# SYNAPTICS & GOOGLE RESEARCH: ADDRESSING EDGE AI FRAGMENTATION CHALLENGES

*Paul Schell, Senior Analyst*
*Malik Saadi, Content Manager*

## CONTENTS



# EDGE AI BACKDROP: CONTEXT & ISSUES WITH TODAY'S EDGE AI OFFERINGS

## OVERVIEW

The edge is a significant area of disruption as Artificial Intelligence (AI) models are shrinking in size and can deliver richer experiences without relying on cloud compute. This is a significant area of disruption and will deliver the next wave of AI innovation. Nonetheless, outside of Personal and Work Devices (PWD) such as smartphones and laptops, there are several issues affecting the proliferation of AI at the edge in form factors powered by batteries, which includes Internet of Things (IoT) devices, smart home, industrial automation, and wearables that interact with other devices. These factors increase the risk for developers and hardware vendors by creating frictions between hardware platform transitions, which locks developers into single vendors, and dilutes the Return on Investment (ROI). Some of the overarching issues in the edge AI market include:

- **The market is notably more fragmented** than in larger form factors such as PWD and the cloud, with a diversity of hardware and systems architectures, and therefore, software developer environments.

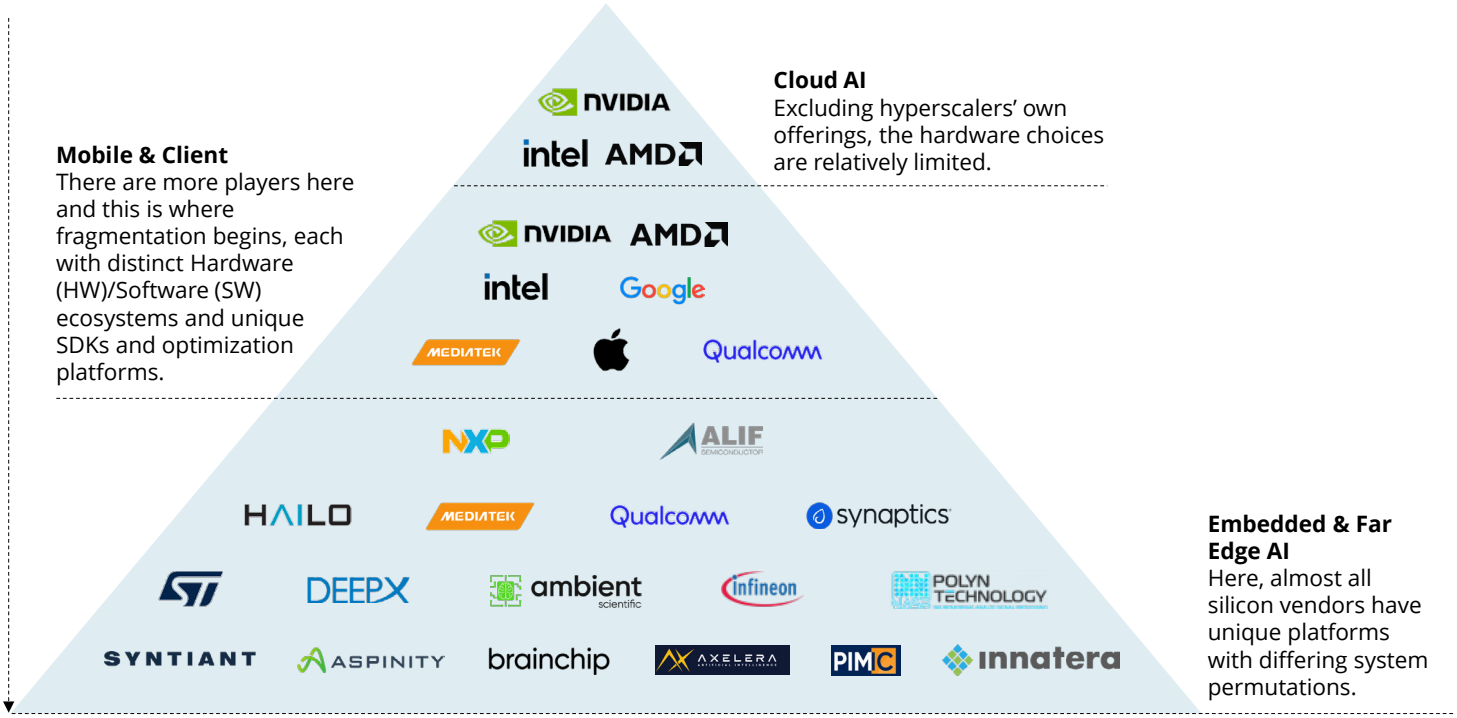
- **Security concerns about private, user-specific data—including biometrics**—decrease demand particularly from consumers for agentic use cases leveraging multiple data sources for highly personalized experiences, for example.
- **The emergence of transformers and Convolutional Neural Networks (CNNs) at the edge** require tight hardware and software roadmap alignment to eke out optimizations and enable value-add applications on restricted power budgets (as occurs in the cloud, but in a more fragmented hardware landscape).

## THE ISSUES SURROUNDING FRAGMENTATION

This fragmentation is demonstrated by the large number of hardware vendors operating in this space, from incumbent, embedded System-on-Chip (SoC) players, including Synaptics, ST Microelectronics, Renesas, Infineon, Qualcomm, Texas Instruments, and NXP Semiconductors, with a legacy in embedded systems that include connectivity and other critical parts of the module, to the startup ecosystem, which has seen large growth, but also consolidation over the past 5 years. While many of the startups (e.g., POLYN, Ambient Scientific, BrainChip) target always-on use cases running on Application-Specific Integrated Circuits (ASICs), many incumbent vendors have partnered to leverage the Intellectual Property (IP) for their own designs. Although there have been a number of commercial successes in this domain, the economics of silicon development (and associated proprietary software tooling) are a restriction on progress, given the multi-million-dollar budgets needed to refresh silicon for the next generation of transformer-based Large Language Models (LLMs) and other models.

*Figure 1: Fragmentation of AI Silicon from Cloud to Edge*

*(Source: ABI Research)*



Furthermore, vendors operate within largely closed Machine Learning Operations (MLOps) software ecosystems, restricting investments by developers to each individual platform. While multi-vendor software tools exist, this area has seen significant consolidation, exemplified by Qualcomm and Infineon's acquisitions of Edge Impulse and Imagimob, respectively. Overall, this holds back the broader adoption of AI applications that would drive greenfield demand and refresh cycles of existing IoT deployments, as Original Equipment Manufacturers (OEMs) experience frictions transitioning between silicon vendors, and model developers hold back investment due to uncertainties around uptake and hardware compatibility.

## RISC-V AT THE AI EDGE

Another notable trend at the edge is the emergence of RISC-V processors, which extends to AI workloads. The appeal to developers stems from the flexibility of the architecture to address specific workloads, as well as its scalability. The open Instruction Set Architecture (ISA) is unified under RISC-V International, which nurtures the ecosystem, and the RISE software project, fostering collaboration between technology leaders, including Google. However, the risk of fragmentation also exists in the open-source RISC-V space, as custom extensions to the ISA for vector and matrix calculations, for example, which are essential for AI workloads, mean that AI/Machine Learning (ML) software optimized for one vendor's specific extensions may be incompatible or perform poorly on another's hardware. This potentially forces developers to rely on vendor-specific compilers and libraries, complicating portability and hindering the development of a unified software stack. While the vector extension has been ratified under the common RISC-V ISA, there are several vendor-specific matrix extensions, including from SiFive and Andes, while efforts are still under way to ratify the first standard under the open ISA.

## CONCLUSION

Fragmentation is rarely addressable by single vendors and requires the formation of unified ecosystems under an open approach where diverse interests can be reconciled. A best practice approach can be observed in the Android initiative for smartphones—eventually under the direction of Google—which created a more updatable and uniform ecosystem, reducing the impact of both hardware and software variations. This was broadly achieved through a strategic modularization of the Operating System (OS). Thus, application developers can avoid the issues arising from hardware idiosyncrasies and deploy across a range of devices and Android versions, reducing the development and testing burden. This allows for more innovation and richer end-user experiences.

# INTRODUCING SYNAPTICS ASTRA SL2610 PRODUCT LINE

## OVERVIEW

Taking into account the issues inherent to today's edge AI ecosystem outlined above, it is clear that what is needed is a fundamental rethink of the developer experience. What is required is a new class of edge silicon to address the fragmentation that is further complicated by the rapid pace of AI innovation and model development.

*Click here for additional information on the Coral Neural Processing Unit*

Synaptics' latest Astra SL2610 product line—consisting of five families—features a series of systems designed from the ground up in collaboration with Google Research to address edge AI use cases and rectify the issues of fragmentation in silicon and software stacks that have, to date, inhibited the progress of sustainable edge AI productization. By integrating the Coral Neural Processing Unit (NPU) ML core from Google Research with Synaptics' Astra AI-Native hardware and open-source software, this partnership accelerates the development of intelligent, context-aware devices and machines in consumer and industrial markets.

## THE ASTRA SL2610

The Astra SL2610 is Synaptics' strategic next step into AI-native edge computing, offering scalable compute tailored for multimodal AI workloads in IoT environments for both consumer and industrial applications. With a 1 Tera Operations per Second (TOPS) NPU optimized for transformer models, dual-core Arm® Cortex®-A55 Central Processing Units (CPUs), and a low-power Cortex-M52 core with Helium Digital Signal Processor (DSP) extensions, it is engineered to deliver high-performance AI inference at minimal power and cost. Its architecture supports real-time applications like object detection, predictive maintenance, and voice recognition, targeting smart homes, industrial automation, and healthcare. The inclusion of PSA Level 3 security, ECC memory, and open-source Linux Software Development Kits (SDKs) ensures not only robust protection, but also developer flexibility. Ultimately, its AI value lies in its ability to bring contextual, multimodal intelligence directly to the device edge—reducing latency, enhancing privacy, and enabling autonomous decision-making without cloud dependency.

| Category | Feature Details |
|---|---|
| CPU & Architecture | Dual-core Arm® Cortex-A55 @ 2 GHz<br>64-bit architecture<br>Cache: 256 KB shared L3 |
| AI Processing | 1 TOPS secure NPU<br>Supports multiple DNN frameworks<br>Transformer & CNN support<br>FP16 & FP32 support |
| Memory Support | 16-bit LPDDR4/LPDDR4x/DDR4/DDR3L<br>Up to 3200 MT/s<br>ECC support |
| Graphics & Vision | Arm® Mali G31 3D GPU<br>MIPI-CSI (2-lane), 8-bit RGB<br>Soft ISP on A-core/NPU<br>MIPI-DSI® 1080p60 |
| Audio Capabilities | AI-based noise reduction & AEC<br>3x I2S (16 channels), 4x PDM, 1x SPDIF |
| Security | PSA Level 3 (target) with secure island<br>DPA/SCA cryptography protection<br>Secure boot & runtime integrity |
| Connectivity & I/O | 2x GbE with TSN/1588<br>2x CAN-FD<br>eMMC 5.1<br>2x USB 2.0<br>ADC (8 channels) |
| Software Ecosystem | Astra Yocto Linux SDK<br>Android (vertical markets)<br>AI MCU SDK: FreeRTOS & Zephyr |

## TORQ EDGE AI PLATFORM

The Synaptics Torq™ Edge AI platform amplifies this value by providing a unified software and hardware ecosystem purpose-built for AI scalability. At its core are the Torq "T" line NPUs—a multi-generational joint development integrating Google Coral NPU IP—paired with open-source AI software compilers, runtimes, and build frameworks. Supporting frameworks LiteRT (formerly known as TensorFlow Lite), TensorFlow, ONNX, and JAX, and leveraging the IREE compiler, Torq enables seamless deployment of advanced models across diverse edge devices, such as smart home hubs, industrial machines, and consumer electronics. Its modular design and open-source, developer-first philosophy empower rapid innovation, while curated multimodal AI models and public SDKs accelerate time to market. Torq's integration with Astra SL processors ensures a future-ready foundation for building intelligent, context-aware IoT systems that adapt to users and environments in real time, while also leveraging Synaptics' established legacy in wireless connectivity.

One of the most important elements of the Torq Edge AI Platform is its IREE compiler, originally developed at Google, and championed by industry leaders, such as AMD. The Multi-Level Intermediate Representation (MLIR) and Intermediate Representation Execution Environment (IREE) are open-source projects using a compiler structure developed under the LLVM project to streamline and unify AI software development across heterogeneous hardware platforms. They seek to address the fragmentation in AI frameworks by offering a modular and extensible Intermediate Representation (IR) that supports multiple levels of abstraction within a single compilation pipeline.

## GOOGLE CORAL NPUs: HARDWARE & SOFTWARE

Google Research's Coral NPU IP is based on the open-source ISA RISC-V, which can be accessed freely by developers and silicon vendors. It is a product of the Coral project, building on over a decade of experience and learnings within Google, in particular from the edge Tensor Processing Unit (TPU), a mass-produced custom ASIC with a vertical stack. With developer and SoC vendor enablement in mind, Google Research turned to a 100% RISC-V architecture with Coral NPU, which is open and extensible to distribute a common ISA front end for developers to streamline development, and for silicon vendors—of which Synaptics is the first—to productize.

Crucially, the work does not stop there, and this is a multi-generational commitment to improve and scale the capabilities of the Coral NPU IP, which will be freely distributed among those who wish to leverage its capabilities. This includes the ongoing work to ratify matrix extensions within the wider RISC-V ecosystem to expand the AI capabilities of all those building on RISC-V, as well as the accompanying AI software toolkits essential for optimizations.

## TARGET USE CASES AND FORM FACTORS

The unified, scalable platform will enable the execution of Generative Artificial Intelligence (Gen AI) on battery-powered devices, addressing multiple modalities, including vision, audio, and sensor data. This will primarily target the smallest form factors for consumers and is particularly appropriate for the emerging Agentic AI space that will leverage data from wearables such as smartwatches and earphones. Synaptics will also deploy this platform for larger form factors such as smart home gateways and hubs for use in IoT applications in the ambient environment, enabling connectivity with the cloud for hybrid AI experiences and leveraging the more

performant compute found in data center environments, combining several AI models, including for data processing and cleansing.

Crucially, all upcoming Stock Keeping Units (SKUs) can work on battery power alone and can scale up to 1 TOPS with the first generation NPU, with AI workloads still remaining on-device. Moreover, Google Research is also working to enable compatibility with its Gemma family of lightweight, state-of-the-art open models, designed to run on a wide variety of platforms, including mobile devices.

Focus applications include the following key target markets, which represent a sizable market opportunity:

- **Smart Home:** Home control and automation applications in devices such as gateways, intelligent hubs, control panels, thermostats, and appliances.
- **Industrial & Infrastructure:** Industrial process and factory control applications in machines and cameras, as well as Electric Vehicle (EV) charging infrastructure.
- **AI Hub and Point of Sale (POS):** Centralized hub to aggregate, manage, and distribute AI workloads, as well as POS and scanners in retail settings.
- **Additional Areas:** Audio, healthcare, drones, and other consumer goods such as e-bikes, scooters, and toys.

The resulting applications will be capable of use cases such as object detection, image classification, facial recognition, predictive maintenance, noise reduction, and echo cancellation.

# ABI RESEARCH COMMENTARY: HOW SYNAPTICS AND GOOGLE RESEARCH ADDRESS EDGE AI PAIN POINTS TODAY AND GOING FORWARD

Synaptics and Google Research are partnering to directly confront the fragmentation that has been observed by ABI Research, and that stifles innovation in the edge AI market. The current landscape is defined by a multitude of vendors offering powerful, but siloed solutions, each with its own proprietary software toolchains and closed MLOps ecosystems. This forces developers into vendor lock-in, increases development costs, and prevents the creation of portable AI applications that can run across different hardware. Indeed, one of the biggest complaints among edge AI developers today is the proprietary toolchains for different SoC vendors, which typically slow innovation. There are specific compilers and runtimes for specific classes of devices from individual vendors. This is exactly what Synaptics and Google Research are partnering to address: open the "black box" from the developer experience and allow models to be optimized and ultimately run on diverse hardware.

The partnership's GTM strategy represents a fundamental departure from these closed, vertical approaches. This new approach is to create an open and horizontal ecosystem, analogous to the strategy that made Android successful. Google Research is providing its Coral NPU IP, based on

the open RISC-V ISA, to be freely distributed, with Synaptics serving as the first silicon partner to productize it in their Astra SL2610 platform. This GTM approach lowers the barrier to entry for both developers and other hardware vendors, aiming to establish a de facto standard. By offering a common hardware foundation and a unified, open-source software platform (Torq), they reduce the risk for OEMs, which are no longer tied to a single silicon vendor's roadmap and proprietary tools.

This collaborative solution has the potential to become superior to alternatives because it directly solves the core problem of cross-vendor incompatibility, rather than just optimizing within a single ecosystem. The technical cornerstone of this approach is the MLIR compiler, which acts as a universal translator for AI models, allowing frameworks like LiteRT and ONNX to target a diverse range of hardware through a single, unified pipeline. For developers, this eliminates the need to learn and maintain multiple vendor-specific toolchains, enabling true software portability. Ultimately, this open ecosystem fosters faster innovation, accelerates the deployment of advanced Gen AI models on low-power devices, and creates a collaborative environment where the entire industry can work to solve challenges like standardizing RISC-V matrix extensions, a goal that is difficult within the confines of competing, closed systems.

For Synaptics, this will be an evolution in its second-generation Astra line of SoCs, which will progress in lock-step with Google Research for future iterations, after being first-to-market with its Coral NPU IP, also acting as a blueprint for success to bring other silicon vendors into the fold. For Google Research, this is about edge AI partnerships, developer enablement, creating best practices, and, ultimately, ecosystem building.

October 2025
157 Columbus Avenue
New York, NY 10023
Tel: +1 516-624-2500
www.abiresearch.com

**WE EMPOWER TECHNOLOGY INNOVATION AND STRATEGIC IMPLEMENTATION**

ABI Research is uniquely positioned at the intersection of end-market companies and technology solution providers, serving as the bridge that seamlessly connects these two segments by driving successful technology implementations and delivering strategies that are proven to attract and retain customers.

©2025 ABI Research. Used by permission. ABI Research is an independent producer of market analysis and insight and this ABI Research product is the result of objective research by ABI Research staff at the time of data collection. The opinions of ABI Research or its analysts on any subject are continually revised based on the most current data available. The information contained herein has been obtained from sources believed to be reliable. ABI Research disclaims all warranties, express or implied, with respect to this research, including any warranties of merchantability or fitness for a particular purpose.