Will AI PCs Change How We Work?
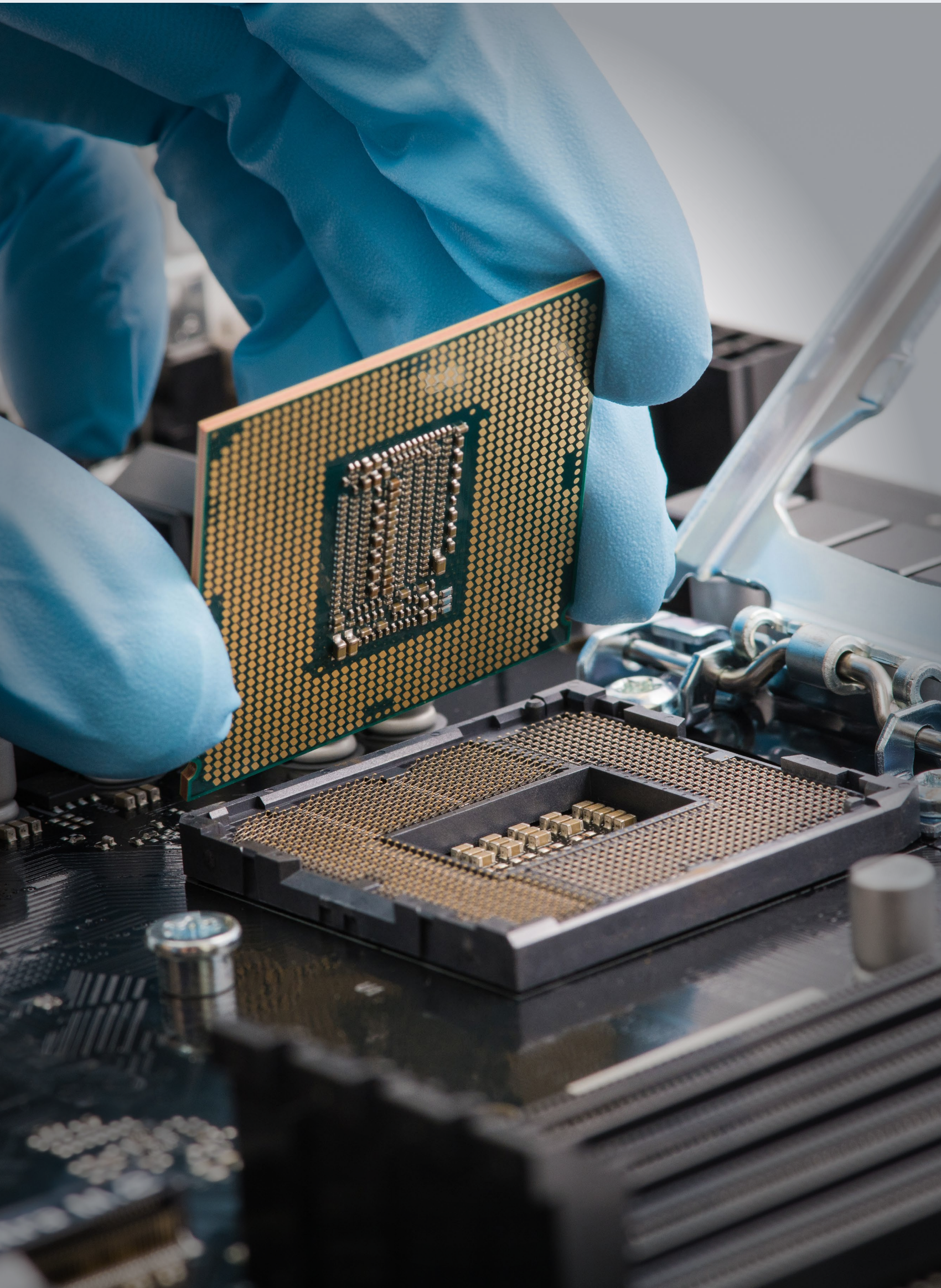
Foreword and Contents

The PC industry looks set to capitalize on the many benefits of bringing AI capabilities from the cloud to the client, with AI PC shipments predicted to hit 167 million by 2027[1]. Faster performance equating to enhanced user productivity, lower inferencing (the process of transforming a prompt into a response) costs, and better on-device privacy and security are just some of the promises fueling strong IT decision-maker interest in this technology.

But as the AI PC hype ramps, what key applications will actually move the technology from niche to a majority? How does cloud stack up against local AI applications? What are the security benefits, and what are the storage implications when it comes to keeping pace with this evolving technology? This eBook will address these questions and explore whether AI PCs really will change the way we work, with the help of leading AI experts.

1. PureAI.com:
   https://pureai.com/Articles/2024/02/14/AI-PC-Shipments-to-Hit-167-Million-by-2027.aspx

Contributors

This eBook has been created by two industry experts in AI, IT, and emerging technologies.

Rafael Bloom

Rafael has spent his career in senior Technology Product, Marketing Communications, and Business Development roles. His advisory practice focuses on the new organizational, product, and communications challenges of technological and regulatory changes.

This highly diverse work involves subject matter expertise on information governance and compliance by design, data privacy, and emerging technologies such as AdTech, mobile & 5G, AI, and machine learning.

Rob May

Rob May is the founder and Executive Chair of ramsac. He has a team of over 120 consultants working with him to deliver technology and cybersecurity services/support (and has done so for over 30 years).

Rob is the UK Ambassador for Cyber Security for the Institute of Directors and is placed #5 in the global rankings for cybersecurity thought leaders/influencers.

He is on the Advisory Board of The Cyber Resilience Centre for the South East, working with industry, academia, and law enforcement.

AI PCs in Today's Workplace

Today's organizations need to account for the major shift in capabilities of dedicated artificial intelligence (AI) tools and ensure their hardware can execute AI task loads both now and in the future. Which is one reason AI PCs are growing in usage. Designed to handle complex AI tasks locally, AI-optimized PCs offer substantial benefits in terms of processing speed, data privacy, and real-time analytics. They enable organizations to perform intensive AI computations in-house without relying on cloud services. This capability is crucial for tasks requiring immediate data processing and high levels of security.

These computers are adept at handling large datasets and complex algorithms with speed and efficiency, supporting activities such as real-time data analysis, on-the-fly content creation, and sophisticated simulation or modeling tasks. They are particularly beneficial in environments where quick decision-making is critical, such as in financial trading or emergency response scenarios.

Beyond enhancing productivity, AI-optimized PCs cater to several niche applications such as healthcare, where the ability to quickly process high-resolution images and complex diagnostic data directly on an AI-optimized

PC can help healthcare providers deliver faster and more accurate patient care. Scientific research is another area where massive datasets are common, in fields like genomics or climatology. In this case, researchers can benefit from AI PCs that can process data locally, enabling faster iterations and real-time anomaly detection without compromising data privacy.

> " In terms of productivity, AI-optimized PCs transform several workplace dynamics. They allow for quicker processing of data-intensive tasks, which is vital in roles that rely on data analytics or engineering simulations. For example, financial analysts can run complex models and simulations directly on their local machines, drastically reducing the time spent waiting for cloud processing and data transfers. "
>
> **Rob May**

Personalizing PC User Experiences with AI

AI also plays a transformative role in enhancing personalized user experiences on PCs, leveraging its capabilities to tailor computer interactions according to individual user preferences, habits, and needs. This personalization not only improves the overall user experience but also boosts productivity and efficiency by adapting the PC environment to better suit the user's specific workflows and preferences.

> "
> In essence, AI significantly enriches the user experience on PCs by making interactions more intuitive, efficient, and customized to individual needs. As AI technology continues to advance, these personalized experiences are expected to become even more sophisticated, further blurring the lines between user and technology in a seamless and productive manner.
> **Rob May**
> "

> "
> A lot of the new functionality is about playfulness, like automatically recognizing a gesture like a thumbs-up in a video call and overlaying graphics, or being able to erase an unwanted element of a photograph without having to train as a graphic artist first. As AI experiences become more 'baked-in' to operating systems, I suspect the experience will be more like that of a capable concierge.
> **Rafael Bloom**
> "

AI can analyze how a user interacts with their PC, including commonly used applications, frequently accessed files, and preferred system settings. Using this data, AI can dynamically adjust the user interface to simplify accessibility and align with the user's most common tasks. This can reduce the time taken to navigate PCs and increase a user's overall efficiency. From predictive text and intelligent search to voice-activated assistants to enhanced accessibility features; by learning from individual needs and usage patterns, AI can adjust features to provide a more personalized and accessible computing experience.

Local AI Virtual Assistants Enhance Productivity on PCs

Local AI virtual assistants, integrated within AI-optimized PCs, represent a significant evolution in how computers augment the daily workflow of professionals. By running directly on local machines, these assistants can leverage the full power of specialized AI hardware such as GPUs and NPUs, to deliver faster and more responsive assistance.

This immediacy allows for real-time interactions between the user and the assistant, significantly smoothing the workflow. AI is also capable of taking accurate minutes from a meeting and creating a shared document without a human being involved in notetaking, summarizing, and distribution. This greatly assists in multitasking environments, where users often switch between tasks rapidly and need instant responses to maintain their flow of work.

With data processing and storage kept local, sensitive information does not need to be transmitted to the cloud, reducing the risk of data breaches. Local AI virtual assistants can handle sensitive tasks, such as drafting confidential documents or analyzing personal data, without the data ever leaving the device.

This ensures compliance with stringent data protection regulations and building trust. But in a world of ever-increasing security risks, how else can AI PCs solve key security and privacy challenges?

> " This is already happening via the existence of AI tools like Copilot. Essentially, productivity gains occur when an AI can safely be trusted to engage in certain tasks, with an assurance layer of a human being to validate the quality of the end result.
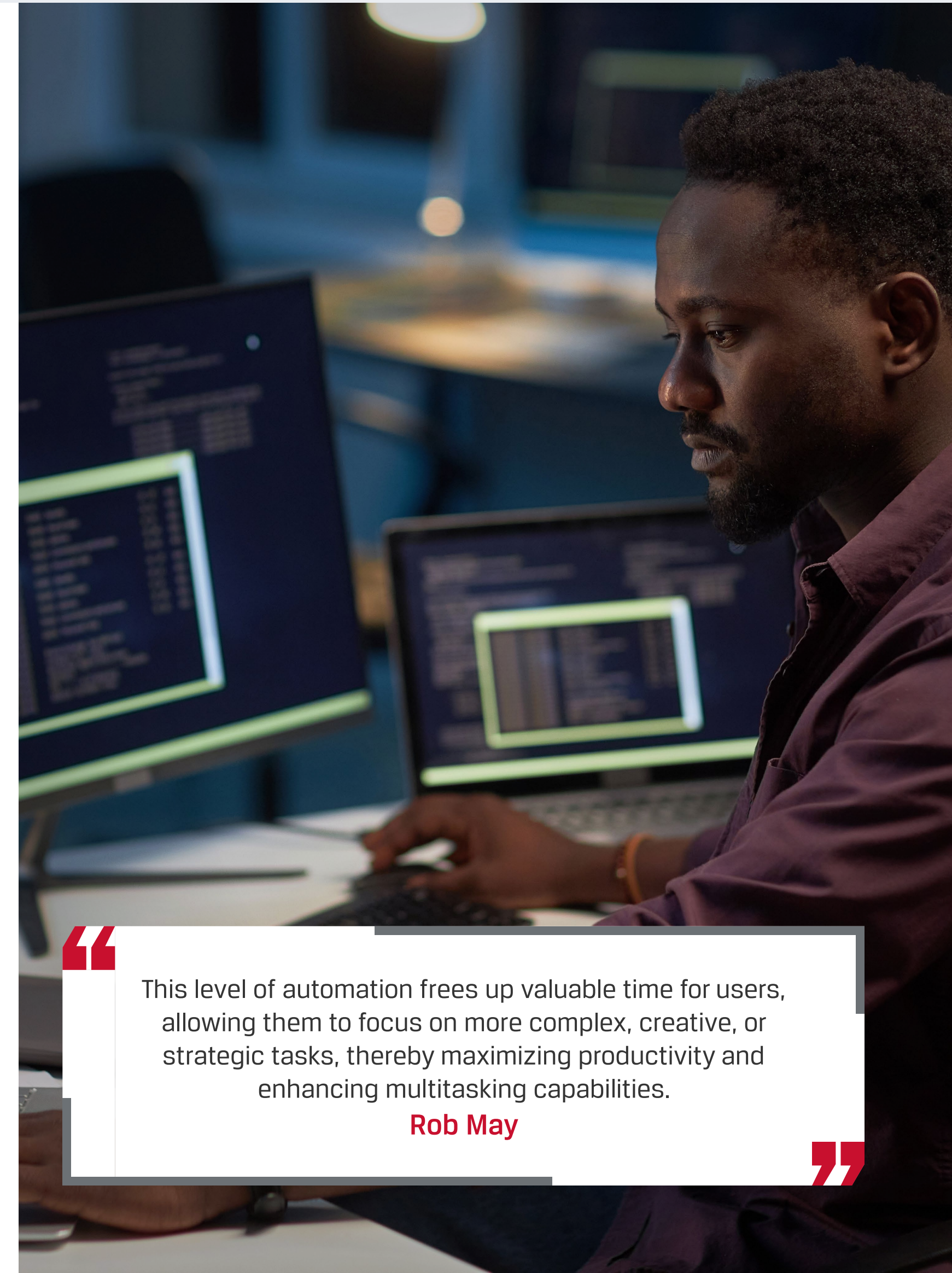> **Rafael Bloom**

> " This level of automation frees up valuable time for users, allowing them to focus on more complex, creative, or strategic tasks, thereby maximizing productivity and enhancing multitasking capabilities.
> **Rob May**

Security and Privacy Challenges Solved by AI PCs

With their enhanced processing capabilities for AI tasks, the integration of advanced AI tools directly into local hardware can significantly bolster an organization's cybersecurity measures, mitigate security risks, and enhance data privacy. Here are a few examples of how:

Enhanced threat detection and response: AI PCs are equipped to perform sophisticated real-time monitoring and analysis of system data, which allows for the early detection of anomalies that could indicate a cybersecurity threat, such as malware, ransomware, or intrusion attempts.

Improved data privacy: By processing and storing data locally, AI PCs minimize the need to transmit sensitive information over networks or store it in cloud environments, which can be susceptible to breaches. Local processing ensures that personal and sensitive data is handled within the confines of the user's physical environment, offering a higher degree of control and security.

Secure authentication mechanisms: AI PCs can implement advanced biometric authentication mechanisms, such as facial recognition or fingerprint scanning, which are more secure and less susceptible to theft or forgery than traditional passwords or PINs. The local processing capabilities of AI-optimized PCs ensure that biometric data does not need to leave the device, thus maintaining the privacy and integrity of the authentication data.
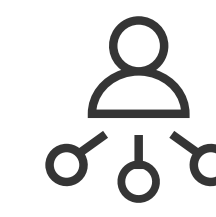
Network security: AI PCs can be utilized to monitor network traffic locally and identify potential threats or irregular behavior within an organization's network. This includes detecting unusual data transmissions that could signify data leakage, or spotting patterns that deviate from normal operational behavior. AI-driven systems can also help in segmenting the network, ensuring that only authenticated and authorized devices and users can access sensitive parts of the network, thus preventing lateral movement by potential attackers.

Given the privacy that AI PC workloads offer, should organizations only favor AI PC's vs. their cloud counterparts? Or does cloud still have a role to play?

While AI PCs provide significant privacy advantages by processing and storing data locally, there are scenarios where cloud or server solutions might be advantageous or even necessary.

As highlighted, the local processing on AI PCs ensures sensitive data remains within the organizational perimeter, which is crucial for industries governed by strict data protection laws. This reduces the risk of data breaches associated with external data transmission and storage. AI PCs can handle real-time data processing without the latency associated with data transmission to and from the cloud, and relying on local AI processing ensures critical operations are not disrupted by connectivity issues.

That said, cloud environments offer unparalleled scalability options compared to local solutions. Cloud solutions offer a pay-as-you-go model that can be more financially accessible. And cloud providers often offer cutting-edge AI capabilities that are constantly updated and improved, providing access to more powerful and diverse AI tools vs. on local machines alone.

Taking a hybrid approach, where AI PCs are used in conjunction with cloud or server solutions, combines the strengths of both paradigms and mitigates their individual limitations. This approach allows organizations to localize sensitive workloads, balance cost and performance, and enhance disaster recovery.

> While AI PCs offer substantial benefits in terms of privacy and real-time processing, a hybrid approach provides a flexible framework that allows organizations to leverage the strengths of both local and cloud computing. This approach is not only strategic in managing resources but also crucial in adapting to the varied and evolving needs of modern enterprises.
> **Rob May**

Memory and Storage Within AI PCs: The Next Five Years

While the role of cloud looks set to remain, the evolution of memory and storage within AI PCs over the next five years is likely to be influenced by the increasing demands of AI and machine learning workloads, which require rapid data access and substantial processing power. The development of specialized AI storage solutions, integration of non-volatile memory technologies, adoption of software-defined memory (SDM), and increasing memory capacity and speed will all have a significant impact.

For example, as AI models become more complex, the amount of data they need to process and store grows exponentially. For RAM, we anticipate advancements such as higher-density chips that can hold more data closer to the processor, reducing latency and increasing speed. Similarly, for storage solutions like SSDs, we will likely see continued improvements in read/write speeds and data throughput, alongside overall capacity increases, to facilitate faster data retrieval and handling of larger datasets. And, as environmental considerations become more pressing, energy efficiency will also be a focus. This includes the development of components that consume less power and generate less heat, thereby reducing the overall energy footprint of AI systems.

> The next five years will likely see significant technological advancements in memory and storage for AI PCs, aimed at supporting the increased complexity and scale of AI applications. These developments will not only enhance the capabilities and efficiency of AI systems but also address broader needs such as security and environmental sustainability.
>
> **Rob May**

Compared to their server AI counterparts, the current absence of HBM (High Bandwidth Memory) in AI PCs means DRAM is the most critical component that significantly influences the ability of AI PCs to handle demanding applications like deep learning or complex simulations. It affects not only the number of tasks that can be processed simultaneously but also the efficiency and speed with which these tasks are completed.
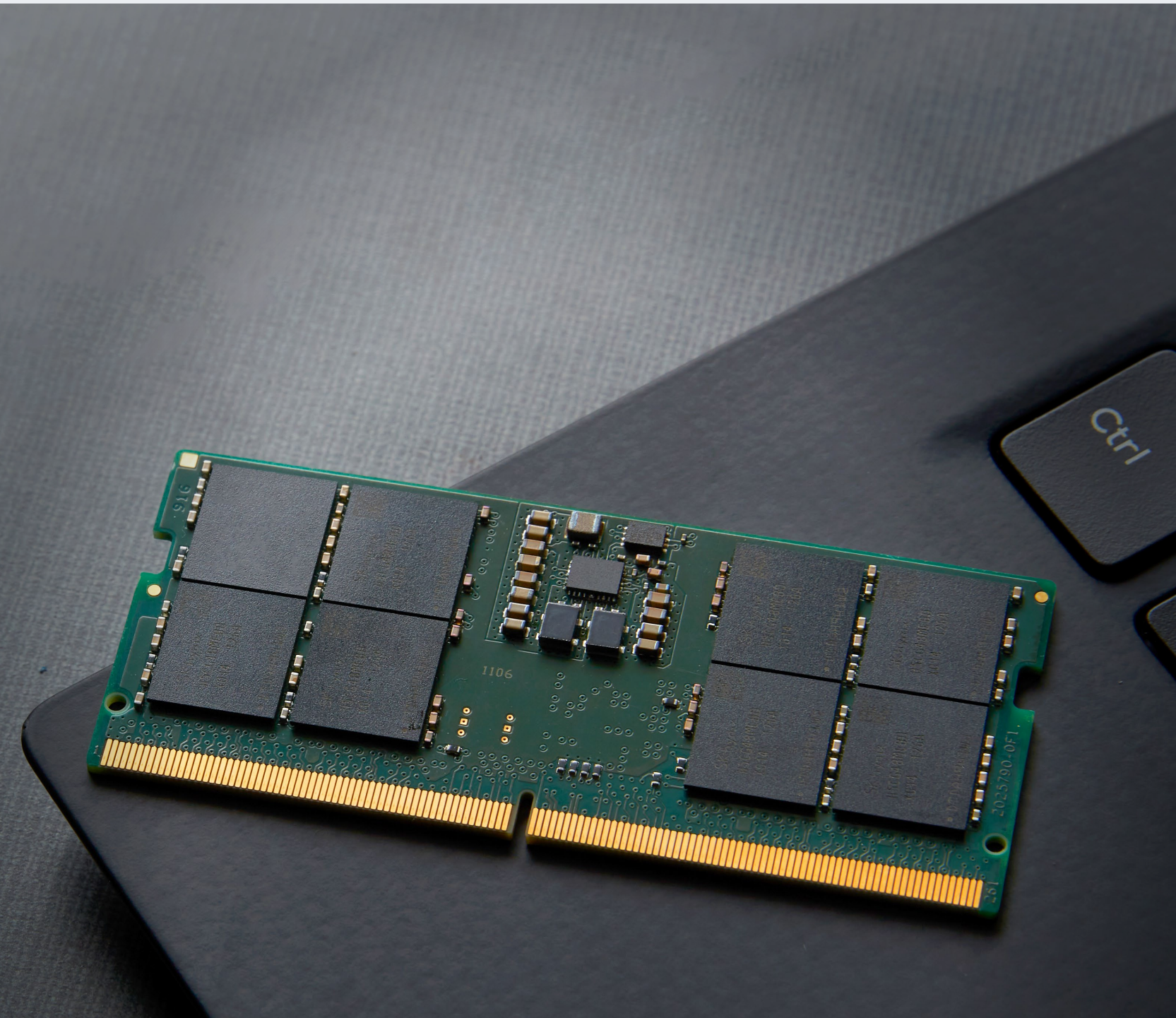
> Five years ago, no one spoke about NPUs. Even up until recently, it was largely GPUs taking the burden of the multidimensional computation necessary in AI. Similarly, the modalities of addressing storage and memory destinations in AI contexts are very different. All that being said, I believe we will see memory and storage media optimized for AI workloads occurring locally rather than in the cloud and AI PCs are the obvious platform for that.
>
> **Rafael Bloom**

How SSD Type Impacts AI PC Performance
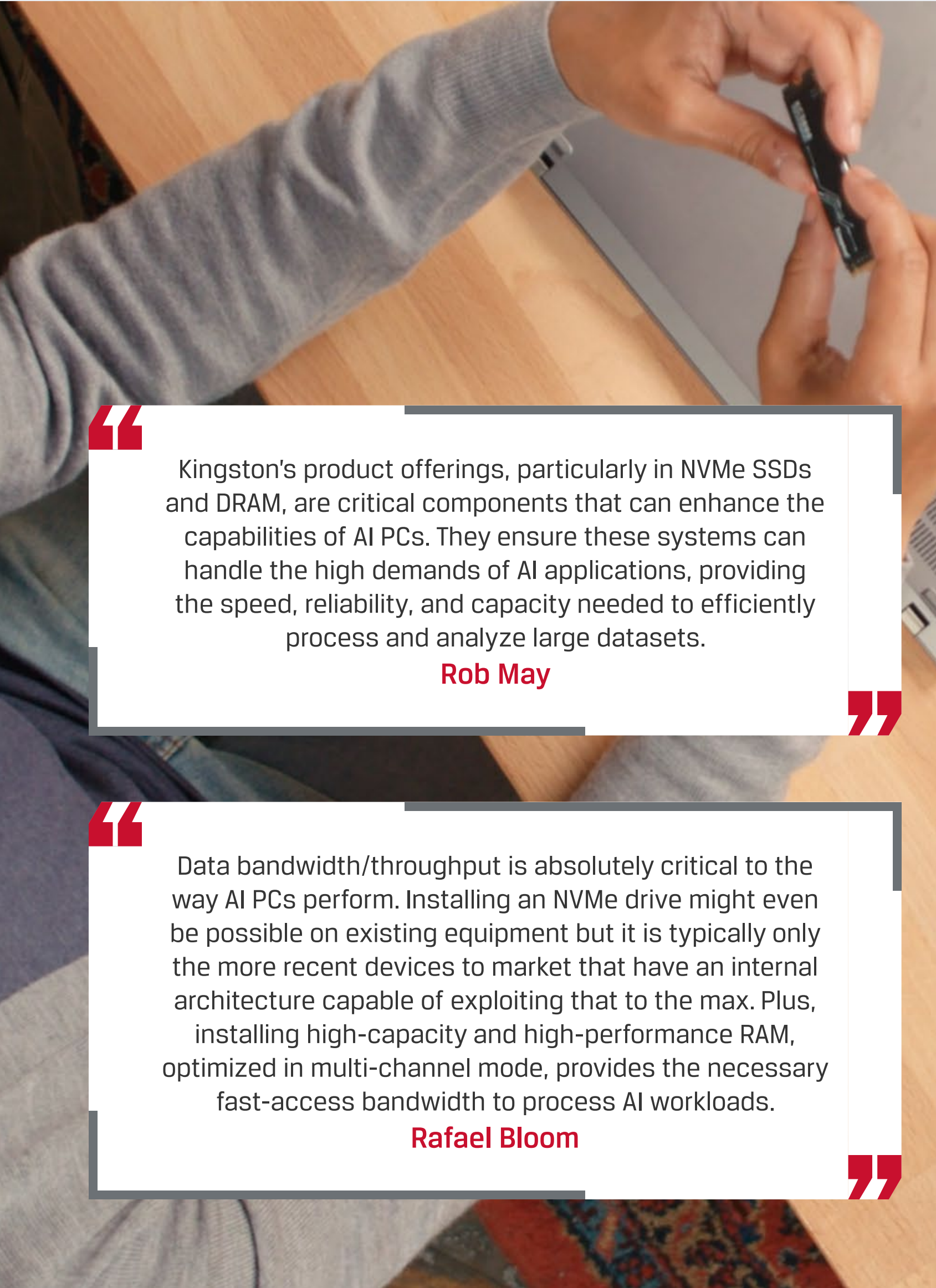
> " Kingston's product offerings, particularly in NVMe SSDs and DRAM, are critical components that can enhance the capabilities of AI PCs. They ensure these systems can handle the high demands of AI applications, providing the speed, reliability, and capacity needed to efficiently process and analyze large datasets.
> **Rob May** "

> " Data bandwidth/throughput is absolutely critical to the way AI PCs perform. Installing an NVMe drive might even be possible on existing equipment but it is typically only the more recent devices to market that have an internal architecture capable of exploiting that to the max. Plus, installing high-capacity and high-performance RAM, optimized in multi-channel mode, provides the necessary fast-access bandwidth to process AI workloads.
> **Rafael Bloom** "

Scalability is another critical factor for AI PCs, especially given the rapidly evolving nature of AI technology. As AI applications continue to grow in complexity and capability, the hardware supporting these applications must also adapt to meet increasing demands. Upgrading the storage, particularly to faster SSDs or larger capacities, means AI applications can run more efficiently. Larger and faster storage ensures that data-heavy AI operations, such as machine learning and data analytics, are not bottlenecked by data retrieval speeds.

The type of SSD used in AI PCs, particularly the adoption of NVMe (Non-Volatile Memory Express) drives over traditional SATA SSDs or HDDs (Hard Disk Drives), can significantly impact the performance and efficiency of these systems, especially in the context of AI and machine learning workloads. NVMe drives are designed to exploit the full potential of high-speed PCIe (Peripheral Component Interconnect Express) lanes, offering much faster data transfer rates compared to SATA drives, which are limited by the SATA interface's lower bandwidth.

Kingston Technology offers a wide range of memory and storage solutions that are particularly relevant and beneficial for AI PCs. Kingston's DDR5 memory solutions

provide the high-speed data transfer rates needed to support AI calculations for the CPU, NPU, and GPU. And for storage, Kingston's NVMe SSDs are designed to provide high-speed data access and storage, essential for the intensive workloads typical in AI applications. These drives offer faster read and write speeds compared to SATA SSDs, which is crucial for handling large datasets and accelerating the data processing tasks necessary for training machine learning models or conducting real-time analytics.

Kingston SSDs are also designed with a focus on reliability and endurance, which are critical for AI PCs that operate continuously and handle large volumes of data transactions. They are built to withstand the rigors of non-stop environments, with technologies to extend the life of the SSDs through advanced wear-leveling and error correction code algorithms. Beyond SSDs, Kingston's range of DRAM solutions complements its storage offerings to optimize the overall performance of AI PCs.

Best Practice Guidance for AI PC Investments

If you're aiming to deploy AI capabilities effectively, ensuring that an AI PC is compatible with specific software tools and frameworks is crucial. This involves a thorough understanding of its software and hardware capabilities and requirements, such as the need for specific types of processors (GPUs or NPUs), and minimum requirements for memory (RAM) and storage.

You should also consider software dependencies, scalability and performance needs, vendor support and community resources, and testing and validation before finalizing your choice. As yet, neither major updates like Windows 11 24H2 nor details on Windows 12 have been released, so details about its impact on AI PCs are speculative. But ultimately, these could provide new opportunities and tools for organizations using AI PCs, improving their performance, security, and ease of use, provided they are prepared to adapt to and adopt the new OS versions efficiently.

> For organizations deploying AI PCs, investing in sufficient and scalable DRAM capacity is essential to ensure their systems can meet current and future demands, thus safeguarding their investment and enabling them to leverage the full potential of AI technologies.
> **Rob May**

> A dedicated AI PC couples the appropriate kind of AI-optimized chipset with modern DRAM's capabilities to perform vast, multi-dimensional data tasks with speed and without impacting any other tasks and applications running in parallel.
> **Rafael Bloom**

Summary

The advent of AI-optimized PCs marks a significant step forward in computational capabilities at the edge, offering organizations the ability to harness powerful AI tools directly on their local hardware. This not only enhances productivity across various functions, but it also opens up possibilities for innovative applications in specialized fields.

With the ability to drive efficiencies and new capabilities in diverse sectors, AI PCs have the potential to truly transform the way we work. And, as AI technology continues to advance and be integrated into various sectors, Kingston's solutions remain pivotal in supporting the growth and effectiveness of these powerful computing platforms.

We can accelerate your AI journey with industry leadership offering proven expertise and best practices, and the memory and storage solutions needed to take your next steps with confidence.

About Kingston

With over 35 years of experience, Kingston has the knowledge, agility, and longevity to enable both data centers and enterprises to respond to the challenges and opportunities presented by the emergence of AI, 5G, IoT, and edge computing.

#KingstonIsWithYou