Deeptech becomes sexy as digital and physical tech converges

2024 marks the dawn of a new industrial revolution fuelled by converging digital and physical technologies. Deeptech barriers are crumbling, and startups are poised to lay the foundation for a more intuitive and accessible future.

Generative AI leads the charge, with natural voice interfaces becoming the norm. Think ChatGPT-powered robots and machines seamlessly interacting with the physical world. This convergence will be a game-changer for industrial tech, promising democratised interfaces – meaning anyone can shape the digital-physical world, not just tech giants, and enhanced efficiency – imagine voice-controlled factories and AI-powered supply chains.

However, 2024 will also be a year of reckoning for Gen AI, as we leave behind peak hype and unrealistic expectations. Tourist investors who flocked to inflated valuations will see their bets fizzle, leading to bankruptcies and low-ball acquisitions. At the same time, a lot of value has already been created and true product-market fit winners who prove real, quantifiable ROI will continue to emerge, creating the next generation of tech monopolies. As a result, the AI compute race will continue to heat up, with billions set to be invested in AI infrastructure as both established hyperscalers and ambitious startups vie for dominance.

In short, 2024 is a year of disruption and opportunity. Deep tech is democratising the future while the AI gold rush climaxes.

Inferencing at the edge: The AI revolution starts where data happens

Data is exploding, and AI models need to react in real time. Cloud computing needs to catch up. 2024 is the year inferencing at the edge takes off. What does that mean?

Local smarts: Instead of sending data to the cloud, AI models "live" on devices, analysing data on the spot. Think cameras reacting intelligently in real-time to information or faster robotic manipulation.

Fine-tuning on the fly: Edge models adapt to specific situations, making them more accurate and efficient—no more waiting for updates from the cloud.

Resourceful revolution: Edge devices have limited power, so AI models must be lean and mean. This opens the door for innovative startups specialising in efficient AI for the edge.

Why is this important? Inferencing at the edge unlocks the true potential of AI, driving regenerative AI and efficient edge models that minimise energy consumption, creating a virtuous cycle of data-driven improvement and new industrial applications, which provide real-time insights at the factory floor that translate to increased efficiency and productivity.

So, who's winning this race? Big tech like Amazon will provide the infrastructure, but the real action is at the edge. Innovative startups with efficient, adaptable AI models are poised to dominate. Big tech will likely snap them up, recognising the edge as the true frontier of AI. 2024 is the year AI moves beyond the cloud, and startups with their boots on the ground are ready to lead the charge.

Nvidia's throne under threat in 2024, as the AI chip landscape heats up

In 2024, we will see an avalanche of news related to the following:

New AI chips: big tech companies like Amazon, Google, and others are rolling out their own purpose-built AI chips, aiming to dethrone Nvidia's GPU dominance in training and running complex AI models.

Vertical integration: Tech giants vertically integrate, design, and manufacture their own AI chips to optimise cloud systems for AI workloads. This gives them greater control and potentially lowers costs.

Increased accessibility: With more players in the field, AI computing may become more readily available and affordable for a broader range of users beyond big tech.

Think of it as a replay of the automotive and semiconductor industries, but this time for AI in the cloud. The competition will be fierce, and 2024 is poised to be a pivotal year in the AI chip race.

Taming the Wild West: App Stores for Generative AI emerge

Right now, the Generative AI app world feels like the Wild West – chaotic and disorganised. But 2024 promises the rise of "app stores" or marketplaces that bring much-needed structure. These platforms will:

Organise the chaos: By curating and categorising various applications and foundational AI models, these platforms will make it easier for businesses to find and leverage relevant tools.

Increase accessibility: Like app stores democratised smartphones, these marketplaces will make powerful AI tools more accessible to smaller companies and startups.

Offer flexibility: Users can mix and match apps and models, creating custom solutions tailored to their specific needs.

Startups starting to fill this gap are likely to be highly attractive to investors, recognising the immense potential of a well-organised AI app ecosystem. While

initial platforms might offer broader categories like "Industrial Tech," the trend will likely move towards vertical specialisation, with dedicated marketplaces for specific industries catering to the unique needs of sectors like X or Y. The wild west of Generative AI apps is about to get a sheriff – AI app stores will bring order and accessibility, creating a gold rush for startups who can tame the frontier.

Christian Noske is a Partner at *NGP Capital*.

Article by CHRISTIAN NOSKE