

TRENDING

Are we in a generative AI bubble?

Join the ITPro Network

How to get a grip on sustainability

Cyber security tip:

When you purchase through links on our site, we may earn an affiliate commission. [Here's how it works.](#)

Business > Policy and Legislation

Bletchley Declaration draws cautious approval

By [Emma Woollacott](#) published about 21 hours ago

AI safety agreement garners international support, but some believe it needs more specificity



(Image credit: Getty Images)

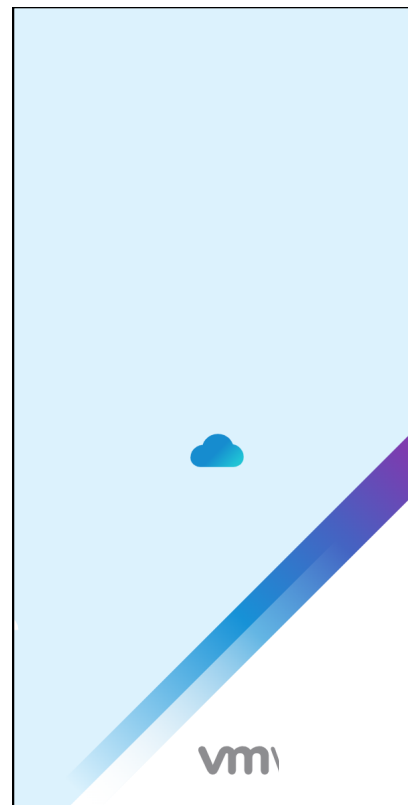
The UK, US, European Union (EU), India and China are amongst 28 countries and governments to sign up to the Bletchley Declaration on AI safety, announced on 1 November at the UK's AI Summit.

The agreement recognises that so-called frontier AI technology, such as generative AI chatbots and image generators, has "potential for serious, even catastrophic, harm, either deliberate or unintentional". It outlines three key areas of risk: Cybersecurity, biotechnology, and disinformation.

It says risks are best addressed through international cooperation, with participating nations agreeing to work together to support a network of scientific research on AI safety.

"This is a landmark achievement that sees the world's greatest AI powers agree on the urgency behind understanding the risks of AI – helping ensure

Advertisement



Advertisement

the long-term future of our children and grandchildren," said UK prime minister Rishi Sunak.

RECOMMENDED VIDEOS FOR YOU...

The Most Common Password-cracking Techni...

We detail some of the most common, and most effective methods hack...

PLAY SOUND

00:10 / 02:58

"The UK is once again leading the world at the forefront of this new technological frontier by kickstarting this conversation, which will see us work together to make AI safe and realise all its benefits for generations to come."

The agreement has received tentative approval from the UK tech sector, with Rashik Parmar, CEO of BCS, saying that it takes a more positive view of the potential of AI than expected.

"I'm also pleased to see a focus on AI issues that are a problem today – particularly disinformation, which could result in personalised fake news during the next election – we believe this is more pressing than speculation about existential risk," he added.

"The emphasis on global cooperation is vital to minimise differences in how countries regulate AI."

Amanda Brock, CEO of OpenUK, said the agreement will need hard practicalities baked in, and is calling for the open source software and open data business communities to be included.

RELATED RESOURCE



Analyzing the economic benefits of Trend Micro Vision One
Discover how the XDR Workbench enables alert prioritization based on security scores

■ DOWNLOAD NOW

"Their input will help to appropriately shape the transparency called for in the Declaration alongside the internationally inclusive, collaborative cross border research and innovation and international co-operation in an open manner," she said.

The advertisement features a light blue background with a purple-to-blue diagonal gradient at the bottom. It includes the headline 'Apps and clouds, paired perfectly.', the sub-headline 'Maximise the value of multi-cloud with VMware.', a 'LEARN MORE' button, and the slogan 'The smarter way to cloud™'. The graphic shows stylized clouds with server icons and a network diagram.

MOST READ



1 How mixed reality is transforming city and urban planning

2 Conquering technology risk in banking

3 Banking Dive: How hyperautomation can bolster security and reduce risk in banking

4 Customer insights: Book of knowledge 23

5 Surging cyber threats exacerbating security staff burnout

Advertisement

The advertisement features the text 'Go faster. Spend less.' on the left and the VMware logo on the right, set against a light blue background with a white diagonal line.

Joseph Thacker, researcher at SaaS security firm AppOmni, was less confident, however.

"The biggest challenge is that the open source ecosystem is really close to enterprises when it comes to making frontier AI," he said.

"And the open source ecosystem isn't going to adhere to these guidelines – developers in their basement aren't going to be fully transparent with their respective governments."

The Bletchley Declaration isn't the only game in town, though. US secretary of commerce Gina Raimondo used the summit to announce a new US-led international AI safety institute.

Just two days before the conference kicked off, on 30 October, US president Joe Biden signed an executive order, described by the White House as "the most significant actions ever taken by any government to advance the field of AI safety", requiring AI firms to report the risks their technology could pose.

The EU is already in the process of passing its AI Act and, on the same day Biden signed his executive order, the G7 agreed a set of guiding principles for AI and a voluntary code of conduct for AI developers.

The challenge, most agree, will be to ensure that future international efforts are focused on practical measures.

"The test will be seeing cooperation on the concrete governance steps that need to follow," said Seán Ó hÉigeartaigh, programme director of AI: futures and responsibility at the University of Cambridge.

"I hope to see the next summit in South Korea forge a consensus around international standards and monitoring, including establishing a shared understanding of which standards must be international, and which can be left to national discretion," he added.

Advertisement

Go faster.
Spend less.
Be free.
Make the most of
multi-cloud
with VMware.

The smarter way.
cloud

vn

Click Here for More Infor

Ad By Sponsor

Get the ITPro. daily newsletter

Receive our latest news, industry updates, featured resources and more. Sign up today to receive our FREE report on AI cyber crime & security - newly updated for 2023.

SIGN ME UP

- Contact me with news and offers from other Future brands
- Receive email from us on behalf of our trusted partners or sponsors

By submitting your information you agree to the [Terms & Conditions](#) and [Privacy Policy](#) and are aged 16 or over.



Emma Woollacott  

Emma Woollacott is a freelance journalist writing for publications including the BBC, Private Eye, Forbes, Raconteur and specialist technology titles.

MORE ABOUT POLICY AND LEGISLATION

LATEST





G7 outlines "innovation friendly" AI code of conduct ▶



UK needs "regulatory certainty" to become global AI leader ▶



How mixed reality is transforming city and urban planning ▶

[SEE MORE LATEST ▶](#)

TOPICS

ARTIFICIAL INTELLIGENCE REGULATION

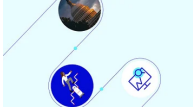
SPONSORED FINANCIAL CONTENT

dianomi



Whisky Casks: Diversify With A Tax-Free Investment. Get Free Guide

VCL Vintners



Big headlines: client-friendly infographic for your market updates.

Invesco



Looking to stabilise your portfolio with a Tax-Free asset? Get Guide

Hackstons



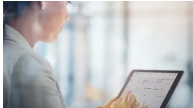
Explore opportunities in the digital future now. Capital at risk

For professional clients only



Five high-profile shares I am avoiding - Free report

Wealth Club Limited



LYNQS: manage your own investment strategy

Leonteq - [Disclaimer]



City Index: Spread bets & CFDs. 70% of retail CFD accounts lose money.
CITY INDEX



Quality Dividend Growth: the 'all-weather' core for a portfolio

WisdomTree

MOST POPULAR



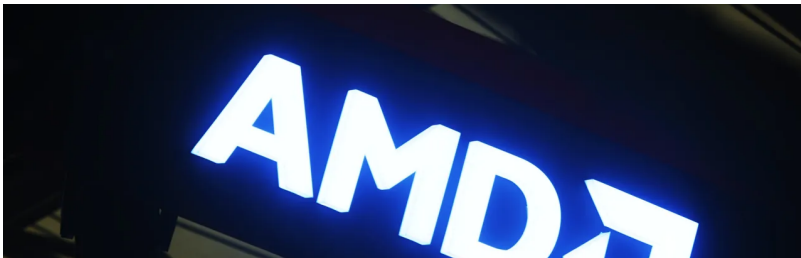
So long WeWork, it's been a blast

By Ross Kelly 1 November 2023



Windows 11 23H2 update brings full AI Copilot integration

By Solomon Klappholz 1 November 2023



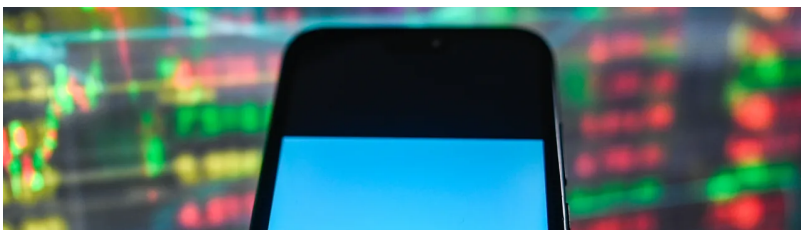
Why AMD could struggle to topple Nvidia's AI lead

By Rory Bathgate 1 November 2023



Warning issued over "inadvertent" spread of malvertising on Google Dynamic Search Ads

By Solomon Klappholz 1 November 2023



Logitech concludes CEO search with appointment of Hanneke Faber

By Daniel Todd 1 November 2023



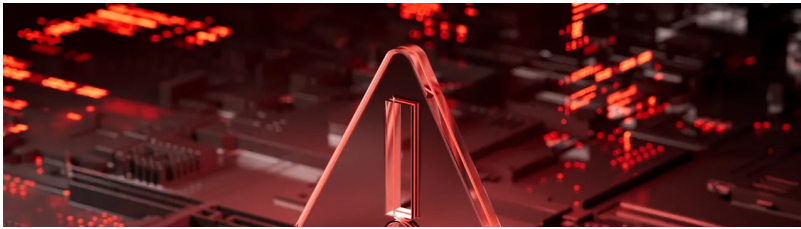
Can the UK's AI Safety Summit succeed?

By Rory Bathgate 31 October 2023



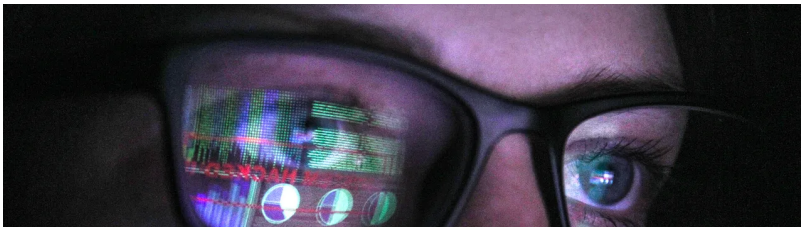
£118 million training package to 'future-proof' UK's AI skills

By Solomon Klappholz 31 October 2023



US-led coalition of nations agrees to end ransomware payments to hackers

By Ross Kelly 31 October 2023



Firms still feeling strain of cyber skills deficit despite workforce surge

By Solomon Klappholz 31 October 2023



Vantage Data Centers snags €2.5 billion investment to fuel EMEA growth

By Daniel Todd 31 October 2023



Harmful effects of “cloud concentration” now a key concern for IT leaders

By Ross Kelly 31 October 2023

RESOURCES

WHITEPAPER



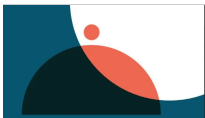
The ultimate guide to the best restaurant POS systems

Posted



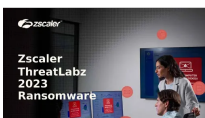
The threat prevention buyer's guide

Posted



The state of restaurants in 2023

Posted



2023 ThreatLabz state of

**ransomware
report**

Posted

[VIEW MORE WHITEPAPERS](#)

IT Pro is part of Future plc, an international media group and leading digital publisher. [Visit our corporate site.](#)

[Terms and conditions](#)

[Contact Future's experts](#)

[Privacy policy](#)

[Cookies policy](#)

[Accessibility statement](#)

[Careers](#)

[About Us](#)

[Contact Us](#)

[GDPR consent](#)

© Future Publishing Limited Quay House, The Ambury, Bath BA1 1UA. All rights reserved. England and Wales company registration number 2008885.