Deep Integration of Human and Machine Intelligence for Accessibility

“Time Cook”

Jeffrey P. Bigham
What do blind people want to know?
73. What is this can
74. Hi is there an error message on this screen; and what does the screen say; or at least as much of it as you can see. Thank you
75. Hello could you please tell me what's on this box? Thanks a lot
76. Could you please tell me what is in this jar thank you
77. (no audio)
78. What's this?
79. What is that please
80. -this envelope addressed to
145. Can you identify this for me please

146. What color is this shirt

147. What's the title of this essay

148. What is the color

149. Is this set to 'fan'

150. Is the air conditioner set to air; fan; or dehumidifier

151. What is the phone number for the Contact Centre

152. -s how to cook these
157. What is this?
158. Are you a computer?
159. What does this shirt say
160. What's on this page?
161. What is in this can.
162. - his name.
163. What does this part of the sky look like
164. What's this and how long do I cook it for?
185. What's in this box?

186. What is the top button on the left?

187. I'd like to know what this can is; please

188. Is this diet or regular coke? Thanks.

189. - card number.

190. Okay I'm just curious what the controls are for this machine that's in front of me; I can feel that there are two rows of buttons; so can you tell me what the buttons are from left to right for each row?

191. What's the weather like outside on this photo? Thank you.

192. What is in this box?
VizWiz is an iPhone app that allows blind users to receive quick answers to questions about their surroundings. VizWiz combines automatic image processing, anonymous web workers, and members of the user’s social network in order to collect fast and accurate answers to their questions.

VizWiz dataset was collected using the VizWiz application, which was released in May 2011.

- Users have asked more than 100,000 questions, 48,169 of which were asked by users who agreed to allow their questions to be anonymized and shared. We carefully and redundantly checked each and removed those that contained personally identifying information, resulting in 43,543 remaining questions.
- Each question includes an image, a transcription of the question, and several answers.
- Answers come primarily from Amazon Mechanical Turk, if chosen by the user; other resources include IQ Engines, Facebook, Twitter, or email.

Sample JSON for an image:
```
{
  "id": "ag12aXo3aXo3c9jeWFscgQLEgLgYWryeR3Z8BA",
  "sequence": "ag12aXo3aXo3c9jeWFscgQLEgLgYWryeR3Z8BA",
  "question": "How long do I cook this?",
  "answers": [
    {
      "answer": "9 minutes",
      "correct": true,
      "source": "web workers"
    },
    {
      "answer": "meatloaf",
      "correct": false,
      "source": "IQ engines"
    }
  ],
  "canAnswer": true,
  "blurry": false,
  "subjectOutOfFrame": false,
  "textQuestion": true,
  "identificationQuestion": false,
  "colorQuestion": false,
  "subjectiveQuestion": false
}
```

Downloads
- Download training set (15Gb, 33,543 images)
- Download training JSON (33,543 questions & answers)
Using Devices
User + AI + Crowds

A. Initial Crowdsourced Segmenting and Labeling

B. Real-Time Recognition and Control

- Object Localization
- Fingertip Detection
- Information Lookup
- Providing Feedback
- Providing Guidance
Choosing "Add Interface" in app..
3D Hubs printing, 20x speed
User aligning overlay with control panel..
What do blind people want to know?
Working with AI
Working with AI

Target What People Want to Know

Working with AI

Errors are Important and Not Created Equally

Measure and Express Confidence Well

…a man doing a trick on a skateboard at night.

The image shows Hillary Clinton walking onto a stage in front of a large crowd of people.

Working with AI

* Amplify Users with AI Support  (blind photography)

* Solve Real Problems  (understand user needs)

* Embrace and Understand Errors  (all errors are not equal)

* Measure Confidence and Express it Well
* Amplify Users with AI Support (blind photography)

* Solve Real Problems (understand user needs)

* Embrace and Understand Errors (all errors are not equal)

* Measure Confidence and Express it Well

Funded by: Google, Microsoft, NSF, NIDILRR