

CSE518A: Office Hour (till 4pm)

Lecture 13

Real-time Crowdsourcing

Instructor: Chien-Ju (CJ) Ho

Logistics: Assignment 3 and Project Milestone 1

- Assignment 3: Due Oct 27 (Tue)
- Project milestone 1: Due Oct 30 (Fri)
 - Initial literature survey (know what other works are out there)
 - A plan on what you want to do for the remaining of the semester
 - **Formalize your research question and approaches**, e.g.,
 - Theory/simulation project: formalize your models
 - Data-analysis project: figure out where and how to get data and what you plan to do with it
 - Experiment/application project: have a prototype design and an evaluation plan
 - **Include a timeline** (weekly or biweekly) on what you plan to do

Logistics: More about Assignment 3

- Understanding conditional probability
- For problem 2 (b,c)
 - Grader 1 and grader 2 are reviewing the same assignment
 - Grader 1 receives signal B
 - What's the probability of grader 2 also receiving signal B?
- It's different from $\Pr(B)$, the probability for a grader to receive signal B for an assignment randomly drawn from all assignments

Logistics: Project Milestone 2

- Milestone 2: Due Nov 20
 - Summarize your progress
 - Last chance to convert a research project to an extensive literature survey
- What's the expected amount of literature survey?
 - Case by case
 - As a general guideline, the minimum number of papers
 - For research project: 3~5 papers (# papers in one lecture)
 - For literature survey project, 15~20 papers (# papers in one “topic”)

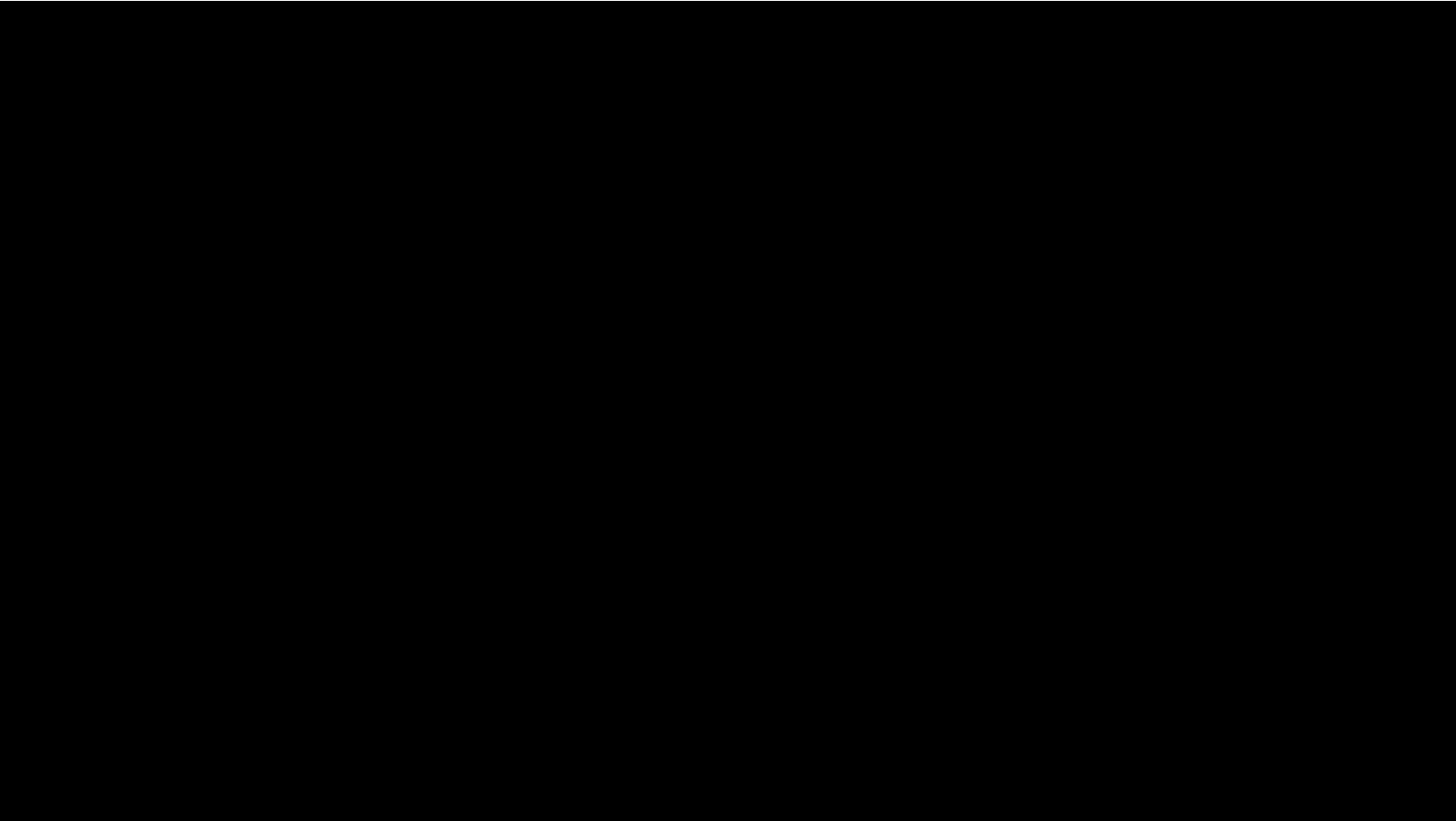
Logistics: Project Midterm Pitch and Discussion

- Milestone 2: Due Nov 20 (Friday)
- Midterm Pitch: Nov 19 (Thu), in lecture
 - More details to come
 - Tentative formats:
 - Every team gives a spotlight presentation for 1~2 min
 - The entire class splits into 2~3 big groups
 - Each team in turn obtains the entire attention of the big group
 - Using this time to collect data, obtain feedback, elicit suggestions, etc
 - Try to think about what you can do with this time in your milestone 1 plan and utilize it
 - Let me know early if it's helpful to utilize the entire class for help
 - Might be able to allocate a small amount of time in a need basis

Crowds in Two Seconds: Enabling Real-time Crowd-Powered Systems.

Michael Bernstein, Joel Brandt, Robert Miller, and David Karger. UIST 2011

Crowds in Two Seconds



Real-time Crowdsourcing: Why

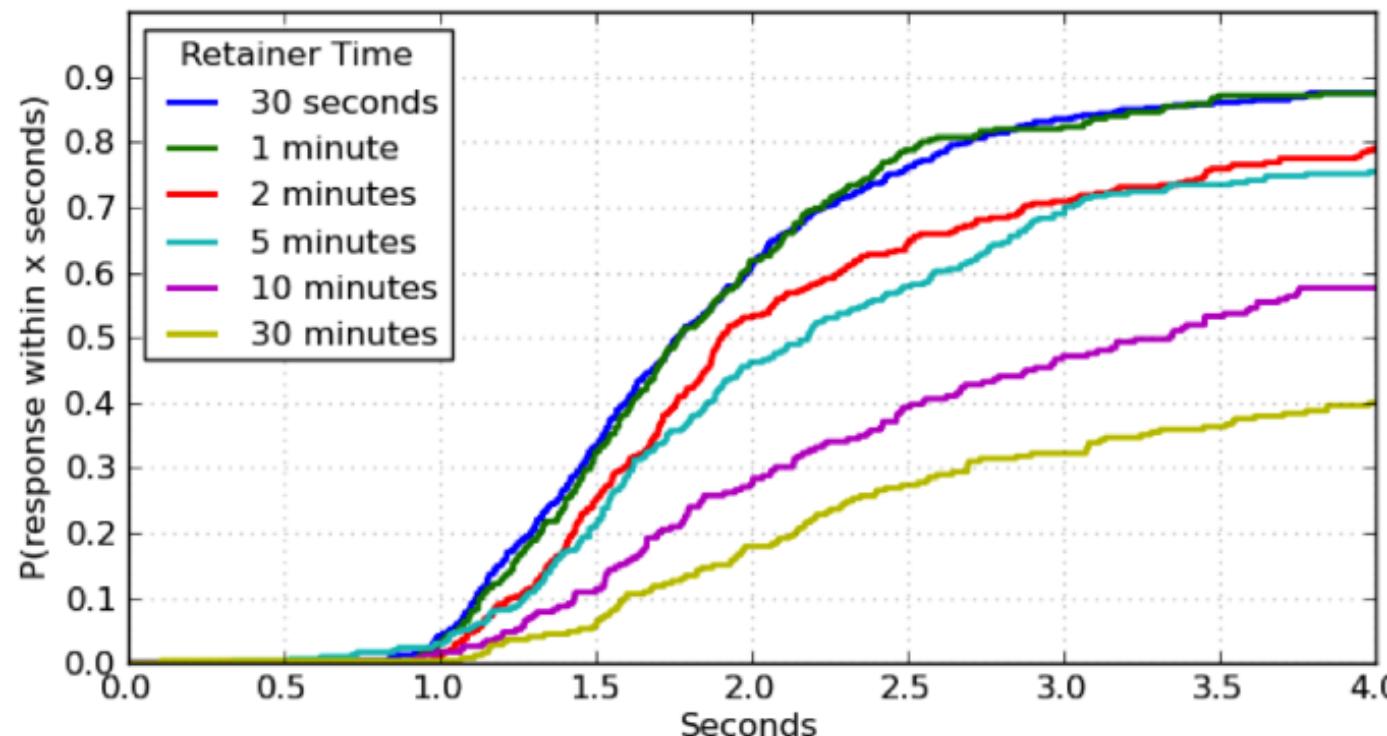
- Enable **interactive** systems
 - Users can't stand long waiting time



- Enable **synchronous** crowds
 - Multiple workers can work together for tasks
 - There might be other ways to achieve this:
 - Set a task start time, invite workers back at the given time

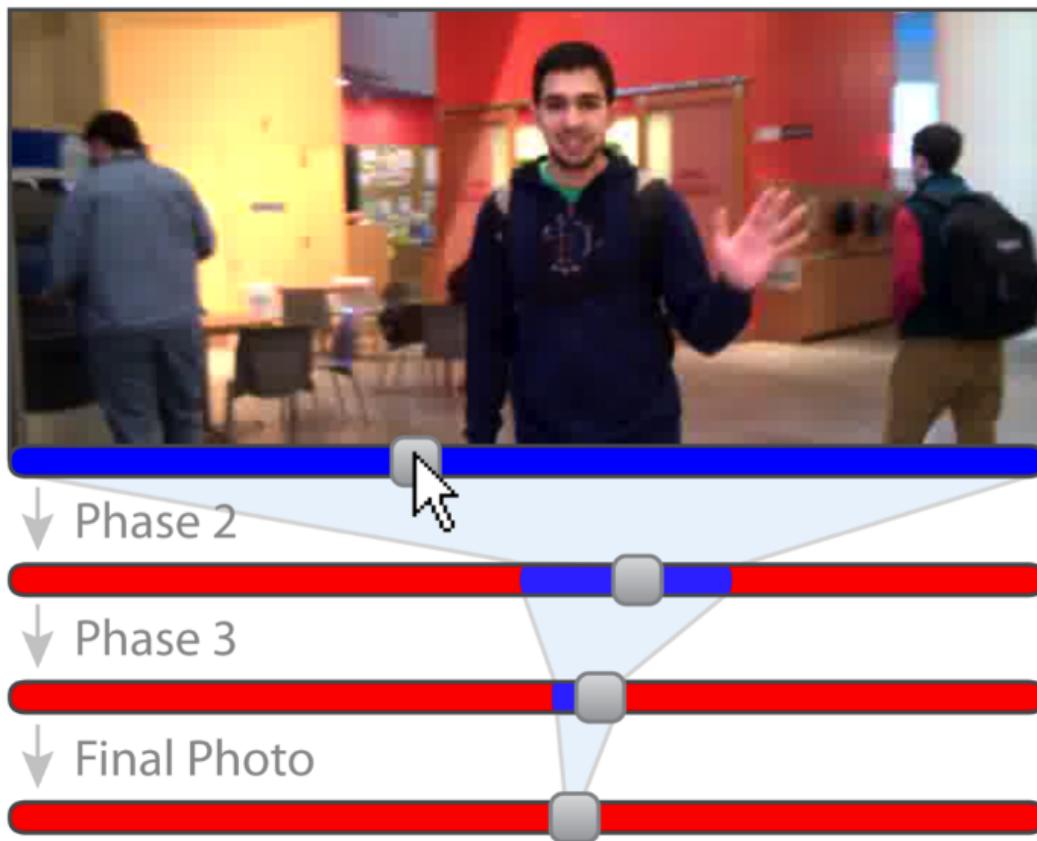
Real-time Crowdsourcing: How

- Retainer models
 - Pay workers to wait
 - Better interface design (alerts)



Real-time Crowdsourcing: How

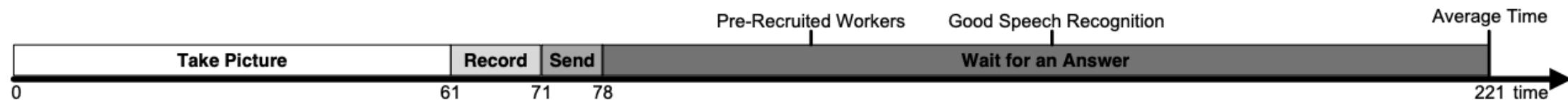
- Rapid refinement
 - Using AI to help focus on more “important” tasks



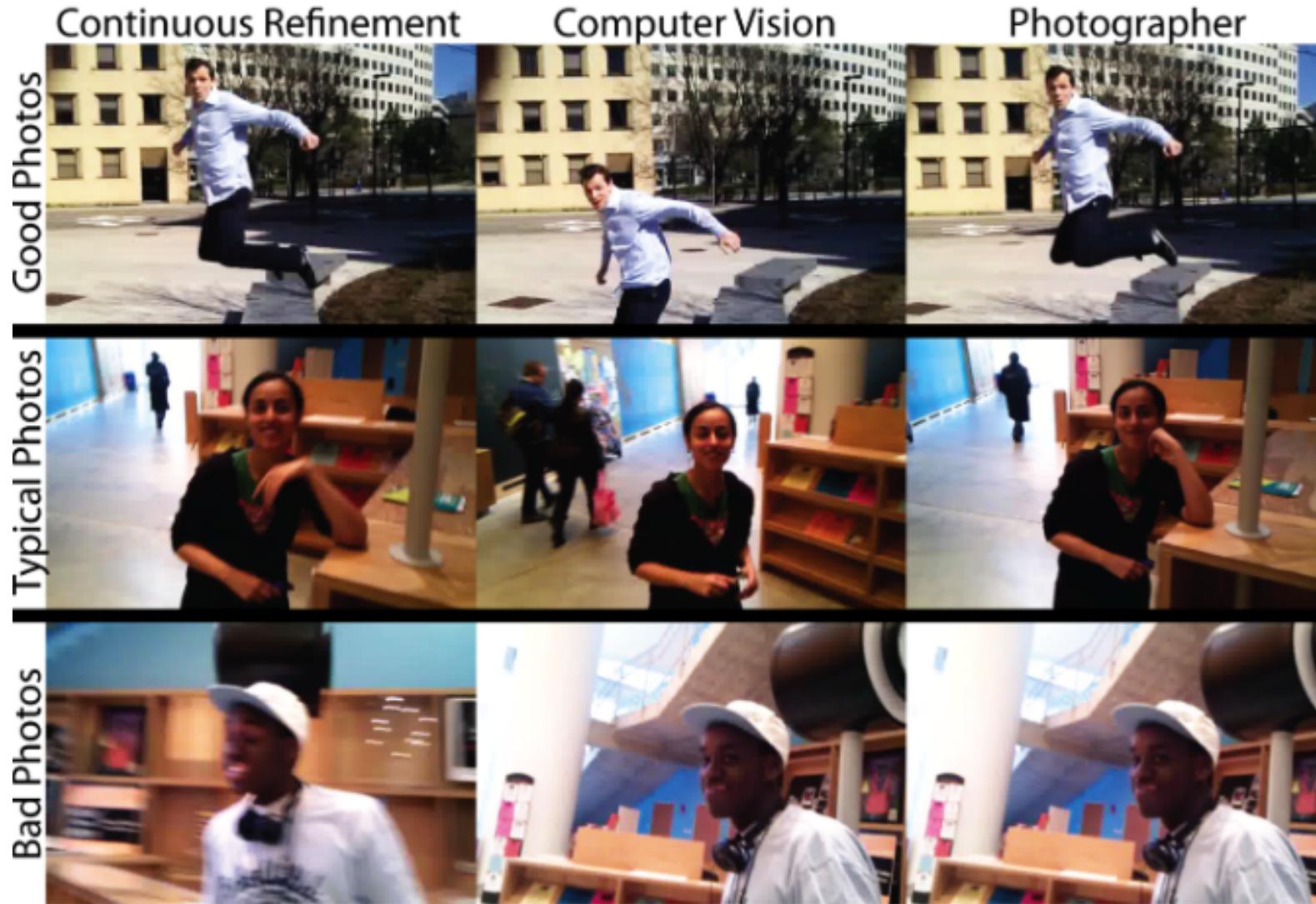
We might apply similar ideas for annotation tasks.

Real-time Crowdsourcing: How

- Predict the need [Bigham et al., UIST 2010]
 - Tasks involve a sequence of actions
 - Start recruiting the crowd before they are needed



Real-time Crowdsourcing: What



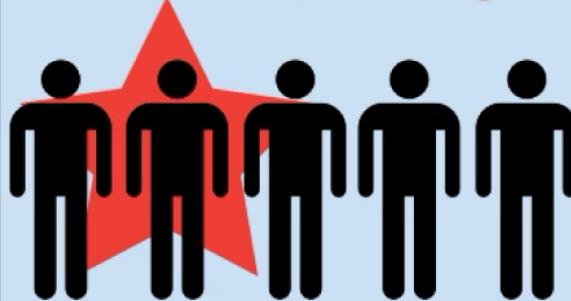
Real-time Crowdsourcing: What

Which font looks better?

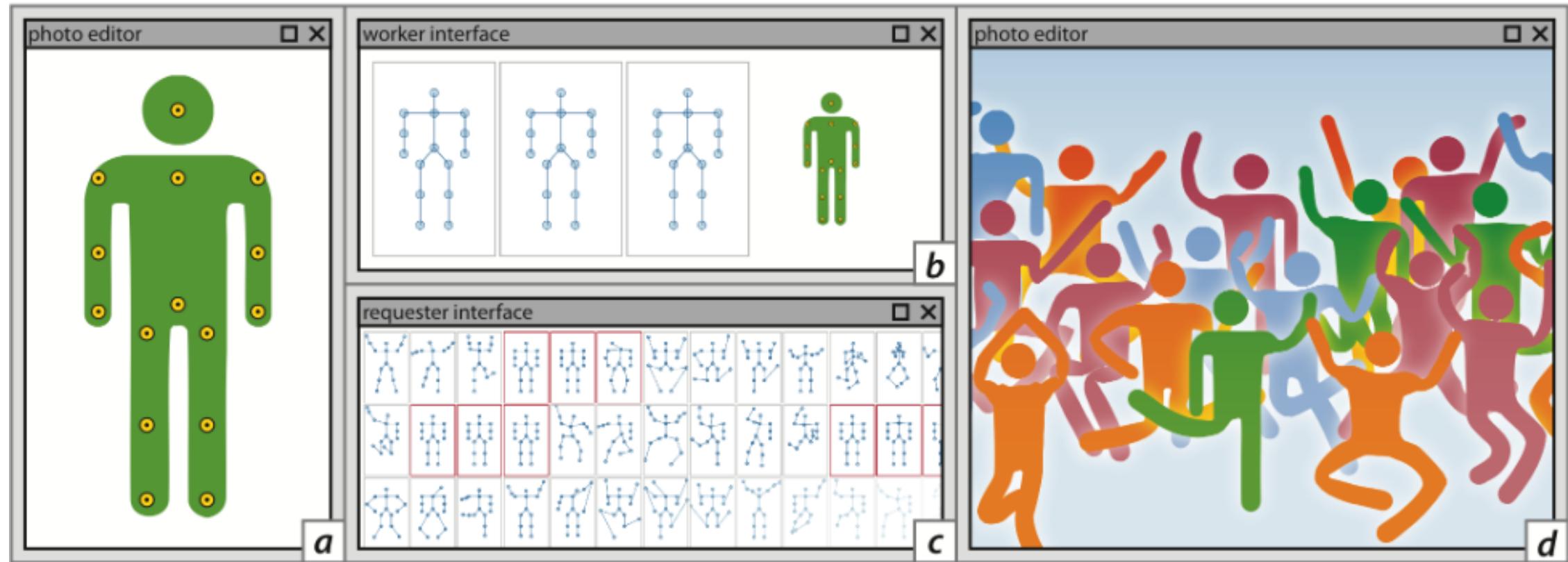
SuperStar
Shopping



SuperStar
Shopping



Real-time Crowdsourcing: What



Discussion

- Review questions
 - What are the other applications that could benefit with real-time crowdsourcing?
 - What do you think workers' opinions would be on the retainer model? What kind of worker population do you think this method would attract? And what might be the impacts to the data?
- Criticisms/improvements on the approaches they take?
- Other approaches to enable/improve real-time crowdsourcing?

A Few More Applications

VizWiz [Bigham et al. UIST 2010]

- Answer Blind People's Visual Questions

What color is this pillow?



What denomination is this bill?



Do you see picnic tables across the parking lot?



What temperature is my oven set to?

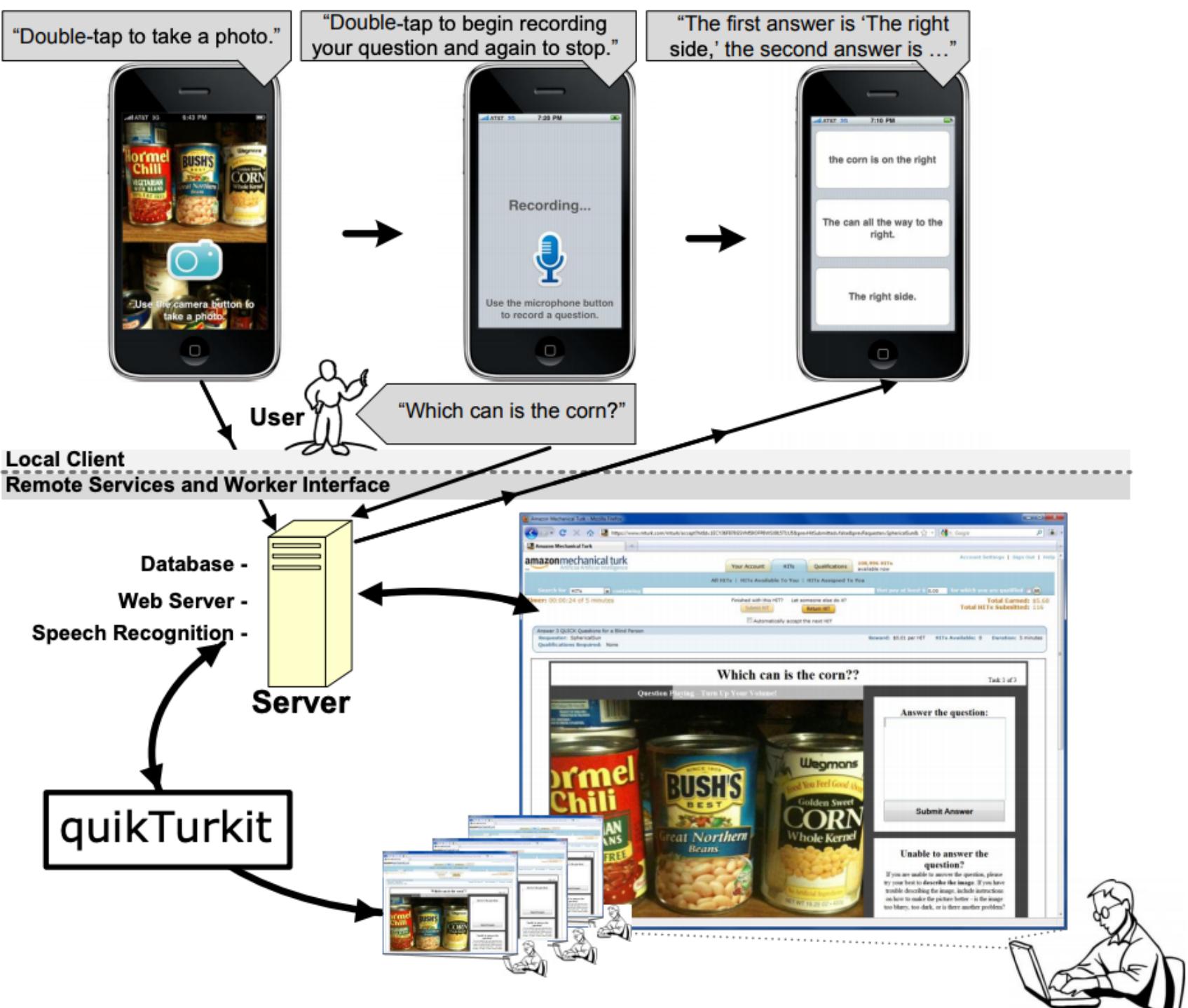


Can you please tell me what this can is?



What kind of drink does this can hold?





Enhancement to Real-time Crowdsourcing

- Predict the need
 - Recruit workers before the questions are asked



VizWiz [Bigham et al. UIST 2010]

- Leads to a long line of research to assist blind people
 - <https://vizwiz.org>



[Home](#) [Browse Dataset](#) [Tasks & Datasets](#) [Workshops](#) [Acknowledgments](#)

Overview

A common goal in computer vision research is to build machines that can replicate the human vision system; for example, to recognize and describe objects/scenes. A natural grand challenge for the computer vision community is to design such technology to assist people who are blind to overcome their real daily visual challenges. Towards this aim, we introduce the first datasets and artificial intelligence challenges originating from people who are blind to encourage a larger community to collaborate on developing algorithms for assistive technologies. In particular, we built the datasets with data submitted by users of a mobile phone application, who each took a picture and (optionally) recorded a spoken question about that picture. Ultimately, we hope this work will educate more people about the technological needs of people who are blind while providing an exciting new opportunity for researchers to develop assistive technologies that eliminate their accessibility barriers.

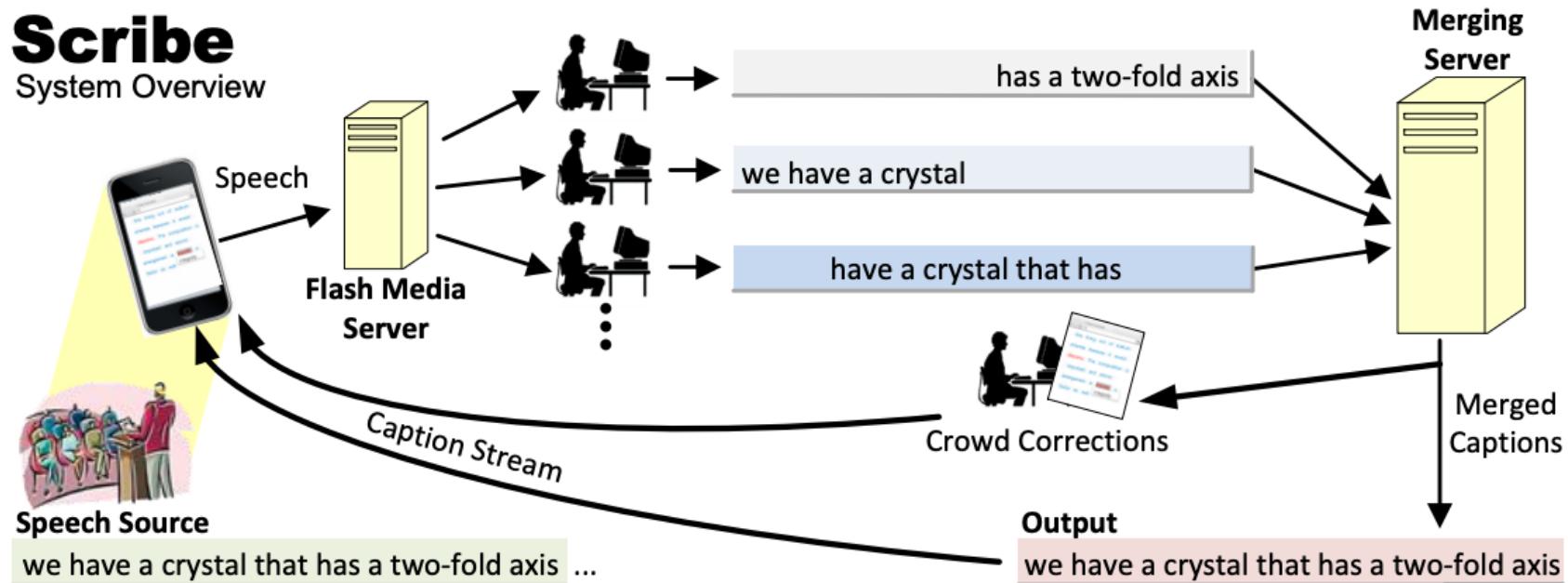
Publications

Real-time Captioning [Lasecki et al. UIST 2012]



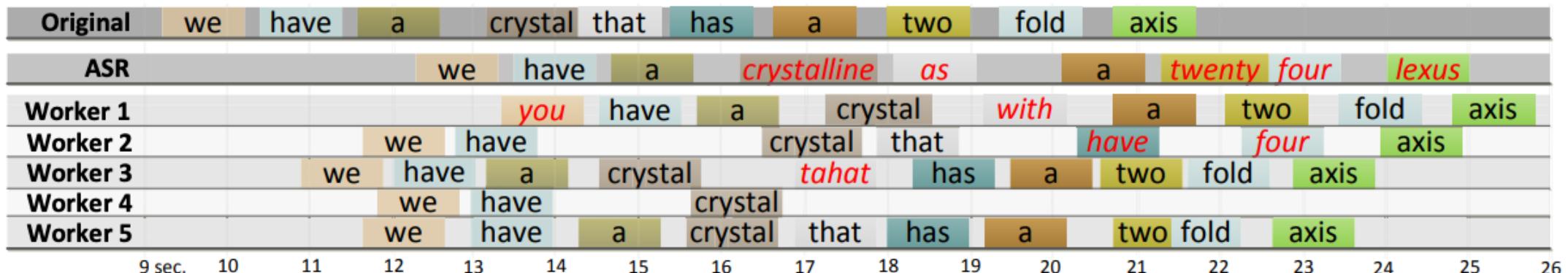
Real-time Captioning [Lasecki et al. UIST 2012]

- Trained expert to type over 225 words per minute
- Hard for a single un-trained expert to do real-time captioning
 - Utilize multiple workers
 - Each of them working on **overlapping** small segments (for verification)



Real-time Captioning [Lasecki et al. UIST 2012]

- Need to aggregate time-series of data



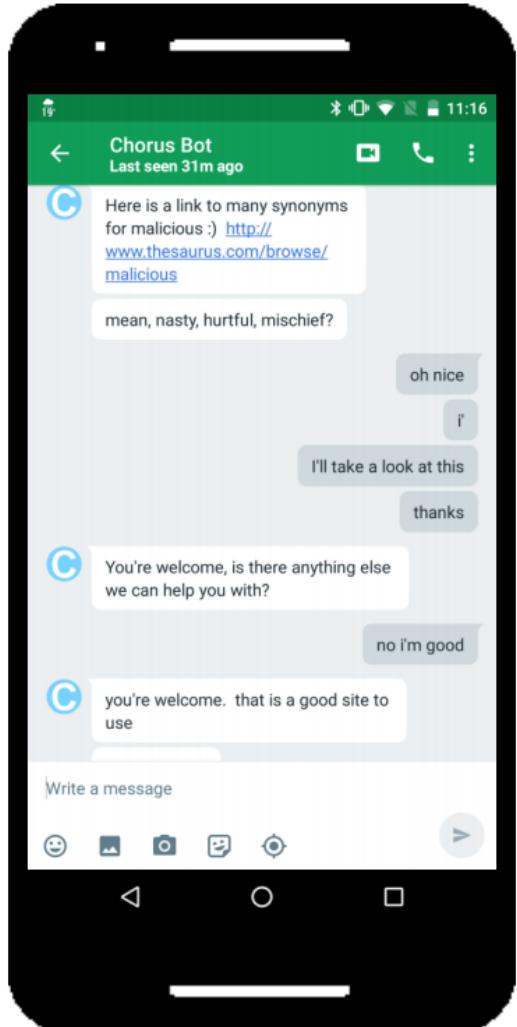
- Multiple sequence alignment

```
1: ---learn---g is such      a suitcase word though right so ----- has a lot of there ---s a lot
2: -o learning is such      a----- learning has ----- is a lot
3: learning ss such        a suitcase word though ----- learning has ----- is a lot
4: lea ning is su h        a----- right so learning ----- a lot
5: so learning is such     a suitcase ----- though learning has ----- lot
6: learning is such        a suitcfse word though right ----- this in a lot
F: so learning is such    a suitcase word though right so learning has a lot of there is a lot
```

Real-time Captioning [Lasecki et al. UIST 2012]

- Captioning might be solvable by AI now (especially for English)
- Main ideas might still apply for other applications:
 - Real-time translation, etc

Crowd-Powered Chatbot [Huang et al. HCOMP 2016]



- There are impressive progress in Chatbots, but it's still not good enough

Amazon launches \$2.5 million Alexa Prize for college students building bots

- A crowd play as a chatbot by
 - Propose responses
 - Vote for the best one
 - Take notes
- Intermediate solutions; dataset generation

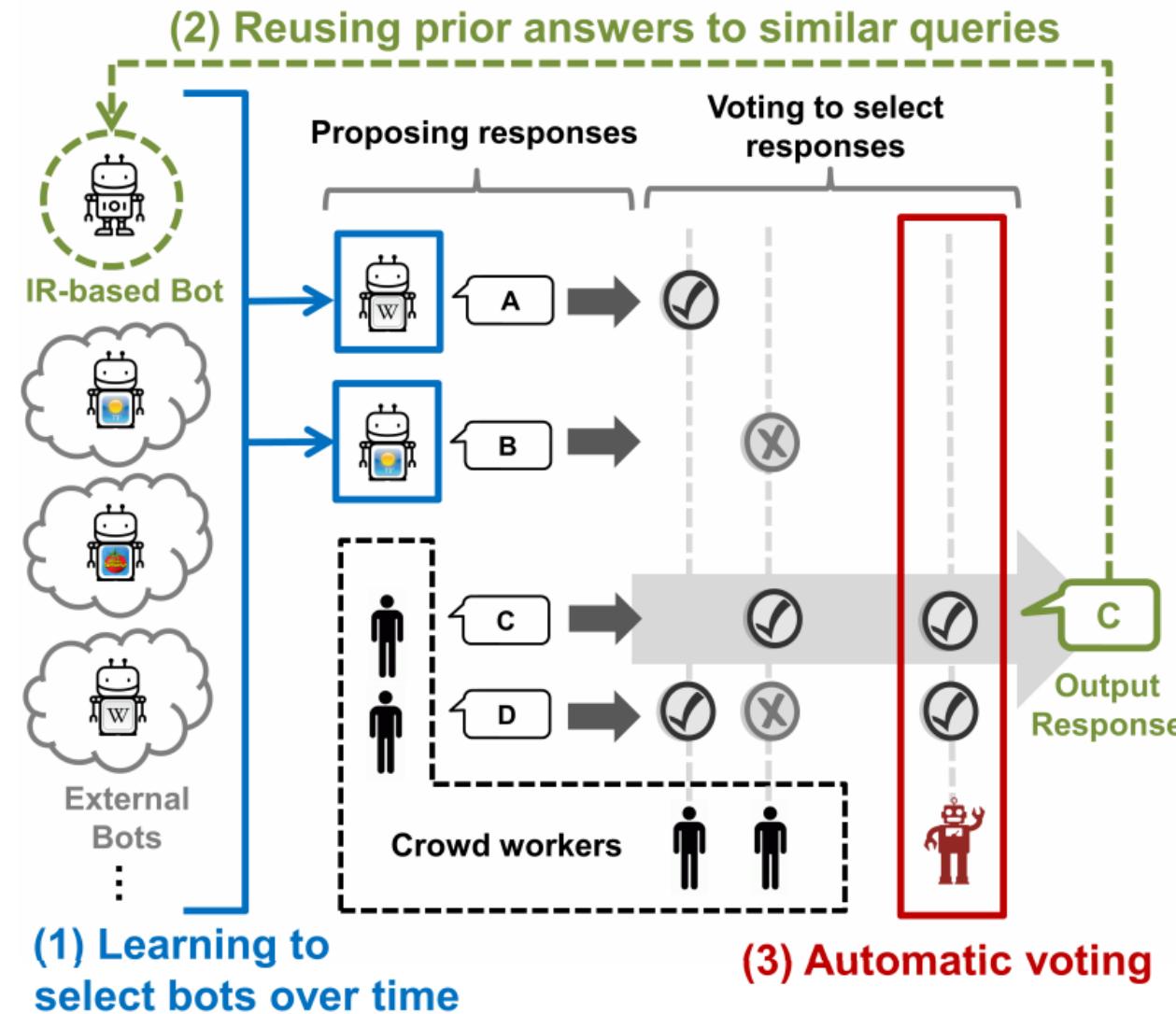
Crowd-Powered Chatbot [Huang et al. HCOMP 2016]

User	How many suitcases can I take on a flight from the US to Israel?	Up to four bags in Main Cabin on Delta and in all cabins on Delta Connection flights. Up to five bags in Delta One™, First and Business Class on Delta aircraft only. Approved Personal Items: 1 purse, briefcase, camera bag or diaper bag, or 1 laptop computer (computers cannot be checked), or 1 item of a similar or smaller size to those listed above	or all other international flights You may check up to two bags that meet our size & weight restrictions at no extra charge	are you flighing [sic] on Delta One™, First, Business class?
Chorus	Let me check. Can I ask you from where are you planning to board the flight?			here is the detail - Size & Weight Restrictions. To avoid extra charges for oversize or overweight baggage, your checked bag must: weigh 50 pounds (23 kg) or less. not exceed 62 inches (157 cm) when you total length + width + height.
User	Pittsburgh			No, just a regular flight (economy)
Chorus	with which company are you flying? and which air services are you using?			Then yes, you will have one bag for free I am afraid you'll have to pay extra. for the other bag. AirTran Airways
User	Delta airlines			First Checked Bag: \$20 each way for all economy-class reservations. Second Checked Bag: \$25 each way for all economy-class reservations. and the second is \$100
Chorus	You may bring one carry-on item onboard the aircraft, plus one personal item If you are: 1) active duty U.S. military personnel on orders to or from duty stations and dependents traveling with them; or 2) active U.S. military dependents traveling on relocation orders, you may check the following at no charge:	User What about checked in baggage? I should be able to bring at least one suitcase, no? I'm not any of those Chorus one personal item mentioned above is allowed. it could be 1 purse, briefcase, camera bag or diaper bag or 1 laptop computer (computers cannot be checked) or 1 item of a similar or smaller size to those listed above.	User Oh, so I get two bags for free to Israel? Chorus < URL of Delta's Web Page of Baggage Policy > Yes you can check all th [sic] details here Southwest allows two (2) checked pieces of baggage per ticketed Customer Yes. hi how can I help u? User In the link it says that to Israel, the second bag is \$100. I'm confused.... Chorus yes They charge you for extra baggage.	User Too bad. OK thanks!

Auto-Evolving Crowd Chatbot [Huang et al. CHI 2018]

- Train AI chatbots with historical data
- A crowd consists of both humans and AIs
 - Propose responses
 - Vote responses
- May even consider dynamically determining the ratio of AIs

Auto-Evolving Crowd Chatbot [Huang et al. CHI 2018]



AI Task Delegability [Lubars and Tan. NeurIPS 2019]

- Determine what tasks should be done by AI and what tasks should be done by humans.
-

**Ask not what AI can do, but what AI should do:
Towards a framework of task delegability**

Brian Lubars
University of Colorado Boulder
brian.lubars@colorado.edu

Chenhao Tan
University of Colorado Boulder
chenhao.tan@colorado.edu

Twitch Plays Pokemon

- <https://www.twitch.tv/twitchplayspokemon>



Twitch Plays Pokemon

- Guinness World Record:
 - "the most participants on a single-player online videogame" with 1,165,140
- Best Fan Creation in Game Award 2014
- An application “for fun”, not to “solve problems”
- Highlight an interesting question
 - How should we decide which action to take?
 - Challenging for sequential actions under real-time constraints

Example: Crowd Robot Navigation

[Lasecki et al. UIST 2011]

Real-time Crowd Control of Existing Interfaces

Walter S. Lasecki, Kyle I. Murray, Samuel White
Robert C. Miller and Jeffrey P. Bigham



hci.cs.rochester.edu



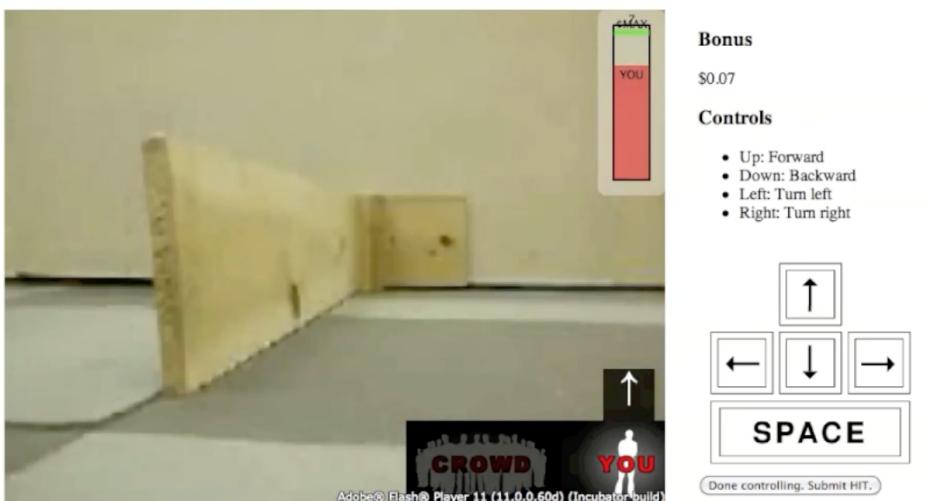
Massachusetts Institute of Technology

groups.csail.mit.edu/uid/

Discussion

- What do you think are the efficient ways for crowd control?
 - You can use the robot navigation as the application.

Goal: Drive into the tea kettle. Complete this task as quickly as you can!



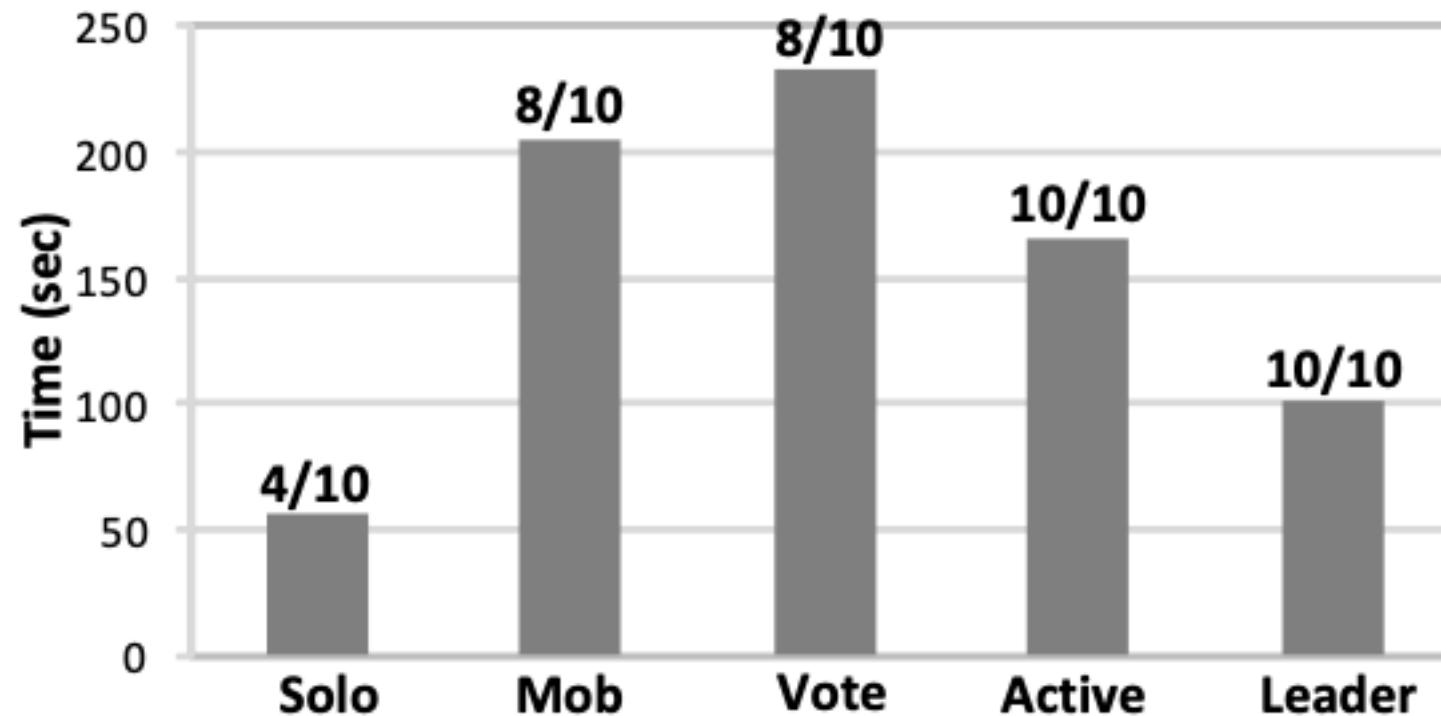
- Or imagine you are together with 10 other people in a critical environment. You together need to make a series of decisions to survive under time pressure. What's the mechanism you think you should use?

Crowd Robot Navigation [Lasecki et al. UIST 2011.]

- Solo: Hire a single worker
- Mob: Pass on all inputs to the original interface (first come first serve)
- Vote: Follow actions by majority vote
- Leader:
 - Update “weights” of users based on agreements with the crowd
 - Follow the actions by the leader (which might change over time)
- Active:
 - Randomly select a worker as the leader, change worker if the input is not in on time

Crowd Robot Navigation [Lasecki et al. UIST 2011.]

- Performance measure
 - Number of success (out of 10)
 - Average time to complete the task when success



Generally, how to aggregate preferences from a crowd?

Social choice theory

Consider Elections

- Three candidates: A, B, C
- Three users: 1, 2, 3
- Users' preferences:
 1. $A > B > C$
 2. $B > A > C$
 3. $C > A > B$
- How should we design a mechanism to decide who wins the election

Arrow's Impossibility Theorem

- Some nice criteria we want to have:
 - **Unanimity**: If every person prefers A over B, then the group prefers A over B
 - **Non-dictatorship**: no person's preference is always strictly preferred than others
 - **Independence**: If for two sets of preferences, A and B have the same order between sets, A and B should have the same order in the group decision
- Arrow's Impossibility Theorem
 - No mechanism satisfies the three criteria when the number of candidates ≥ 3