AMIA
INFORMATICS PROFESSIONALS. LEADING THE WAY.

Amplify
INFORMATICS CONFERENCE

EXACT: LLMs for Structured Eligibility Extraction for Clinical Trial Matching

Session Code: TRI40

Adam Blum

HealthKey

adam@healthkey.ai

Disclosure of Relevant Financial Relationships

- Adam Blum, HealthKey, Inc., CTO, equity

Learning Objectives

At the conclusion of this presentation the learner will be able to:

- Explain why clinical trial matching requires structured eligibility representation
- Describe how the ACE Engine uses LLMs and Prompt Workbench to extract criteria from trial text
- Represent eligibility criteria in Conjunctive Normal Form (CNF) for automated evaluation
- Apply the PATCH Engine to harmonize patient attributes against CNF criteria
- Use the PatientInfo 266-field flat table for fast, stateless patient-trial matching
- Interpret Eligible / Potential / Ineligible verdicts and the Suitability Score
- Embed the open-source EXACT engine in a larger PHR or clinical platform

The Clinical Trial Matching Challenge

- Why matching is hard:
  - ~400,000 active clinical trials globally with complex eligibility criteria
  - Criteria written in natural language — not machine-readable
  - Patient records are fragmented across EHRs, labs, genomics reports
  - Manual matching by coordinators takes hours per patient
  - Result: <5% of eligible cancer patients enroll in trials (NCI estimate)

- What EXACT does:
  - Automatically extracts structured criteria from trial text using LLMs
  - Matches those criteria against a structured patient record
  - Returns Eligible / Potential / Ineligible verdicts with explanations

EXACT Architecture

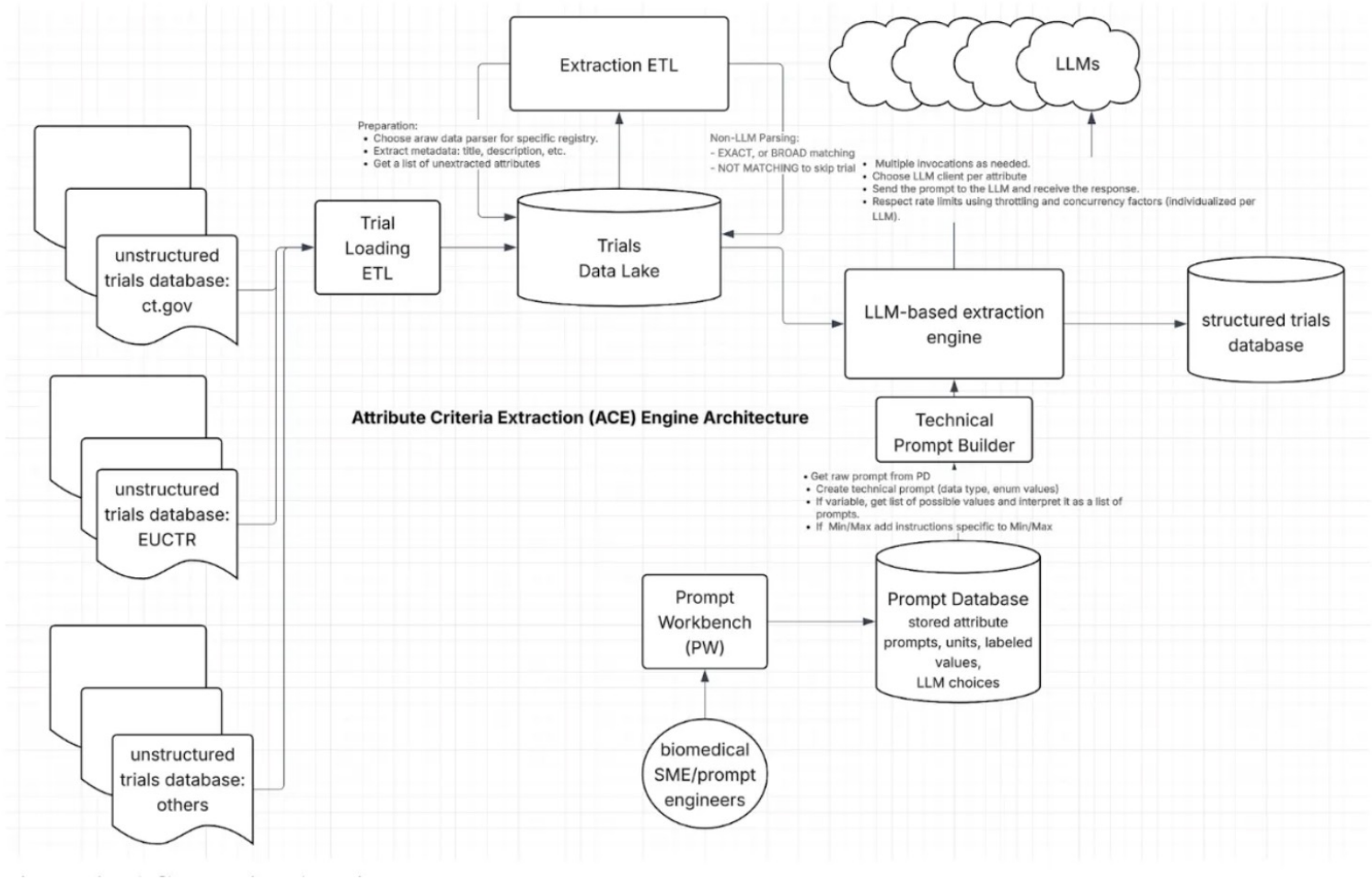

Figure 1 — ACE Engine Architecture

EXACT: A Two-Phase Open-Source Pipeline

- Phase 1 — Criteria Extraction (offline, per-trial):
  - ACE Engine reads free-text eligibility criteria from ClinicalTrials.gov
  - Prompt Workbench sends targeted LLM prompts per attribute type
  - Outputs structured criteria in CNF (Conjunctive Normal Form)
  - Stored in CTOMOP trial tables for reuse across all patients
- Phase 2 — Patient Matching (online, per-patient):
  - PatientInfo flat table provides 266 denormalized patient attributes
  - PATCH Engine evaluates CNF criteria clauses against patient attributes
  - Returns verdict + per-criterion explanations in milliseconds
  - Stateless design: no session state; scales horizontally

Open source: github.com/healthkey-ai/exact

ACE: Attribute-Criteria Extraction with LLMs

- Problem: Eligibility criteria are complex prose
  - E.g. 'Patients must have received ≥2 prior platinum-based regimens, no corticosteroids, ECOG<=2'
  - Must be decomposed into machine-evaluable attribute-value pairs

ACE Engine approach:

- Receives raw inclusion/exclusion criteria text from ClinicalTrials.gov API
- Invokes the Prompt Workbench to generate one LLM call per attribute type
- Each prompt asks: 'Does this criterion constrain attribute X? If so, how?'
- Attribute types: diagnosis, biomarker, stage, prior therapy, performance status, age, lab value, genomic variant, …
- LLM returns structured JSON with operator (>=, <=, =, IN), value, and unit

Output: a set of (attribute, operator, value) triples per criterion clause

These triples are the input to PATCH matching

Prompt Workbench: Precision LLM Prompting

- Challenge: One generic prompt cannot extract all criterion types accurately
- Prompt Workbench solution:
  - Library of purpose-built prompts — one per attribute
  - Prompt outputs measured real-time against labeled criteria from real ClinicalTrials.gov studies
- Key design decisions:
  - Separation of concerns: each prompt focused on one attribute
  - Structured JSON output enforced via function-calling / grammar constraints
  - Prompt versioning enables A/B testing and quality regression tracking
- Result: High precision extraction, auditable per criterion per attribute

CNF: Logic Representation of Eligibility Criteria

- What is CNF (Conjunctive Normal Form)?
  - A boolean formula written as AND of OR-clauses
  - Example: (A OR B) AND (C OR D) AND E
  - DeMorgan's law allows any arbitrarily complex criteria to be transformed to CNF
- Why CNF maps naturally to trial eligibility:
  - ALL major criteria must hold simultaneously — the AND
  - A patient fails the trial if any single AND-clause is unsatisfied
  - Can be executed quickly
  - Patients and doctors can easily understand
- OR criteria are represented as named attributes
  - MeetsCRAB= Hypercalcemia OR renal insufficiency OR anemia OR bone lessions
  - MeetsSLiM
  - Renal insufficiency itself

PATCH : Patient-Trial Criterion Harmonization

- Core function:
  - For each CNF clause, check if the patient satisfies at least one atom
  - A patient satisfies a clause if any OR atom evaluates to true
  - Patient is eligible only if ALL clauses are satisfied (AND logic)
- Harmonization challenges PATCH solves:
  - Unit normalization: mg/dL vs. mmol/L for lab values
  - Temporal reasoning: 'received within last 6 months' vs. absolute date
  - Concept mapping: ICD-10 diagnoses to OMOP standard concepts via Athena
  - Negation handling: 'no prior anthracycline' must invert patient history
- Output per clause:
  - SATISFIED / UNSATISFIED / UNKNOWN (insufficient patient data)
  - Explanation string for each evaluated atom (shown to patient/coordinator

PatientInfo: 266-Field Longitudinal Patient Record

- Challenge: OMOP normalized tables require many JOINs for matching
- Solution: PatientInfo — a denormalized flat projection for fast attribute lookup
- PatientInfo field categories (266 total):
  - Demographics: age, sex, race, ethnicity, geographic region
  - Diagnosis: primary cancer type, histology, ICD-10/SNOMED codes
  - Stage & Grade: TNM staging, AJCC edition, grade at diagnosis
  - Genomics: EGFR, ALK, KRAS, BRAF, HER2, MSI, TMB, PD-L1, 50+ biomarkers
  - Treatment History: all prior drug exposures with start/end dates
  - Lines of Therapy: automatic LOT assignment from episode table
  - Lab Values: CBC, CMP, LFTs, creatinine, most recent values + dates
  - Performance Status: ECOG, Karnofsky, most recent scores

- Key design properties:
  - Computed once per patient update, cached for sub-millisecond access
  - Versioned: each PHR update creates a new PatientInfo snapshot
  - Stateless matching: PatientInfo row is the only input PATCH needs

Matching Verdicts & Suitability Score

- Three-tier verdict system:
  - ELIGIBLE — satisfies all inclusion + no exclusion criteria
  - POTENTIAL — most criteria met; some attributes unknown/borderline
  - NELIGIBLE — definitively fails one or more hard exclusion criteria

- Suitability Score (0-100)
  - Rates trials on patient preferences of risk, benefit, burden, distance
  - Benefit: ESMO-MCBS scaled 1-20
  - Burden: Getz' framework (less distance)
  - Risk: see https://medium.com/cancerbot/what-makes-a-trial-good-for-a-patient-a43e5b651754
  - Distance: thresholds ~= ln(distance)

Evaluation: Matching Accuracy Results

- 98 multiple myeloma trials labeled attributes from SMEs to assess accuracy
- Results at accuracy.cancerbot.org

Category	Micro F1	Micro Precision	Micro Recall	Micro Mcc
Treatment	0.8047	0.8747	0.7451	0.758
Labs	0.9702	0.9744	0.9661	0.1941
Disease	0.8882	0.8565	0.9222	0.8211
Blood	0.9386	0.9427	0.9345	0.064
Behavior	0.9667	0.9812	0.9526	0.6304

EXACT in Large Clinical Ecosystem

- Upstream (feeds EXACT):
  - EHR importers (Epic, Cerner, VA, One Up Health) -> FHIR R4 -> OMOP ETL
  - Genomics (Foundation Medicine, Tempus) -> OMOP Genomics Extension
  - Patient-reported outcomes -> PatientInfo fields
- EXACT core (runs in cloud functions):
  - ACE Engine runs nightly to update criteria for new/amended trials
  - PATCH Engine triggered on each PatientInfo update
- Downstream (consumes EXACT output):
  - HealthTree
  - CancerBot
  - Harvard DCI Network

Open Source EXACT: Getting Started

- Repository: github.com/healthkey-ai/exact
- License: Apache 2.0 — free to use, modify, and embed
- What is included:
  - ACE Engine
  - PATCH Engine with unit normalization and OMOP concept mapping
  - DockerHub.com/healthkey/exact
- How to contribute:
  - Add diseases with identified eligibility attributes
  - Extend trial schema in existing diseases for new trial eligibility attributes
  - Submit benchmark evaluation datasets
- Contact: adam@healthkey.ai