SigSecure AI

1. **Team Details**

- **Team Name:** Genesis
- **Team Members:**
  - Nilay Sharma
  - Ayesha Sheikh

2. **Problem Understanding and Scope:**

- **Problem Statement:** The core problem is that current deidentification tools treat all visual elements equally, failing to understand the contextual significance and temporal nature of personally identifiable information in documents. "Linked visual data" means documents where textual PII (names, addresses) is semantically connected to visual elements (signatures, photos, stamps) - creating complex privacy relationships that require intelligent, context-aware protection rather than blanket redaction.

- **Target Documents & Formats:**
  - ***Legal contracts*** *with multiple signatures and witness information.*
  - ***Medical consent forms*** *with patient signatures and ID photos.*
  - ***Financial loan applications*** *with applicant signatures and identity documents.*
  - ***Government forms*** *(passport applications, visa documents) with photos and signatures.*
  - ***Insurance claim documents*** *with signature verification and supporting photos.*
  - ***Real estate documents*** *with multiple party signatures and property photos.*

- **Types of Identifiable Data (PII) to Detect & Redact:**
  - ***Signature-specific PII:*** *Handwritten signatures, digital signatures, initials, witness signatures.*
  - **Connected visual PII:** ID photos linked to signature holders, stamps/seals with personal information.
  - **Contextual textual PII:** Names, addresses, phone numbers when linked to signatures.
  - **Temporal PII:** Dates that establish signature validity and legal significance.
  - **Relationship PII:** Information that becomes identifying when combined (signature + address + date).
  - **Metadata PII:** Document creation timestamps, digital signature certificates.

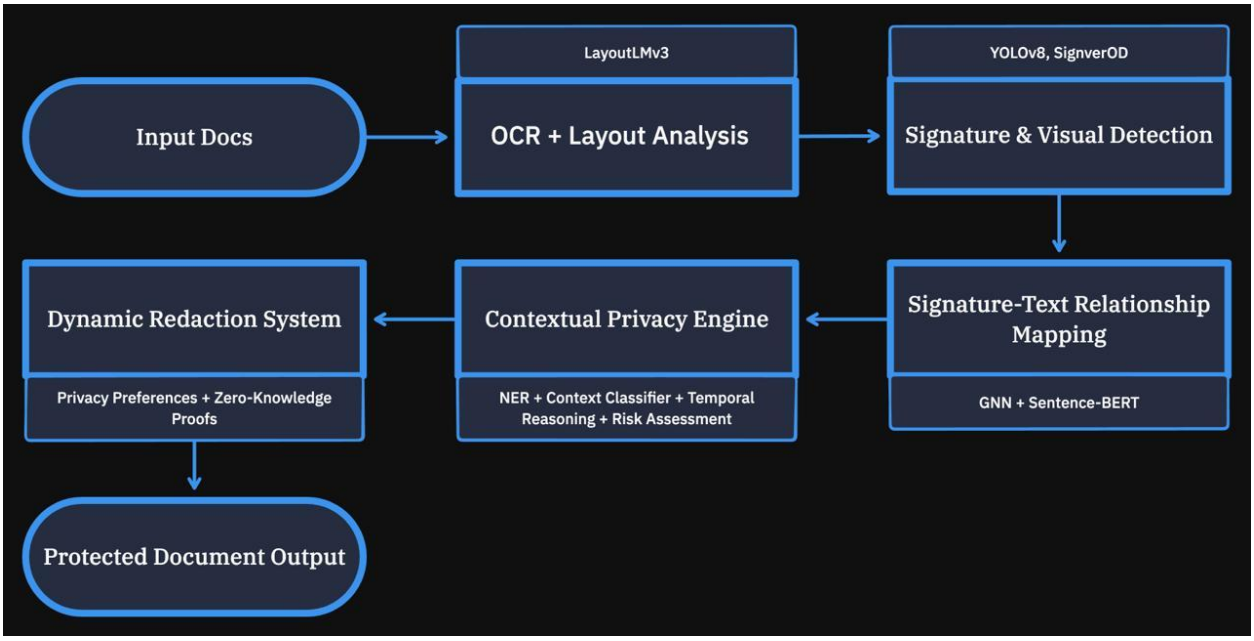
- **User Personas / End Users:**
  - ***Legal firms*** *handling confidential client documents with multiple*

signatories.

- ***Healthcare institutions*** *processing patient consent forms and medical records.*
- ***Financial institutions*** *managing loan applications and account opening documents.*
- ***Government agencies*** *processing citizen applications and official documents.*
- ***HR departments*** *handling employment contracts and background verification documents.*
- ***Insurance companies*** *processing claims with signature verification requirements.*

3. Proposed Solution & Approach:

- **High-Level Architecture (with Diagram):**



- **AI/ML Models Considered:**
  - ***SignverOD dataset models*** *for signature detection and classification (4-class detection).*
  - ***LayoutLMv3*** *for document structure understanding and text visual relationship mapping.*
  - ***YOLO v8*** *for real-time object detection of stamps, seals, and photos.*
  - ***spaCy + Custom NER models*** *for PII detection with signature context awareness.*
  - ***Sentence-BERT*** *for semantic similarity between signatures and related text.*
  - ***Custom Graph Neural Networks*** *for modeling relationships between visual and textual elements.*
  - ***Transformer based models*** *for temporal reasoning and document classification.*

- **Data Strategy:**
    - ***Primary:*** *Leverage SignverOD dataset for foundational signature detection training.*
    - ***Augmentation:*** *Create synthetic documents with known signature-text relationships for training.*
    - ***Crowdsourcing:*** *Develop privacy-preserving annotation platform for community contributions.*
    - ***Simulation:*** *Generate realistic document layouts with embedded signature-PII relationships.*
    - ***Validation:*** *Create test suites with ground truth for signature privacy scenarios.*
    - ***Continuous learning:*** *Implement federated learning for model improvement without data sharing.*

- **Innovation / Unique Selling Point (USP):**
    - ***World's first signature-contextual privacy system*** *that understands relationships between signatures and associated PII.*
    - ***Temporal signature intelligence*** *that adapts privacy protection based on document lifecycle and legal significance.*
    - ***Multi-party privacy orchestration*** *allowing different signers to set individual privacy preferences on shared documents.*
    - ***Signature authenticity preservation*** *that maintains legal validity while protecting privacy.*
    - ***Zero-knowledge signature verification*** *enabling authentication without identity exposure.*

4. Proposed Solution & Approach:

- ***Intended UI/UX Design (if applicable):***
    - ***Interactive privacy control panel*** *where users can adjust protection levels for different signature types.*
    - ***Collaborative dashboard*** *for multi-party documents showing each signer's privacy preferences.*
    - ***Privacy visualization tools*** *showing data flow and protection levels applied.*
    - ***Drag-and-drop*** *web interface with real-time signature detection preview.*
    - ***API-first architecture*** *with comprehensive developer documentation and SDKs.*

- ***Input & Output Format Expectations:***
    - ***Input formats:*** *PDF, PNG, JPG, TIFF, multi-page scanned documents, mobile camera captures.*
    - ***Output formats:*** *Privacy-protected PDFs with preserved formatting, redacted images with metadata.*
    - ***Audit outputs:*** *Detailed privacy application logs, compliance reports, signature analysis summaries.*
    - ***API outputs:*** *JSON responses with signature locations, privacy levels applied, and confidence scores.*

    - ***Batch processing:*** *Support for bulk document processing with progress tracking.*

- ***Accessibility / Ease of Use Considerations:***
    - ***No-code interface*** *for non-technical users with preset privacy templates.*
    - ***Multilingual support*** *with culturally-aware signature recognition (different signature styles globally).*
    - ***Low-bandwidth optimization*** *for users in regions with limited internet connectivity.*
    - ***Offline processing capability*** *for sensitive environments that cannot use cloud services.*
    - ***Screen reader compatibility*** *and keyboard navigation for accessibility compliance.*
    - ***Progressive disclosure*** *of advanced features to avoid overwhelming casual users.*
    - ***One-click privacy presets*** *for common document types (legal, medical, financial).*