Crowd-in-the-loop

Towards real-time emergency response using crowd supported analysis of social media

Vassilis Kostakos & Jakob Rogstadius

13 May 2011

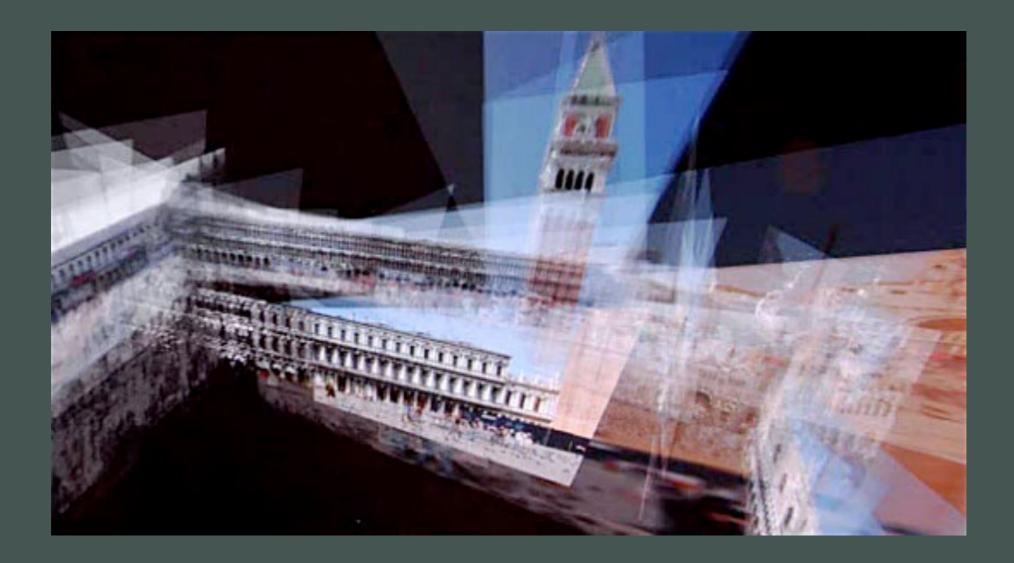


Motivation

- A major crisis happens, such as a natural or man-made disaster
- Situational awareness is distributed
- People share their local knowledge through social media
- Can we collect and organise this information?
- Better and more timely information improves emergency coordination



Synthesising the big picture from multiple perspectives





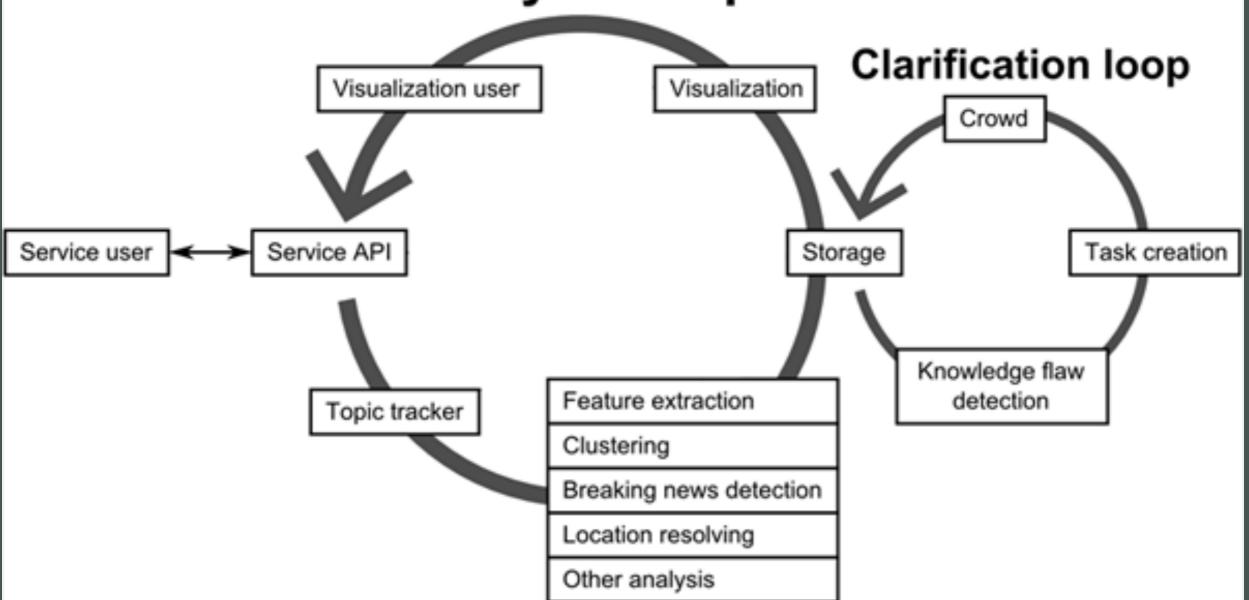
Related work

- Disaster relief & event tracking using social media
 - Hurricane Katrina craigslist
 - Haiti earthquake ushahidi
 - North Africa unrest twitter
 - Hudson river plane crash twitter
- Media aggregation systems
 - Twitris, the Europe Media Monitor, Breakingnews.com
- Clustering
 - named entity extraction and relevance classification, in particular methods adapted for social media content.
- Crowdsourcing platform design
 - Amazon's Mechanical Turk (mturk.com) and CrowdFlower (crowdflower.com) are directly relevant to this work



Conceptual architecture

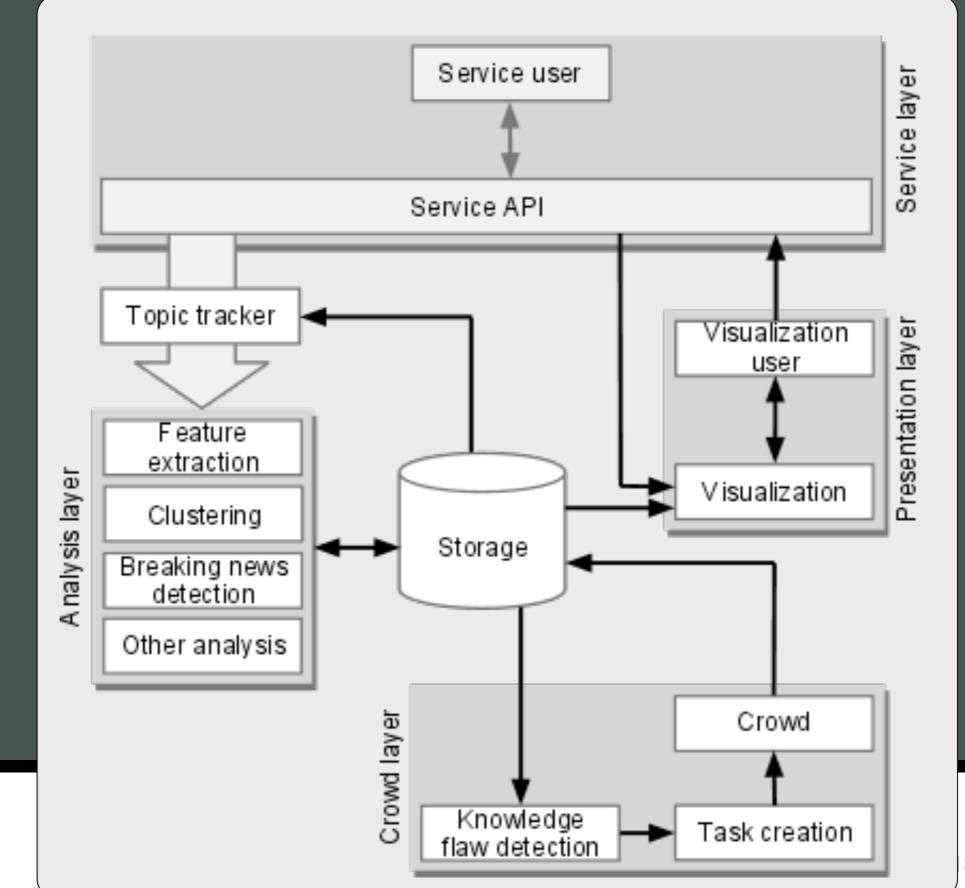
Analysis loop



Rogstadius, J., Kostakos, V., Laredo, J., and Vukovic, M. (2011). **Towards Real-time Emergency Response using Crowd Supported Analysis of Social Media**. In adjunct proceedings of the conference on Human factors in computing systems (CHI 2011), Workshop on Crowdsourcing and Human Computation: Systems, Studies and Platforms. Vancouver, Canada.



System architecture



Slide 6

3, 2011

Prototype - Demo?

| MAIN TOPICS | ACTIVITY TIME | ELINE | | TWEETS | ; |
|------------------------------------------------------|----------------|------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------|-------------------------------------------------------------------------------------------------------|
| Test topic about Fukushima Test topic about Libya | 00 000 00 | | | E. | BentRamallah RT @NABEELRAJAB: Candle protest in Manama http://t.co/8Gg50dx #bahrain #feb14 #bhn |
| Test topic about the Middle East | 1. WI | M . | | | 4 hours ago 85 similar |
| Test topic about Syria | My Maller | MM Lunphan | and the all all the and the and the and the second and the second the se | | inferno0o RT @Trackerinblue: Candles Protest in |
| | 15 April 13:35 | | 15 mins ag | O FILLING | Manama http://t.co/j2t674I via @youtube #feb14 #lulu #bahrain |
| | STORIES | _ | | | 4 hours ago 3 similar |
| | Time | Activity | Story | _ | |
| | 3 hours ago | 35/48 | Was he killed because of 2007 CNN interview with his parents ple | | NABEELRAJAB Candle protest in Manama http://t.co/8Gg50dx #bahrain #feb14 #bhn |
| | 4 hours ago | 30/110 | Candle protest in Dreaz Candles rally in bahrain 23/3/2011 http | | 4 hours ago 2 similar |
| | 6 hours ago | 7/30 | #3aisha #daraa #egypt #Libya #Bahrain #ksa #q8 #UAE #Qata | | nasser668 @NABEELRAJAB: Candle protest in Manama |
| | 9 hours ago | 0/40 | #Libya #Iraq #Saudi http://bit.ly/gto0Yw #syria #Yemen #qata | | http://t.co/epqkSZU #bahrain #feb14 #bhn |
| | 11 hours ago | 12/109 | : : http://www.al-akhbar.com/node/10825 #Bahrain #Egypt # | A soft | 4 hours ago 2 similar |
| | 16 hours ago | 0/45 | Bahrain : Friday protests 22-04-2011 http://t.co/21658nE #bahra | | SalchAlbahrani RT @Moawen: Candles protests continue |
| | 17 hours ago | 0/28 | CNN: Opinion: #Bahrain in the shadow of #Iran, #Saudi Arabia and | | in several places in #Bahrain . Manama protest |
| | 18 hours ago | 0/64 | Protesters killed in Yemen: Witnesses: Protesters took to the streets | | http://www.youtube.com/watch?v=Otg6TTDEUSI 1 hours ago 2 similar |
| | 19 hours ago | 0/26 | The border with #Jordan is totally shut. #deraa under siege. But an | | |
| | 19 hours ago | 0/62 | UNACCEPTABLE!! RT @khabarist: #Bahrain regime seeks death sent | | Moawen Candles protests continue in several places in |
| | 19 hours ago | 0/157 | #Libya #Iraq #Saudi http://bit.ly/hqlkBC #syria #Yemen # | J. | #Bahrain . Manama protest http://www.youtube.com/watch?v=Otg6TTDEUSI |
| | 20 hours ago | 0/27 | RT#Saudi #Kuwait #Bahrain #Egypt #Iraq # Jordan | | 1 hours ago |
| | 20 hours ago | 0/25 | RT @byshr: AP: #Bahrain confirms 3 players from national side deta | | yemenwatch Yemen Observer: Yemeni tribesman |
| | 21 hours ago | 0/58 | Yemen protests kill 3 as opposition haggles http://t.co/cLh9eZD | - | kidnaps Saudi diplomat http://bit.ly/f8hScL |
| | 22 hours ago | 0/439 | Bahrain seeks death sentence for protesters on trial: MANAMA (Reul | - AD MA | 4 hours ago |
| | 23 hours ago | 0/32 | Yemen troops kill 2 in new clashes with protesters $n (AP) n: AP$ - | 1000 | NABEELRAJAB Candle protest in Dreaz Candles rally |
| | 23 hours ago | 0/87 | Deaths reported in fresh Yemen protests: Sources say two people de | 100 | in bahrain 23/3/2011 http://t.co/I82fCqO #bahrain |
| | 4 | c II | | | #feb14 #bhn 4 hours ago |





Rogstadius, J., Kostakos, V., Kittur, A., Smus, B., Laredo, J., and Vukovic, M. (2011). **An Assessment of Intrinsic and Extrinsic Motivation on Task Performance in Crowdsourcing Markets**. In proceedings of the AAAI Conference on Weblogs and Social Media, Barcelona, Spain.



Study I

- Can we motivate people to donate their time to a "good" cause?
 - -Findings: Complex relationship between rewards (payments), time-to-complete crowdsourced tasks, and the nature of the task itself (is it for a "good" cause?).
 - Application: Develop a better crowdsourcing strategy.
 We will have to pay people if we want faster, but not necessarily better, results



Ways to motivate people

- Intrinsic = personal enjoyment, desire to help
- Extrinsic = payment, social status, threats
- Study design
 - -Posted work on Amazon's Mechanical Turk
 - -Varied payment and cover story (cause)
 - -Measured completion speed and accuracy

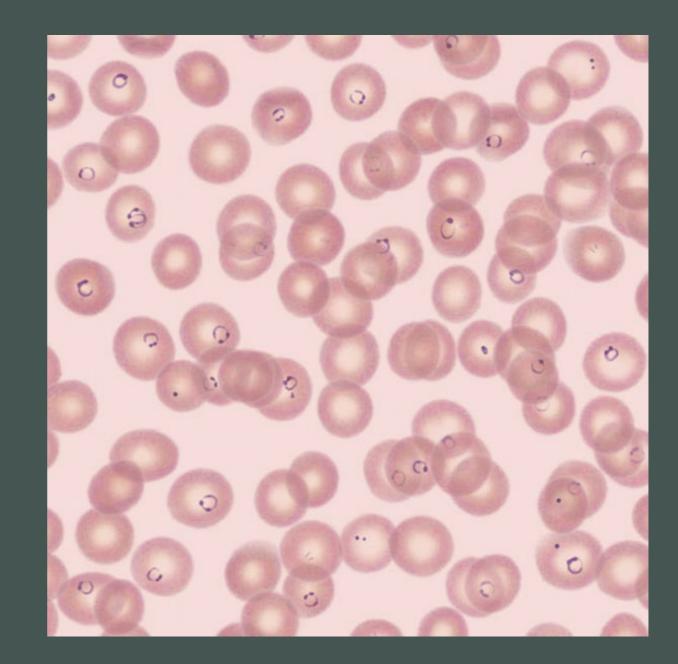


The task

 Generated artificial images of infected blood cells

• Task

- -Count blood cells
- Count parasites in double-dotted ring form
- -Accuracy = weighted
 average





Conditions

- For charity: the work is posted on behalf of the World Health Organisation
- For profit: the work is posted on behalf of a marketleading pharmaceutical company.
- Reward manipulation:
 - -Pay: 0, 3, 10 cents per task
- Task manipulation
 - -Number of cells in image
 - -Number of parasites in image



Recruitment

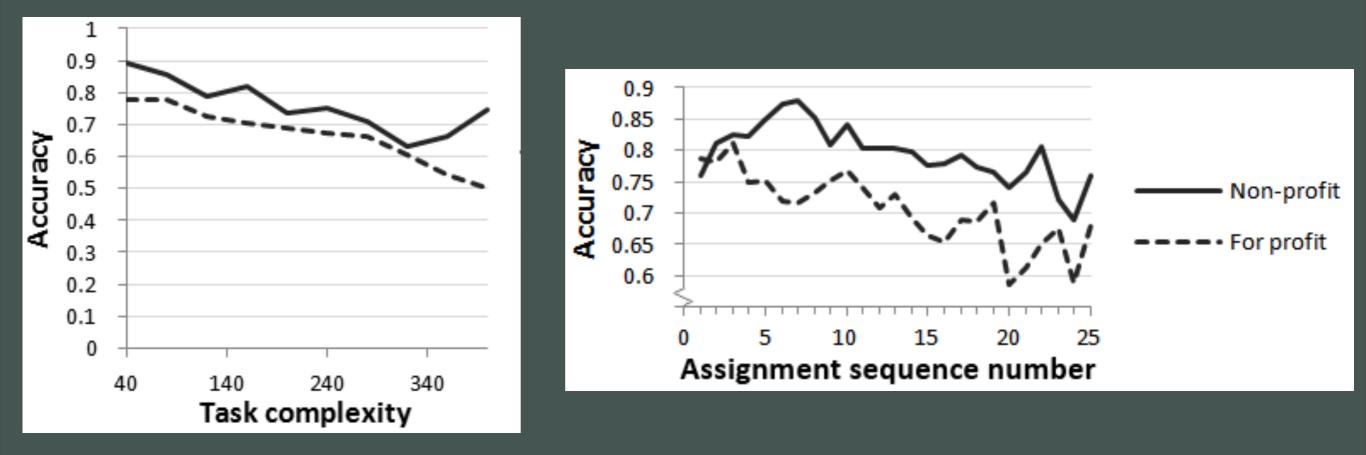
- MTurk's "Qualifications" used to allocate people to conditions
- Initial questionnaire => person allocated to one of 6 conditions
 - -2 (cover story) x 3 (payment)
- A person can only see & do tasks from their condition

 Tasks varied in terms of difficulty
- Ensure that people don't see the same task for different reward.



Varying cover story

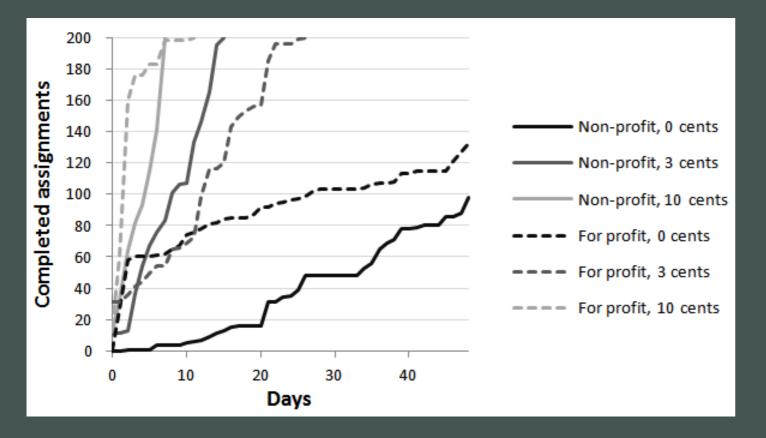
- Working for charity = generally more accurate
- No effect on completion speed

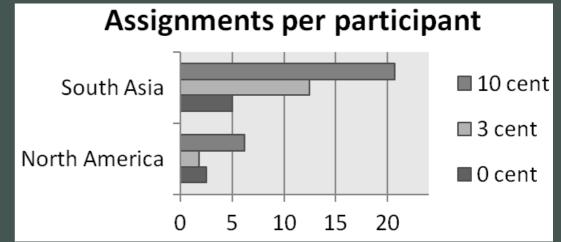




Varying payment

- More money = faster results
- Size of effect varies (with level of income?)
- Non-significant effects on accuracy







Interaction effects

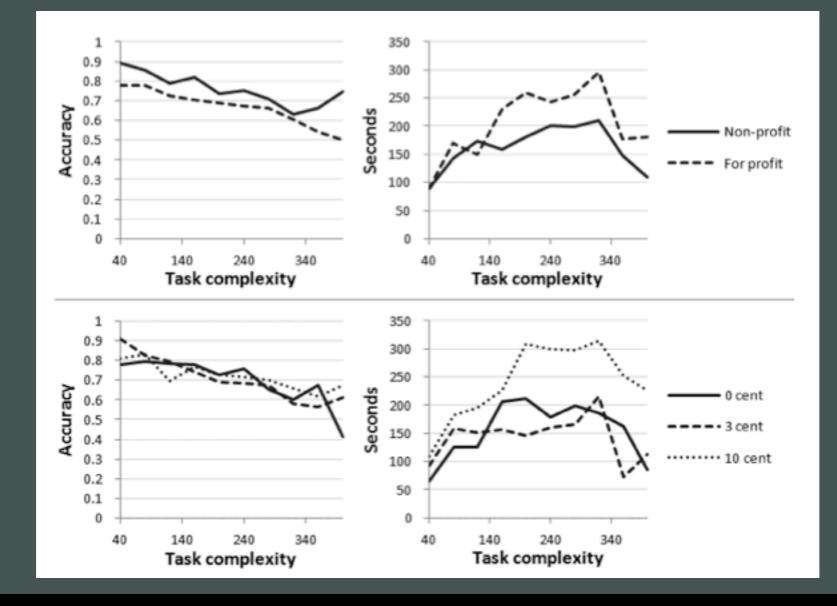
- Primary motivation source? Other explanation?
- Worker bias?
- We observed two different reactions to charity cause





Difficulty threshold

• Beyond some difficulty threshold, people gave up





Study 2



Study 2

- If people decide to donate their time, how can we help them make best use of their time while contributing to our system?
 - -Findings: An understanding of how people judge "similarity" between tweets, what information they find useful, what information they discard (as noise)
 - -Application: Avoid overloading our users, give them the information they need to complete the task quickly. Give them the right kind of task.



Rate the topic similarity of tweet pairs

Instructions:

Below are five pairs of tweets (status updates from the microblogging service Twitter). The tweet on the left hand side (Reference tweet) is the same for all five pairs.

- 1. Read the tweets in each pair.
- 2. Rate how similar the topics of these two tweets are.
- Tell us which of the available information the score you gave was based on.

It may help to imagine that you are using a search engine, where the scores assigned by you are used to retreive tweets that are **of similar topic** to the reference tweet. Tweets from pairs which you have assigned a higher score will appear as more relevant than pairs which you have given a lower score. In some HITs, all tweets will be of different topics, in which case all pairs should get low scores.

This HIT is part of a research project conducted at the University of Madeira.

| Reference twe | et: | Tweet 1: | | | | | | | | | |
|-------------------|------------------------------|----------------------------------------------|---------------------|-------------------------|--|--|--|--|--|--|--|
| Bout time to head | l off to sonisphere yo :D | Capturas de @Direngrey en el Sonisphere 2010 | | | | | | | | | |
| In reply to: | | | http://bit.ly/afJmP | <u>a</u> | | | | | | | |
| Tweet time: | July 30 2010 | Ir | n reply to: | | | | | | | | |
| Tweet location: | | T | weet time: | August 5 2010 | | | | | | | |
| User (nickname): | Connie Raitt (ConnieEvarose) | יד | weet location: | | | | | | | | |
| User location: | Banbury | U | Jser (nickname): | AniRockers (AniRockers) | | | | | | | |
| User followers: | 69 | U | Jser location: | Venezuela | | | | | | | |
| User friends: | 81 | U | Jser followers: | 146 | | | | | | | |
| | | U | Jser friends: | 143 | | | | | | | |

How similar are the topics of these two tweets?

Very different

Very similar

| Reference twe | et: | Tweet 1: | | | | | | | | | |
|---------------------------------------------------------------------------------------------------------|---------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------|---------------|--|--|--|--|--|--|--|--|
| Bout time to head | d off to sonisphere yo :D | Capturas de @Direngrey en el Sonisphere 2010 | | | | | | | | | |
| In reply to: Tweet time: Tweet location: User (nickname): User location: User followers: | July 30 2010 Connie Raitt (ConnieEvarose) Banbury 69 | http://bit.ly/afJmF In reply to: Tweet time: Tweet location: User (nickname): User location: | August 5 2010 | | | | | | | | |
| User friends: | 81 | User followers: User friends: | 146 143 | | | | | | | | |

How similar are the topics of these two tweets?

| Ver | y dif | ferei | nt | | | | | | | | | | | | | | | | | | | | | | | | Ver | ry sin | nilar |
|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|
| \bigcirc |

For this pair of tweets, which of the following information had any influence (positive or negative) on your topic similarity rating?

- Hashtags (e.g. #christmas2010)
- Destination pages of URLs (links)
- Named entities that were mentioned (e.g. World Cup, Obama, New York, Xbox, Twitter)
- Sentiment of tweet(s) (e.g. positive/negative)
- Text in the tweet, excluding hashtags, URLs, sentiment and explicitly mentioned named entities
- Any text in a reply-to tweet
- Tweet time(s)
- Tweet location(s)
- User name(s)
- User location(s)
- User followers
- 📃 User friends

Overview

- Collected tweets relating to different events
 - Music festivals
 - Hurricanes
 - Oil spill
 - Random (control group)
- See how people assess the "similarity" of pairs of tweets
 - Understand how clustering of tweets can be distributed to a crowdsourcing environment
 - Ask them to indicate what information they relied on to complete their assessment
 - 12 independent ratings for each single pair of tweets

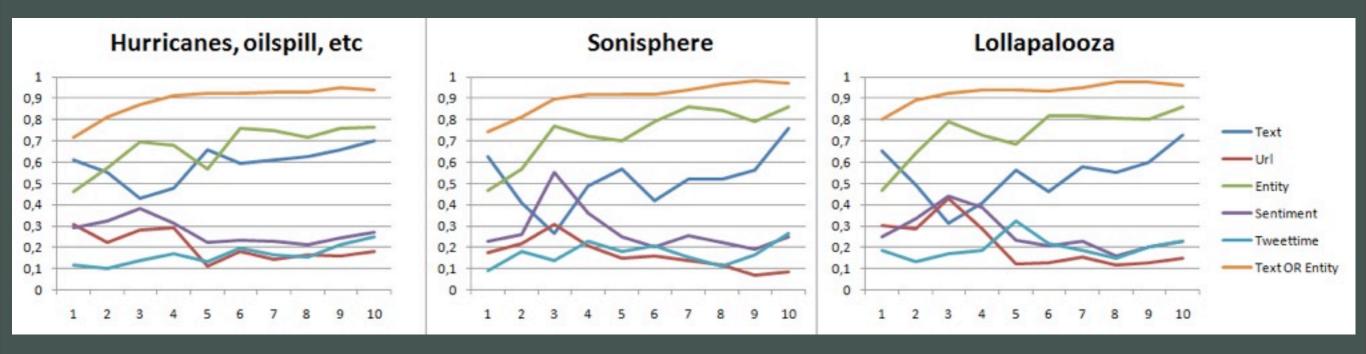


How do people interpret tweets?

- Text (except URLs, entities, sentiment or hash tags) (61.9%)
- Named entities (54.5%)
- Sentiment (30.1%)
- URLs (20.3%)
- Time of tweet (14.1%)
- Hash tags (4.1%)
- Location of poster (3.5%)
- Number of followers (3.3%)
- Number of friends (3.0%)
- Location of tweet (2.9%)
- Any text in the "reply to" tweet (as indicated by Twitter) (2.9%)
- Name of poster (2.7%)



How do people interpret tweets?

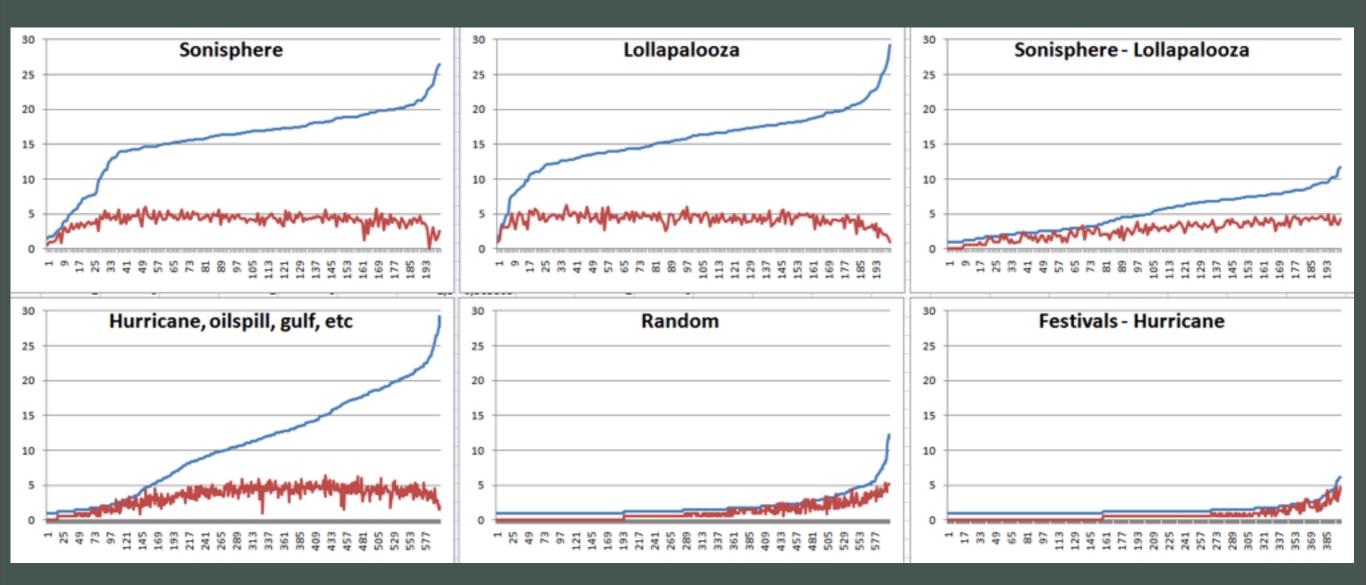


Iow Tweet similarity rating





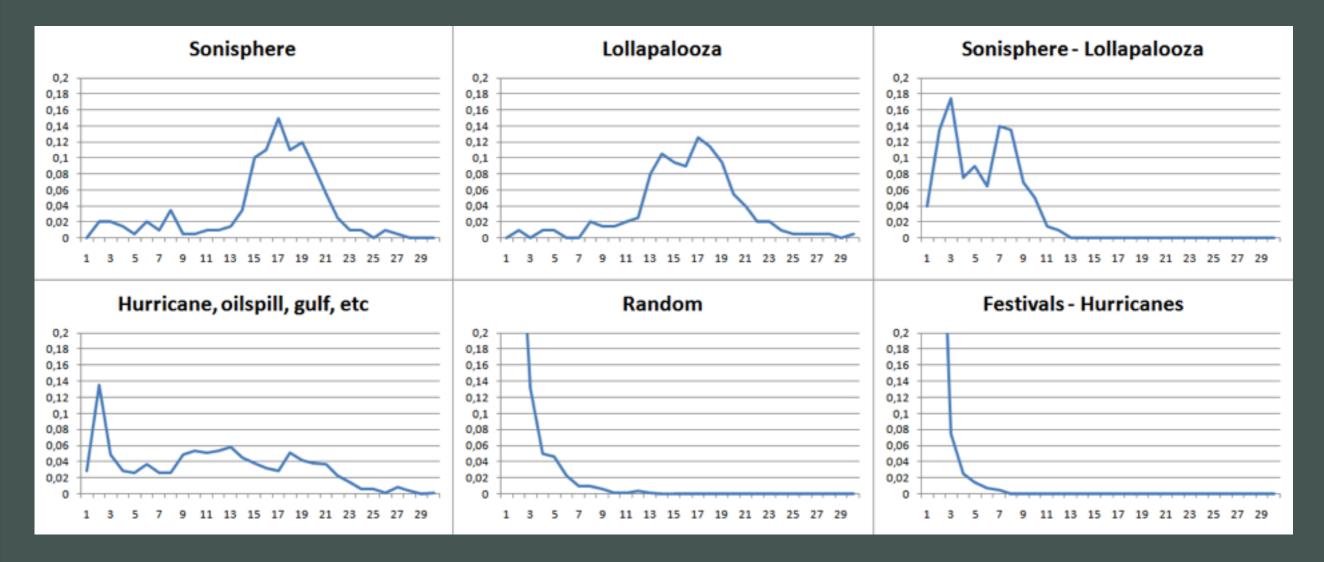
Similarity ratings



Similarity (blue) and stderr (red) for pairs of tweets



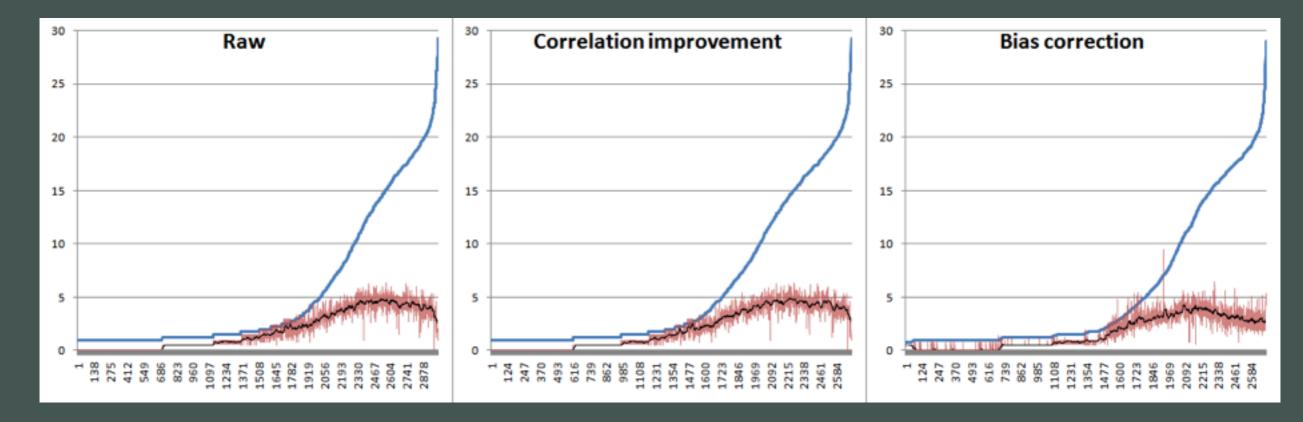
Similarity ratings



Probability function of similarity rating for pairs of tweets



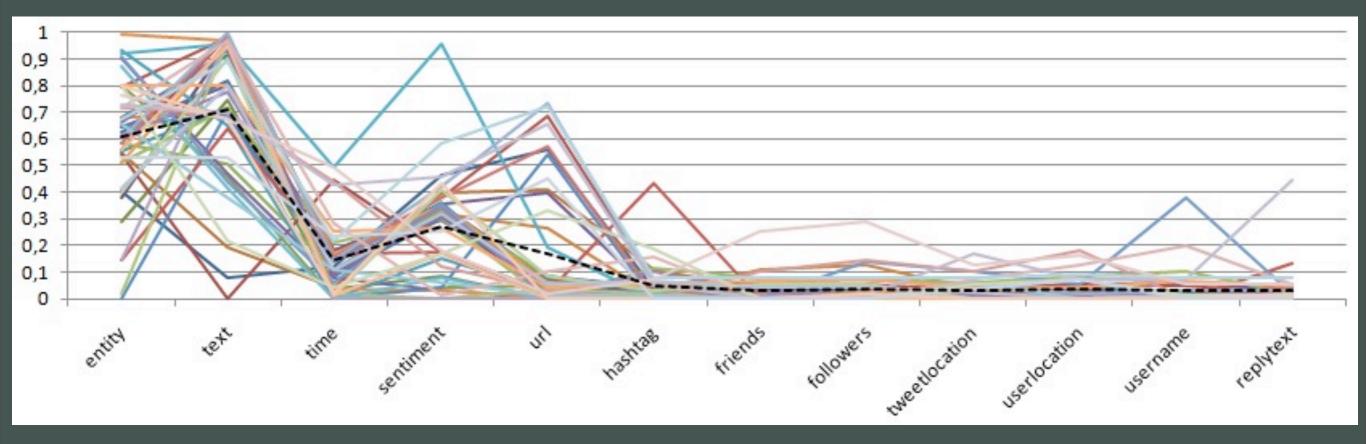
Removing "bad" workers



Removing "bad" workers has insignificant effect



Individual worker strategies



Usefulness of information (y-axis) varies by worker



In conclusion



Conclusion

- Ongoing work on
 - -Understanding better crowdsourcing markets and tasks
 - -Integrating crowdsourcing (in-situ?)
 - -Deploying crisis tracking software
- Results
 - -Study I provides evidence regarding worker motivation
 - -Study 2 provides evidence regarding worker capabilities



Publications

- Our work is already being cited by the community (ECSCW Workshop cites our work as motivation)
- Interview on New Scientist (to appear)
- Rogstadius, J., Kostakos, V., Laredo, J., and Vukovic, M. (2011). Towards Real-time Emergency Response using Crowd Supported Analysis of Social Media. In CHI 2011, Workshop on Crowdsourcing and Human Computation: Systems, Studies and Platforms.
- Rogstadius, J., Kostakos, V., Kittur, A., Smus, B., Laredo, J., and Vukovic, M. (2011). An Assessment of Intrinsic and Extrinsic Motivation on Task Performance in Crowdsourcing Markets. In proceedings of the AAAI Conference on Weblogs and Social Media, Barcelona, Spain.

