# A new global standard for AI ethics

Context:
- Artificial intelligence (AI) is more present in all spheres of life, from predicting choice of using social media to helping one understand weather patterns to manage agriculture.
- AI algorithms can also be partially credited for the rapidity with which vaccines were developed to tackle COVID-19.
- The algorithms crunched complex data from clinical trials being undertaken in all corners of the world, creating global collaborations that could not have been imagined even a decade ago.
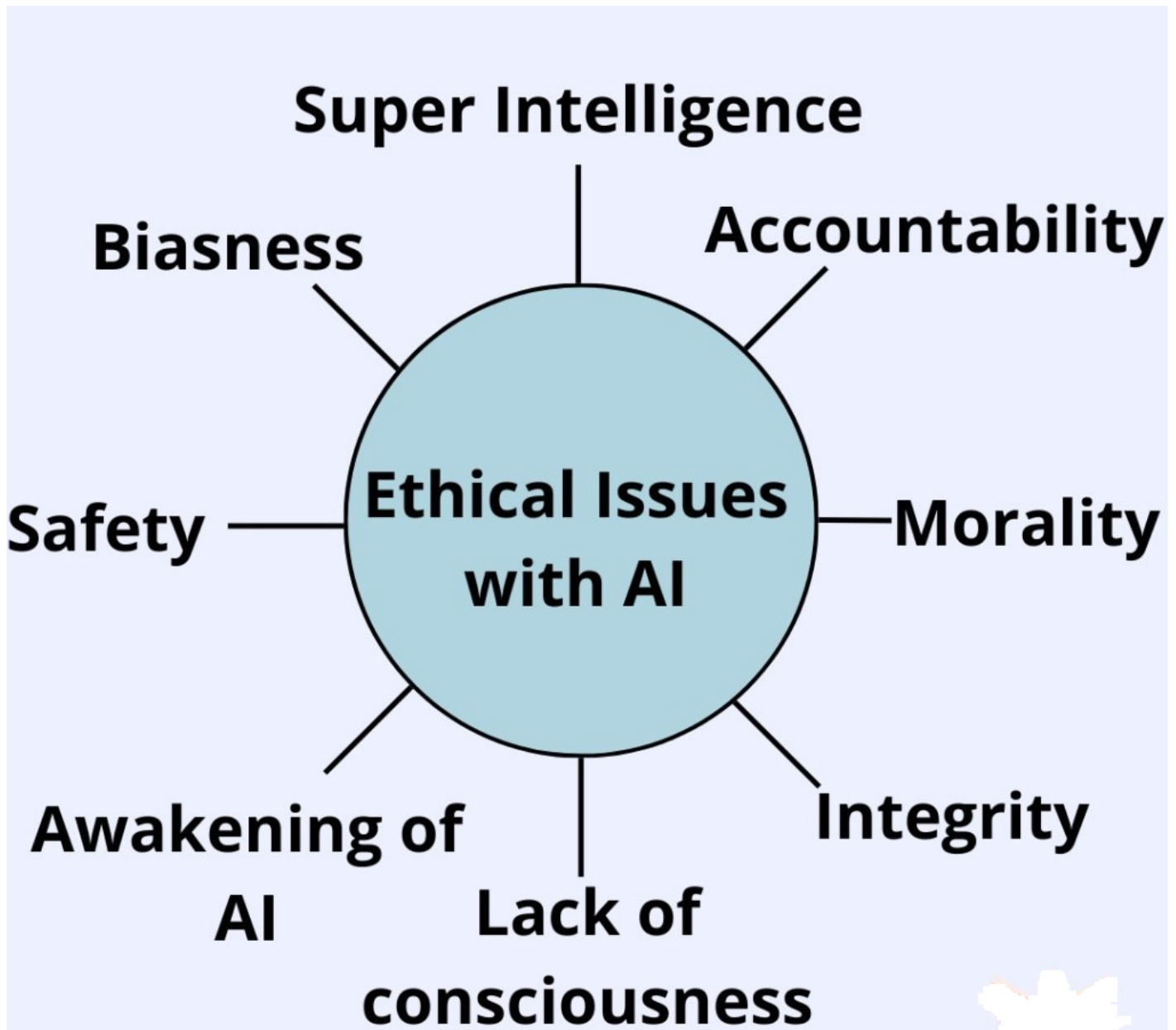
Issues in AI:
- But AI-related technology cannot be said to always be beneficial. The data used to feed into AI often aren't representative of the diversity of our societies, **producing outcomes that can be said to be biased or discriminatory.**

Biases in facial recognition technologies:

- For instance, while India and China together constitute approximately a third of the world's population, Google Brain estimated that they form just 3% of images used in ImageNet, a widely used dataset. Similarly, there are **problems emerging in facial recognition technologies**, which are **used to access phones, bank accounts and apartments**, and are increasingly **employed by law-enforcement authorities**, in identifying women and darker-skinned people.
- For three such programs released by major technology companies, **the error rate was 1% for light-skinned men, but 19% for dark-skinned men, and up to 35% for dark-skinned women**.
- Biases in facial recognition technologies have led to wrongful arrests.

Why it matters more in case of India?
- These issues are of particular importance to India, which is one of the **world's largest markets for AI-related technologies**, valued at over $7.8 billion in 2021.
- Indeed, the National Strategy on Artificial Intelligence released by NITI Aayog in 2018 highlights the massive potential of AI in solving complex social challenges faced by Indian citizens **across areas such as agriculture, health, and education**, in addition to the **significant economic returns** that AI-related technologies are already creating.

Ethical AI governance:
- To ensure that the full potential of these technologies is reached, the **right incentives for ethical AI governance** need to be established in national and sub-national policy.
- India has made great strides in the development of responsible and ethical AI governance, starting with NITI Aayog's #AIForAll campaign to the many corporate strategies that have been adopted to ensure that AI is developed with common, humanistic values at its core.

UNESCO's Recommendation on the Ethics of Artificial Intelligence:
- However, until recently, there was **no common global strategy to take forward this importance agenda**. This changed in November 2021, when 193 countries reached a groundbreaking agreement at UNESCO on how AI should be designed and used by governments and tech companies.
- UNESCO's Recommendation on the Ethics of Artificial Intelligence took two years to put together and involved thousands of online consultations with people from a diverse range of social groups.
- It aims to **fundamentally shift the balance of power between people, and the businesses and governments developing AI**.
- Countries which are members of UNESCO have agreed to implement this recommendation by enacting actions to regulate the entire AI system life cycle, ranging from research, design and development to deployment and use.
- This means they must use affirmative action to make sure that women and minority groups are fairly represented on AI design teams.

Data protection:
- The recommendation also underscores the importance of the **proper management of data, privacy and access to information.** It establishes the need to keep control over data in the hands of users, allowing them to access and delete information as needed.
- It also calls on member states to ensure that appropriate safeguards schemes are devised for the processing of sensitive data and effective accountability, and redress mechanisms are provided in the event of harm. All of this takes enforcement to the next level.

Addressing broader socio-cultural impacts of AI:
- Additionally, the **broader socio-cultural impacts of AI-related technologies** are also addressed, with the Recommendation taking a strong stance that AI systems should not be used for social scoring or mass surveillance purposes; that particular attention must be paid to the psychological and cognitive impact that these systems can have on children and young people; and that member states should invest in and promote not only digital, media and information literacy skills, but also socio-emotional and AI ethics skills to strengthen critical thinking and competencies in the digital era.

- This is all critical for ensuring accountability and transparency of AI-related technologies, underpinning a strong rule of law that adapts to new digital frontiers.

Way Forward:
- The new agreement is broad and ambitious. It is a recognition that AI-related technologies cannot continue to operate without a common rulebook.
- Over the coming months and years, the recommendation will serve as a compass to guide governments and companies, to voluntarily develop and deploy AI technologies that conform with the commonly agreed principles it establishes – similar moves happened after UNESCO's declaration on the human genome set out norms for genetic research.
- Second, it is hoped that **governments will themselves use the Recommendation as a framework to establish and update legislation**, regulatory frameworks, and policy to embed humanistic principles in enforceable accountability mechanisms.
- To accompany countries in the realisation of the full potential of AI and with the aim of building the institutional capacity of countries and all the relevant stakeholders, **UNESCO is in the process of developing tools to help them assess their readiness in the implementation** of the Recommendation and identify, monitor and assess the benefits, concerns and risks of AI system.