

D4: Responsible AI – SmartPay AP

Responsible AI Principles

- **Fairness & Bias Mitigation:**

- Monitor for vendor, region, or amount-based bias in training data.
- Use explainable AI to surface reasons for mismatch detection.
- Regularly re-train model with diverse, updated invoice datasets.

- **Data Privacy & Security:**

- Mask sensitive vendor details in logs & training.
- Enforce role-based access control (RBAC) for invoice data.
- Encrypt all data at rest (AES-256) and in transit (TLS).

- **Human-in-the-Loop:**

- No auto-emailing to vendors without AP officer review.
- Explainability ensures humans understand AI decisions.

Governance & Model Operations

- **Audit Trail:**

- Log all model predictions, explanations, and approvals.
- Provide traceability for compliance & vendor disputes.

- **Model Monitoring:**

- Track drift in invoice/PO distributions.
- Monitor key metrics: precision, recall, false positives.
- Implement alerting for performance degradation.

- **Cost & Sustainability:**

- Optimize LLM usage (cache frequent prompts, batch requests).
- Scale infra dynamically via cloud auto-scaling.

■ Ensures solution is **Trustworthy, Secure, Auditable, and Scalable**.