



INTRODUCTION TO EDGE AI

— CVPR 2026 Tutorial —

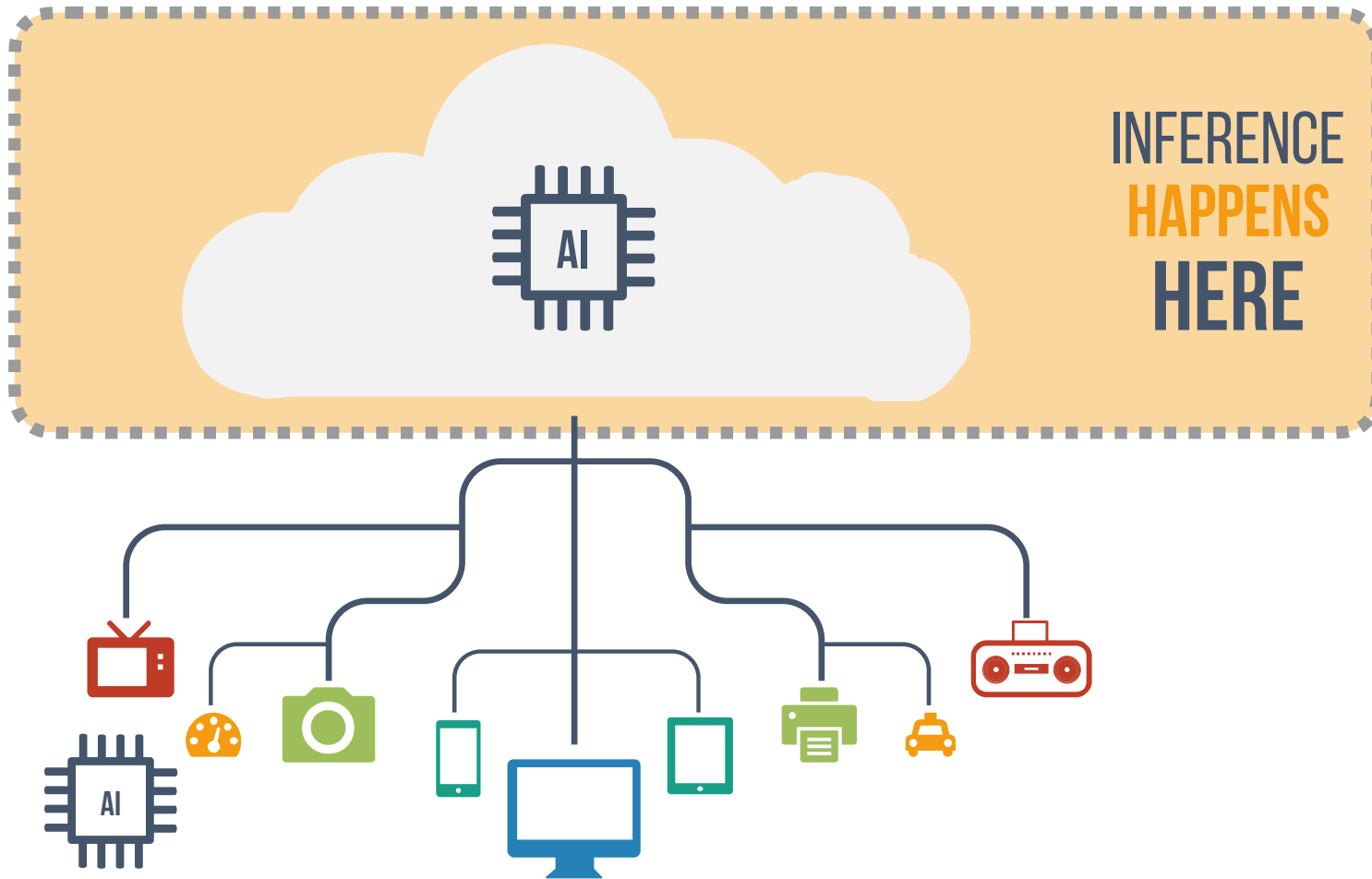
The IEEE/CVF Conference on Computer Vision and Pattern Recognition 2026

Denver, CO, USA

INTRODUCTION: WHAT IS **EDGE AI**?

WHAT IS CLOUD AI?

Introduction



1



Scalability

Cloud AI systems are highly scalable, allowing for adjustments based on the workload and user demand.

2



Accessibility

Users can access these technologies from anywhere in the world, requiring only an internet connection.

3



Cost-Effectiveness

You can utilize AI tools and computing power on a pay-as-you-go basis, which helps manage costs effectively.

4



Integration and Collaboration

The integration enables seamless data flow and collaboration across different platforms and teams.

5

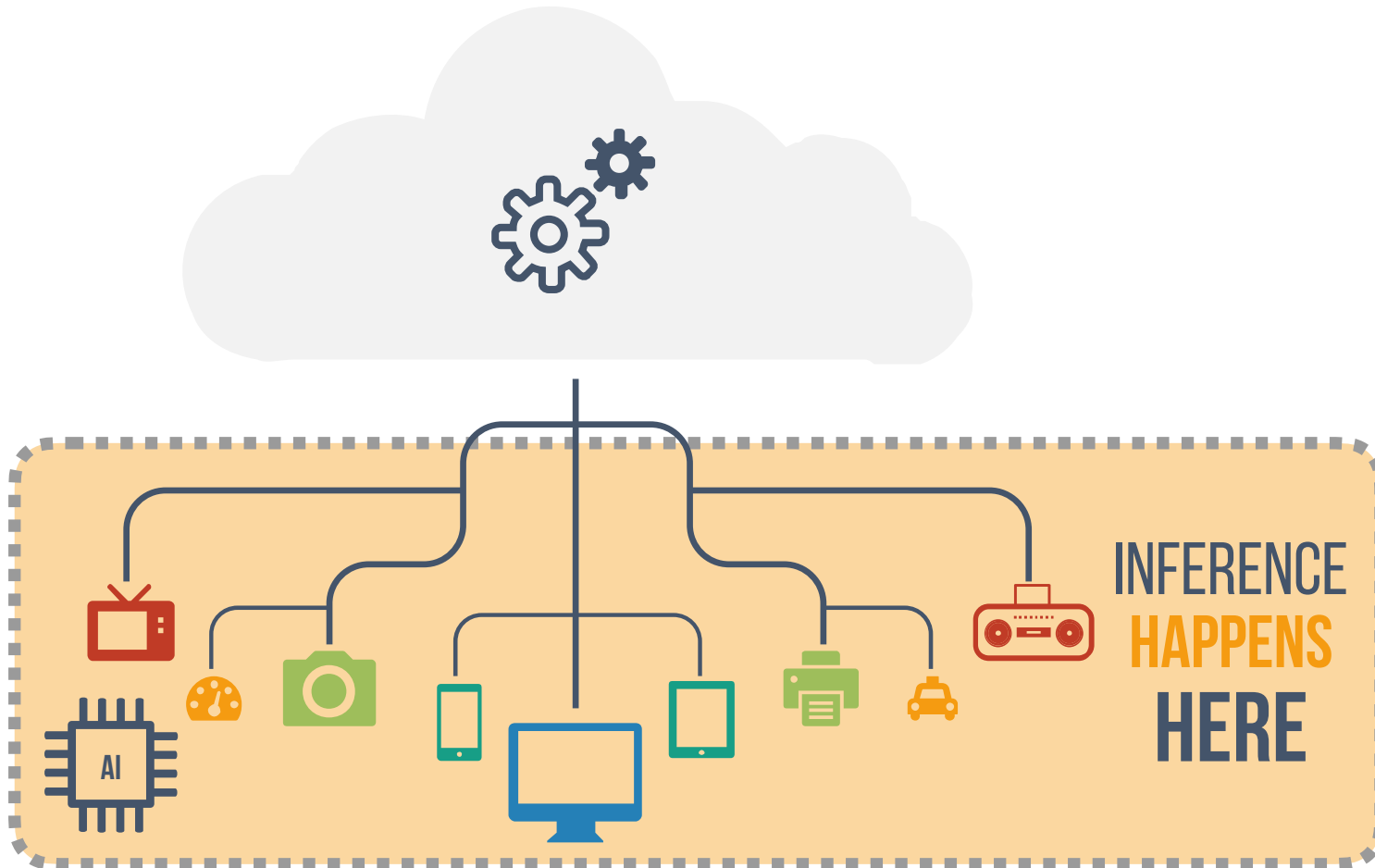


Continuous Improvements

Cloud AI services are maintained by providers who ensure that the AI models are continuously updated.

WHAT IS EDGE AI?

Introduction



1



Low Latency

Local processing significantly reduces response times and improves the performance of real-time applications.

2



Reduced Bandwidth

By processing data on the device itself, Edge AI decreases the volume of data transmitted over the network.

3



Enhanced Privacy and Security

Local data processing means sensitive information does not have to leave the device, enhancing data privacy.

4



Operational Reliability

Edge AI allows devices to operate uninterrupted, independently of the cloud or central servers.

5



Energy Efficiency

Processing data locally can be more energy-efficient than sending data to a cloud for analysis.

EDGE AI EXAMPLES

Example in different industries



Tesla Full Self Driving

By Tesla



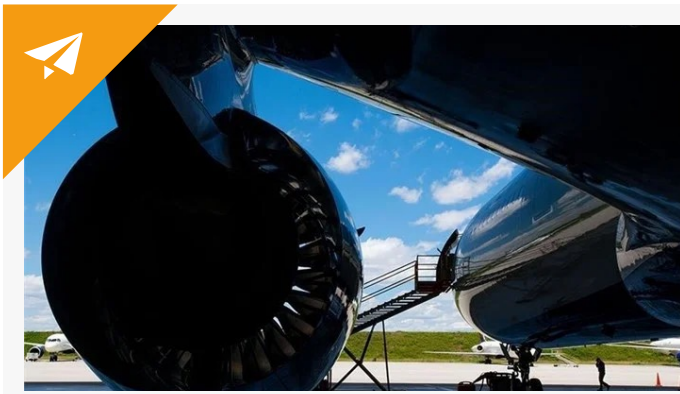
See and Spray

By John Deere



Apple Watch

By Apple



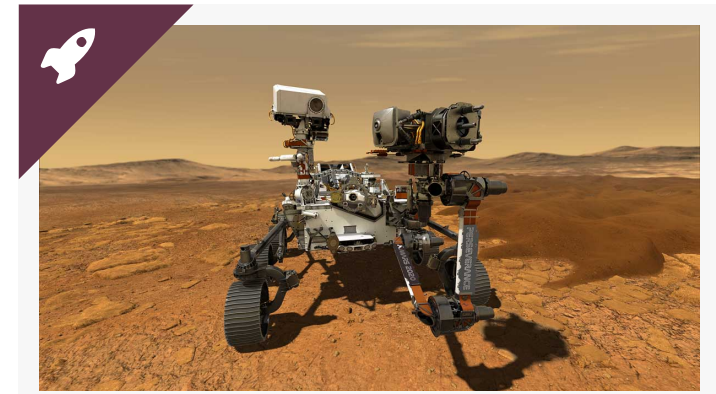
Delta Airlines Predictive Maintenance

By Delta Airlines



Jabra PanaCast 40 VBS

By Jabra

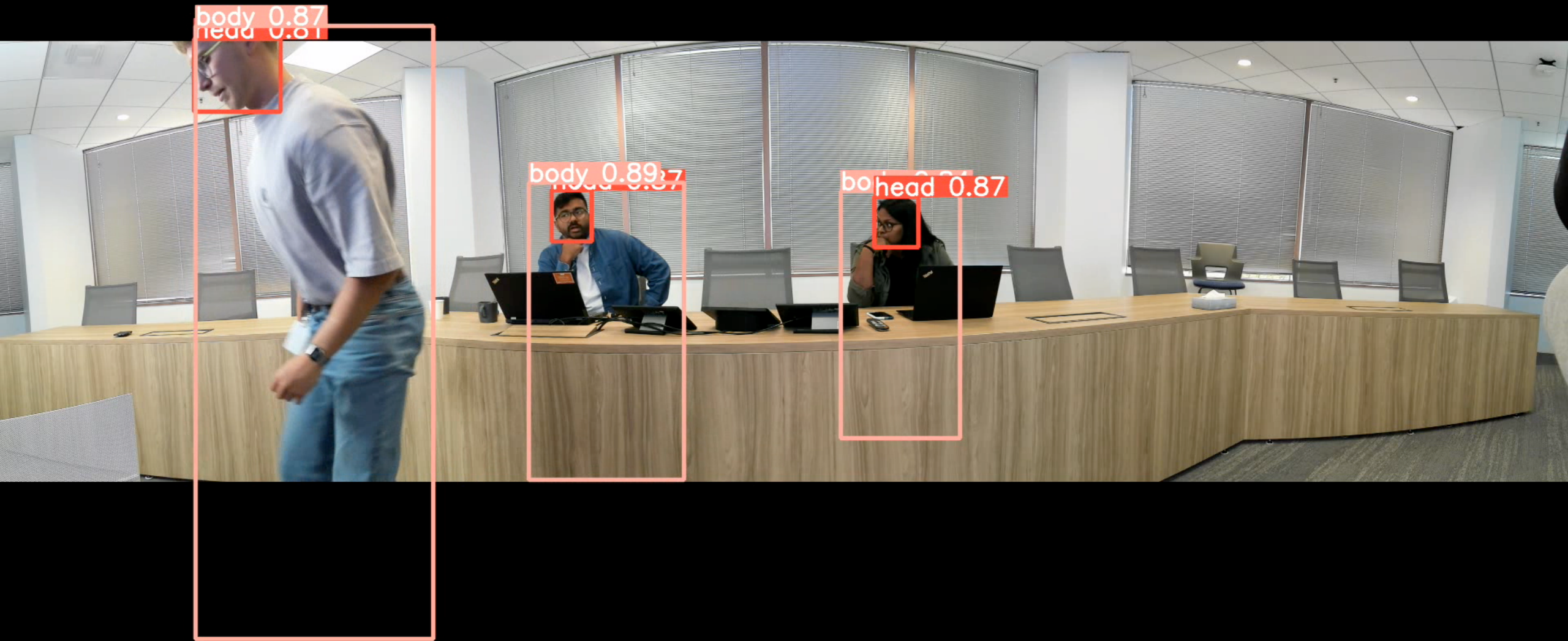


Perseverance Mars Rover

By NASA

MODEL IN ACTION

Jabra PanaCast 50 VBS



INTELLIGENT MEETING SPACES

Jabra PanaCast 50





SECURITY AND PRIVACY

Security & Privacy in Edge AI

As we integrate AI into devices at the edge, we can take robust actions to protect sensitive information and maintain user trust.



Data Encryption

Ensuring data remains encrypted during processing and storage.



Firmware Updates

Protecting devices from unauthorized access and ensuring they run trusted software.



Data Anonymization

Processing data in ways that prevent identification of individuals.



Regulatory Compliance

Meeting standards such as GDPR and AI Act (European Union) by keeping data processing local.

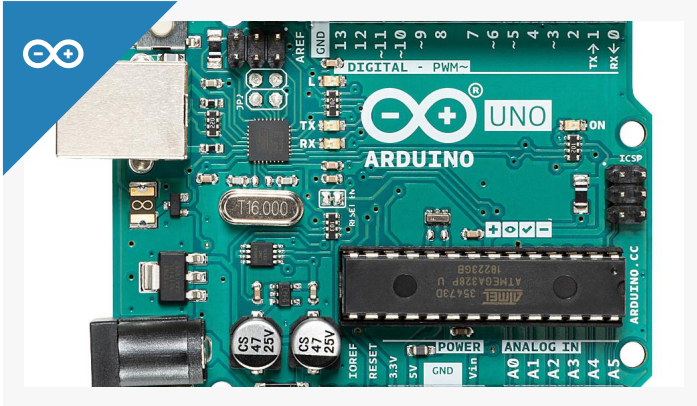
COMPONENTS OF EDGE AI

Specialized Hardware and Software



EDGE AI HARDWARE

Examples of Hardware for Edge AI



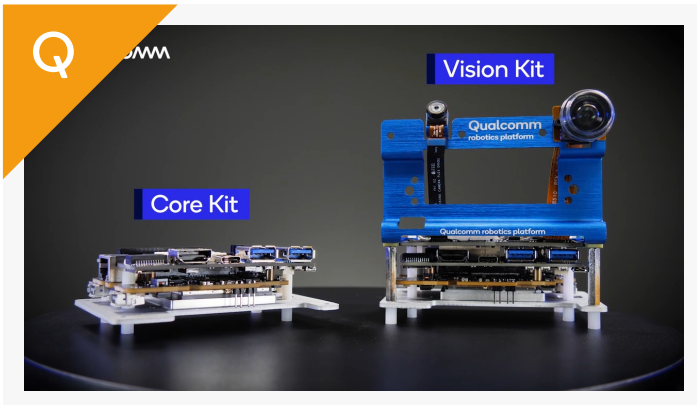
Arduino Microcontroller
By arduino.cc



Intel Neural Compute Stick 2
By Intel



Hailo 8L AI Accelerator
By Hailo



Qualcomm RB3 Development Kit
By Qualcomm



NVIDIA Jetson
By NVIDIA

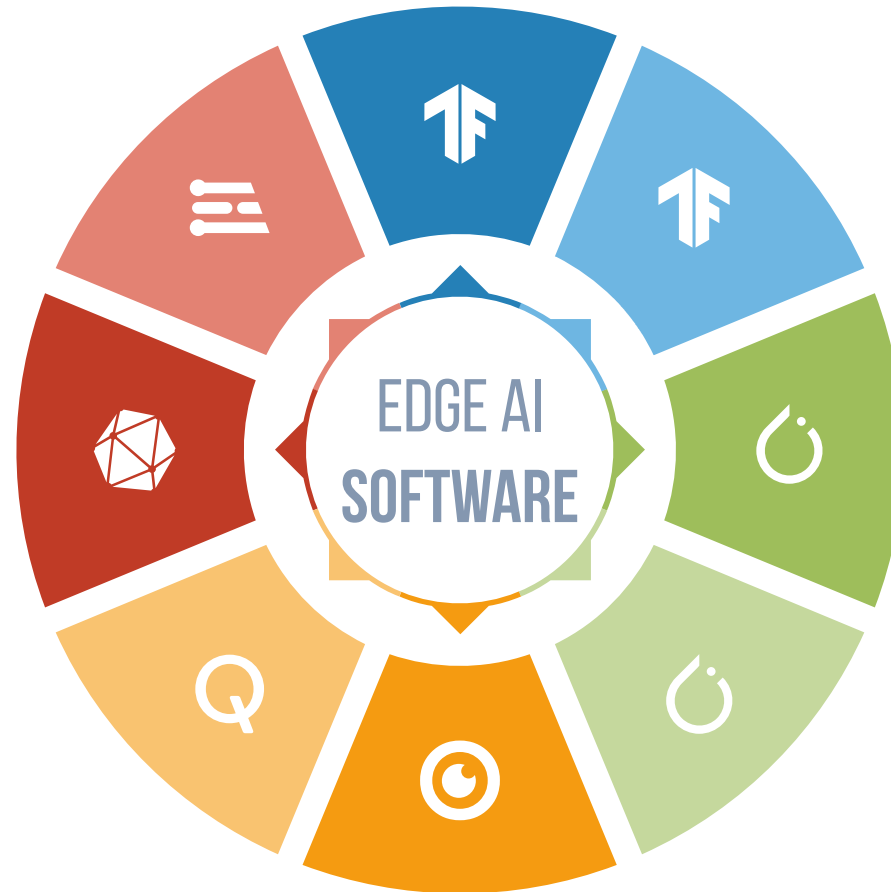


Google EdgeTPU
By Google

EDGE AI SOFTWARE

Frameworks for Edge AI

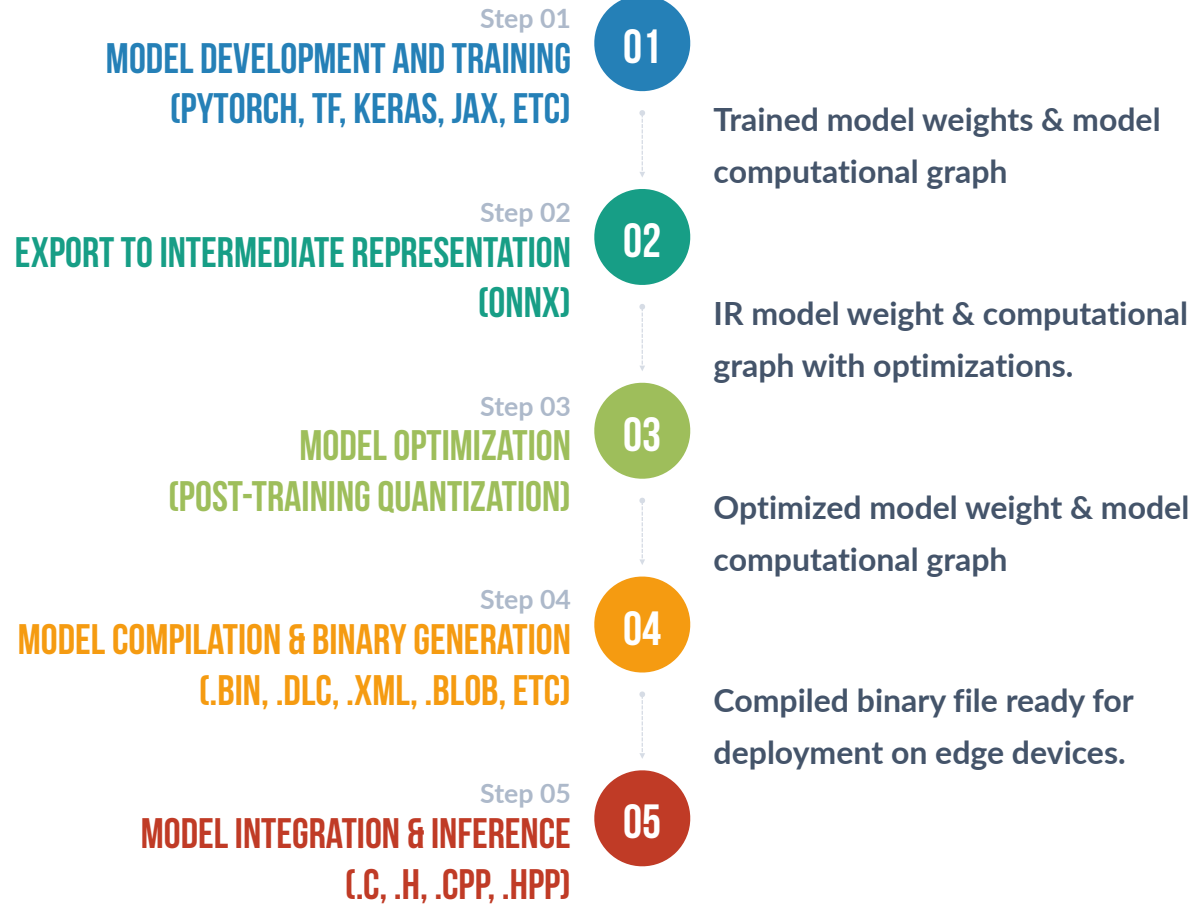
- Edge Impulse** 8
A platform for developing, optimizing, and deploying AI models to edge devices.
- ONNX Runtime** 7
Cross-platform, high-performance scoring engine for ONNX models.
- Qualcomm QAIRT** 6
It allows run DL models on Qualcomm Snapdragon mobile platforms.
- Intel OpenVINO** 5
A toolkit designed to optimize ML and DL models for Intel hardware.



- 1 LiteRT**
Lightweight version of TensorFlow optimized for mobile and edge devices.
- 2 TensorRT**
Runs ML models efficiently on any NVIDIA device.
- 3 Hailo RT**
Efficient runtime for Hailo AI chips.
- 4 PyTorch ExecuTorch**
Enables on-device inference capabilities across mobile and edge devices.

BRING IT ALL TOGETHER

Workflow for Edge AI Model Deployment



T H A N K Y U !