Health Guide – Link to POC

AI Symptom Intake + Medical Transcriber/GenAI IDP + Insurance-Aware Provider Matching + Personalized Post-Visit Actions + Billing + Signature/Consent Detection

---

1) Thesis (Problem + Scope) - Jennifer

Healthcare is fragmented across **pre-visit**, **during visit**, and **post-visit**, with critical information spread across **audio recordings, unstructured documents, consent forms, and separated billing statements**. Patients commonly struggle to:

- Decide **which doctor/specialty** to see from symptoms
- Find **in-network providers** and understand coverage constraints
- Interpret technical language in notes, referrals, discharge documents
- Know **what to monitor after** a visit/procedure and when to seek help
- Reconcile **multiple bills** for one episode of care
- Understand what they are signing (consent, authorization, financial responsibility) **during or after** the visit

Scope: Build an **AWS-based Generative AI IDP + Transcription system** that supports an end-to-end patient workflow with these core outcomes:

A) Before the visit: Symptom intake → suggested specialty + in-network provider shortlist

- User inputs symptoms, severity, duration, and constraints (location, timing, language, preferences)
- System generates a **triage-oriented summary** (non-diagnostic) and suggests which provider type/specialty to consider
- Matches to providers based on **insurance network** and (where available) **provider availability**

B) During & after the visit: Audio + docs → plain-language + structured outputs

- Ingest audio (doctor dictation/explanation) + documents (intake/referral/clinical notes/discharge)
- Produce:

    - Plain-language summary
    - Action items + next steps
    - “Things to look out for” aftercare watch-outs pulled from source
    - Structured fields (symptoms, conditions as written, meds, tests, follow-up dates)

C) Signature/consent detection (during visit or post-visit)

- Detect and summarize **signature-required documents** (procedure consent, privacy authorizations, financial responsibility, trial consent)
- Output a “Before You Sign” section with:
    - key commitments/obligations
    - confusing/vague clauses to clarify
    - missing fields (dates, costs, provider names)
    - questions to ask before signing

D) Billing reconciliation: explain separated bills + group related charges

- Ingest billing statements and EOB-style documents
- Group and explain charges by visit/episode (facility vs professional vs lab/imaging vs anesthesia)
- Highlight what to verify (dates of service, provider entity, potential duplicates)

E) Personalized post-visit action plan (tailored to the individual)

After a visit, generate a **tailored checklist and monitoring plan** based on:

- the patient’s extracted instructions + diagnosis context (from documents/audio)
- the patient’s self-reported preferences/constraints (time, language, caregiving, reminders, comfort level)
- risk flags and follow-up timelines in the discharge plan

Examples of personalized actions

- Medication schedule formatted to the patient’s routine (“morning/evening” reminders)
- Follow-up appointment tasks (call, portal message template, what to ask)
- Symptom tracking prompts tailored to that condition (“track fever, pain score, wound redness”)
- Personalized red-flag thresholds (“if you experience X, contact provider/urgent care” as stated in the discharge instructions)
- Diet/activity restrictions summarized into a simple, customized do/don’t list
- Checklist ordering (most urgent first) and “what can wait” guidance (based on document timing)

Safety boundary: The system summarizes and personalizes organization/formatting of instructions already present in the source; it does **not** create new medical or legal advice.

---

2) Dataset (Public + Synthetic) -

Public datasets

- Clinical trial protocols from **ClinicalTrials.gov**
- Public medical intake form templates (PDFs)

Synthetic datasets (PHI-free, realistic)

- Simulated patient histories + symptom descriptions with severity levels
- Referral documents, physician notes, discharge instructions
- Synthetic visit transcripts (audio-style text)
- Consent/authorization/financial responsibility forms
- Billing statements + EOB-style documents with multiple billers
- Patient preference profiles for personalization (language, schedule, reminder frequency, caregiving constraints)

Data characteristics

- Dense text + semi-structured layouts (tables, bullets, paragraphs)
- Multi-document “episode of care” linking (one visit ⟶ many docs + signatures + bills)
- Insurance/provider metadata (plan-network mappings, provider specialty, location, availability)

---

3) Concepts & Data Science Models - Kritika

- Symptom normalization + specialty routing (rules + ML)
- OCR + layout detection
- Speech-to-text (ASR)
- Document classification (intake/referral/note/discharge/consent/billing)
- Named entity extraction (symptoms, conditions, meds, tests, follow-ups, warnings)
- LLM summarization + simplification (patient-friendly + structured JSON)
- RAG over glossary + billing explainer + consent clause patterns
- Confidence scoring + HITL validation
- Episode linking (tie documents/bills/consents to the same visit)
- Personalization logic (template + constraints + priority ranking) for post-visit action plans

---

4) AWS Tools Used (AWS-only) -

- **Amazon S3** – raw + processed + results storage
- **Amazon Textract** – OCR/layout extraction for PDFs/images
- **Amazon Transcribe Medical** – audio transcription
- **Amazon Bedrock** – summarization, simplification, structured extraction, watch-outs, personalized action plan generation
- **Bedrock Embeddings** + **Amazon OpenSearch Service** – RAG/vector retrieval for glossary + billing + consent explainers
- **AWS Step Functions** – orchestration of the full workflow
- **AWS Lambda** – validation, confidence scoring, rules, formatting, routing
- **Amazon DynamoDB** – provider directory + insurance network mapping + patient preference profiles + episode metadata
- *(Optional)* **Amazon SageMaker** – custom doc classifier or specialty-routing model
- **Amazon A2I** – human review on low-confidence or high-risk outputs
- **CloudWatch + IAM + KMS** – monitoring + permissions + encryption
- *(Optional)* **Amazon Cognito** – user authentication

---

5) Future Expansion - Kritika

- **Signature documentation reading during/after visits** (real-time detection in patient portals/tablet forms)
- **Real provider availability + scheduling integrations** (health system APIs)
- **Real insurance eligibility checks + prior authorization workflows** (where permitted)
- **EHR interoperability (FHIR export/import)**
- **Operational automation** (referral packet assembly, post-visit follow-up automation)
- **Compliance & audit tooling** (traceability from extraction → model outputs → reviewer edits → final output)
- **Personalized aftercare monitoring** (reminders + symptom check-ins tied to discharge plan, escalations based on documented red flags)