Deep Integration of Human and Machine Intelligence for

Jeffrey P. Bigham

Human-Computer Interaction Institute
CMU ACCESS

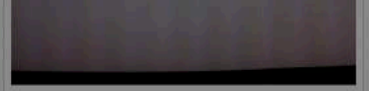
What do blind people want to know?

Jeffrey P. Bigham
@jeffbigham

Human-Computer Interaction Institute

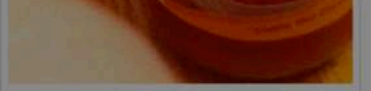
CMU ACCESS

73. What is this can



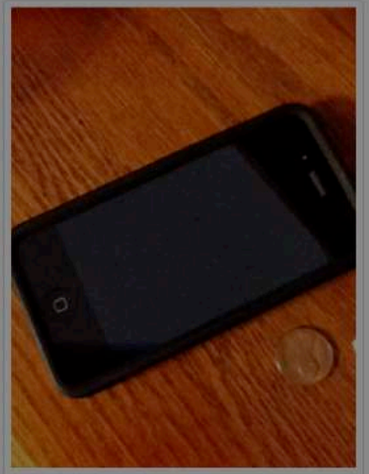
74. Hi is there an error message on this screen; and what does the screen say; or at least as much of it as you can see. Thank you

75. Hello could you please tell me what's on this box? Thanks a lot

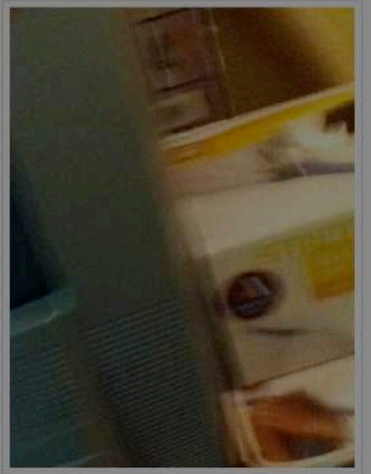


76. Could you please tell me what is in this jar thank you

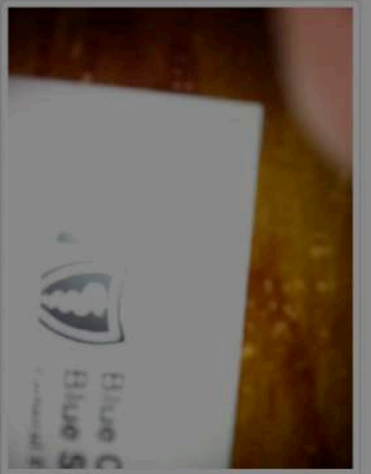
77. (no audio)



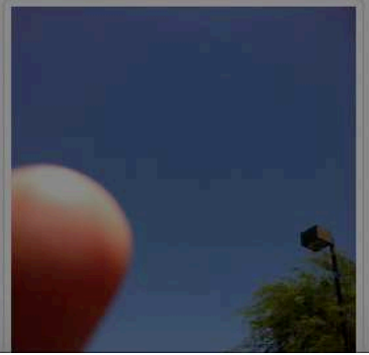
78. What's this?

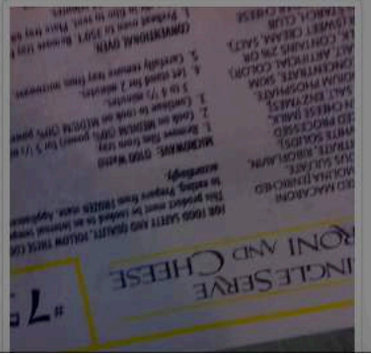


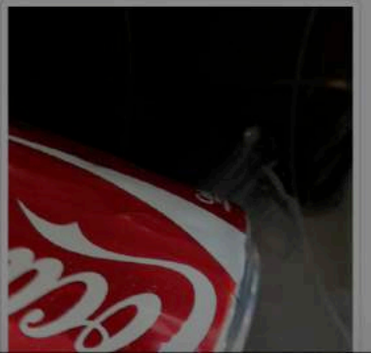
79. What is that please

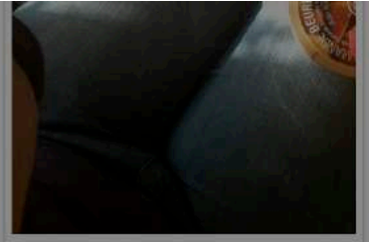


80. -this envelope addressed to

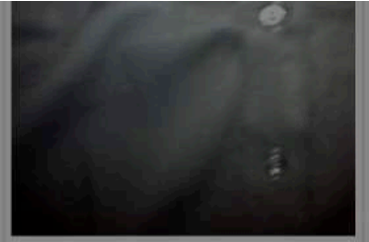




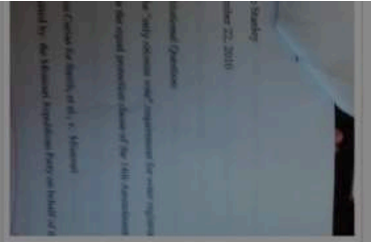




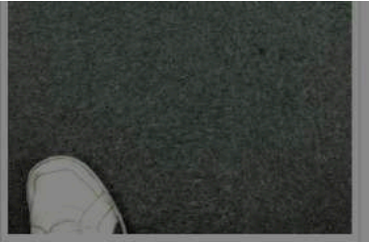
145. Can you identify this for me please



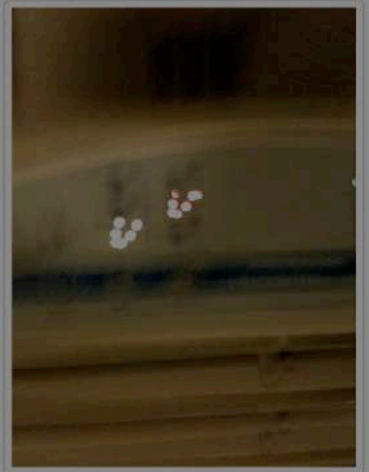
146. What color is this shirt



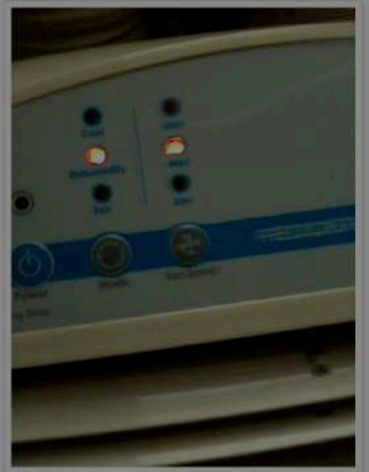
147. What's the title of this essay



148. What is the color



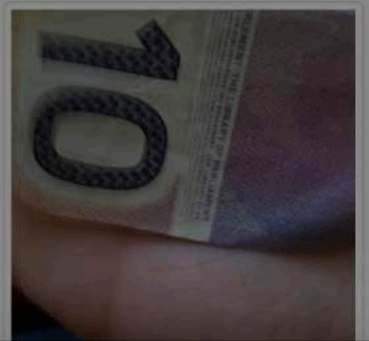
149. Is this set to 'fan'

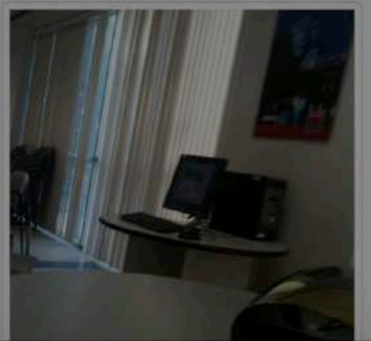


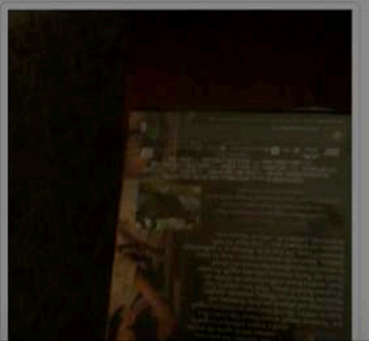
150. Is the air conditioner set to air; fan; or dehumidifier

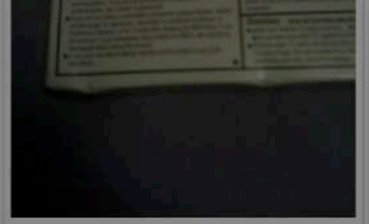
151. What is the phone number for the Contact Centre

152. -s how to cook these

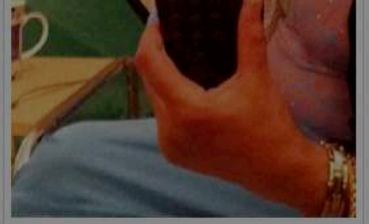




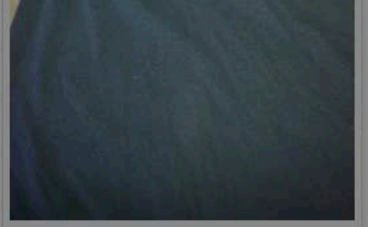




157. What is this?

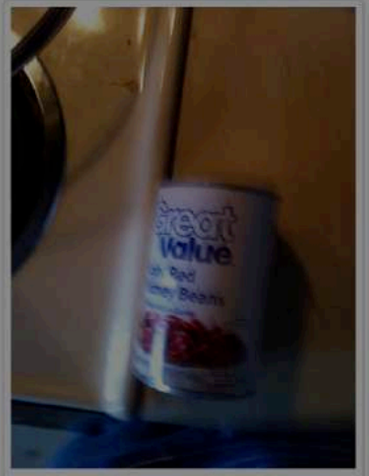


158. Are you a computer?



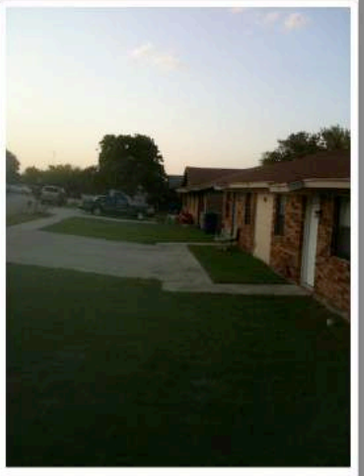
159. What does this shirt say

160. What's on this page?

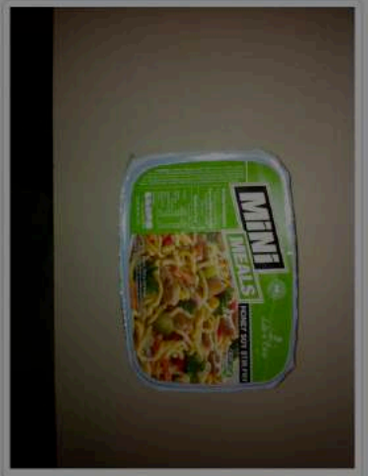


161. What is in this can.

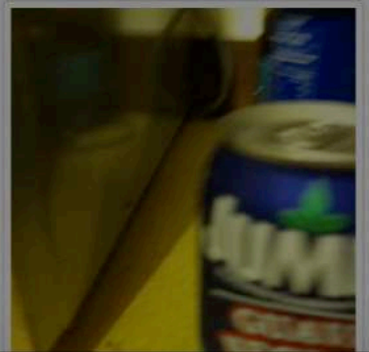
162. - his name.

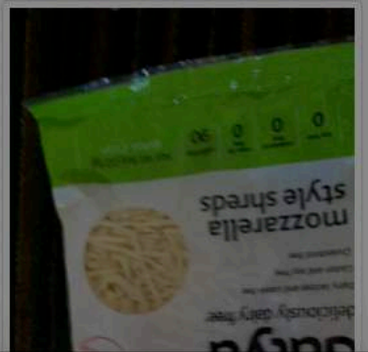


163. What does this part of the sky look like



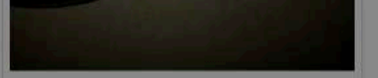
164. What's this and how long do I cook it for?



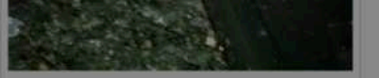


185. What's in this box

186. What is the top button on the left?



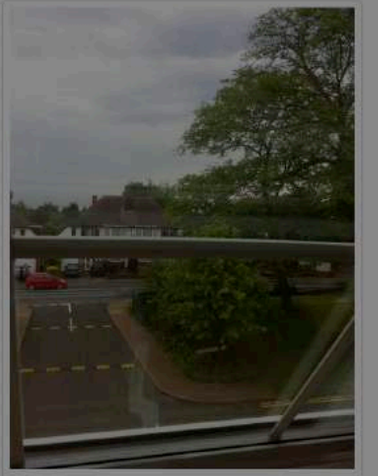
187. I'd like to know what this can is; please



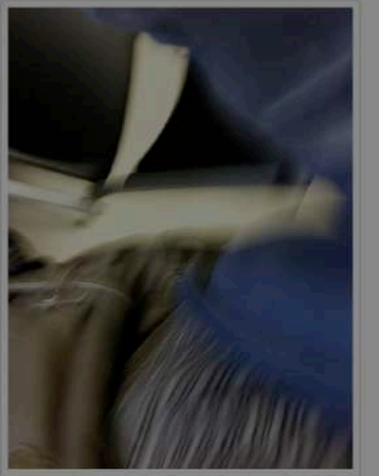
188. Is this diet or regular coke? Thanks.

189. - card number.

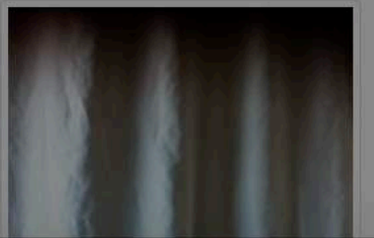
190. Okay I'm just curious what the controls are for this machine that's in front of me; I can feel that there are two rows of buttons; so can you tell me what the buttons are from left to right for each row?

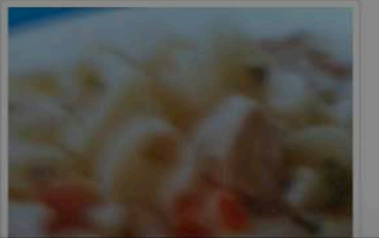


191. What's the weather like outside on this photo? Thank you.



192. What is in this box?





VizWiz Dataset

vizwiz.org/data

Check out Microsoft's Seeing AI!

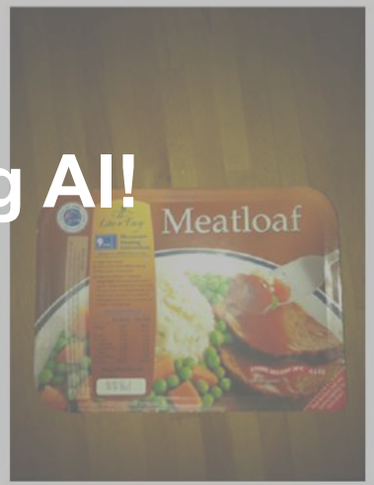
VizWiz Dataset | Dataset | Research | Examples | Contact

VizWiz Dataset

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:** "ag12aXp3aXotc29jaWFscg0LEgVRdWVyeRi9g1oM",
sequence:
["ag12aXp3aXotc29jaWFscg4LEgVRdWVyeRjT3Z8BDA"],
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)

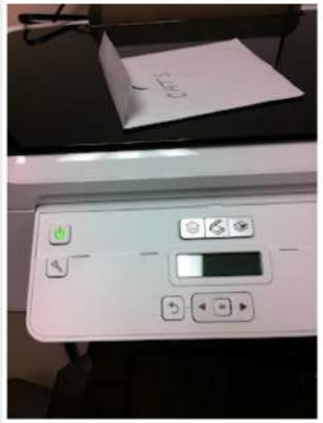
Jeffrey P. Bigham
@jeffbigham

Human-Computer Interaction Institute

CMU ACCESS

Using Device

User + AI + Crowd



186. What is the top button on the left?

20. Which knob; is it the left or the right; is for temperature control?

362. Which button says power? And have I got it set to delicate or daily?

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?

221. Could you tell me what the time remaining on this microwave is; it should be in seconds; like 33 seconds or 1 minute and 33 seconds or something like that. Thank you.

92. What is this machine and what are the buttons labeled as

233. Okay let's try this a second time; can you tell me all the buttons on this crockpot; what they are; left to right; up and down; and also -

665. (no audio)

Bigham, Jeffrey, et al. "VizWiz: nearly real-time answers to visual questions." UIST 2010.

Brady, Erin, et al. “Visual challenges in the everyday lives of blind people.” CHI 2013.

X

Jeffrey P. Bigham
@jeffbigham

Human-Computer Interaction Institute

CMU ACCESS

"Take photo"

Web Server & Database

A. Initial Crowdsourced Segmenting and Labeling

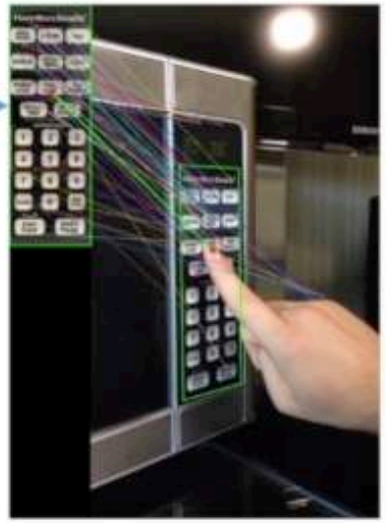
Segmenting

Labeling

B. Real-Time Recognition and Control

"Weight Defrost"

Computer Vision Server



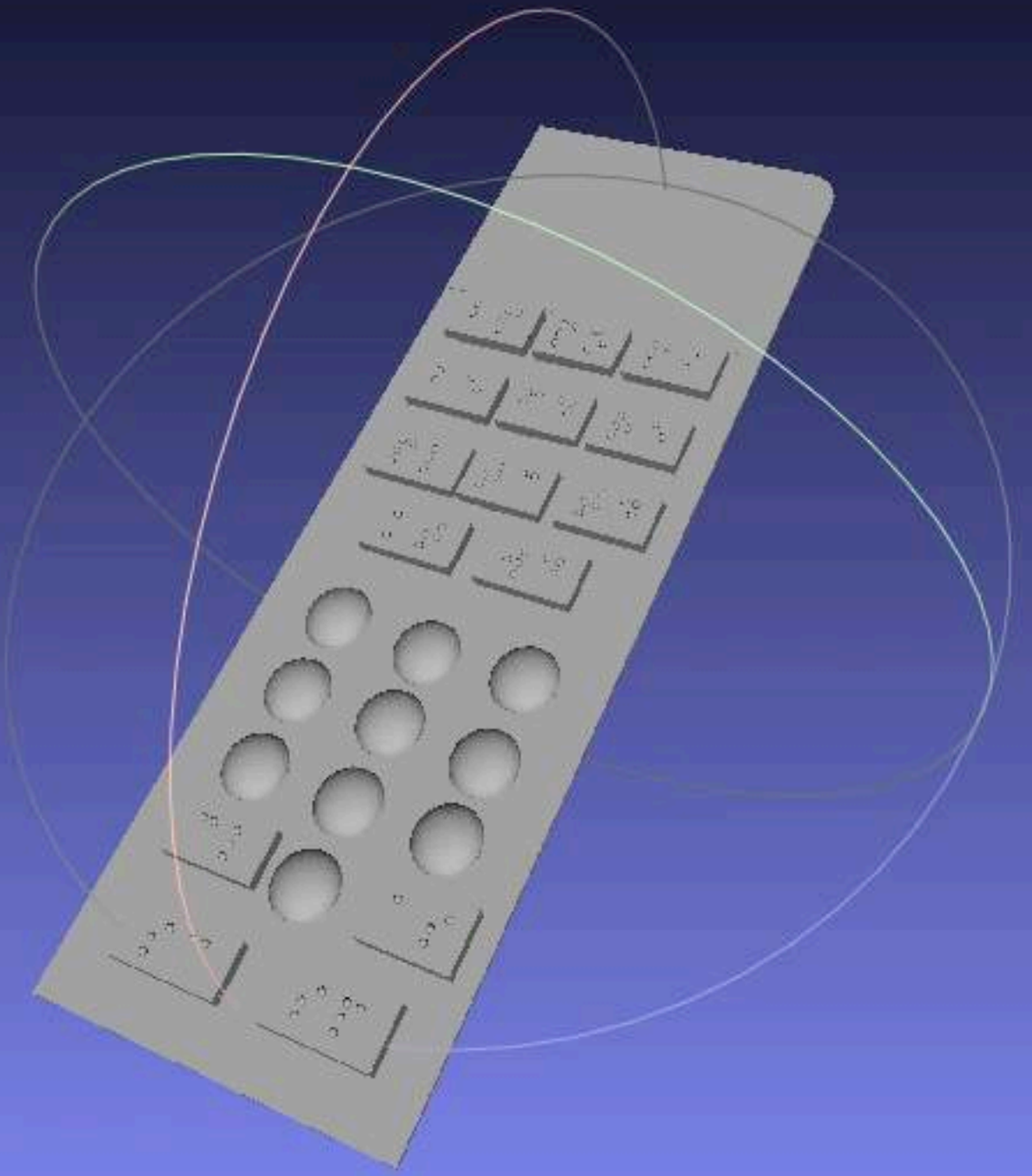
- Object Localization
- Fingertip Detection
- Information Lookup
- Providing Feedback
- Providing Guidance

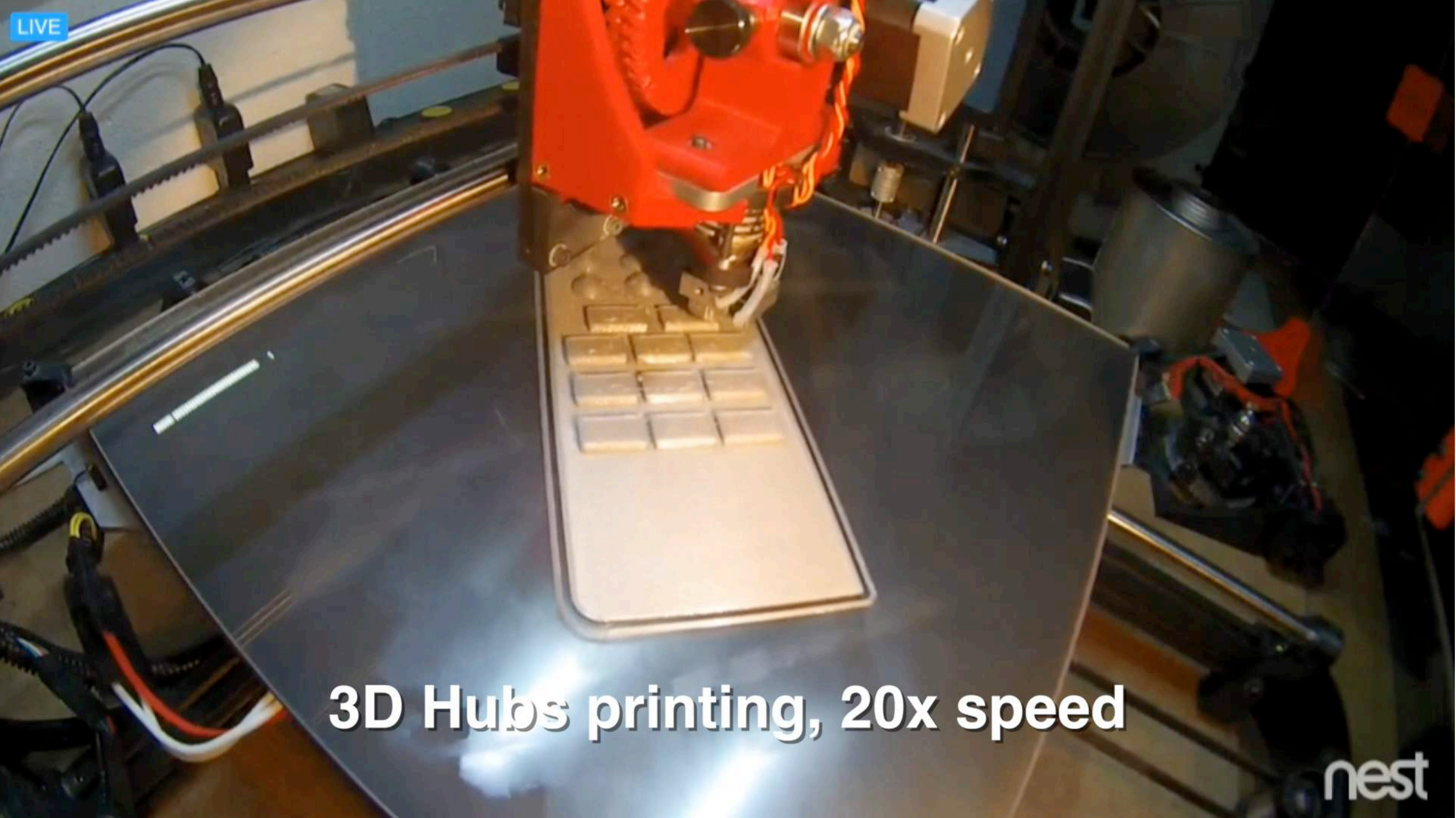
Local Client

Remote Services

Vending machine

Choosing "Add Interface" in app..




LIVE
3D Hubs printing, 20x speed
nest

User aligning overlay with control panel..

What do blind people want to know?

Jeffrey P. Bigham
@jeffbigham

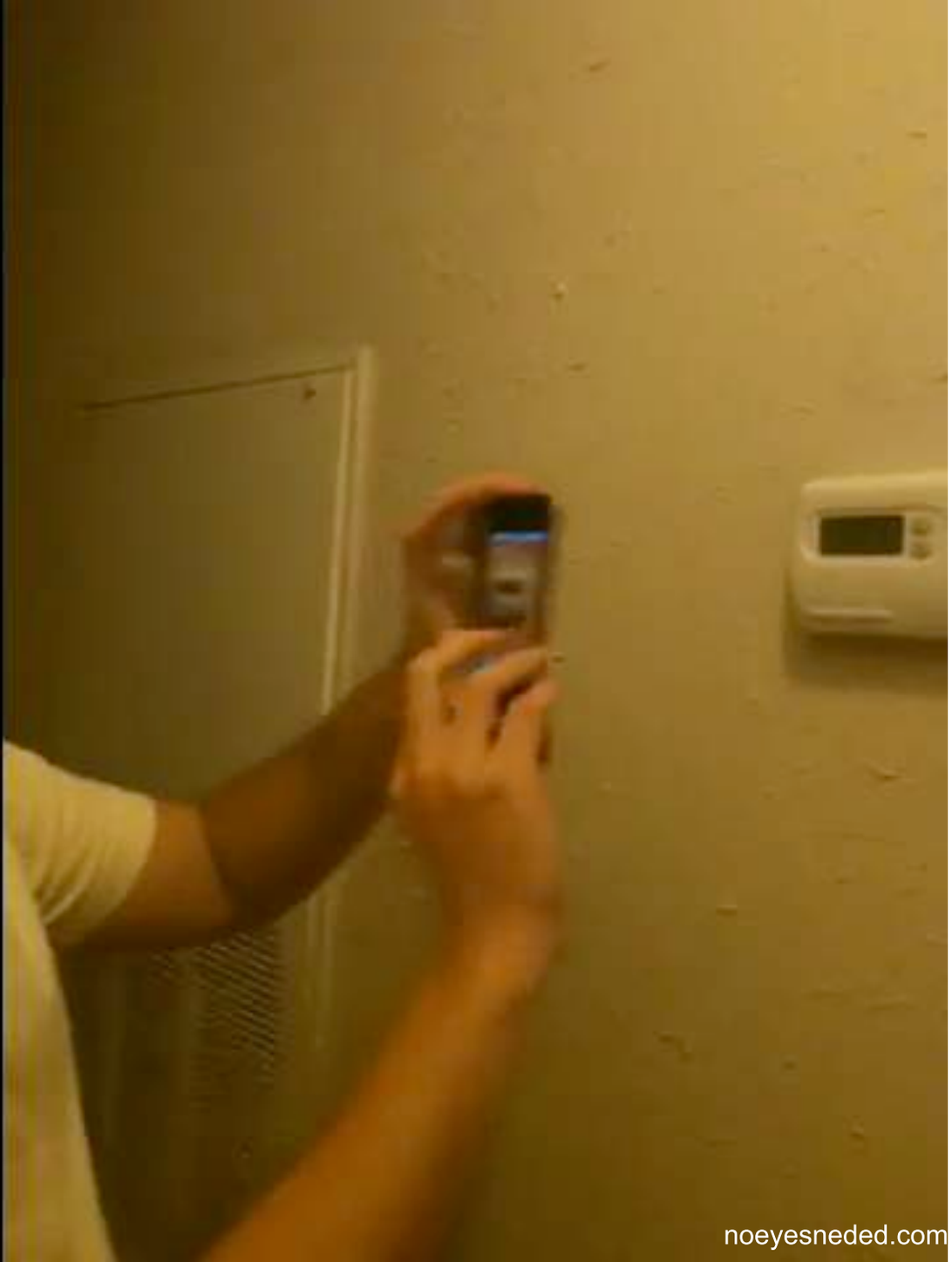
Human-Computer Interaction Institute

CMU ACCESS

Working with AI

Human-
Computer
Interaction
Institute

CMU ACCESS



Working with AI

Supporting Blind Photograph

Supporting Blind Photography. *Jayant, Chandrika, Ji, Hanjie, White, Samuel, Bigham, Jeffrey P.* ASSETS 2011.

Human-Computer Interaction Institute

CMU ACCESS

Working with AI

Target What People Want to Know

Wu, *et al.* Automatic Alt-text: Computer-generated Image Descriptions for Blind Users on a Social Network Service. CSCW 2017.

Human-Computer Interaction Institute

CMU ACCESS

Working with AI

Errors are Important and Not Created Equ
Measure and Express Confidence Well

…a man doing a trick on a skateboard at night.		The image shows Hilary Clinton walking onto a stage in front of a large crowd of people.

MacLeod, Bennett, Morris, Cuttrell, *et al.* Understanding Blind People's Experiences with Computer-Generated Captions of Social Media Images. CHI 2017.

Human-Computer Interaction Institute

CMU ACCESS

Working with AI

* **Amplify Users with AI Support** (blind photograp
* **Solve Real Problems** (understand user needs)
* **Embrace and Understand Errors** (all errors ar
* **Measure Confidence and Express it Well**

Human-Computer Interaction Institute

CMU ACCESS

The End

VizWiz Dataset:
http://www.vizwiz.org/data

VizWiz Publications:
http://www.jeffreybigham.com

*** 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

Jeffrey P. Bigham
@jeffbigham

Human-Computer Interaction Institute

CMU ACCESS