

Lesson Summary

Congratulations! You have completed this lesson. At this point in the lesson, you know:

- Ethical considerations and responsible use of AI refer to the principles, practices, and guidelines aimed at ensuring that artificial intelligence technologies are developed, deployed, and used in ways that are beneficial, fair, and respectful of human rights and societal values.
- The ethical considerations of AI include data privacy and security, bias and fairness, transparency and accountability, autonomous systems and human oversight, and access and equality.
- AI can be used responsibly by:
 - Using diverse datasets
 - Designing fully autonomous systems that include human-in-the-loop mechanisms
 - Ensuring AI's accessibility and benefits to all segments of society
 - Optimizing algorithms for energy efficiency
 - Utilizing renewable energy sources for data centers
- Generative AI involves the use of algorithms to create new content, such as text, images, audio, and video, presenting unique opportunities and challenges.
- Some considerations around using generative AI are:
 - Copyright and ownership: It is important for developers and users to familiarize themselves with legal frameworks and stay updated on evolving laws.
 - Privacy and confidentiality of user data: It is beneficial to use private AI models and a combination of technical and legal measures.
 - Deepfakes and misinformation: It is essential to think critically in ensuring the accuracy of generative AI outputs, and how it can mitigate AI flaws like hallucinations.
 - Ethical considerations: It is crucial for developers to prioritize fairness, accountability, and social benefits in AI development.

- Examining the perspectives of the key players can help one gain a comprehensive understanding of the diverse approaches to AI ethics, the challenges they aim to address, and the strategies they employ to promote the responsible use of AI technologies.
- IBM's pillars of trust are central to its approach to ensuring that AI systems are explainable, fair, robust, transparent, and privacy-preserving.
- Microsoft's approach to ethical AI involves human-in-the-loop systems, continuous monitoring and improvement, regular audits using the artificial intelligence act (AIA), internal review bodies, and responsible AI standards.
- Google's AI principles focus on benefiting society, ensuring fairness and accountability, and upholding high standards of scientific excellence.
- Premature deployment and adoption of AI systems could put the companies at a huge risk of reputational and financial loss.
- AI governance refers to a set of rules, standards, and processes that have been established to ensure the responsible and ethical development and deployment of artificial intelligence systems.
- Risks associated with AI:
 - Bias in AI models:
 - Human-generated data can contain hidden biases.
 - AI models can pick up and reflect these biases in their outcomes.
 - Privacy and copyright infringement:
 - Sensitive data can be inadvertently included in models.
 - Unstructured data may contain copyrighted material.
 - Lack of transparency:
 - Black box models offer high accuracy but lack explainability.
 - Lack of transparency can erode trust in AI systems.
 - Model deterioration:
 - AI models require continuous monitoring and updates.

- Incoming data may differ from training data, causing performance degradation.
- Continuous monitoring ensures AI models produce consistent, high-quality outcomes and prevents deterioration in model performance over time.
- Organizations worldwide are creating regulations and guidelines for AI.
 - Examples include the NIST AI Risk Management Framework and the EU AI Act.
 - Regulations can penalize non-compliant companies, emphasizing the importance of adhering to ethical guidelines.
- AI promises significant benefits but also poses real risks.
- Proper governance is crucial to fully realize the potential of AI while mitigating associated risks.