## **Security Policy for LLM ‘Red Team Agent’**

## **1. Introduction**

This security policy outlines the measures taken to ensure the confidentiality, integrity, and availability of data processed by the AI-powered Bid Agent. The policy covers data security, compliance, AI-specific concerns, and system-level best practices.

## **2. Data Security Measures**

### **2.1 Secure Data Transmission**

* All data transmitted between users and the AI Bid Agent is encrypted using **TLS 1.2+ (HTTPS)**.
* Any document uploads and API calls follow secure data transfer protocols.

### **2.2 Data Storage & Retention**

* The system follows a **"No Storage" Policy**, ensuring bid documents are processed in-memory and deleted after analysis.
* If storage is required, **AES-256 encryption** is applied.
* Users are informed about data retention policies and can request deletion of their data.

### **2.3 Secure API Authentication**

* API endpoints require **OAuth 2.0 authentication and JWT token-based access controls**.
* Rate limiting and IP filtering prevent unauthorized or excessive requests.

### **2.4 Access Control**

* **Role-Based Access Control (RBAC)** ensures users access only relevant documents and system functionalities.
* User authentication is enforced using **multi-factor authentication (MFA)** where applicable.

## **3. Compliance & Legal Considerations**

### **3.1 GDPR Compliance (EU Users)**

* Users must provide **explicit consent** before data processing.
* A **"Delete My Data" feature** allows users to erase their records.
* Personal and bid data processing complies with **Article 5 of the GDPR (Principles of Processing)**.

### **3.2 ISO 27001 & NIST Compliance**

* All sensitive data is encrypted at rest and in transit using **AES-256 encryption**.
* Access logs are maintained for **audit and monitoring purposes**.
* The system undergoes **regular security assessments and penetration testing**.

## **4. AI-Specific Security Concerns**

### **4.1 Preventing AI Bias & Manipulation**

* AI models are trained on **curated datasets** to ensure unbiased and reliable feedback.
* Model updates are reviewed periodically to prevent unintended biases.

### **4.2 Guardrails for AI Usage**

* The AI Bid Agent is restricted from **generating bid content**; it only provides feedback.
* Ethical AI guidelines are enforced to prevent unauthorized or unfair bid assessments.

### **4.3 Ethical AI Monitoring**

* AI outputs are monitored for accuracy, fairness, and compliance with industry standards.
* Users can provide feedback to improve AI-generated insights.

## **5. System-Level Security Best Practices**

### **5.1 Zero-Trust Architecture**

* The system follows a **Zero-Trust security model**, ensuring every request is authenticated and authorized before processing.

### **5.2 Regular Security Audits**

* Security audits and **penetration testing** are conducted quarterly.
* Automated vulnerability scanning is implemented to detect and remediate threats.

### **5.3 Audit Logs & Monitoring**

* Detailed logs are maintained to track document access, user interactions, and AI feedback requests.
* Anomaly detection systems flag suspicious activity for investigation.

## **6. Key Takeaways**

* **Encryption & Secure Access:** TLS 1.2+, AES-256, OAuth 2.0.
* **Regulatory Compliance:** GDPR, ISO 27001, NIST security standards.
* **AI Ethics:** No bid content generation, only structured feedback.
* **Ongoing Monitoring:** Regular audits, role-based controls, and real-time anomaly detection.

This policy will be reviewed and updated periodically to align with industry best practices and evolving security requirements.

**Effective Date:** [Insert Date]

**Review Date:** [Insert Date]

**Approved by:** [Insert Name]