UK, US, EU and China sign declaration of AI's 'catastrophic' danger

The UK, US, EU, Australia and China have all agreed that artificial intelligence poses a potentially catastrophic risk to humanity, in the first international declaration to deal with the fast-emerging technology. Twenty-eight governments signed up to the so-called Bletchley declaration on the first day of the AI safety summit, hosted by the British government. The countries agreed to work together on AI safety research, even amid signs that the US and UK are competing to take the lead over developing new regulations. Rishi Sunak welcomed the declaration, calling it "quite incredible". In remarks ahead of his own appearance at the summit on Thursday, the prime minister added: "There will be nothing more transformative to the futures of our children and grandchildren than technological advances like AI. "We owe it to them to ensure AI develops in a safe and responsible way, gripping the risks it poses early enough in the process." Referring to the risks posed by the most advanced AI systems, the declaration stated: "There is potential for serious, even catastrophic, harm, either deliberate or unintentional, stemming from the most significant capabilities of these AI models." Michelle Donelan, the UK technology secretary, told reporters: "For the first time we now have countries agreeing that we need to look not just independently but collectively at the risks around frontier AI." Frontier AI refers to the most cutting-edge systems, which some experts believe could become more intelligent than people at a range of tasks. Speaking to the PA news agency on the sidelines of the summit, Elon Musk, the owner Tesla and SpaceX, and of X, formerly Twitter, warned: "For the first time, we have a situation where there's something that is going to be far smarter than the smartest human … it's not clear to me we can actually control such a thing." The communique marks a diplomatic success for the UK and for Sunak in particular, who decided to host the summit this summer after becoming concerned with the way in which AI models were advancing rapidly without oversight. Donelan opened the summit by telling her fellow participants that the development of AI "can't be left to chance or neglect or to private actors alone". She was joined onstage by the US commerce secretary, Gina Raimondo, and the Chinese vice-minister of science and technology, Wu Zhaohui, in a rare show of global unity. Matt Clifford, one of the British officials in charge of organising the summit, called the appearance of Raimondo and Wu together on stage "a remarkable moment". China signed the declaration, which included the sentence: "We welcome the international community's efforts so far to cooperate on AI to promote inclusive economic growth, sustainable development and innovation, to protect human rights and fundamental freedoms, and to foster public trust and confidence in AI systems to fully realise their potential." Wu told fellow delegates: "We uphold the principles of mutual respect, equality and mutual benefits. Countries regardless of their size and scale have equal rights to develop and use AI." South Korea has agreed to host another such summit in six months' time, while

France will host one in a year. So far, however, there is little international agreement over what a global set of AI regulations might look like or who should draw them up. Some British officials had hoped other countries would agree to beef up the government's AI taskforce so that it could be used to test new models from around the world before they are released to the public. Instead, Raimondo used the summit to announce a separate American AI Safety Institute within thecountry's National Institute of Standards and Technology, which she called "a neutral third party to develop best-in-class standards", adding that the institute would develop its own rules for safety, security and testing. Earlier this week, the Biden administration released an executive orderrequiring US AI companies such as OpenAI and Google to share their safety test results with the government before releasing AI models. Kamala Harris, the vice-president, then gave a speech on AI in London in which she talked about the importance of regulating existing AI models as well as more advanced ones in the future. Clifford denied any suggestion of a split between the US and UK on which country should take the global lead on AI regulation. "You'll have heard Secretary Raimondo really praise us in a full-throated way and talk about the partnership that she wants to have between the UK and the US safety institute," he said. "I really think that that shows the depth of the partnership." Sunak said the summit had proved "the appetite from all of those people for the UK to take a leadership role". The EU is in the process of passing an AI bill, which aims to develop a set of principles for regulation, as well as bringing in rules for specific technologies such as live facial recognition. Donelan suggested the government would not include an AI bill in the king's speech next week, saying: "We need to properly understand the problem before we apply the solutions." But she denied the UK was falling behind its international counterparts, adding: "We have called the world together – the first ever global summit on AI at the frontier – and we shouldn't minimise or overlook that."