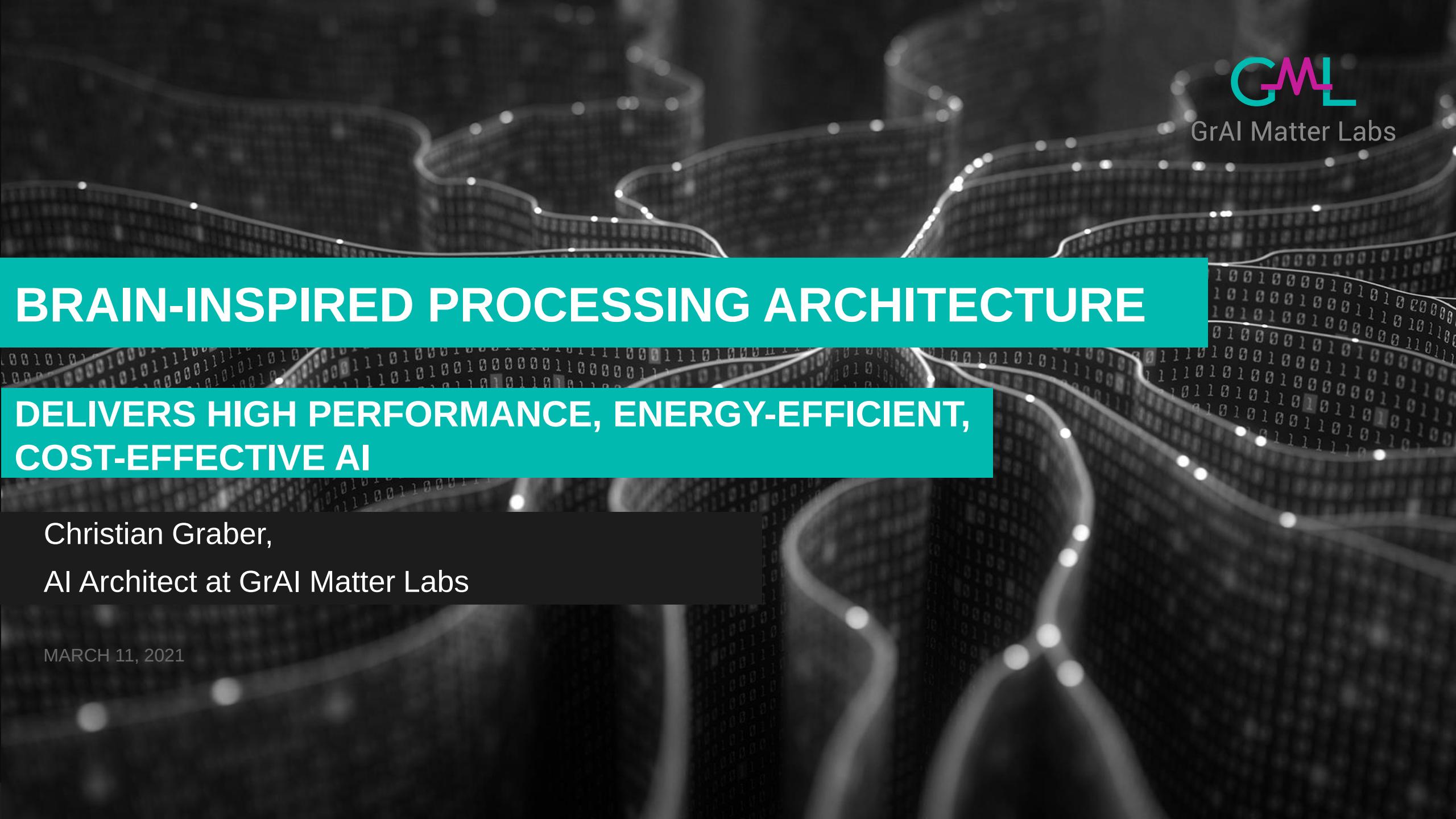


GrAI Matter Labs

BRAIN-INSPIRED PROCESSING ARCHITECTURE

DELIVERS HIGH PERFORMANCE, ENERGY-EFFICIENT, COST-EFFECTIVE AI

Christian Graber,

AI Architect at GrAI Matter Labs

MARCH 11, 2021

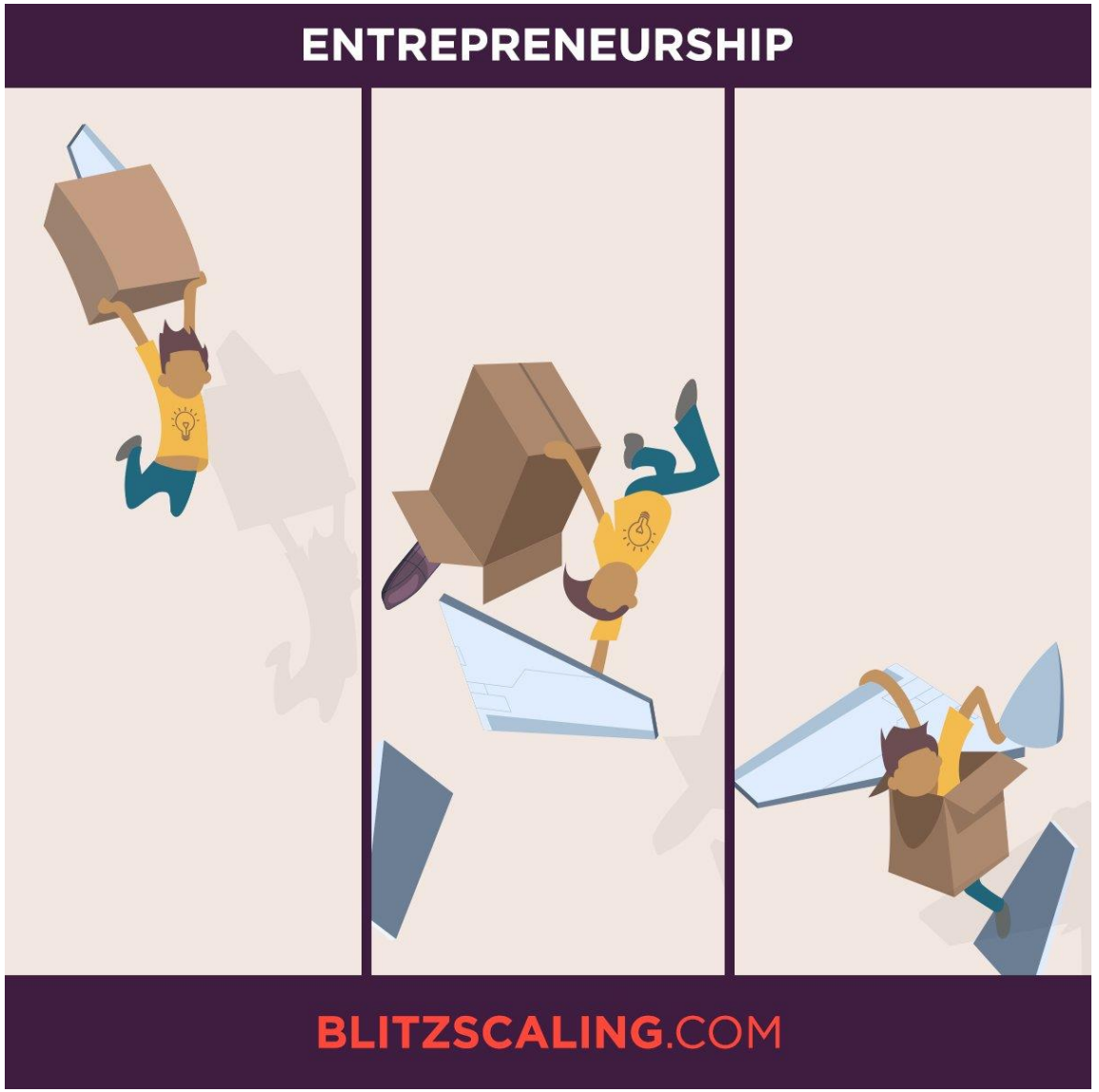
REID HOFFMAN

@REIDHOFFMAN

.

OCT 19, 2018

I'VE OFTEN SAID THAT STARTING A COMPANY IS LIKE JUMPING OFF A CLIFF AND ASSEMBLING A PLANE ON THE WAY DOWN.



@GML

We have followed this model

2019

NeuronFlow

Invented the brain-inspired technology

2020

GrAI One

Proof-of-concept for NeuronFlow technology

2021

GrAI VIP, Vision Inference Processor

Enabling our customers to create magic

Agenda

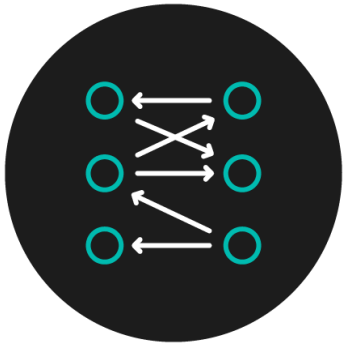
- NeuronFlow: Brain-inspired computing
- GrAI One: World's first Sparsity enabled accelerator
- Use Case: Leveraging GrAI One to enable Intelligent User Interface
- Key Learnings
- GrAI VIP: World's first Sparsity enabled SOC
- Q/A

NEURONFLOW

BRAIN-INSPIRED

COMPUTING

Neuromorphic Engineering



Dataflow Computing

NeuronFlow

- Values vs Events
- Compute in network
- Compute on demand
- Compute near memory

EXPLOITING SPARSITY IN NEURONFLOW

SPARSITY in **TIME**

Real world signals are correlated in time. NeuronFlow only computes on changing data (events) due to its dataflow nature.

SPARSITY in **SPACE**

Real world signals are sparse in all dimensions. NeuronFlow does not compute on values that are zero or close to zero.

SPARSITY in **CONNECTIVITY**

NeuronFlow only computes on graph edges with significant weight (exploits “small world” connectivity)

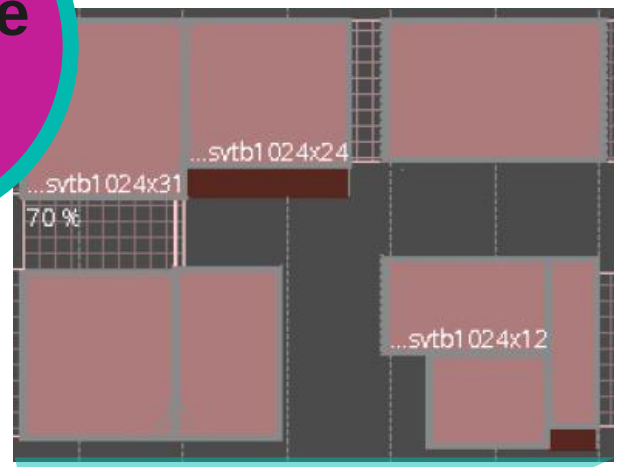
SPARSITY in **ACTIVATION**

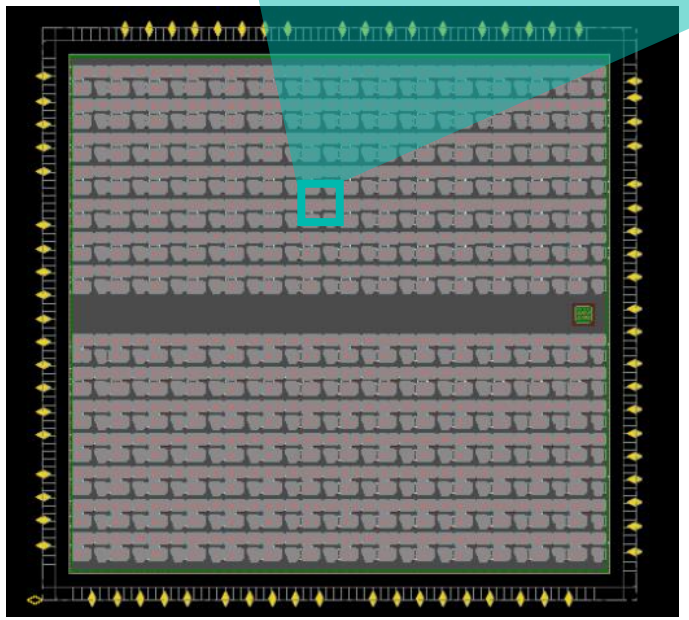
NeuronFlow allows reducing network activity by adjusting neuron thresholds.

GrAI ONE ACCELERATOR

Available Today

Neuron Core - Detailed





[14 x 14] Array of Neuron Cores

Configuration	GrAI One
# of Neuron Cores	196
# of Neurons	up to 200,704
Technology	TSMC 28 HPC+
Silicon Size	20 mm2

Flexible

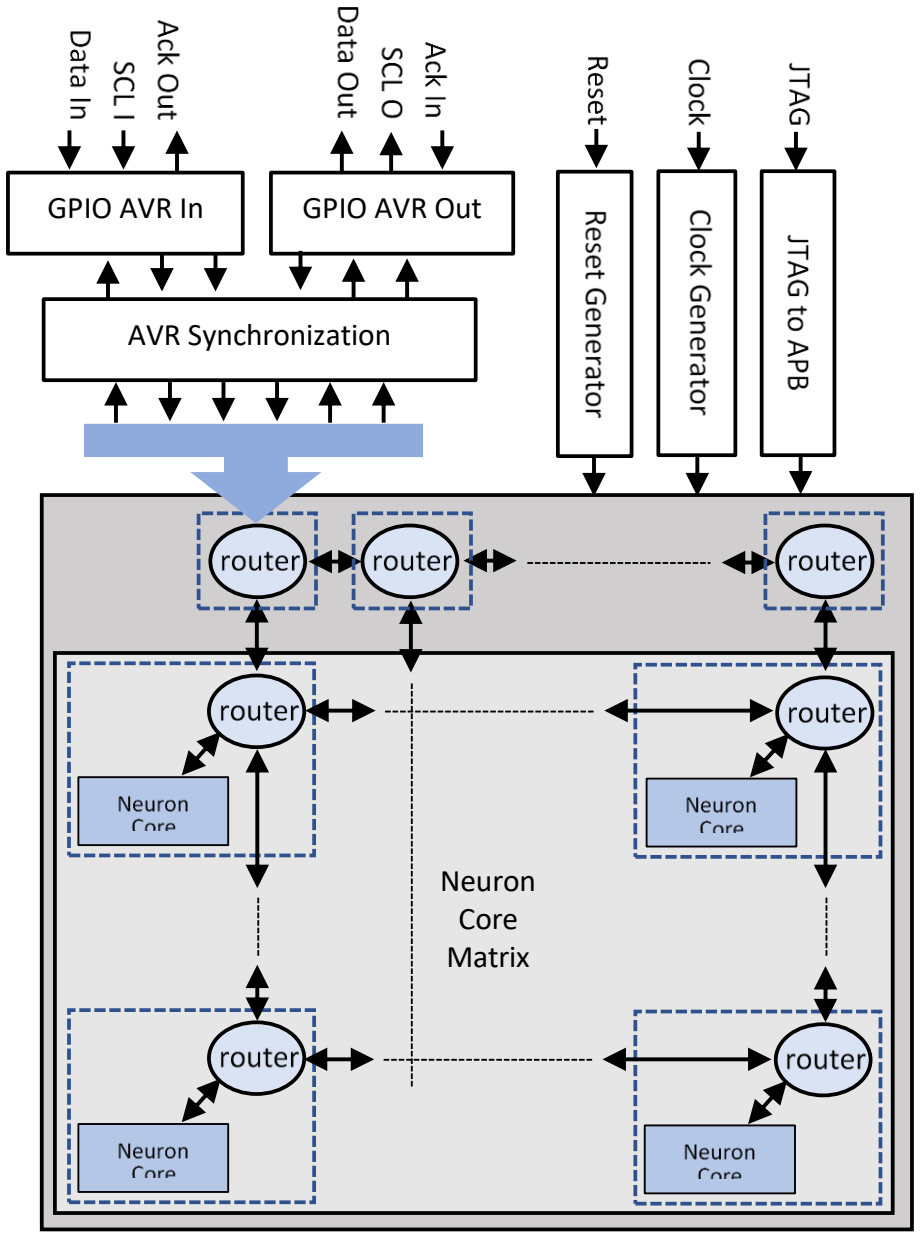
Fully programmable / SDK

C++ / Python / TensorFlow

Self-Contained

No external DRAM

GrAI One



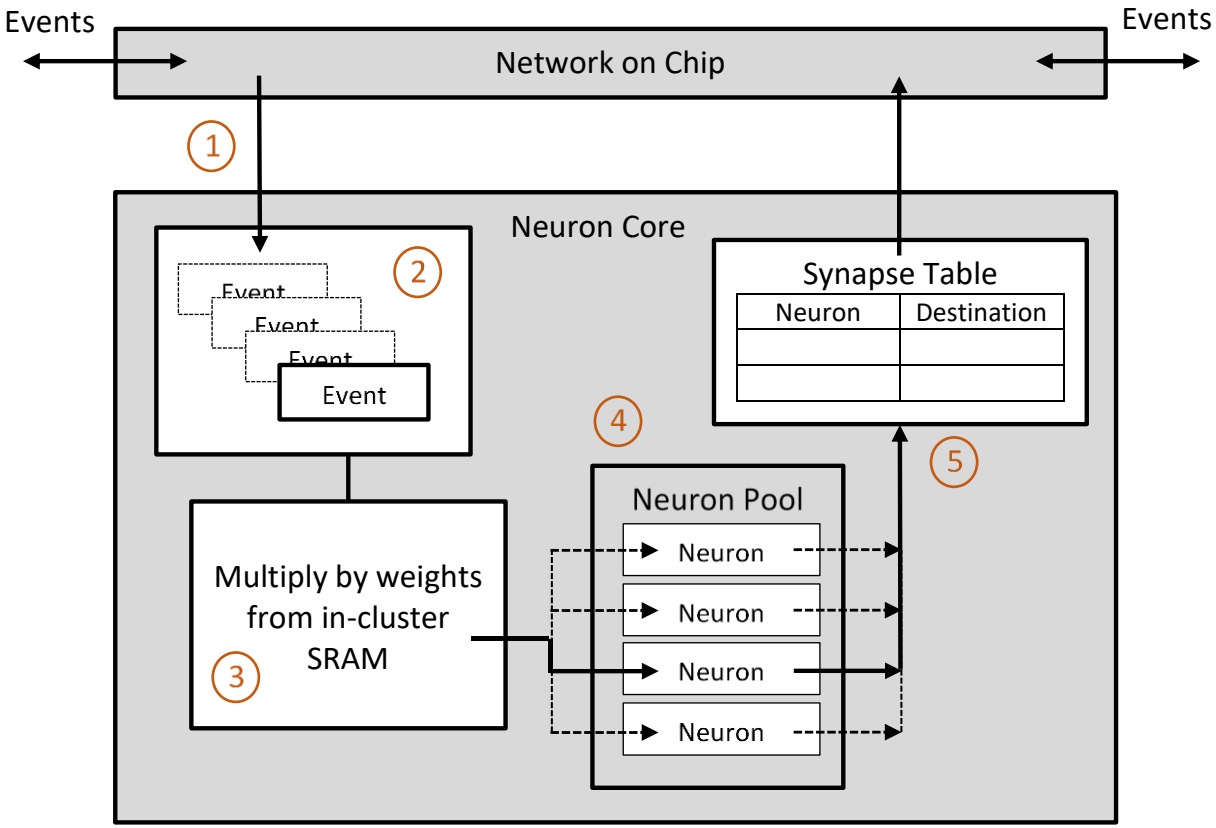
Interfaces

- Power, clock, reset, JTAG
- Data interface: Address-Value Representation, with input data routed to inputs nodes of network
- Relative addressing allows tiling

Configuration

- Fabric of cores, each of which represents 1024 neurons
- Cores connected by proprietary Network-on-Chip.

GrAI One



Processing Flow in Neuron Core

1. Events arrive via NoC based on destination address
2. Events are processed in FIFO queue
3. Weights stored in local SRAM and can be shared.
4. Data is weighted and passed to neurons
5. Neurons have state in local SRAM and perform basic neural and ALU functions
6. Events and mathematical output values sent to destinations via synapse table and NoC

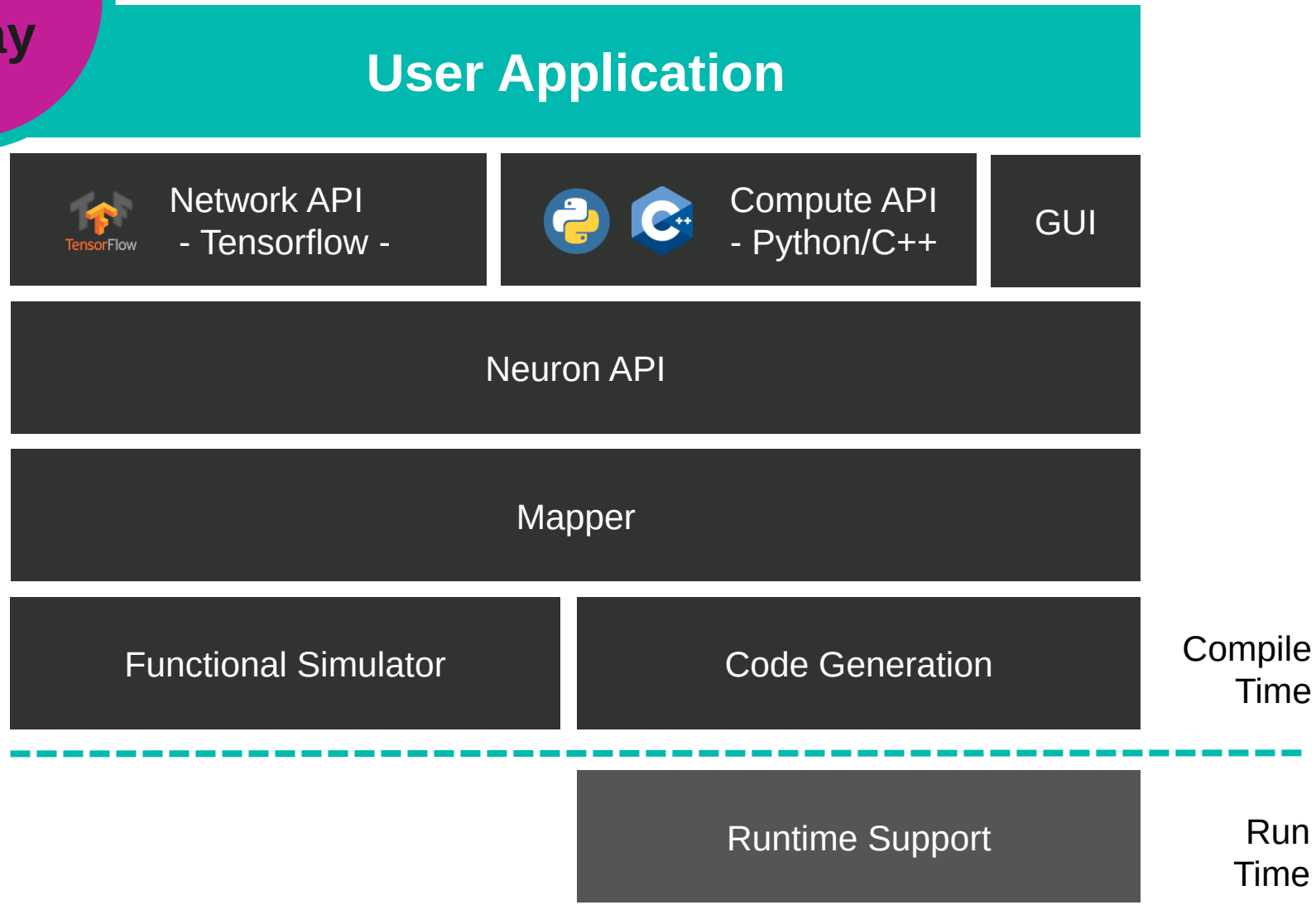
GrAIFLOW

SDK

Available Today

Key Features

Conventional Programming & Machine Learning
Direct Network Import
Integrated Simulator
Integrated Debugger
Graphical Editor



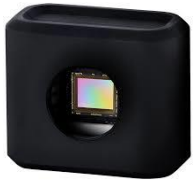
GRAI ONE HDK @ Customer

GrAI One

Audio Mic

RGB Camera

Event Camera



POC development with GrAI One

Reference Apps:
Keyword Spotting
Gestures with event camera
Face Detect with RGB camera
PilotNet with RGB camera

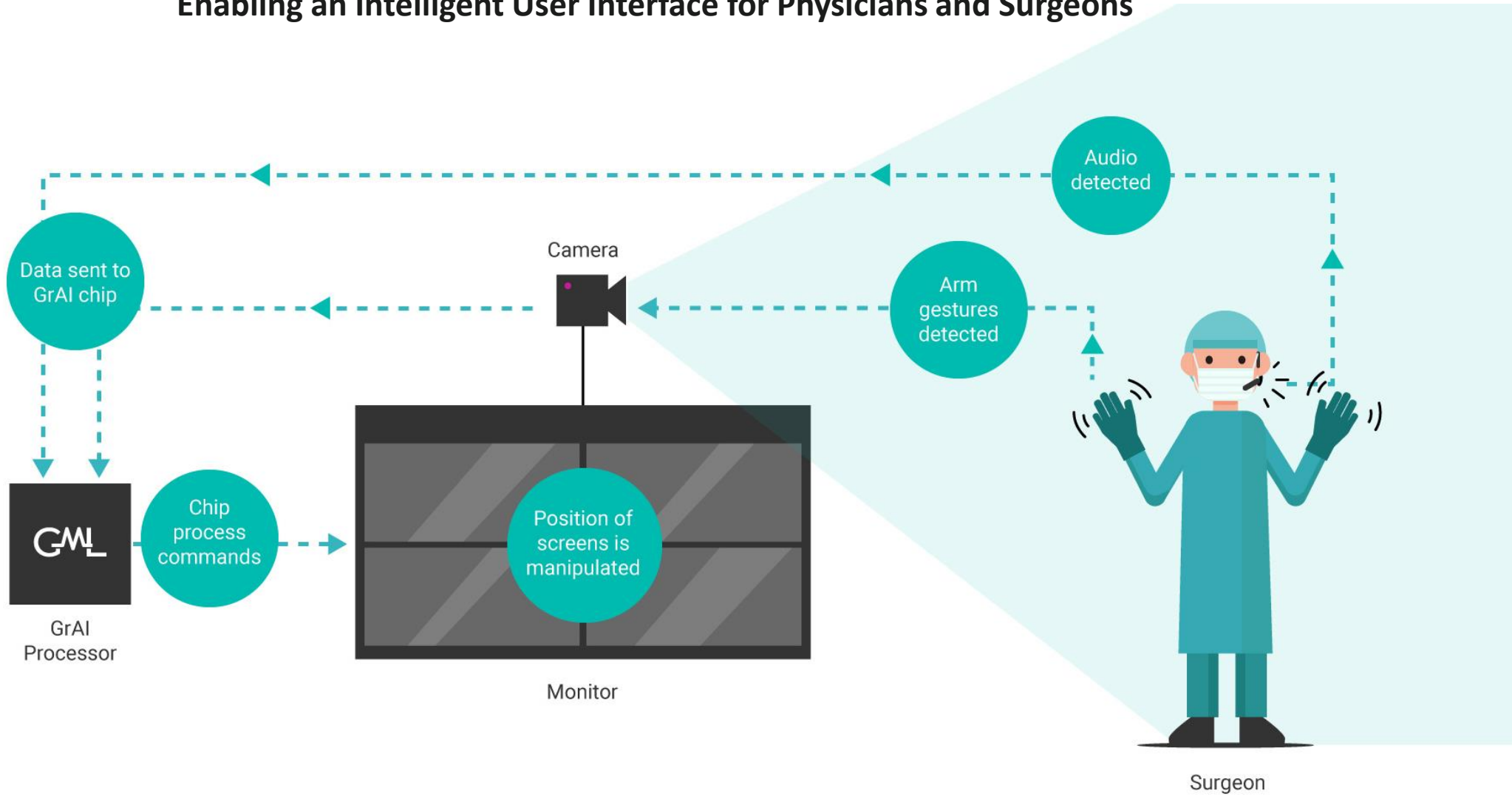
INTEGRITECH

Introduction

Enabling an Intelligent User Interface for Physicians and Surgeons

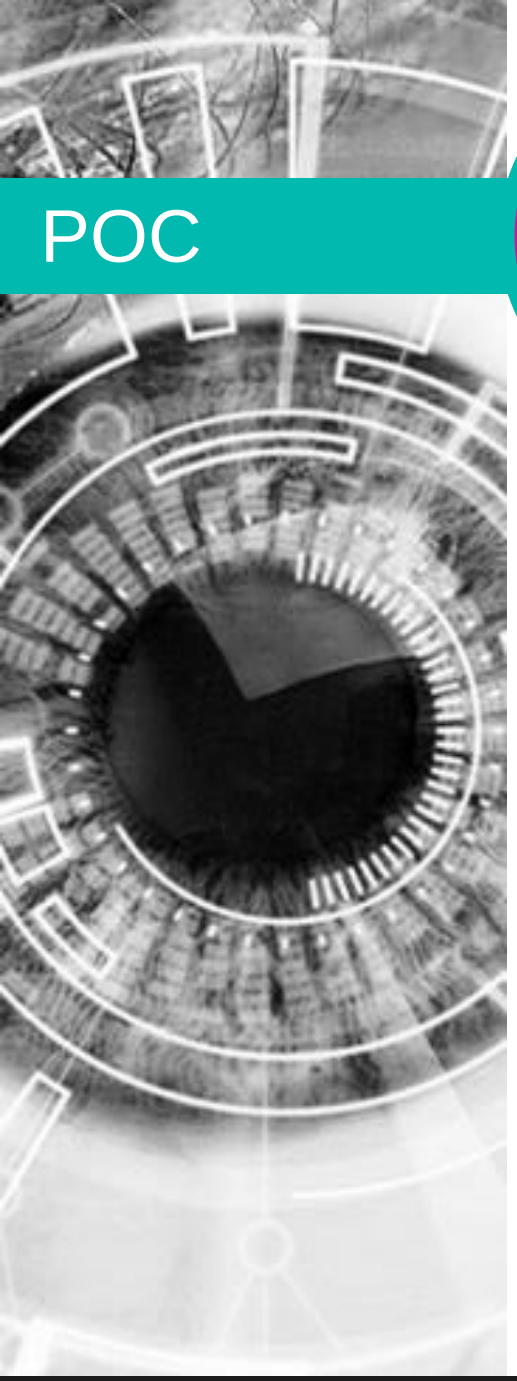
Customer

Use-Case



GrAI Matter Labs

INTEGRITECH



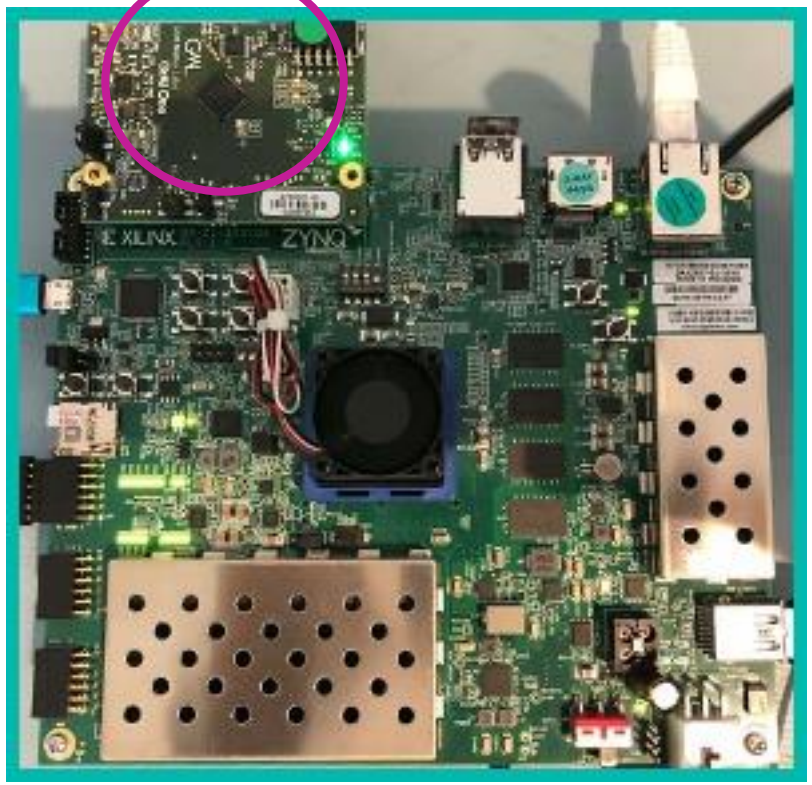
POC

Intelligent Ux

GrAI One

Audio Based Keyword Spotting

Video Based Gestures



Audio Mic

RGB Camera

POC development with GrAI One

Key Learnings

Latency of Response Time

Training for Accents and Alternatives

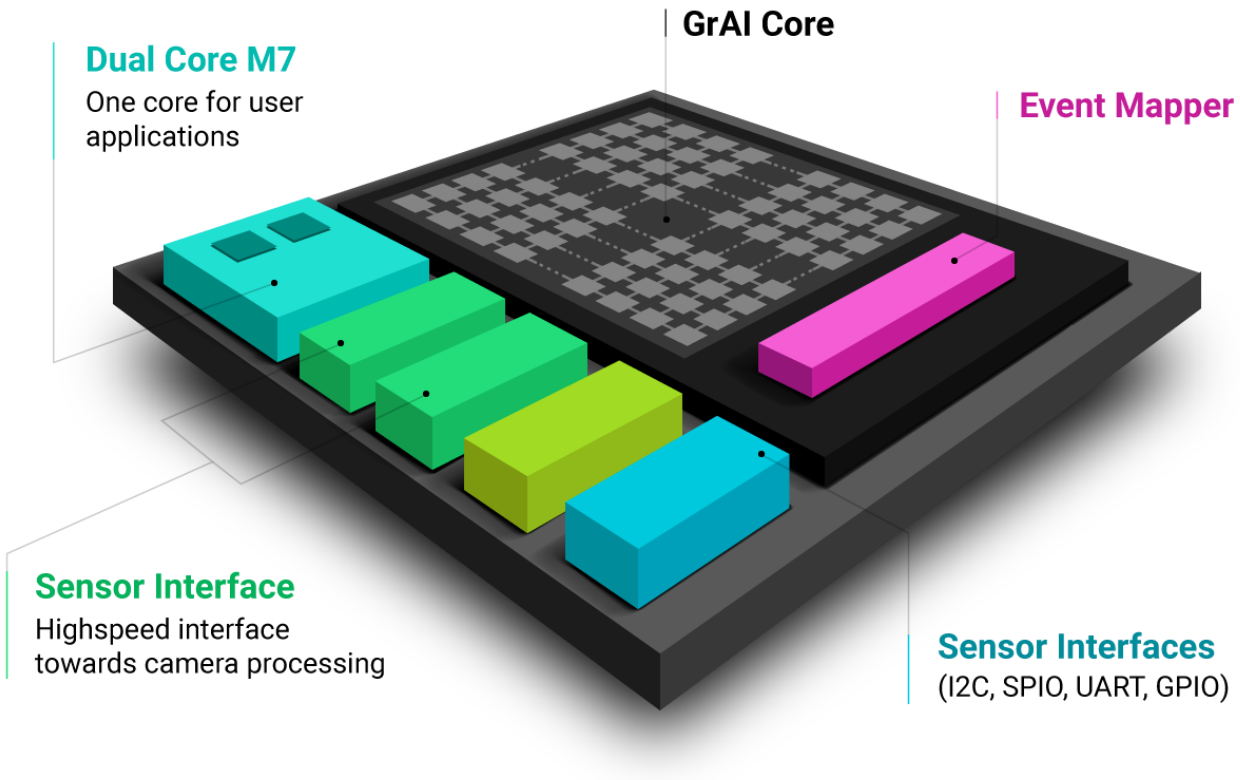
Lighting conditions for Video

Need for simplicity combining Audio and Video

Ease of Deployment

GrAI VIP

System On Chip



Fastest Edge AI Processing per Watt

Accommodate Audio and Video On Same Chip

High Performance: Latency, Size

Ease of Programmability

Interfaces to multiple sensors

Development Board Form Factor

Transform AI Compute

GLOBAL TEAM

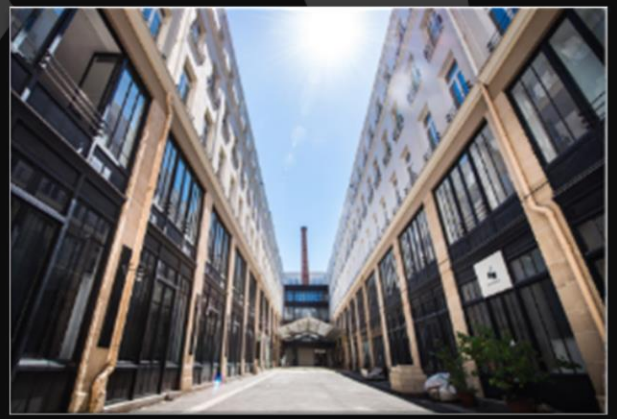
GrAI Matter Labs

Silicon Valley

- Product Marketing & Sales, CEO
- Customer Solutions

Paris

- Science Center
- System & Applications Engineering



Eindhoven

- Silicon Design Center
- SDK Engineering

AT A GLANCE

GrAI Matter Labs

Live AI	**Low Power & Ultra Low Latency** Real-time sensor data
NeuronFlow	**Processor Technology** Uniquely optimized to exploit sparsity @ edge
GrAI One	**Accelerator for Live AI** Samples and HDK currently deployed
GrAI VIP	**System on Chip for Live AI** Get started today


GrAI Matter Labs
THANK YOU

EMPOWERING PRODUCT CREATORS TO HARNESS EDGE AI AND VISION

The Edge AI and Vision Alliance (www.edge-ai-vision.com) is a partnership of 100+ leading edge AI and vision technology and services suppliers, and solutions providers

Mission: To inspire and empower engineers to design products that perceive and understand.

The Alliance provides low-cost, high-quality technical educational resources for product developers

Register for updates at www.edge-ai-vision.com

The Alliance enables edge AI and vision technology providers to grow their businesses through leads, partnerships, and insights

For membership, email us: membership@edge-ai-vision.com

edge ai + vision
ALLIANCE™

JOIN US AT THE EMBEDDED VISION SUMMIT
MAY 25-28, 2021—ONLINE

The only industry event focused on practical techniques and technologies for system and application creators

- *"Awesome! I was very inspired!"*
- *"Fantastic. Learned a lot and met great people."*
- *"Wonderful speakers and informative exhibits!"*

Embedded Vision Summit 2021 highlights:

- **Inspiring keynotes** by leading innovators
- High-quality, practical **technical, business and product talks**
- Exciting **demos**, **tutorials** and **expert bars** of the latest applications and technologies

Visit www.EmbeddedVisionSummit.com to learn more and register (use promo code EARLYBIRD21 by 4/16 to receive your 15%-off Early Bird Discount!)