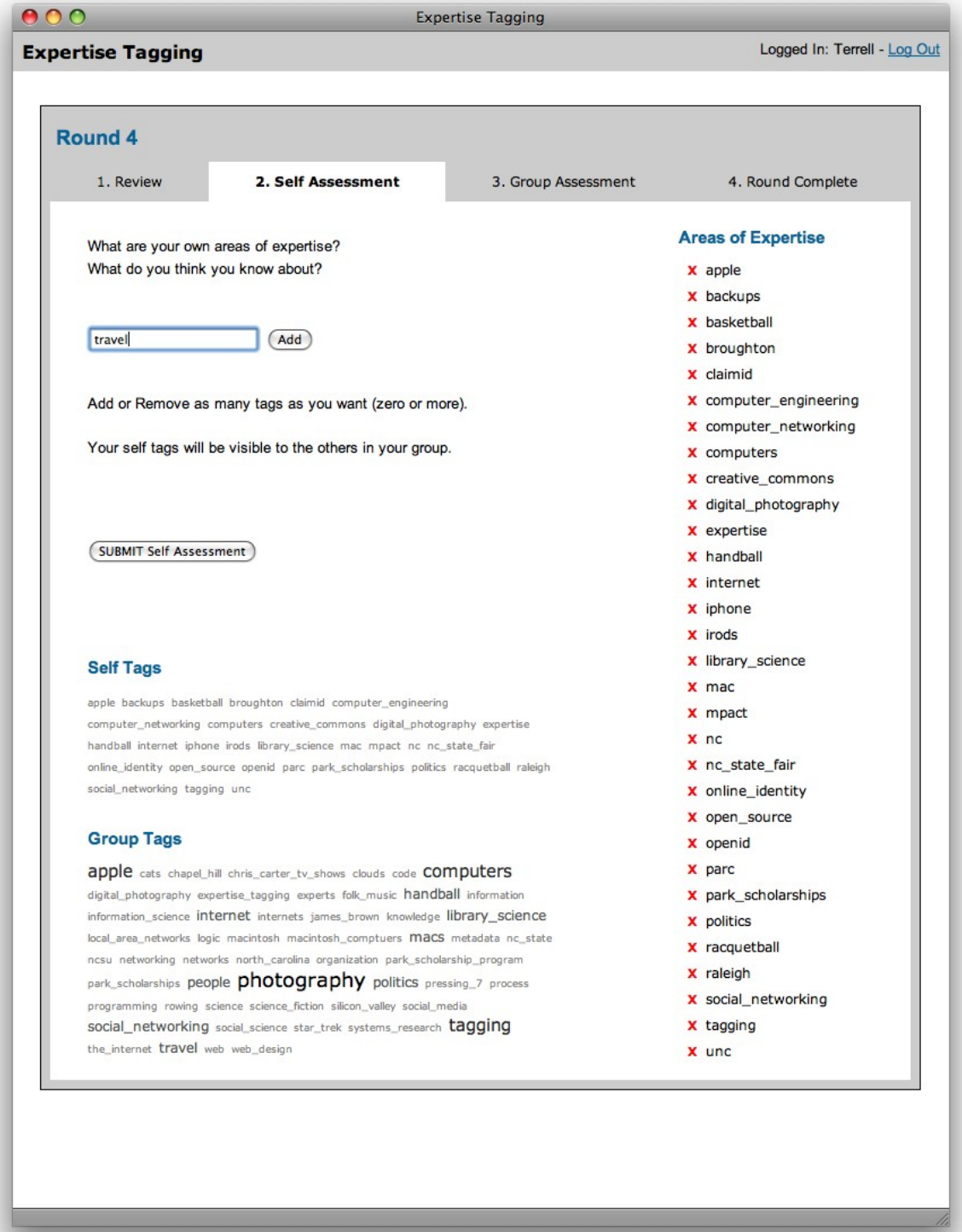
- Modified Delphi
  - Group Members
  - Unattributed
  - 5 Rounds

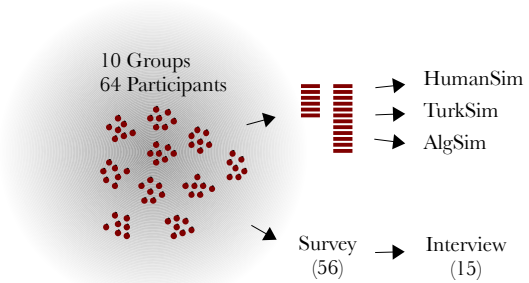
Goal is to collectively label members' areas of expertise.

Helmer and Rescher, 1959

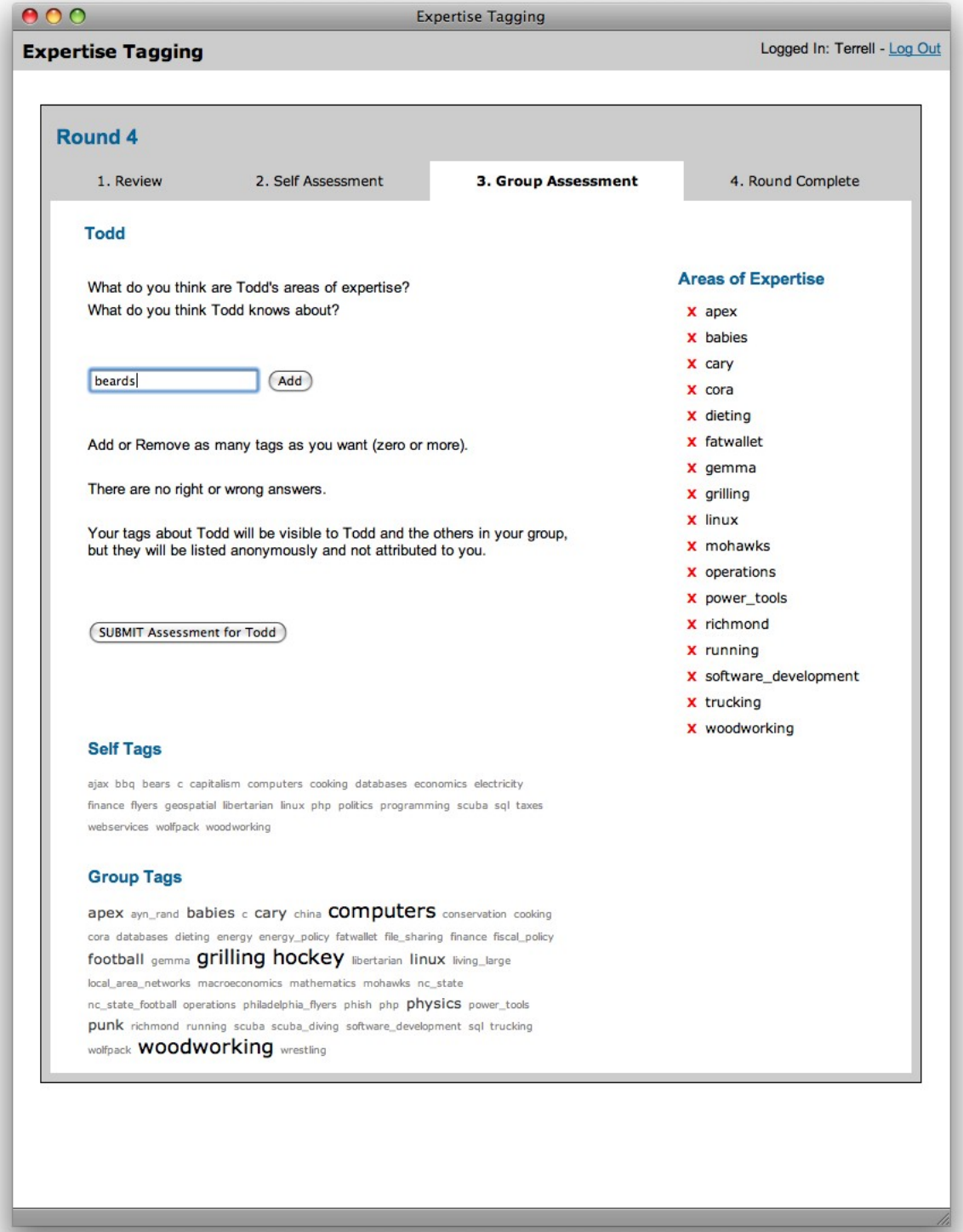


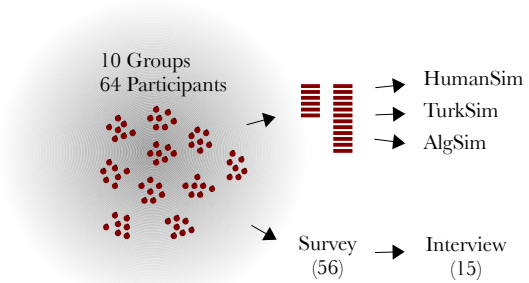
Generating Tags





Generating Tags





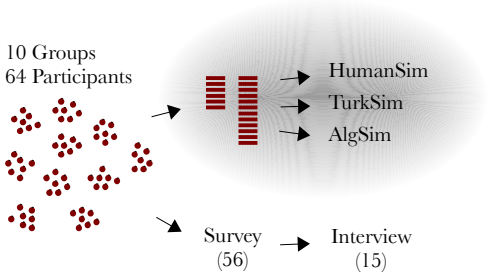
Generating Tags

Expertise Tagging

	Self	Group
Kelly	art crafts design fabric fonts literature popular_culture webby_things	**4** design **4** pop_culture **3** art **3** journalism **2** crafts **2** fabric **2** fonts **2** internet **2** park_scholarships **2** politics **2** quilting (Show Single Tags)

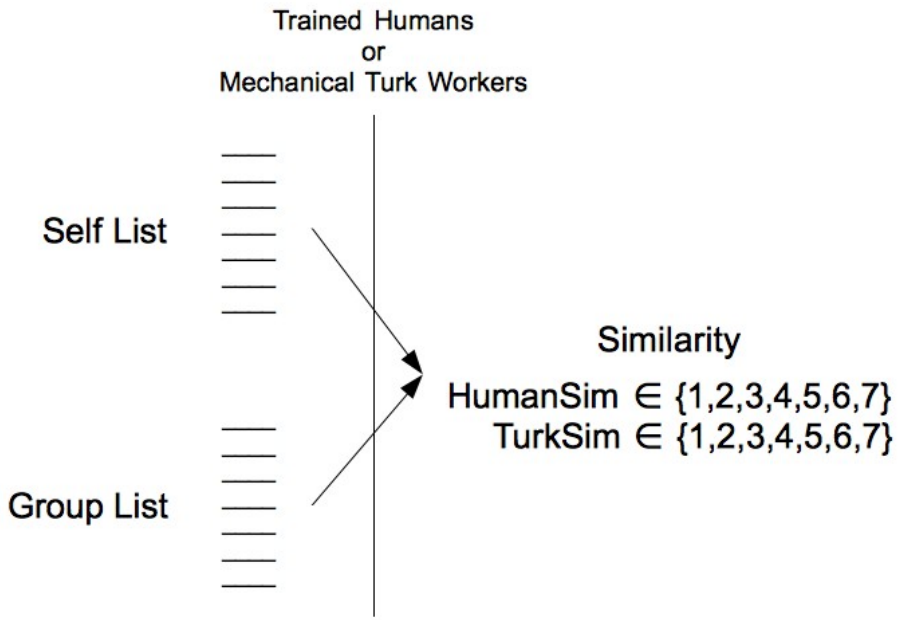
	Self	Group
Simpson	asia c china computers dinosaurs games linear_algebra linux lisp mandarin mma philosophy probability reverse_engineering science taiji taiwan tea videogames	**4** china **3** c **3** linux **3** skateboarding **2** economics **2** graphics **2** hacking **2** libertarian **2** punk **2** tai_chi **2** taiwan **2** thailand (Show Single Tags)

	Self	Group
Todd	ajax bbq bears c capitalism computers cooking databases economics electricity finance	**3** computers **3** grilling **3** hockey **3** woodworking **2** apex **2** babies **2** cary **2** football **2** linux **2** physics **2** punk

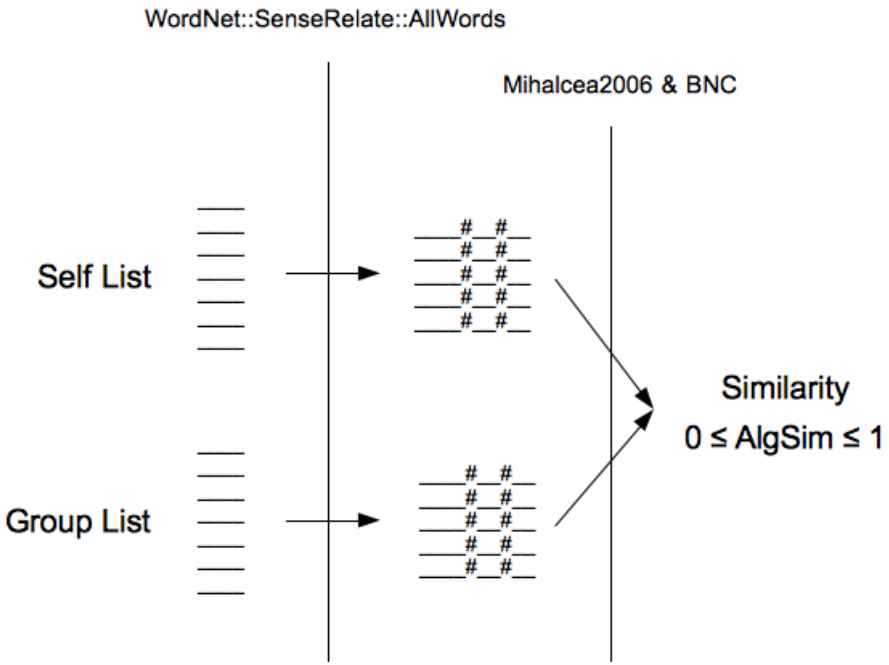


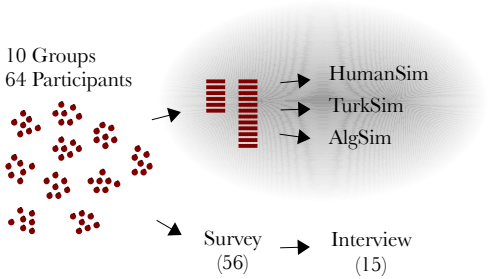
Calculating Similarity

HumanSim and TurkSim



AlgSim





Calculating Similarity

HIT Preview

Similarity Rating

The following two lists of words come from different sources.

They were generated in two different ways and one list may have more words than the other.

We are interested in how similarly they describe the same concepts and ideas.

Please examine these two lists of words:

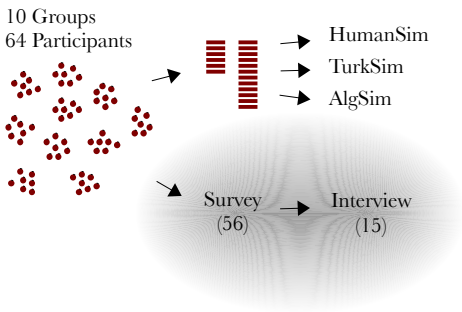
fabric	journalism
popular_culture	art
crafts	park_scholarships
literature	politics
webby_things	fabric
design	fonts
art	design
fonts	crafts
	pop_culture
	quilting
	internet

Please rate your level of agreement with the following statements:

I think these two lists describe similar concepts and ideas.

Extremely Disagree	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree	Extremely Agree
○	○	○	○	○	○	○

Submit



Gathering Sentiment

- Survey
  - All Participants
  - Estimated 95% participation
- Interview
  - Self-selected for further discussion
  - Estimated 10% participation

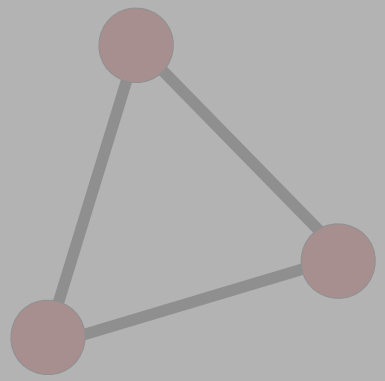
Motivation

Proposal

Methodology

- **Findings**

Summary



Demographics

10 groups consisting of 64 participants

Group	Interaction	Primary Employment	Location
family retail business	daily	yes	physical
dentist's office	daily	yes	physical
distributed software development	daily	no	virtual
distributed software development	daily	yes	virtual
museum education staff	daily	yes	physical
writer's network	not daily	no	virtual
legal non-profit	not daily	no	physical
global engineering firm	daily	yes	physical
academic faculty	daily	yes	physical
academic administrative office	daily	yes	physical

Age	Responses	%
21-30	16	28.6%
31-40	15	26.8%
41-50	9	16.1%
51-60	7	16.1%
Over 60	7	12.5%
Total	56	100%

Sex	Responses	%
M	24	43%
F	31	55%
N/A	1	2%
Total	56	100%

Time in Group	Responses	%
Less than 6 months	7	12.5%
6-12 months	7	12.5%
1-3 years	13	23.2%
3-5 years	9	16.1%
More than 5 years	20	35.7%
Total	56	100%

Research Questions

R1. Does CAT Work?

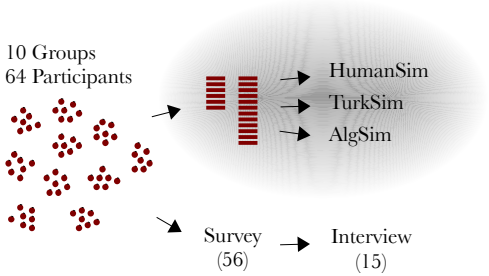
R2. How acceptable is CAT?

Research Question 1

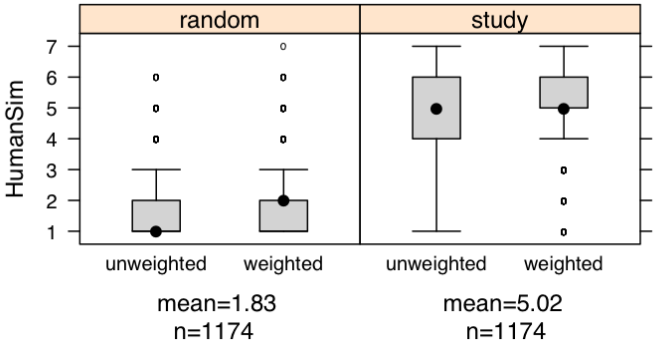
Does CAT Work?

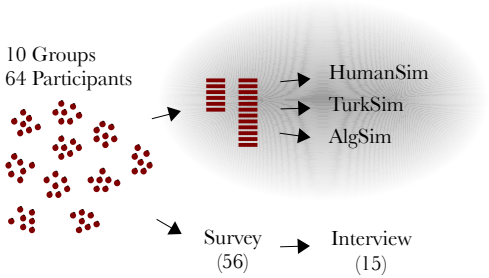
(a) **Similarity** – How similar are a group member's opinion of his/her own areas of expertise and the group's opinion of his/her areas of expertise?

(b) **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?

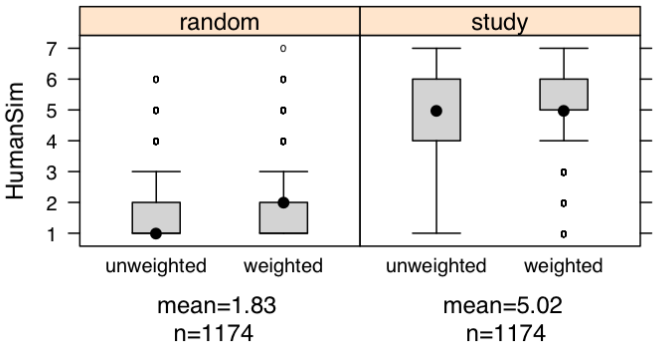


Similarity and Convergence

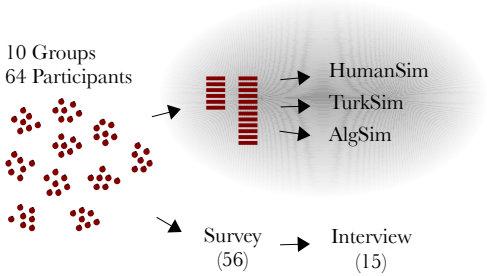




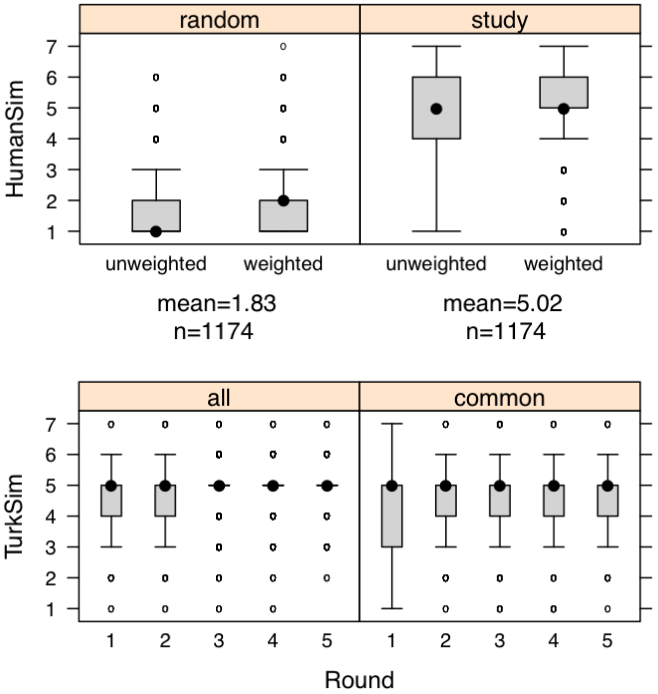
Similarity and Convergence



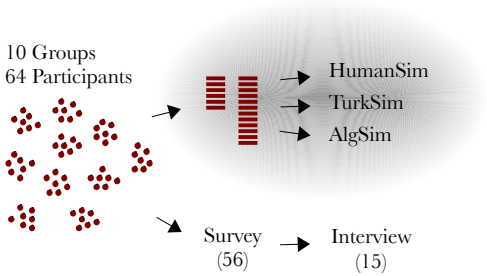
- Humans can differentiate CAT pairings from random pairings
- Humans rate self/group CAT pairings about a person as similar



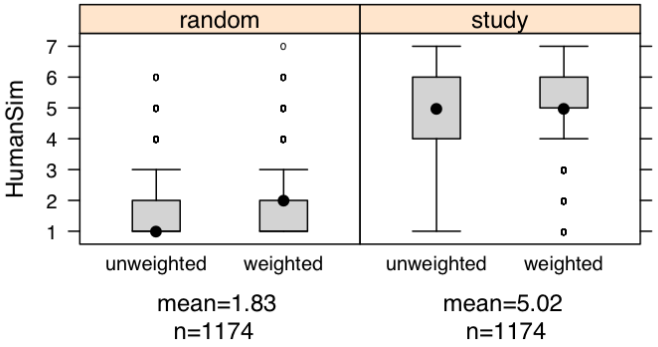
Similarity and Convergence



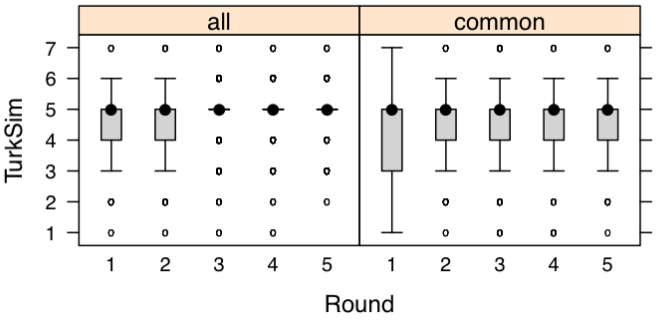
- Humans can differentiate CAT pairings from random pairings
- Humans rate self/group CAT pairings about a person as similar



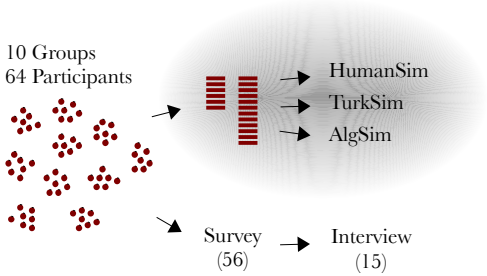
Similarity and Convergence



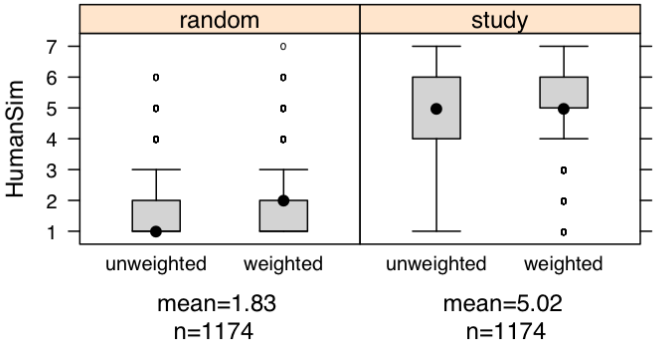
- Humans can differentiate CAT pairings from random pairings
- Humans rate self/group CAT pairings about a person as similar



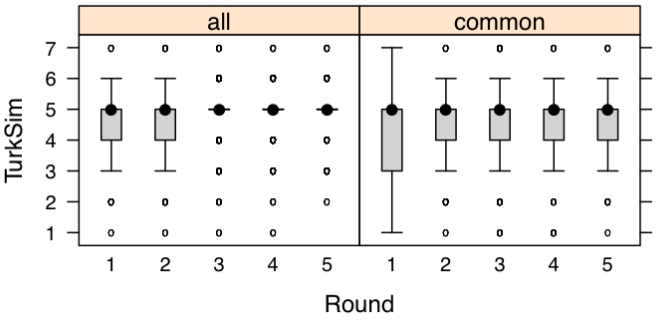
- Turkers can differentiate CAT pairings from random pairings
- Turkers rate self/group CAT pairings about a person as similar
- Turker-rated similarity ratings decrease in variability over time



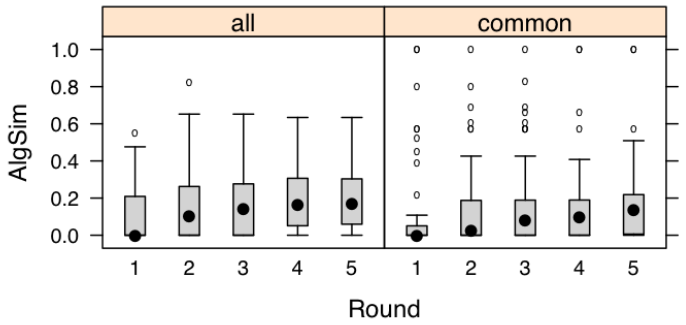
Similarity and Convergence

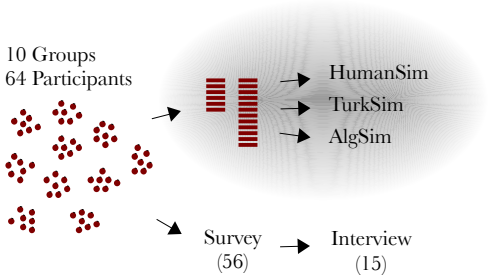


- Humans can differentiate CAT pairings from random pairings
- Humans rate self/group CAT pairings about a person as similar

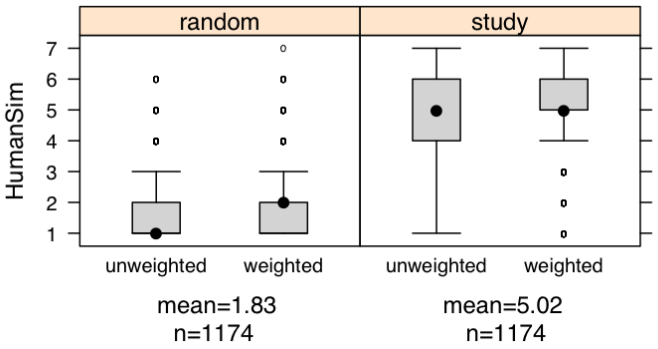


- Turkers can differentiate CAT pairings from random pairings
- Turkers rate self/group CAT pairings about a person as similar
- Turker-rated similarity ratings decrease in variability over time

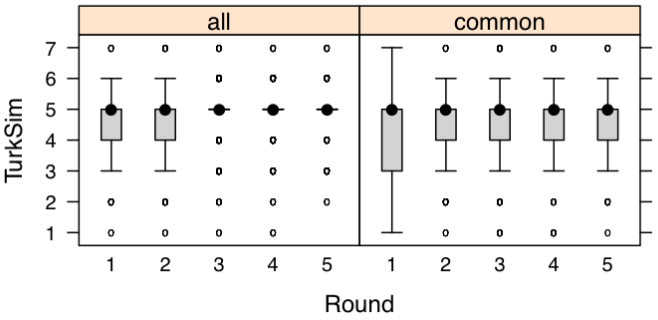




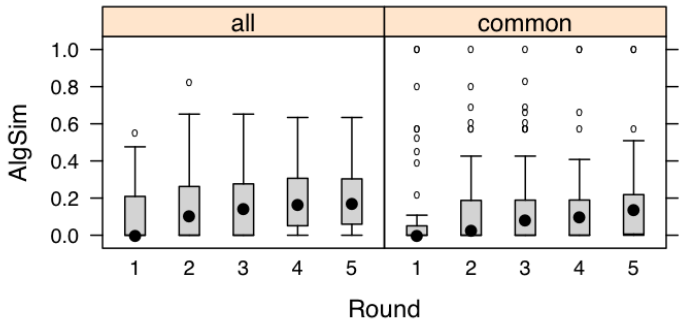
Similarity and Convergence



- Humans can differentiate CAT pairings from random pairings
- Humans rate self/group CAT pairings about a person as similar



- Turkers can differentiate CAT pairings from random pairings
- Turkers rate self/group CAT pairings about a person as similar
- Turker-rated similarity ratings decrease in variability over time



- Algorithm can differentiate CAT pairings from random pairings
- Algorithm rates self/group CAT pairings about a person as non-zero
- Algorithm similarity ratings increase over time
- Algorithm similarity ratings level off after initial round(s)

Research Question 1

Does CAT Work?

(a) **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?

(b) **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?

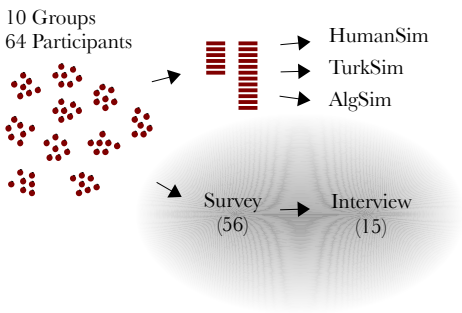
Research Question 2

How acceptable is CAT?

(a) **Comfort** – How comfortable are group members in participating? What are the main factors influencing their comfort level?

(b) **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?

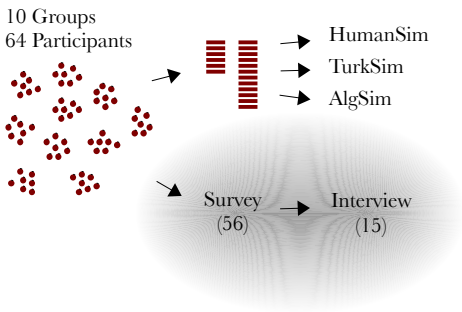
(c) **Usefulness** – What is useful about a system like this? What did participants learn? How would using this system affect participants' decision making?



Survey Responses

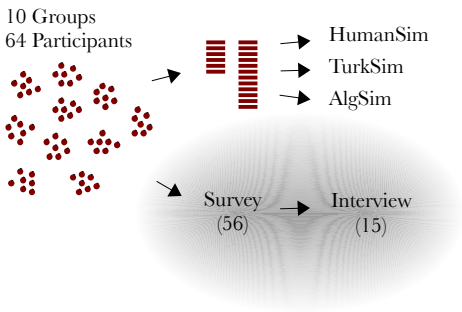
(1-7 scale)

Original Items	Average Rating
I am comfortable with my group's tags about my areas of expertise.	5.439
I am happy with my group's tags about my areas of expertise.	5.351
I am familiar with my group members' areas of expertise.	5.333
This was an interesting exercise.	5.196
My group members are familiar with my areas of expertise.	5.175
My group did not list important areas of my expertise.	4.764
I am confident that this system gives me new information.	4.696
This was a useful exercise.	4.679
I am confident that this system gives me good information.	4.643
I am willing to incorporate output from this system into my decision making.	4.607
I would be more comfortable with my group's tags if the tags were not anonymous.	3.298
Scales	**Average Rating**
Data Quality	4.709
Effort Expectancy	4.670
Result Demonstrability	4.299
Facilitating Conditions	4.250
Performance Expectancy	3.836
Relative Advantage	3.742
Anxiety (reverse coded)	3.036 (4.964)



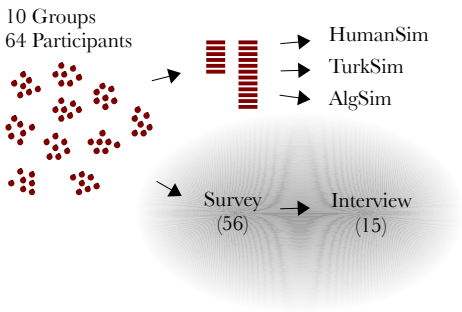
Survey Responses

Favorite Part	Count
thinking about specific strengths of others	12
what people thought of me	11
more awareness	8
seeing others' self claims	8
how others see others	7
good to reconnect	2
self assessment	2
making connections / learning about others	2
thinking about friends / uplifting / feel better	2
non-job related interests	2
not time consuming	1
similarity and consensus	1
got to know people faster	1
tag clouds of expertise	1
the challenge of listing explicitly	1
help learn about colleagues, otherwise limited contact	1



Survey Responses

Least Favorite Part	Count
redundancy of multiple rounds (3 was enough)	29
nothing disliked	4
yet another email / feeling tardiness	2
talking about myself / "not very modest"	2
non-uniformity of terms	2
phrasing of tags is hard	2
everyone has a different view	1
no semantic equivalence	1
fear of future reduced group dynamics because of exclusion	1
defining "expertise"	1
trying to determine whether someone was an expert	1
when others did not reciprocate	1
vulnerability	1
stressful	1
nervous	1
realizing I know very little about 3 group members	1
concern over "doing it wrong"	1
being asked if I was sure	1
could not go back and modify	1
entering passcodes manually	1



Interview Responses

"I need to be better about promoting"

"it would be more beneficial if we talked about it as an office"

"wanted something more at the end"

"I want people to know more about what I'm doing"

"helpful"

talking about oneself was "weird", "awkward", or "advertisey"

"really interesting"

"learned a bit about how I like to be viewed by others"

Research Question 2

How acceptable is CAT?

(a) **Comfort** – How comfortable are group members in participating? What are the main factors influencing their comfort level?

(b) **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?

(c) **Usefulness** – What is useful about a system like this? What did participants learn? How would using this system affect participants' decision making?

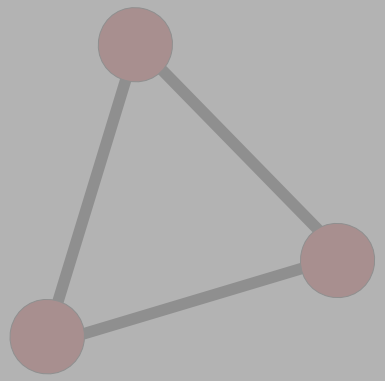
Motivation

Proposal

Methodology

Findings

- **Summary**



Conclusions

- CAT succeeds in identifying the areas of expertise of group members.
- CAT provides insight that is most relevant to group members who are not as “established” (i.e. new members).
- CAT is complementary and should be deployed alongside or integrated into existing knowledge management infrastructure.
- CAT needs to be accompanied by guidelines for interpretation. Raw data is not enough.

Limitations

- only 10 groups, 64 participants
- small groups with well-known members
- recruiting and the self-selection of groups
- simple algorithm
- WordNet database
- subject level expertise of similarity raters

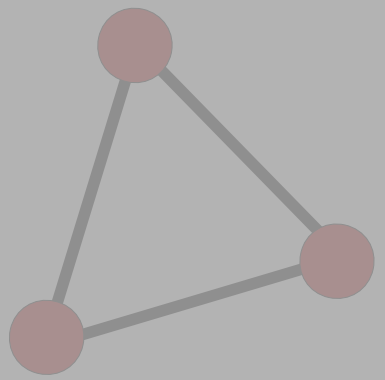
Contributions

- A validated, relatively inexpensive method for generating quality expertise assessments from group members
- An automated Modified Delphi study

Future Work

- Larger groups/organizations
  - Multiple groups within a single organization
- Incorporation with existing personnel tools
- Open Internet
  - Attribution
  - Weighting
  - Incentivization
  - Recursion

Thank You



References

- Helmer, O. and Rescher, N. (1959). On the epistemology of the inexact sciences. *Management Science*, 6(1):25–52.

- 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.

- Moore, G. C. and Benbasat, I. (1991). Development of an instrument to measure the perceptions of adopting an information technology innovation. *Information Systems Research*, 2(3):192–222.

- Nonaka, I. (1994). A dynamic theory of organizational knowledge creation. *Organization Science*, 5(1):14–37.

- Venkatesh, V., Morris, M. G., Davis, G. B., and Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. *Management Information Systems Quarterly*, 27(3):425–478.

- Wang, R. Y. and Strong, D. M. (1996). Beyond accuracy: What data quality means to data consumers. *Journal of Management Information Systems*, 12(4):5–34.

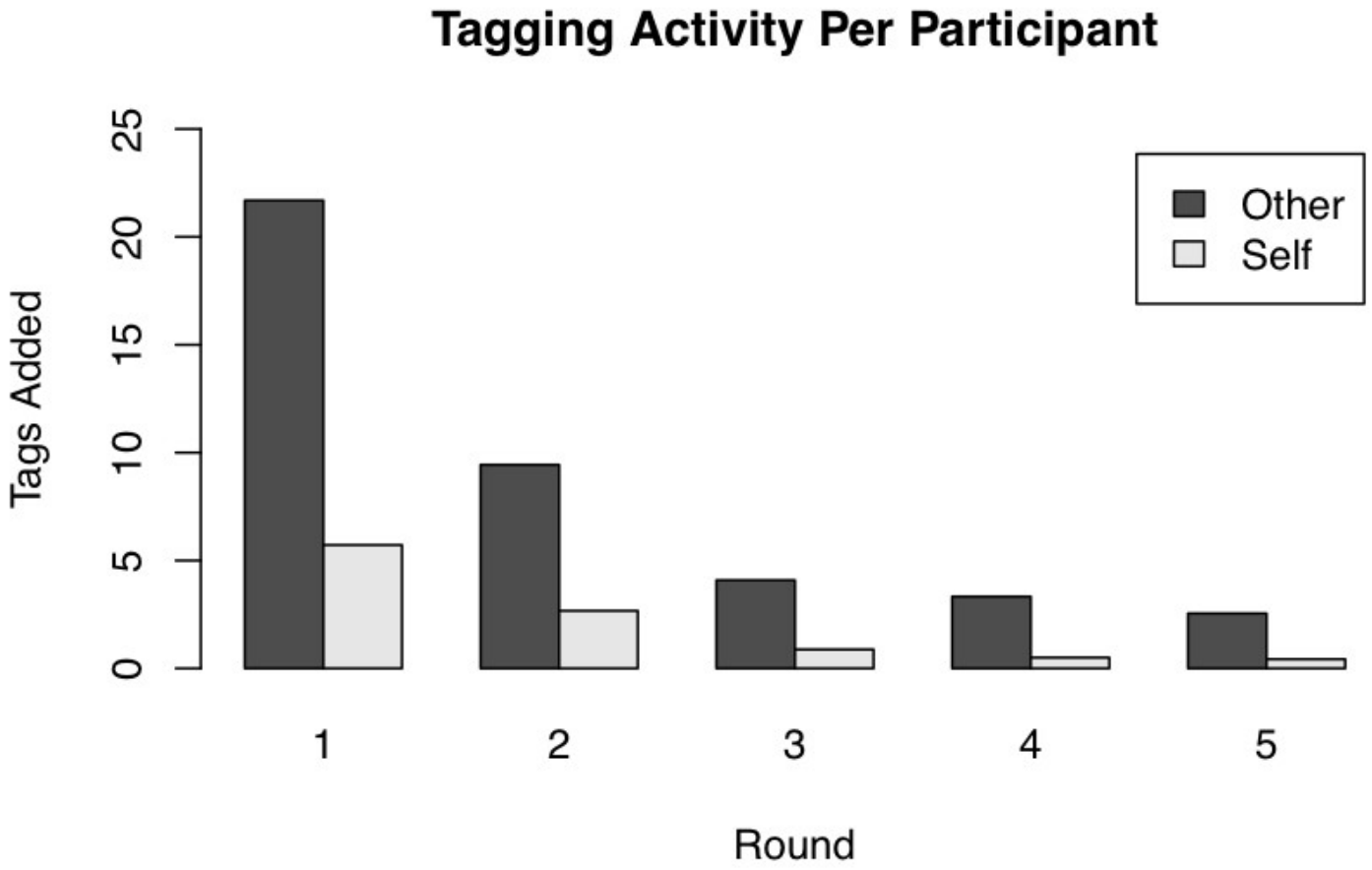
Survey Items from Selected Scales

4.709	**Data Quality (Wang and Strong, 1996)**
	This system produced data in conformance with the actual or true values.
	This system produced data that is applicable and relevant to my job.
	This system produced data that is intelligible and clear.
	This system produced data that is easily accessible.
4.670	**Effort Expectancy (Venkatesh et al, 2003)**
	My interaction with this system would be clear and understandable.
	It would be easy for me to become skillful at using this system.
	I would find this system easy to use.
	Learning to operate this system would be easy for me.
4.299	**Result Demonstrability (Moore and Benbasat, 1991)**
	I would have no difficulty telling others about the results of using this system.
	I believe I could communicate to others the consequences of using this system.
	The results of using this system are apparent to me.
	I would have difficulty explaining why using this system may or may not be beneficial. (reverse coded)
4.250	**Facilitating Conditions (Venkatesh et al, 2003)**
	I have the resources necessary to use this system.
	I have the knowledge necessary to use this system.
	This system is not compatible with other systems I use. (reverse coded)
3.836	**Performance Expectancy (Venkatesh et al., 2003)**
	I would find this system useful in my job.
	Using this system enables me to accomplish tasks more quickly.
	Using this system increases my productivity.
3.742	**Relative Advantage (Moore and Benbasat, 1991)**
	Using this system would enable me to accomplish tasks more quickly.
	Using this system would improve the quality of work I do.
	Using this system would make it easier to do my job.
	Using this system would enhance my effectiveness on the job.
	Using this system would give me greater control over my work.
3.036	**Anxiety (Venkatesh et al, 2003)**
	I feel apprehensive about using this system.
	It scares me to think that I could lose a lot of information using this system by hitting the wrong key.
	I hesitate to use this system for fear of making mistakes I cannot correct.
	This system is somewhat intimidating to me.

Comparison Matrix

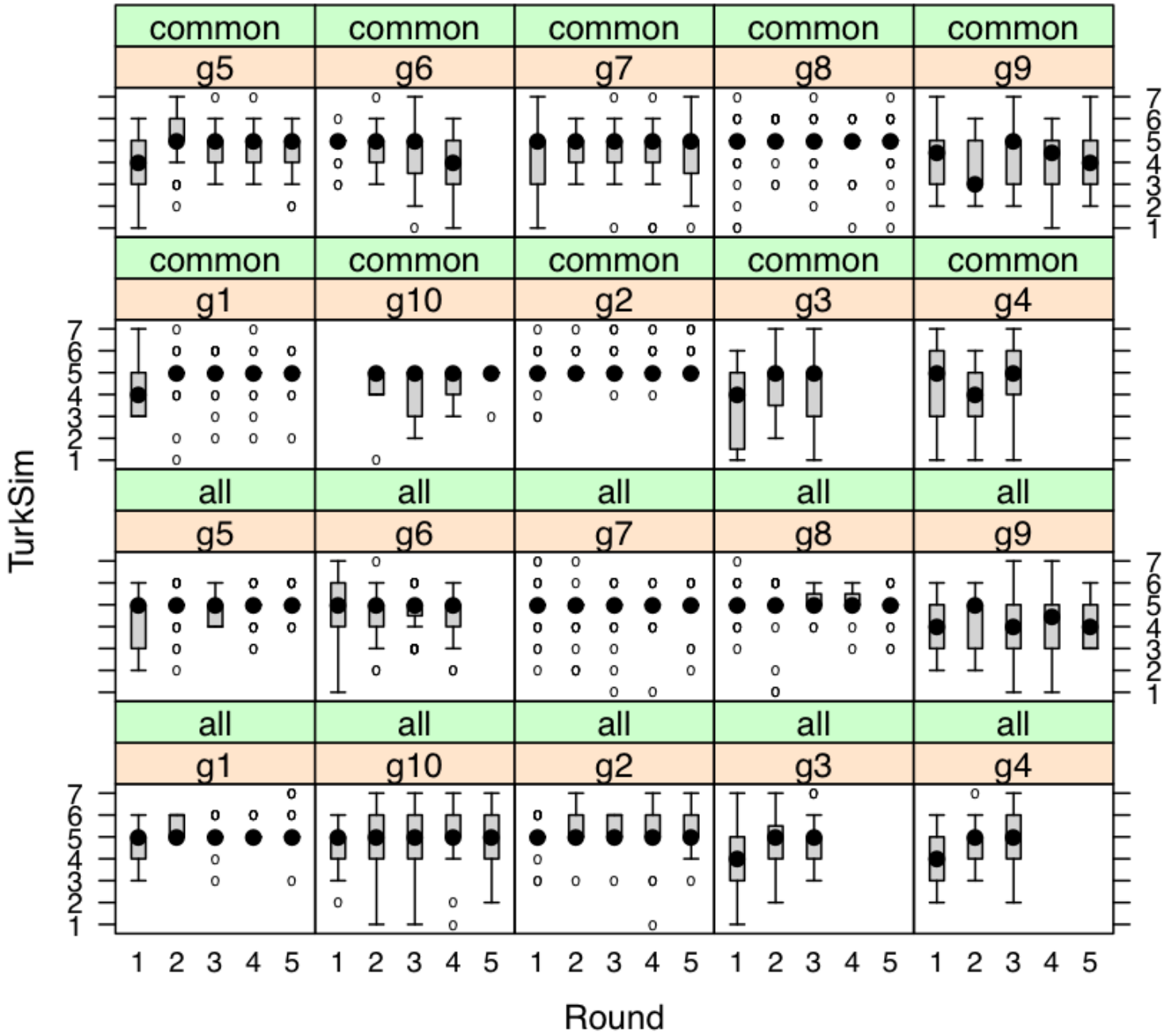
All Possible Similarity Comparisons

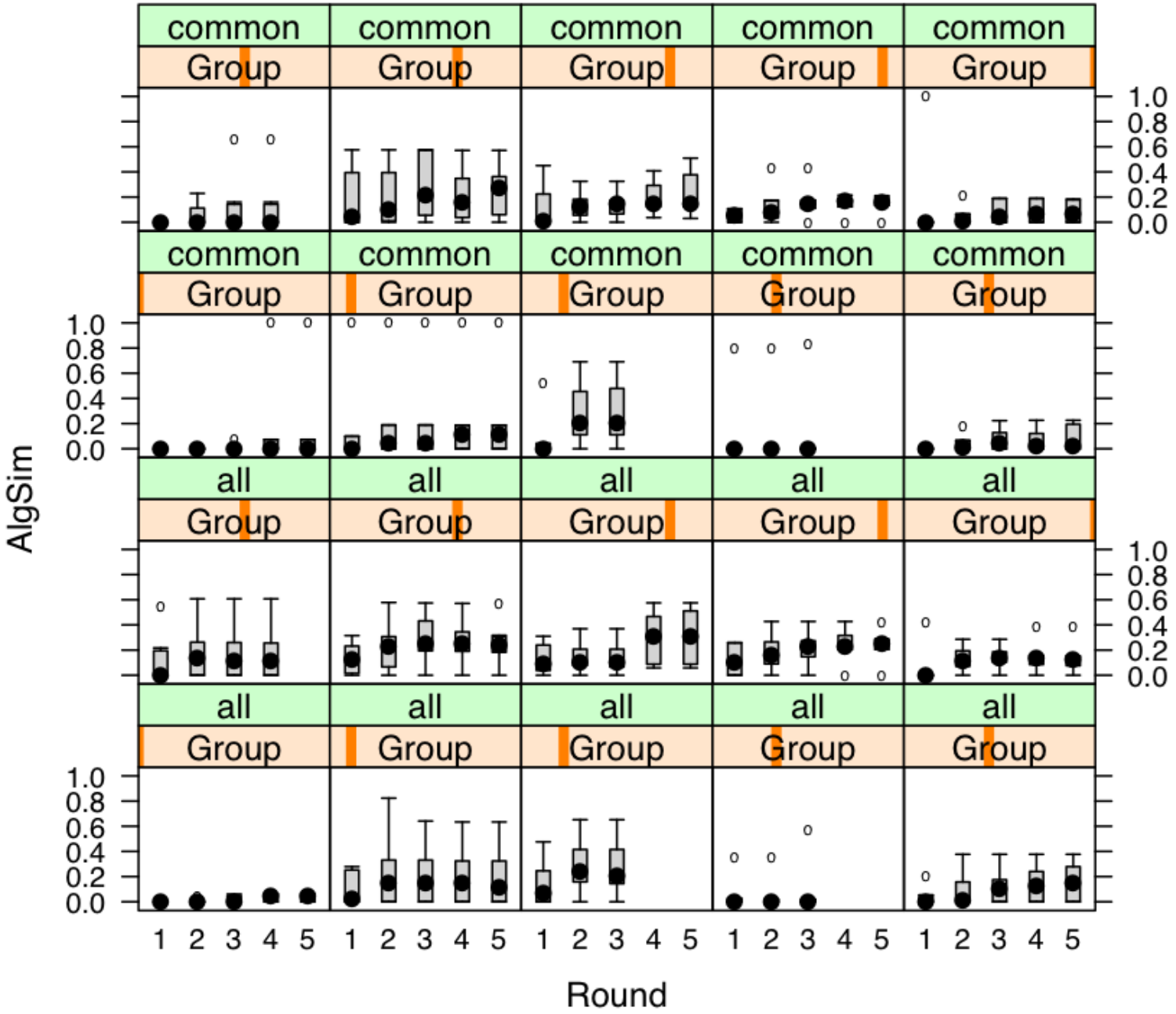
Cleaned	Random	Group/Study	WordNet	Weighted	**HumanSim**	**TurkSim**	**AlgSim**
-	random	group	matching	unweighted	-	-	-
-	random	group	matching	weighted	-	-	-
-	random	group	all	unweighted	-	-	-
-	random	group	all	weighted	-	-	-
-	**Random**	study	matching	unweighted	-	4.4	4.13
-	random	study	matching	weighted	-	-	-
-	**Random**	study	**All**	unweighted	-	4.5	-
-	**Random**	study	**All**	**Weighted**	-	4.6	-
-	-	**Group**	matching	unweighted	-	4.7	4.14
-	-	group	matching	weighted	-	-	-
-	-	**Group**	**All**	unweighted	-	4.8	-
-	-	**Group**	**All**	**Weighted**	-	4.9	-
-	-	**Study**	matching	unweighted	-	4.10	4.15
-	-	study	matching	weighted	-	-	-
-	-	**Study**	**All**	unweighted	-	4.11	-
-	-	**Study**	**All**	**Weighted**	-	4.12	-
cleaned	random	group	matching	unweighted	-		-
cleaned	random	group	matching	weighted	-	-	-
cleaned	random	group	all	unweighted	-	-	-
cleaned	random	group	all	weighted	-	-	-
Cleaned	**Random**	study	matching	unweighted	-	-	4.16
cleaned	random	study	matching	weighted	-	-	-
Cleaned	**Random**	study	**All**	unweighted	4.2	-	-
Cleaned	**Random**	study	**All**	**Weighted**	4.2	-	-
Cleaned	-	**Group**	matching	unweighted	-	-	4.17
cleaned	-	group	matching	weighted	-	-	-
cleaned	-	group	all	unweighted	-	-	-
cleaned	-	group	all	weighted	-	-	-
Cleaned	-	**Study**	matching	unweighted	-	-	4.18
cleaned	-	study	matching	weighted	-	-	-
Cleaned	-	**Study**	**All**	unweighted	4.2	-	-
Cleaned	-	**Study**	**All**	**Weighted**	4.2	-	-


Tagging Activity Per Participant
Tags Added
0
5
10
15
20
25
1
2
3
4
5
Round
Other
Self

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

Mihalcea, 2006


common
g5
g6
g7
g8
g9
common
g1
g10
g2
g3
g4
all
g5
g6
g7
g8
g9
all
g1
g10
g2
g3
g4
TurkSim
Round
1 2 3 4 5 6 7


common
Group
all
AlgSim
Round
1.0
0.8
0.6
0.4
0.2
0.0
1 2 3 4 5