

Solar Industries - Responsible AI Development Framework

Version: 2.1 **Effective Date:** July 1, 2025 **Governing Body:** AI Ethics & Governance Committee

Preamble

This document outlines the core principles that must guide all Artificial Intelligence (AI) and Machine Learning (ML) development at Solar Industries. Our commitment is to create technology that is not only powerful but also fair, transparent, and accountable. Adherence to this framework is mandatory for all AIML projects.

Principle 1: Fairness and Bias Mitigation

- **Requirement:** All models trained on demographic data must be rigorously tested for bias across protected categories (e.g., race, gender, age).
- **Action:** Project teams must document their bias detection and mitigation strategies. Tools such as fairness indicators and counterfactual analysis are to be used where applicable. Models exhibiting unacceptable levels of bias will not be deployed.

Principle 2: Transparency and Explainability

- **Requirement:** For any AI system that directly impacts customers or employees (e.g., credit scoring, hiring recommendations), a mechanism for explaining its decisions must be implemented.
- **Action:** "Black box" models are discouraged in high-stakes applications. Techniques like SHAP (SHapley Additive exPlanations) or LIME (Local Interpretable Model-agnostic Explanations) should be integrated to provide decision-level insights. The existence and general logic of the AI system must be disclosed to affected individuals.

Principle 3: Accountability and Human Oversight

- **Requirement:** All autonomous AI systems must have a designated human overseer and a clear "off-switch" or manual override mechanism.
- **Action:** AI systems are not permitted to make critical, irreversible decisions without a human in the loop. The role and responsibilities of the human overseer must be clearly defined and documented for each system.

Principle 4: Security and Privacy

- **Requirement:** All data used for training and inference must be handled in accordance with GDPR, CCPA, and internal data privacy standards.

- **Action:** Personally Identifiable Information (PII) must be anonymized or pseudonymized wherever possible. Models must be robust against adversarial attacks, and a data retention policy must be established for all training datasets.

Review Process

All new AIML projects must undergo an Ethical Review by the AI Ethics & Governance Committee before deployment.