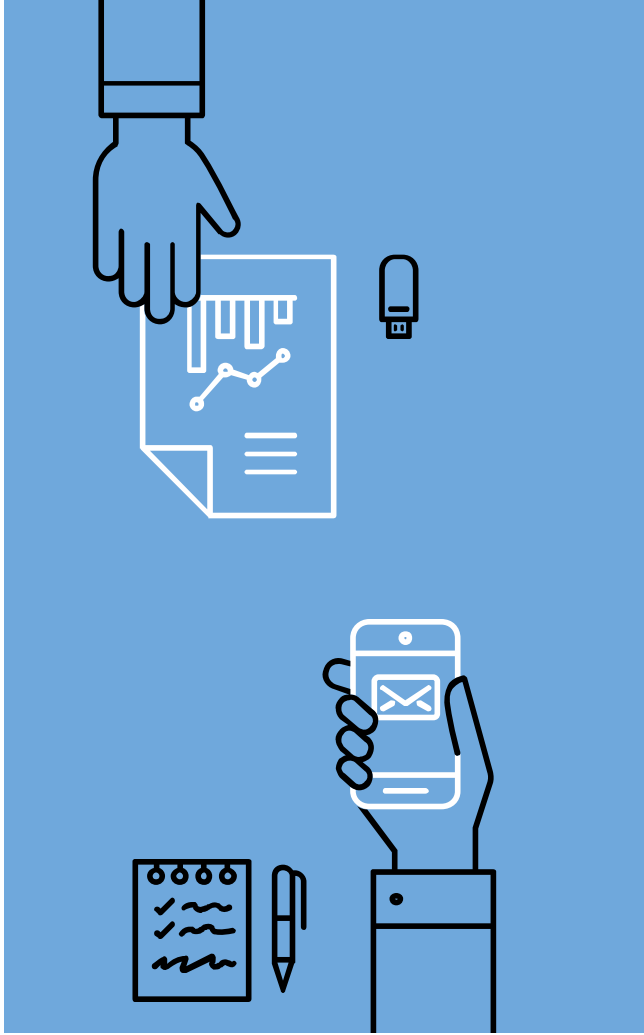
Living on the Cloud's Edge

Tannaz N. Roshandel
@tannaznvr

Rustam Mehmandarov
@rmehmandarov

Building something...

- Comply with Strict Privacy Policy
- High Performance
- Quick (Low Latency)
- Cheap to Run



HELLO!

Rustam Mehmandarov

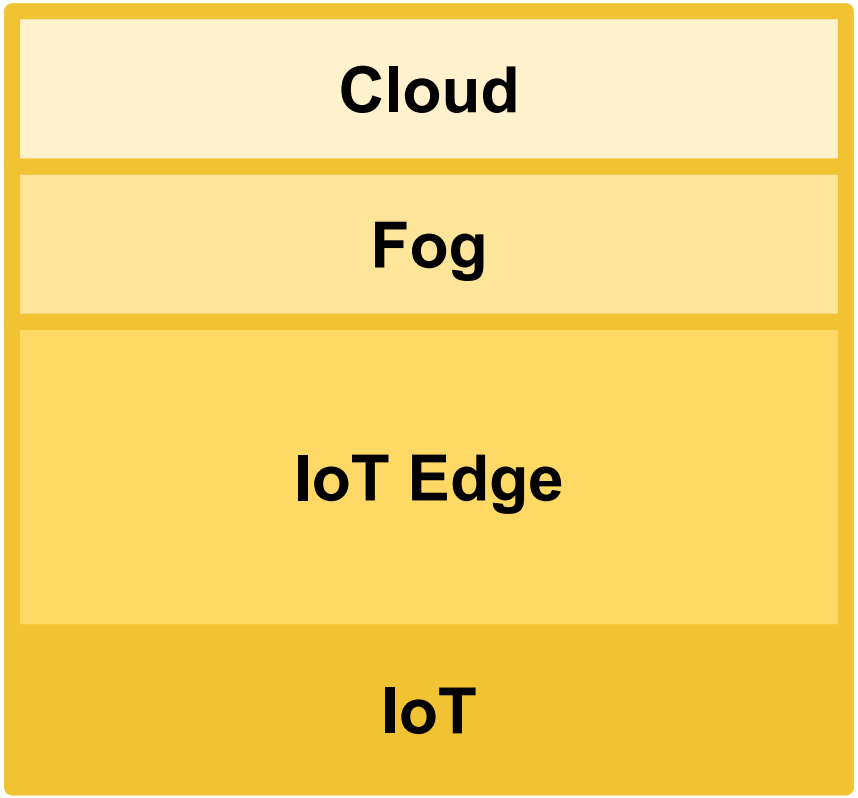
@rmehmandarov

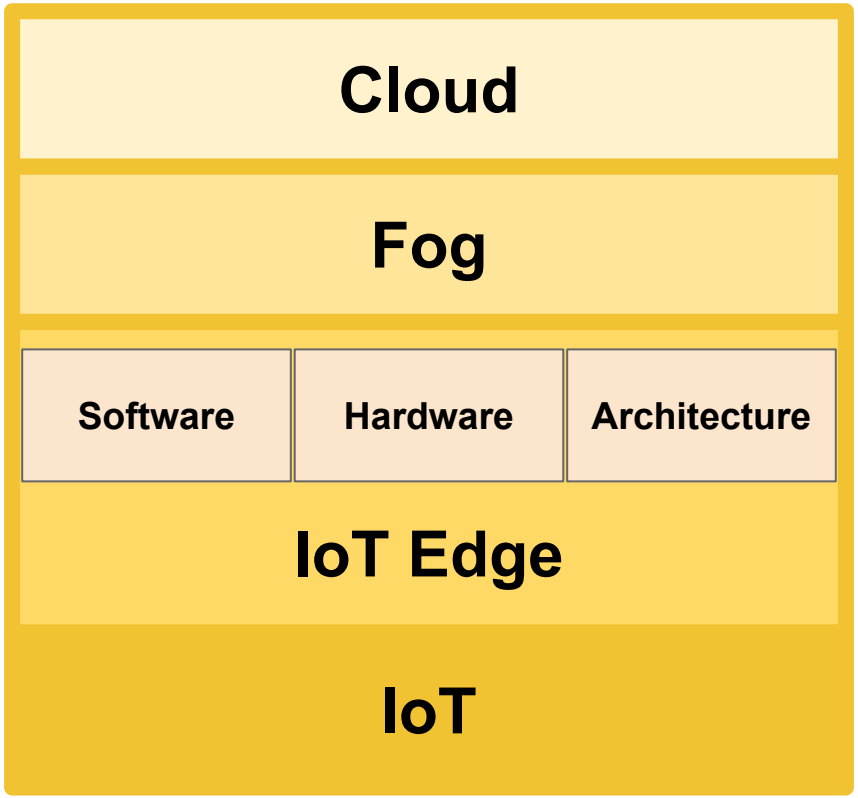
computas

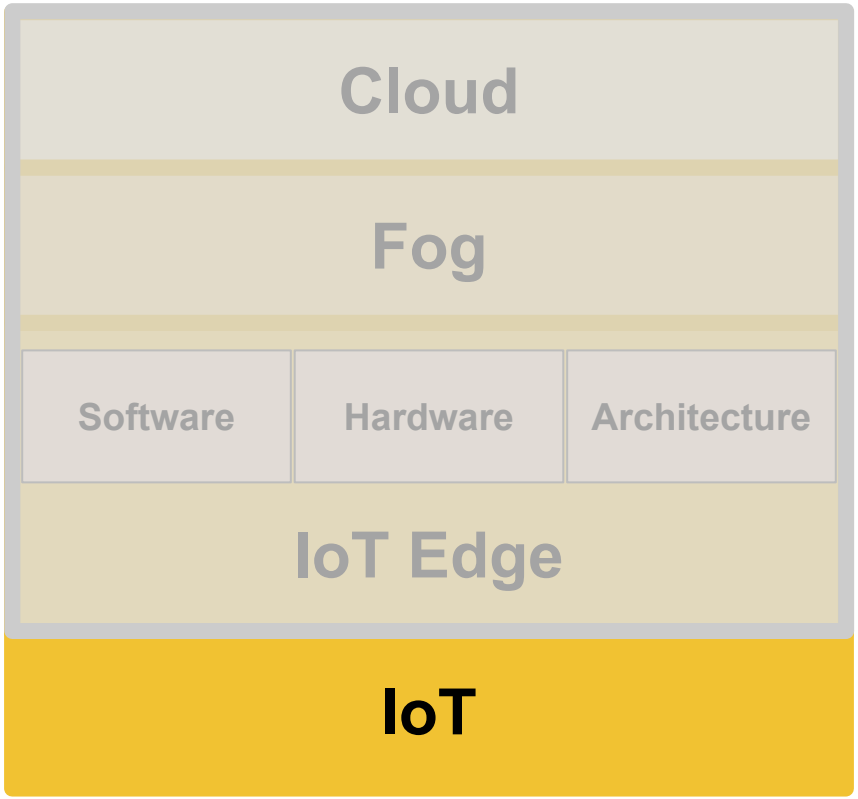
Tannaz N. Roshandel

@tannaznvr

UiO : University of Oslo


Cloud
Fog
IoT Edge
IoT


Cloud
Fog
Software
Hardware
Architecture
IoT Edge
IoT


Cloud
Fog
Software
Hardware
Architecture
IoT Edge
IoT

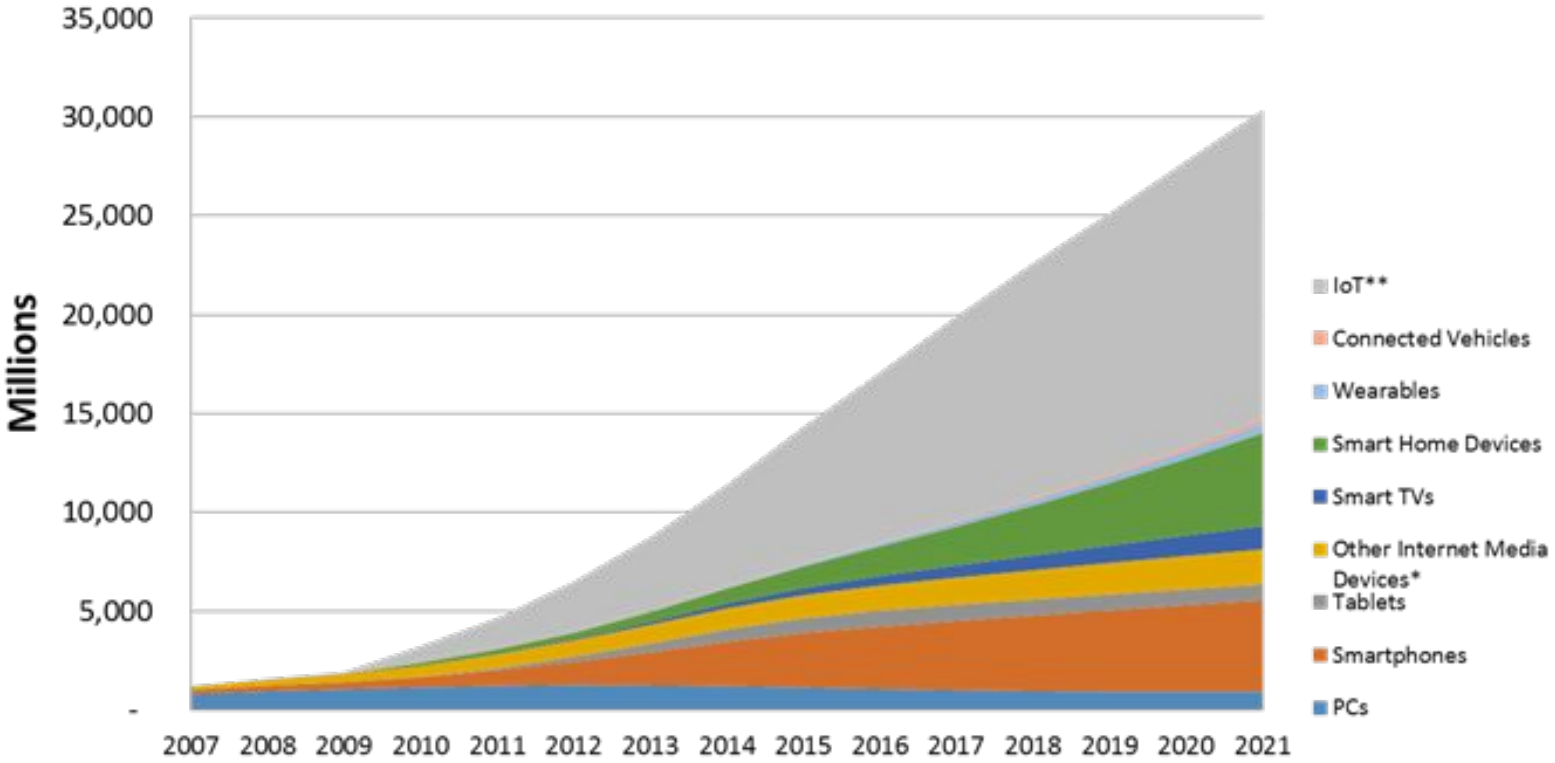
IoT

Internet of Things

Electronic devices with Internet connectivity that can be remotely monitored and controlled

STRATEGY ANALYTICS

Global Connected and IoT Device Installed Base Forecast



Source – Strategy Analytics research services ,October 2017: IoT Strategies , Connected Home Devices, Tablet and Touchscreen Strategies, Wireless Smartphone Strategies, Wearable Device Ecosystem, Smart Home Strategies

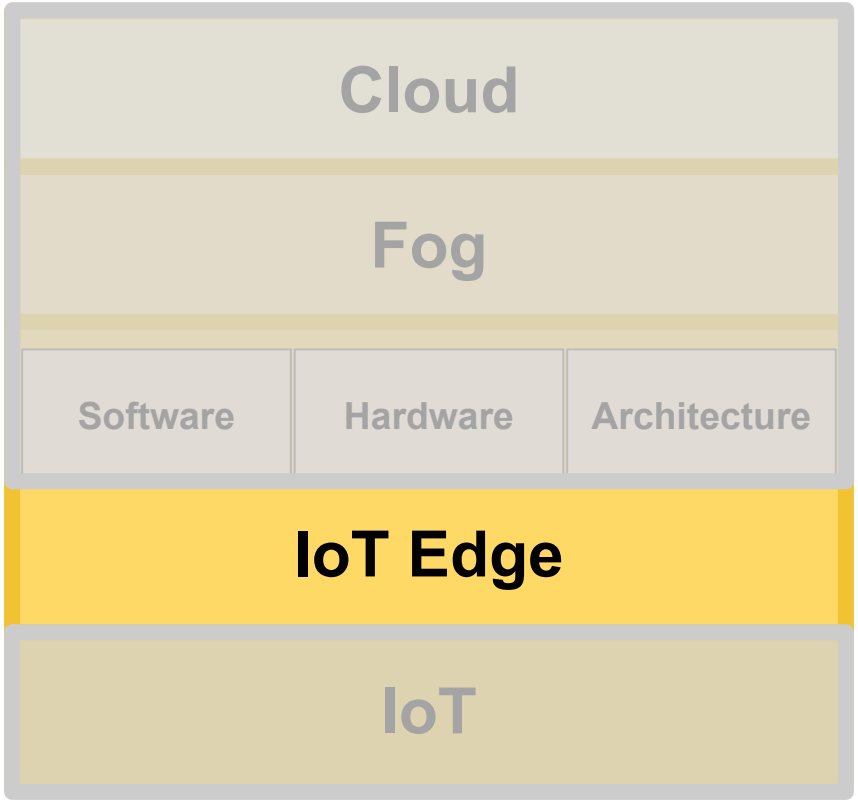
Internet of Things

\+

- Lead time processing
- Accessibility
- Mobility
- Tracking Ability
- Efficiency
- Cost

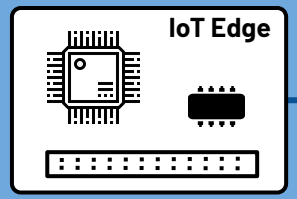
\-

- Data Security
- Privacy
- Massivity of Data
- Environment Issues
- Complexity
- Network dependent


Cloud
Fog
Software
Hardware
Architecture
IoT Edge
IoT

Edge Computing

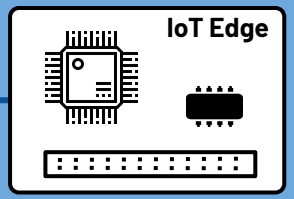
Improve operations
and cut costs



Edge computing is a mesh network of ***micro data centers*** *that* ***process or store*** *critical data locally and* ***push*** *it to a central data center or Cloud.*

— IDC

International Data Corporation



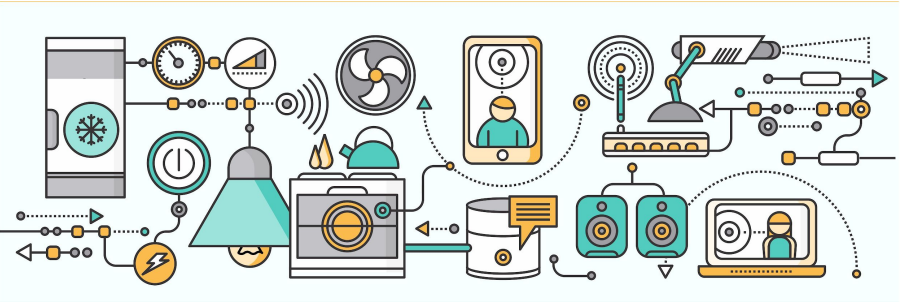
Edge Computing

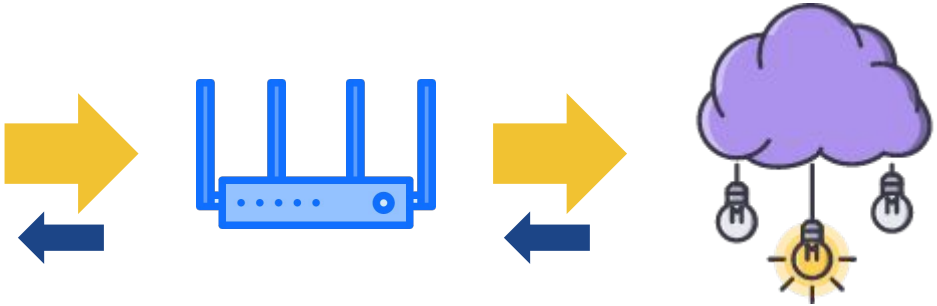
+

- Low Latency
- Privacy
- Real-Time Availability
- Real-Time Data Transmission
- Productivity Increase

-

- Limited (or no) Redundancy
- Potential Loss of Data
- Higher Risk in general





IoT “Smart Devices”, Assistants, etc.
(not Edge devices)

Gateway

Cloud

Rich Rogers
@RichRogersIoT

Følg

My wife asked me why I was speaking so softly at home.

I told her I was afraid Mark Zuckerberg was listening!

She laughed. I laughed.

Alexa laughed. Siri laughed.

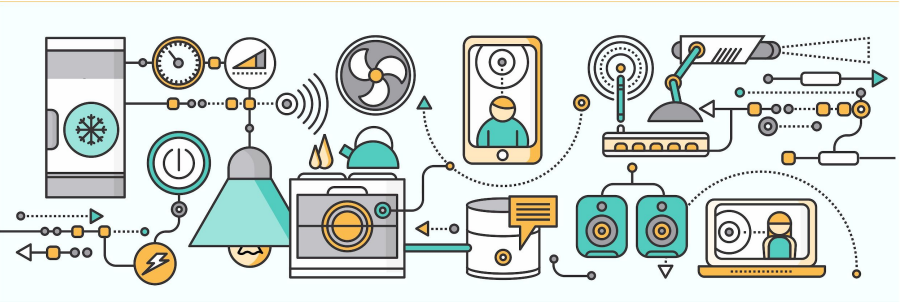
17:00 - 1. jun. 2019

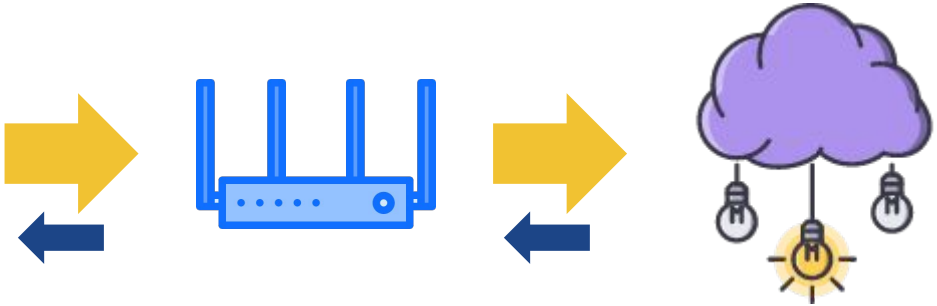
12 998 retweets **48 437** liker

346 13 K 48 K

https://xkcd.com/1807/

https://xkcd.com/1807/

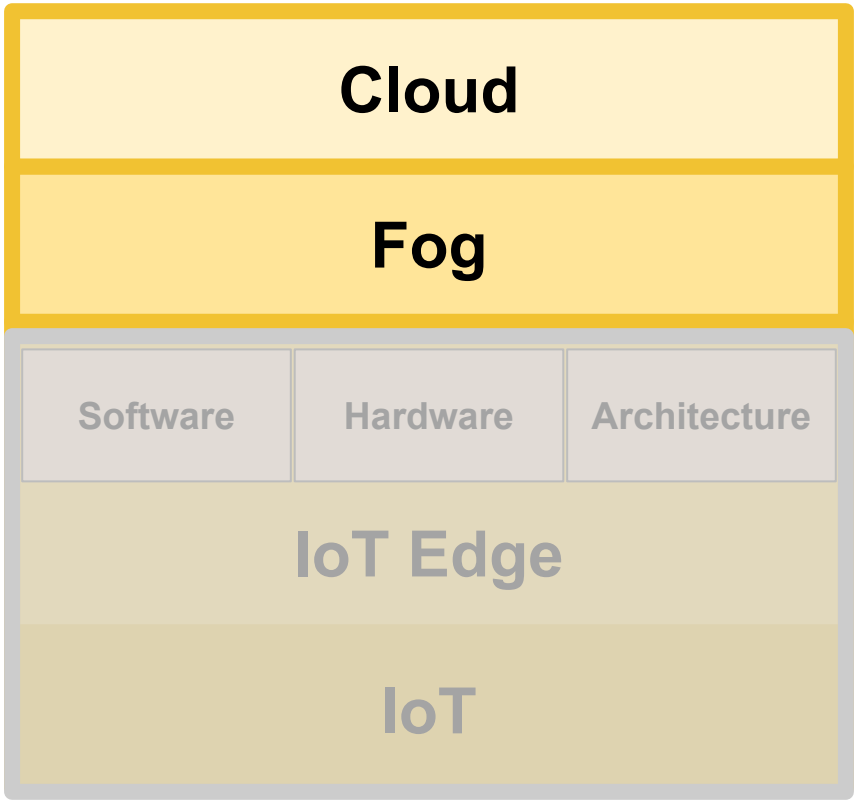




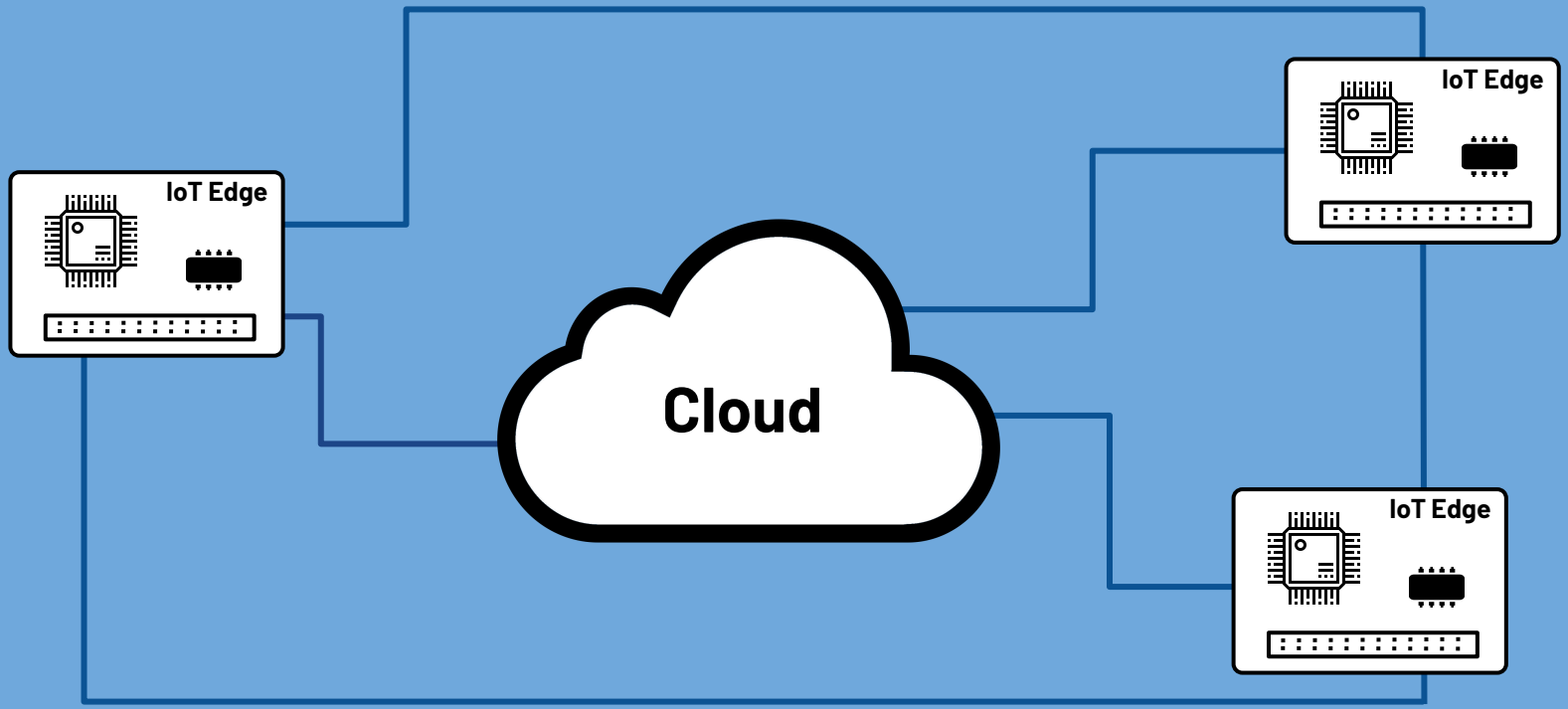
IoT “Smart Devices”, Assistants, etc.
(not Edge devices)

Gateway

Cloud

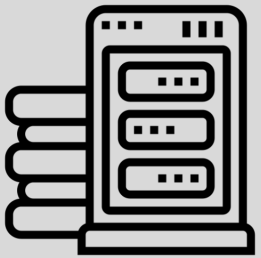

Cloud
Fog
Software
Hardware
Architecture
IoT Edge
IoT

What is Fog Computing?

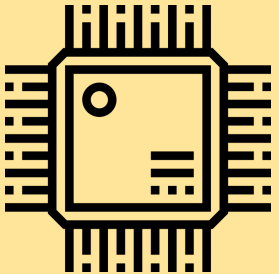


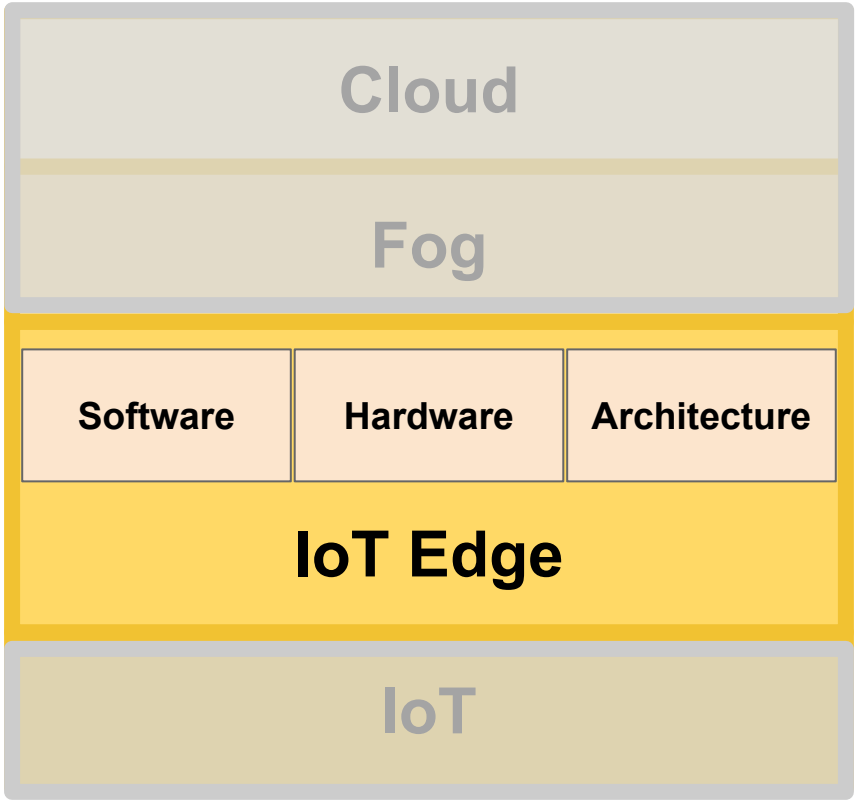
Cloud Computing

Fog Computing

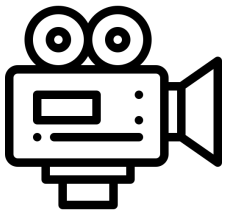


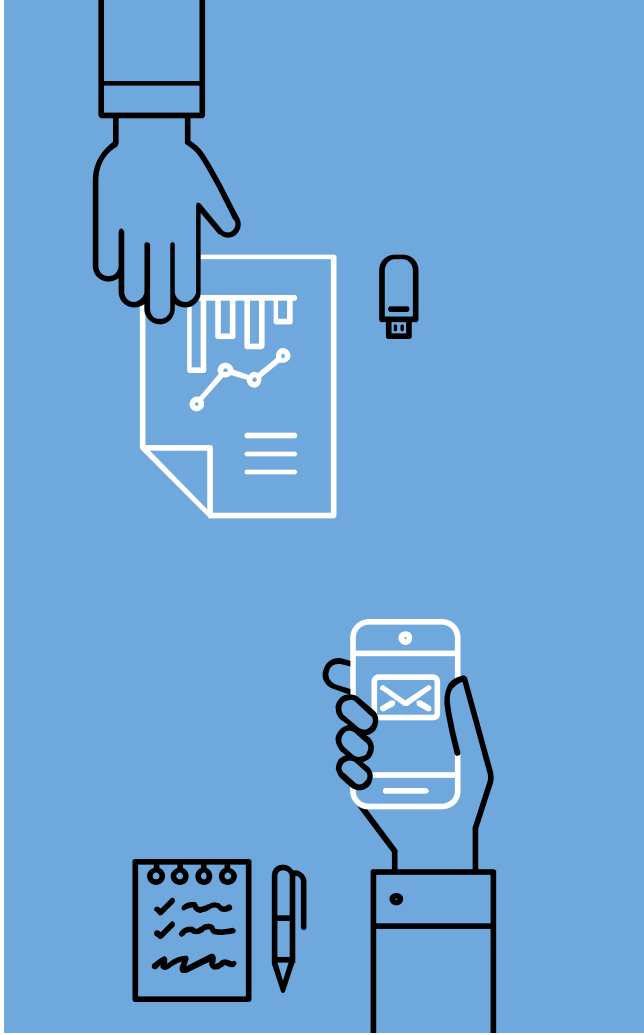
Edge Computing




Cloud
Fog
Software
Hardware
Architecture
IoT Edge
IoT

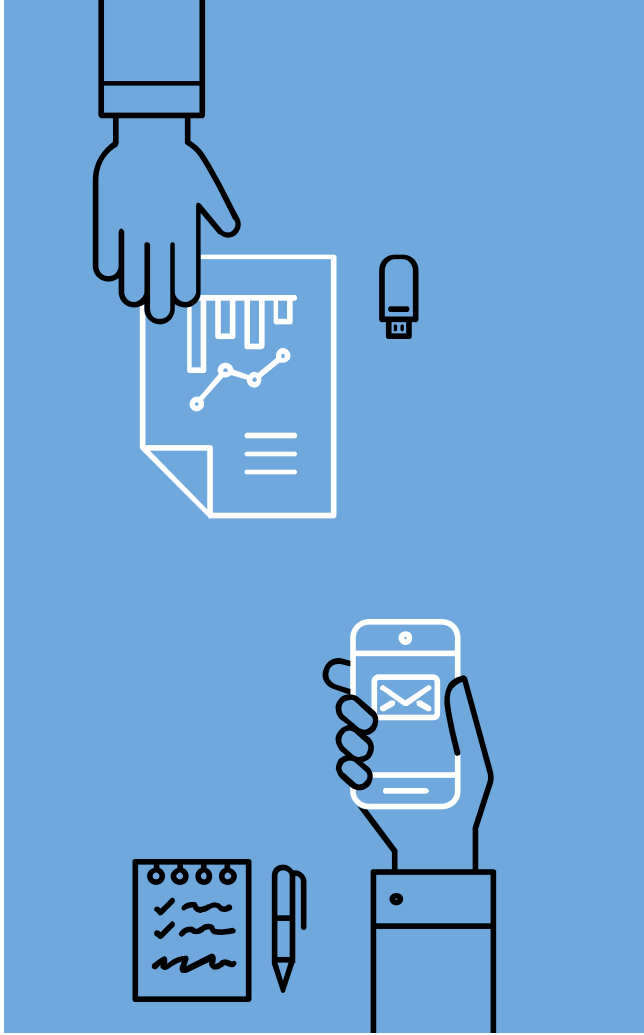
Real-Life Example





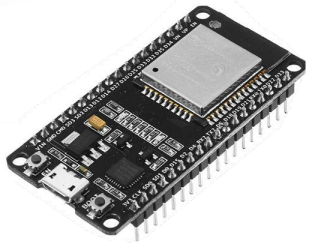
ML-based People Counter

- Privacy Concerns
- Performance
- Latency
- Cost



Edge Computing Devices

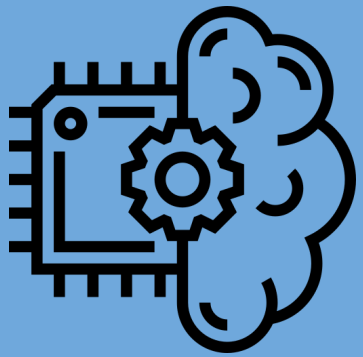
- Azure Sphere
- AWS FreeRTOS
- Google Coral



Raspberry Pi, Banana Pi, Intel Edison ++

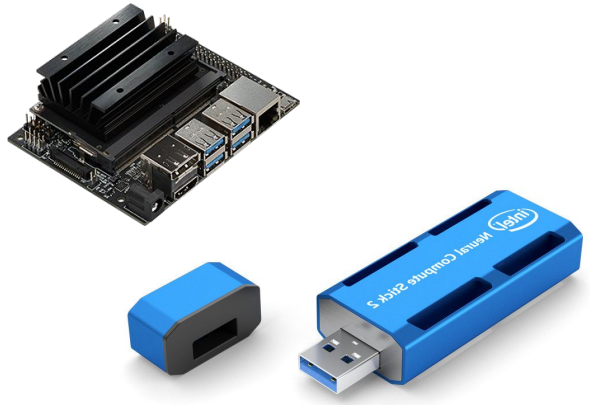
Edge AI Computing

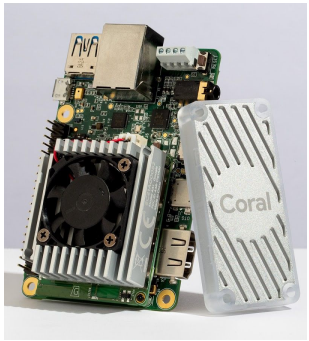
Edge Computing + AI =



AI Accelerators

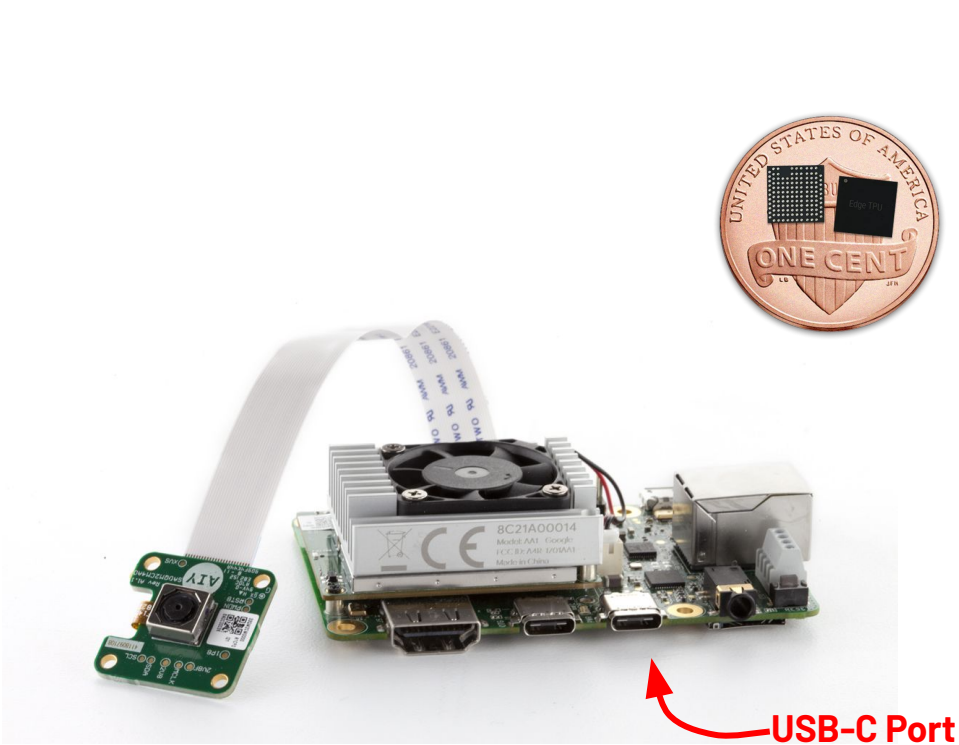
- NVIDIA Jetson
- Intel Movidius Myriad
- Google Edge TPU

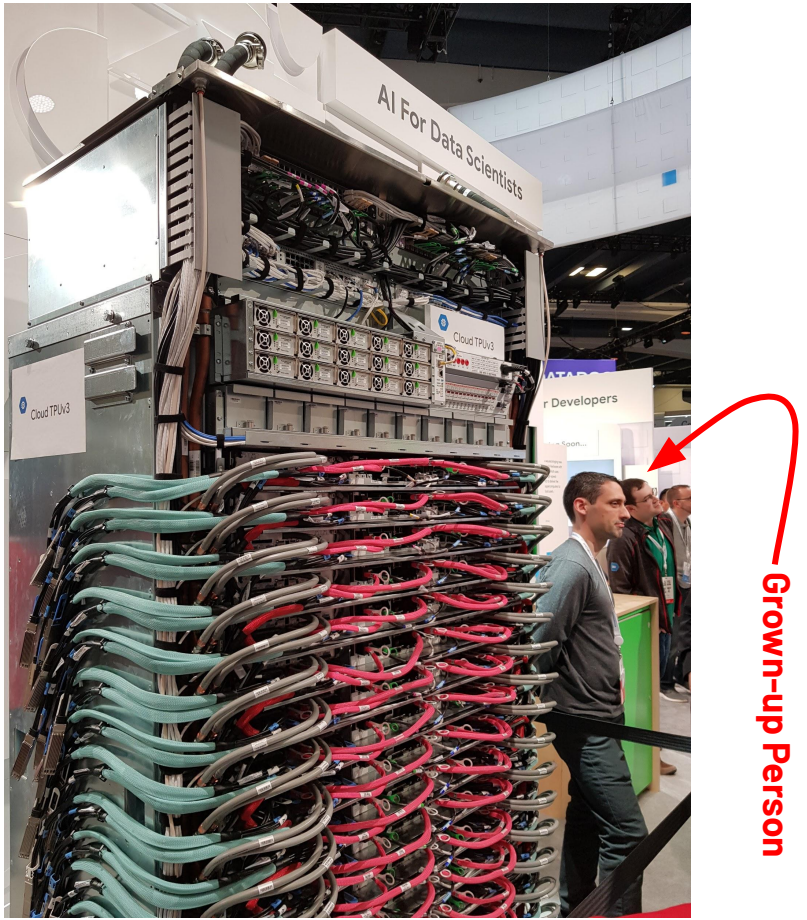




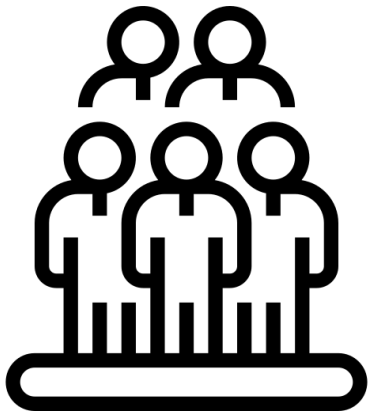
Google Coral Dev Board

Not all TPUs are alike...

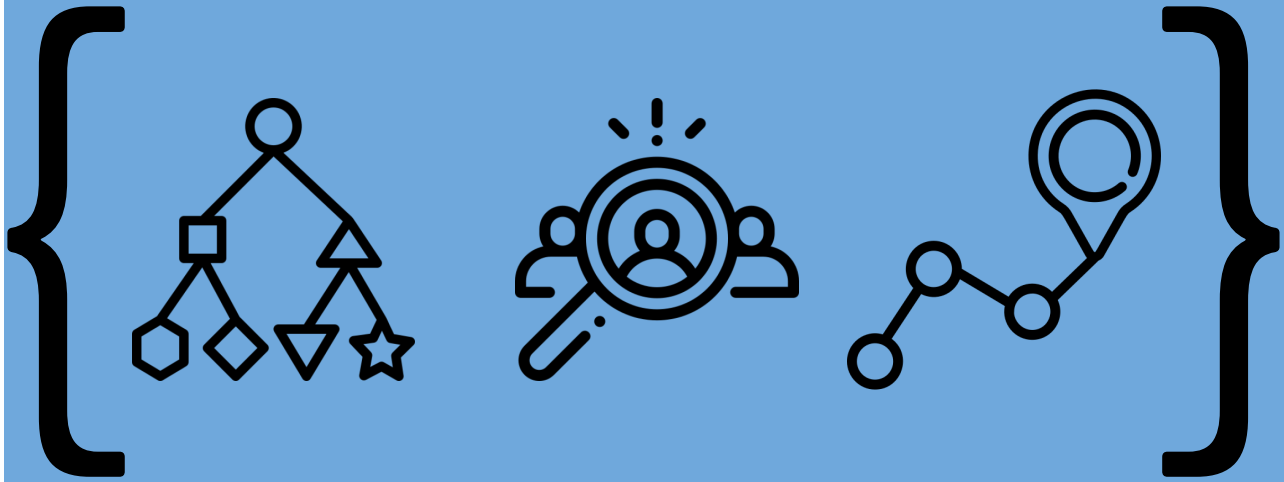




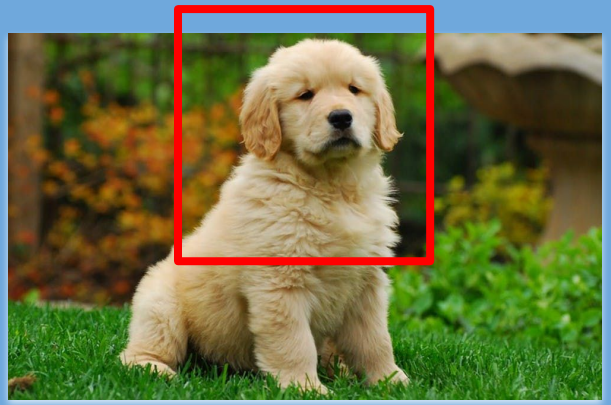
https://cloud.google.com/tpu/



People Counter



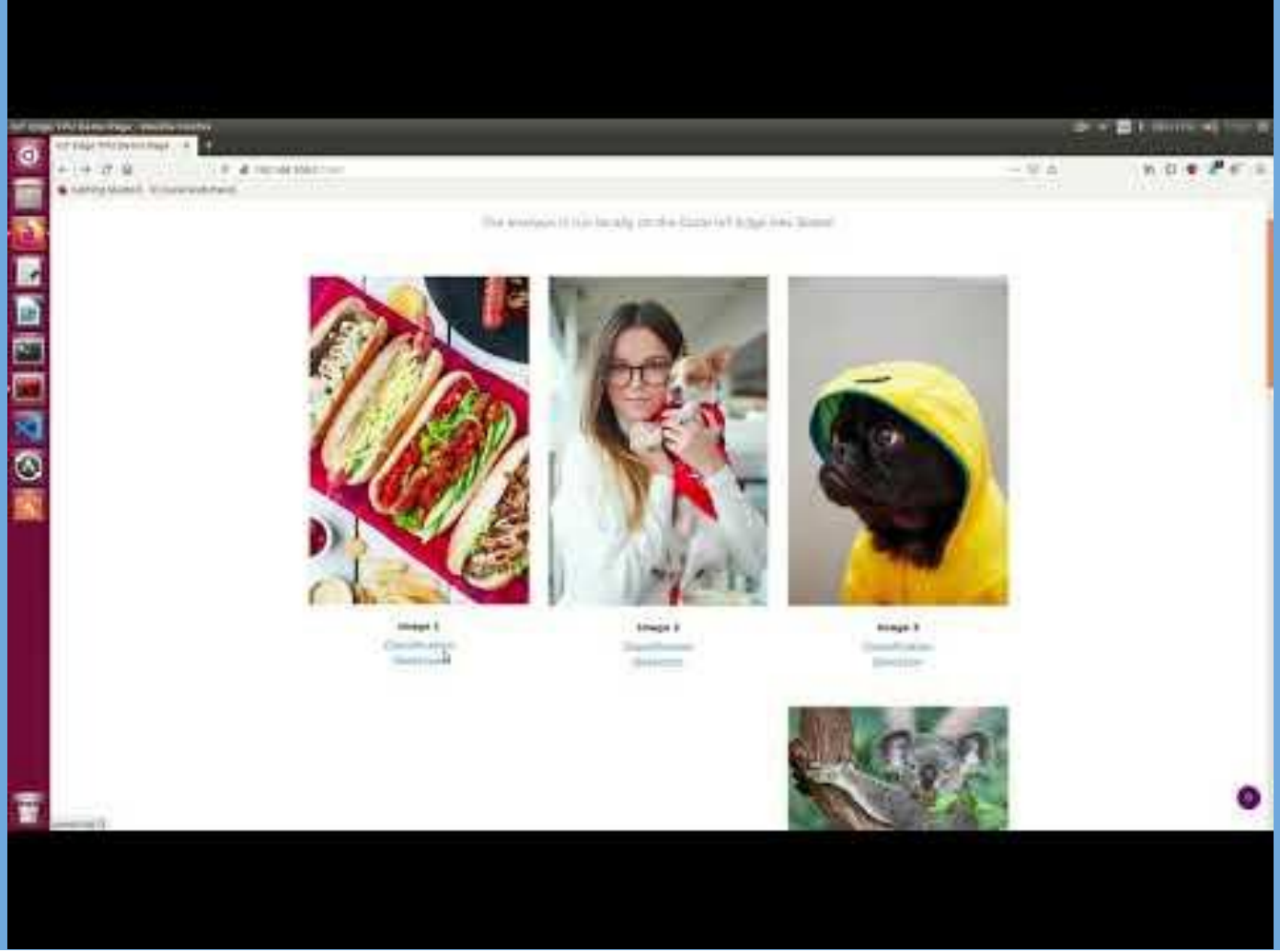
Classification Detection Tracking



Classification

Detection

Tracking



Detection vs. Tracking

- Video -> stream of images (fps)

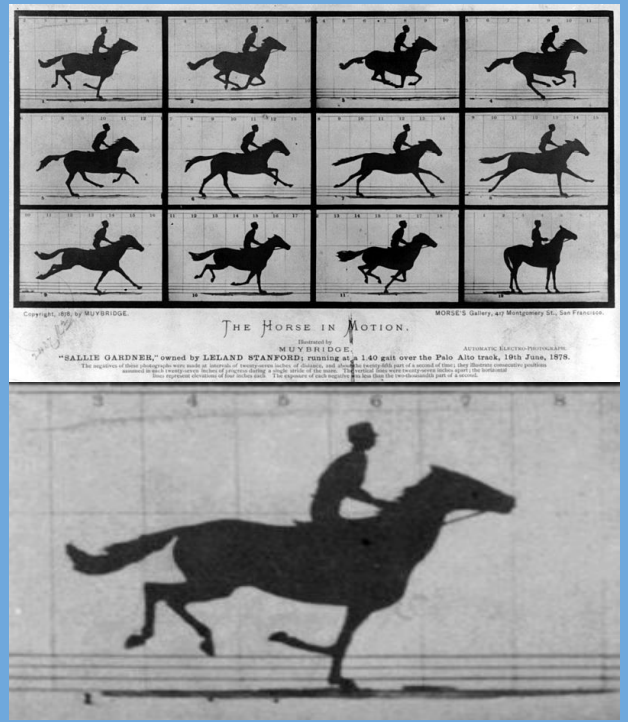
Lots of images!

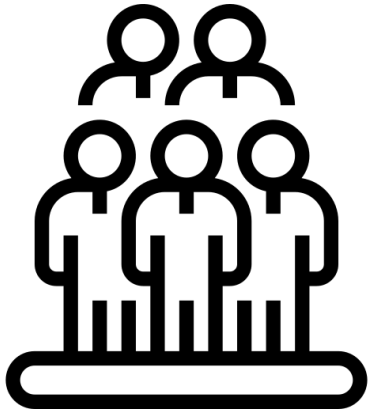
- Detect an object in each image.

ML is CPU intensive, but not for TPU!

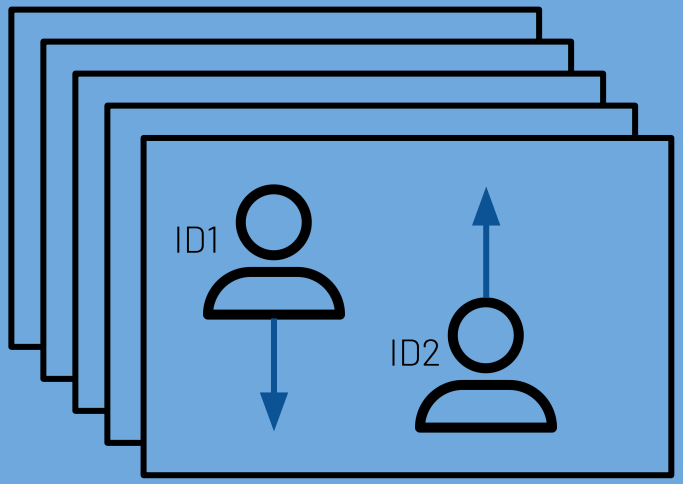
- Tracking = correlating objects between frames

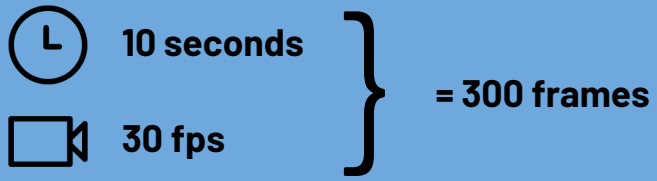
Quite CPU intensive and tricky!

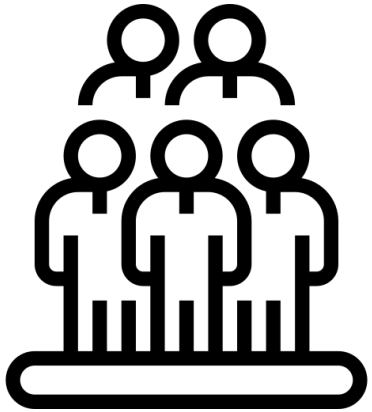




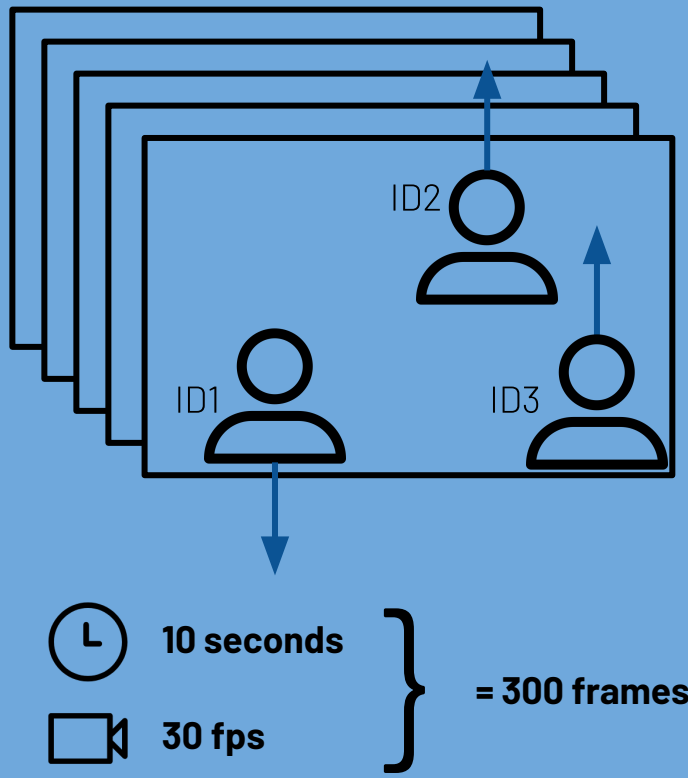
People
Counter

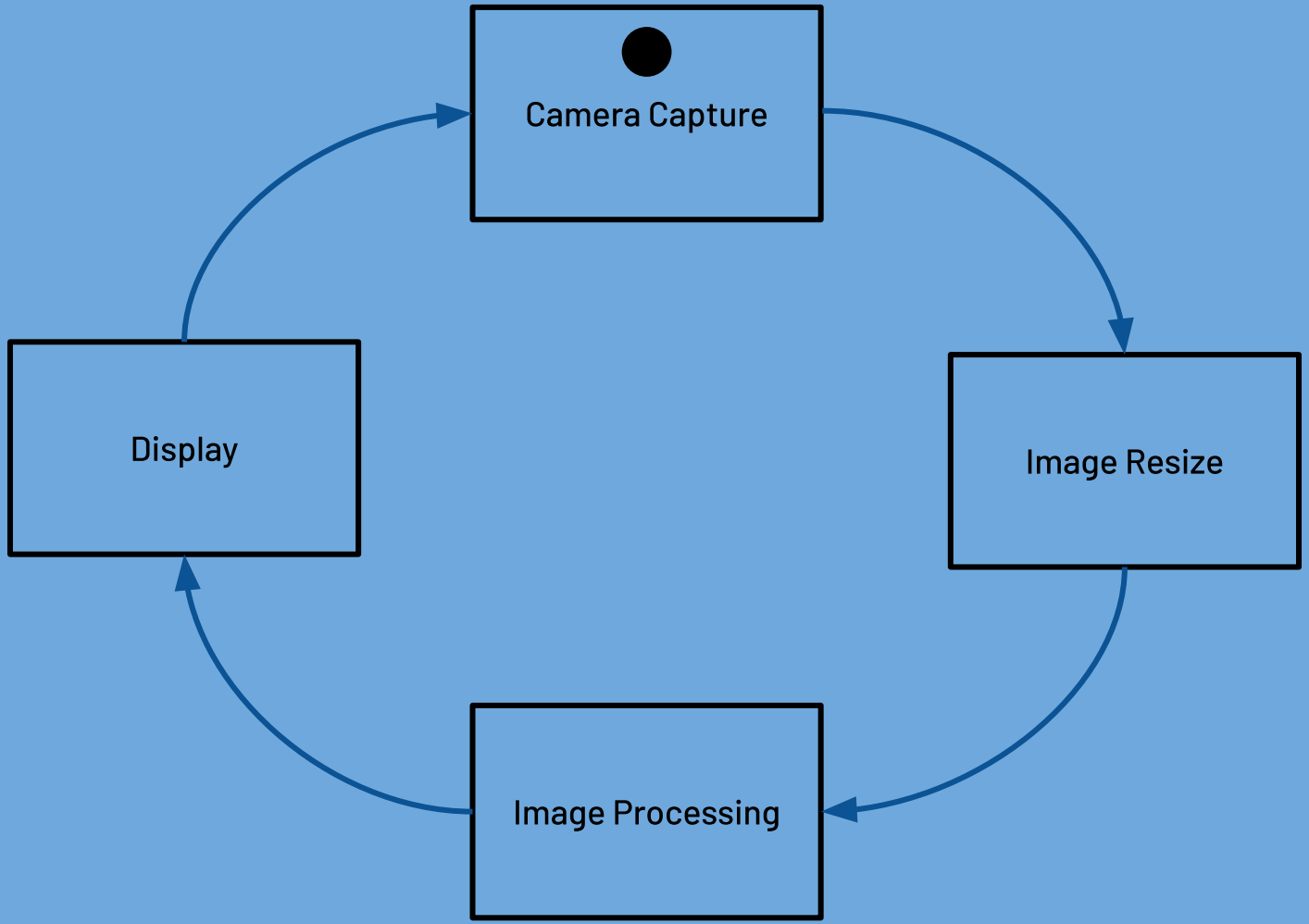




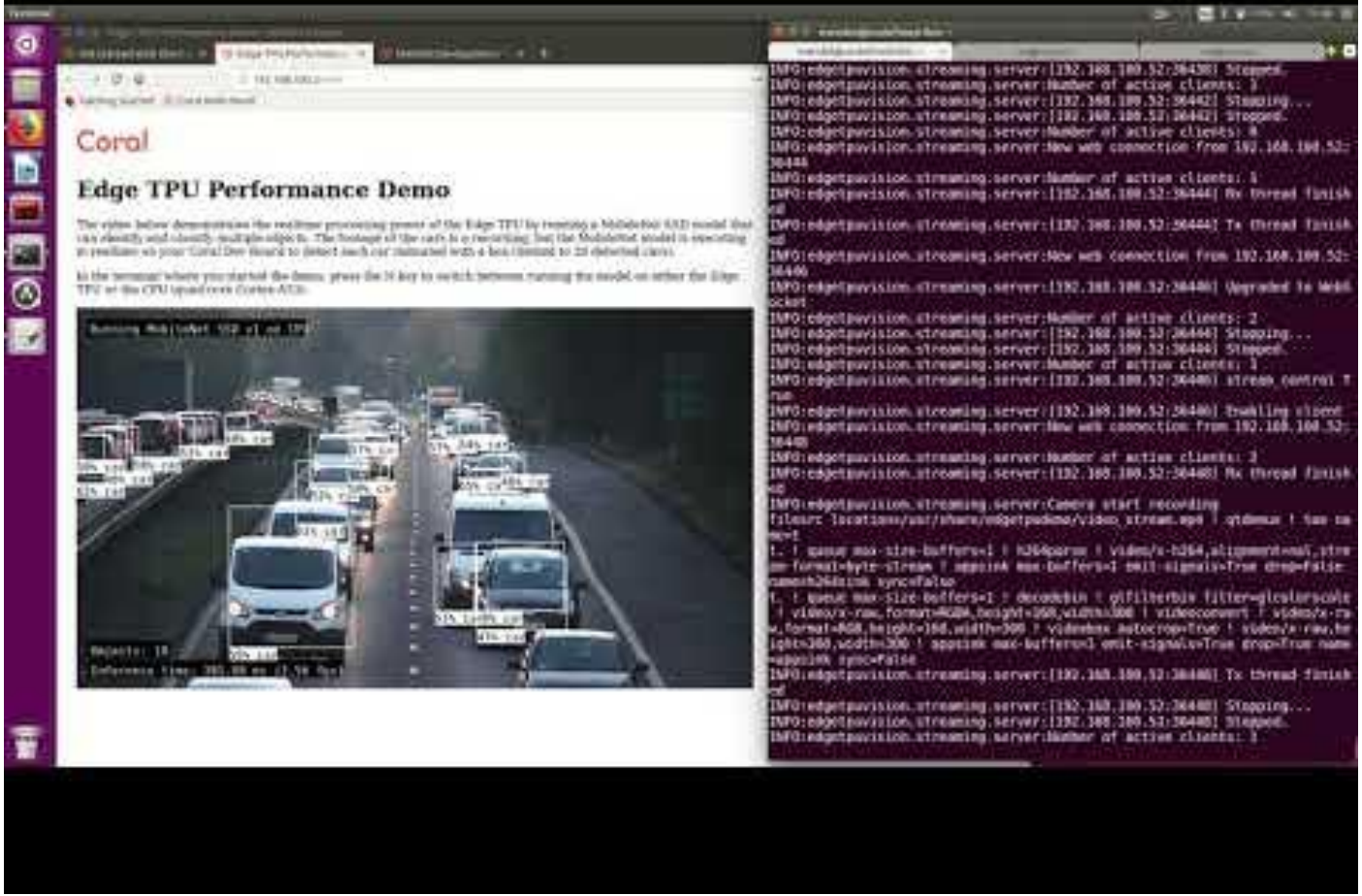


People
Counter



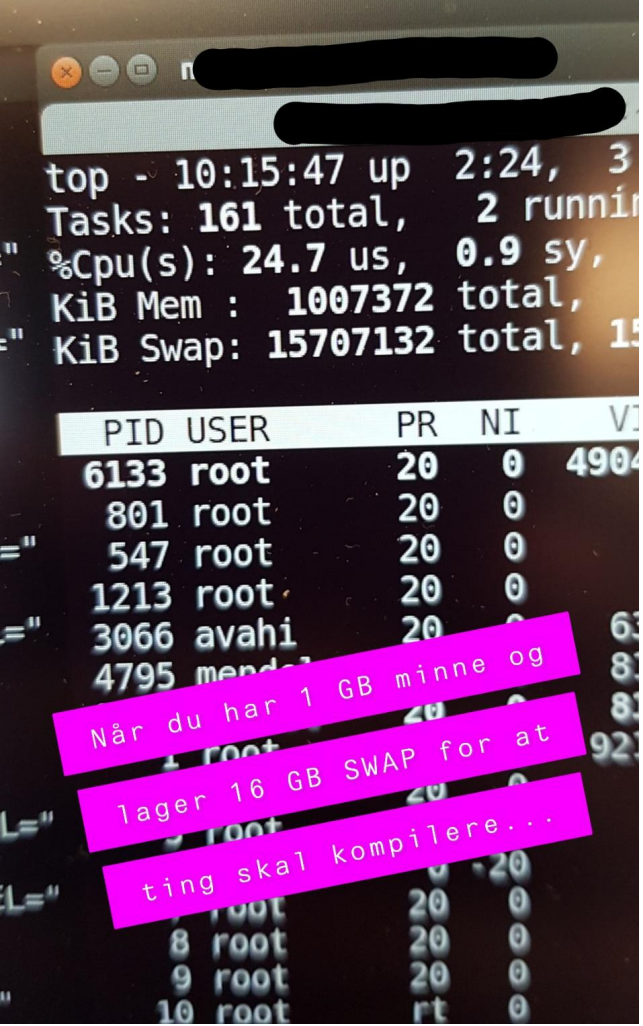

Camera Capture
Image Resize
Image Processing
Display

@rmehmandarov @tannaznvr

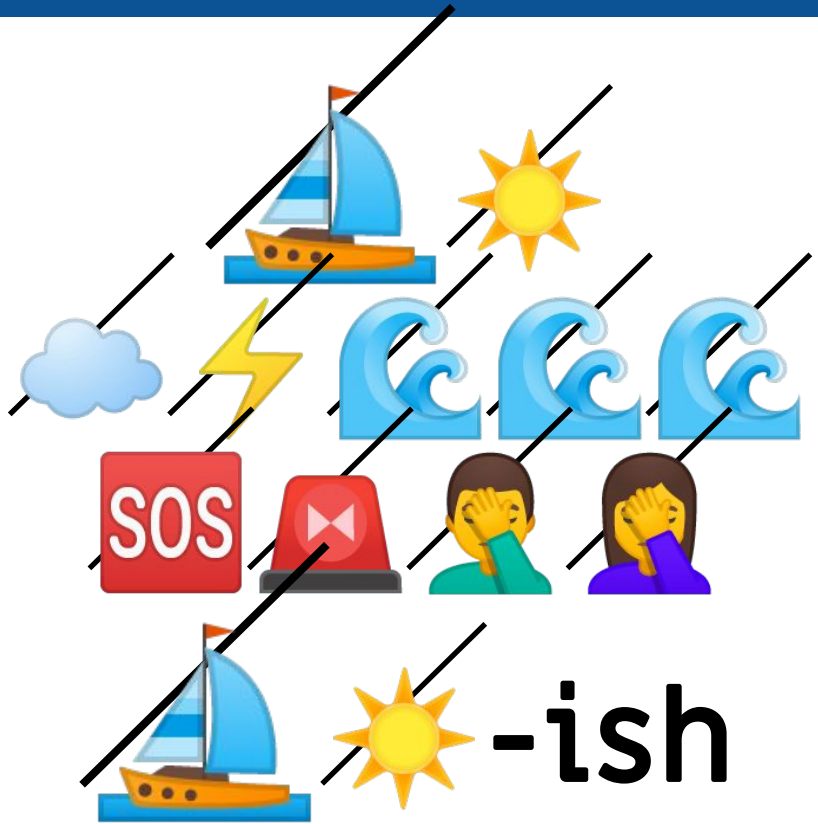

Coral
Edge TPU Performance Demo

Our Experiences

- Various models have difference performance
- Don’t put much code in synchronous functions
- Optimize, optimize, optimize



Smooth Sailing?



DEMO

THANK YOU!

Tannaz N. Roshandel
@tannaznvr

Rustam Mehmandarov
@rmehmandarov

UiO : University of Oslo

computas