Flash | Citi Research

01 Jun 2026 16:17:18 ET | 8 pages

Japan IT services

NVIDIA Computex keynote takeaways: AI ROI and physical AI initiatives

Moriya Koketsu[AC]

CITI'S TAKE
We summarize the key takeaways from NVIDIA's Computex Keynote. Key points for the IT services sector include the following: 1) With the transition to the Vera Rubin generation and accelerating enterprise AI adoption, the AI phase is shifting from training → inference → agents. NVIDIA stated that considering costs, enterprises need to design their own dedicated agents at low cost and high efficiency to adopt agents. The solution NVIDIA announced is the Nemotron 3 Ultra model, emphasizing the importance of affordable and fast open models. The implication is that SIer operations may transform from system development → enterprise AI agent design, management, and security. 2) NVIDIA is strengthening its physical AI initiatives through models like Cosmos 3, and we will be closely monitoring developments at partner Fujitsu.

- **AI and ROI** – A critical recent issue for IT services is AI ROI. Uber's COO stated that it remains unclear whether productivity gains sufficiently justify AI token costs, while Amazon is shifting from maximizing token usage to optimizing token usage. In fact, costs for OpenAI API and other API are rising as model performance improves. While this is unavoidable due to factors like computing power price increases being passed through, some enterprises actually paying these costs can no longer justify them. However, NVIDIA's announced Nemotron 3 Ultra addresses this ROI issue. The significance of NVIDIA releasing Nemotron 3 Ultra as an open model – a cost-effective and excellent model for enterprises developing and using agents – is substantial. NVIDIA stated it will release Nemotron 4 and 5 at a later stage, and we want to track NVIDIA's position as an AI model provider to enterprises. Additionally, the current focus on optimizing model usage will become a critical issue as enterprises adopt AI agents. Which models to use for which tasks will likely be determined by task difficulty and importance, and this work will likely become SIer business in the future. In other words, enterprises will need to differentiate usage, and we see SIers providing this support. Software companies mentioned as building agent AI based on NVIDIA include ServiceNow, Palantir, SAP, and CrowdStrike.

- **Physical AI** – NVIDIA's Keynote covered many topics, including the company's CPU expansion into PC business and the start of Vera Rubin full production. Among these, NVIDIA concluded by highlighting its physical AI initiatives. The company appears focused on physical AI as the next phase after agent AI, announcing initiatives related to robotics and autonomous driving. Among SIers, we see direct implications for Fujitsu, which has a business partnership with NVIDIA. In physical AI, robots will likely use rule-based software for specific task processing like agent AI (for example, when performing tax calculations, using rule-based software is more efficient than having AI verify tax calculation methods each time). Therefore, we expect interest to grow in Fujitsu, which collaborates with NVIDIA and possesses business-specific software, data, and expertise, as physical AI advances (see Fujitsu (6702.T) – Physical AI expectations).

See Appendix A-1 for Analyst Certification, Important Disclosures and Research Analyst Affiliati