

Embracing Ethical AI

Monoceros looks at the Bletchley Park Declaration

1. What is the Bletchley Park Declaration?

On 1 November 2023, the [Bletchley Park Declaration](#) (the “Declaration”) was signed at the AI Safety Summit. The Declaration is a collective pledge by 28 leading AI nations to create a framework of ethical AI standards in [artificial intelligence](#) (“AI”) development.

It champions transparency, fairness, accountability, and human rights for responsible AI innovation. It also warns about the *“potential for serious, even catastrophic, harm, either deliberate or unintentional, stemming from the most significant capabilities of these AI models.”* These risks extend beyond “frontier AI” to data bias and data privacy.

2. Key points:

The Declaration addresses:

- **Frontier AI risks:** The declaration acknowledges the potential benefits and challenges of frontier AI, and the need for international cooperation to mitigate the risks posed by AI systems that may be harmful, unpredictable, or misaligned with human values.
- **AI safety principles:** The declaration outlines some general principles for AI safety, such as ensuring human oversight, transparency, accountability, and ethical standards. The declaration also calls for fostering public awareness and education on AI safety issues.
- **AI safety research:** The declaration supports the advancement of scientific research on AI safety, and the sharing of best practices and data among countries. The declaration also encourages the development of technical tools and methods to ensure the safety and reliability of AI systems.
- **AI safety governance:** The declaration proposes the establishment of a global AI safety forum, which will be a platform for dialogue and collaboration among countries, experts, and stakeholders on AI safety matters. The declaration also urges the government and the industry to work together to create appropriate regulations and policies for AI safety.

3. Are there other International AI Guidelines?

Yes, in October 2023 in the United States of America, the Biden-Harris Administration issued a landmark [Executive Order on AI](#). The Order establishes new standards for AI safety and security, seeks to protect Americans’ privacy, seeks to advance equity and civil rights, consumer and worker protections and promote innovation.

The European Union also issued (and subsequently updated) a draft **AI Act** that proposes to regulate the use of AI within the EU and establish regulatory principles within a risk-based framework. The draft act focuses on rules around data quality, transparency, human oversight and accountability and addresses ethical matters. The June 2023 amendments seek a ban on the use of AI technology in biometric surveillance and for generative AI systems to disclose AI-generated content.

4. For more see

- <https://www.gov.uk/government/news/countries-agree-to-safe-and-responsible-development-of-frontier-ai-in-landmark-bletchley-declaration>
- <https://www.whitehouse.gov/briefing-room/statements-releases/2023/10/30/fact-sheet-president-biden-issues-executive-order-on-safe-secure-and-trustworthy-artificial-intelligence/>
- <https://www.weforum.org/agenda/2023/06/european-union-ai-act-explained/>

Contact us

For more information about:

- data ethics, governance and privacy
- how AI can transform your business
- data trusts

contact Monoceros Innovation at: Info@monocerosinnovation.com



Emma German
Founder, Director



Monoceros Innovation Advisory Limited

Monoceros Innovation Advisory Limited is registered as a limited company in Jersey with registered number 133604, of Les Marias, La Route du Marais, St Ouen Jersey JE3 2GG, carrying on innovation services.

©Monoceros Innovation Advisory Limited 2023

All content featured on the Monoceros Innovation website, including text, graphics, logos, images, audio clips, digital downloads, and software, is the property of Monoceros Innovation Advisory Limited and is protected by international copyright laws.