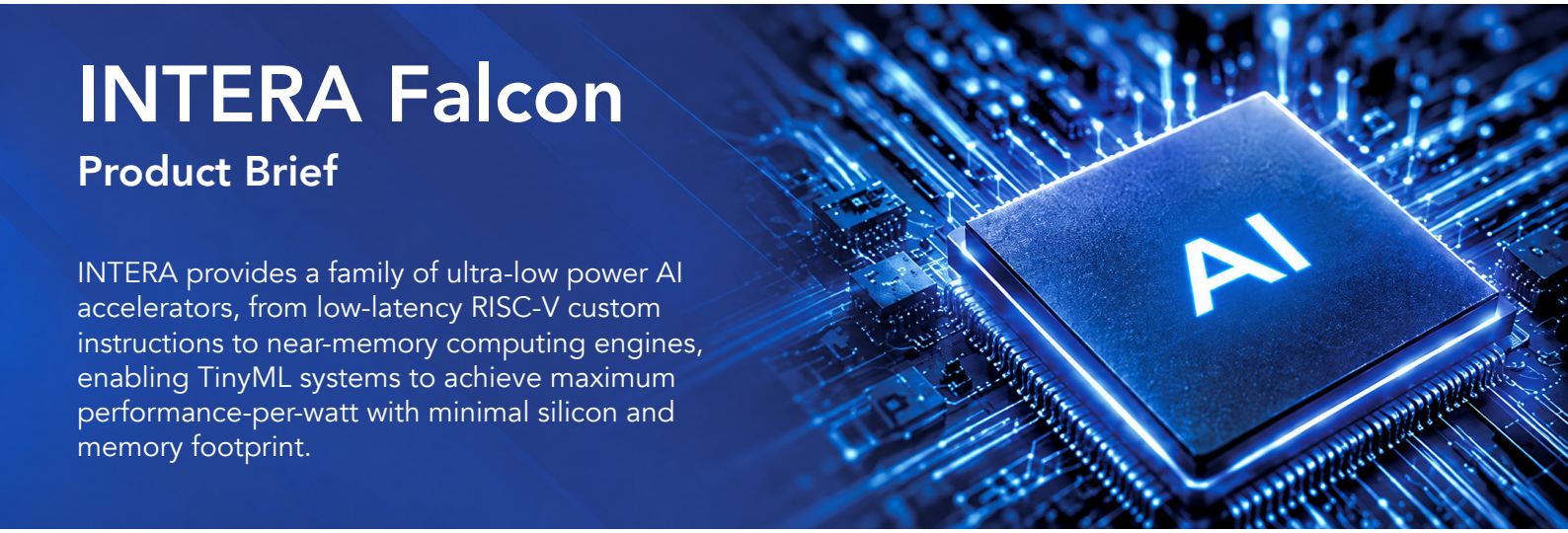
INTERA
SEAMLESS AI

INTERA Falcon

Product Brief

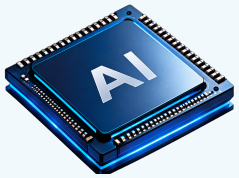


INTERA provides a family of ultra-low power AI accelerators, from low-latency RISC-V custom instructions to near-memory computing engines, enabling TinyML systems to achieve maximum performance-per-watt with minimal silicon and memory footprint.

Key features

- Ultra-low-power matrix-vector acceleratort
- Always-on availabilityt
- Smallest practical memory-mapped AI enginet
- Excellent cost/performance ratio

Key features



- Sovereignty
- Data privacy
- Low latency
- Ideal for battery-powered applications

Optimized memory mapped core

INTERA Cognit IP Selection Guide

AI Accelerator	Architecture	Operation accelerated	Compute size (*PE) and scalability
Falcon	**Monolithic, streaming processor**	**MxV**	**Single PE, 2 to 32-bit implementation**
Hawk Advance	RISC-V custom instruction	MxV / MxM[(2)]	Single PE, 2 to 32-bit implementation Entry-level version, scales to Hawk Pro
Hawk Pro	RISC-V custom instruction	MxV / MxM[(2)]	1, 2, or 4 PEs Configurable 2 to 32-bit
Eagle	Near-memory computing	MxV[(2)]	4, 16 PEs; tested up to 32 PEs Configurable 2 to 32-bits
Harrier	Monolithic, streaming processor	MxM	Scalable matrix array; 8x8 = 64 PEs; configurable as 32x32, 32x64, 64x64, ...

**PE = processing element*
[(2)] Indirect MxM multiplication by software arrangement

Typical target applications

- Health monitoring wearables
- Wake-word detection
- Acoustic event detection (AED)
- Neuromorphic vision
- Occupancy sensing
- Predictive maintenance
- Fall detection
- Radar perception
- Smart security sensors
- BMS (state of charge)

C/Sabino de Arana 14-16, 2nd - Barcelona, Spain - 08028 info@intera-group.com www.intera-group.com

INTERA
SEAMLESS AI

Software Suite for Classical AI

(CNN, LSTM, MLP/FCNN)



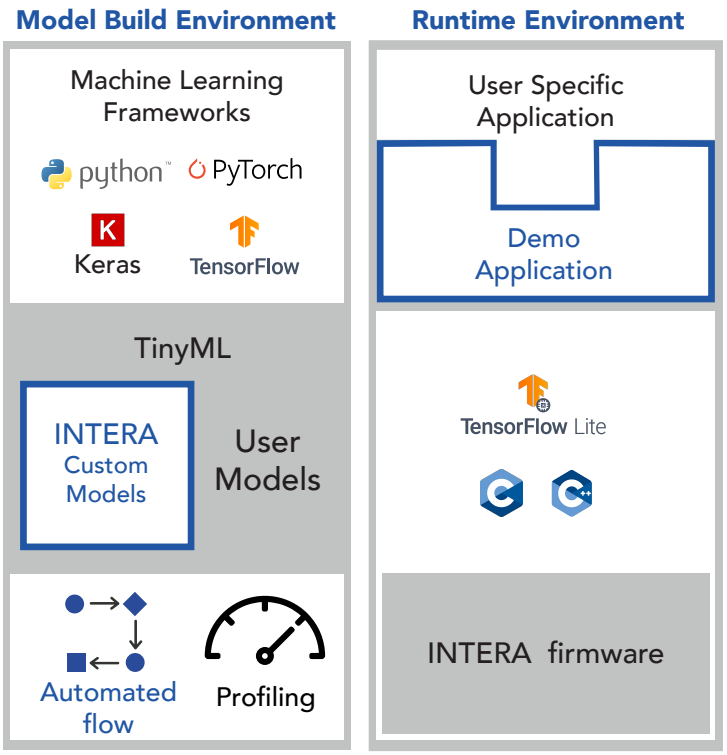
INTERA provides an open-source tool suite which seamlessly interoperates with existing deep learning development frameworks, enabling fast and efficient deployment of AI models on INTERA accelerators.

Development Workflow

Seamless integration with existing deep learning frameworks

Pre-trained benchmark models

Maximizing AI compute performance by efficient use of NN core resources



Ultra-low-power applications

Near sensor AI computing

INTERA SW component

Third party SW component

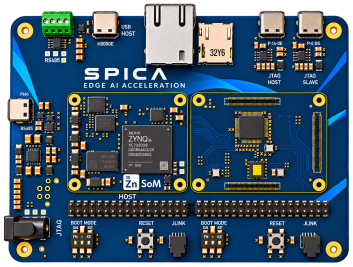
Neural Network Models & Application Support

AI Frameworks
TensorFlow, TensorFlow Lite, Keras
PyTorch & ONNX, snnTorch, NIR

Comprehensive dataflow compiler & open frameworks enable customers to easily port their models (CNN, LSTM, MLP/FCNN, SNN) or use INTERA custom models

Starter development kit

Hardware(*)



Documentation

User guide
Datasheet
Examples

Software

Plug & Play software SDK
Classical AI (CNN, LSTM, MLP/FCNN)
SNN (spiking neural networks)

C/Sabino de Arana 14-16, 2nd - Barcelona, Spain - 08028 info@intera-group.com www.intera-group.com