Training Course on AI Ethics and Responsible Development

Professional Development Training Course Outline

Category	Data Science	**Duration**	5 Days
Base Fee	$1,500 USD	**Delivery Mode**	Online / On-site Hybrid

Course Introduction

Artificial Intelligence (AI) is rapidly transforming industries and societies, bringing unprecedented opportunities but also raising critical ethical questions. As AI systems become more autonomous and impactful, ensuring their development and deployment align with human values and societal good is paramount. Training Course on AI Ethics & Responsible Development dives deep into the core principles of **fairness, accountability, and transparency (FAT)**, equipping participants with the knowledge and practical tools to build trustworthy, equitable, and human-centric AI solutions.

In today's rapidly evolving technological landscape, the ethical implications of AI are no longer a fringe concern but a central pillar of successful innovation and long-term sustainability. This course moves beyond theoretical discussions to provide actionable strategies for identifying and mitigating algorithmic bias, implementing robust governance frameworks, and fostering a culture of responsible AI development. By mastering the principles of FAT, organizations can build public trust, navigate complex regulatory landscapes, and unlock the full, positive potential of AI for a more equitable and inclusive future.

Programme Curriculum

Training Course on AI Ethics & Responsible Development: Principles of Fairness, Accountability, and Transparency in AI

Introduction

Artificial Intelligence (AI) is rapidly transforming industries and societies, bringing unprecedented opportunities but also raising critical ethical questions. As AI systems become more autonomous and impactful, ensuring their development and deployment align with human values and societal good is paramount. Training Course on AI Ethics & Responsible Development dives deep into the core principles of **fairness, accountability, and transparency (FAT)**, equipping participants with

the knowledge and practical tools to build trustworthy, equitable, and human-centric AI solutions.

In today's rapidly evolving technological landscape, the ethical implications of AI are no longer a fringe concern but a central pillar of successful innovation and long-term sustainability. This course moves beyond theoretical discussions to provide actionable strategies for identifying and mitigating algorithmic bias, implementing robust governance frameworks, and fostering a culture of responsible AI development. By mastering the principles of FAT, organizations can build public trust, navigate complex regulatory landscapes, and unlock the full, positive potential of AI for a more equitable and inclusive future.

Course Duration

5 days

Course Objectives

1. Grasp core AI ethical frameworks and their application in real-world scenarios.
2. Identify and mitigate algorithmic bias across diverse datasets and model architectures.
3. Implement fairness metrics and explainable AI (XAI) techniques for enhanced interpretability.
4. Develop robust AI governance strategies and accountability mechanisms.
5. Navigate the evolving global AI regulatory landscape and compliance requirements.
6. Understand data privacy and security best practices in AI development, including privacy-preserving AI.
7. Assess the societal impact of AI technologies and promote human-centric AI design.
8. Foster an ethical AI culture within organizations, emphasizing responsible innovation.
9. Apply ethical by design principles throughout the AI development lifecycle.
10. Conduct comprehensive AI impact assessments to anticipate and address potential harms.
11. Explore human-in-the-loop strategies for enhanced AI oversight and control.
12. Design and implement trustworthy AI systems that build public confidence.
13. Drive sustainable AI development practices considering environmental and social factors.

Organizational Benefits

- Building public and stakeholder confidence through transparent and ethical AI practices.
- Proactively mitigating the risks of algorithmic bias, data breaches, and non-compliance with emerging AI regulations.
- Differentiating from competitors by demonstrating a commitment to responsible innovation and ethical AI.
- Ensuring AI systems produce fair, accurate, and explainable outcomes, leading to more reliable business intelligence.
- Creating a purpose-driven work environment that aligns with the values of ethical AI professionals.
- Encouraging interdisciplinary dialogue between technical teams, legal, and ethics departments.
- Developing adaptable and resilient AI strategies that can navigate evolving ethical and regulatory challenges.
- Contributing to the development of AI that serves humanity and promotes equitable outcomes.

Target Audience

1. AI Developers, Engineers, and Data Scientists

2. Product Managers and Business Leaders
3. Legal and Compliance Professionals.
4. Ethicists and Policy Makers.
5. Researchers and Academics
6. Auditors and Risk Management Professionals.
7. Human Resources Professionals.
8. Anyone interested in the societal and ethical implications of AI.

Course Outline

Module 1: Foundations of AI Ethics and Responsible AI

- Defining AI Ethics: Principles, values, and the moral compass for AI.
- The Urgency of Responsible AI: Why ethical considerations are critical now.
- Key Ethical Theories in AI: Consequentialism, deontology, virtue ethics, and their application.
- Global AI Ethics Initiatives and Guidelines: A review of leading frameworks (e.g., OECD, EU AI Act).
- Understanding the AI Development Lifecycle from an Ethical Perspective.
- **Case Study:** The rise of facial recognition technology and debates around privacy vs. public safety.

Module 2: Fairness and Bias Mitigation in AI

- Sources and Types of Algorithmic Bias: Data bias, algorithmic bias, deployment bias.
- Quantifying Fairness: Statistical parity, equalized odds, and other fairness metrics.
- Techniques for Bias Detection: Auditing datasets and models for discriminatory patterns.
- Strategies for Bias Mitigation: Pre-processing, in-processing, and post-processing techniques.
- Fairness in Practice: Trade-offs and challenges in achieving equitable outcomes.
- **Case Study:** COMPAS (Correctional Offender Management Profiling for Alternative Sanctions) algorithm and its racial bias in predicting recidivism.

Module 3: Transparency and Explainable AI (XAI)

- The Importance of Transparency in AI: Building trust and accountability.
- Defining Explainable AI (XAI): Why and how AI systems make decisions.
- Techniques for Model Interpretability: LIME, SHAP, feature importance.
- User-Centric Explainability: Designing explanations for different stakeholders.
- Balancing Explainability with Performance and Privacy.
- **Case Study:** AI in loan applications: Explaining rejection decisions to applicants to ensure fairness and compliance.

Module 4: Accountability and AI Governance

- Defining Accountability in AI: Who is responsible for AI outcomes?
- Establishing AI Governance Frameworks: Policies, processes, and oversight.
- Roles and Responsibilities in Responsible AI: From developers to leadership.
- AI Auditing and Impact Assessments: Proactive risk identification and management.
- Legal and Ethical Liability in AI: Exploring frameworks for responsibility.
- **Case Study:** Autonomous vehicle accidents: Assigning accountability to manufacturers, developers, or operators.

Module 5: Data Privacy and Security in AI

- Fundamental Principles of Data Privacy: GDPR, CCPA, and global regulations.
- Privacy-Preserving AI Techniques: Federated learning, differential privacy, homomorphic encryption.
- Ethical Considerations in Data Collection and Usage for AI Training.
- Security Risks and Vulnerabilities in AI Systems (e.g., adversarial attacks).
- Implementing Robust Data Governance for AI.
- **Case Study:** Use of patient health data in AI diagnostics and ensuring HIPAA compliance and ethical data handling.

Module 6: Societal Impact and Human-Centric AI

- AI and Employment: Job displacement, reskilling, and the future of work.
- Addressing Discrimination and Inequality Amplified by AI.
- AI and Human Autonomy: The role of human oversight and decision-making.
- AI for Sustainable Development Goals (SDGs): Leveraging AI for good.
- Fostering Public Dialogue and Engagement on AI's Societal Implications.
- **Case Study:** AI in hiring platforms: Discussing concerns about perpetuating gender and racial biases in recruitment.

Module 7: Ethical AI Development Practices

- Integrating Ethics throughout the AI Development Lifecycle: From conception to deployment.
- Ethical Design Principles and Methodologies: Value-sensitive design, participatory design.
- Tools and Resources for Ethical AI Development: Open-source libraries and frameworks.
- Building an Ethical AI Culture: Training, awareness, and incentives.
- Promoting Interdisciplinary Collaboration: Bridging the gap between technical and ethical expertise.
- **Case Study:** Designing an AI-powered content moderation system with built-in ethical safeguards to prevent censorship and promote free speech.

Module 8: Emerging Trends and Future of AI Ethics

- The Evolving AI Regulatory Landscape: Anticipating future policies and standards.
- Advanced AI Safety and Alignment: Ensuring AI systems align with human values.
- Ethical Considerations of Generative AI and Large Language Models (LLMs).
- AI in Critical Infrastructures: Ethical implications and risk management.
- The Role of AI Ethics in Global Geopolitics and International Cooperation.
- **Case Study:** The ethical challenges of deepfakes and generative AI in misinformation campaigns.

Training Methodology

This training course will employ a highly interactive and practical methodology designed to foster deep understanding and application of AI ethics principles. The approach will include:

- **Interactive Lectures and Discussions:** Engaging presentations followed by open forums to explore concepts and perspectives.
- **Real-World Case Study Analysis:** In-depth examination of prominent ethical dilemmas and successful responsible AI implementations.
- **Group Exercises and Collaborative Problem-Solving:** Applying learned principles to hypothetical and real-world scenarios.

- **Practical Tools and Frameworks:** Hands-on experience with bias detection tools, explainability libraries, and governance templates.
- **Guest Speakers:** Insights from industry leaders, ethicists, and policymakers.
- **Role-Playing and Simulations:** Experiencing ethical dilemmas from different stakeholder perspectives.
- **Q&A Sessions and Debates:** Encouraging critical thinking and diverse viewpoints.

Register as a group from 3 participants for a Discount

Send us an email: info@fineskilltrainingcenter.org or call *+254769199797*

Certification

Upon successful completion of this training, participants will be issued with a globally- recognized certificate.

Tailor-Made Course

We also offer tailor-made courses based on your needs.

Key Notes

a. The participant must be conversant with English.

b. Upon completion of training the participant will be issued with an Authorized Training Certificate

c. Course duration is flexible and the contents can be modified to fit any number of days.

d. The course fee includes facilitation training materials, 2 coffee breaks, buffet lunch and A Certificate upon successful completion of Training.

e. One-year post-training support Consultation and Coaching provided after the course.

f. Payment should be done at least a week before commence of the training, to Fineskill Training Center account, as indicated in the invoice so as to enable us prepare better for you

Upcoming Scheduled Sessions

Start Date	End Date	Location	Price
23 Nov 2026	27 Nov 2026	Online	$1,000
23 Nov 2026	27 Nov 2026	Physical	$1,500

Start Date	End Date	Location	Price
23 Nov 2026	27 Nov 2026	Physical	$3,000
23 Nov 2026	27 Nov 2026	Physical	$3,000
23 Nov 2026	27 Nov 2026	Physical	$6,000
23 Nov 2026	27 Nov 2026	Physical	$4,500
23 Nov 2026	27 Nov 2026	Physical	$6,000
23 Nov 2026	27 Nov 2026	Physical	$6,000

Ready to enroll or request a customized program? Book online or contact us:

Register Online Now

Email: info@fineskilltrainingcenter.com | Website: www.fineskilltrainingcenter.com