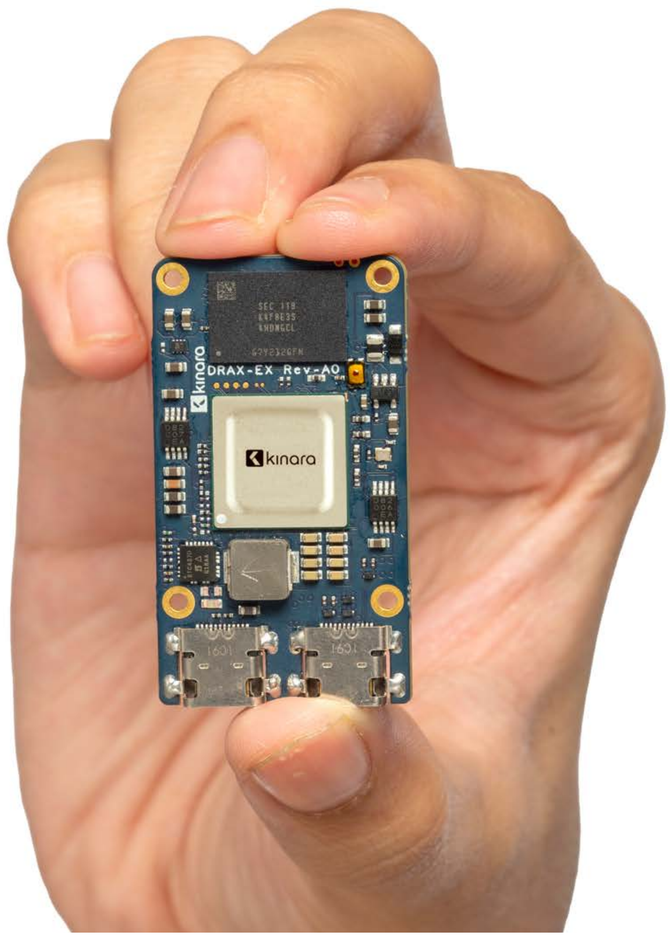
kinara
Leading Edge AI

PRODUCT BRIEF

KU-1 USB Module with Ara-1 Processor

Best-in-class AI Performance with Plug-and-play Convenience

Featuring the Kinara Ara-1 AI processor with latency optimized architecture, seamless multi-model execution and state-of-the-art model support including vision transformers



Available now as part of the Ara-1 Starter Kit

Inquire Now

When combined with the Kinara SDK, the KU-1 USB module enables rapid deployment of state-of-the-art AI models on the high-performance, low power Kinara Ara-1 processor.

Versatile Form Factor and Host System Integration

- Seamless support for multiple KU-1 modules based on system requirement
- Supports desktop PCs, laptops and Arm-based embedded systems
- Supports Linux or Windows runtime
- Supplement or completely offload the AI inference requirement of the host system

AI Acceleration for Many Applications

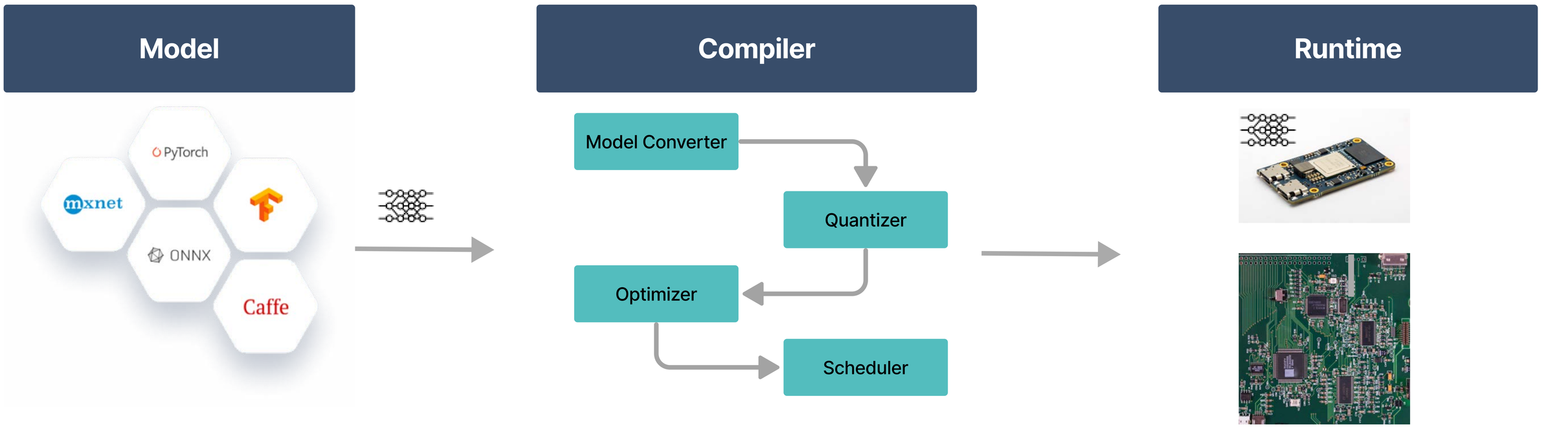
- Smart Retail
- Physical Security
- Factory Automation
- Robotics

kinara
Leading Edge AI

Key Features

AI Model Frameworks Supported	TensorFlow, PyTorch, MXNet, ONNX, Caffe2
Board Size (W x L x H)	27mm x 50mm x 5mm (board only) 48.3mm x 91.5mm x 23.8mm (with heatsink)
Interface	USB Type-C port (3.2 Gen 1) 2nd Type-C port (additional power, if needed)
Memory	1GB LPDDR4 (stores all user models)
Performance	ResNet50-v1 : 100 fps MobileNet-v1 : 546 fps
Latency	ResNet50-v1 : 10 msec MobileNet-v1 : 1.8 msec
Module Power Consumption (Typical)	4.5 W
Operating Voltage (DC Supply)	5V ±10%
Thermal Management	Passive cooling with heat sink
Operating temperature (ambient)	0°C to 40°C
Operating System Support	Linux, Windows
Certification	CE / FCC Class B
Ordering Information (part number)	ARA-U314 (board only) ARA-U312-C (with heatsink)

Software flow and SDK



Kinara's end-to-end software **seamlessly migrates** trained AI models and runs them on the hardware. Our **'software first'** approach enables users to run their own custom models without requiring any retraining. In addition, the software can **run multiple models** on the same stream without any model switching cost - resulting in low latency inferences for edge AI deployments. The software supports all the **state-of-the-art models** including vision transformers.