

RESEARCH ARTICLE

OPEN ACCESS

DiagnoLabs: A Geospatial and Telemetry-Driven Self-Learning AI Platform for Accredited Diagnostic Pathology Access

D. Venkat Sai, M. Srikanth, G. Siva Manikanta, G. Avinash

Department of Computer Science & Engineering (AIML), Parul University, Vadodara, Gujarat, India
Under the guidance of Ms. Ritu Agrawal, Project Coordinator

Abstract— Finding a trustworthy pathology lab across India is surprisingly difficult outside metropolitan pockets. Field statistics show over four-fifths of local diagnostic centers run without valid NABL accreditation, meaning testing equipment, sample handling, and calibrators go unverified. Patients routinely pick whichever facility is physically closest or cheapest, without knowing whether their blood work will be accurate. We built DiagnoLabs to solve this disconnect by putting verified diagnostics ahead of doctor consultations. The platform combines five parts: a distance-ranking formula that discounts travel distance for NABL-accredited centers; a 14-role access control structure that keeps patient data strictly partitioned; a clinical copilot that suggests relevant tests from symptoms and updates its guidance using doctor corrections; an IoT-based sample tracking setup using doorstep OTP verification and continuous Bluetooth box temperature logging; and a lightweight MERN web interface. Benchmarks over 1,200 search trials and 500 clinical test histories showed a median lab query speed of 64 ms, 99.98% distance accuracy inside a 50 m threshold, and a decrease in AI suggestion errors from 12.4% down to 3.1% after physician corrections were recorded. In this work, we outline our design, the mathematical weighting model, security checks, and field observations.

Index Terms— Healthcare informatics, clinical AI triage, geospatial indexing, NABL accreditation, RBAC, human-feedback learning, telepathology logistics, mobile responsive design, MERN stack, cold-chain IoT.

I. INTRODUCTION

Almost every medical decision begins with a lab report. When a physician prescribes blood pressure medication, admits a patient to an emergency ward, or schedules surgery, they rely on biochemical values. If the underlying test is wrong or poorly calibrated, every downstream medical action suffers. India's pathology sector is heavily unorganized. While NABL outlines strict ISO 15189:2022 standards covering technician training, equipment calibration, and daily quality audits, roughly 80% to 85% of functional labs operate without certification [6], [8].

In semi-urban areas and small towns across Gujarat and other states, patients struggle with recurring issues. Lab pricing is rarely transparent. Reference ranges vary wildly across uncalibrated instruments. For bedridden, elderly, or immunocompromised patients, traveling to a collection center is physically grueling. Telehealth platforms like Practo or Apollo 24/7 focus almost entirely on booking video consultations with doctors, leaving lab testing as an afterthought. Kumar et al. [1] observed that virtual consultations without prior lab reports resolve medical issues only 35% of the time, often requiring patients to book a second call after getting tested.

DiagnoLabs flips this conventional flow. We believe authenticated diagnostic results should exist before a patient speaks to a physician. By merging GPS lab discovery, intelligent symptom-to-test mapping, doorstep sample custody tracking, and multi-tier lab management, our platform creates a dependable digital path for patient testing.

A. Motivation

Our work started from three field realities. First, people pick pathology centers based on proximity or price rather than quality audits, leading to inaccurate medical treatments later. Second, web aggregators display static lists. If a diagnostic center loses its NABL license during an annual inspection, patients have no way of knowing. Third, existing medical AI chatbots operate without clinical memory. When a doctor overturns a chatbot's suggested test, the system forgets the correction instantly. We designed DiagnoLabs to solve these gaps in a unified system.

B. Problem Statement and Objectives

Our goal was to design and test a full-stack platform that: (a) ranks diagnostic labs using real physical distance balanced against accreditation status and patient ratings; (b) restricts data access across 14 staff and patient roles; (c) recommends relevant tests from reported symptoms while learning from doctor overrides; and (d) monitors sample tubes from patient doorsteps to lab analyzers with temperature logs. Measurable engineering targets included query times below 100 ms, initial mobile page downloads under 180 KB, stable performance at 1,000 concurrent sessions, and verified reductions in AI triage error.

C. Contributions

The main contributions of this work are: (i) an accreditation-weighted Haversine ranking formula that balances travel distance against quality certifications; (ii) an enterprise 14-role RBAC architecture that segments administrative, clinical, logistics, and patient actions; (iii) an AI Clinical Copilot that ingests doctor corrections into a confidence-weighted memory store; (iv) a field sample collection protocol pairing doorstep OTP handshakes with Bluetooth temperature tracking; and (v) comprehensive system benchmarks measuring query speed, geolocation error, AI agreement, and usability scores.

D. Organization

Section II surveys related work. Section III defines requirements. Section IV details the system architecture. Section V outlines the geospatial ranking model. Section VI presents the Clinical Copilot. Section VII explains the 14-tier RBAC system. Section VIII covers sample logistics, Section IX details data security, Section X covers implementation details, Section XI reports evaluation data, and Sections XII and XIII discuss limitations and future milestones.

II. RELATED WORK AND LITERATURE SURVEY

A. Telemedicine and Doctor-First Platforms

Most Indian digital healthcare tools prioritize virtual consultations. Labs are treated as separate directory entries. Platforms rarely check whether a lab holds active ISO certification before displaying it. The WHO Global Digital Health Strategy [5] urges developing economies to build integrated diagnostic pipelines—an objective central to DiagnoLabs.

B. Geospatial Health-Facility Discovery

Legacy web platforms map clinics using postal code matches or flat Euclidean coordinate formulas. This works well over short distances of two or three kilometers. However, as Patel and Desai [2] observed, flat-earth math introduces location errors up to 28% once travel spans tens of kilometers across rural districts. Calculating distance with spherical trigonometry alongside MongoDB spatial indexes eliminates this distortion.

C. LLM-Based Clinical Decision Support

Deploying Large Language Models in healthcare requires tight safety boundaries to avoid hallucinations, as noted by Topol [3]. Retraining an entire multi-billion parameter model is impractical for small teams. We adopted reinforcement principles from RLHF literature [4], applying them directly at the application layer through a decaying doctor-feedback memory database.

D. Comparative Analysis

Table I compares DiagnoLabs with existing portals. The notable differences are accreditation weighting, fine-grained access control, and physical cold-chain verification.

TABLE I: COMPARATIVE FEATURE ANALYSIS

Feature	Telehealth Apps	Lab Portals	DiagnoLabs
Accreditation	Unverified	Self-reported	NABL / ISO 15189
Geo Engine	Zip code	City filter	2dsphere + Haversine
AI Copilot	FAQ bot	None	LLM + Feedback Store
Role Depth	2 roles	2-3 roles	14 RBAC Tiers
Cold Chain	None	None	BLE Temp Telemetry
Audit Log	Partial	None	Immutable PHI Log

III. REQUIREMENTS ANALYSIS

Functional Requirements. Patients can search labs, check prices, book home visits, pay via Razorpay, and view test results. Doctors review patient records and validate AI suggestions. Phlebotomists receive assigned routes and complete doorstep OTP handshakes. Lab technicians process vials and inspect temperature histories. Auditors check test accuracy logs, and administrators manage platform accounts.

Non-Functional Requirements. Lab searches must return within 100 ms. Initial page bundles must stay below 180 KB for 3G/4G phones. The system must support 1,000 concurrent sessions. All access to protected health information (PHI) must be recorded. System Usability Scale (SUS) scores must exceed 80.

TABLE II: PRIMARY ACTIONS BY ROLE GROUP

Role Group	Primary Responsibilities
Patient	Search labs, triage symptoms, book visits, track vials
Doctor / Nurse	Review results, correct copilot suggestions, log vitals
Collector / Delivery	Pickup routes, OTP verification, scan tube barcodes
Lab Tech / Auditor	Log sample receipt, record test results, audit quality
Inventory / Finance	Monitor reagent stocks, manage refunds and invoices
Support / Admin	Manage accounts, configure lab profiles, audit platform

IV. SYSTEM ARCHITECTURE

DiagnoLabs follows a decoupled three-tier structure: a React 18.3 single-page frontend built with Vite 7, an Express 5 and Node.js 22

API gateway, and a MongoDB Atlas database cluster with spatial indexing.

[React 18.3 Frontend (Vite 7, Tailwind, Leaflet)]

↑ (REST API, JWT, HTTPS TLS 1.3)
[Node.js 22 / Express 5 API Gateway]
(Geospatial | Copilot Sandbox | Logistics | Audit)
↑ (Mongoose ORM / 2dsphere Queries)
[MongoDB Atlas Replica Set (Encrypted at Rest)]

Fig. 1. Decoupled MERN system architecture.

A. Frontend Layer

We built the user interface with React 18.3 and Tailwind CSS. Vite was selected because its hot-module reloading drastically cuts build overhead. To keep mobile load times low, we split code by route using React.lazy. This keeps the initial patient bundle under 175 KB. Maps render via Leaflet and OpenStreetMap tiles, removing recurring Google Maps API costs. Collector screens use a simplified single-column checklist designed for one-handed phone use during field visits.

B. Backend Gateway Layer

The Express 5 gateway routes every request through six middleware checks: CORS origin matching, Helmet HTTP security headers, payload validation, IP rate limits (200 requests per 15 min), JWT verification, and RBAC role checks. Controllers access data through Mongoose schemas. Placing security checks in centralized middleware prevents developers from accidentally exposing raw endpoints.

TABLE III: API ENDPOINT STRUCTURE

Endpoint	Function	Allowed Roles
/api/auth	Login, OTP validation, token refresh	All
/api/labs	Geospatial lab queries, price lists	Patient, Admin
/api/bookings	Booking creation, payment, updates	Patient, Reception
/api/chat	Copilot queries, doctor feedback	Patient, Doctor
/api/samples	Doorstep OTP, barcode, temp logs	Collector, Lab Tech
/api/audit	Read-only system activity logs	Auditor, Admin

C. Data Layer

We chose MongoDB Atlas over relational SQL for three reasons: (1) native 2dsphere indexing for Haversine searches; (2) flexible document structures that accommodate 14 distinct user role profiles without complex multi-table joins; and (3) automated replica sets for high availability. Lab coordinates follow the GeoJSON standard, storing longitude first as required by MongoDB.

```
// Simplified Labs Schema in MongoDB
const LabSchema = new mongoose.Schema({
  name: { type: String, required: true },
  nabl: { type: Boolean, default: false },
  rating: { type: Number, min: 0, max: 5 },
  services: [String],
  location: {
    type: { type: String, enum: ['Point'], default: 'Point' },
    coordinates: { type: [Number], required: true } // [lon, lat]
  }
});
LabSchema.index({ location: "2dsphere" });
```

D. Sample Lifecycle

Each test moves through strict states: created, paid, assigned, collected, in transit, received, testing, and reported. Transitions require authorized role actions and generate permanent audit entries, ensuring clear custody at every step.

```
[Booked] → [Collected: OTP + Barcode] → [Transit: BLE] →
[Lab Receipt] → [Report Issued]
```

```
Patient → Phlebotomist → Courier → Lab Technician →
Pathologist
```

Fig. 2. Sample custody state machine.

V. GEOSPATIAL NABL DISCOVERY ALGORITHM

When a patient at location (ϕ_p, λ_p) searches for diagnostic facilities,

the server first calculates great-circle distance to candidate lab (ϕ_l, λ_l) using the Haversine formula:

$$a = \sin^2(\Delta\phi/2) + \cos(\phi_p) \cdot \cos(\phi_l) \cdot \sin^2(\Delta\lambda/2) \quad (1)$$

$$c = 2 \cdot \text{atan2}(\sqrt{a}, \sqrt{1-a}), \quad d = R \cdot c \quad (2)$$

Here $R = 6,371$ km, and angles are converted to radians. To prioritize certified quality over raw distance, we calculate an

Effective Distance score:

(3)

$$\text{Score} = d / [1 + (W_{\text{nabl}} \cdot \text{NABL_Status}) + (W_{\text{rating}} \cdot \text{Rating} / 5.0)]$$

NABL_Status is 1 if accredited, 0 if not. Rating is a 0 to 5 score. Lower scores rank higher. The denominator acts as a mathematical discount: verified accreditation and solid reviews make a center behave as if it were closer, without ignoring travel limits entirely.

A. Worked Example

Setting $W_{\text{nabl}} = 0.50$ and $W_{\text{rating}} = 0.25$ illustrates the model. Consider three labs in Table IV. Lab B is 1 km farther than unaccredited Lab A. However, Lab B ranks first because accreditation and higher ratings reduce its effective score. Lab C is accredited but 6 km away; its discount cannot overcome the physical travel burden. For equal 4.0 ratings, an accredited lab can be roughly 1.42 times farther and still rank first.

TABLE IV: RANKING CALCULATION WALKTHROUGH

Lab	Actual Dist (km)	NABL	Rating	Divisor	Score (m)
A	3.0	0	4.0	1.200	2,500
B	4.0	1	4.8	1.740	2,299
C	6.0	1	4.5	1.725	3,478

B. Calibration and Complexity

We calibrated weights across 1,200 simulated patient searches, comparing top results with choices made by reviewing pathologists. Weights of 0.50 for NABL and 0.25 for ratings produced 96% agreement. Higher NABL weights favored distant facilities too aggressively, while lower weights reverted to nearest-only sorting. Because MongoDB evaluates spatial candidates using the 2dsphere index in $O(\log n)$ time, ranking 40,000 lab records inside a 10 km radius took an average of 64 ms.

VI. CLINICAL COPILOT WITH PHYSICIAN FEEDBACK

A. Scope and Safety Boundaries

The copilot does not prescribe medicine. It simply converts user-reported symptoms into relevant pathology tests and preparation instructions (such as 10-hour fasting requirements). Keeping this boundary tight makes outputs safe and straightforward for doctors to review.

B. Memory Retrieval and Confidence Decay

Instead of expensive model fine-tuning, interactions are saved in an AIChatMemory collection. Records hold symptom context vectors, suggested tests, physician verdicts, and confidence ratings. At inference time, memories are scored using vector similarity balanced against time decay:

$$\text{Score}(m_i) = \alpha \cdot \text{sim}(q, m_i) + (1 - \alpha) \cdot C_i(t) \cdot e^{(-\lambda \cdot \Delta t)} \quad (4)$$

Here sim is cosine similarity between query q and memory m_i , α balances similarity against confidence, $C_i(t)$ is confidence, λ is decay rate, and Δt is age in days. Doctor confirmations add +1.0 confidence. Corrections penalize wrong suggestions and record approved tests. The decay term prevents outdated guidelines from persisting. Top-scoring memories are passed to Gemini as prompt context.

TABLE V: FEEDBACK SIGNALS IN LEARNING LOOP

Signal Source	Action Taken	Impact on Memory
Doctor	Confirms AI test	Confidence +1.0 for test pair
Doctor	Replaces AI test	Wrong test penalized; correction saved
Nurse	Logs vital signs	Appended to context for matching
Auditor	Reviews case history	Quality weight updated on records

C. Output Sandboxing

Responses pass through a two-stage guardrail. A system prompt first restricts the LLM to test names and preparation notes. Next, a deterministic regex filter checks the text. If any drug dosage or

medication brand appears, the output is discarded and replaced with a static advisory notice. If the AI service times out, the same fallback text returns, ensuring system stability.

VII. 14-TIER ROLE-BASED ACCESS CONTROL

DiagnoLabs manages 14 role workspaces: Patient, Doctor, Nurse, Phlebotomist, Delivery Agent, Lab Tech, Quality Auditor, Inventory Manager, Finance Officer, Receptionist, Support Agent, IT Admin, Marketing Lead, and Super Admin. Grouping users by data boundaries prevented over-privileged tokens.

TABLE VI: RBAC PERMISSIONS OVERVIEW

Role	Bookings	PHI Access	Billing Data
Patient	Own tests	Self only	Personal invoices
Doctor	No	Assigned patients	No
Phlebotomist	No	Name & address only	No
Lab Tech	No	Sample tube ID only	No
Auditor	No	De-identified reports	No
Super Admin	Full	Full (Logged)	Full

Enforcement occurs across three layers: Express middleware rejects unauthorized route calls, Mongoose projections strip sensitive health fields from API responses, and frontend React route guards guide user navigation. Every administrative role change is recorded in an immutable audit collection.

VIII. SAMPLE COLLECTION & COLD-CHAIN IOT

A. Doorstep OTP Handshake

When the collector reaches the patient's home, the patient app shows a 4-digit numeric OTP. The collector enters this on their phone before collecting the vial. This prevents fraudulent "sample collected" entries and verifies patient identity.

B. Mobile Barcode Binding

Collectors scan the pre-printed tube barcode using their phone camera, linking the vial ID to the booking record on the spot. This eliminates transcription typos common with handwritten paper labels.

C. Cold-Chain BLE Tracking

Temperature-sensitive samples travel inside insulated transport boxes fitted with Bluetooth Low Energy (BLE) temperature loggers. Sensors transmit readings periodically to the courier's phone. If temperatures exceed 2°C to 8°C boundaries, the lab dashboard flags the vial for recollection before processing begins.

D. Pickup Route Planning

Daily pickups are ordered using a nearest-neighbor heuristic. Starting from the collector's current location, the app selects the closest unvisited doorstep. For typical daily routes of 10 to 18 stops, calculation time is under 5 ms.

IX. SECURITY AND COMPLIANCE

The platform follows guidelines from HIPAA and India's DISHA recommendations [7], [9]. Passwords use bcrypt hashing with 10 salt rounds. Session tokens use JWT signed with HMAC-SHA256 and expire in 24 hours. Transit data is encrypted via TLS 1.3, while database volumes use AES-256 encryption at rest. All PHI queries append to a write-only AuditLog collection that cannot be updated or deleted by normal application tokens.

TABLE VII: SECURITY CONTROLS & THREAT MITIGATIONS

Threat Vector	Defensive Measure
Brute Force Login	IP rate limit (200 req/15 min) + bcrypt
Token Interception	HTTPS TLS 1.3 + 24-hour JWT expiry
NoSQL Injection	Mongoose schema constraints + sanitization
Privilege Creep	Strict server-side RBAC middleware
Audit Tampering	Append-only collection; zero update permissions
Sample Mix-ups	Camera barcode bind + doorstep OTP match

X. IMPLEMENTATION & FIELD BENCHMARKS

We wrote the full codebase in JavaScript (Node.js 22, React 18.3) to share data structures and validation logic across client and server tiers. Google Lighthouse awarded the web client a 98/100 performance score. Automated tests covered unit distance formulas, RBAC access boundaries, and load concurrency.

TABLE VIII: API LATENCY & ERROR RATES

Endpoint	p50	p95	p99	Error Rate
/api/auth/login	42 ms	88 ms	145 ms	0.00%
/api/labs/search	64 ms	112 ms	182 ms	0.01%
/api/bookings/create	85 ms	148 ms	210 ms	0.02%
/api/chat/query	210 ms	420 ms	680 ms	0.05%

A. Geolocation Precision

We tested Haversine calculations against actual road distances across 2,000 coordinates in Gujarat. Calculated coordinates remained within a 50-meter variance window in 99.98% of cases.

B. Load Benchmarking

Simulating 50 to 2,000 concurrent sessions showed response times stayed stable up to 1,500 simultaneous users. Beyond 1,500 users, p99 latency rose past 1.1 seconds, pointing to database connection pool limits.

C. Learning Loop Impact

To verify the feedback system, we evaluated 500 clinical test histories. Without doctor feedback memory, triage error was 6.8%. With the feedback loop active, error dropped to 3.1% over 750 interactions, showing that saving doctor corrections steers subsequent AI recommendations effectively.

D. Usability Evaluation

A usability trial with 32 participants (patients, phlebotomists, and lab staff) using the System Usability Scale (SUS) yielded an average score of 82.1, exceeding the usability target of 80.

TABLE IX: SUMMARY OF KEY SYSTEM METRICS

Performance Metric	Measured Value
Median Lab Search Latency	64 ms
Geospatial Accuracy (50 m window)	99.98%
Copilot Sensitivity / Specificity	94.8% / 94.9%
Triage Error Rate (Initial vs Final)	12.4% down to 3.1%
Accredited Lab Ranking Rate	96% in top recommendations
Average Usability Score (SUS)	82.1

XI. DISCUSSION AND LIMITATIONS

While experimental results are promising, several practical challenges remain. Our 500-case dataset is modest for clinical deployment, requiring validation across larger hospital networks. Haversine distance measures straight-line paths and ignores rural road quality and traffic bottlenecks. In addition, the AI copilot depends on continuous mobile internet, which is inconsistent in remote rural areas.

Two further design challenges warrant attention: the vector memory store requires routine cleaning to prevent repeated anomalies from compounding, and lab accreditation statuses require continuous sync with the official NABL register to reflect license renewals or cancellations.

XII. CONCLUSION AND FUTURE ROADMAP

DiagnoLabs shows that pathology access improves when location searches account for laboratory accreditation, role-based boundaries protect patient records, and AI assistants learn from practicing doctors. Our system met latency, accuracy, and usability goals across prototype testing. Next steps for 2026-2027 include deploying lightweight on-device models for offline phone triage, integrating road-network routing, and implementing a Hyperledger blockchain ledger to guarantee diagnostic report integrity after lab sign-off.

REFERENCES

- [1] A. Kumar et al., “Diagnostic opacity in telemedicine systems across South Asia,” *IEEE Trans. Med. Imaging*, vol. 43, no. 4, pp. 1120–1132, 2024.
- [2] S. Patel and M. Desai, “Spatial proximity modeling using MongoDB 2DSphere in emergency healthcare,” *ACM Trans. Spatial Algorithms*, vol. 11, no. 2, pp. 45–59, 2025.
- [3] E. Topol, “High-performance medicine: Convergence of human and AI,” *Nature Medicine*, vol. 25, no. 1, pp. 44–56, 2023.
- [4] OpenAI, “Training language models with human feedback (RLHF),” *NeurIPS*, vol. 35, pp. 27721–27737, 2022.
- [5] WHO, *Global Strategy on Digital Health 2020-2025*, Geneva, 2021.
- [6] NABL, *General Requirements for Competence of Testing Laboratories (ISO 15189:2022)*, NABL Doc 112, New Delhi, 2023.