



By: TA Analysis

Who decides what constitutes dangerous artificial intelligence: the state or Nvidia?



Jensen Huang, the founder and chief executive of Nvidia, the world's leading artificial-intelligence chipmaker, said in an **interview with Bloomberg** Television on 14 October that Nvidia would suspend deliveries to any customer whose use of artificial intelligence would harm people.

He was responding to a hypothetical question and added that the decision would be preceded by a conversation with the customer, acknowledging that there are limits to a supplier's right to impose its values on others.

Nvidia generated \$89 billion in revenue from data-centre equipment alone in the quarter ended 26 July, and its chips power most of the world's artificial-intelligence systems.

Any customer denied deliveries would struggle to find a replacement in the quantities required, so such a decision could effectively halt its development.

Nvidia is now linking the sale of its computing equipment to the development of safety tools and support for researchers examining AI safety.

Huang's willingness to judge the harms caused by the use of its chips would extend that influence to the selection of eligible customers.

The company would therefore independently decide under which conditions the technology may spread, while its director claims that the industry does not need new laws.

China's test of Huang's promise

China is the strongest test of Huang's promise, as there is already a case there in which the state concluded that the sale of Nvidia's chips posed a danger.

Washington has been restricting exports of the most advanced chips to China since 2022, fearing that they will be used to strengthen its military and to develop its own artificial

intelligence capabilities.

In May 2025, Huang called the controls a failure, arguing that they merely help China's competitors grow stronger.

His argument is supported by the latest data from research company IDC, which shows that Nvidia's share of the **Chinese market for artificial intelligence** chips fell in the first half of 2026 to 34 per cent, while China's Huawei reached 28 per cent.

Meanwhile, Washington relaxed but maintained controls. On 8 December 2025, President Donald Trump announced that Nvidia could sell H200 chips, the previous generation of its most powerful processors, to selected Chinese customers.

On 13 January 2026, the **Department of Commerce** mandated that each request be considered individually, with security requirements and mandatory chip testing in the United States.

Nvidia received the first licence, for limited quantities and specified customers, in February.

China shows how Nvidia reacts when the threat assessment comes from outside and starts to cost it money

The chips are manufactured in Taiwan, so before being shipped to China they must pass through the United States, where they are subject to a 25 per cent tariff. The US government thus takes a quarter of the value of each sale.

Beijing has not made it easy either, restricting domestic companies from buying American chips to encourage Huawei and other Chinese manufacturers.

That is why Nvidia delivered only a fraction of the approved quantities by the end of July, and those sales brought in less than 1 per cent of its data-centre equipment revenue.

Restrictions cost Nvidia money even when they are relaxed. In the same report, the company took a \$400 million write-down because of an inventory of H200 chips that it could not sell.

When the US government banned the sale of **H20 chips**, a model made specifically for the Chinese market, in April 2025, Nvidia was left with stock it could no longer sell and recorded a loss of \$4.5 billion.

Export controls concern the military and technological power of another country, and Huang's October pledge concerns the harmful use of products, so the two cases are not quite the same.

China, however, shows how Nvidia reacts when the threat assessment comes from outside and starts to cost it money.

The company has the right to challenge government policy, but the argument that the restrictions do not work simultaneously protects its access to one of the world's largest markets.

Huang's announcement therefore carries no real weight until it is clear how much loss Nvidia would accept if the harmfulness of using its chips were decided by someone outside the company.

Engineering solution and political dispute

On 15 September, at the Dreamforce conference in San Francisco, the annual gathering of software company Salesforce, **Huang said** the industry "doesn't need new laws or new regulations" and that AI security is an "engineering problem".

Pope Leo XIV responded on 28 September at a press conference on the plane, on his way back from France.

He said he had read Huang's claim that he had found a way to place technical limits on

artificial intelligence and pointed out that this was the same man who said it did not need any restrictions or government regulation. Nvidia presented precisely that **control system** on the same day.

The Pope's remark draws on the encyclical **Magnifica Humanitas**, his letter on artificial intelligence published in May.

In it, he warns that decision-making is being concentrated in the hands of a small number of technology companies and insists that affected communities have a say. He specifically calls for legal frameworks and independent oversight.

The dispute with Huang thus centres on who sets the limits of artificial intelligence: the companies that create it, or the society that lives with its consequences.

Huang does not reject all state control, however. In a conversation with American journalist Ezra Klein at the end of September, he said that AI companies should not be exempt from competition laws or from responsibility for damage caused by their products.

He also supported special rules where needed, for example for autonomous cars.

The more secure agents become, the more companies will deploy them, and most of these agents will run on Nvidia chips

He draws the line at the models themselves, insisting that companies should decide whether a new artificial intelligence system is safe enough to enter the market.

Nvidia's system is designed for so-called AI agents, programmes that autonomously perform a range of tasks, from sending emails to working with confidential company data.

The OpenShell programme determines which data and systems the agent is allowed to access, with rules set by the company that

uses it.

Sentry, a dedicated hardware watchdog, monitors the agent's operation and, according to Nvidia, can stop it within milliseconds if it tries to exceed the permitted limits.

The more secure agents become, the more companies will deploy them, and most of these agents will run on Nvidia chips. Protection is therefore a new way for Nvidia to sell more equipment.

Part of that offering is open. The OpenShell is publicly available and can be adapted to third-party equipment.

Sentry, however, runs only on Nvidia's BlueField-4 hardware. A company that builds protection around both tools will be able to move software easily, but not monitoring, so switching to another supplier will become increasingly expensive as it relies more on Nvidia's equipment.

UK support returns to Nvidia's system

Nvidia announced that it will provide **British artificial intelligence security** researchers with computing capacity worth an estimated \$250 million, or about \$50 million per year, over the next five years.

That funding will remain within Nvidia's system, as universities gain the right to use the computers of its British partners, equipped with its chips.

Among the first users will be the BOLD laboratory at Oxford and the SOFAIR laboratory, UK national AI research laboratory.

The announcement explicitly states that researchers will also work with the OpenShell and Sentry tools.

In the coming years, British security experts will, at Nvidia's expense, examine and improve the very protections that Nvidia sells to companies

In the coming years, British security experts will, at Nvidia's expense, examine and improve the very protections that Nvidia sells to companies.

Universities will receive equipment for costly experiments that they could hardly fund themselves, and Nvidia will gain the opportunity to establish its solutions as a common benchmark for assessing the safety of artificial intelligence.

Key terms have not yet been agreed. Nvidia has not announced which partners will provide the services, how the computers will be organised, under what commercial terms, or how much individual universities will receive.

Those contracts will also depend on whether researchers are allowed to freely publish findings that are unfavourable to the donor.

Part of the programme could also work against Nvidia's interests. The SOFAIR laboratory investigates new ways of building artificial intelligence, with the aim that development does not depend on a few large companies.

Nvidia's funding could therefore support technologies that may one day reduce dependence on its products, provided the results can be used on other organisations' equipment.

The British package seems more modest than the American one. On 8 October, Nvidia announced \$1 billion over five years for US science, health, energy and advanced computing research – four times more than for Britain.

However, Britain is not tied to a single supplier. Nvidia's main competitor, AMD, announced in June that it would invest up to two billion pounds over five years in British

computing infrastructure and scientific collaboration, and it is also participating in Oxford's BOLD laboratory.

This freedom of choice has its price, because every programme written for one set of equipment must be adapted before switching to another, and that price increases with each year of operation on the same chips.

State verification with industry consent

The British government has its own institute to evaluate artificial intelligence, but it has no authority to examine systems without the consent of the companies that develop them.

The **AI Security Institute** was founded in November 2023, at the first world summit on artificial intelligence safety at Bletchley Park, under the name AI Safety Institute.

The name was changed in February 2025, when the government shifted its focus to national security threats and criminal misuse.

Access to the most advanced systems still depends on the voluntary cooperation of companies, as the government reaffirmed in its reply to Parliament on 12 October.

Minister for Artificial Intelligence Kanishka Narayan announced on the same day that the institute is testing the latest models.

Access was secured only after discussions with the laboratories and the US government, so that British experts can check the systems before they go on the market.

Narayan also stated that the government is considering stronger protection through cyber-security laws.



The state is cultivating experts who know how to assess artificial intelligence, but what they will examine still depends on what the companies choose to show them

The state is thus cultivating experts who know how to assess artificial intelligence, but what they will examine still depends on what the companies choose to show them.

The question has also come before MPs. On 13 October, the parliamentary committee first heard from former employees of leading artificial intelligence companies, and then from representatives of current management and state institutions.

MPs were therefore able to compare what those who left those companies say about safety with what those who run them claim.

In the coming years, Nvidia is likely to be simultaneously the largest seller of artificial-intelligence equipment, a manufacturer of tools to control it, and one of the main funders of British research into its safety.

For the time being, this suits the British government, because it gains access to equipment that it could hardly afford on its own.

The price of this arrangement is increasing dependence on suppliers whose products British researchers need to assess.

The real test will come with the first serious disagreement: when British researchers conclude that a system is dangerous, but Nvidia or its customers judge existing protections sufficient.

At that point, it will become clear whether the researchers are allowed to publish their findings and whether their access to the equipment remains open afterwards.

China has already shown that the limits of acceptable risk shift when money and politics are brought into the equation.

Huang wants safety to be decided by those who make the technology, and Britain will soon discover the value of a government assessment that relies on the same company's equipment.