SurgiSense: AI-Powered Surgical Recovery Companion

Problem Statement

The surgical care journey is currently broken by two critical "points of failure" that cost hospitals millions and endanger patient lives:

- **The "Cancellation Crisis":** Up to **15% of surgeries are cancelled last-minute** simply because patients fail to comply with complex pre-op instructions (e.g., stopping blood thinners on time). This operational inefficiency costs hospitals millions in lost Operating Room (OR) utilization.
- **The "Monitoring Black Box":** Upon discharge, patients lose clinical oversight. Surgical Site Infections (SSIs) are often detected too late because patients cannot distinguish between normal healing and dangerous infection, leading to preventable and costly readmissions.

Motivation

Current post-operative care relies on static, confusing paper instructions that patients often lose, misunderstand, or ignore.

- **The Gap:** There is no "digital bridge" between the hospital and the home. Generic health apps lack the specific context of the patient's medical history, while hospital portals are too complex and passive for daily recovery management.
- **The Opportunity:** The emergence of **Multimodal GenAI** allows us to finally solve this. We can convert passive, unstructured paper records into an active, 24/7 "Digital Twin" that guides the patient, reducing anxiety and improving clinical outcomes without requiring expensive human intervention.

Application

- **Real-World Use Case: SurgiSense** acts as an AI-powered recovery companion. It ingests the patient's specific physical discharge summary and converts it into an interactive, gamified recovery timeline.
- **Target Users:**
    - **Primary:** Elderly and post-surgical patients recovering at home, specifically those with low health literacy or language barriers.
    - **Secondary:** Hospital administrators seeking to reduce readmission penalties and OR cancellations.
- **Where it Applies:** The solution is **"Hospital-Agnostic"**. It integrates into the patient's daily routine via a mobile-first web interface, working seamlessly with discharge papers from *any* facility.

Proposed Method

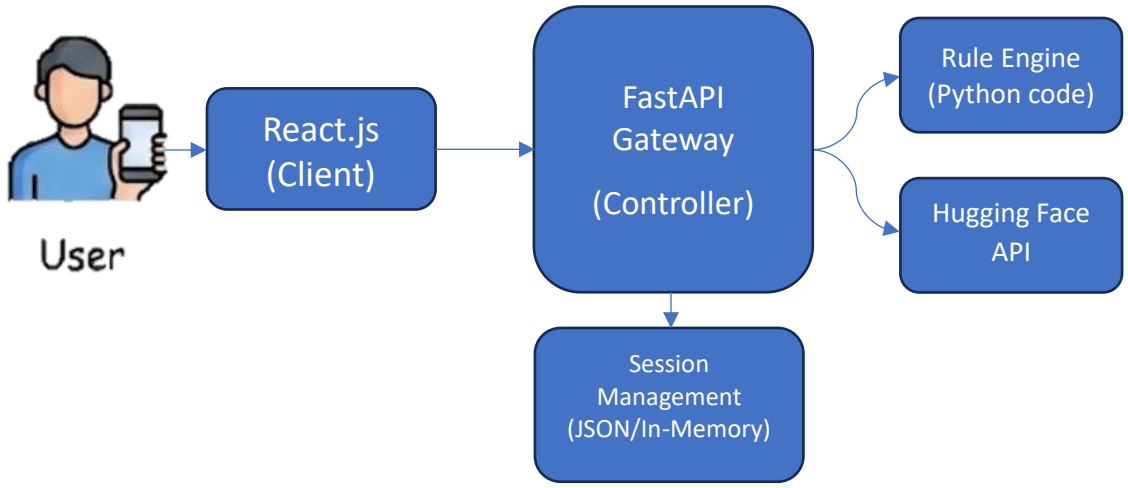

Figure 1: ***Privacy-First Architecture*** *Decoupled design using React, FastAPI, and Hugging Face for secure, local-ready deployment.*

We utilize a **Privacy-First, Neuro-Symbolic GenAI Architecture** to ensure safety, accuracy, and data sovereignty. Our approach decouples probabilistic AI from deterministic safety checks.

1. **Semantic Record Digitization (The Brain)** We use **Meta-Llama-3-8B-Instruct** (via Hugging Face Inference API) to ingest raw PDF discharge summaries. The model extracts unstructured data—dates, medications, dosage instructions—and normalizes it into a structured **JSON dashboard**. This model was chosen for its state-of-the-art instruction-following capabilities.

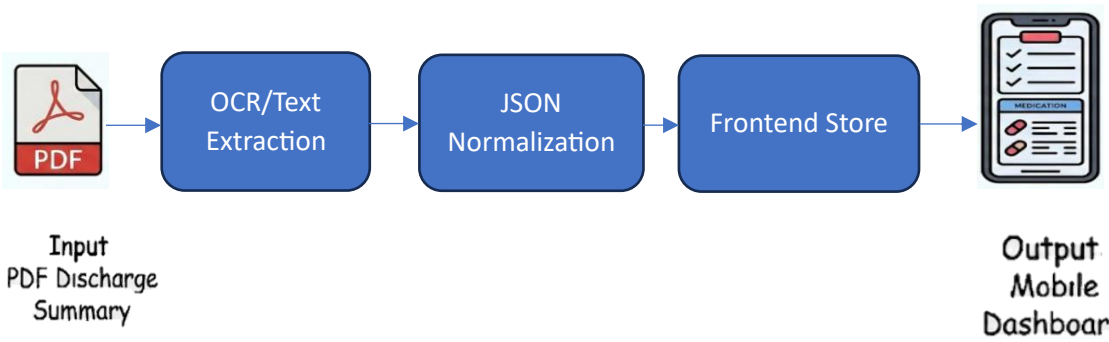

Figure 2: ***The "Insta-Digitize" Flow*** *converts messy PDF summaries into a structured, interactive JSON dashboard*

2. **Visual Wound Analysis (The Eyes)** We employ **LLaVA-v1.5 (Large Language-and-Vision Assistant)** to analyze patient-uploaded photos of surgical incisions. The model scans for

visual biomarkers of **Erythema (redness)** and discharge, acting as an early warning system for infections.

3. **"Insta-Digitize" RAG Chat** A **Retrieval-Augmented Generation (RAG)** system answers patient queries (e.g., *"Can I shower today?"*) by strictly citing the uploaded PDF context, ensuring zero hallucinations.

4. **Safety Layer (Deterministic Rule Engine)** To prevent medical errors, a Python-based **Rule Engine** sits between the AI and the user. It validates critical pre-op constraints (e.g., *"Aspirin use < 7 days pre-op = ALERT"*) using deterministic logic rather than probabilistic generation.

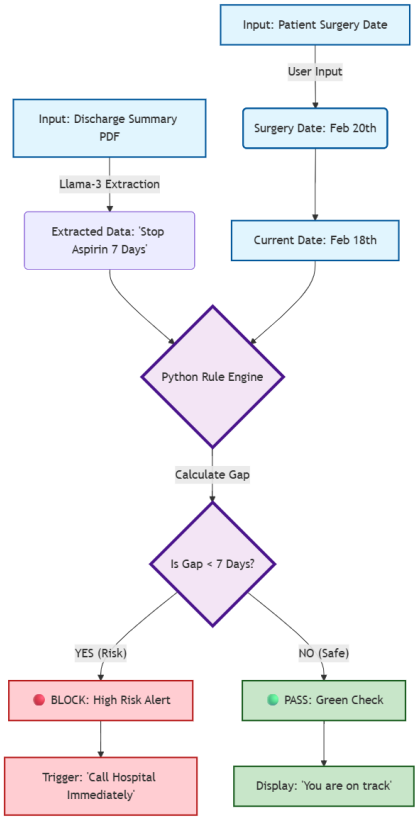

Figure 3: Hybrid Neuro-Symbolic Safety Logic

5. **Voice-First Vernacular Interface** To democratize access, we utilize a browser-native **Speech-to-Text engine** that supports local languages (Hindi/Tamil), allowing elderly users to speak symptoms naturally.

6. **Affiliate Pharmacy Integration** A "One-Click Refill" feature redirects users to partner pharmacies, creating a sustainable B2B2C revenue model.

Datasets / Data Source

- **Input Data:** The system is designed to ingest **unstructured PDF Discharge Summaries** provided by hospitals. It does not require a pre-labeled dataset for training, as it uses zero-shot extraction.

- **Model Weights:** We utilize pre-trained open-source models (**Llama-3-8B** and **LLaVA-v1.5**) accessible via the **Hugging Face Inference API**, ensuring we do not rely on closed, paid APIs.
- **Validation Data:** For the hackathon, we are using a synthetic dataset of de-identified discharge summaries and standard wound progression images to benchmark model performance.

Experiments

We validate the solution through a three-stage evaluation framework:

1. **Extraction Accuracy:** Measuring the **F1-score** of Llama-3 in correctly extracting medication names, dosages, and dates from a test set of 50 diverse PDF layouts.
2. **Safety Verification:** Testing the Rule Engine against "adversarial" inputs (e.g., contradictory dates or dangerous medication combos) to ensure it triggers "High Risk" alerts **100% of the time**.
3. **Visual Sensitivity:** Evaluating LLaVA's ability to detect redness in wound images compared to ground-truth labels, establishing a confidence threshold (e.g., >80%) for triggering a "Consult Doctor" alert.

Novelty and Scope to Scale

- **Novelty:** Unlike generic chatbots that act as wrappers for OpenAI, SurgiSense uses a **"Privacy-First Architecture."** By leveraging lightweight, open-source models, we prove that medical AI can be **HIPAA-compliant** and deployable on private hospital servers. Our **Neuro-Symbolic** approach (AI for perception + Code for safety) addresses the critical "hallucination" barrier in healthcare.
- **Scope to Scale:**
    - **Technical Scalability:** The decoupled microservices architecture (FastAPI + React) allows for easy horizontal scaling.
    - **Market Scalability:** The solution is **Hospital-Agnostic**. It requires **zero integration** with hospital IT systems initially—we simply ingest the PDF they already give patients—allowing for rapid, friction-free adoption.
    - **Future Scope:** The roadmap includes **Edge AI** (running models locally on the phone) and **EHR Integration** (HL7/FHIR) to write alerts back to hospital systems.