Smart Real-Time Monitoring and Inspection Mobile Application for DoSJE Schemes: A Traceable, Human-In-The-Loop Digital Monitoring Framework

Preethi S.[1], Ramya Shree S.[2], Pavin Sundar P.[3], Rajesh S.[4], Giriprasath G.[5], Karthika S.[6]
[1,2,3,4,5,6]*Department of Computer Science and Engineering, SNS College of Technology, Coimbatore, Tamil Nadu, India*

***Abstract*—Government welfare programs require reliable workflows for recording field activity, documenting inspections, and reviewing implementation evidence. The Department of Social Justice and Empowerment (DoSJE) lists ground-level monitoring, project monitoring, surveillance, social audit, and evaluation among the components of its Information-Monitoring, Evaluation and Social Audit (I-MESA) scheme. This paper proposes a design-oriented mobile-and-dashboard framework that supports these activities through a shared, traceable workflow. The proposed system comprises a mobile client, an authenticated backend application programming interface (API), validation and evidence services, a relational database with private object storage, and a government review dashboard.**

Its scope includes attendance events, configurable geofence checks, inspection assignment, time- and location-associated evidence, exception handling, audit logging, and optional biometric or closed-circuit television (CCTV) analysis. Global Positioning System (GPS) readings, biometric comparisons, photographs, and model-generated detections are treated as fallible signals, not conclusive proof of identity, presence, or compliance. The backend remains authoritative for authentication, authorization, state transitions, and audit records.

Contributions are an integration architecture, a workflow and data-provenance model, a least-privilege access matrix, a risk-oriented treatment of artificial intelligence (AI) outputs, and a controlled pilot protocol covering functional correctness, latency, synchronization reliability, geofence error, usability, and security. No measured performance, completed pilot, deployment, or field-impact results are reported; the paper defines a reproducible evaluation plan rather than claiming effectiveness. The framework is intended to assist authorized personnel and does not replace formal inspection or social-audit procedures.

***Index Terms*—digital government, evidence provenance, geofencing, human-in-the-loop AI, inspection workflow**

I. INTRODUCTION

A. Administrative context

Public welfare programs are delivered through administrative arrangements involving government departments, state and district authorities, institutions, staff, and intended beneficiaries. Monitoring is both an information-management task and an accountability task: reviewers need to know what activity was reported, who recorded it, when the record was received, what evidence supports it, and whether an authorized person reviewed it. The Department of Social Justice and Empowerment (DoSJE) describes Information-Monitoring, Evaluation and Social Audit (I-MESA) as a Central Sector Scheme intended to examine implementation at ground level and lists information dissemination, a Project Monitoring Unit, a Central Smart Surveillance Unit, social audit, and evaluation and studies among its components [1]. The published guidelines provide a policy reference for this context [2]. These sources establish context; they do not validate the proposed application or imply departmental endorsement.

B. Technical problem

Attendance registers, inspection schedules, photographs, video meetings, and narrative reports may be recorded in separate systems or communication channels. Fragmentation can increase the effort required to reconstruct an inspection history, locate missing evidence, compare statuses, and route exceptions. A shared digital workflow may improve retrieval and traceability, but digitization alone does

not guarantee record accuracy. Research on artificial intelligence (AI)-assisted computational optimization in solar-absorber design demonstrates a cross-domain approach in which candidate configurations are explored systematically and model predictions are checked against simulation conditions [3]. This is methodological inspiration only: the physical design variables and performance results in that study do not transfer to government inspection workflows. A reported location can be imprecise or manipulated; a photograph can be authentic yet incomplete; biometric comparison can produce false matches or false non-matches; and video detectors can misclassify a scene. The design problem is not merely to automate collection, but to preserve provenance, expose uncertainty, apply access controls, and support accountable review.

C. Research gap and positioning

The cited policy material establishes the monitoring context, while prior work provides methods for face representation and object detection [4]–[6]. These sources do not evaluate this proposed end-to-end workflow. The paper addresses a system-integration problem: how attendance, assignment-driven inspection, evidence storage, reviewer decisions, and optional automated signals can be represented in one auditable workflow without treating any single signal as conclusive.

The study follows a design-oriented approach in which an artifact is specified and its evaluation criteria are defined before empirical validation [7]. It does not claim a new biometric algorithm, geofencing algorithm, or object detector, nor does it claim novelty relative to every existing government platform. A systematic review and direct comparison with deployed systems would be needed to support a broader novelty claim.

D. Aim and objectives

The aim is to specify a modular mobile-and-dashboard framework for traceable digital monitoring and inspection of activities associated with DoSJE schemes. Objectives are to: (i) define system layers and trust boundaries; (ii) model workflow states and evidence provenance; (iii) specify role-based permissions and data-retention controls; (iv) define safe integration points for location, biometric, and video-analysis signals; and (v) provide a controlled, reproducible evaluation protocol with explicit reporting requirements.

E. Research questions

RQ1: How can attendance, inspection assignment, evidence submission, and reviewer decisions be represented as a single traceable workflow? RQ2: Which controls are required to prevent location, biometric, and video-analysis outputs from being mistaken for definitive evidence? RQ3: Which functional, performance, usability, privacy, and security measurements are needed before a field pilot can support claims about operational effectiveness? These questions guide the proposed design; they are not presented as answered by experimental results.

F. Contributions and scope

The paper makes four bounded contributions: (i) a system-integration architecture; (ii) a workflow and data-provenance model for evidence capture, synchronization, review, and audit; (iii) a risk-oriented approach in which automated outputs are reviewable signals; and (iv) a controlled pilot protocol with defined metrics and reporting fields. The system is intended to assist authorized personnel. It does not independently determine beneficiary eligibility, establish misconduct, certify scheme compliance, or replace formal inspection and social-audit procedures.

II. RELATED WORK AND TECHNICAL CONTEXT

A. Government monitoring

DoSJE's I-MESA page describes ground-level examination of scheme implementation and lists monitoring, surveillance, social audit, and evaluation components [1]. The I-MESA guidelines are the primary policy reference used here [2]. The proposed application is framed as a potential technical support layer. Adoption would require the responsible department to confirm the applicable process, roles, evidence requirements, data fields, retention rules, and review authority.

B. Location and mobile evidence

Connected-device research illustrates the use of Internet of Things (IoT) sensing and machine-learning methods for monitoring physical systems. Mubaraali et al. describe an IoT-based induction-motor diagnosis

approach using vibration sensing and convolutional neural networks [8]. Nirmaladevi et al., including Mubaraali, combine sensing, signal processing, deep learning, and sensor-placement optimization for structural-damage localization [9]. These studies provide adjacent examples of monitoring pipelines, but their industrial and structural-health settings differ from public-sector inspection. They do not validate the proposed DoSJE workflow or its attendance and evidence-verification controls.

Mobile devices can combine forms, cameras, timestamps, and location services in one field interface. A geofence compares a reported location with a configured boundary and may return inside, outside, or uncertain.

Global navigation satellite system (GNSS) error, indoor reception, device settings, permission denial, mock-location tools, and device compromise can affect coordinates. Empirical tests of smartphone positioning show that error differs substantially between assisted Global Positioning System (GPS), Wi-Fi, and cellular methods, so a reported fix should be interpreted together with its source and accuracy [10]. Location is contextual evidence rather than proof of identity or conduct. A defensible implementation records the source and accuracy when available, applies a documented tolerance, and offers an exception path for poor reception or denied permission.

C. Biometric verification

ArcFace proposes an additive angular-margin loss for learning discriminative face representations [4]. It is a technical example, not evidence that another model or deployment will perform similarly. The Face Recognition Vendor Test of the National Institute of Standards and Technology (NIST) reports demographic differentials across evaluated algorithms and notes that image quality and algorithm choice affect errors [5].

NIST Digital Identity Guidelines discuss probabilistic biometric comparison, false-match and false-non-match risks, presentation attacks, and alternative authentication routes [11]. These are technical references, not Indian legal advice. Biometric verification is optional, must be locally validated, and must not be the sole basis for adverse administrative action.

D. Video analytics

You Only Look Once (YOLO) formulates object detection as prediction of bounding boxes and class probabilities from an image [6]. Performance reported for a model and benchmark does not establish fitness for a particular institutional camera, lighting condition, or monitoring policy. Camera angle, occlusion, compression, class definitions, and annotation quality affect results. Video analytics should create reviewable candidate events with model/version metadata and contextual evidence, not automatic findings of misconduct or compliance.

E. Design implications

The literature supports cautious use of technical components, not a presumption of end-to-end effectiveness. AI-assisted computational optimization offers a general methodological lesson: define the design space, test candidate configurations systematically, and validate predictions under varying conditions [3]. Sensor-placement research similarly illustrates that AI-assisted monitoring depends on sensing configuration and the conditions under which signals are collected [9]. These studies address different technical domains and do not evaluate this application. For the proposed framework, the transferable principle is to document assumptions, test across realistic operating conditions, and validate model outputs before operational use—not to import the solar absorber's performance claims. Integration should preserve independent verification signals, record policy and model versions, expose uncertain states, restrict evidence access, and require authorized human review for consequential decisions. The complete workflow must be evaluated in addition to individual algorithms.

III. SYSTEM ARCHITECTURE AND WORKFLOW

A. Five-layer architecture

Layer 1 — Presentation: Flutter mobile client(s) for field users and a government-facing dashboard. Clients render tasks and statuses but are not authoritative for permissions or final state transitions.

Layer 2 — Authenticated application programming interface (API) and orchestration: a FastAPI service authenticates requests, checks permissions, validates

payloads, applies workflow rules, and coordinates downstream services.

Layer 3 — Domain and validation services: attendance, geofence evaluation, inspection assignments, evidence metadata, exception handling, and optional biometric/video adapters. Each adapter returns a typed result with a reason code and version information.

Layer 4 — Persistence: a relational database (for example, PostgreSQL) stores users, assignments, events, metadata, and audit records; private object storage holds photographs and other large evidence files.

Layer 5 — Observability and governance: structured logs, metrics, health checks, access reviews, retention jobs, incident response, and audit review. Real-time subscriptions may update the dashboard, but every read and state-changing action must still pass server-side authorization.

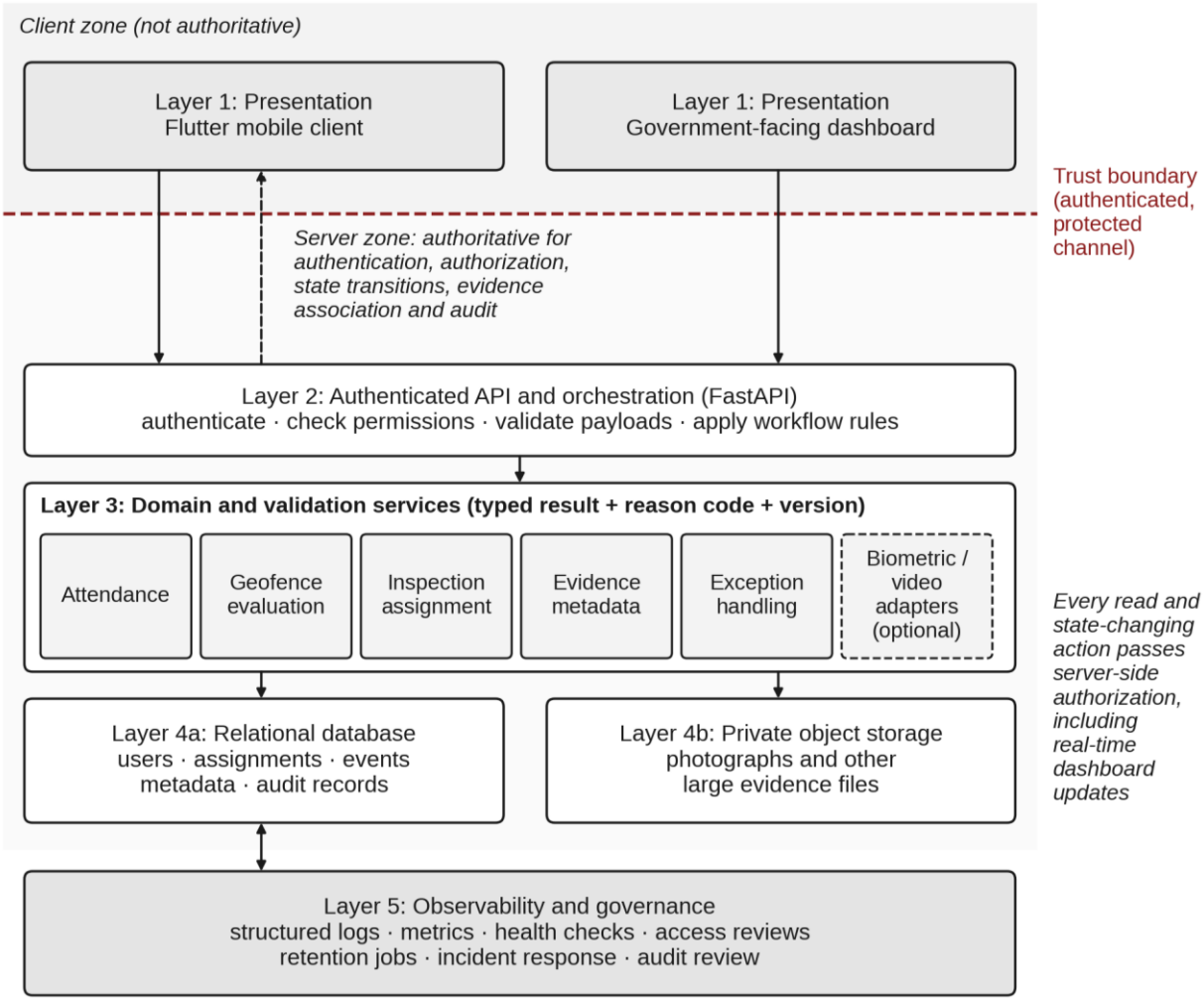

Fig. 1. Five-layer architecture of the proposed framework, showing the client–server trust boundary.

B. System flow

The principal flow is: field user → mobile app → authenticated API → validation/geofence/evidence services → database and private object storage → dashboard reviewer → audit trail. The backend is authoritative for authentication, authorization, accepted state transitions, evidence associations, and audit logging. A client-side message such as "verified" must never substitute for server-side validation.

C. Evidence-upload sequence

(1) The user opens an assigned inspection and captures evidence through the app, where permitted. (2) The client records available capture metadata and

associates the file with an assignment/inspection identifier. If offline, the file and metadata enter a protected local queue with a locally generated identifier and "pending synchronization" status. (3) When connectivity returns, the client requests an upload session or submits the file through an authenticated endpoint. Retries use an idempotency key to avoid duplicate records. (4) The server validates the authenticated user, assignment status, file size/type, required metadata, and upload policy. (5) The file is transferred over a protected channel to private object storage. The service computes or verifies a cryptographic checksum for integrity checking and stores the object key, checksum, server-received time, and metadata in the database. A checksum can detect later alteration; it cannot prove the original image is truthful. (6) The backend commits the record and returns a server-generated identifier and status. The client removes its pending marker only after confirmation. (7) The dashboard retrieves evidence through an authorized API or time-limited access mechanism. The reviewer records a decision and rationale; the backend appends an audit event. Failed uploads remain visibly pending or failed and can be retried without silently overwriting evidence.

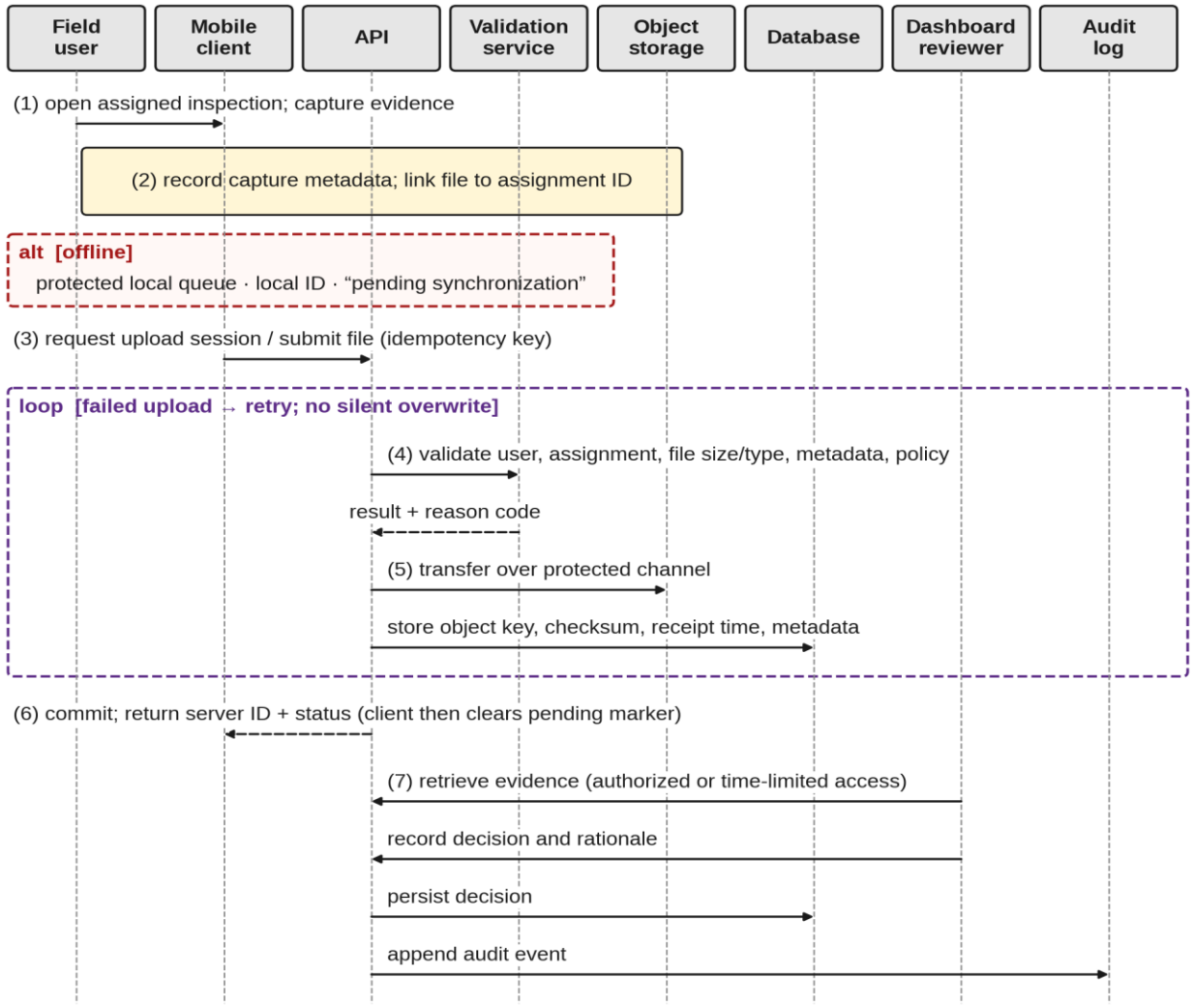

Fig. 2. Evidence-upload sequence, including the offline queue and retry loop.

D. Offline and synchronization semantics

Offline capture should be limited to explicitly approved data and protected by device-level controls. Device time and server receipt time must be stored separately. On synchronization, the server rechecks authorization and assignment state; it must not assume

that a record was valid merely because it was created offline. Conflict resolution should preserve prior records and log corrections. The protocol should define maximum offline duration, local encryption, logout behaviour, device-loss response, duplicate handling, and circumstances in which delayed submission requires manual review. These parameters must be agreed before a pilot.

E. Data model and figures

The core entities are users, institutions, roles, assignments, attendance events, inspections, evidence objects, alerts/exceptions, reviewer decisions, and audit events. Records should use stable identifiers and explicit foreign-key relationships. Evidence metadata should distinguish client capture time, server receipt time, source device/application version, checksum, storage object key, and review status. Fig. 1 shows the five-layer architecture and trust boundaries. Fig. 2 is a sequence diagram with field user, mobile client, API, validation service, object storage, database, dashboard reviewer, and audit log, including the offline queue and retry loop. Fig. 3 is an entity-relationship diagram for the core data model.

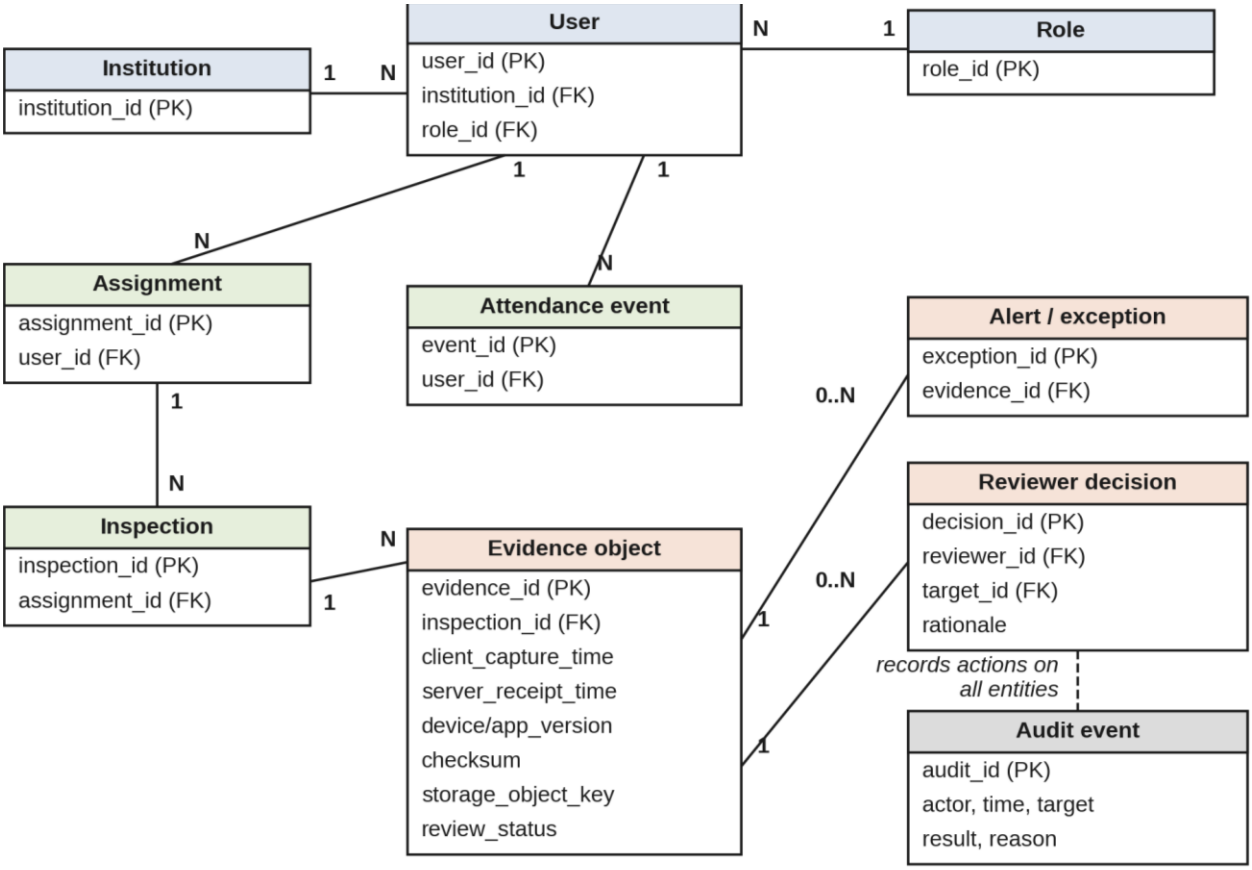

Fig. 3. Conceptual entity-relationship diagram of the core data model.

IV. IMPLEMENTATION MAPPING AND SYSTEM STATUS

A. Status interpretation

Table I is a provisional mapping based on the project description and available draft context, not a code audit. “Implemented” means an artifact is reported to exist; it does not mean that the module has passed systematic testing.

“Partially implemented” means a component or workflow is present but gaps remain or end-to-end behavior has not been established. “Planned” means

the capability is specified but no integrated implementation has been verified. "Not integrated" means the capability may exist separately or be under consideration but is not established as connected to the end-to-end workflow.

B. Implementation mapping

Table I maps the modules to artifacts that should be verified. Implementation status must not be confused with validated field effectiveness. A module can be coded and still fail functional, security, usability, or field validation. No status in Table I establishes deployment or operational effectiveness

TABLE I Provisional implementation mapping of system modules and the evidence needed to verify each status.

Module	Implementation artifact to verify	Provisional status	Evidence required
Mobile client	Flutter app source, build/run evidence	Partially implemented	Commit/build ID; end-to-end tests for stated workflows
Authentication	FastAPI authentication/JWT middleware and role checks	Partially implemented	Valid, invalid, expired, and unauthorized-session tests
Attendance workflow	Attendance API, mobile screen, event schema	Partially implemented	End-to-end success and failure-path test logs
Geofence validation	Geofence config, server-side rule, location payload	Partially implemented	Inside/outside/low-accuracy/permission-denied tests
Inspection assignment	Assignment allocation and status transitions	Partially implemented	Tests linking assignment to authorized user and inspection
Evidence uploads	Mobile capture/upload path and evidence metadata	Planned	Authenticated upload, retry, duplicate, invalid-file, checksum tests
Object storage	Private bucket/policy and object-key handling	Not integrated	Private-access, expiry, deletion and backup tests
Dashboard	Dashboard source and API integration	Partially implemented	Role-specific view and record-consistency tests
Exception queue	Exception states, reviewer UI, resolution endpoint	Planned	Creation, assignment, resolution, audit tests
Audit trail	Audit events for privileged/state-changing actions	Planned	Tamper-resistance and unauthorized-edit tests
Optional biometrics	Face/voice enrolment and matching adapters	Partially implemented	Version, threshold, test set, error metrics, fallback
Optional CCTV analytics	YOLO/video ingestion, event schema, dashboard integration	Planned	Approved footage, labels, model version, reviewer metrics

V. ACCESS CONTROL, DATA GOVERNANCE, AND RESPONSIBLE AI

A. Role-based access control

The backend should enforce least privilege on every request, not merely hide controls in the user interface. Role-based access control, which grants permissions to roles rather than to individual users, is a long-established model for administering such rules [12]. Roles should be scoped by institution, assignment, and jurisdiction where applicable. Read access to evidence should be separate from the ability to approve it; user administration should be separate from inspection review; and exports should require a documented purpose. High-impact actions should record actor, time, target, result, and reason. Table II is a proposed default and must be reconciled with departmental policy

TABLE II Proposed least-privilege role-based access-control matrix.

Permission	Field personnel	Institutional representative	Government reviewer	System administrator
Create attendance	Own events	Only if policy permits	No by default	No by default
Read attendance	Own records/status	Own institution, limited	Assigned jurisdiction	Only for support need
Update attendance	Correction request	Correction request	Review/annotate; no silent rewrite	Exceptional correction with audit
Delete attendance	No	No	No; request retention action	Controlled policy workflow
Upload evidence	Assigned inspections	Assigned response only	No by default	No by default
Read evidence	Own submitted evidence	Own institution/assignment	Assigned jurisdiction	Only when needed
Review evidence	No	No	Yes, with reason	No by default
Resolve exceptions	Respond/clarify	Respond/clarify	Yes, within assigned scope	Configuration/support only
Manage users	No	No	No by default	Yes, with audit
Configure geofences	No	No	Request/approve if policy allows	Yes, with approval/audit
Export data	Own records if enabled	Institution-scoped if approved	Scoped export with purpose	Restricted, logged export
View audit logs	Own-action receipt if appropriate	Own-action receipt if appropriate	Scoped read-only access	Administrative access; no silent alteration

B. Privacy and ethics

The application may process identifiers, attendance events, location metadata, photographs, biometric templates, closed-circuit television (CCTV)-derived events, and reviewer notes. Purpose limitation requires each field to be linked to a documented operational purpose. Data minimization means collecting only what is necessary and avoiding continuous location collection when an event-based check is sufficient. Users should receive clear notice and consent should be obtained where applicable; consent alone does not resolve every legal or public-sector governance question.

Access, correction, retention, and deletion procedures must be defined, along with incident notification and breach response. India's Digital Personal Data Protection Act, 2023 is a relevant legal source to review with the responsible institution and qualified counsel [13]; this manuscript does not provide legal advice or determine definitive applicability to a particular deployment. Service providers should have documented processing instructions, security obligations, breach cooperation, subcontractor controls, and secure deletion requirements.

C. Data retention

No retention duration is invented here. Table III identifies retention decisions that must be approved before a pilot. Periods should be set by the responsible authority using applicable law, departmental records policy, purpose, and risk. Deletion should be logged and include backups and vendor-held copies where technically and contractually applicable.

TABLE III Retention decisions to approve before a pilot; durations are intentionally not invented.

Data class	Purpose	Retention rule to approve	Controls
Attendance metadata	Attendance and dispute resolution	Approved records period; define start/end events	Role-scoped; corrections as new events; deletion logged

Data class	Purpose	Retention rule to approve	Controls
Location traces	Point-in-time geofence check	Prefer event-level location; delete raw traces when no longer needed	Limit precision/access; avoid unnecessary background tracking
Inspection photos	Evidence for a specific inspection	Through review/appeal and approved records period	Private storage; time-limited access; deletion of replicas where feasible
Biometric templates/samples	Optional identity verification	Raw samples not retained longer than necessary; explicit approval for templates	Strong restriction/encryption; alternative method; deletion verification
CCTV-derived events/clips	Authorized review of defined events	Approved event window; no blanket indefinite retention	Purpose-scoped access; model version; reviewer status; export logging
Alerts	Route records for review	Until resolution plus approved audit/dispute period	Status history; no automatic punitive action
Audit logs	Accountability and investigation	Documented period consistent with records/security policy	Restricted access; integrity monitoring; approved deletion only

D. Responsible AI and human review

Biometric matches and CCTV detections are probabilistic or context-dependent signals. They must be shown with status, source, model or policy version, and relevant uncertainty. A reviewer must be able to mark an event incorrect, request more information, or use a documented alternative process. Automated outputs must not alone trigger punitive action, benefit denial, or a formal finding of misconduct. Model updates require version control and regression testing. If biometric verification is enabled, evaluation must report false-match and false-non-match errors, enrollment failure, subgroup results where lawful and feasible, and availability of a non-biometric alternative. Voice-biometric use requires separate review of technical guidance and legal/ethical suitability; it should not be assumed acceptable merely because a library supports it.

VI. PILOT EVALUATION PROTOCOL

A. Study design and setting

The proposed evaluation is a controlled, staged software-and-usability pilot, not a field-impact trial. Stage A uses synthetic accounts, synthetic inspection records, and test media in a controlled test environment. Stage B, if separately approved, uses a small group of authorized staff and non-sensitive or appropriately governed test records in a supervised setting. No real beneficiary data or operational CCTV should be used without written authorization, a documented purpose, access controls, retention rules, and appropriate notices/consents. A practical initial duration is two weeks: setup and orientation, structured test sessions, defect correction, and repeat tests. This is a proposal, not a claim that a pilot has occurred.

B. Participants and roles

Recruit a documented group representing field personnel, institutional representatives, government reviewers, and a system administrator/test operator. Approximately 8–12 participants may be a feasible usability sample, but it is not sufficient to establish population-wide effectiveness or demographic fairness. Final sample size should be set before collection based on available staff, risk, and required precision. Participation should be voluntary where applicable, and results should be reported in aggregate unless individual reporting is necessary and authorized.

C. Tasks and test conditions

Participants execute scripted tasks: authenticate; open an assigned inspection; submit attendance; respond to permitted and denied location states; capture and upload evidence; interrupt connectivity during upload; capture offline and synchronize later; retry a failed upload; open an exception; review evidence; request clarification; resolve an exception; and inspect audit

history. Test cases cover normal connectivity, weak/intermittent connectivity, offline capture and later synchronization, denied location permission, low-accuracy location, duplicate retry, invalid or oversized file, expired/unauthorized session, and reviewer correction or exception handling. Biometric and CCTV tests are conditional on integration and approval. Fig. 4 shows protocol approval → version freeze → functional/resilience tests → optional biometric/CCTV evaluation → usability sessions → security assessment → remediation/retest → independent review → go/no-go decision

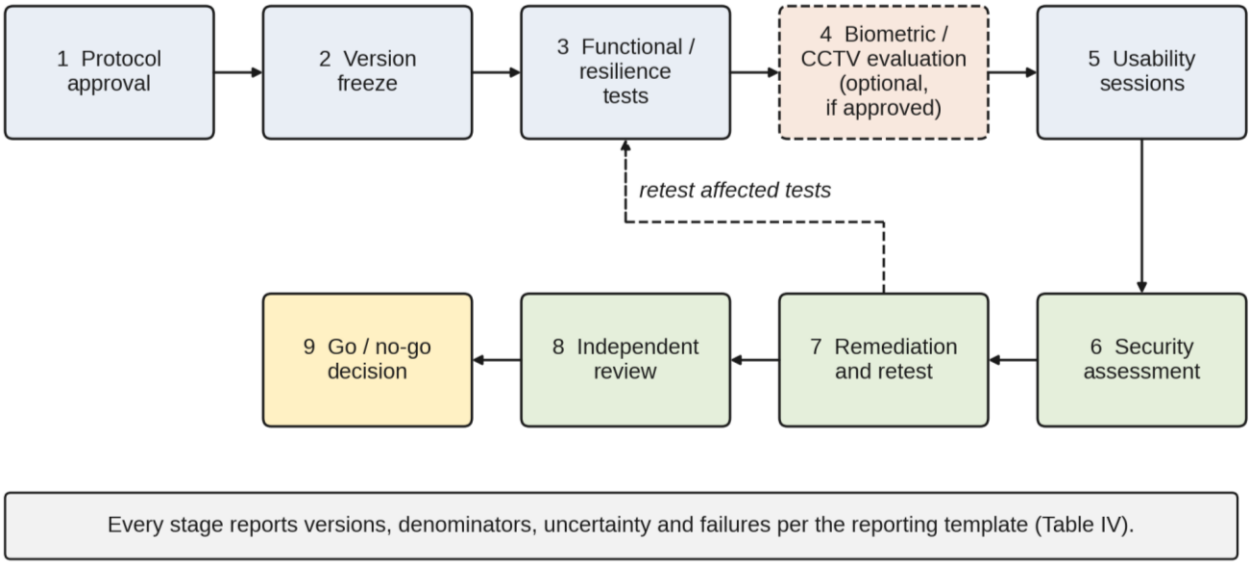

Fig. 4. Gated pilot-evaluation protocol from protocol approval to the go/no-go decision.

D. Data collection

Record test-case ID, expected and observed outcomes, client/server timestamps, device model and operating system (OS) version, app build, API version, database/storage configuration, network condition, request/response status, retry count, file size, checksum result, geofence reference and reported accuracy, exception reason, reviewer action, and defect ID. Avoid collecting real biometric samples or identifiable CCTV footage for convenience. If biometric testing is approved, use a documented consent and security protocol and suitable test set. For video analytics, document sample provenance, annotation rules, model version, confidence threshold, and adjudication procedure.

E. Metrics and success criteria

Functional correctness is the proportion of scripted cases producing the predeclared expected state; all authorization and data-integrity failures are release-blocking even if aggregate success is high. API latency should be reported as median, 95th percentile, and 99th percentile by endpoint and workload; acceptance targets must be declared before testing based on requirements, not selected after seeing results. Evidence-upload reliability is completed uploads divided by attempted uploads, with duplicates, checksum mismatches, retry recovery, and pending records reported separately. Geofence error should be measured against reference locations and predeclared inside/outside/uncertain ground truth; report false-inside and false-outside counts and rates and uncertainty due to sample size. If biometric verification is enabled, report false-match rate, false-non-match rate, failure-to-enrol rate, threshold, sample composition, and confidence intervals; do not report a single accuracy figure alone. If CCTV analytics is enabled, report precision, recall, F1 where appropriate [14], false alerts per camera-hour, missed events, and processing delay against independently labelled data.

Usability should include task completion, time-on-task, assistance requests, user-reported difficulty using a specified instrument (for example, the System Usability Scale [15]), and missing responses. Security testing should report findings by severity, reproducible evidence, affected component, remediation status, and residual risk.

F. Analysis and reproducibility

Pre-register or timestamp the test plan, hypotheses, acceptance criteria, test-case definitions, exclusions, and analysis approach before evaluation. Report numerator and denominator for every rate, uncertainty intervals where appropriate, software/hardware versions, environment, workload, network profile, and all material deviations. Include failure cases and negative results. If the sample is too small for stable inference, report descriptive statistics and avoid claims of statistical significance or generalizability. Separate software correctness from operational effectiveness: a controlled test can establish whether the workflow behaves as specified, but cannot by itself establish more accurate, timely, fair, or beneficial inspections.

G. Reporting template

Table IV specifies the minimum outputs expected from the pilot. Complete the rows only after the corresponding tests have actually been run.

TABLE IV

Evaluation area	Required outputs
Functional correctness	Passed/total cases; failure count/severity; failed case IDs; build and test-suite version.
API latency	Endpoint/workload; request count; median, p95, p99; throughput; error rate; hardware/network.
Evidence uploads	Attempted/completed uploads; denominator; retry recovery; duplicates; checksum mismatches; pending/failed uploads; file sizes.
Geofence behavior	Ground truth; inside/outside/uncertain counts; false-inside/outside rates; sample size; accuracy distribution; intervals where appropriate.
Biometrics (if enabled)	Model/version; threshold; sample; FMR/FNMR; failure-to-enroll; subgroup analysis where lawful/feasible; uncertainty intervals; fallback usage.
CCTV (if enabled)	Model/version; annotated sample; precision/recall/F1 as appropriate; false alerts per camera-hour; missed events; delay; reviewer adjudication.
Usability	Participant roles and n; task completion/total; time distribution; assistance requests; instrument/scoring; missing responses.
Security	Scope/threat model; findings by severity; reproducible evidence; affected component; remediation state; residual risk.
Privacy/ operations	Data classes; retention policy; access-review findings; deletion test; incident process; protocol deviations.
Reproducibility	App/API/model versions; device/OS; DB/storage; scripts; network conditions; exclusions; denominators; failures.

VII. RESULTS AND DISCUSSION

A. Current evidentiary status

This manuscript reports no completed pilot, measured performance, deployment, user-feedback, or field-impact findings. There are therefore no empirical results to interpret. This section describes the evidence required before claims can be made. The architecture and status table are design artifacts and provisional project mapping, not validation results.

B. Trade-offs

Centralization can make records easier to retrieve and compare, but concentrates risk: a misconfigured storage bucket or compromised privileged account may expose multiple institutions' records. Real-time visibility can shorten the time to notice missing information, but aggressive alerts can increase reviewer workload and create alert fatigue. More verification signals may reduce some errors while increasing privacy intrusion, user friction, and failure rates under poor device or network conditions. Offline support improves continuity but complicates timestamp interpretation, conflict resolution, device-loss response, and duplicate prevention. These trade-offs should be measured rather than assuming more data or automation is always better.

C. Operational and procedural fairness

A system that labels a location or biometric check "failed" without context may disadvantage users with

poor connectivity, disabilities, device limitations, or difficult capture conditions. A reviewable "uncertain" state and documented exception path can preserve accountability while avoiding automatic adverse conclusions. Reviewers need sufficient context to correct errors, but access to sensitive information should remain restricted. The operational burden of reviewing false positives must be measured alongside technical performance.

VIII. LIMITATIONS

First, this is a design-oriented proposal and does not demonstrate implementation completeness or field effectiveness. Second, the provisional implementation mapping requires direct verification against the current repository, configuration, and test evidence. Third, a small controlled sample may not generalize across institutions, locations, devices, languages, network conditions, camera installations, or staff populations. Fourth, operational workflows depend on departmental procedures, legal review, records policy, and authorization not established by this paper. Fifth, location and biometric performance vary with hardware, environment, thresholds, and adversarial conditions.

Sixth, optional CCTV analytics may be infeasible where camera access, image quality, legal authority, or ground-truth labels are unavailable. Seventh, a short usability pilot cannot establish improved inspection quality or beneficiary outcomes. Finally, independent review is needed to reduce confirmation bias and prevent acceptance criteria being retrofitted to observed results.

IX. CONCLUSION

This paper specifies a traceable, human-in-the-loop framework for supporting monitoring and inspection workflows associated with DoSJE schemes. Its principal design elements are a five-layer architecture, server-authoritative workflow control, provenance-aware evidence upload, least-privilege permissions, explicit exception handling, retention governance, and cautious integration of biometric and video-analysis signals. The paper makes no deployment or performance claim. Its immediate value is to define what must be implemented, governed, and measured before the system can be evaluated credibly.

A. Actionable future work

Future work should proceed in a gated sequence: (1) conduct a code/configuration audit and replace provisional module statuses with evidence-backed labels; (2) finalize data schemas, access rules, retention durations, offline synchronization semantics, and audit-event definitions; (3) complete unit, integration, permission, upload-integrity, and recovery tests; (4) conduct privacy and security review before sensitive data use; (5) run the controlled pilot using predeclared test cases and acceptance criteria; (6) if results and approvals justify it, design a separately reviewed field study comparing the workflow with existing processes; and (7) publish versions, denominators, uncertainty intervals, failures, and limitations for reproducibility.

DECLARATION

A generative AI tool was used solely for grammar checking of the manuscript text. The authors reviewed all content and take full responsibility for the manuscript.

REFERENCES

[1] Department of Social Justice and Empowerment, Government of India, "Information-Monitoring, Evaluation & Social Audit (I-MESA)," official scheme page. [Online]. Available: I-MESA Official Scheme Page (accessed Oct. 10, 2026).

[2] Department of Social Justice and Empowerment, Government of India, *Information-Monitoring, Evaluation & Social Audit (I-MESA): Scheme Guidelines*, effective Oct. 4, 2023. [Online]. Available: I-MESA Official Scheme Page (accessed Oct. 10, 2026).

[3] R. Rasappan Senthil Kumar, U. Arun Kumar, A. Flah, and H. Kraiem, "Computational design and AI optimization of a symmetry-engineered ultra-broadband metasurface solar absorber for renewable energy systems," *Energy Rep.*, vol. 16, Art. no. 109524, 2026, doi: 10.1016/j.egyr.2026.109524.

[4] J. Deng, J. Guo, J. Yang, N. Xue, I. Kotsia, and S. Zafeiriou, "ArcFace: Additive angular margin loss for deep face recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 44, no. 10, pp.

5962–5979, 2022, doi: 10.1109/TPAMI.2021.3087709.

[5] P. Grother, M. Ngan, and K. Hanaoka, "Face recognition vendor test (FRVT) part 3: Demographic effects," Nat. Inst. Standards Technol., Gaithersburg, MD, USA, NIST Interagency/Internal Rep. 8280, 2019, doi: 10.6028/NIST.IR.8280.

[6] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 2016, pp. 779–788, doi: 10.1109/CVPR.2016.91.

[7] A. R. Hevner, S. T. March, J. Park, and S. Ram, "Design science in information systems research," *MIS Q.*, vol. 28, no. 1, pp. 75–105, 2004, doi: 10.2307/25148625.

[8] L. Mubaraali, N. Kuppuswamy, and R. Muthukumar, "Internet of Things-based induction motor diagnosis using convolutional neural network," *Elect. Power Compon. Syst.*, vol. 52, no. 2, pp. 265–276, 2024, doi: 10.1080/15325008.2023.2217434.

[9] J. Nirmaladevi, L. Mubaraali, A. H. Bagem, and A. A. Priya, "Deep learning-guided optimization of piezoelectric sensor placement for high-fidelity mechanical damage localization," *Mech. Adv. Mater. Struct.*, vol. 33, no. 1, Art. no. 2716725, 2026, doi: 10.1080/15376494.2026.2716725.

[10] P. A. Zandbergen, "Accuracy of iPhone locations: A comparison of assisted GPS, WiFi and cellular positioning," *Trans. GIS*, vol. 13, no. s1, pp. 5–26, 2009, doi: 10.1111/j.1467-9671.2009.01152.x.

[11] National Institute of Standards and Technology, "Digital identity guidelines: Authentication and authenticator management," NIST Special Publication 800-63B, Rev. 4. [Online]. Available: NIST SP 800-63B (accessed Oct. 10, 2026).

[12] R. S. Sandhu, E. J. Coyne, H. L. Feinstein, and C. E. Youman, "Role-based access control models," *Computer*, vol. 29, no. 2, pp. 38–47, 1996, doi: 10.1109/2.485845.

[13] Ministry of Electronics and Information Technology, Government of India, *Digital Personal Data Protection Act, 2023*, 2023. [Online]. Available: India Code – Digital Personal Data Protection Act, 2023 (accessed Oct. 10, 2026).

[14] M. Sokolova and G. Lapalme, "A systematic analysis of performance measures for classification tasks," *Inf. Process. Manage.*, vol. 45, no. 4, pp. 427–437, 2009, doi: 10.1016/j.ipm.2009.03.002.

[15] J. Brooke, "SUS: A 'quick and dirty' usability scale," in *Usability Evaluation in Industry*, P. W. Jordan, B. Thomas, B. A. Weerdmeester, and I. L. McClelland, Eds. London, U.K.: Taylor & Francis, 1996, pp. 189–194.