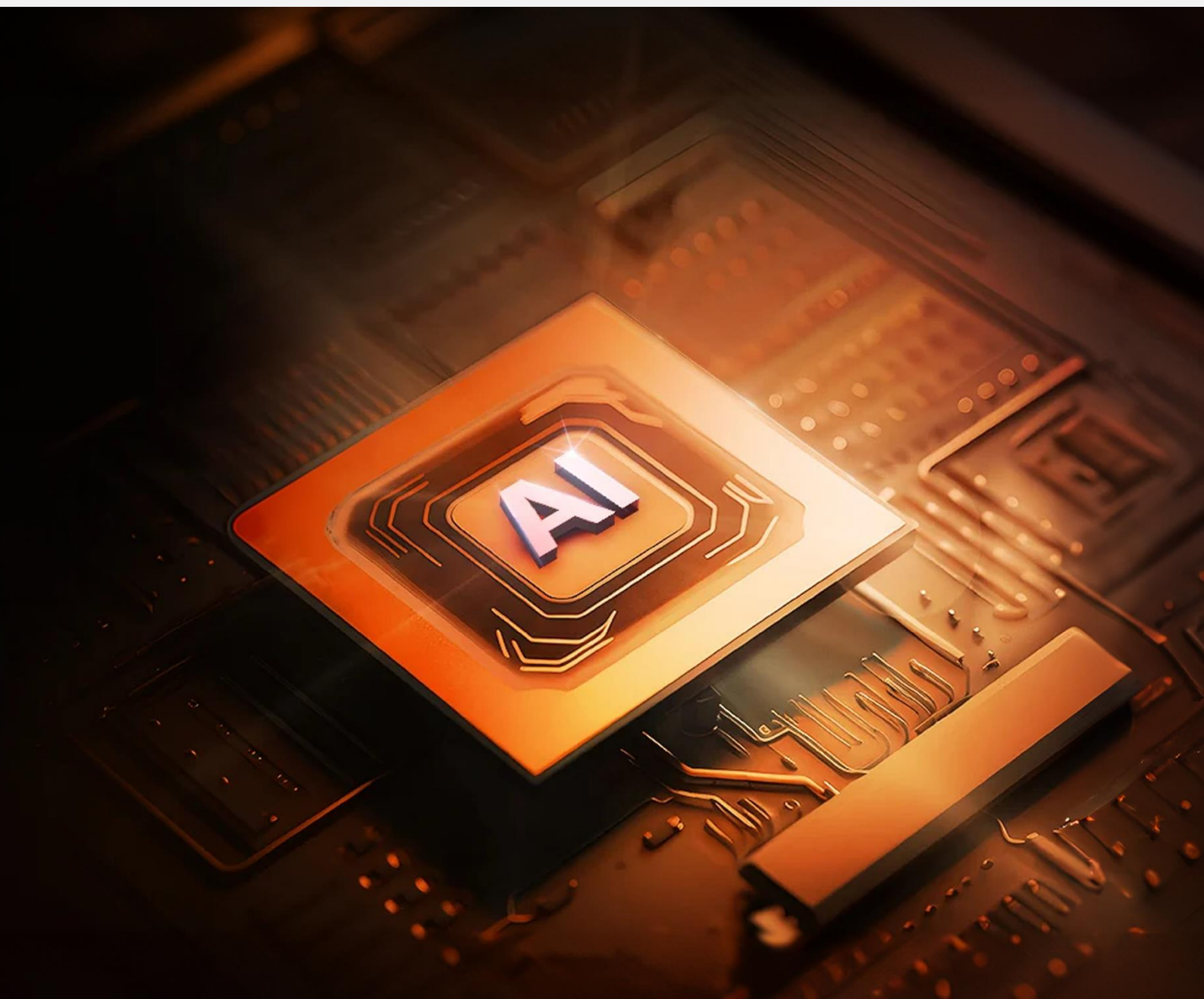


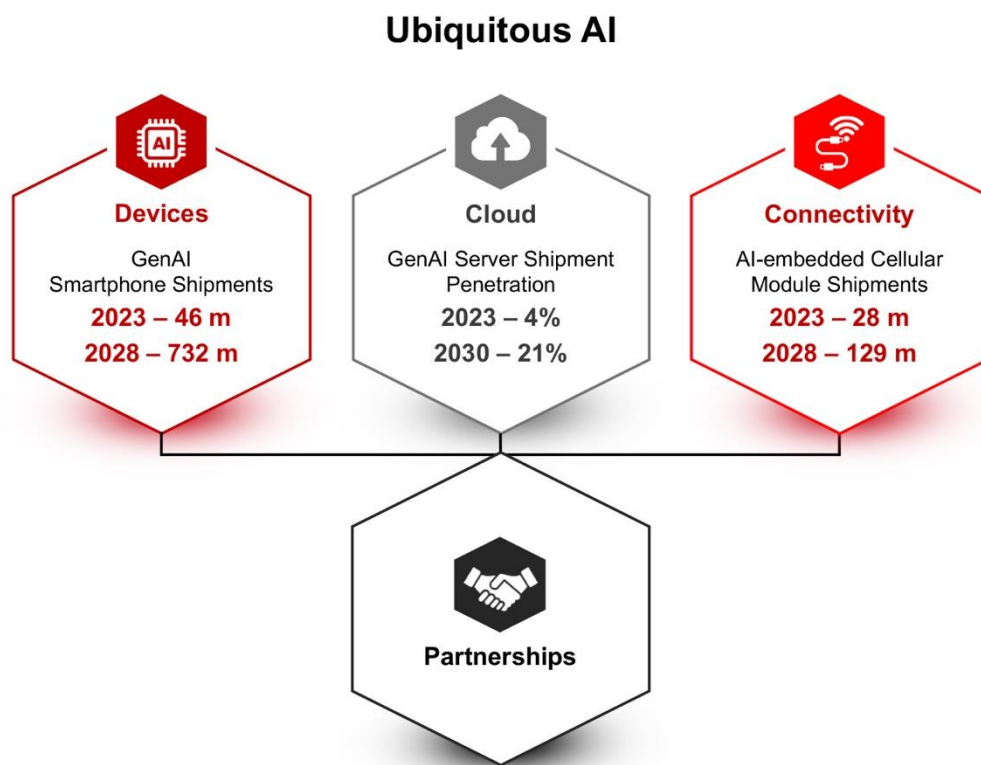
The Future is Here: Ubiquitous AI Computing



Executive Summary

- *The era of AI is upon us, with endless new products, services and applications demonstrating advanced AI capabilities.*
- *The future lies in bringing all these disparate segments together under one umbrella of ubiquitous AI to enable intelligence everywhere, all the time.*
- *The four pillars underpinning this vision are edge devices, cloud, connectivity and partnerships.*
- *MediaTek's vision of the future combines all four pillars, enabling ubiquitous AI computing.*

Almost every day we are seeing new products, services and applications demonstrating advanced AI capabilities, in particular Generative AI. The future, though, lies in bringing all these disparate segments together under one umbrella of ubiquitous AI.



There are four key pillars required to realize this vision:

1. **Devices:** The demand for faster performance is increasingly utilizing the speed, latency and privacy benefits of AI processing at the edge or “on-device. Whether for smartphones, inside an automobile or at home using a new gateway with Generative AI, AI computing is opening up new applications in the areas of productivity, entertainment and safety.
2. **Cloud:** As the performance requirements for AI processing increase, there is a growing demand for custom silicon, especially for AI accelerators in the cloud, to support ever increasing performance for training, compute and inference.

3. **Connectivity:** Edge computing will play a pivotal role in AI applications, but some processing will need to be sent to the cloud. This requires a hybrid compute model where cloud and edge devices work together. High-performance hybrid AI computing will require great connectivity for both wireless and wired technology, enabling a seamless, always-connected experience.
4. **Partnerships:** Underpinning the other three pillars is the need for companies to work together to accelerate AI innovation and harness the potential of AI, whether by supporting LLMs, pushing performance and efficiency to new heights, or exploring new AI-enabled use cases and verticals.

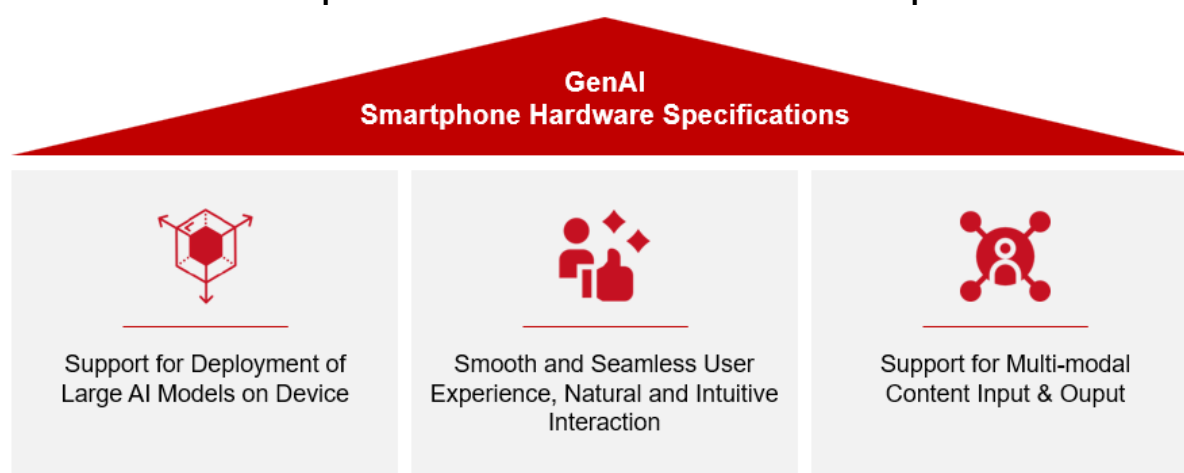
Through these pillars, MediaTek is at the forefront of ubiquitous AI everywhere, with its robust product offerings.

MediaTek is leading the charge in how we engage with AI in our everyday lives, from the home to the car and our mobile devices. Much of this is driven by greater demand for edge processing in everyday devices, where MediaTek has established itself as a proven leader in both technology and market presence. And AI computing on the edge now means data processing can take place on the device, increasing speed, decreasing latency and improving privacy.

Devices

For **smartphones**, MediaTek introduced the Dimensity 9300 flagship chipset last year with the industry's first All Big Core design that combines industry-leading performance and power efficiency. MediaTek followed up this year with an even more powerful Dimensity 9400. This new CPU+GPU+NPU platform is capable of over 70 total TOPs – with the NPU being over 50 TOPs, more than that in the iPhone 16 Pro Max with 35 TOPs – and provides significant improvements for on-device generative AI performance, which includes support for LLMs up to 33 billion parameters. In addition to high performance, MediaTek has also improved the power consumption of the platform by designing the Dimensity 9400 in TSMC's 3nm process technology. These capabilities usher in new possibilities for Generative AI applications across productivity, entertainment and creativity for users in a natural and personalized way.

Counterpoint's Definition of a Generative AI Smartphone



Generative AI is also transforming the **automotive** industry by powering more personalized and intuitive user experiences. More and more vehicles are using voice assistants, which add a layer of safety by giving drivers more control without taking their hands off the wheel. Generative AI will take this experience to the next level by enabling personal assistants that will create a more conversational experience. MediaTek's AI-enabled Dimensity Auto Cockpit platform runs LLMs as well as multimodality models in the car, allowing for in-vehicle personal assistant support, driver alertness detection, and many other leading AI-based safety and entertainment applications.

In-vehicle User Experiences Redefined with AI

Reimagining How We Interact with Our Vehicles



Augmented Reality Driver Info



AI Assistants



In-car Entertainment

Source: MediaTek

This is why earlier this year MediaTek announced its CX-1 platform, the highest end of its four Dimensity Auto Cockpit chipsets built using MediaTek's latest manufacturing process to maximize performance, feature integration and improve power efficiency, all while adhering to the automotive industry's high-quality standards. These will power the AI cockpits of the future.

In the home, many **smart home** devices already incorporate voice assistants for entertainment, information retrieval, calendar settings and more. But Generative AI will add

Smart Home Use Cases



Source: MediaTek

more personalized interactions, such as in the case of MediaTek's "Home Hub", which offers faster and more accurate responses and the ability to learn a user's preferences over time. With this innovation, we see three use cases emerging – Home Search, Home Safety and Home Diagnosis/Troubleshooting, relying on Generative AI running inside a home gateway with the lowest latency possible while maintaining total home privacy.

Cloud

Increasing AI processing and performance requirements are driving demand for custom silicon, such as AI accelerators in the cloud, which can lower overall costs while at the same time improving performance and power efficiency. Data centers are driving this boom for custom AI accelerators.

MediaTek's leadership in the ASIC market helps companies maximize overall value – from chip design, manufacturing and packaging to product support – and its computing expertise enables maximized performance in extremely power-efficient chips.

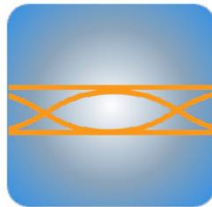
MediaTek is one of the very few companies with the knowledge and expertise to design highly integrated silicon solutions using the latest process technology. It has experience in designing products at 4nm, 3nm and beyond. Furthermore, MediaTek has a portfolio of technologies to allow die-to-die interconnects such as 224G (and up to 400G in the future) SerDes, D2D and UCIe. These are critical in supporting the increasing complexities of AI computing that require high bandwidth interconnects.

And critically, MediaTek is among the very few with the capability to manufacture highly integrated silicon solutions using advanced packaging options such as 2.5D and 3D designs with CoWoS. In addition, MediaTek has designs for embedded high-density, low-power SRAM which play a critical role in AI accelerator applications as well.

MediaTek Leading-edge Capabilities for AI Acceleration



Computing xPU



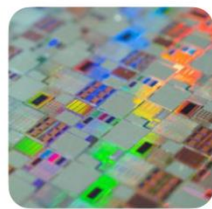
Interconnect SerDes,
D2D, UCIe



Advanced Process



Chiplet 2.5/3D
Packaging, HBM



High-Density
Low-Power SRAM



Broad Supply
Chain Ecosystem

Source: MediaTek

Connectivity

Edge computing itself is pivotal in running AI applications to reduce power consumption, lower inference cost, improve latency and protect user privacy. However, there will be cases where some applications will need to be sent to the cloud, which is why hybrid computing is essential to making ubiquitous AI a reality. This highlights the need for great connectivity between the edge device and the cloud. MediaTek has the most expansive portfolio of connectivity technologies for both wireless and wired technologies.

5G, for example, is essential for keeping users connected on the go, but Wi-Fi and Bluetooth will also be important, as will 10Gigabit and above fiber and ethernet. Thanks to having the industry's broadest connectivity portfolio, MediaTek has the expertise to deliver the latest connectivity technologies seamlessly.

Fast and Reliable Connectivity Key to Hybrid AI Computing



Partnerships and Collaboration

Underlying the above three pillars are partnerships. The acceleration of AI requires companies from different fields to work together to truly harness the potential of ubiquitous AI. This strong partner ecosystem is critical, whether for supporting new models, pushing performance and efficiency to new heights, or for exploring new AI-enabled verticals and new end-user applications.

MediaTek is a great example of an industry leader working closely with a wide range of partners to drive AI innovation. For example:

- MediaTek chipsets offer support for cutting-edge models, including Meta Llama, Google Gemini, Baichuan AI and Alibaba Cloud Qwen.
- Leveraging Arm's technologies, MediaTek is also part of the Arm Total Design ecosystem, expanding the footprint of Neoverse CAA in cloud AI and infrastructure applications.
- MediaTek also works closely with NVIDIA to provide advanced AI that powers the connectivity and computing capabilities for next-generation smart cabins. It has also integrated NVIDIA TAO into MediaTek's NeuroPilot SDK, helping companies harness the potential of edge AI in diverse IoT verticals.

MediaTek's Strong Global Ecosystem and Partnerships

 Baichuan AI	 Gemini Nano	 	
 	 	 Meta Llama 3	 Alibaba Cloud Qwen LLM

Source: MediaTek

Alone, each of these pillars can bring advanced AI capabilities to a wide range of devices and applications. But only by combining these pillars can the true potential of ubiquitous AI be realized and lead to intelligence everywhere.

As an industry leader in developing high-performance and power-efficient system-on-chip (SoC) products, MediaTek is uniquely positioned to drive this vision and support the adoption of Generative AI applications and technologies, whether in its diverse product ecosystem, on the cloud, at the edge in devices, or as a hybrid approach that takes advantage of both the cloud and the edge together. By pushing the boundaries of what AI can do, MediaTek is unlocking new opportunities, discoveries and progress that will shape the future.

Authors, Copyright, User Agreement and Other General Information



Jan Stryjak

Associate Director

✉ jan@counterpointresearch.com



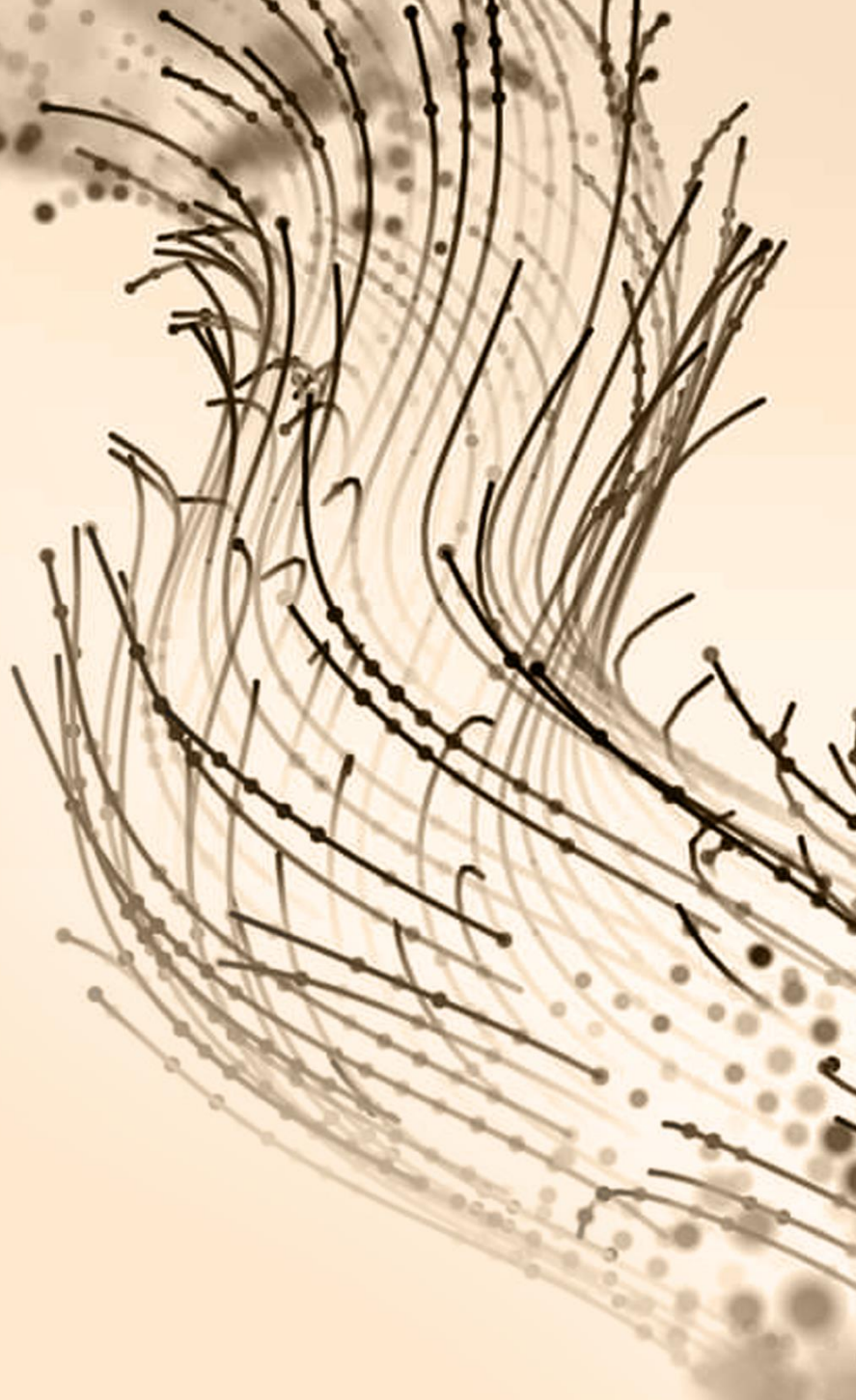
COUNTERPOINT TECHNOLOGY MARKET RESEARCH

Hong Kong | USA | South Korea | India | UK | Argentina | China

info@counterpointresearch.com



©2024 Counterpoint Technology Market Research. This research report is prepared for the exclusive use of Counterpoint Technology Market Research clients and may not be reproduced in whole or in part or in any form or manner to others outside your organization without the express prior written consent of Counterpoint Technology Market Research. Receipt and/or review of this document constitutes your agreement not to reproduce, display, modify, distribute, transmit or disclose to others outside your organization the contents, opinions, conclusions or information contained in the report. All trademarks displayed in this report are owned by Counterpoint Technology Market Research and may not be used without prior written consent.



COUNTERPOINT TECHNOLOGY MARKET RESEARCH

Hong Kong | USA | South Korea | India | UK | Argentina | China | Taiwan | Japan

info@counterpointresearch.com

Contact Us:

www.counterpointresearch.com

info@counterpointresearch.com

✕ @CounterPointTR

in @Counterpoint Technology Market Research