

# Why Edge Computing is a Disruptive Technology

**Mahadev Satyanarayanan**  
**School of Computer Science**  
**Carnegie Mellon University**

# Value Proposition of Edge Computing

## 1. Highly responsive cloud services

*“New applications and microservices”*

**Latency**  
(mean and tail)

## 2. Edge analytics in IoT

*“Scalable live video analytics”*

**Bandwidth**  
(peak and average)

## 3. Exposure firewall in the IoT

*“Crossing the IoT Chasm”*

**Privacy**

## 4. Mask disruption of cloud services

*“Disconnected operation for cloud services”*

**Availability**

# What is a Cloudlet?

*aka “micro data center”, “mobile edge cloud”, “fog node”*

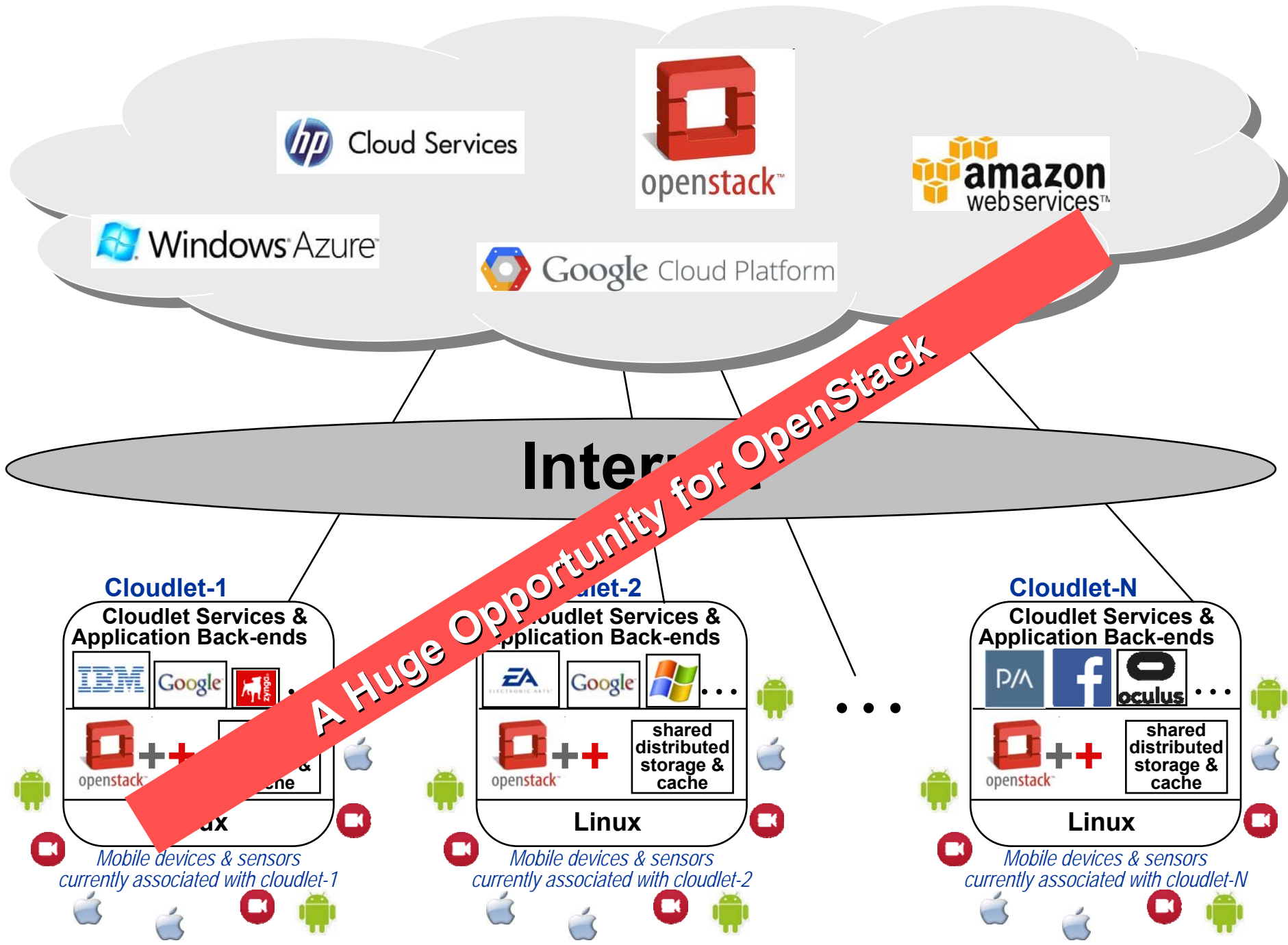
## ***Small data center at the edge of the Internet (many sizes & forms)***

- **one wireless hop (+fiber or LAN) to mobile devices**  
(Wi-Fi or 4G LTE or 5G)
- **multi-tenant, as in cloud**
- **good isolation and safety (VM-based guests)**
- **lighter-weight containers (e.g. Docker within VMs) also possible**

## ***Non-constraints*** (relative to mobile devices)

- **energy**
- **weight/size/heat**

***Catalyst for new mobile applications***



# Does Latency Really Matter?

***"The Impact of Mobile Multimedia Applications on Data Center Consolidation"***

Ha, K., Pillai, P., Lewis, G., Simanta, S., Clinch, S., Davies, N., Satyanarayanan, M.

Proceedings of IEEE International Conference on Cloud Engineering (IC2E), San Francisco, CA, March 2013

***"Quantifying the Impact of Edge Computing on Mobile Applications"***

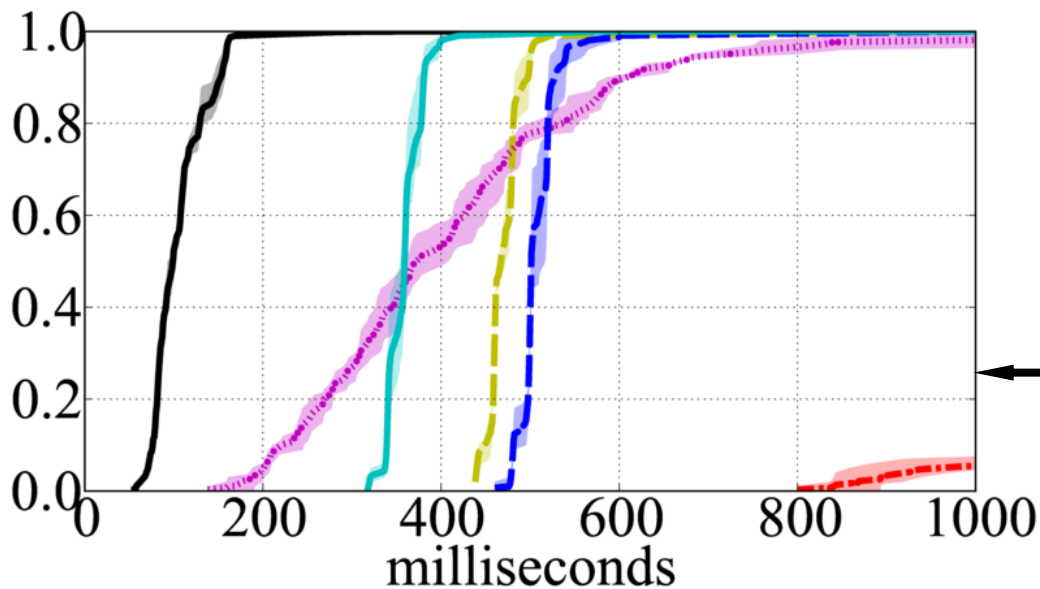
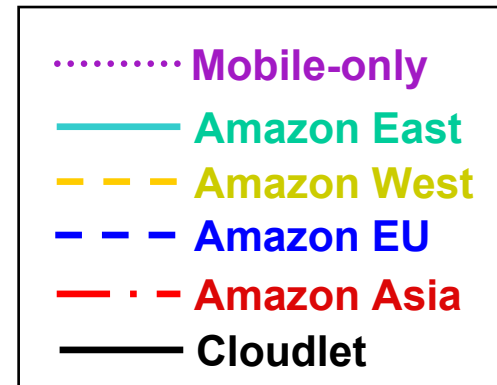
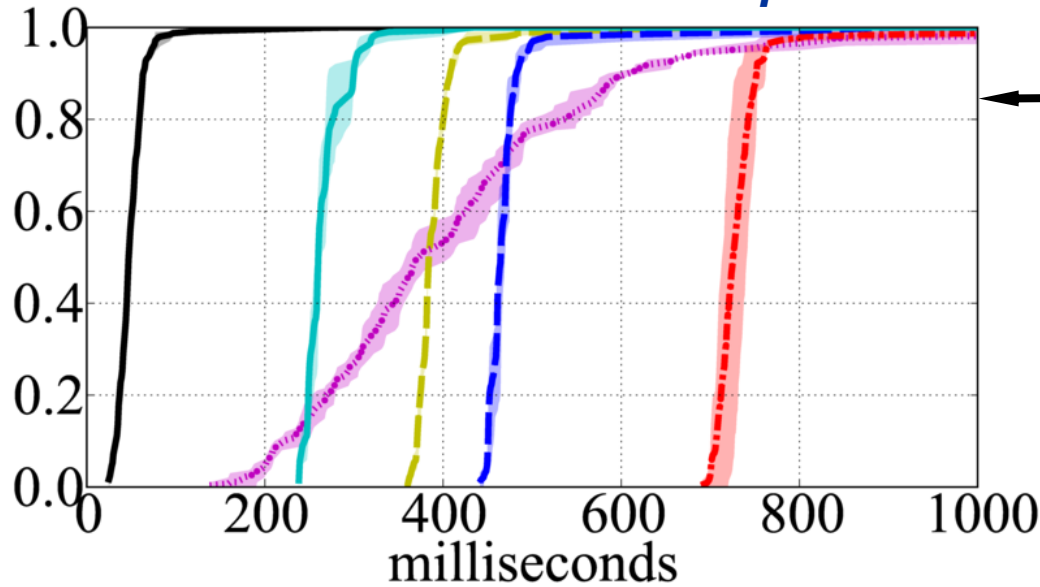
Hu, W., Gao, Y., Ha, K., Wang, J., Amos, B., Pillai, P., Satyanarayanan, M.

Proceedings of ACM APSys 2016, Hong Kong, China, August 2016

# Augmented Reality

## E2E Response Time CDF

1. Send JPG image from device to cloud/cloudlet
2. Recognize landmark buildings using computer vision
3. Send labels & coordinates back to device



# Per-Operation Energy Use by Device

Face Recognition		Augmented Reality
12.4 J	..... Mobile-only	5.4 J
2.6 J	—— Cloudlet	0.6 J
4.4 J	—— Amazon East	3.0 J
6.1 J	- - Amazon West	4.3 J
9.2 J	- - Amazon EU	5.1 J
9.2 J	— . Amazon Asia	7.9 J

# What is the Killer Use Case?

## ***“Towards Wearable Cognitive Assistance”***

Ha, K., Chen, Z., Hu, W., Richter, W., Pillai, P., Satyanarayanan, M.

Proceedings of the Twelfth International Conference on Mobile Systems, Applications, and Services (MobiSys 2014), Bretton Woods, NH, June 2014

## ***“Early Implementation Experience with Wearable Cognitive Assistance Applications”***

Chen, Z., Jiang, L., Hu, W., Ha, K., Amos, B., Pillai, P., Hauptmann, A., Satyanarayanan, M.

Proceedings of WearSys 2015, Florence, Italy, May 2015

## ***“An Empirical Study of Latency in an Emerging Class of Edge Computing Applications for Wearable Cognitive Assistance”***

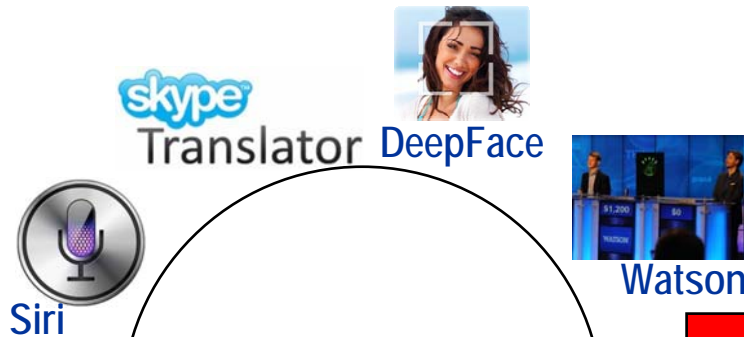
Chen, Z., Hu, W., Wang, J., Zhao, S., Amos, B., Wu, G., Ha, K., Elgazzar, K., Pillai, P., Klatzky, R., Siewiorek, D., Satyanarayanan, M.

Proceedings of SEC 2017, San Jose, CA, October 2017



# A Unique Moment in Time

*Convergence of  
Advances in 3  
Independent  
Arenas*



**Cognitive  
Algorithms**

**This  
Research**

**Edge  
Computing**

**Wearable  
Hardware**



**Cloudlets**



# Wearable Cognitive Assistance

*A new modality of computing*

Entirely new genre of applications

Wearable UI with wireless access to cloudlet

**Real-time cognitive engines** on cloudlet

- scene analysis
- object/person recognition
- speech recognition
- language translation
- planning, navigation
- question-answering technology
- voice synthesis
- real-time machine learning
- ...

Low latency response is crucial



*Seamlessly integrated into inner loop of human cognition*

# Task-specific Assistance

## Example: cooking

### passive recipe display



### versus active guidance



“Wait, the oil is not hot enough”

# Inspiration: GPS Navigation Systems

## Turn by turn guidance

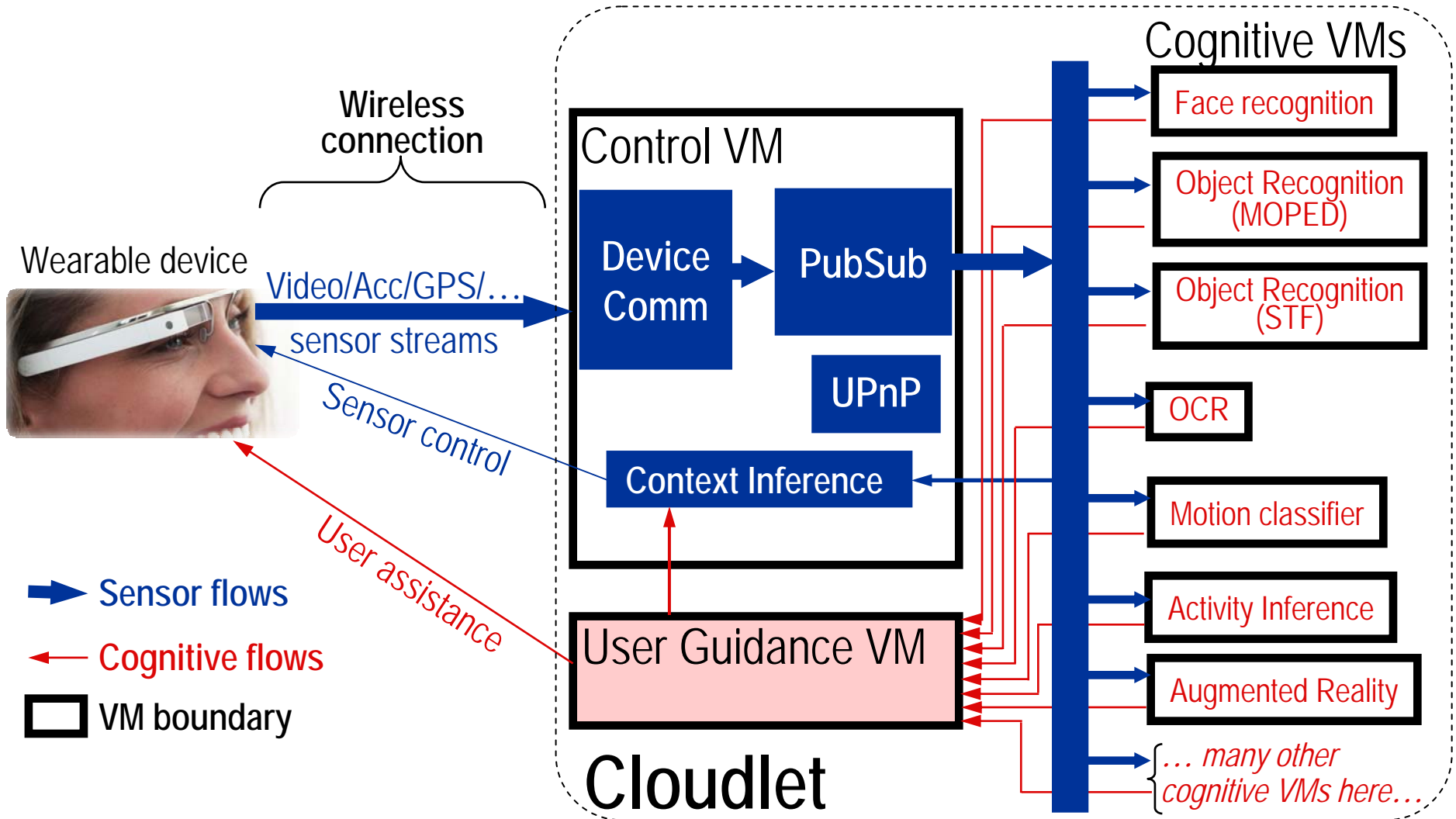
- Ability to detect and recover
- Minimally distracting to user

Uses only one type of sensor: location from GPS

*Can we generalize this metaphor?*

# Gabriel Architecture

*(PaaS for Wearable Cognitive Assistance)*



# Baby Steps: 2D Lego Assembly

Very first proof-of-concept (September 2014)

Deliberately simplified task to keep computer vision tractable

[2D Lego Assembly](http://youtu.be/uy17Hz5xvmY) (YouTube video at <http://youtu.be/uy17Hz5xvmY>)

# On Each Video Frame



(a) Input image



(b) Detected dark parts



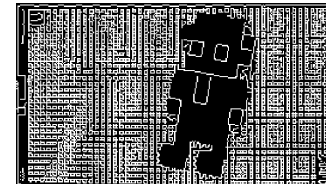
(c) Detected board



(d) Board border



(e) Perspective corrected



(f) Edges detected



(g) Background subtracted



(h) Side parts added



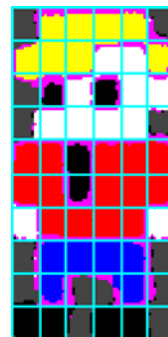
(h) Lego detected



(i) Unrotated



(i) Color quantized



(j) Partitioned

```
[[0, 3, 3, 3, 3, 0],
 [3, 3, 3, 1, 1, 3],
 [0, 6, 1, 6, 1, 1],
 [0, 1, 1, 1, 1, 0],
 [4, 4, 6, 4, 4, 4],
 [4, 4, 6, 4, 4, 4],
 [1, 4, 4, 4, 4, 1],
 [0, 5, 5, 5, 5, 0],
 [0, 5, 0, 0, 5, 0],
 [6, 6, 0, 6, 6, 0]]
```

(j) Matrix



(k) Synthesized

# When Milliseconds Matter

**Ping-pong assistant**

([https://www.youtube.com/watch?v=\\_lp32sowyUA](https://www.youtube.com/watch?v=_lp32sowyUA))



# Assembling an IKEA Kit

## IKEA kit assistant

([https://www.youtube.com/watch?v=qDPuvBWNIUs&index=5&list=PLmrZVvFtthdP3fwHPy\\_4d61oDvQY\\_RBgS](https://www.youtube.com/watch?v=qDPuvBWNIUs&index=5&list=PLmrZVvFtthdP3fwHPy_4d61oDvQY_RBgS))

# Many Monetizable Use Cases ...



Assembly instructions



Industrial troubleshooting



Medical training



Correct Self-Instrumentation



Strengthening willpower

# AR Meets AI

Low latency of AR + Compute intensity of AI

Has the “look and feel” of AR, but the functionality of AI

October 9, 2016: CBS “60 Minutes” special on AI

[Short \(90 seconds\) video clip on Gabriel](#)

YouTube video at [https://youtu.be/dNH\\_HF-C5KY](https://youtu.be/dNH_HF-C5KY)

Full 60 Minutes special (~30 minutes) at CBS web site:

<http://www.cbsnews.com/videos/artificial-intelligence>

# In Closing

***Edge Computing is transformative***

***It enables revolutionary new applications***

***It is a truly disruptive technology!***