Internet of Things and Devices Poised for Edge AI-Driven Upgrades

Ido Caspi
icaspi@globalxetfs.com

Date: July 25, 2024
Topic: Thematic

For AI to fully integrate into everyday consumer life, end-user devices must be able to process AI workloads, rather than solely relying on cloud computing. On-device AI can reduce costs and latency, improve personalization and privacy, and enable real-time processing, which are crucial to bringing AI to everyday applications and use cases. As a result, we expect edge AI, which is closely associated with the Internet of Things (IoT), to continue to gather momentum.

Edge AI spans various applications across IoT and consumer devices, such as smartphones, wearables, headsets, drones, industrial technology, and automobiles. By 2027, an estimated 62% of data will be processed on edge devices. And by 2028, an estimated 26 billion short-range IoT connections will be driven by on-device AI, benefitting various associated technologies that enable connected devices infrastructure.[1,2] As a result, we believe edge AI presents an attractive investment case for investors looking to find AI opportunities outside of semiconductor players and cloud hyperscalers that have dominated the AI story thus far.

GLOBAL IOT DEVICES ARE EXPECTED TO APPROACH 30 BILLION IN 2024 AND TOP 50 BILLION IN 2029

Sources: Global X ETFs with information derived from: Ericsson. (n.d). Ericsson Mobility Visualizer. Accessed on July 1, 2024.
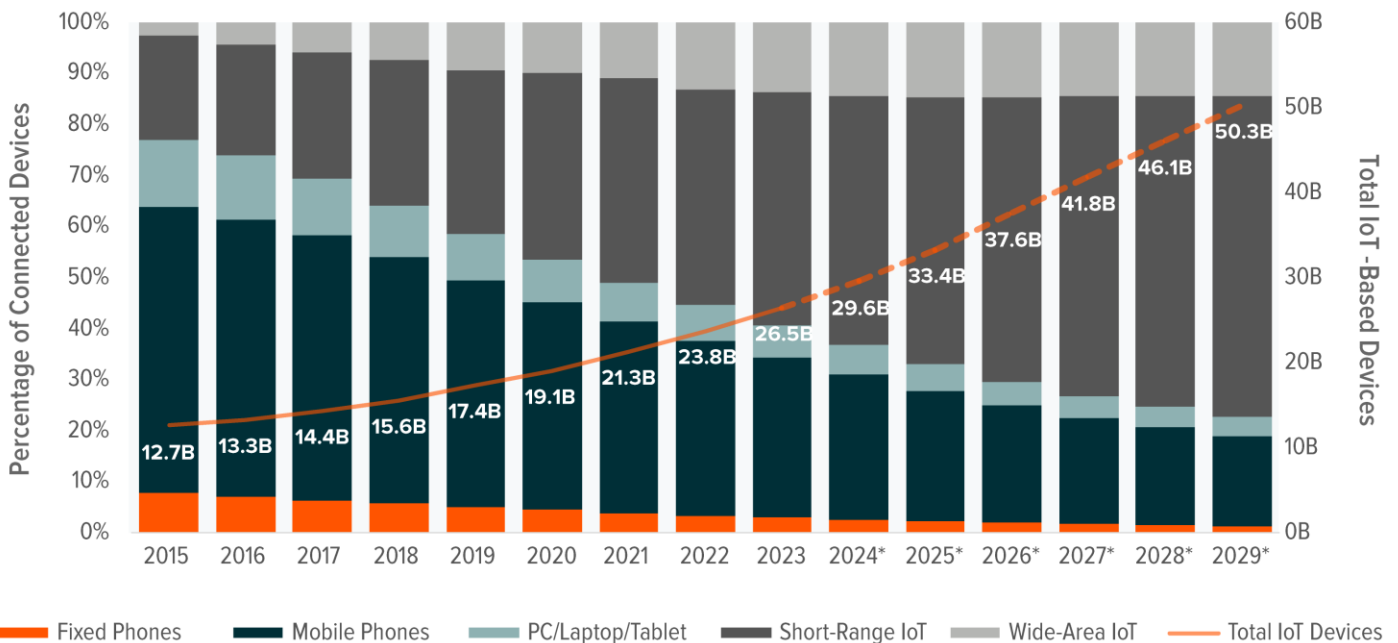
*Forecast.
Note: Connected devices classified within IoT include items such as wearables, machines, sensors, connected cars, and more. These devices are further classified into Short-Range IoT and Wide-Area IoT based on connectivity range.

Key Takeaways

- With the rise of generative AI, the push for greater computational power directly on consumer hardware is accelerating and expanding.
- On device AI integrations are expected to reverse the downtrend in smartphone sales and drive the next major upgrade cycle, particularly for Apple's iPhones.
- As consumer behaviors increasingly embrace AI technology, various devices across multiple industries will require upgrades to integrate AI capabilities, ensuring all connected devices are equipped with advanced functionality.

Smartphone Giants Usher in the Era of Generative AI-Powered Devices

AI offers smartphones, a $500 billion annual market, the greatest near-term monetization opportunity among consumer devices by accelerating replacement cycles and increasing AI-related premiums.[3] Virtually all leading smartphone providers already have dedicated AI strategies in place, including Apple and Samsung.

Apple recently introduced Apple Intelligence, a suite of AI-powered features for new iPhones, iPads, and Macs.[4] Unlike the chatbot-centric approach that many AI players have taken thus far, Apple's AI focuses on personalization and app integration, enhancing the overall user experience. These tools include an improved Siri, powered by in-house large language models (LLMs) and OpenAI's GPT-4o.[5] In our view, the new generative AI-enabled Siri with enhanced natural language processing (NLP) and in-app command capabilities that generate more accurate responses and task execution across iPhone apps, positions Apple as the most sophisticated consumer digital agent. Additional features include enhanced predictive text, advanced photo and video editing, and generation tools that are likely to redefine how users interact with their devices.[6]

Samsung's AI approach for smartphones emphasizes communication and productivity, focusing on specific tasks such language translation. In January 2024, Samsung introduced Galaxy AI, a collection of AI tools designed to improve the user experience through a combination of on-device AI and cloud-based AI tools from third parties.[7] For example, Galaxy AI enables the Samsung Keyboard to offer translation features across every app and Galaxy phones to handle live phone call translations, acting as personal interpreters.[8] Additionally, several photo editing features let users quickly edit, change, or reposition objects within images. Another tool is Circle to Search, which enables users to launch a Google search for anything on screen by circling it with their finger.[9]

On-Device AI to Fuel the Next Big Smartphone Upgrade Cycle

We expect generative AI integrations to reverse the downturn in smartphone sales and propel a major upgrade cycle. For Apple, the impact may be comparable to the introduction of the iPhone 12 in 2021, the first 5G iPhone, and the iPhone 6 series, which featured larger screens. Importantly, Apple Intelligence will be limited to devices with the A17 Pro or M-series chips, or just 5% of the current iPhone installed base, meaning 95% of iPhone users will need to upgrade their devices to experience Apple Intelligence.[10]

Historically, most U.S. users upgrade their smartphone every 2-3 years.[11] In recent years, though, Apples iPhone replacement cycle lengthened, likely due to a combination of macroeconomic headwinds that impacted consumer spend and a slower pace of innovation. Nonetheless, the most recent iPhone super cycle was roughly 3 years ago with the iPhone 12 and 5G, in line with the historical average.[12]

The iPhone 12 surpassed 100 million units sold within just seven months, two full months faster than the iPhone 11, which was released a year prior.[13] And the iPhone 12 nearly matched iPhone 6 series sales, which drove Apple's first major super cycle.[14] We expect accelerated replacement cycles to occur with the AI iPhone as well, given that users will be heavily incentivized to upgrade to a personalized, handheld device equipped with AI.

INNOVATIVE FEATURES AND ENHANCED FUNCTIONALITY HAVE ACCELERATED IPHONE REPLACEMENT CYCLES
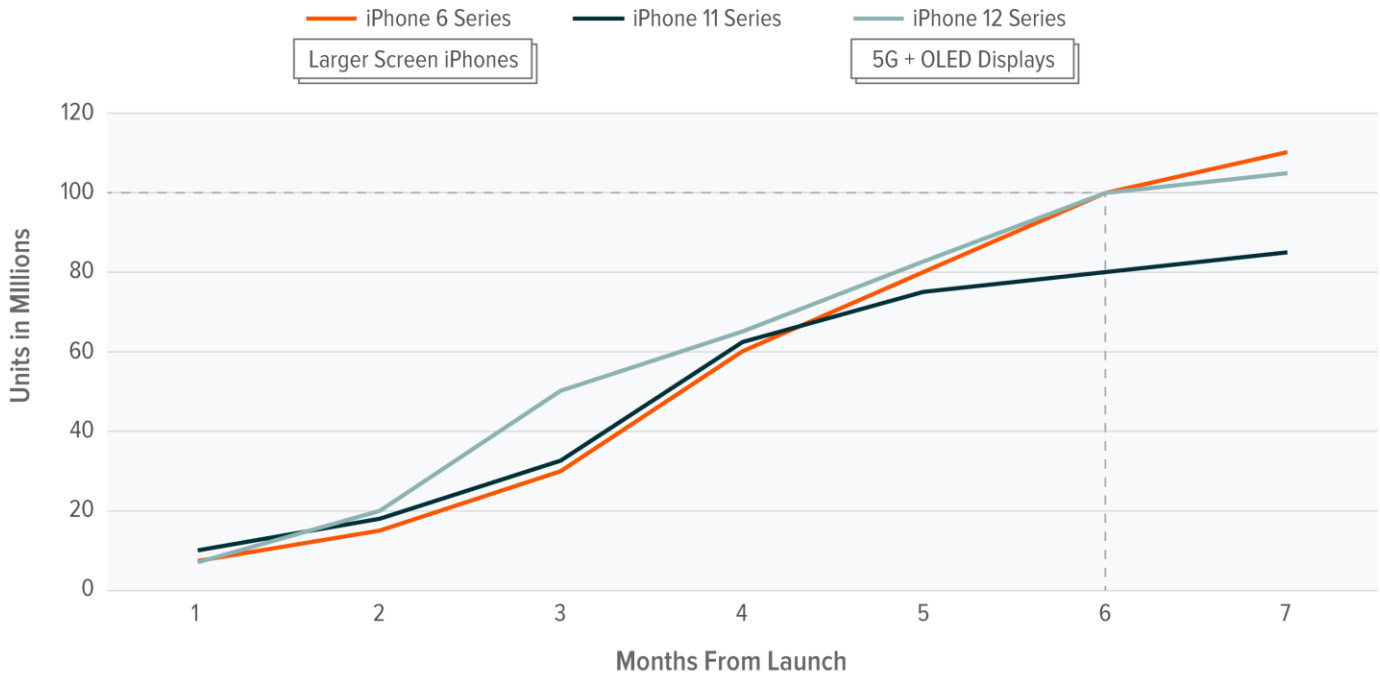
As the AI smartphone story plays out, we believe that IoT companies in the smartphone supply chain are well-positioned, including system-on-a-chip (SoC) and memory players. Smartphones have advanced from processing AI models with a few hundred million parameters to models exceeding 1 billion parameters.[15] Future models are expected to process, or run inference, on models with more than 10 billion parameters.[16] High inference costs have hindered adoption of early consumer-facing generative AI offerings. But with the ability to handle more parameters, model providers can shift low level tasks from data centers to smartphones to reduce inference costs and allow developers to leverage edge computing for a diverse set of front-end consumer experiences.

AI-Driven Shifts in Consumer and Corporate Behavior Expected to Propel Connected Device Upgrades

Beyond smartphones, generative and edge AI have the potential to unlock demand for a variety of more powerful machines across various industries and use cases over time. In Q2 2024, Microsoft introduced its Copilot + PCs, the first 40+ trillion operations per second (TOPS) AI PCs with an on-device AI assistant.[17] In Q1 2024, AMD's chips sales increased by a substantial 85% year-over-year (YoY), driven by robust demand for its Ryzen 8000 series CPUs tailored for desktop and laptop applications.[18] An estimated 54.5 million AI-embedded PCs are expected to ship in 2024 as manufacturers ramp up production, accounting for nearly one-quarter of total PC spend.[19] By 2027, 60% of PC shipments could have AI capabilities.[20]

Wearables that measure biometric and healthcare data are also likely to benefit from on-device AI. Between 2016 and 2022, the number of global connected wearable devices surged 240% from 325 million to 1.1 billion.[21] However, much of the data processing that these devices mediate occurs on the cloud or on a connected smartphone. Computing power costs are high, and latency is an issue, which can negatively impact feedback loops and the user experience. As AI models become more sophisticated and personalized, on device data processing should improve and better the relationship between device and user. Among the most promising potential use cases is remote patient monitoring.

ON-DEVICE AI ADVANCEMENTS COULD IMPROVE REMOTE PATIENT MONITORING

Sources: Global X ETFs with information derived from: Prevounce. (2024, January 19). 27 Remote Patient Monitoring Statistics Every Practice Should Know.
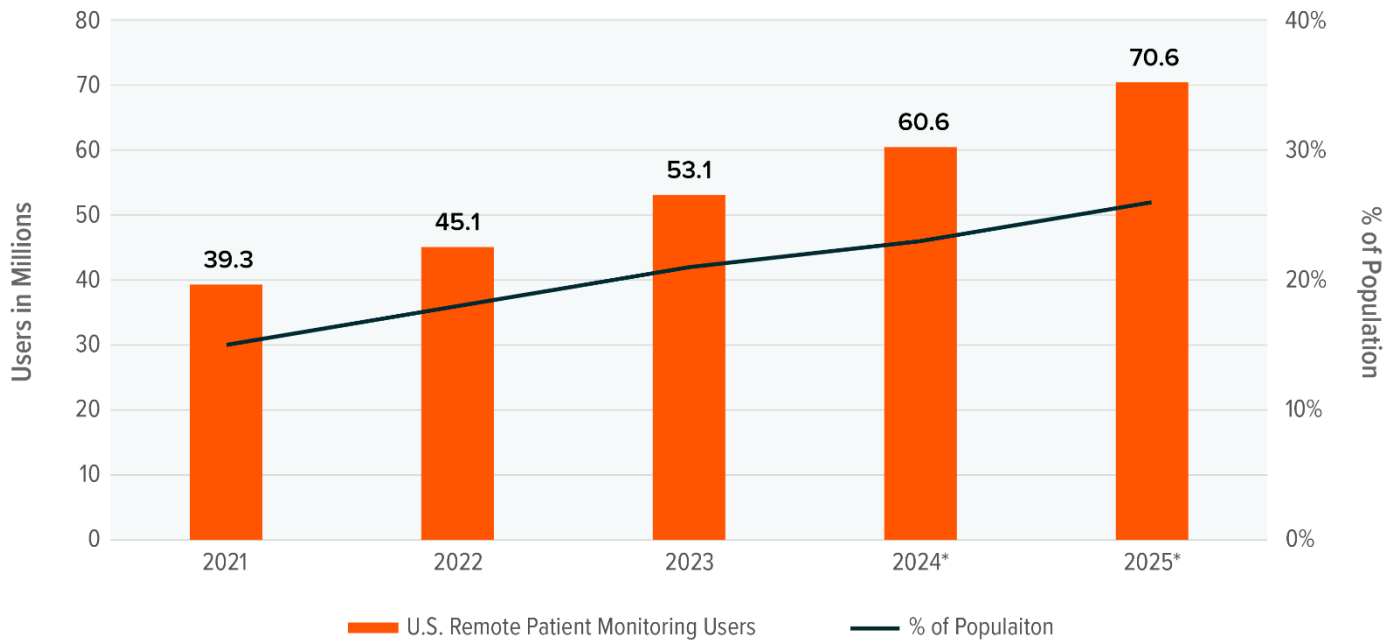
*Forecast

Industrial IoT offers significant growth potential, with one forecast showing the potential for the global market to grow at a 23% compounded annual growth rate (CAGR) from $394 billion in 2023 to $1.7 trillion by 2030.[22] Advancements in AI, particularly through automation of manual tasks, are expected to help, as humans still feed significant information into machines, leading to inefficiencies. Even in our technology-driven world, in 2020, only 31% of manufacturing companies had a single fully automated process.[23] For AI to drive more efficient processes and increase productivity, IoT nodes will be critical in providing the hardware layer necessary to capture and process data. IoT sensors produce data that AI can gather and analyze for predictive maintenance, quality control, and workplace safety. Sensor data can detect anomalies early, predict machine failures, and alert employees to necessary maintenance in advance. Facilitating microsecond-decisions can allow manufactures to catch issues in real-time on assembly lines, ensuring higher quality in large-scale production.

Growing investments to support reshoring of manufacturing is also driving demand for Industrial IoT. AI chipsets used directly on industrial equipment can reduce costs, network traffic, and enhance security, as they reduce the amount of data going to the cloud for processing. With exposure to diverse end markets, a company like NXP Semiconductor, reported that revenues for its Industrial & IoT segment grew a strong 14% YoY in Q1 2024.[24]

Edge computing is also crucial for autonomous vehicles (AVs) to make real-time decisions without relying on a stable internet connection or cloud servers. AVs use video cameras and sensors like LiDAR and ultrasonic to understand their surroundings, as training a computer to operate a vehicle requires vast amounts of pre-processed real-world data. Tesla and Mobileye are among the companies developing the necessary AI infrastructure and advanced driver-assistance systems (ADAS) to facilitate this process. High-speed connectivity Is also needed for continuous updates and on-road condition notifications. Tesla designed its own computer chips for these purposes, and other companies are expected to follow suit, benefitting the custom silicon market.

Conclusion: 2024 Expected to be a Catalyst Year for On-Device AI and the IoT Ecosystem

In 2023, the generative AI buildout largely benefited compute hardware and cloud infrastructure providers in the bottom and middle of the enterprise stack. In 2024, AI is gaining prominence and momentum at the edge, given its expected impact. IoT companies provide the connective layers between the physical and digital worlds and make computing at the edge possible via processing and sensing capability. These layers are essential for AI to fulfill its promise in driving efficiency, controlling costs, and mitigating labor shortfalls. For investors, we believe the potential for edge AI-driven applications and consumer use cases is an emerging AI opportunity set.

Related ETFs

SNSR – Global X Internet of Things ETF

AIQ – Global X Artificial Intelligence & Technology ETF

Click the fund name above to view current performance and holdings. Holdings are subject to change. Current and future holdings are subject to risk.

Footnotes

1. Medium. (2024, January 22). Edge Computing in the Age of AI: An Overview | Dell Technologies Info Hub.
2. RCRWirelessNews. (2023, June 22). Total IoT connections to surge, but cellular IoT to lose share with 2G/3G switch off.
3. Statista Market Insights. (2024, June). Smartphones – Worldwide.
4. CNET. (2024, June 11). Apple Intelligence Dominates WWDC: iOS 18 and Everything Else Revealed.
5. Ibid.
6. Ibid.
7. ZDNet. (2024, January 20). Samsung Unpacked 2024 recap: Galaxy AI, S24 Ultra, Smart Ring, and more.
8. Ibid.
9. Ibid.
10. Barrons. (2024, June 11). Apple Debuts a Clear AI Strategy. It Hopes to Spur iPhone Sales.
11. Consumer Affairs. (2023, December 12). Cellphone statistics 2024.
12. Counterpoint. (2021, June 1). Monthly Market Pule – April 2021.
13. Ibid.
14. Ibid.
15. Qualcomm. (2023, September 6). Democratizing on-device generative AI with sub-10 billion parameter models.
16. Ibid.
17. Microsoft. (2024, May 20). Introducing Copilot + PCs.
18. AMD. (2024, April 30). AMD Reports First Quarter 2024 Financial Results.
19. CIO Dive. (2024, February 9). AI-enabled PCs to infiltrate enterprise IT this year.
20. Ibid.
21. Tech Report. (2024, May 30). 15 Wearable Technology Statistics [2023 Edition].
22. Grand View Research. (n.d). Global Industrial Internet of Things Market Size & Outlook. Accessed on July 1, 2024.
23. McKinsey & Company. (2020, August). The imperative for automation success.
24. NXP Semiconductor. (2024, April 29). NXP Semiconductors Reports First Quarter 2024 Results.