AI Safety Summit 2023

Why in news?

Recently, the World's 1st Artificial Intelligence (AI) Summit was held at Bletchley Park, a historic site in UK where the Enigma code was cracked during World War II.

Bletchley Park- The Birthplace of Artificial Intelligence (AI)
• It is a historic site in England that was the headquarters of the *Government Code and Cypher School* during World War II. • It is considered the birthplace of modern computing and artificial intelligence. • **Enigma code** - Bletchley Park is most known for cracking the Enigma code, the *complex cipher system* used by the Nazis during World War II to encrypt their radio messages • It was cracked by ***Alan Turing*** who was often considered as the *father of modern computer science.* • **Colossus**- A machine called Colossus was built at Bletchley Park to break the *Lorenz cipher*, used by the German High Command. • It was the world's first programmable, electronic, digital computer, but it was destroyed and kept secret after the War. • Bletchley Park houses the world's largest collection of working historic computers at the *National Museum of Computing.*

What is AI?

- Artificial intelligence (AI) is the simulation of human intelligence processes by machines, especially computer systems.
- Specific applications of AI include expert systems, natural language processing, speech recognition and machine vision.

What is the London Summit on AI?

- **AI safety summit**- A meeting of 28 nations, including the U.S. and China, to discuss the risks and benefits of AI and how to regulate its development and use.
- The summit is centred around '***frontier AI',*** which is defined as "highly capable foundation models that could possess dangerous capabilities sufficient to pose severe risks to public safety.
- **Aim**- To address the safety and ethical issues of AI, especially the advanced large language models (LLMs) that can generate natural language texts.
- **Two pronged agenda**-
    - Identifying risks of shared concern and building the scientific understanding of them
    - Building cross-country policies to mitigate the concerns
- **Key takeaways of the Summit**

- ***Bletchley Declaration*** was signed at the summit, pledging to work together to understand and manage the potential catastrophic effects of AI, especially the cutting-edge frontier AI that could threaten humanity's existence.
- The summit discussed the establishment of an *international register of frontier AI models* that will allow governments to assess the risks involved.
- The ***world's first AI Safety Institute*** has been launched in the ***UK,*** tasked with testing the safety of emerging types of AI.

What are the key highlights of Bletchley Declaration?

- **International Cooperation**- The declaration underscores the necessity for a united front in addressing the challenges and leveraging the opportunities that AI presents on a global stage.
- **Safety Standards**- It includes shared commitment to reducing risks associated with AI and ensuring that these technologies are developed with a safety-first approach.
- **Ethical AI**- The declaration ensures that AI technologies respect human rights, privacy, and democratic values, fostering a human-centric approach to AI.
- **Transparency and Accountability**- Both are seen as a cornerstone for building public trust and understanding, essential for the successful integration of AI technologies into society.
- **Knowledge Sharing**- It aims at accelerating the global understanding and mitigation of AI-related risks, promoting a culture of shared learning and continuous improvement in AI safety practices.

To know more about ethical artificial intelligence click here

What are the global regulatory landscapes in Artificial Intelligence?

- **USA**- US had recently released a *Blueprint for an AI Bill of Rights.*
- **European Union-** It is discussing the promulgation of world's 1st comprehensive framework for regulation, which will define rules to govern the development and use of AI across EU.

  EU plans to set European Board for Artificial Intelligence to audit and administer the new rules.

- **China**- It outlined a broad set of principles for international cooperation and global governance of AI.
- **India-** Niti Aayog published a series of papers on the subject of Responsible AI for All.
- **United Nations** - UN Secretary-General Antonio Gueterres has announced the creation of an advisory body that would address the international governance of AI with technology executives, government officials and academics.

References

1. Indian Express- London Summit and how to make AI responsible
2. The Hindu- UK AI summit to tackle catastrophic risks
3. Indian Express- World’s first AI summit
4. Image | AI

