Responsible AI Evaluation Checklist

1. Purpose & Use Case Clarity

- Is the business objective clearly defined?

- Does the AI use case align with company strategy or customer value?

- Are stakeholders (business, legal, technical) aligned on the use case?

2. Fairness & Bias

- Is the training data representative of the population it will serve?

- Have potential biases been identified and mitigated (e.g., gender, race)?

- Have you tested the model for disparate impact across key groups?

3. Explainability & Transparency

- Is the model explainable to non-technical users (e.g., via LIME, SHAP)?

- Can the key features influencing predictions be understood and communicated?

- Is documentation (model cards, datasheets) maintained for the model?

4. Privacy & Consent

- Does the model use personal or sensitive data?

- Have you ensured compliance with data protection laws (e.g., GDPR, CCPA)?

- Is data anonymized or minimized when possible?

5. Security & Resilience

- Has the model been tested for adversarial risks or spoofing attacks?

- Are model access controls and audit trails in place?

- Are there safeguards to detect model drift or data quality issues?

6. Accountability & Governance

- Is there a clear model owner or responsible party?

- Is the model reviewed and approved through an AI/ML governance process?

- Are model performance and risk metrics tracked post-deployment?

7. Human Oversight

- Is human-in-the-loop review applied where needed (e.g., hiring, finance, healthcare)?

- Can end users override or contest automated decisions?

- Are users trained to use AI outputs responsibly?

8. Metrics & ROI

- Are success metrics (accuracy, precision, recall, business impact) defined?

- Is the ROI clearly measurable and linked to business outcomes?

- Are unintended consequences or opportunity costs considered?