

*Is the coffee machine in use?*



*Is the trashcan full [...]?*



*Is someone using a printer?*



*How many people are in the line [...]?*



# Crowd-AI Camera Sensing in the Real World

*Ubicomp 2018, Singapore / October 8-12*

*Anhong Guo, Anuraag Jain, Shomiron Ghose, Gierad Laput, Chris Harrison, Jeffrey Bigham*

**Carnegie Mellon**









How many stools are occupied?

A A A A





Are any of the sofa's available?

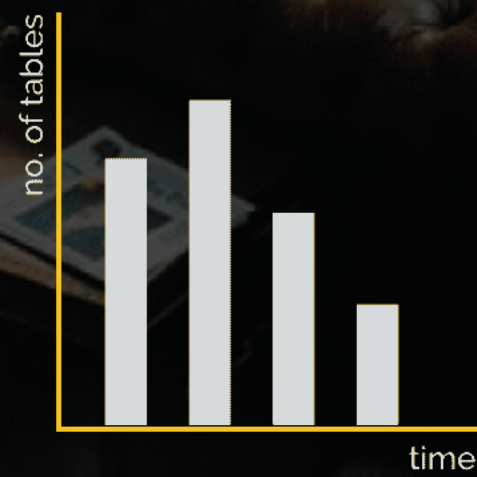
No







How many people are seated outside?





How many empty cups are there?















America's best network.  
Can your network say that?  
Better matters.  
verizon

MISSION

ONE WAY

Click Here  
to get started



# Current Solutions



Affordable



Flexible



Specific

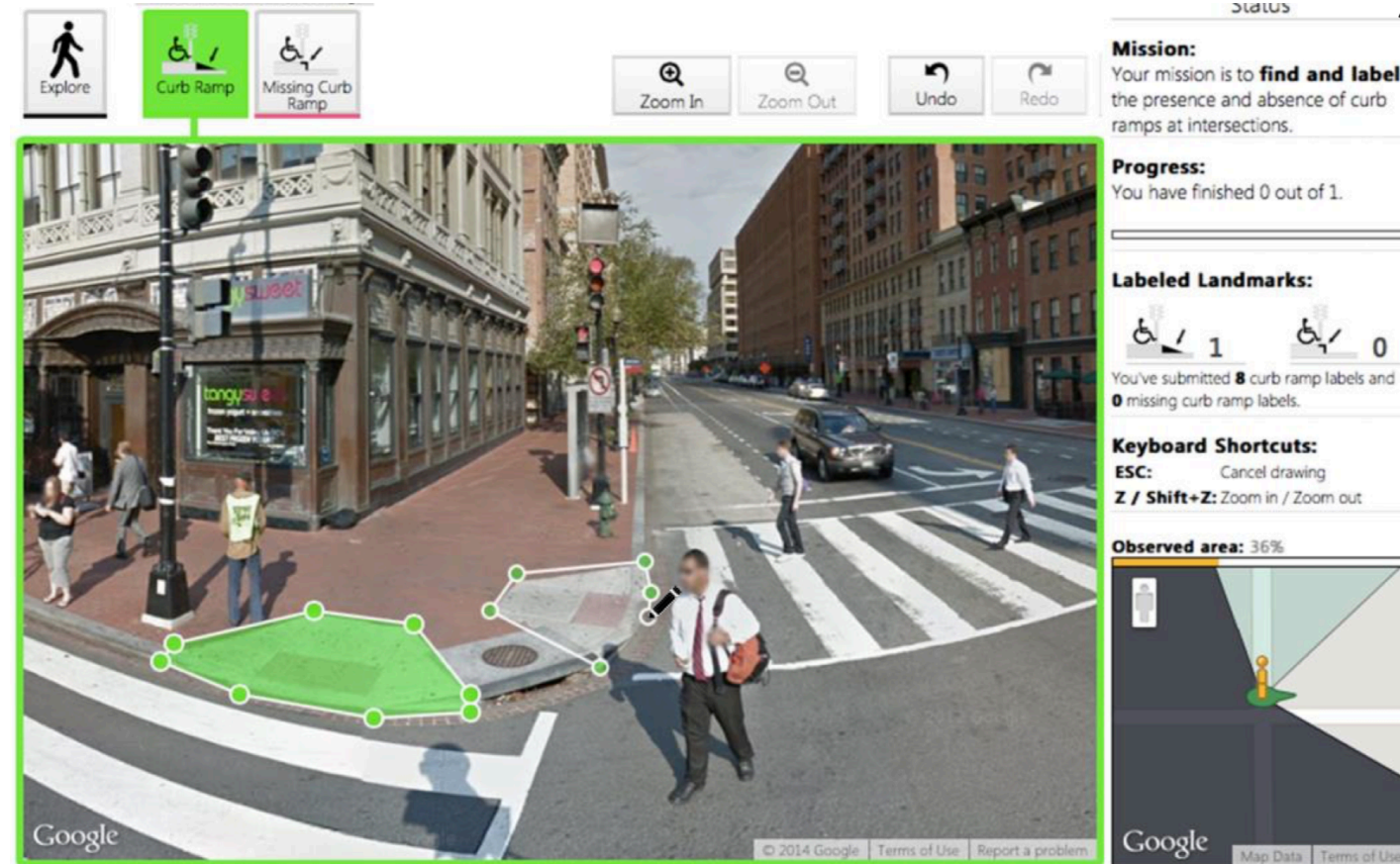


Expensive

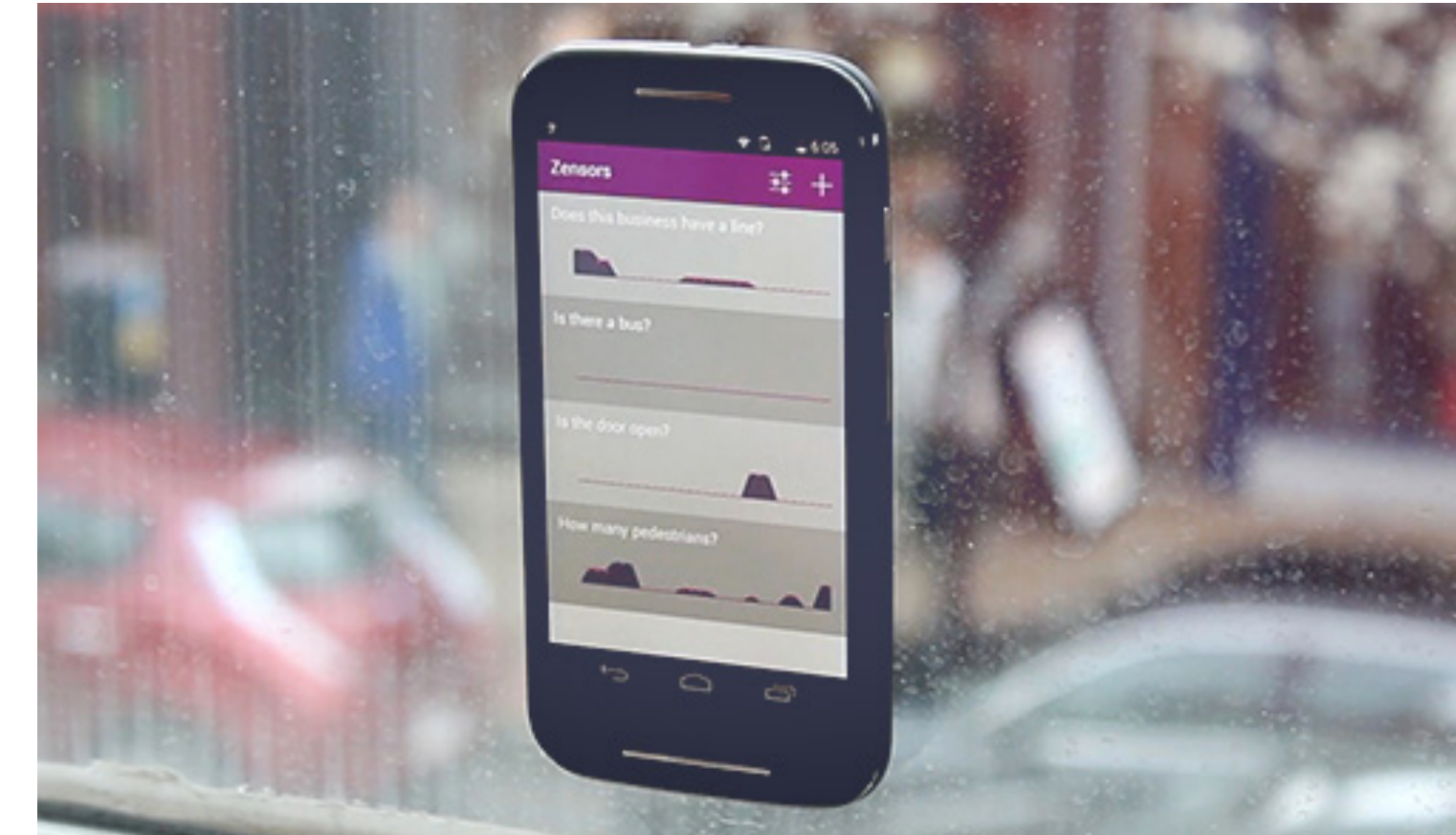




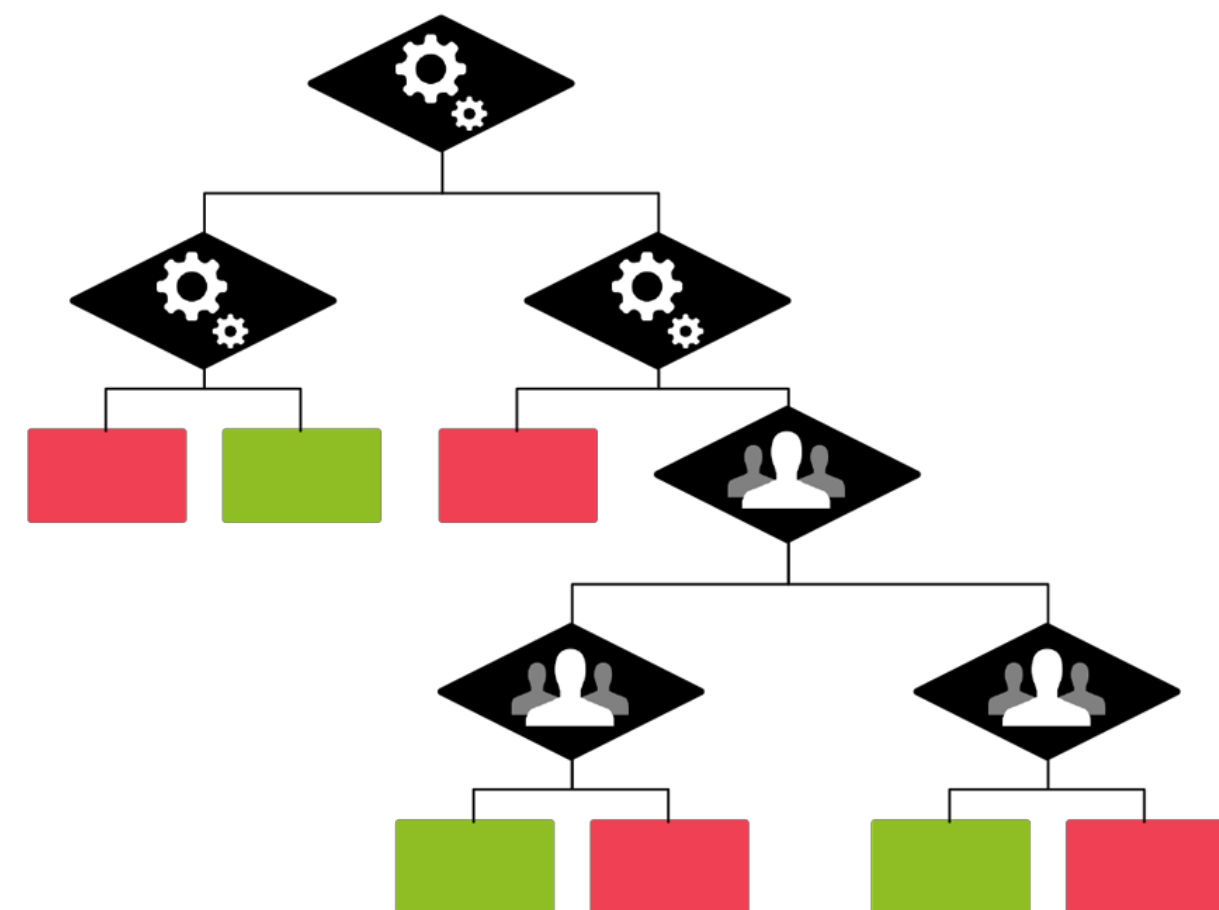
# Crowd-AI Systems



Tohme, Hara et al. 2014



Zensors, Laput et al. 2015



Flock, Cheng et al. 2015



VizLens, Guo et al. 2016

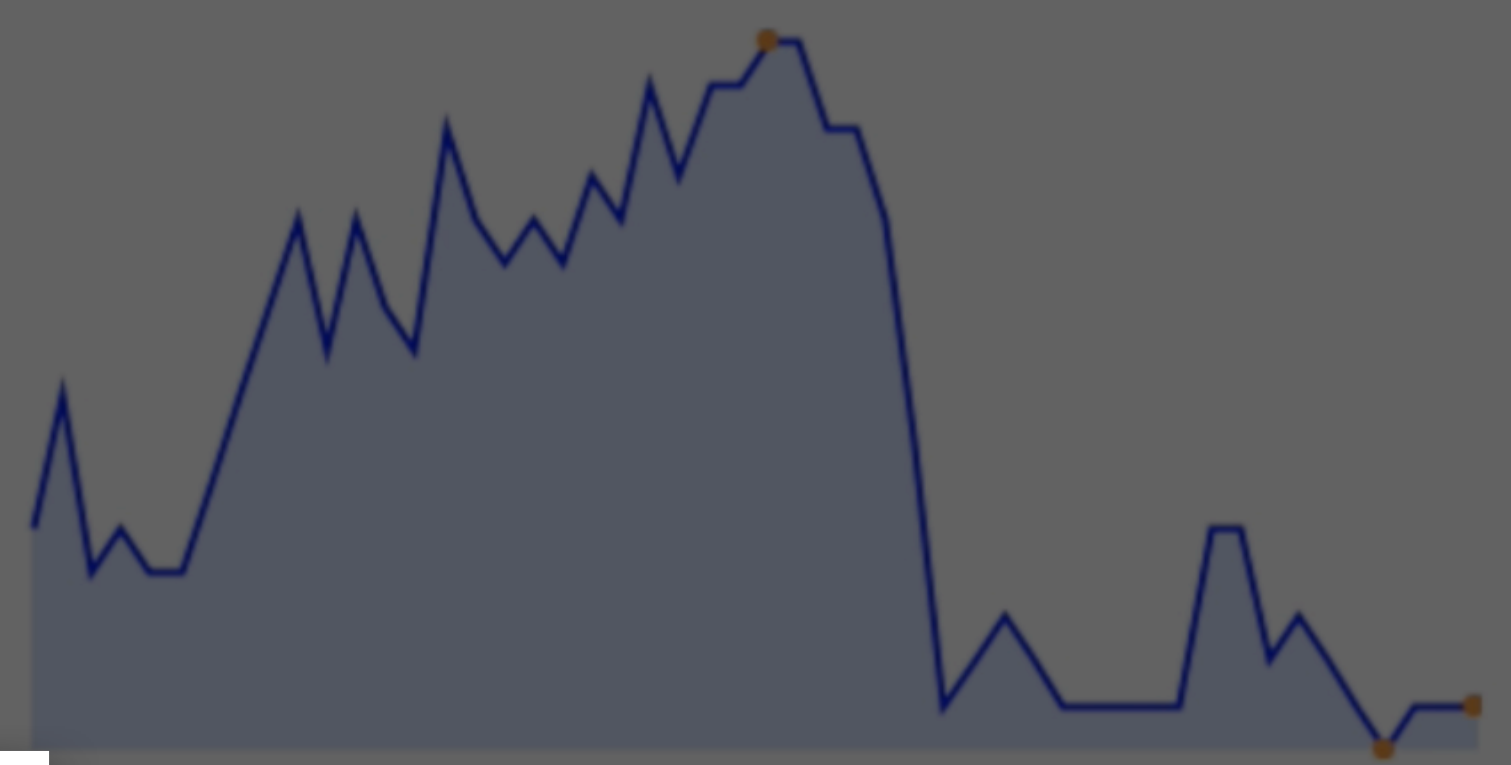


- Home
- Devices
- Sensors
- Notifications
- User Profile
- Logout

Welcome back, test

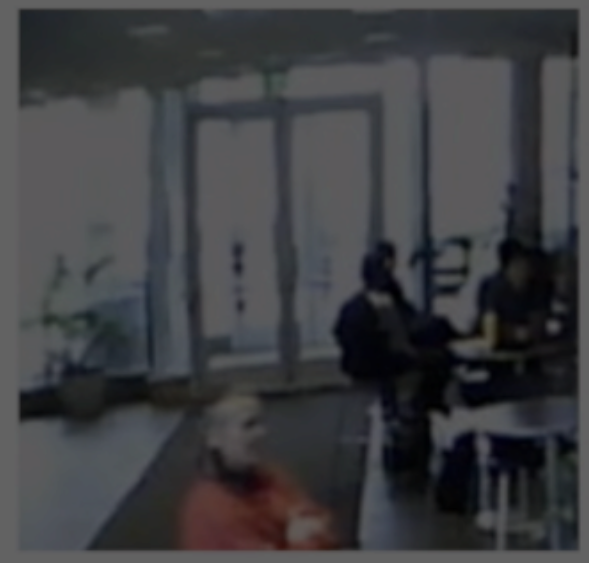
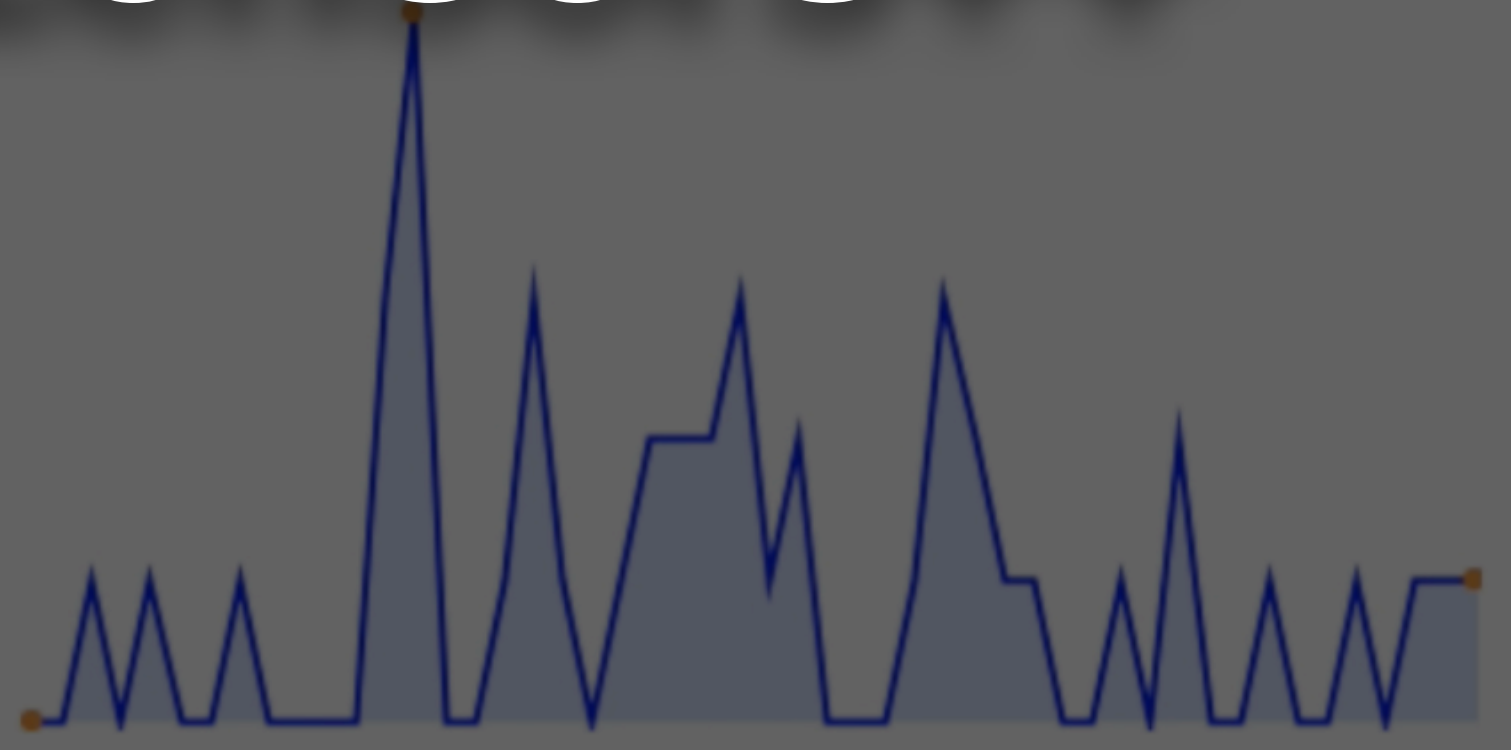
Current time is Feb. 6, 2018, 2:27 p.m.

Restaurant line\*

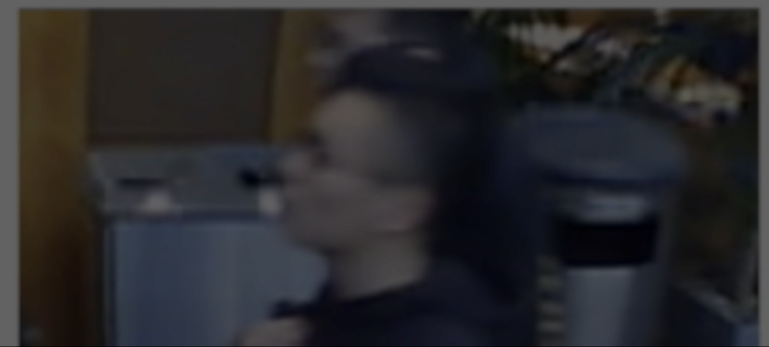
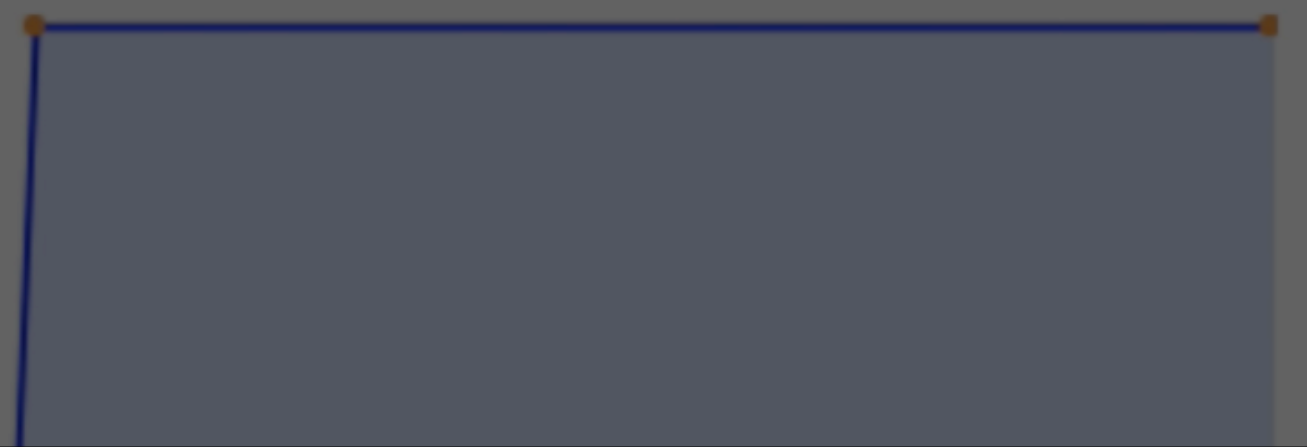


# Zensors++

Door Usage



Trash Can Usage\*



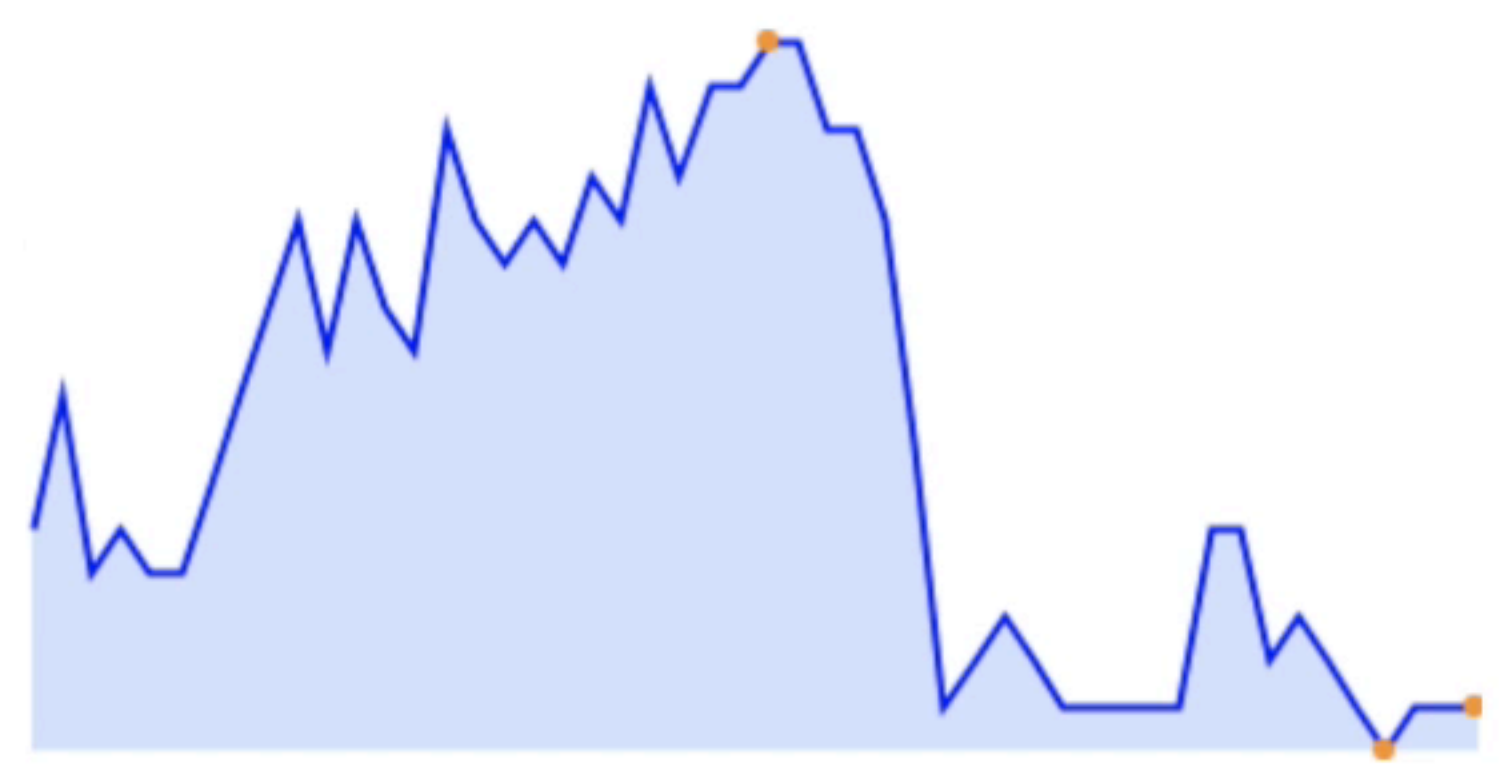


- Home
- Devices
- Sensors
- Notifications
- User Profile
- Logout

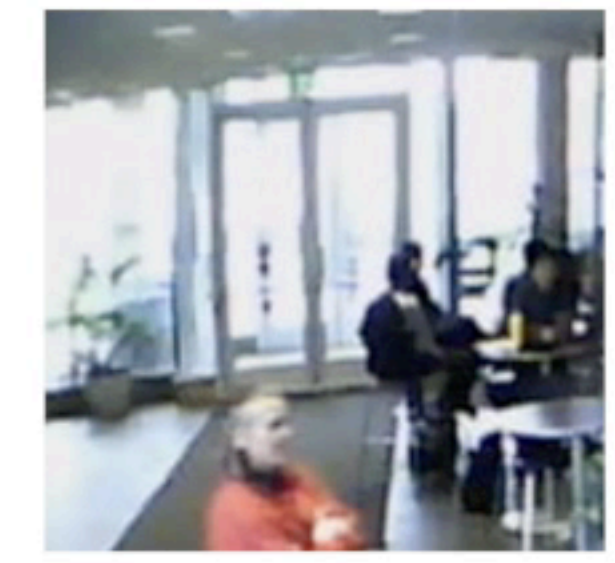
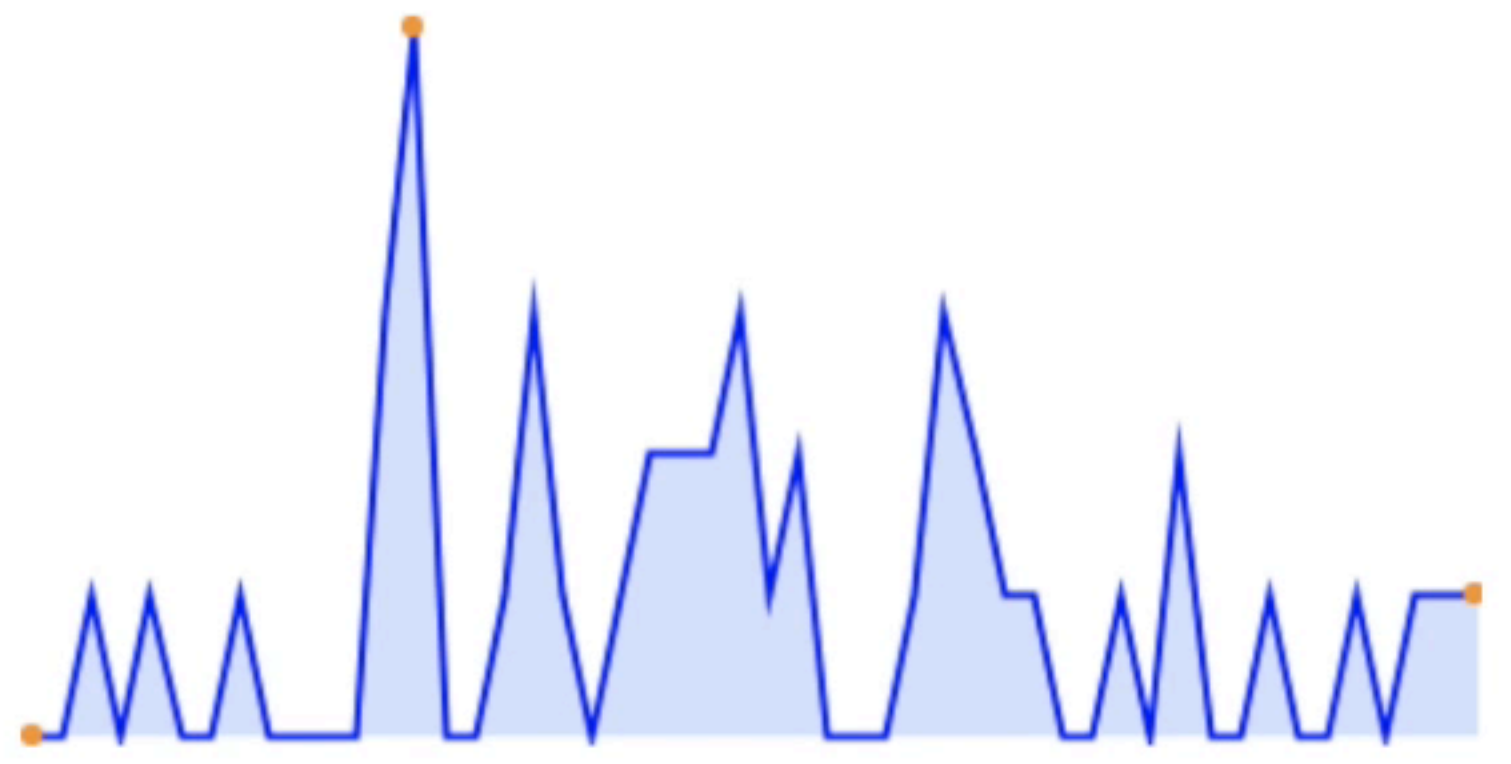
Welcome back, test

Current time is Feb. 6, 2018, 2:27 p.m.

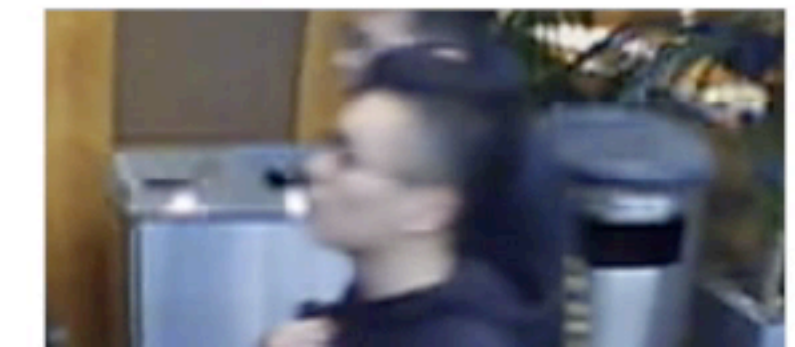
Restaurant line\*



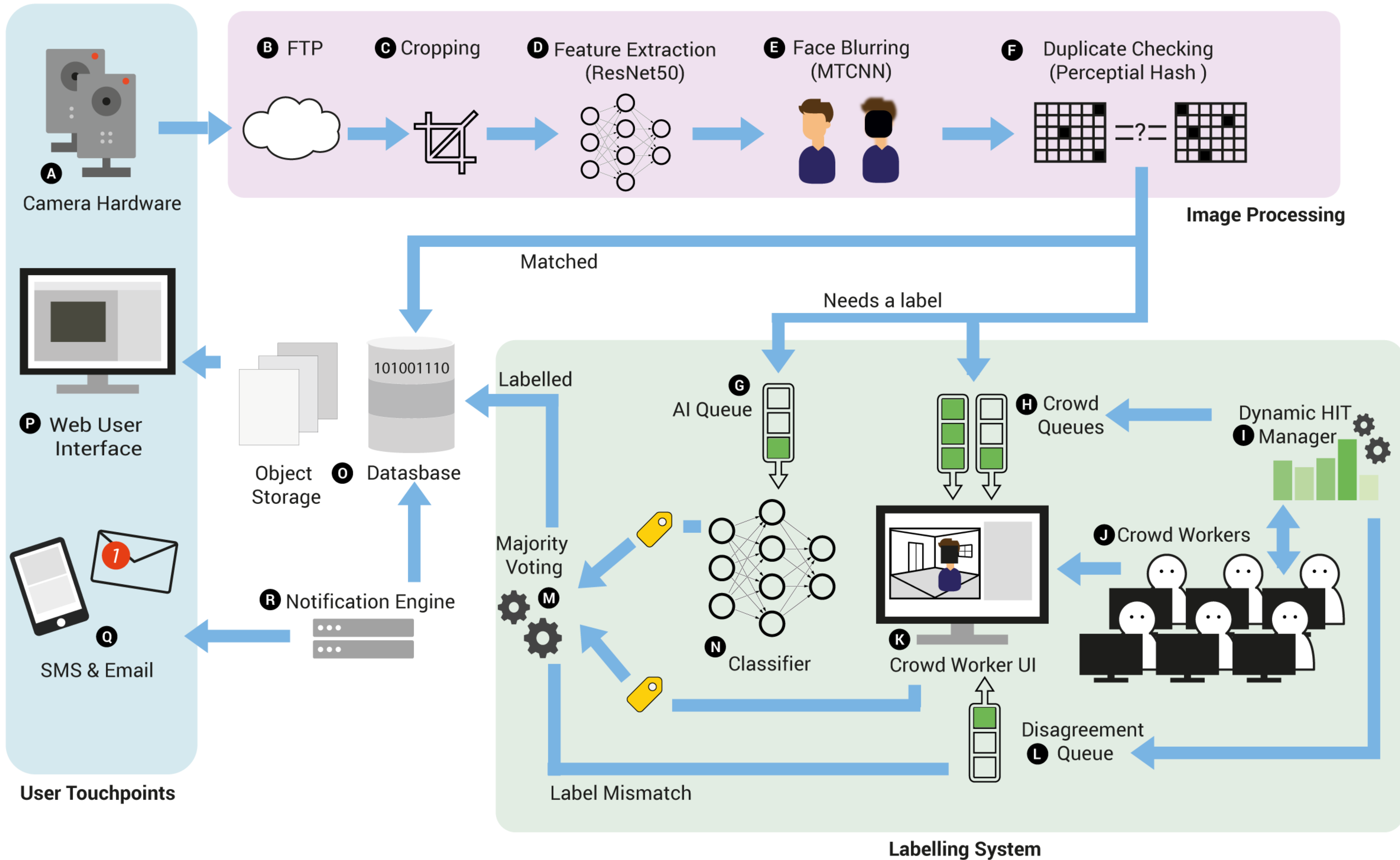
Door Usage



Trash Can Usage\*



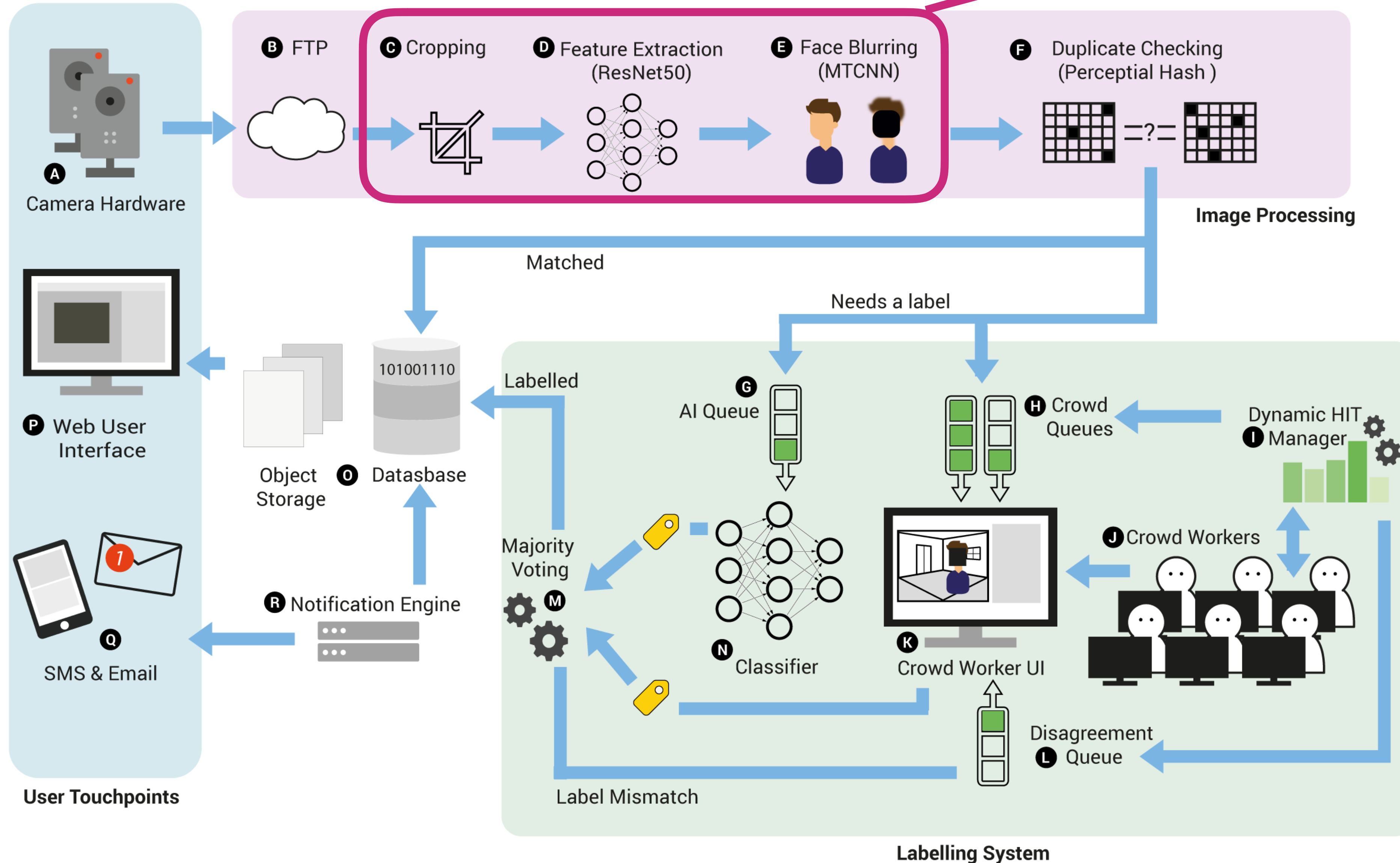






# Zensors++ Architecture

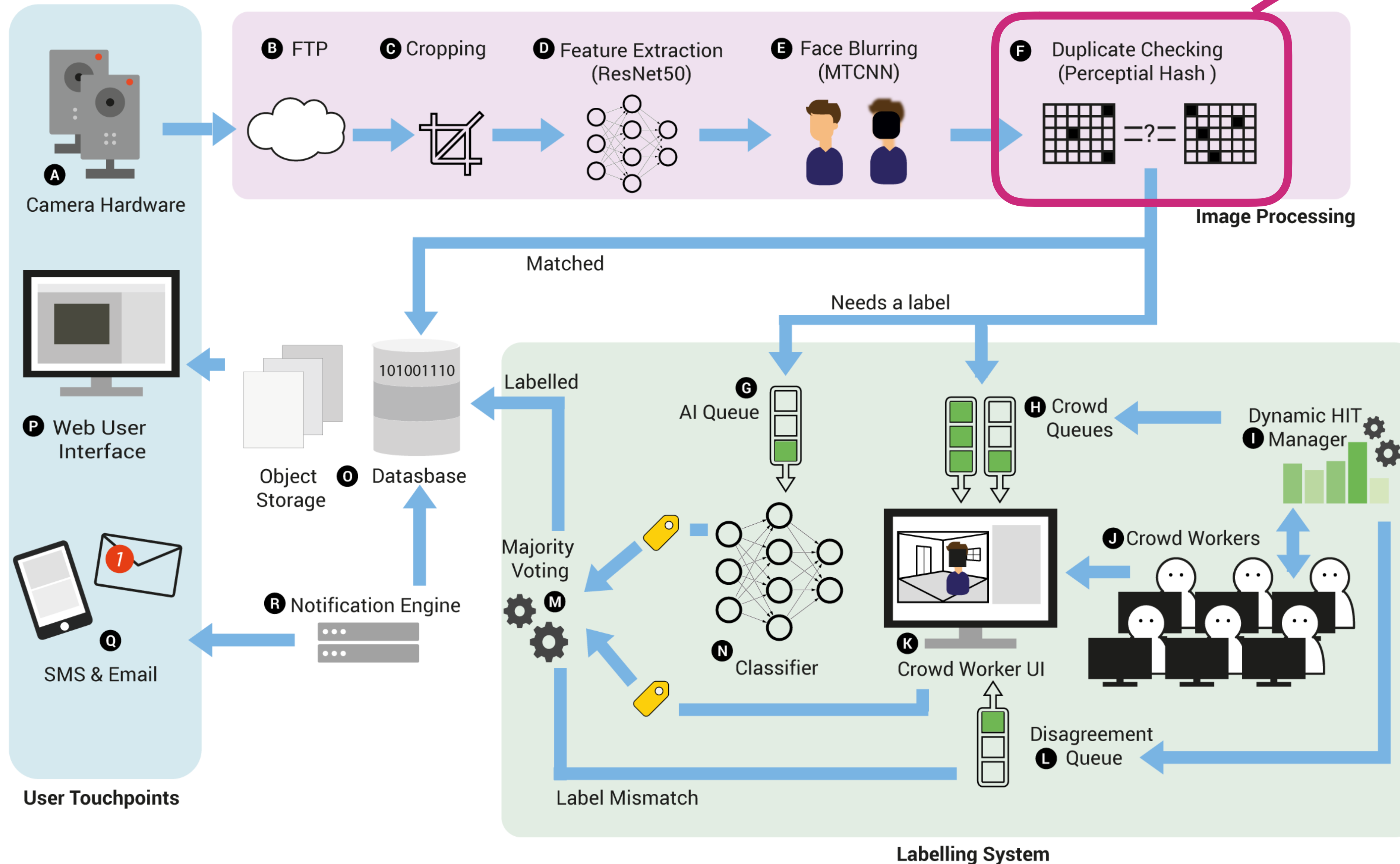
Privacy Preservation





# Zensors++ Architecture

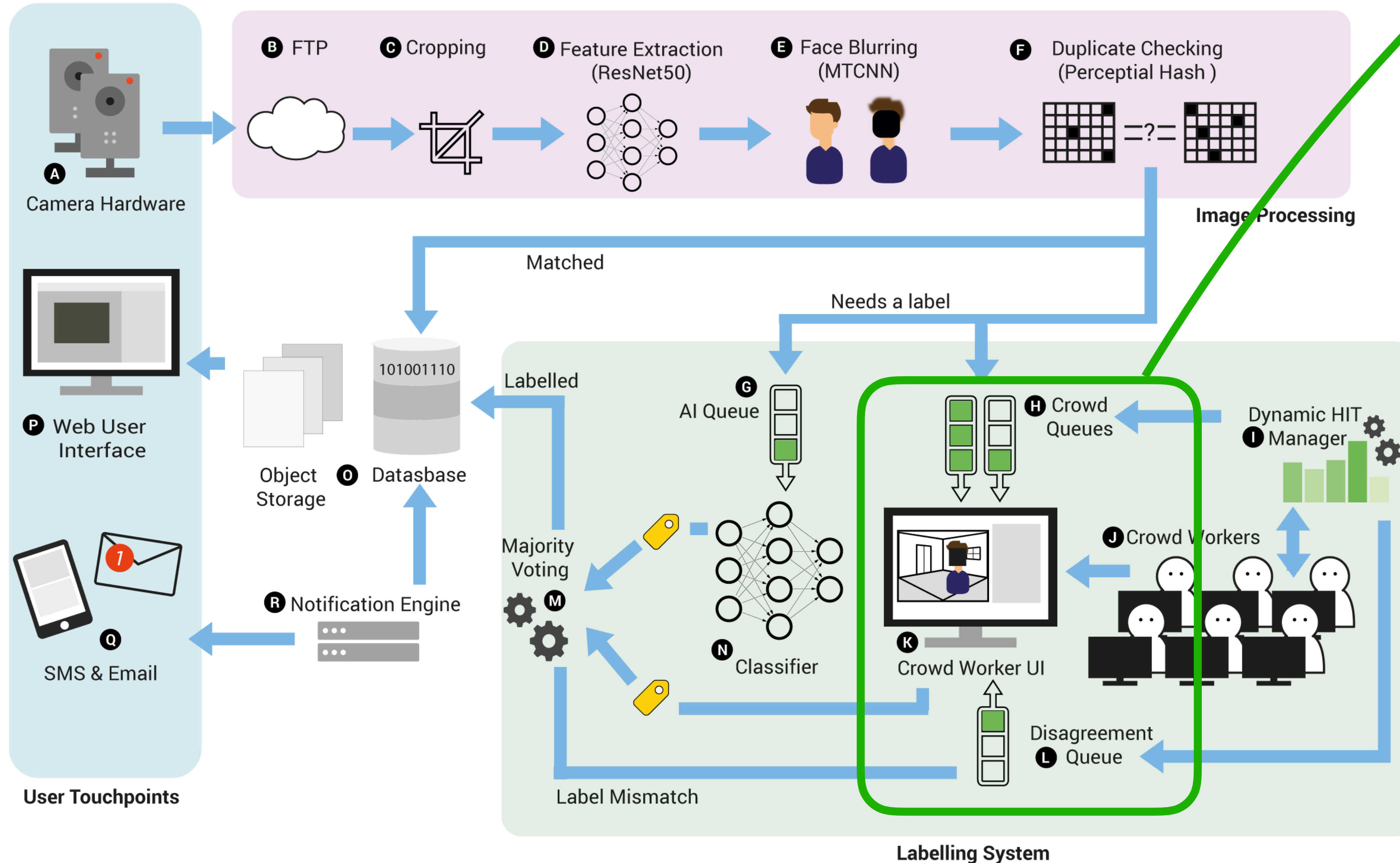
Reducing Redundancy





# Zensors++ Architecture

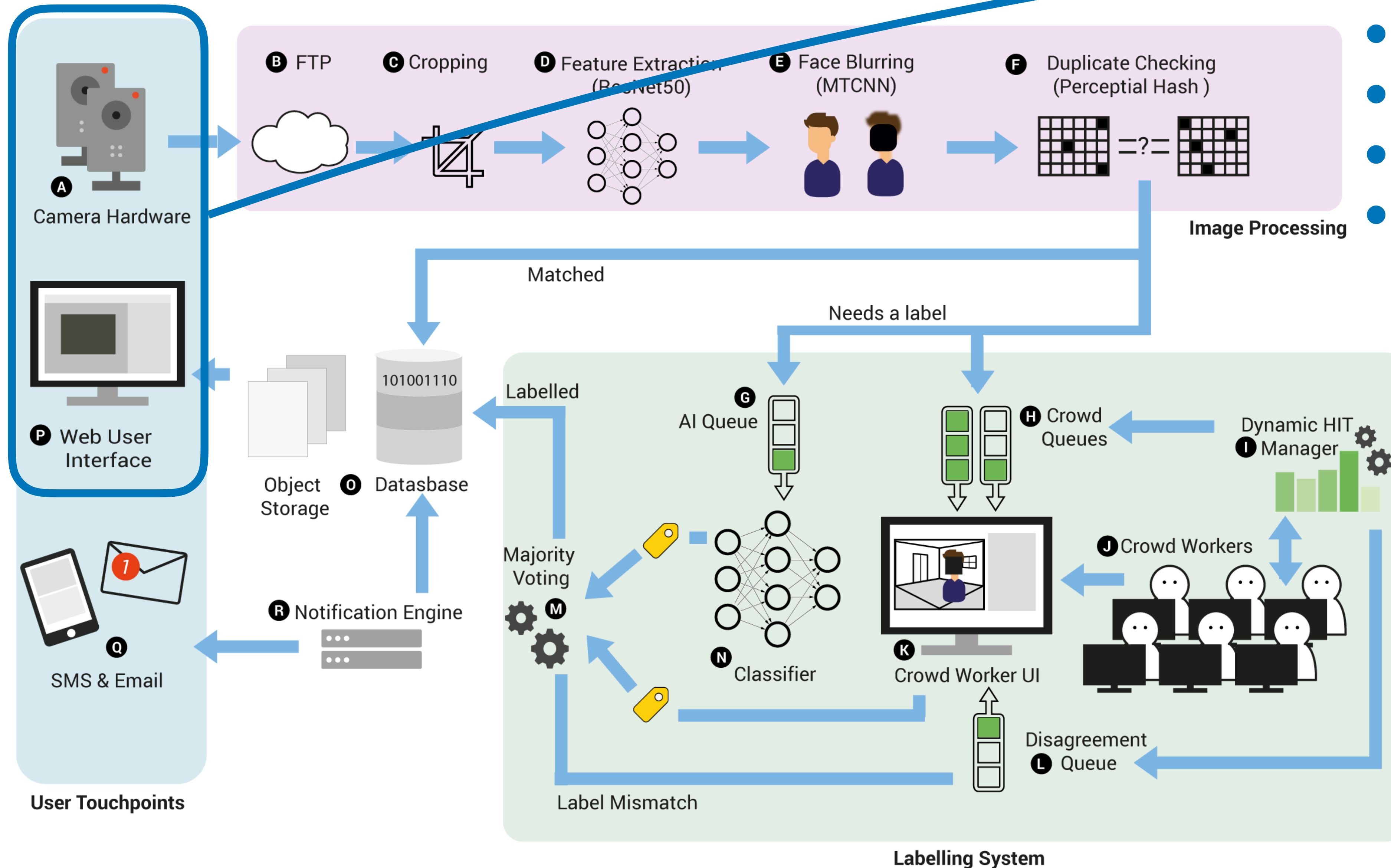
## Crowd Majority Voting





# Zensors++ Architecture

## Sensor Authoring Support



- Poor image cropping
- Ambiguous language
- Missing context
- Poor image quality



# Evaluation Deployment

- **Evaluation deployment:** 17 users, 4 weeks, 63 sensors, 937,228 answers
- **Occupations:** department and program directors, administrative coordinators, facility and lab managers, professors and students
- **Locations:** homes, offices, labs, cafes, food courts, parking lots, classrooms, workshops, and shared kitchens



A high-angle, slightly dimly lit photograph of a busy cafe or coffee shop. The scene is filled with people: some are seated at tables, some are at the bar counter, and others are standing. The cafe has a rustic feel with brick walls, a wooden bar, and various coffee-making equipment. Large windows on the right side offer a view of the street outside. The overall atmosphere is one of a bustling, social environment.

# Applications of Zensors++

How do end users apply crowd-AI camera sensing in their domestic and work lives?  
What are the perceived value?



**Do you see a motorcycle here?**



**Is the light on?**



**Are there people in the classroom?**



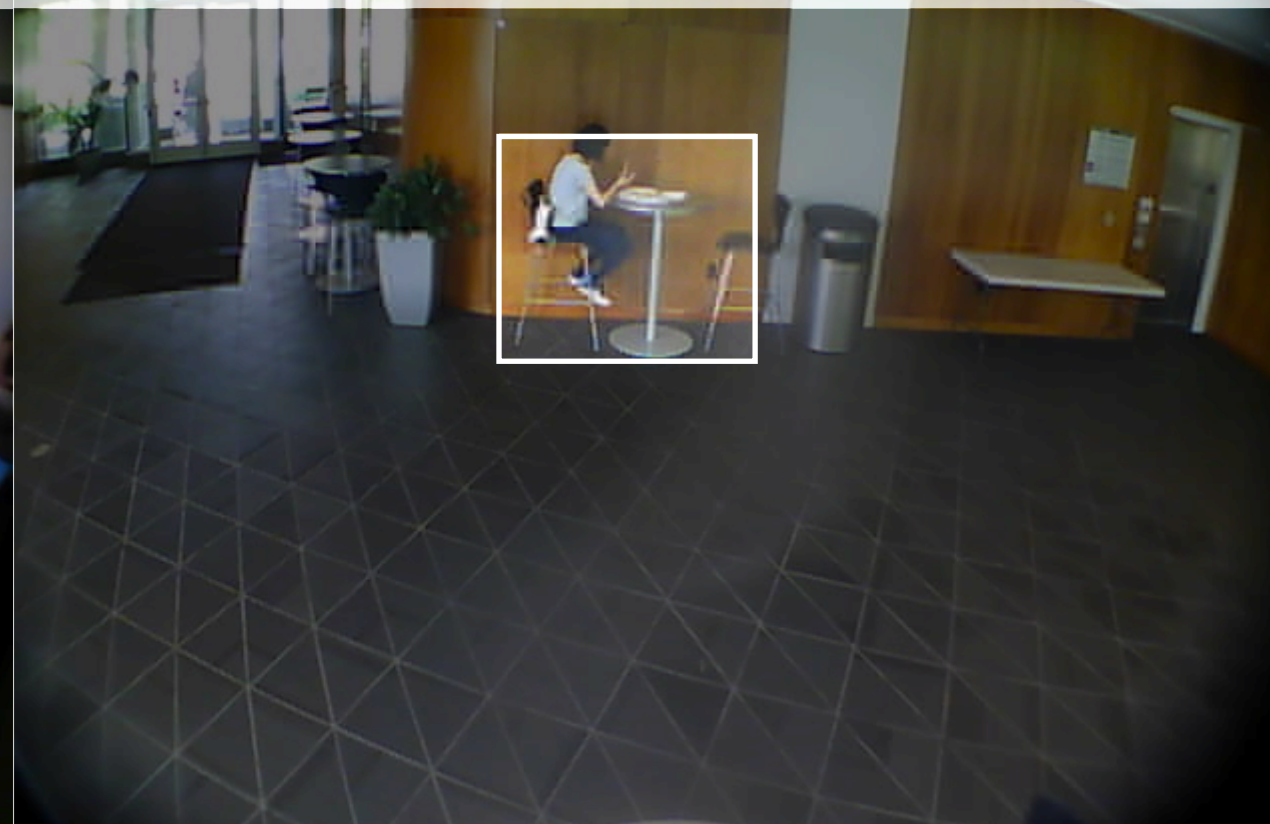
**Do you see any part of a car here?**



**Is a person entering this door?**



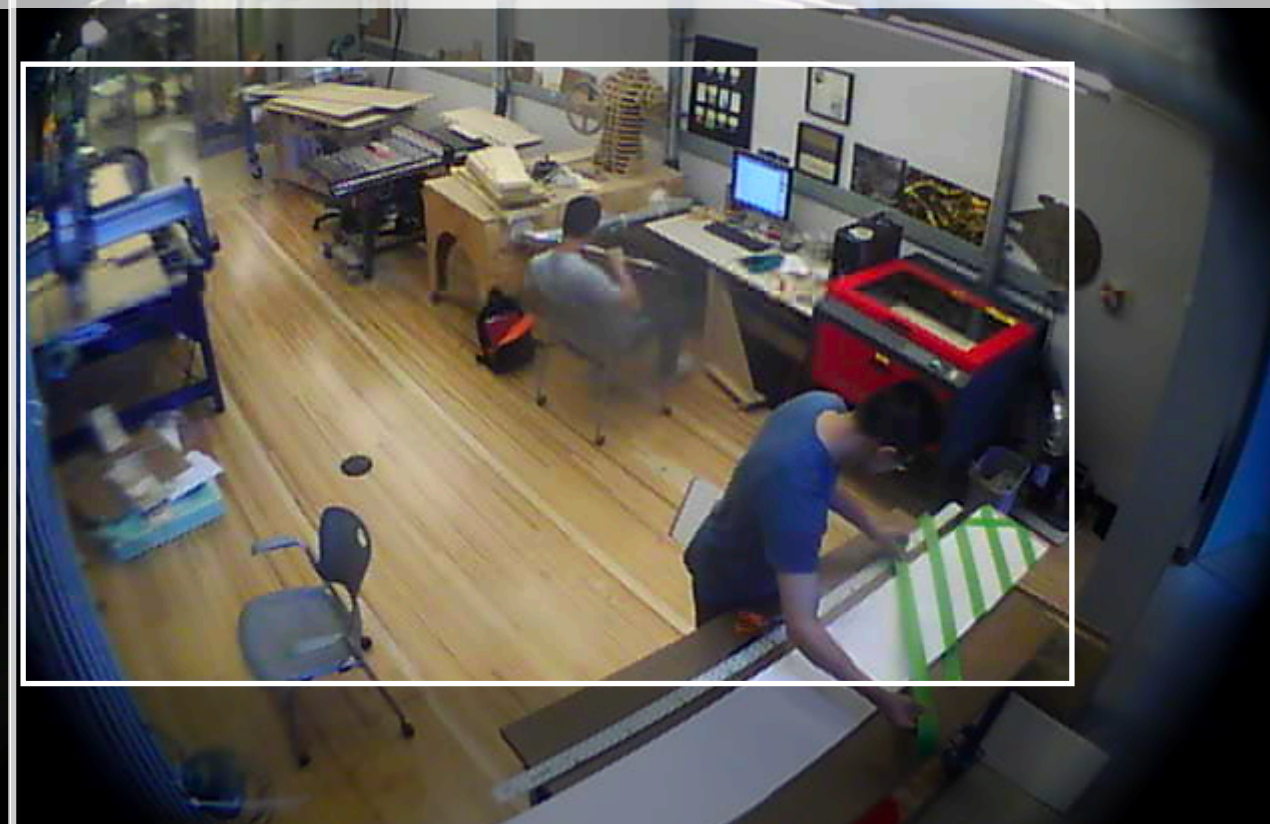
**Is the [...] table occupied?**



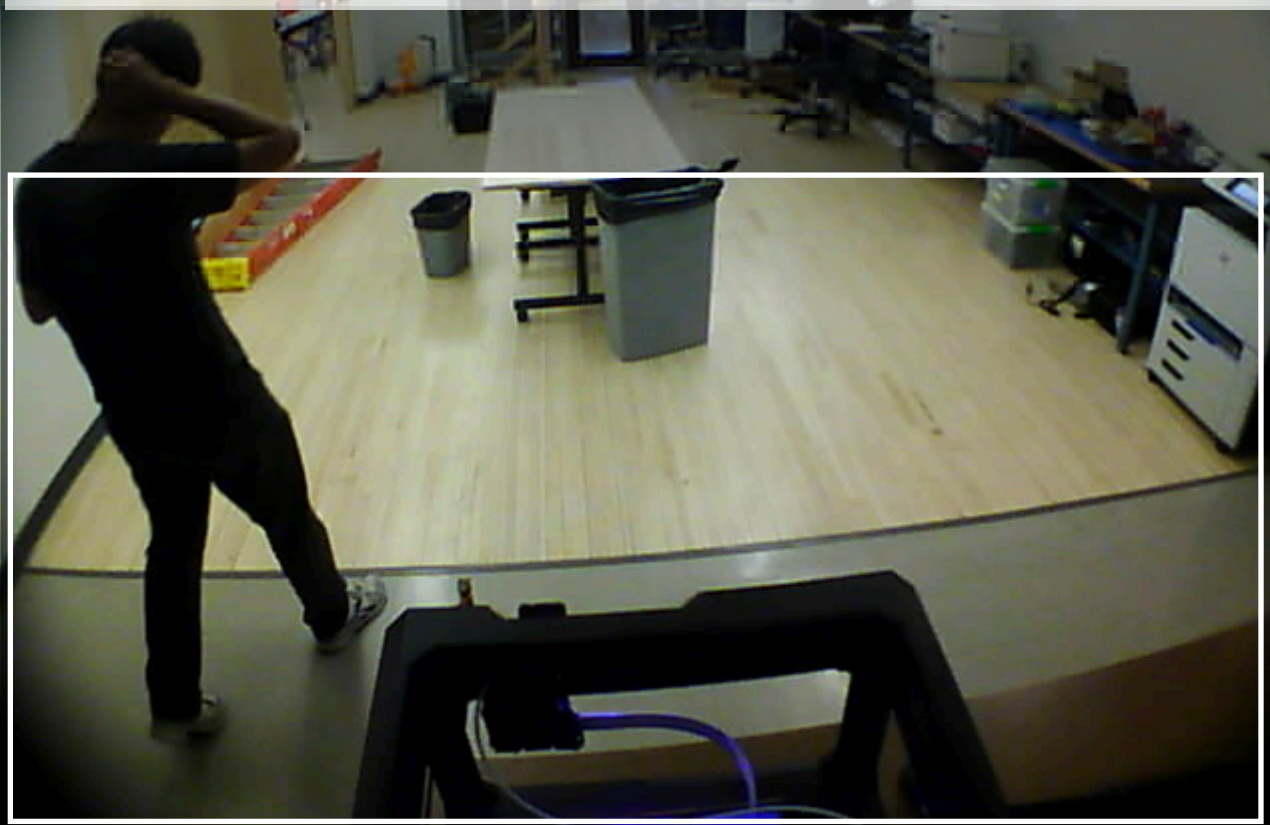
**Is there any paper or mail here?**



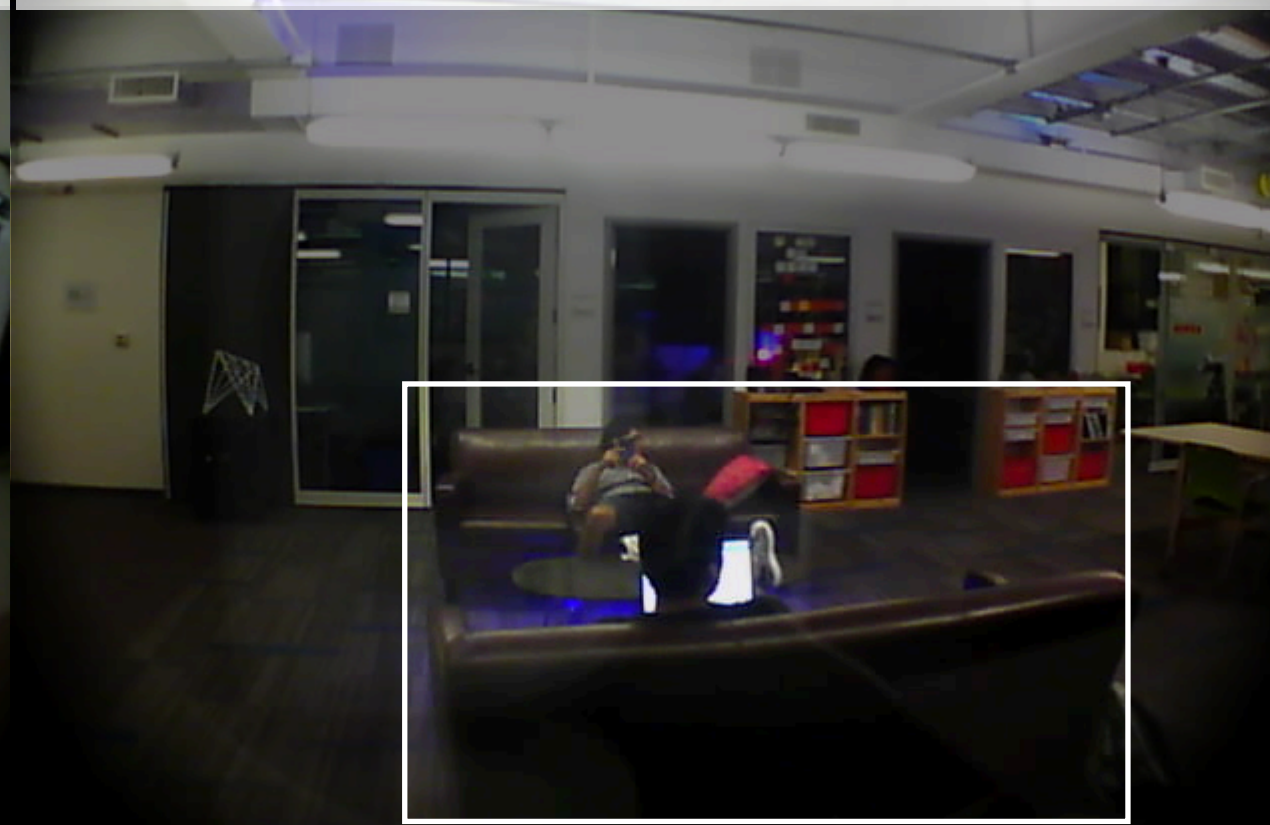
**Is anyone using the tools or equipment?**



**How many people are in the room?**



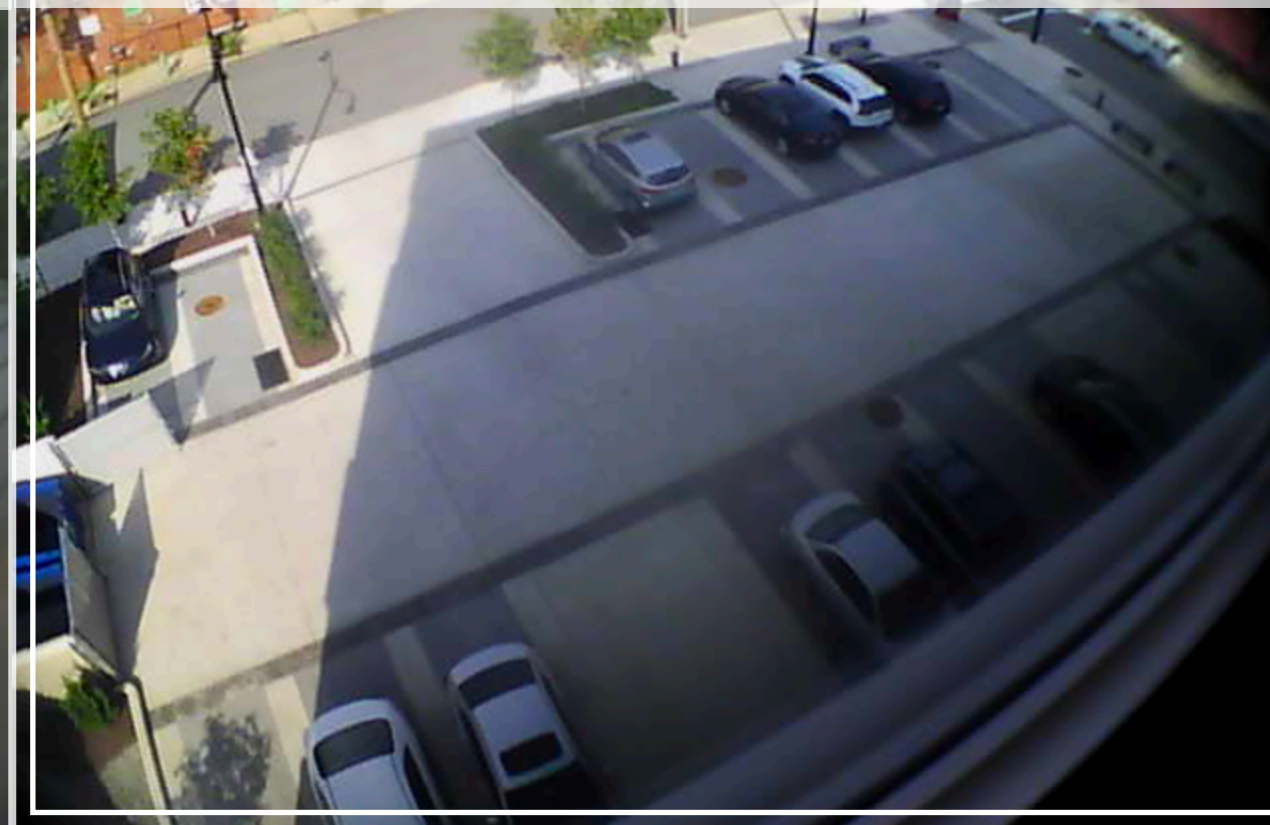
**How many people are [...] on two sofa?**



**How many people [...] on the benches?**



**How many cars are parked here?**





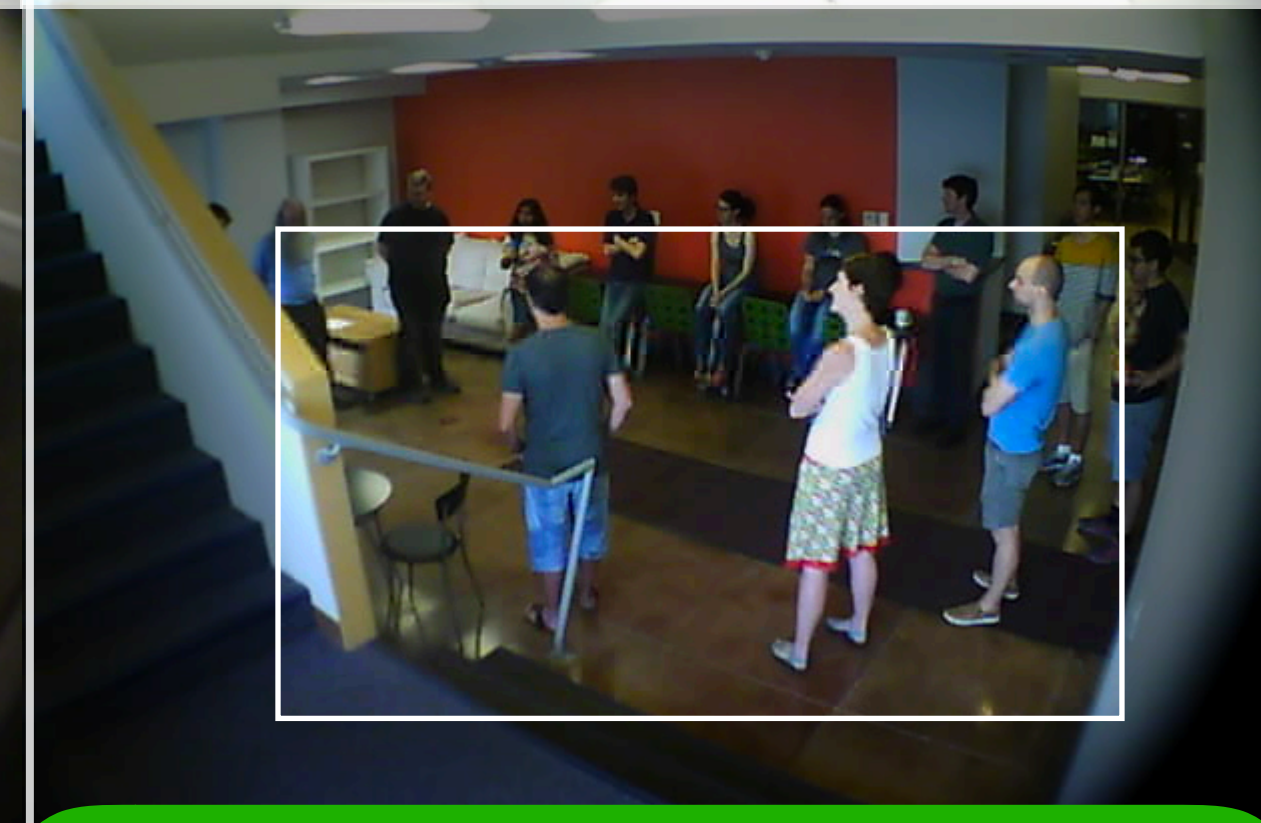
**Is the kitchen fridge open?**



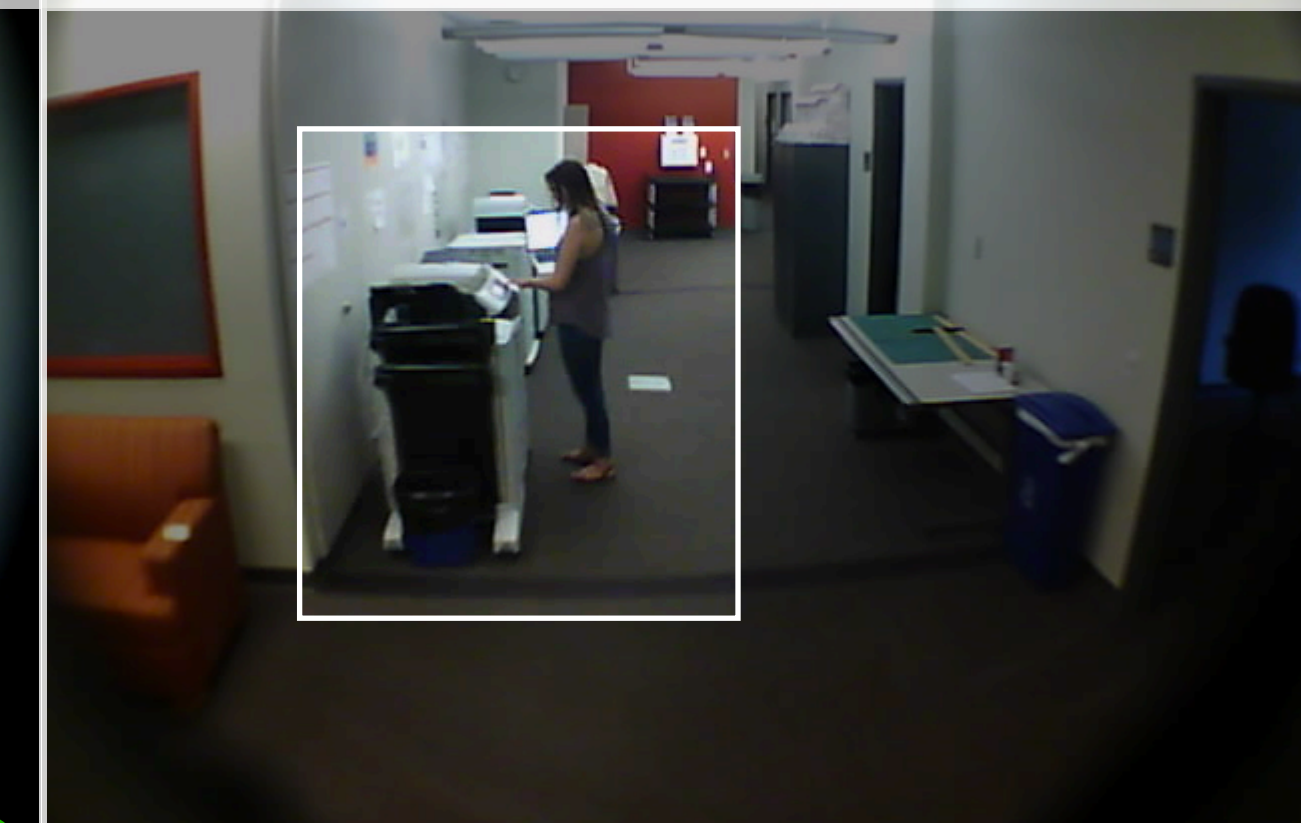
**Is the coffee machine in use?**



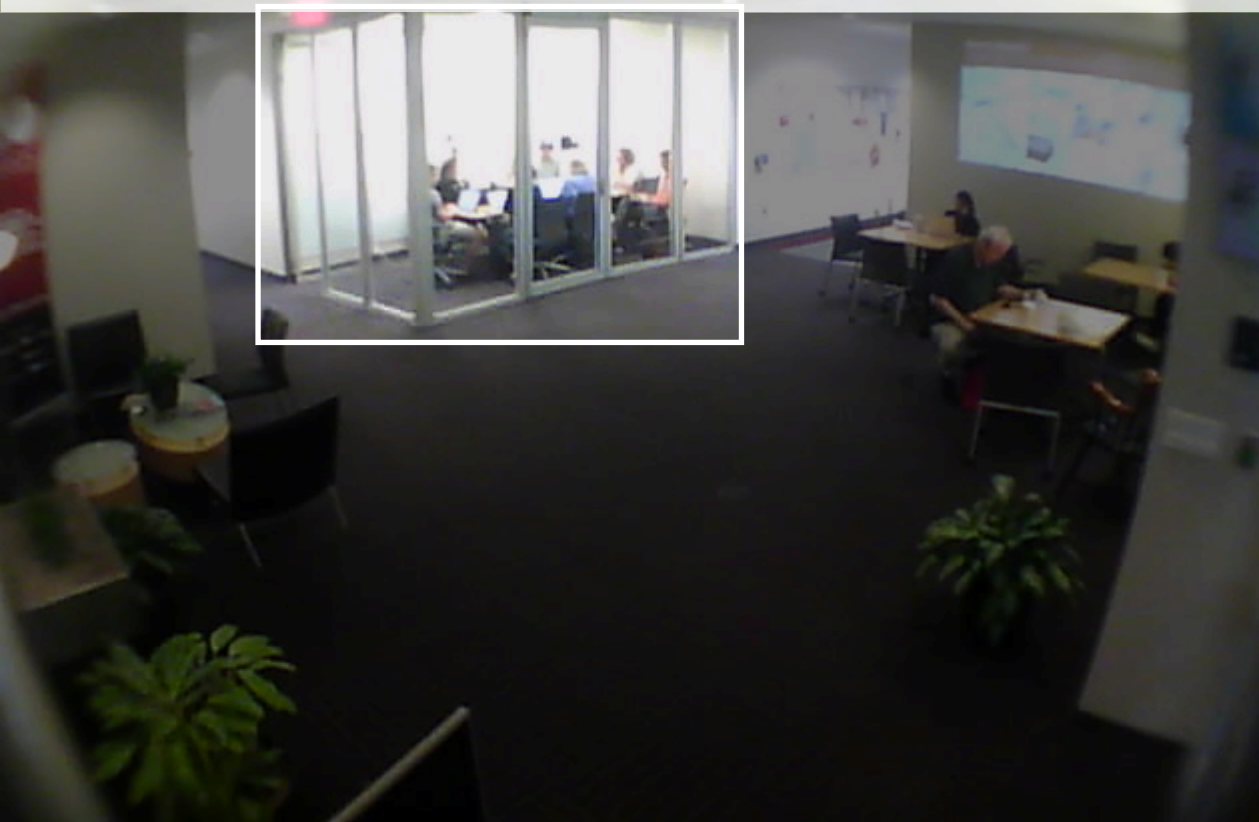
**Is there a gathering of people [...]?**



**Is someone using a printer?**



**Is there anyone in the glassed-in room?**



**Is anything written on the whiteboard?**



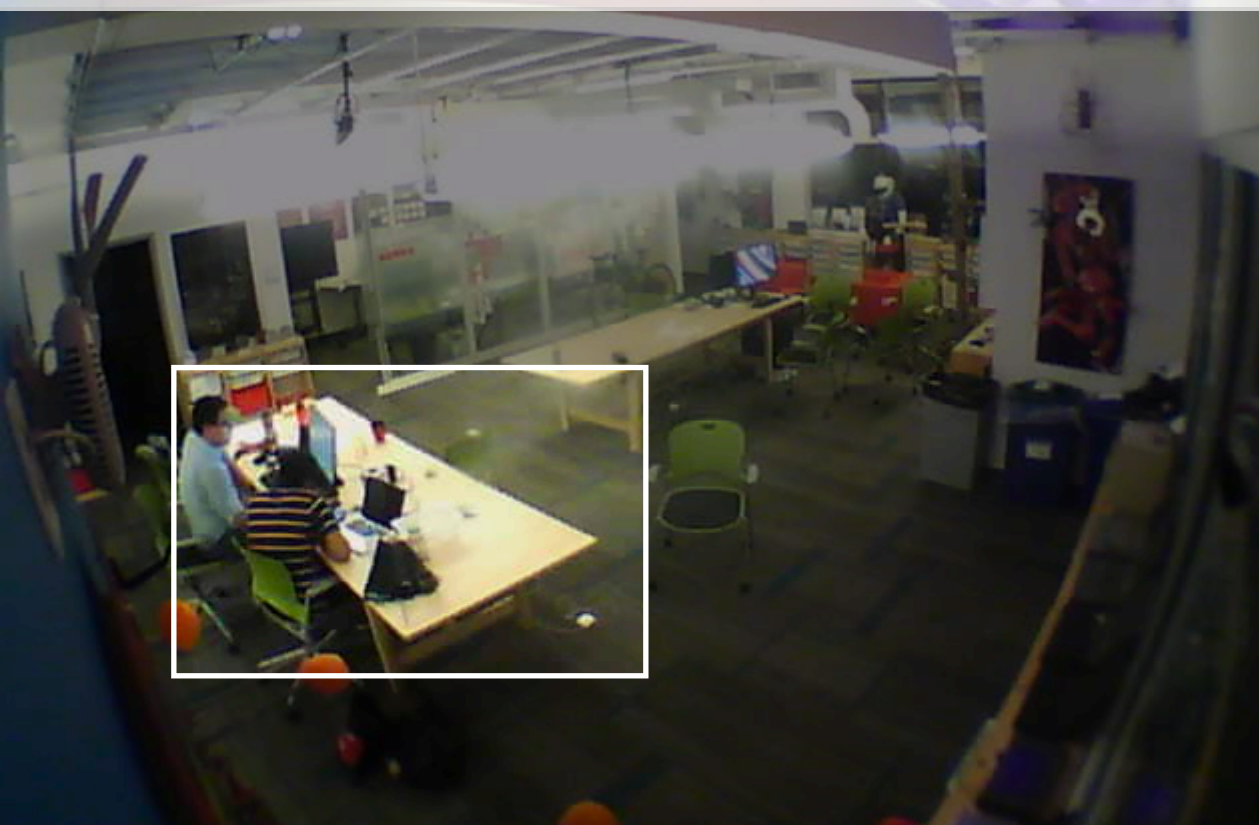
**Is someone sitting on this furniture?**



**Is the trashcan full [...]?**



**How many people are sitting [...]?**



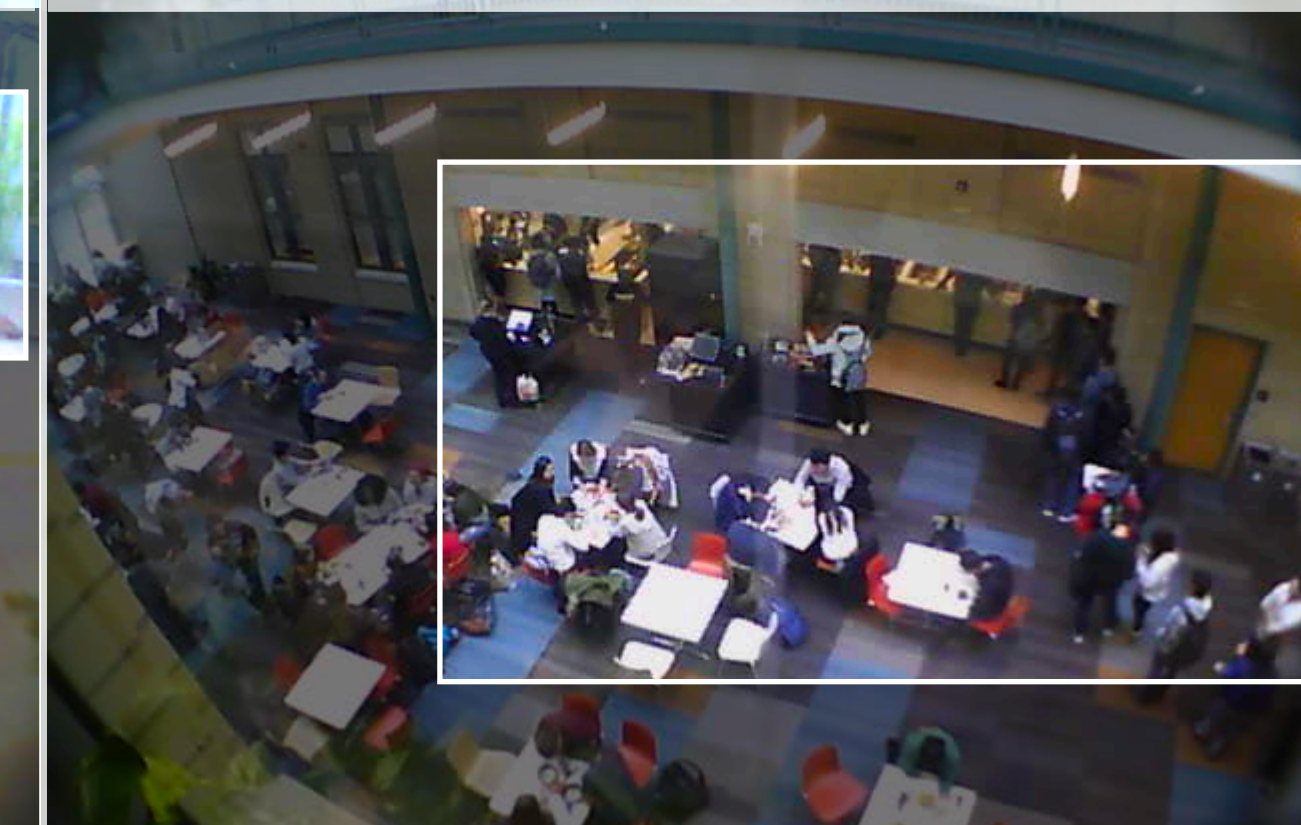
**How many cars do you see?**



**How many pedestrians do you see?**



**How many people are in the line [...]?**

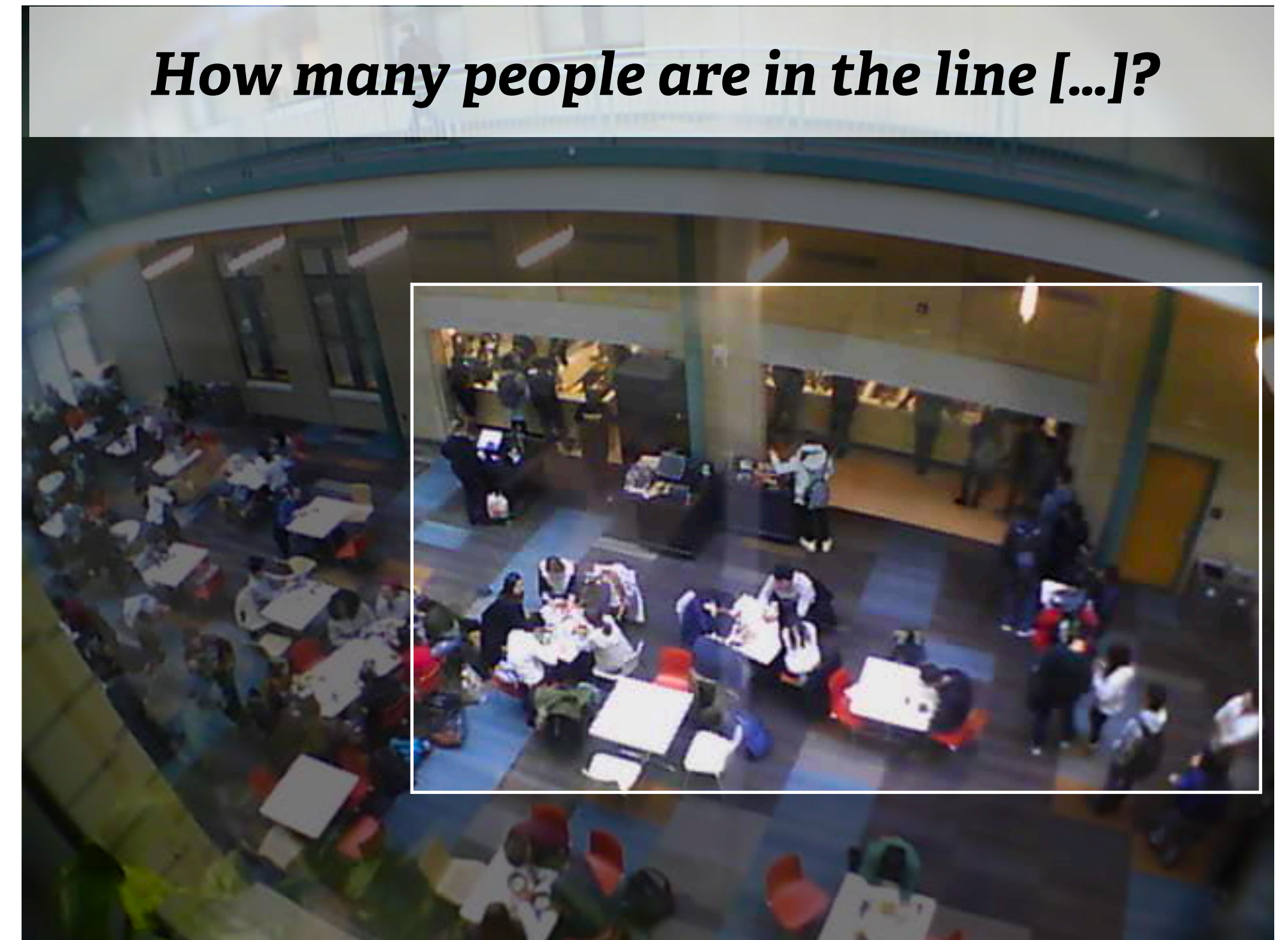




# Use of Data

*How do end users apply crowd-AI camera sensing in their domestic and work lives?*

- **Users in commercial settings more interested in longitudinal data**
- **Users in personal capacity cared more about in-the-moment state**





# Proxy Questions

*How do end users apply crowd-AI camera sensing in their domestic and work lives?*

- **Alternative framing that was context free to answer a different question**
- **Future work to support users define and formulate questions**



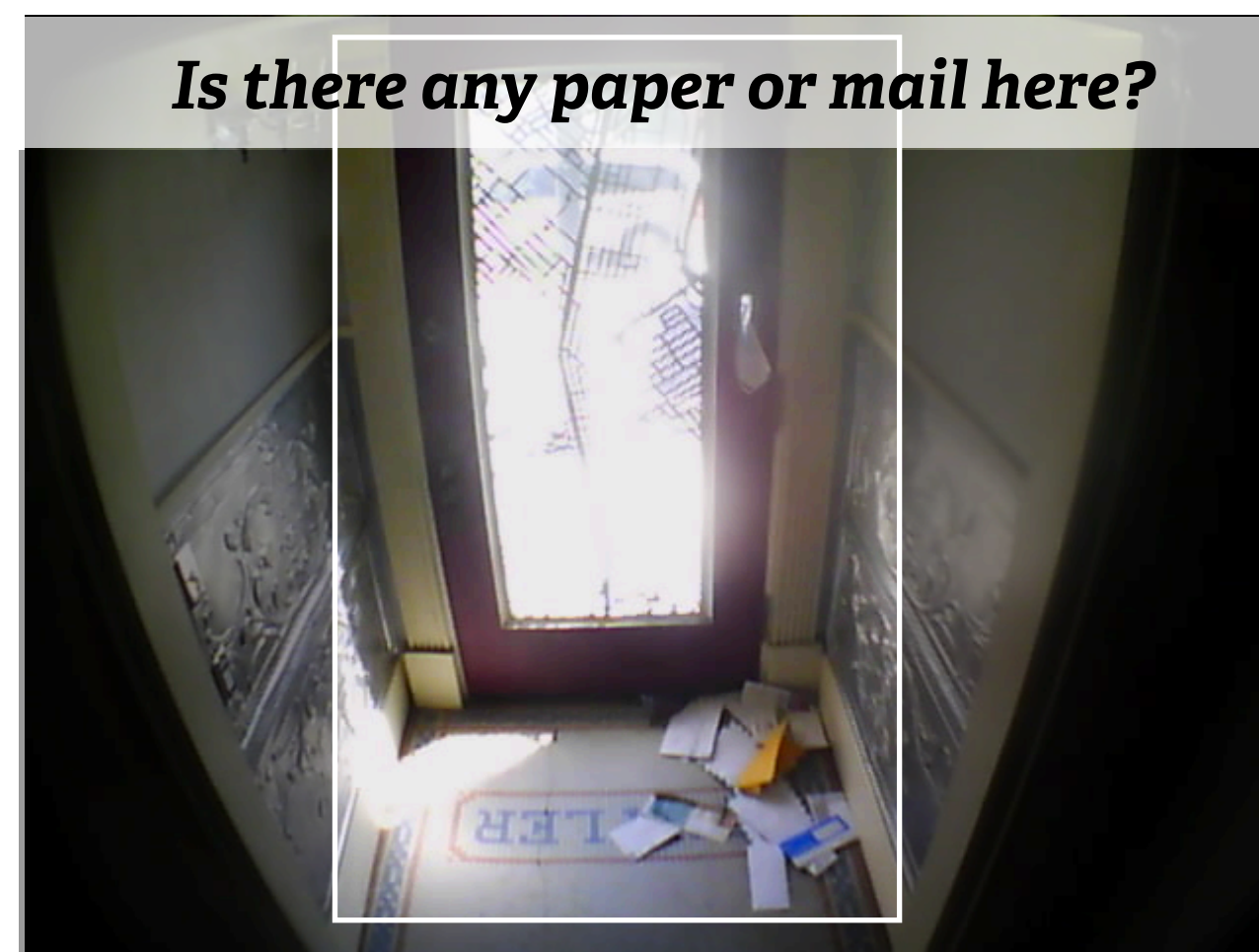


# Perceived Value

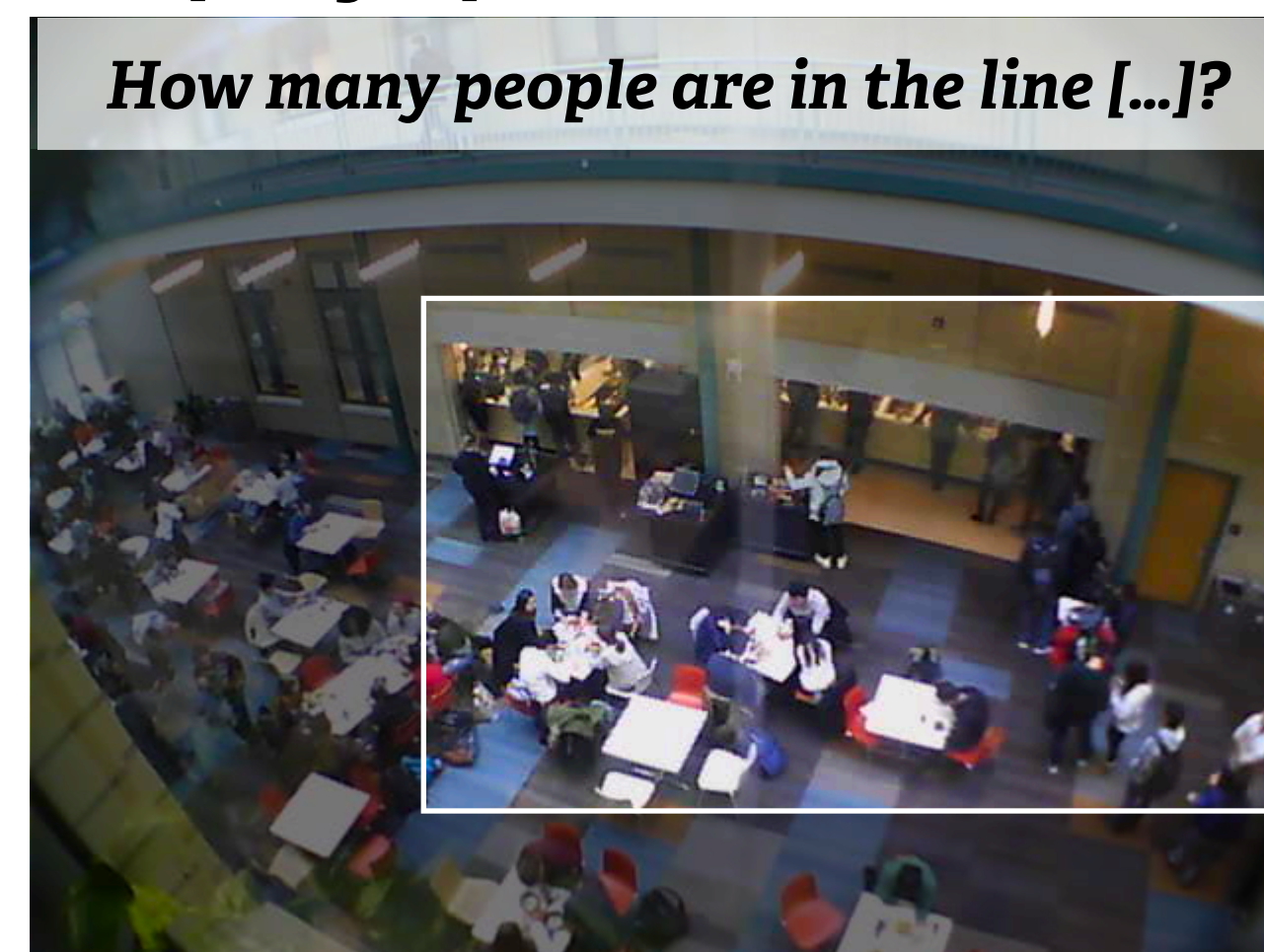
*What are the perceived value?*

- Participants in professional roles are willing to pay large amounts for questions that directly complemented or augmented existing practices

**Personal use: \$1-10**



**No pay, prefer business**



**Thousands annually**







# Performance of Zensors++

What is the accuracy, latency, cost, and automation that can be achieved in real world deployments?



# Accuracy

*What is the accuracy, latency, cost, and automation that can be achieved in real world deployments?*

- **~80% for yes/no questions, 0.2 unit error for count questions**
- **Sources of errors:**
  - Malicious crowd behavior
  - Ill-defined user questions



# Latency

*What is the accuracy, latency, cost, and automation that can be achieved in real world deployments?*

- **Worker labeling duration:** 5.8s for yes/no, 6.6s for count
- **User receiving answer duration (crowd):**
  - First Answer:** 2 mins
  - Majority Vote:** 5 mins
- **User receiving answer duration (hashing):**
  - zero-latency



# Hashing

*What is the accuracy, latency, cost, and automation that can be achieved in real world deployments?*

- **Hashing rate: 74.4% (697,345) out of 937,228 answers**  
Near-zero latency and cost, saving us approximately \$17,500
- **Hashing accuracy: 99%**  
Error cause: selected region was large,  
question asking about small changes



# Cost

*What is the accuracy, latency, cost, and automation that can be achieved in real world deployments?*

- **Average of \$0.006 per answer**

\$6,069 for crowd labels for 937,228 answers

- **Average of 2.5 labels per answer**

Saved \$1,127 in crowd cost, demonstrating effectiveness in voting scheme

- **Average per-day cost: \$2.4 for yes/no, \$4.5 for count**

60% hashed for yes/no, 45% hashed for count

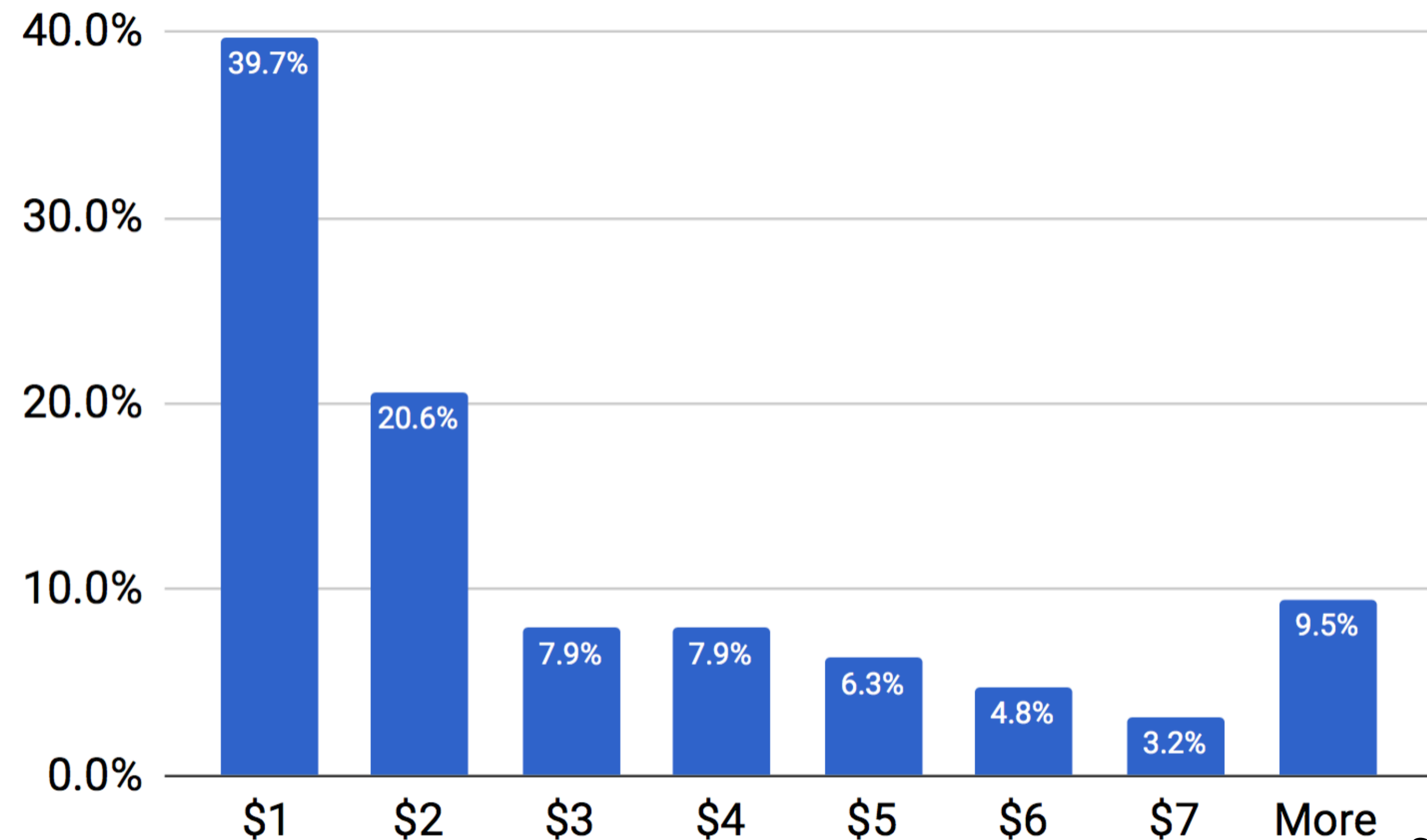
Might be that count questions inherently more dynamic and complex



# Cost

*What is the accuracy, latency, cost, and automation that can be achieved in real world deployments?*

- **60% cost < \$2 per day**
- **Cost can be further reduced**
- **Suggest long term viability of many use cases**





# Future Work

- **Applying private crowds**

Higher quality answers, less need for majority voting

Long-term crowd engage with users to refine questions, curate automation

- **Zensors++ dataset**

Sequential examples, as apposed to one-off questions, *e.g.*, VQA dataset

Transfer learning across environments, similar question types

- **Deeper exploration of sharing (devices, sensors, data)**

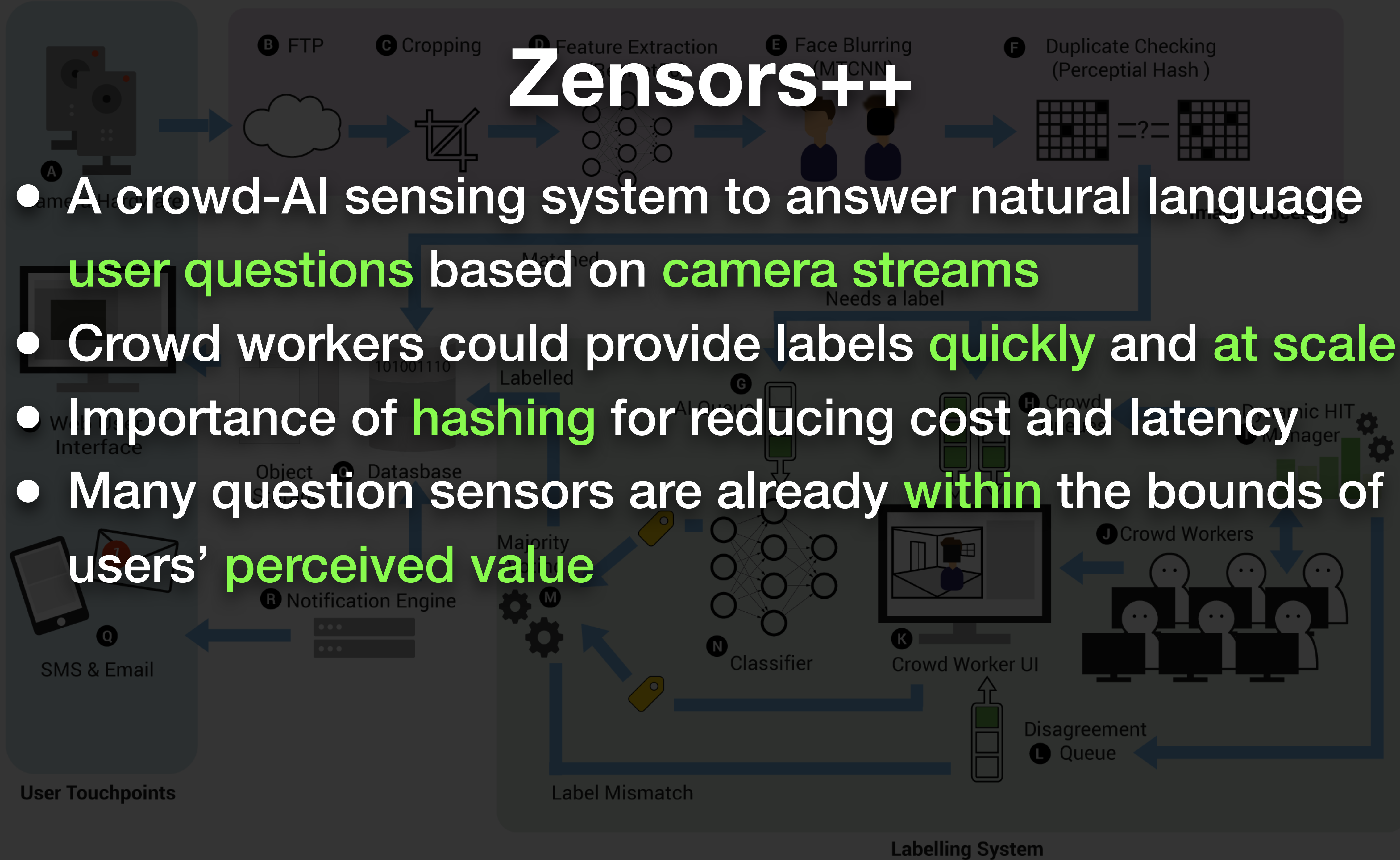
Cameras aiming outdoors can be utilized by third-parties

The ubiquity of cameras offers unique opportunities



# Zensors++

- A crowd-AI sensing system to answer natural language user questions based on camera streams
- Crowd workers could provide labels quickly and at scale
- Importance of hashing for reducing cost and latency
- Many question sensors are already within the bounds of users' perceived value







ZENSORS

## Penn Parking (South)

How many cars are parked here? ● **RUNNING** on camera Penn & Gross (West)



Dashboard



Questions



Create Question



Cameras



Register Camera

Accounts

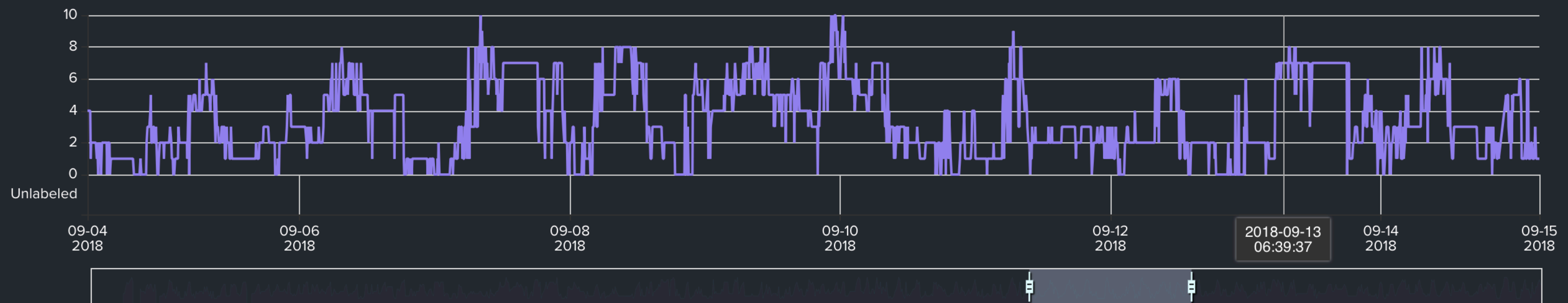
Log Out

### Actions



7

at 2018-09-13 06:39:37





*Is the coffee machine in use?*



*Is the trashcan full [...]?*



*Is someone using a printer?*



*How many people are in the line [...]?*



# Crowd-AI Camera Sensing in the Real World

*Anhong Guo, Anuraag Jain, Shomiron Ghose, Gierad Laput, Chris Harrison, Jeffrey Bigham*

*anhongg@cs.cmu.edu guoanhong.com*

*Sponsored by:*

