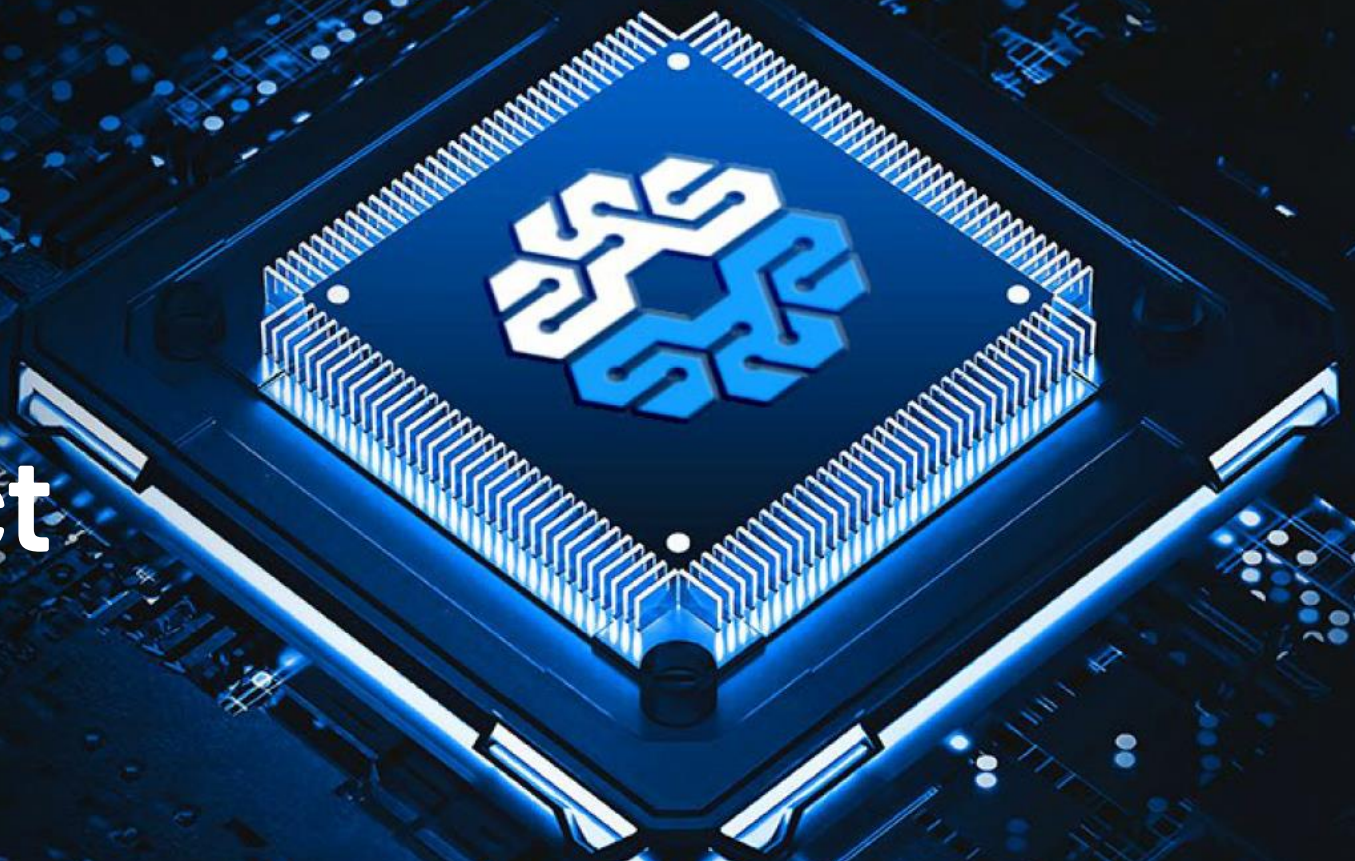
neuroSoC

NeuroSoC Project

Month 18 status



Funded by the European Union

Innovate UK

Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

This work was supported by European Union (Horizon Europe Grant Agreement n°101070634), Swiss State Secretariat for Education, Research and Innovation (SERI) under contracts number SBFI 22.00202 and 23.00205 and UK Research and Innovation (UKRI) under the UK government's Horizon Europe funding guarantee [grant number 10040829]

About NeuroSoC

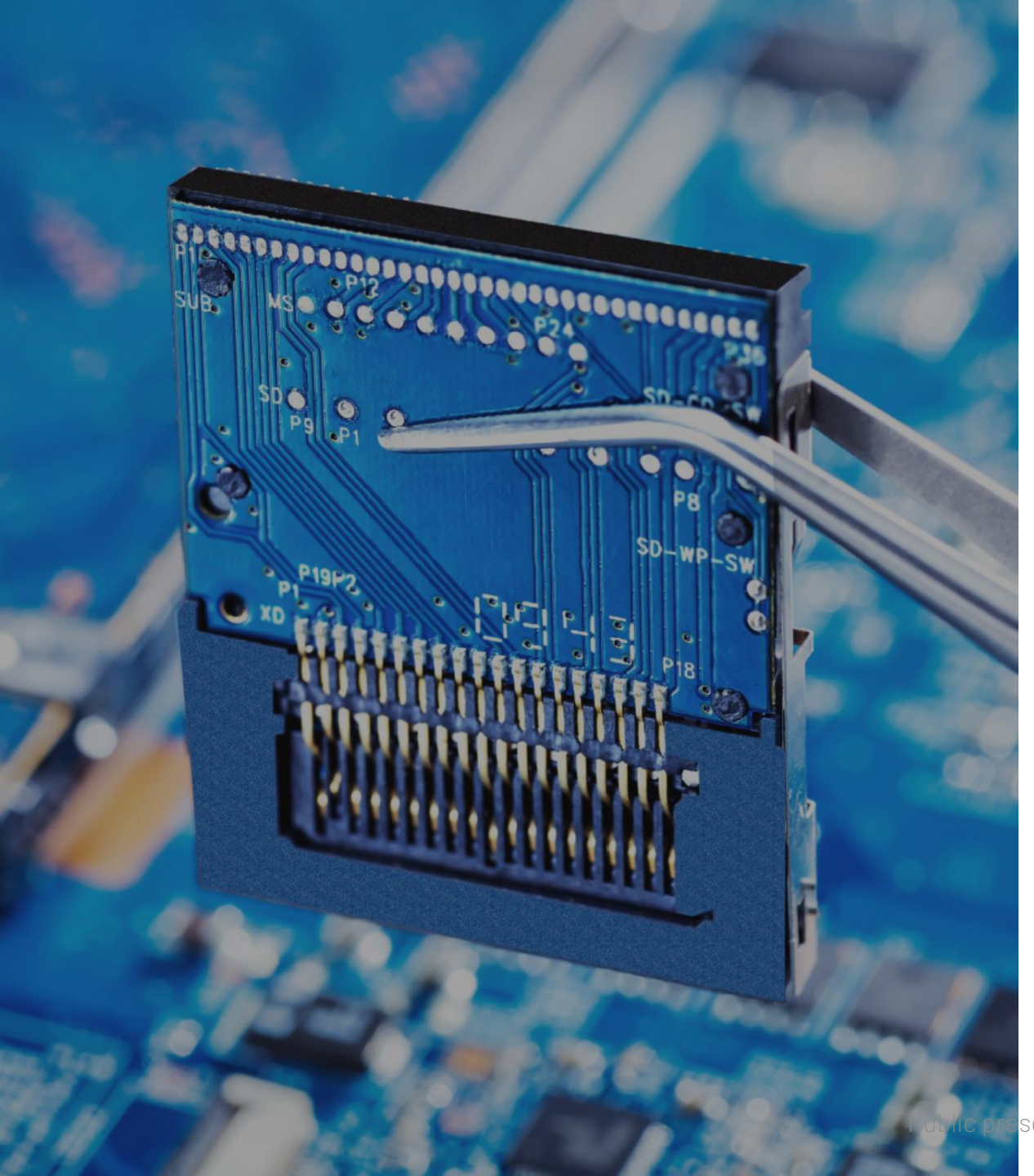
NeuroSoC stands for:

A multiprocessor System-on-Chip with In-Memory neural processing unit

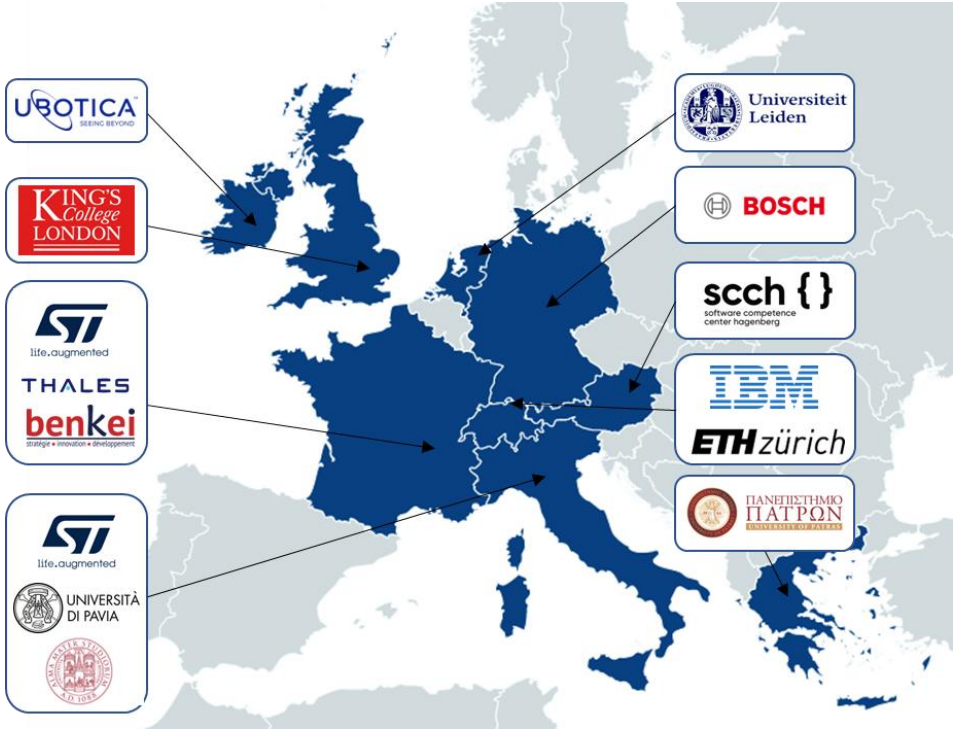
A 42-month EU/UKRI/Switzerland funded project aiming at using Phase Change Memory and FD-SOI 28 nm technologies to develop an advanced multiprocessor System-on-Chip

NeuroSoC at a glance

Call and Topic/Activity:	HORIZON-CL4-2021-DIGITAL-EMERGING-01-01 — Ultra-low-power, secure processors for edge computing (RIA)
GA number:	101070634
Type of action:	RIA (Research & Innovation Action)
Project cost:	7 952 677 EUR (only beneficiaries)
Duration:	42 months; start 1 September 2022
Website:	www.neurosoc.eu



An European strong value chain



NeuroSoC Rationale

The explosive growth of artificial intelligence

Its movement to the edge and end devices

Significant research on highly energy efficient and low-latency non-von Neumann computing paradigms such as in-memory computing (IMC)

neurSoC answer

Develop a **flexible computing system where an Analog IMC-based neural processing unit (AIMC)** is integrated into a **multi-processor functional safe and secure system-on-chip**

Relying on a solid, mature, and qualified reliable **Phase Change Memory** technology

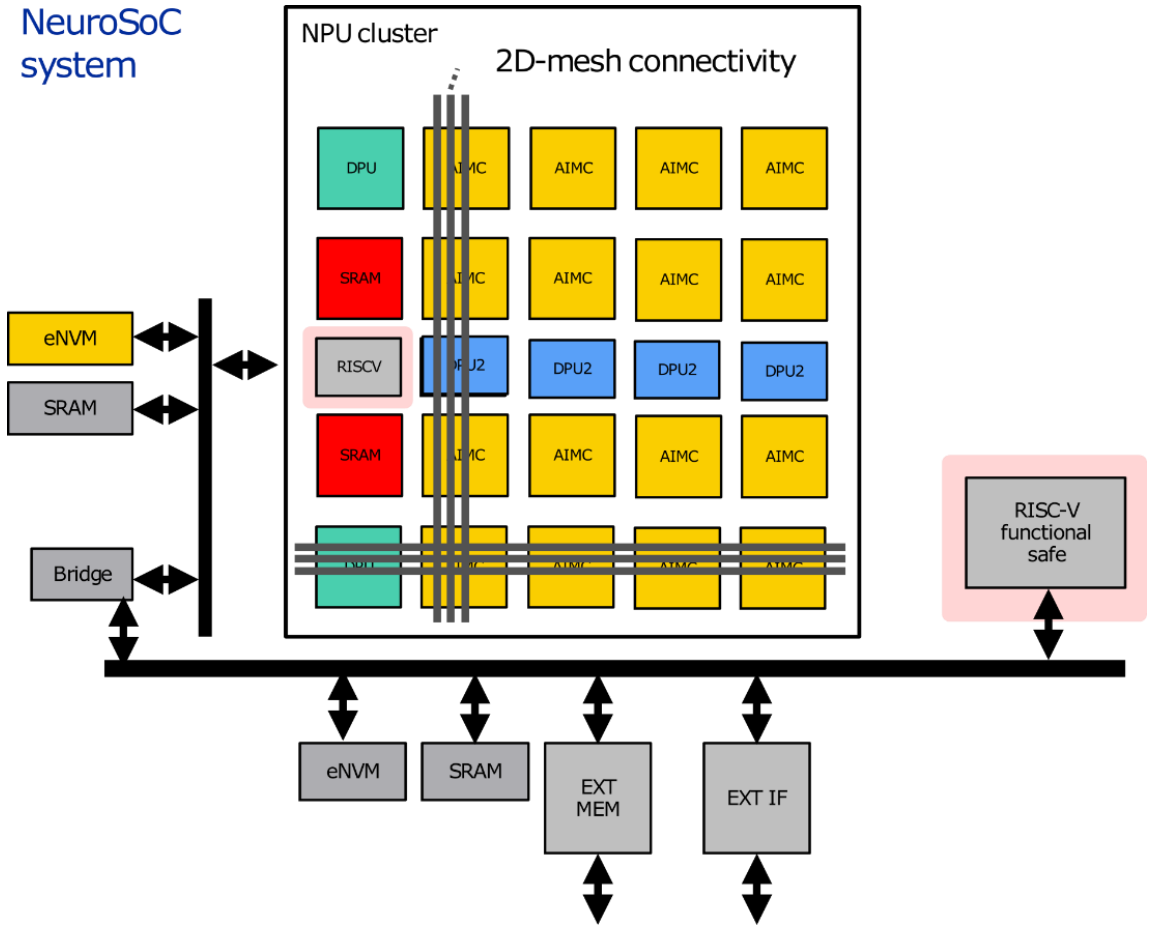
To tackle the requirements of a wide set of edge-AI applications.

Will enable the creation of an industrially proven path answering to the level of maturity need compatible with a mass volume production and cost

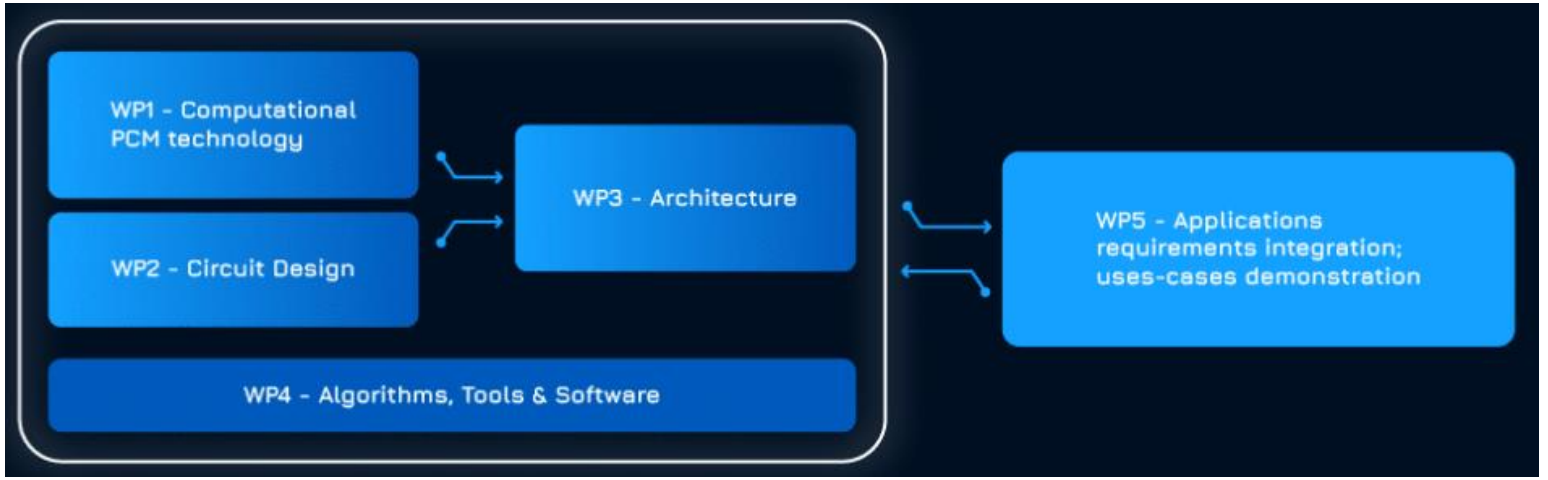
High-level Design view of the NeuroSoC System

- NeuroSoC System on Chip system level architecture comprises of:
  - Cluster of PCM analog in-memory computing tiles
  - Non-volatile memory and SRAM memory support
  - Functional safe host processor
  - Specialized digital processing units
  - RISC-V co-processor

Non-volatile memory (NVM) - Analog IMC-based neural processing unit (AIMC) - Static random-access memory (SRAM) – Data Processing Unit (DPU)



NeuroSoC organisation



- The project structure is based on **2 work packages devoted to the development of the brick of the NeuroSoC architecture, WP1 (Computational PCM technology) and WP2 (Circuit design).**
- The brick and various components are integrated in the **SoC architecture in WP3.**
- **WP4 is devoted to the development of the middle layer software** needed for the usage of the platform and the implementation of the application.
- **WP5 is devoted to the definition and implementation of the use case applications** required for the validation and demonstration of the solution in the different fields of applications.

Main achievements First 18M

Workpackage 1 - Computational PCM technology

Leader: IBM

- **Objective: characterization and modelling of a Phase Change Memory (PCM) device developed by ST-I in FD-SOI 28nm technology to serve as the building block of the In-Memory Computing (IMC) tile.**
  - Programming optimization to desired analog conductance values.
  - Accurate compact models' development to facilitate the design of the IMC tiles in WP2.
  - Statistical models' development based on array-level characterization,
  - Evaluation of the overall compute precision based on the conductance fluctuations and other non-idealities associated with these devices.

- **M18 Results**
  - Two batches of functional rheostatic cells taped-out.
  - Successful programming of the cells to various intermediate states.
  - Technology Computer Aided Design (TCAD) and compact models developed successfully to capture device characteristics.
  - Drift measurements carried out successfully from first batch cells.
  - Statistical model developed to extract rheostatic cell mean and SD of drift coefficient, read noise and Q factor.

Workpackage 2 - Circuit design

Leader: ETHZ

- **Objective: leveraging the multilevel PCM device to design an analog IMC tile.**
  - Definition of the unit cell and a suitable array structure.
  - Design of the associated digital and analog circuits.
  - Anticipate inputs from security analysis to make the resulting IMC tile more robust against side channels attacks and for improved security.

- M18 Results
  - 2 tiles were defined : a Main Tile and an Exploratory Tile
  - The Analog-to-Digital A/D Converters (ADC) for the Main Tile has been designed
  - Start of investigation on an efficient ADC for the Exploratory Tile
  - Main Tile RTL code has been released
  - AIMC tile hard block completed with CAD views
  - AIMC SRAM macro integrated on silicon in P18
  - Initial definition of the tile wrapper completed

Workpackage 3 - Architecture

Leader: ST-Italy

- **Objective: design of a modular IP implementing a Neural Processing Unit (NPU) with a hybrid digital and analog computational model.**
  - Demonstration of a design-time parametric IP supporting multi-tile implementations and scalable from ultra-low cost and power applications.
  - Integration in a Multi-Processor System on Chip (MPSoC) system level architecture.
  - Development and integration of an enhanced version of a RISC-V microprocessor core.

- M18 Results
  - IMNPU baseline architecture defined.
  - First version RTL for main components of the IMNPU nodes are already present.
  - First synthesis for DPU, LSU and RISC-V on p28 done.
  - AIMC emulation activity on FPGA has started.
  - Potential approaches for power analysis attacks have been identified.

Workpackage 4 - Algorithms, Tools & Software

Leader: ST-Italy, IBM, KCL

- **Objective: Development of software extensions, primitives, deployment tools and customized algorithms to fully exploit the NeuroSoC system-level architecture.**
  - Exploration of novel Instruction-Set Architecture (ISA) extension techniques.
  - Development of vertically integrated deployment flows targeting both IMNPU-based and software-based computation.
  - Mapping use-cases and investigating strategies to exploit the whole NeuroSoC system-level architecture.

- **M18 Results**
  - Software stack architecture defined
  - Defined RISC-V cores role and ISA extensions
  - Inference pipeline for accuracy simulations
  - Regularisation techniques for accuracy improvement
  - Evaluation of various neural networks

Workpackage 5 - Applications requirements, integration, and use- cases demonstrations
Leader: UBOTICA

- **Objective: investigate edge-applications where NeuroSoC can offer a compelling advantage.**
  - Selection and qualification of applications.
  - Benchmarking of SoA and emerging solutions.
  - Proposition of an evaluation framework.
  - Assessment of performances vs requirements.

- M18 Results
  - Developed emulator platform
  - Initial benchmarking completed

Public deliverables and publications

Public deliverables

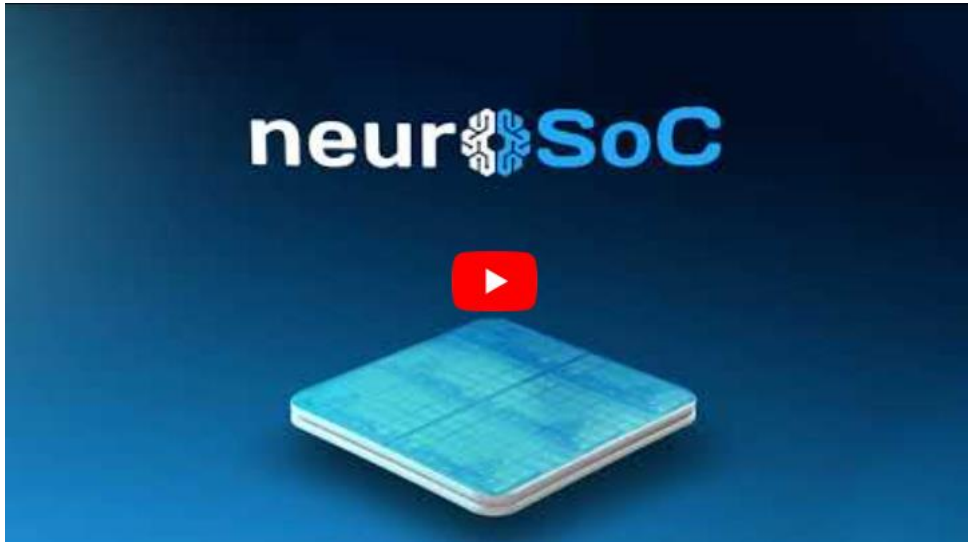
- RISC-V co-processor unit for IMNPU specifications, functional safe RISC-V host core specifications
- Tool-flow for performance evaluation
- Competitors, benchmarks and KPIs
- Report on AI-optimised ISA extensions

Publications

- A Mixed-Precision Matrix-Matrix Operation Engine for Flexible and Energy-Efficient On-Chip Linear Algebra and TinyML Training acceleration
- ITA: An Energy-Efficient Attention and Softmax Accelerator for Quantized Transformers
- Exploiting the State Dependency of Conductance Variations in Memristive Devices for Accurate In-Memory Computing
- Designing Circuits for AiMC Based on Non-Volatile Memories: a Tutorial Brief on Trade-offs and Strategies for ADCs and DACs Co-design

Follow us on our website and social media

NeuroSoC pedagogical video



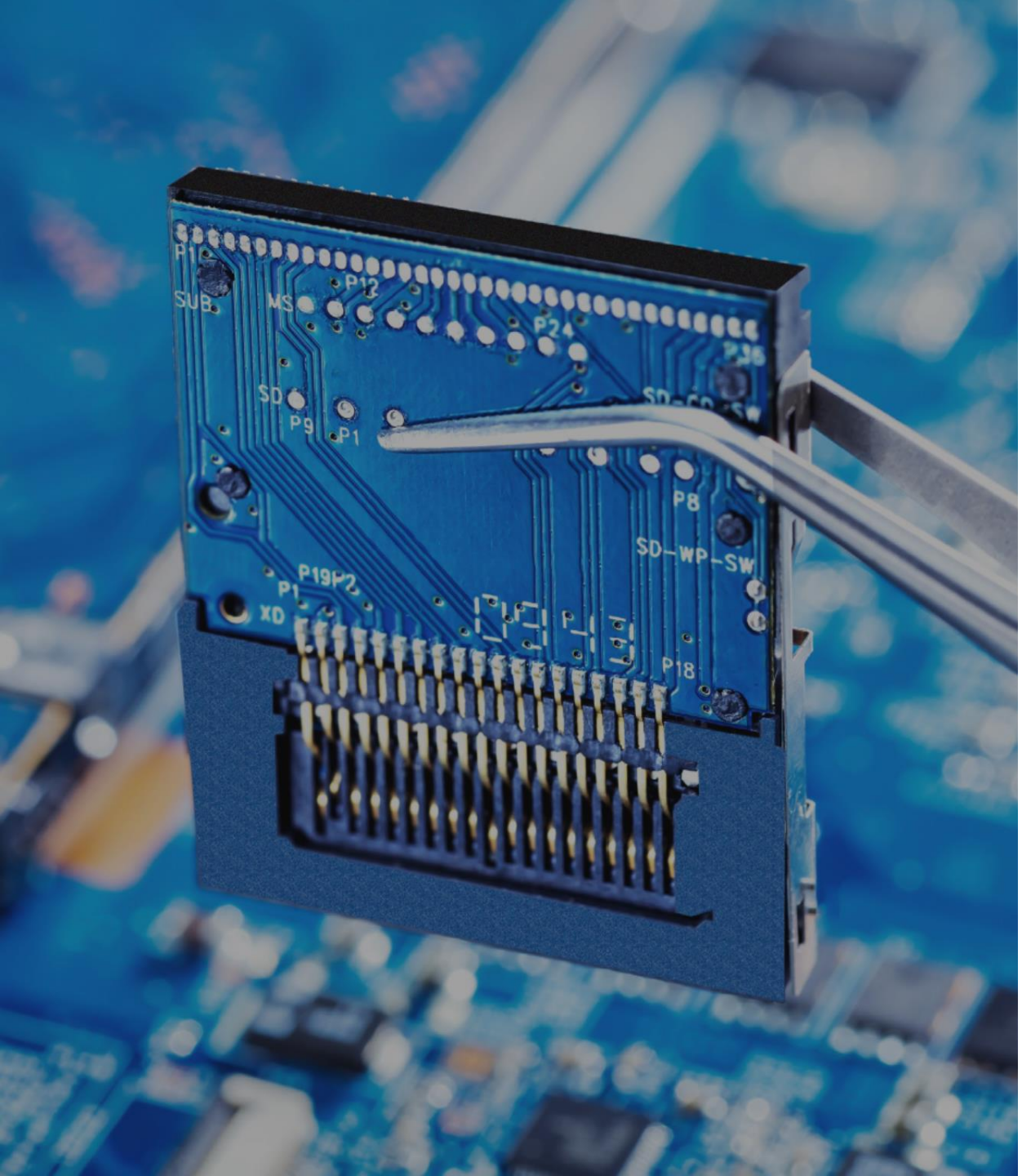
https://youtu.be/UkPTrqwCrAk

neurSoC

https://neurosoc.eu/

LinkedIn

https://www.linkedin.com/company/neurosoc/



Contact details

STMicroelectronics

- M. Giulio Urlini, Project Coordinator
- giulio[dot]urlini[at]stmicroelectronics[dot]com

Benkei

- Mrs Fabienne Brutin, Project administrative manager,
- fabienne[at]benkei[dot]eu

neur SoC

Acknowledgments

- The NeuroSoC project is supported by European Union (Horizon Europe Grant Agreement n°101070634), Swiss State Secretariat for Education, Research and Innovation (SERI) under contracts number SBFI 22.00202 and 23.00205 and UK Research and Innovation (UKRI) under the UK government's Horizon Europe funding guarantee [grant number 10040829]