

Guidelines on ethical and responsible AI:

Ethical considerations play a major role in shaping the development, deployment, and impact of AI technologies. As we harness the power of AI to drive innovation and solve problems, it is essential to uphold certain values and responsibilities. The following section describes principles of responsible AI identified by Intel and should be used as a guideline to ensure the ethical and responsible use of AI in festival entries:

1. Protect the Environment

Recognize that AI applications consume considerable energy and materials. Consider the environmental impact of AI technologies and strive to minimize energy consumption and resource usage throughout the AI lifecycle. Try to adopt eco-friendly development techniques for both software and hardware to help create less waste and emissions.

Example: Artificial intelligence systems can drastically lower their carbon footprint by implementing energy-efficient algorithms. For instance, optimizing algorithms to minimize computational redundancy can lead to less energy consumption during processing, thus lowering environmental impact.

2. Promote Equity and Inclusion

Recognize, address, and minimize potential biases effectively without compromising the AI's effectiveness and utility.

Example: Employing "person-detection" technology that is equally accurate across different demographic groups ensures fairness and inclusivity. By regularly testing and adjusting algorithms to perform well across a variety of ethnicities and genders, biases in identification processes can be minimized.

3. Design for Privacy

Ensure that AI applications uphold the rights and freedoms of individuals, utilizing lawfully obtained data in accordance with legal obligations and policy standards. This includes transparency regarding the necessity of personal data collection, empowering users with choice and control, and implementing robust measures to safeguard personal data in the development and deployment of our products.

Example: While developing a healthcare AI system that analyzes patient data, anonymizing personal information before processing ensures privacy protection. By removing markers

such as names or social security numbers from datasets, the privacy rights of individuals can be upheld while still allowing effective data analysis and insights.

4. Enable Transparency and Explainability.

Document and convey the purpose, limitations, and design outcomes of the program. Maintain accountability for changes, versions, and iterations made to the model. Employ methods that are explainable and comprehensible, enabling users, supervisors, and the public to understand how and why the AI program generates its outputs.

Example: A chatbot could disclose the information it is using to generate a response, helping users understand why certain suggestions are made, hence fostering trust in the AI's capabilities.

5. Advance Security, Safety, and Reliability

Test the AI to a level that matches foreseeable risks associated with its use. The project should implement best practices to enhance the reliability, security, and accuracy of AI systems. By adhering to security protocols, the aim is to enhance the resilience of the system and reduce vulnerability to negative external influences.

Example: For a program dealing with financial transactions, encrypting data during transmission and storage prevents unauthorized access and ensures data integrity.

6. Enable Human Oversight

To ensure human oversight across the entire lifespan of an AI system, it is essential to cater to the requirements of different stakeholders. Developers must aim to educate AI systems with suitable supervision and feedback, while users and other involved parties should have the ability to evaluate the outcomes of these systems and step in when necessary.

Example: A content moderation platform could include tools for human reviewers to easily flag and review potentially harmful or inappropriate content flagged by AI algorithms, ensuring human oversight in the moderation process.

7. Respect Human Rights

Prioritize the consideration and respect of human rights for all relevant stakeholders throughout the development, use, and deployment of the project and try to align practices with global human rights principles.

Example: In an AI-powered hiring tool designed to screen job applicants, the tool must ensure fair treatment and equal opportunity for all candidates, regardless of their race, gender, or other protected characteristics. The AI should be programmed to analyze candidates based solely on job-related criteria and avoid biased decision-making that could perpetuate discrimination or prejudice.

8. Ensure Data Integrity and Bias Mitigation

Verify that the data used by AI systems originates from credible sources and reliable sources. Assess the balance and impartiality of the data to prevent biases from influencing AI decision-making processes.

Example: For autonomous vehicle navigation systems relying on AI algorithms to interpret sensor data and make driving decisions, data integrity is paramount for ensuring safe and reliable operation. The data used to train the AI models should be collected from diverse and representative real-world driving scenarios, including varying road conditions, weather patterns, and traffic situations.

9. Avoid Plagiarism and Provide Proper Credits

Refrain from using plagiarized content in AI systems. Give due credit and references to original sources when incorporating external data, models, or algorithms. Uphold intellectual property rights and acknowledge the contributions of others to foster a culture of integrity and accountability within the AI community.

Example: When developing a text generation model, refrain from directly copying or plagiarizing text from existing sources without proper attribution. Instead, leverage techniques such as paraphrasing or summarization to generate original content with citations.

Additional Resources:

- a. Responsible AI Principles - Intel:
<https://www.intel.com/content/www/us/en/artificial-intelligence/responsible-ai-principles.html>