System Design - Medical Chronologies

Design Challenge
Build a system that ingests large volumes of medical records—both text-native PDFs and scanned images—and produces a consolidated medical chronology for personal-injury attorneys.

Functional Requirements

- **Document Ingestion:** The system shall accept bulk uploads of text-based and image-based medical records.
- **Event Extraction:** The system shall identify and extract clinical events (date, encounter type, provider, free-text notes).
- **Chronology Assembly:** The system shall merge, sort (by date), and deduplicate all extracted events into a single timeline.
- **Chronology Interface:** The system shall expose a mechanism (e.g. UI or API) for users to request and view the assembled chronology.

Design Scope

In your solution, cover:

Overall architecture (ingestion, ML pipelines, storage layers)

Interfaces (user-facing UI/API endpoints as well as between internal systems/models)

Data flows (from upload through extraction to timeline)

Storage mechanisms (raw files, structured event store)

Excalidraw: https://link.excalidraw.com/l/7XY6n100Pgb/rokmTcIAXl

Personal Medical Record

PREPARED FOR	BIRTHDAY	GENDER
Ellen Ross	**12/17/1960** (54)	**Female**

💉 IMMUNIZATIONS continued from page 1

DATE	IMMUNIZATION NAME		TYPE
September 23, 2006	**DTaP (diptheria, tetanus, pertussis)**, IM		Intramuscular injection
PROVIDER		PHONE	DOSE QUANTITY (VALUE/UNIT)
Ashby Medical Center		(000) 555-1200	50 / micrograms
INFORMATION/INSTRUCTIONS			BOOSTER
Mild pain or soreness in the local area following injection.			Every 10 years (Sept 2012)

📋 ENCOUNTERS: PAST APPOINTMENTS

DATE	ENCOUNTER	
April 13, 2013	**ED visit for Ankle Sprain**	
PROVIDER	LOCATION	PHONE
Dr. Henry Seven	Community Hospital	(000) 555-1200

DATE	ENCOUNTER	
September 28, 2012	**Gall Bladder Surgery**	
PROVIDER	LOCATION	PHONE
Dr. Bala Venktaraman	Ashby Medical Center	(000) 555-1200

DATE	ENCOUNTER	
March 21, 2012	**Labor & Delivery**	
PROVIDER	LOCATION	PHONE
Dr. Tiffany Martinez	Ashby Medical Center	(000) 555-1200

DATE	ENCOUNTER	
October 28, 2010	**ED visit for Acute Cholecystitis**	
PROVIDER	LOCATION	PHONE
Dr. Tim Lee	Ashby Medical Center	(000) 555-1200

🔍 RECENT HEALTH OBSERVATIONS

DATE	OBSERVATION	STATUS
March 28, 2014	**Acute asthma**	**Active**
COMMENTS		
Current treatment with Bricanyl inhaler; awaiting results of comprehensive examinations.		

DATE	OBSERVATION	STATUS
November 30, 2010	**Cholecystitis**	**Resolved**
COMMENTS		
Complete recovery from surgical procedure.		

⚕ PROCEDURES

DATE	PROCEDURE	
March 22, 2012	**Cesarian Section**	
PROVIDER	LOCATION	PHONE
Dr. Tiffany Martinez	Ashby Medical Center	(000) 555-1200

DATE	PROCEDURE	
October 28, 2010	**Laparoscopic Cholecystectomy**	
PROVIDER	LOCATION	PHONE
Dr. Bala Venktaraman	Ashby Medical Center	(000) 555-1200

DATE 07/04/14 COMPLAINT _paracitos, vitaminas._

MALE (FEMALE) HT._____ WT._19 los_ B/P_____ TEMP _97._ PULSE _130_

Sibilancias / *Wheezing*	Dolor de cabeza / *Headache*	Dolor abdominal / *Abdominal Pain*	Problemas para respirar / *Breathing Problems*	Tos — well chec Cough — vitamins
El lloro / *Crying*	Fiebre / *Fever*	Dolor de pecho / *Chest Pain*	Problemas menstrual / *Menstrual Problems*	Cargue la perdida / *Weight loss*
Fatiga / *Fatigue*	Diarrea / *Diarrhea*	Nariz Liquida / *Runny Nose*	Apoye el dolor / *Back Pain*	Palidezca/anemia / *Pale/Anemia*
Picazón en los ojos / *Itchy Eyes*	Sudores / *Sweats*	Calambres / *Cramps*	Dolor de hueso / *Bone Pain*	Garganta adolorida / *Sore Throat*

PERTINENT MEDICAL HISTORY: The patient is 1yr old and weigh
If yes, explain   19 lbs. Not eating well
No diarrea. No N/V.

	NO	YES
Heart Disease	✓	
Hypertension	✓	
Diabetes	✓	
Asthma	✓	
Renal	✓	
Pregnant	✓	
Other		
Past History		

Allergies _____   Current Medications _____

PHYSICAL EXAM: Normal Abnormal

Thin apearing but
active 1yr old female
c/w — Patient is noving from
Mother during this office visit

	Normal	Abnormal
EENT	—	—
Heart	—	—
Lungs	—	—
Abdomen	—	—
Genitalia	—	—
Extremities	—	—

LAB: Hb.____ Glucose____ Malaria____ Urine____

IMPRESSION Low ~~birth~~ Weight

TREATMENT Mebendazole use a diretle
Infant formula with Iron
RTC 1 month

Ingestion
1. Text ingest

2. Image ingest

DB tables
1. Patients
   a. Id
   b. Name
   c. Dob
   d. Address
2. Document
   a. Id
   b. Patient_id (fk to patients)
   c. Aws_url
3. Provider
   a. Id
   b. Name
   c. Phone num
   d. Address
   e. license
4. Events
   a. Id
   b. Event_name
   c. Event_notes
   d. Provider_id (fk to providers)
   e. Document_id (fk to docs)
   f. Procedure_date
   g. Duplicate
   h. Duplicate_reason
   i. duplicate_event_id

Ingestion flow (text):
1. Get the information (text, patient info, provider info)
2. Ingestion
   a. LLM prompt with few shot examples
   b. Input would be pdf converted to markdown

Ingestion flow (image):
3. Get the information (text, patient info)
4. Ingestion
   a. OCR
   b. LLM prompt with few shot examples

Interface API
Get api

events/<user_id>/