AI Project Guardrails Agreement Template

EMPYREAN RESEARCH INSTITUTE

This template outlines the boundaries and rules for AI projects, helping AI and legal teams establish a clear framework for compliance, ethical considerations, and innovation. By defining guardrails upfront, teams can safely explore AI's potential without breaching legal or ethical standards.

Project Name:

[Enter the AI project's title]

Brief Description:

[Enter a brief description of the project's purpose, goals, and scope]

Purpose of the Guardrails:

The purpose of this agreement is to establish clear legal and ethical boundaries for AI innovation within the project, ensuring compliance with relevant regulations while promoting responsible AI development.

Guardrails and Restrictions:

1. Data Use:

- **WHAT TO INCLUDE:** *Specify what data can and cannot be used within the AI project. Ensure data handling complies with privacy laws (e.g., GDPR, CCPA)*
- **EXAMPLES:**
  - *The AI model cannot make autonomous decisions related to employment (e.g., hiring/firing)*
  - *AI recommendations must always include an option for human review before final action.*

2. AI Model Limitations:

- **WHAT TO INCLUDE:** *Define limitations on the AI model's usage, especially in decision-making processes that may have legal or ethical implications.*
- **EXAMPLES:**
  - *The AI model cannot make autonomous decisions related to employment (e.g., hiring/firing).*
  - *AI recommendations must always include an option for human review before final action.*

3. Security Measures:

- **WHAT TO INCLUDE:** *Outline security protocols that must be in place to protect data and system integrity.*
- **EXAMPLES:**
  - *Implement end-to-end encryption for all data in transit and at rest.*
  - *Regularly perform security audits and vulnerability assessments on the AI infrastructure.*

4. Regulatory Compliance:

- **WHAT TO INCLUDE:** *Mention any specific laws, industry standards, or regulations the project must comply with.*
- **EXAMPLES:**
  - *Comply with GDPR for any data collected from EU citizens.*
  - *Ensure the AI model adheres to HIPAA standards if handling health-related data.*

Roles and Responsibilities:

1. AI Team:

- *Define the AI team's responsibilities related to the design, development, and deployment of the AI model.*

- ***EXAMPLE:***

  - *Ensure all training data is vetted and approved by the legal team.*
  - *Conduct regular bias audits and model validation to ensure fairness and accuracy.*

2. Legal Team:

- *Define the legal team's responsibilities in terms of compliance reviews and approvals throughout the AI project lifecycle.*

- ***EXAMPLE:***

  - *Review data privacy policies and ensure compliance with relevant regulations.*
  - *Provide sign-off at key stages of the project to ensure legal and regulatory requirements are met.*

Legal Sign-Off Points:

- *Identify critical phases in the project where legal must provide formal approval before moving forward.*

1. Design Phase: [Yes/No]

- *Legal team reviews the initial project design, focusing on data usage and compliance requirements.*

2. Pre-Launch: [Yes/No]

- *Legal reviews the AI model before deployment, ensuring that all guardrails and compliance needs have been addressed.*

3. Post-Launch Review: [Yes/No]

- *After the AI system is live, legal conducts a review to ensure ongoing compliance and address any emerging risks.*

Approval and Agreement:

This section finalizes the agreement, ensuring all parties have acknowledged and accepted the guardrails and responsibilities.

- ***AI TEAM LEAD:*** *[Signature]*
- ***LEGAL TEAM LEAD:*** *[Signature]*
- ***DATE:*** *[Date]*

Additional Features:

1. Ongoing Monitoring:

Add a section for periodic review, where AI and legal teams revisit the guardrails to ensure they remain relevant as the project evolves.

2. Escalation Procedures:

Define clear steps for escalating issues if the AI model operates outside the established guardrails (e.g., triggering an audit, pausing deployment).

3. Ethical Oversight:

Consider adding a sub-section for ethical oversight, where the AI system's performance can be evaluated against ethical AI principles (e.g., fairness, bias mitigation).

ABOUT

The Empyrean Research Institute empowers organizations to transform their workforce through evidence-based, innovative insights and real, actionable strategies for a connected, future-ready culture.