

experiences surveying the crowd

reflections on methods,  
participation, and reliability

Cathy Marshall (Microsoft Research)  
& Frank Shipman (Texas A&M)

Websci 2013, Paris  
2 May 2013

*This talk has been  
annotated and  
modified for the  
web.*

Amazon Mechanical Turk

### HITs containing 'survey'

1-10 of 485 Results

Sort by: HITs Available (fewest first) [GO](#) [Show all details](#) | [Hide all details](#) [1](#) [2](#) [3](#) [4](#) [5](#) > [Next](#) >> [Last](#)

<b>Fast - 25 cents - 2 minute Survey - 15 Questions</b> <a href="#">Request Qualification (Why?)</a>   <a href="#">View a HIT in this group</a>			
<b>Requester:</b> <a href="#">Matthew Rappaport</a>	<b>HIT Expiration Date:</b> Feb 18, 2013 (33 minutes 7 seconds)	<b>Reward:</b> \$0.25	
<b>Time Allotted:</b> 15 minutes		<b>HITs Available:</b> 1	
<b>12-15 Minute Survey -- 50 cents</b> <a href="#">View a HIT in this group</a>			
<b>Requester:</b> <a href="#">VULS</a>	<b>HIT Expiration Date:</b> Feb 18, 2013 (48 minutes 52 seconds)	<b>Reward:</b> \$0.50	
<b>Time Allotted:</b> 45 minutes		<b>HITs Available:</b> 1	
<b>Personality Survey and Creativity - Interesting! (Around 10 minutes)</b> <a href="#">View a HIT in this group</a>			
<b>Requester:</b> <a href="#">HKU social lab</a>	<b>HIT Expiration Date:</b> Feb 18, 2013 (1 hour 58 minutes)	<b>Reward:</b> \$0.20	
<b>Time Allotted:</b> 20 minutes		<b>HITs Available:</b> 1	
<b>Interactions and Perceptions Study</b> <a href="#">View a HIT in this group</a>			
<b>Requester:</b> <a href="#">Nick Camp</a>	<b>HIT Expiration Date:</b> Feb 18, 2013 (2 hours 24 minutes)	<b>Reward:</b> \$0.25	
<b>Time Allotted:</b> 3 hours		<b>HITs Available:</b> 1	
<b>Short survey (5-10 minutes)</b> <a href="#">View a HIT in this group</a>			
<b>Requester:</b> <a href="#">Tatiana Fajardo</a>	<b>HIT Expiration Date:</b> Feb 18, 2013 (2 hours 31 minutes)	<b>Reward:</b> \$0.10	
<b>Time Allotted:</b> 60 minutes		<b>HITs Available:</b> 1	
<b>Thoughts and Opinions (4-8 minute survey)</b> <a href="#">Request Qualification (Why?)</a>   <a href="#">View a HIT in this group</a>			
<b>Requester:</b> <a href="#">Jonathan Corbin</a>	<b>HIT Expiration Date:</b> Feb 19, 2013 (4 hours 25 minutes)	<b>Reward:</b> \$0.50	
<b>Time Allotted:</b> 12 hours		<b>HITs Available:</b> 1	
<b>20 minute Work Experiences Survey</b> <a href="#">View a HIT in this group</a>			
<b>Requester:</b> <a href="#">Erin Gallagher</a>	<b>HIT Expiration Date:</b> Feb 19, 2013 (8 hours 10 minutes)	<b>Reward:</b> \$0.10	
<b>Time Allotted:</b> 60 minutes		<b>HITs Available:</b> 1	
<b>Short Survey for Asian or Asian American</b> <a href="#">View a HIT in this group</a>			

Recruiting study participants is a balancing act. You need diversity. You'd like convenience. You want to keep costs in check. So researchers in different disciplines have been conducting surveys by offering them as Mechanical Turk HITs. If you've logged onto Mechanical Turk as a worker, you might encounter 500 or so different surveys available for you to take for a small payment.

# from Turker Nation

For years I used to be an avid survey taker for money... making anywhere from \$200 to \$300 a month...

...it's such a nice change of pace doing so many different types of surveys/studies on Mturk. I swear, if I did one more survey on my cell phone or on sanitary protection, I was going to pull my hair out.

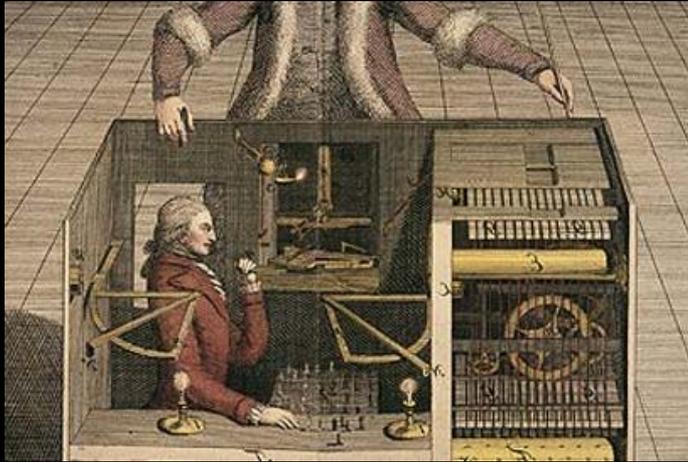
*--sandooch (a woman located in California)*

# US turkers take surveys for a variety of reasons



- for fun  
*"...I do [surveys] quite often, largely for fun..."*  
User [rlrsk8tr1](#), [TurkerNation](#)
- out of interest/to express opinions  
*"...Since both surveys I did were on subjects where I have strong opinions anyway (Global Warming and Privatizing Social Security) they were very easy to write..."*  
User [JoeTurker](#), [TurkerNation](#)
- easy \$ (e.g. while watching tv)  
*"Took me 6 minutes, but I'm watching Judge Mathis. Also, one AC."*  
from [HITsWorthturkingfor](#) subreddit
- also... desperation, community (c.f. Irani and Silberman), competition (c.f. von Ahn) and *because they're already in front of the screen!*

# Mechanical Turk as study venue



3 years, 7 media type-specific surveys offered as a HITs to >1000 qualified participants

We applied and extended best practices from literature (e.g. Kittur et. al., 2008)

In so doing, we learned about:

**methods (our focus here)**

Turkers who take surveys\*

reliability of the crowd\*

*\*These topics are covered in the paper*



using scenarios & hypotheticals to  
elicit social norms



**step 1. characterize the phenomenon of interest**  
**step 2. factor it into parts (actors, actions, concepts)**  
**step 3. make up scenarios and hypotheticals**

# step 1: examples of real reuse phenomena

a convenient illustration

try to avoid social desirability bias  
(Antin and Shaw, 2012)



satire/humor



commercial reuse



## hypotheticals: from legal education and CBR

The pure hypothetical is a counterfactual variation on the fact pattern of an actual case. The hypothetical plays an important role in the Socratic style of law school teaching.

*Solum, Legal Theory Lexicon, 2013*

Hypotheticals may be used to ... *factor a complex situation into component parts (e.g., by exaggerating strengths, weaknesses or eliminating features).*

*Rissland and Ashley, Hypotheticals as Heuristic Device, 1986*

**<Reuser> takes <action> as instance of <concept>**



*Dimension 1*



*Dimension 2*



*Dimension 3*

## Dimension 1: <reuser>

- Photographer (owner)
- Fishermen (subjects, possible rights holders)
- People of varying social distance



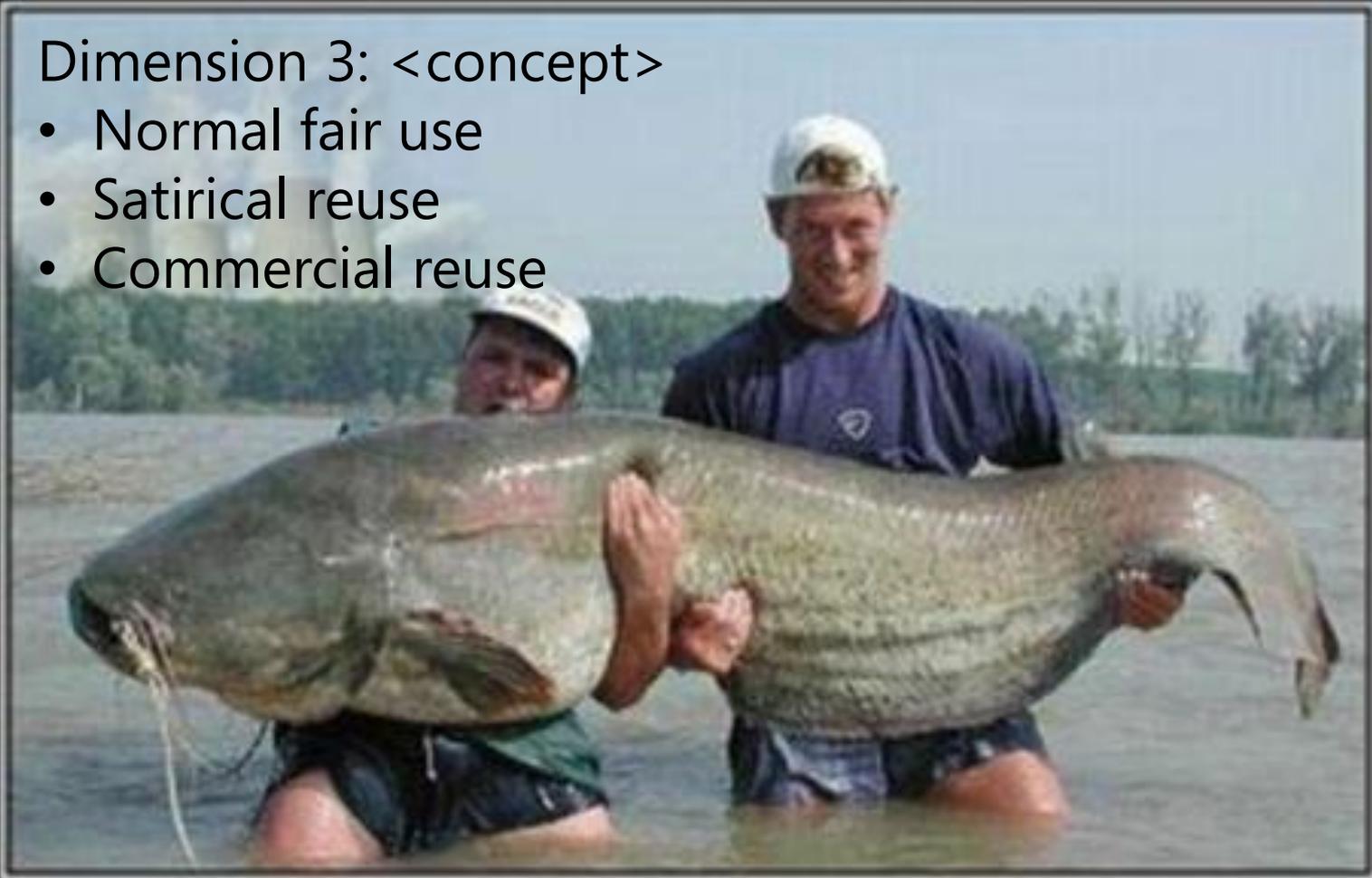
## dimension 2: <action>



- **store**
  - e.g. save to USB key
  - e.g. store to the cloud
- **share**
  - e.g. post to Facebook
  - e.g. attach to email
- **(re)publish**
  - e.g. post to a public blog
  - e.g. upload to You Tube
- **remove**
  - e.g. untag a photo
  - e.g. change privacy settings

Dimension 3: <concept>

- Normal fair use
- Satirical reuse
- Commercial reuse



Достоинства Ядерной Энергетики

# scenario + hypotheticals

On their vacation, Jules and Jim go fishing for catfish in the Mekong river. They make friends with Loan-anh. She snaps their picture after they catch what surely must be a record-breaking fish.



1. Loan-anh can post the photo to her Facebook account.
2. Jules can download the photo from Facebook and store it on his computer.
3. Jules can upload the photo to illustrate his public blog.

# adding a twist

Carlos, a college student, is doing a talk on Chernobyl for class and wants to illustrate it with a funny picture. He finds the photo on Jules' public blog and downloads it. He then adds cooling towers to the photo.

5. Carlos can use the photo for his presentation.
6. Carlos can post his presentation on a public slideshare site.

Carlos edits the photo to add a funny caption.

7. Carlos can submit the captioned photo to an e-postcard site.
8. Carlos can accept royalties from the e-postcard site.

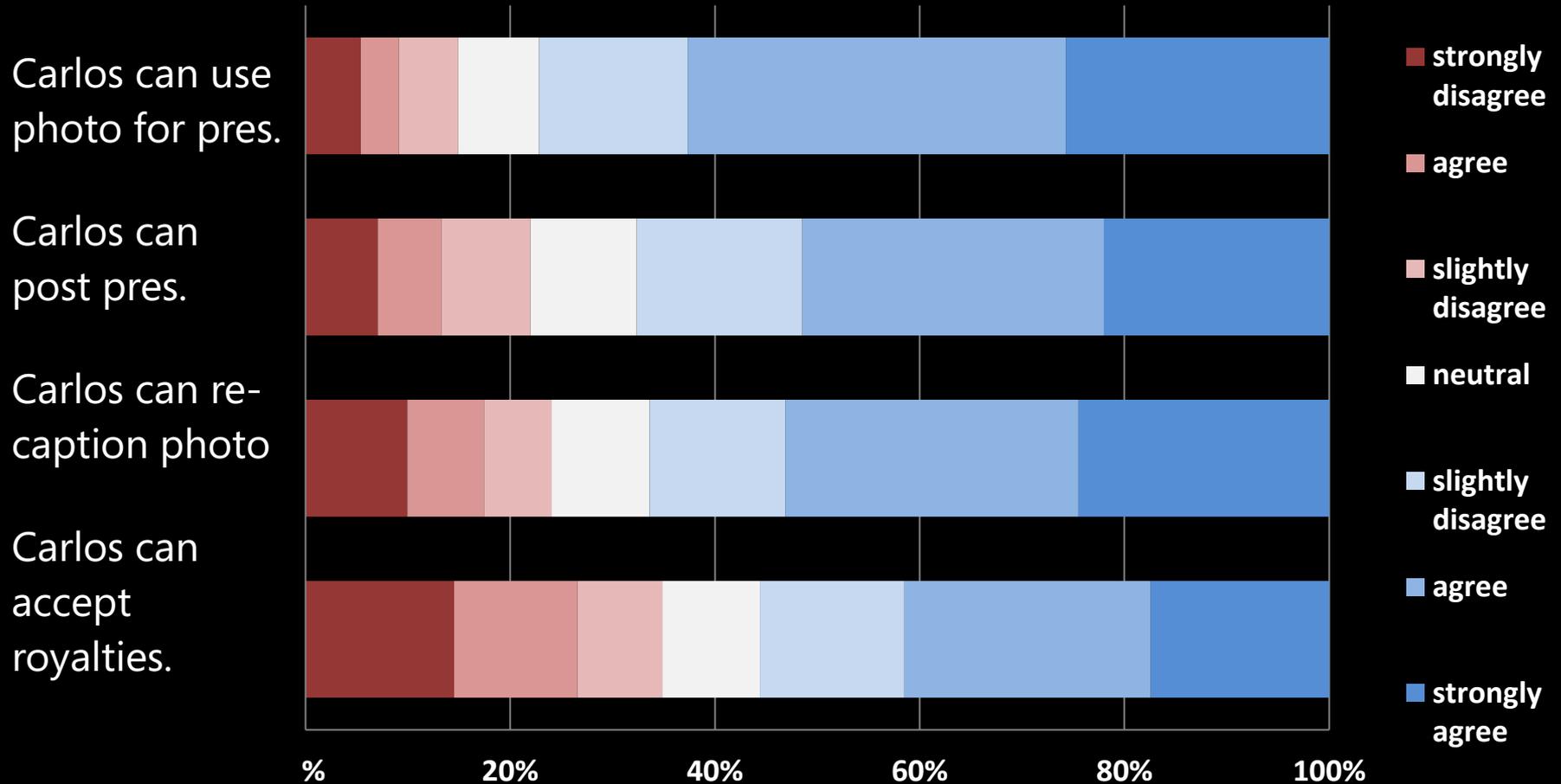


Достоинства Ядерной Энергетики

participants judge hypotheticals  
using a Likert scale

- Strongly Agree
- Agree
- Agree somewhat
- Undecided
- Disagree somewhat
- Disagree
- Strongly disagree

# this gives us the ability to compare across variables



# verdict

hypotheticals are an effective means of eliciting social norms provided:

- the story is real and grounded
- you don't enumerate (3 actors x 4 concepts x 4 actions)
- the distinctions aren't too nuanced

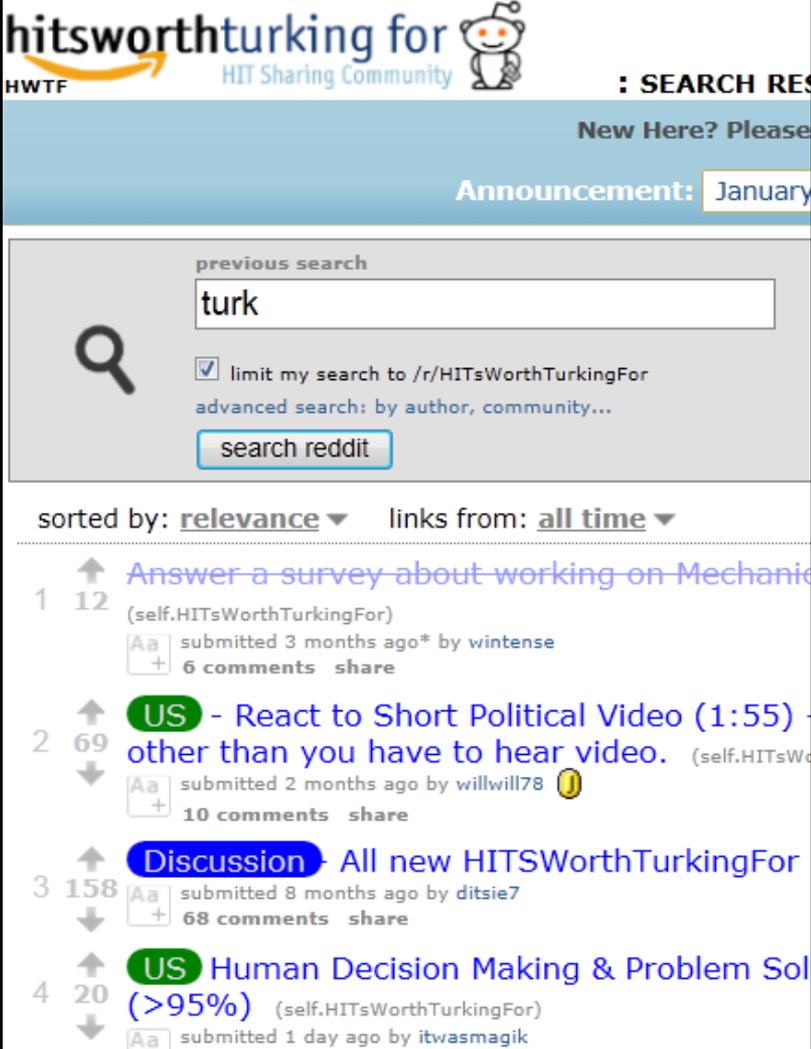
engagement matters! (e.g. Mason & Watts, 2010)



# how to learn more about participants

- take an ethnographic approach to asking questions—Turkers answer them!
- read forums—Turkers talk among themselves
  - turkernation
  - hitsworthturkingfor and mturk subreddits
  - turkopticon (Irani and Silberman, 2013)
  - mturkforum
- read the crowdsourcing literature

*caveat: all crowdsourcing research seems to hit different segments of the worker community.*



The screenshot shows the hitsworthturkingfor subreddit search interface. The header includes the subreddit name, logo, and navigation links. A search bar contains the term 'turk'. Below the search bar, there are options to limit the search to the subreddit and a search button. The search results are sorted by relevance and show four posts:

1. [Answer a survey about working on Mechanic](#) (self.HITsWorthTurkingFor) submitted 3 months ago\* by wintense. 12 upvotes, 6 comments.
2. [US - React to Short Political Video \(1:55\) - other than you have to hear video.](#) (self.HITsWorthTurkingFor) submitted 2 months ago by willwill78. 69 upvotes, 10 comments.
3. [Discussion - All new HITSWorthTurkingFor](#) submitted 8 months ago by ditsie7. 158 upvotes, 68 comments.
4. [US Human Decision Making & Problem Sol \(>95%\)](#) (self.HITsWorthTurkingFor) submitted 1 day ago by itwasmagik. 20 upvotes.

# open-ended questions: a penny for your thoughts?

Media	hypos	demo/practice	MC	open-ended	total
tweets	16	12	3	3	34
photos	18	14	3	6	41
reviews	28	14	3	5	50
podcasts	22	14	2	4	42
videoconferences	20	13	2	5	40
ed. recordings	25	14	2	4	45
mmogs (games)	27	17	2	4	50

*relative structure of the 7 media type specific surveys that we've done so far*

open-ended questions elicit self-absorbed  
and uninhibited responses  
*(see Kiesler and Sproull, 1986)*

e.g. From Games survey: *What do you talk about with other players?*

"Anything. A clan can sometimes be like a second family. If someone had a horrible day at work, you might empathize or give suggestions. Or it might be fun, if your playing a first person shooter to provoke them with political conversation. I'm not a jerk. Really. I just realized it sounds like I am. But i'm not. It can be in good fun! :P"

e.g. online activities

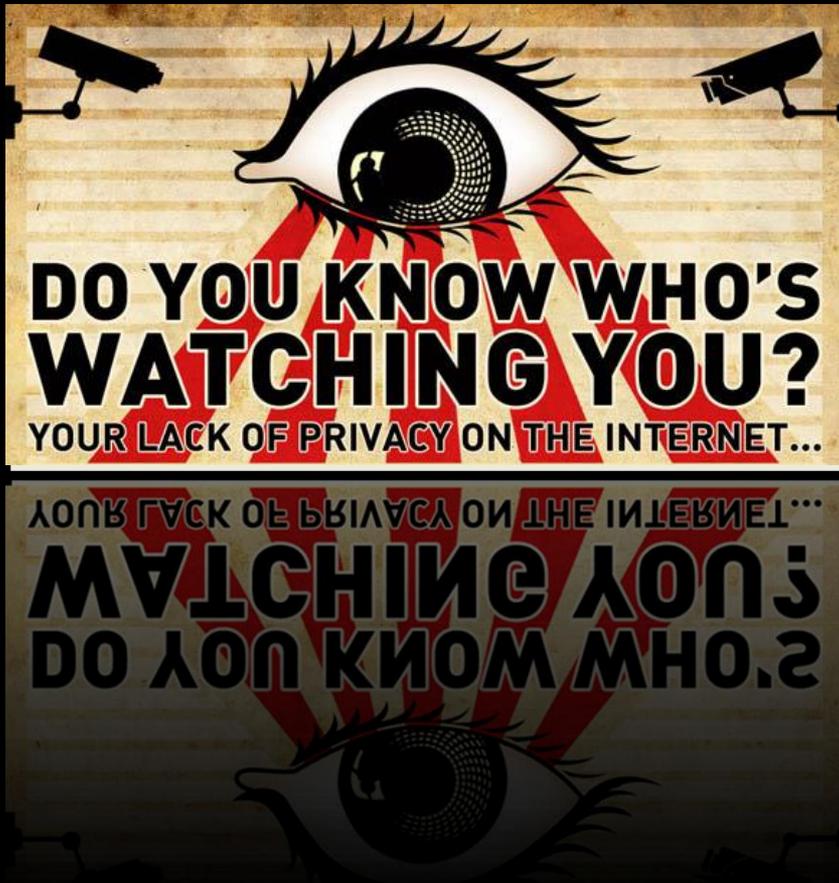
(an open-ended question with open-coded results)

*what else do Turkers do online?*

category	subcategories	total	example (excerpt)
<b>Social media</b>	social networking; Facebook; Twitter; Forums; Reddit; Myspace	1020	"Facebook! I am addicted to letting everyone know what I am doing & reading what they are doing!" [ED117]
<b>Consume</b>	reading; watching videos/tv/movies; listening to music/podcasts/radio	972	"I also like to view the various websites that relate to Asian ball jointed dolls." [ED067]
<b>Communi- cate</b>	email; Skyping; IM/chat	757	"I Skype with so many people even though I could just call them. Lol." [RE112]
<b>Research</b>	researching specific topics; using specific resources	548	"I do use the Internet at work to look up information about medications and diseases. <b>(I am a nurse.)</b> " [RE130]
<b>Work and school</b>	work-related activities; learning or homework; looking for jobs	460	"Mostly I use the internet while I am at work. <b>I am a manager at a law firm</b> so I frequently use Excel and Word." [VC154]
<b>Shopping</b>	shopping (specific stores); shopping for specific items; find coupons	432	"I often shop online, but even when I purchase from a bricks and mortar store I research my purchases online beforehand." [PH159]
<b>Publish media</b>	photo or video sharing; art; blogging; website development; media aggregation	297	" <b>...I work as a freelance illustrator.</b> I use sites such as Blogger, Etsy, and Flickr to share my artwork and look at others' artworks and photos." [PC095]

category	subcategories	total	example (excerpt)
Gaming	gaming (casual, online, multiplayer); fantasy sports; specific games (e.g. WoW)	257	"I also play in some social virtual character websites such as Gaia Online and TinierMe." [PH147]
"Get Paid To..."	Mechanical Turk/HITs; surveys; other; Etsy and eBay selling	185	<b>"I spend a lot of my internet time (about 2-5 hours a day) doing mTurk because I am a single Mom and I need the extra cash."</b> [PH069]
Entertainment	entertainment/fun; killing time/leisure; hobbies and crafts; porn	147	"I also use the internet to get rid of stress and take a break from life." [PH214]; "I spend most of my time looking up craft patterns" [PH039]
non-specific	"everything"; various unspecified (e.g. "etc"); "I don't"	83	"Ever since I was like 13 (I'm 29 now) the internet has been an integral part of my life." [PC144]
Financial	banking; bill-paying; investing/stock trading	48	<b>"The most time consuming activity that i do on the internet is day trade.</b> I actively trade five days per week for four to five hours per day." [TW152]

# a word or two about privacy



We didn't ask about privacy. Respondents talked about it anyway.

e.g what do you publish online?

*"likely too much, hah!"*

Turkers readily surrender private data—they tend to rely on privacy through obscurity.

Data curation and reporting must be done with this in mind.

*Adversarial relationships are reciprocated!*

# wrap-up: why we wrote this paper

- Big companies and researchers now rely on human computation.
- We had gathered extensive data on a large number of participants.
- The WORK is crucial to determining worker reliability. (engagement, purpose, creativity)
- *The changing shape of information work is intriguing (and sometimes alarming).*



fin

contact info:

[cathymar@microsoft.com](mailto:cathymar@microsoft.com)

[shipman@cs.tamu.edu](mailto:shipman@cs.tamu.edu)

<http://www.csd.tamu.edu/~marshall>

<http://www.csd.tamu.edu/~shipman>

<http://twitter.com/ccmarshall>

