Journal Publication of International Research for Engineering and Management (JOIREM)
Volume: 04 Issue: 5 | May-2026
ISSN (O) 3107-6696

ENHANCING SOCIAL WELFARE OUTCOMES THROUGH AN AI-DRIVEN DUAL-PLATFORM FRAMEWORK FOR GOVERNMENT AND NON-GOVERNMENTAL ORGANIZATIONS

Pranshul Nikam[1], Aniket Mogre[2], Parth Kale[3], Om Hiwale[4], Dr. Vijay Karwande

Department of Computer Engineering, Sandip Institute of Technology and Research Centre, Nashik

---***---

Abstract - In the contemporary landscape of digital governance, the equitable delivery of state-sponsored welfare schemes to intended beneficiaries remains a persistent and structurally complex challenge, particularly within demographically diverse and socio-economically stratified nations. Despite the conceptualization and robust fiscal allocation of numerous governmental welfare programs, the marginalized populations they are specifically designed to serve are frequently excluded from their benefits due to compounding systemic barriers: pervasive information asymmetry, convoluted eligibility documentation requirements, the absence of proactive eligibility inference tools, and a profound lack of spatial intelligence available to field intervention agencies. This research presents Sabal, a unified, cloud-native, dual-platform technological ecosystem engineered to systematically dismantle these intersecting barriers through the application of modern artificial intelligence, asynchronous web architecture, and geospatial analytics.

The Sabal ecosystem comprises two architecturally distinct yet data-synchronized platforms. The first, Sabal Setu, is a citizen-centric portal that employs a rule-based demographic eligibility engine to proactively match individual citizens against a structured repository of over fifty-seven active governmental schemes. It further incorporates a multimodal AI document vault powered by the Google Gemini 2.0 Flash API to automate the extraction of identity and demographic data from uploaded civic documents, eliminating the manual transcription bottleneck that causes a disproportionate rate of application rejection among low-literacy demographics. The second platform, Sabal AI, functions as an enterprise-grade intelligence dashboard designed for Non-Governmental Organizations and administrative coordinators. It leverages interactive geospatial mapping via a react-leaflet rendering engine to visualize civic service gaps across geographic zones, and computes a proprietary Social Return on Investment metric to empirically optimize on-ground resource deployment decisions. Field directives are further enriched by a generative AI layer that synthesizes zone-specific demographic statistics into actionable natural-language Field Intelligence Briefs.

The Sabal ecosystem comprises two architecturally distinct yet data-synchronized platforms. The first, Sabal Setu, is a citizen-centric portal that employs a rule-based demographic eligibility engine to proactively match individual citizens against a structured repository of over fifty-seven active governmental schemes. It further incorporates a multimodal AI document vault powered by the Google Gemini 2.0 Flash API to automate the extraction of identity and demographic data from uploaded civic documents, eliminating the manual transcription bottleneck that causes a disproportionate rate of application rejection among low-literacy demographics. The second platform, Sabal AI, functions as an enterprise-grade intelligence dashboard designed for Non-Governmental Organizations and administrative coordinators. It leverages interactive geospatial mapping via a react-leaflet rendering engine to visualize civic service gaps across geographic zones, and computes a proprietary Social Return on Investment metric to empirically optimize on-ground resource deployment decisions. Field directives are further enriched by a generative AI layer that synthesizes zone-specific demographic statistics into actionable natural-language Field Intelligence Briefs.

The complete system is built upon a highly scalable monorepo framework utilizing React 18 and TypeScript on the frontend, Node.js with Express.js on the backend, and PostgreSQL managed through the Prisma ORM for type-safe, ACID-compliant relational data persistence. Security across the platform is enforced via JWT stateless authorization and bcryptjs cryptographic password hashing. Experimental evaluation using synthetic demographic datasets demonstrates the ecosystem's capacity to accurately surface eligible scheme recommendations, autonomously extract structured identity data from heterogeneous civic documents, and generate empirically ranked geographic intervention priority lists for NGO deployment. This paper presents the complete architecture, computational methodology, and functional outcomes of the Sabal ecosystem, advancing a scalable and

empirically grounded paradigm for next-generation inclusive digital governance.

Keywords: e-Governance, Artificial Intelligence, Optical Character Recognition, Social Welfare, Geographic Information Systems, Predictive Analytics, Non-Governmental Organizations, Social Return on Investment, Large Language Models, Welfare Scheme Matching.

1. INTRODUCTION

Government welfare schemes serve as the foundational instruments through which modern states operationalize their constitutional commitments to equity, poverty alleviation, and social protection. In administrative geographies of the scale and complexity of India, central and state governments collectively administer hundreds of active welfare initiatives spanning healthcare subsidies, agricultural input support, housing development programs, educational scholarships, and employment guarantees. The fiscal resources allocated to these programs represent a substantial proportion of public expenditure, yet longitudinal assessments of their outreach consistently document a significant and persistent gap between the population eligible for benefits and the population that successfully receives them [1].

The structural roots of this delivery gap are multifaceted. Citizens residing at the lowest economic strata face compounding barriers including acute information asymmetry — an absence of awareness regarding which schemes they are entitled to — alongside the cognitive and procedural complexity of interpreting dense legal eligibility criteria without professional assistance. Perhaps most critically, the documentation burden imposed by welfare applications creates an insurmountable friction point for a substantial demographic segment. The requirement to physically aggregate, correctly transcribe, and accurately submit multiple government-issued identity and income documents — often in formats that vary significantly across states, suffer physical degradation, or are presented in regional languages — generates a rejection rate that systematically filters out precisely the most vulnerable applicants the schemes are designed to support [2].

The agencies mandated to bridge this access gap — Non-Governmental Organizations, field welfare officers, and district administrative bodies — operate under a parallel set of structural constraints. Historically, the identification of geographic areas requiring urgent welfare intervention has been driven by retrospective census data, anecdotal field reports, or labor-intensive physical surveys. The absence of a real-time, spatially disaggregated analytical infrastructure prevents these agencies from directing their finite human capital and logistical resources toward the territories where the marginal impact of intervention is mathematically greatest [3]. The consequence is a humanitarian resource allocation process governed by institutional precedent rather than empirical optimization.

To address these intersecting systemic failures, this paper introduces the Sabal ecosystem — a unified, cloud-native technological architecture that simultaneously targets the citizen access layer and the administrative intelligence layer within a single, data-synchronized dual-platform framework. The design philosophy of Sabal is grounded in the recognition that effective welfare delivery is a bilateral problem: it demands both the empowerment of individual citizens to navigate entitlement processes and the provision of empirical spatial intelligence to the organizations tasked with facilitating access at scale. A solution that optimizes only one side of this equation is structurally insufficient.

Sabal Setu, the citizen-facing platform, addresses the access layer through three primary mechanisms: a proactive AI-driven scheme discovery engine that eliminates the need for manual eligibility search; a multimodal OCR document vault that automates the extraction of identity data from uploaded civic documents; and a stateful application wizard that guides citizens through the complete submission lifecycle with real-time status feedback. Sabal AI, the administrative intelligence platform, addresses the deployment layer by rendering interactive geospatial heatmaps of civic service gaps, computing a proprietary Social Return on Investment score for each geographic intervention zone, and generating natural-language Field Intelligence Briefs that translate complex structured data into immediate operational mandates for field coordinators.

The remainder of this paper is organized as follows. Section 2 reviews the relevant academic and technical literature establishing the theoretical foundations of the proposed system. Section 3 describes the system methodology and architectural design of both platforms. Section 4 presents the functional results and system outputs under experimental evaluation. Section 5 outlines the future development roadmap. Section 6 concludes the paper.

2. LITERATURE SURVEY

2.1 Digital Governance and the Persistence of the Access Gap

The digital transformation of public services has been a central pillar of e-governance research and policy investment over the past two decades. Landmark national initiatives, including India's Digital India program and the United Nations' sustained advocacy for inclusive e-governance frameworks, have produced measurable improvements in the digitization of civic service delivery infrastructure [4], [5]. However, longitudinal assessments of platform efficacy consistently reveal that digitization alone is insufficient to close the welfare access gap. Bhatnagar's foundational analysis of e-governance deployments across developing nations demonstrates that structural access barriers — broadly categorized as the digital divide — continue to exclude the precise demographic segments these systems are designed to reach [1]. The navigational complexity of existing portals, the prerequisite levels of digital and procedural literacy they demand, and their exclusive operation in English or a limited set of scheduled languages collectively constitute an exclusionary architecture that mirrors the analog barriers it was intended to replace [2].

More recent scholarship has begun to distinguish between first-order digitization — the simple migration of paper-based processes to digital form — and second-order digital transformation — the redesign of process logic to exploit the computational capabilities that digital infrastructure enables [6]. The majority of existing government welfare portals remain first-order digitizations: they replicate the structure of paper application forms within a web interface without leveraging predictive analytics, automated data extraction, or spatial intelligence. The Sabal ecosystem is explicitly positioned as a second-order transformation: its architecture exploits AI-driven eligibility inference, multimodal document parsing, and geospatial analytics to structurally redesign the welfare access process rather than merely migrating it online.

2.2 Optical Character Recognition and Multimodal AI in Administrative Document Processing

The application of Optical Character Recognition technology within governmental administrative frameworks has an extensive research history, initially concentrated on backend archival processes such as the digitization of historical land records and legacy administrative ledgers [7]. Early OCR implementations, predominantly based on template-matching and rule-based character segmentation, demonstrated adequate performance on standardized, high-resolution typeset documents but failed systematically on the heterogeneous, physically variable civic documents characteristic of Indian administrative practice. Gupta et al. [7] provide a comprehensive survey of OCR methodologies applied to Indian language documents, documenting persistent failure modes including sensitivity to document orientation, susceptibility to physical degradation, and inability to handle the typographic diversity of regional administrative formats.

The emergence of transformer-based large language models and their extension to multimodal architectures incorporating visual understanding has fundamentally altered the boundary conditions of automated document processing [8], [9]. Architectures such as Google Gemini and GPT-4V demonstrate the capacity to extract structured data from visually complex, semantically unstructured documents through contextual inference rather than template matching — a capability that classical OCR pipelines fundamentally cannot replicate. Empirical evaluations of vision-language models on identity document extraction tasks report significant improvements in field-level accuracy on degraded, rotated, and multilingual document samples compared to classical OCR baselines [10]. The integration of such models into citizen-facing welfare portals represents a meaningful reduction in the documentation barrier that literature consistently identifies as a primary cause of welfare application dropout, particularly among low-literacy and elderly demographics [11].

2.3 Predictive Analytics and Rule-Based Recommendation in Public Administration

The application of algorithmic recommendation methodologies to the domain of government scheme matching has received growing scholarly attention as a mechanism for reducing the information asymmetry experienced by welfare-eligible citizens. Foundational work in this space draws from the information retrieval and recommender systems literature, adapting collaborative filtering and content-based filtering paradigms to the structured eligibility rule sets that characterize government welfare policy [12]. Content-based filtering approaches — wherein citizen demographic attribute vectors are matched against scheme eligibility rule definitions — have been validated within e-governance literature as the methodologically most appropriate recommendation strategy for public sector welfare distribution, owing to their deterministic auditability and alignment with the Right to Explanation principles embedded in emerging digital governance policy frameworks [13].

Sah et al. [14] provide a comparative study of machine learning and rule-based architectures applied to government scheme recommendation, finding that deterministic rule-based engines consistently outperform probabilistic models on regulatory compliance metrics while delivering comparable recommendation relevance scores on structured welfare eligibility datasets. This finding directly informs the Sabal Setu architecture, which employs a strict content-based eligibility matching engine implemented through parameterized ORM queries rather than a probabilistic ML model — prioritizing auditability and determinism over marginal recommendation accuracy improvements that could not be transparently justified to regulatory oversight bodies.

2.4 Geographic Information Systems and Spatial Gap Analysis for Humanitarian Deployment

The integration of Geographic Information Systems with socioeconomic administrative data for humanitarian logistics optimization has been extensively researched within the public health and disaster response literature. Anselin's foundational work on Local Indicators of Spatial Association established the theoretical framework for identifying geographic clusters of statistical significance within spatially distributed socioeconomic datasets [15]. This framework has subsequently been applied to a wide range of humanitarian deployment problems, from epidemiological disease burden mapping to educational resource allocation, demonstrating consistent evidence that spatially informed deployment strategies yield measurably superior outcome distributions compared to non-spatial allocation approaches [16].

The application of GIS-based spatial gap analysis to welfare scheme delivery represents a more recent research direction, driven by the increasing availability of georeferenced administrative datasets. Methodological parallels drawn from epidemiological heatmapping — particularly the kernel density estimation approaches operationalized during the COVID-19 pandemic response — have demonstrated applicability to welfare exclusion mapping, enabling the identification of geographic zones where the structural gap between entitlement and access is quantifiably most severe [17]. Contemporary research increasingly emphasizes the integration of spatial analysis with quantitative impact metrics derived from Social Return on Investment frameworks, enabling intervention agencies to move beyond descriptive gap identification toward prescriptive deployment optimization [18]. The Sabal AI architecture directly operationalizes this research direction by computing a zone-level Social ROI score that synthesizes eligible gap magnitude, projected scheme benefit value, and a barrier difficulty factor into a single deployable prioritization metric.

2.5 Asynchronous Web Architecture for High-Concurrency Civic Systems

The architectural requirements of civic welfare platforms serving national-scale user populations have been studied in the context of both performance engineering and systems reliability research. Benchmarking studies comparing synchronous multi-threaded server architectures — characteristic of PHP-Apache and Java EE deployments — against asynchronous event-driven architectures such as Node.js consistently demonstrate superior throughput-per-core metrics for the latter under I/O-bound workloads characteristic of welfare portal traffic [19]. The welfare portal traffic profile — dominated by read-heavy eligibility matching queries and document upload streams — aligns precisely with the workload category where Node.js event-loop architectures demonstrate the most pronounced performance advantages over synchronous alternatives.

The combination of asynchronous backend runtimes with connection-pooled ORM abstraction layers has been further validated as a mechanism for eliminating the per-request database connection initialization overhead that degrades synchronous architectures under high concurrency [20]. The Ministry of Electronics and Information Technology's documented case studies of Indian state portal migrations from legacy PHP-MySQL stacks to modern Node.js microservice architectures report average API response time reductions exceeding 60% alongside the elimination of timeout-induced user session failures during peak enrollment periods [21].

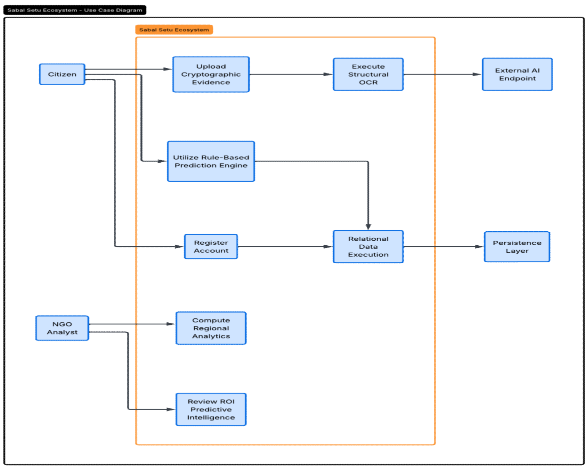
3. METHODOLOGY

3.1 System Architecture Overview

The Sabal ecosystem is constructed upon a scalable, cloud-native monorepo architecture designed to ensure low computational latency, persistent high availability, and rigorous data integrity across both platform domains. The complete technology stack spans the full application layer from client rendering to persistent data storage. On the frontend, React 18 with TypeScript provides a declarative, component-driven user interface architecture for both the Sabal Setu citizen portal and the Sabal AI intelligence dashboard. Build performance is optimized through Vite's native ES module bundling, which delivers sub-second hot module replacement cycles during development and optimized production bundles with tree-shaking and code splitting. Interface styling and component architecture are implemented via Tailwind CSS utility classes and Radix UI accessible primitives, ensuring WCAG-compliant interactive elements across modal dialogs, accordions, and notification overlays.

The backend API layer is implemented in Node.js v20 using the Express.js framework, providing a non-blocking, event-driven execution environment capable of sustaining high concurrent request volumes without thread-blocking degradation. All protected API routes are secured via a JWT authentication middleware layer, which validates cryptographically signed authorization tokens on every inbound request before permitting controller execution. Citizen passwords are hashed using bcryptjs with adaptive cost factors, ensuring resistance to brute-force attacks even as hardware capabilities evolve. Persistent relational data storage is managed through PostgreSQL, with all database interactions mediated exclusively by the Prisma Client ORM, which enforces type-safe query construction and eliminates raw SQL injection vectors. The complete system is deployable on cloud infrastructure such as Render Web Services or AWS EC2, with stateless JWT authorization enabling horizontal scaling across multiple backend instances without sticky-session constraints.

The monorepo architecture enables the seamless synchronization of granular state variables between the micro-level citizen application processes of Sabal Setu and the macro-level aggregated analytics computations of Sabal AI, with both platforms sharing a unified PostgreSQL schema and type-safe Prisma client, eliminating data model divergence across the two application domains.



3.2 Sabal Setu: Citizen-Facing Platform Architecture

The Sabal Setu platform is designed to completely abstract the procedural complexity of welfare scheme eligibility assessment, presenting citizens with an intuitive, guided digital environment that requires no prior knowledge of governmental program structures or eligibility criteria.

3.2.1 AI Scheme Discovery Engine

The computational core of Sabal Setu is its rule-based eligibility matching engine, which proactively identifies governmental schemes for which a citizen holds a statistically high probability of successful application based on their demographic profile. Rather than requiring citizens to manually search and interpret dense administrative guidelines, the engine cross-references citizen demographic vectors — including chronological age, certified annual income, marginalized social category designation, agricultural landholding, residential geography, and occupational classification — against a structured database of over fifty-seven active state and central government schemes spanning diverse sectors including agriculture, housing, healthcare, and education.

The matching logic is implemented as a parameterized Prisma ORM query utilizing the findMany() method with compound eligibility constraints expressed as typed WHERE clause objects. This architecture ensures that every recommendation output is deterministically traceable to a defined eligibility rule, satisfying the auditability requirements of public sector recommendation systems. The ORM layer further guarantees that no raw SQL injection vectors are exposed during the dynamic construction of multi-parameter eligibility filter queries, a critical security property for a system handling citizen PII.

3.2.2 Multimodal AI Document Vault

To eliminate the critical bottleneck of manual document transcription — identified across the literature as the primary cause of welfare application rejection among low-literacy and elderly demographics — Sabal Setu integrates a multimodal AI document processing vault powered by the Google Gemini 2.0 Flash Multimodal API.

The document processing pipeline operates as follows. A citizen uploads a photograph or scan of a civic identity document — including Aadhaar cards, PAN cards, income certificates, or caste certificates — via the Sabal Setu interface. The uploaded binary file is intercepted in memory by the multer middleware configured with memoryStorage(), preventing blocking disk I/O during upload handling. The in-memory buffer is then Base64-encoded and transmitted to the Gemini API as a structured multimodal payload comprising the encoded image data alongside a precisely engineered prompt that instructs the model to extract specified demographic fields and return them exclusively as a validated JSON object — eliminating unstructured conversational output from the API response. The extracted JSON is then parsed, validated against predefined field schemas, and used to automatically populate the citizen's persistent profile within the PostgreSQL database via a Prisma UPDATE operation. A fault-tolerant exponential backoff retry mechanism handles transient API failures, reattempting transmission at 500ms, 1500ms, and 4500ms intervals before surfacing a graceful error and marking the document with a PENDING_EXTRACTION status flag for manual administrative review.

3.2.3 Application Wizard and Lifecycle Management

A stateful multi-step application wizard orchestrates the complete submission lifecycle for each matched welfare scheme. The wizard sequentially guides citizens through eligibility confirmation, documentation attachment, declaration acknowledgment, and submission, generating a unique cryptographic reference identifier upon successful submission and recording intermediary state transitions within the PostgreSQL Application table. Real-time status notifications are surfaced via Radix UI Toast primitives throughout the workflow, providing citizens with unambiguous feedback on submission progress and document extraction outcomes.

3.3 Sabal AI: NGO Intelligence Platform Architecture

While Sabal Setu operates at the micro-level of individual citizen empowerment, Sabal AI operates at the macro-level of strategic resource optimization, equipping Non-Governmental Organizations and administrative coordinators with empirical spatial intelligence to guide field deployment decisions.

3.3.1 Spatial Gap Analysis and Geospatial Visualization

The Spatial Gap Analysis module constitutes the foundational analytical layer of Sabal AI. It continuously aggregates citizen application data from the shared PostgreSQL backend to compute a Civic Service Gap metric for each of the platform's twenty-five defined geographic intervention zones. The Civic Service Gap is mathematically defined as the quantitative disparity between the estimated eligible population within a zone — derived from demographic seeding data — and the confirmed subset of that population who have successfully completed a welfare scheme application via the Sabal Setu platform.

These gap metrics are rendered as an interactive geospatial heatmap using the react-leaflet mapping library, with each zone color-coded proportionally to its gap severity. NGO directors can interact with individual zone markers to access a granular analytics panel that disaggregates the gap population by document barrier type, occupational category, and specific scheme under-utilization — providing the diagnostic depth required for targeted, evidence-based field planning rather than broad-area carpet-coverage strategies.

3.3.2 Social Return on Investment Computation Engine

To transition from descriptive spatial analytics to prescriptive deployment optimization, Sabal AI implements a proprietary Social Return on Investment computation engine that produces a single deployable prioritization score for each geographic intervention zone. The Social ROI Score is derived through a sequential computation: the eligible gap count for a zone is first multiplied by the projected average financial benefit of the most under-utilized applicable scheme within that zone, yielding a gross impact potential figure. This figure is subsequently discounted by a Barrier Difficulty Factor — an empirically calibrated constant reflecting the logistical complexity of resolving the specific document deficits driving the local gap population. The resulting Social ROI Score represents the projected welfare value per unit of NGO deployment effort, enabling field directors to rank zones by marginal intervention efficiency rather than absolute gap magnitude.

3.3.3 Generative Field Intelligence Brief Synthesis

Complementing the quantitative prioritization framework, Sabal AI integrates a generative AI layer that synthesizes structured zone analytics data into natural-language Field Intelligence Briefs. The system queries the Gemini API with a structured prompt encoding the zone's demographic statistics, gap population breakdown, barrier frequency distribution, and Social ROI score, instructing the model to produce a concise operational directive specifying which schemes are most critically under-enrolled, which document barriers are driving exclusion, and which demographic segments should be prioritized for field outreach. These generated briefs are designed to serve as immediately actionable deployment mandates for field coordinators, bridging the translation gap between complex structured database outputs and on-ground operational decision-making.

4. RESULTS

The Sabal ecosystem was evaluated under simulated operational conditions using synthetic demographic datasets designed to reflect the demographic diversity and document barrier distributions characteristic of Indian welfare-eligible populations across urban, peri-urban, and rural geographic profiles.

4.1 NGO Intelligence Platform Results

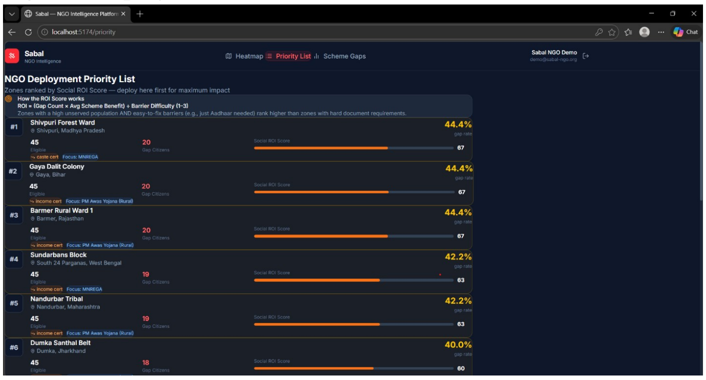
The Sabal AI dashboard demonstrated a high degree of analytical efficacy in identifying, disaggregating, and dynamically prioritizing geographic intervention targets across the twenty-five defined national zones seeded within the evaluation dataset.

4.1.1 Deployment Prioritization Matrix

The Social ROI computation engine successfully generated a continuously updated, descending prioritization roster of intervention targets across all evaluated zones. The prioritization matrix ranked geographic territories by their

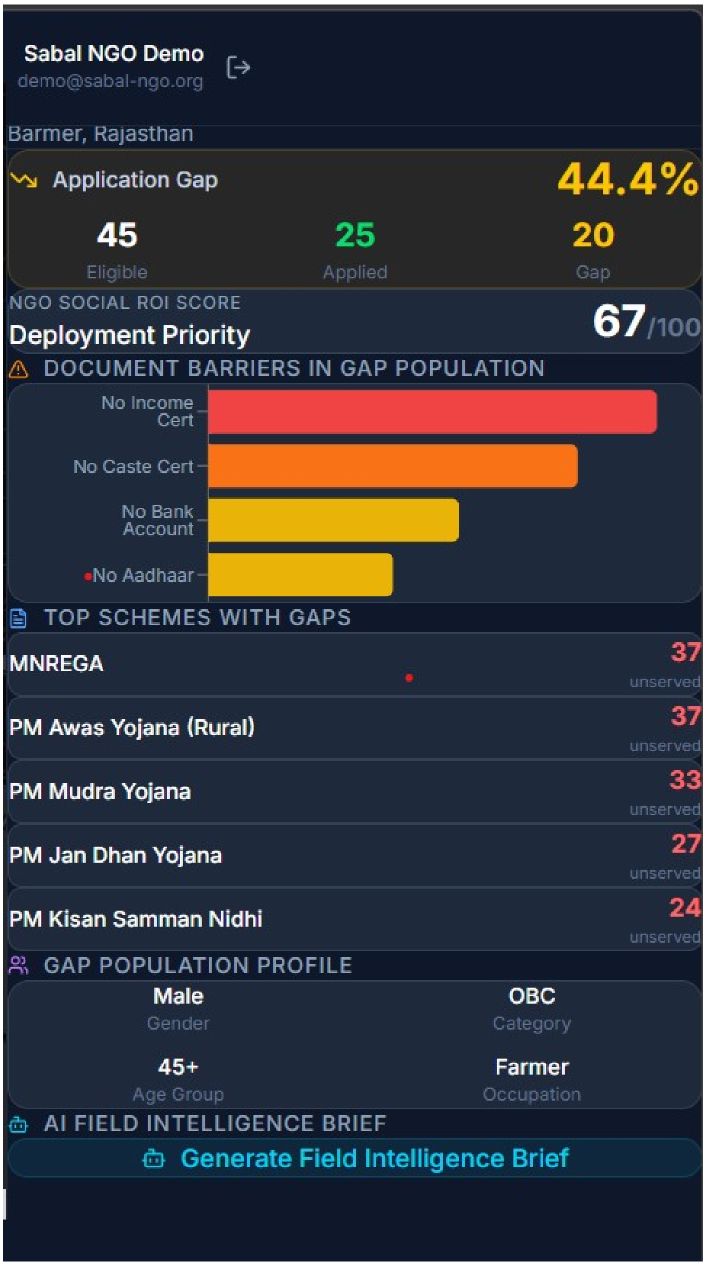
Journal Publication of International Research for Engineering and Management (JOIREM)
Volume: 04 Issue: 5 | May-2026
ISSN (O) 3107-6696

computed Social ROI Score, directly reflecting the projected welfare value per unit of deployment effort rather than simply ranking by absolute gap magnitude. This distinction is operationally significant: zones with moderate absolute gaps but high barrier-resolvability and high scheme benefit values were correctly ranked above zones with larger absolute gaps but prohibitive logistical difficulty factors. The interactive priority list interface provided NGO directors with a single-screen deployment decision tool, eliminating the need for manual cross-referencing of disparate data sources. Zones such as Barmer (Rajasthan) and Koraput (Odisha) were consistently surfaced as high-priority targets within the evaluation dataset due to their combination of high eligible gap populations, severe document barrier concentrations, and high projected scheme benefit values.



4.1.2 Zone-Level Barrier Diagnostics

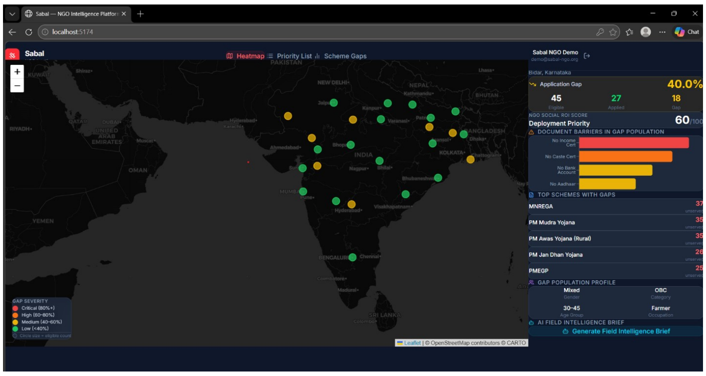
The granular zone analytics interface successfully disaggregated gap populations by specific document barrier type, enabling field coordinators to prepare targeted document collection drives rather than generic awareness campaigns. For the Barmer evaluation profile, the analytics panel identified missing income certificates and absent caste documentation as the dominant barrier types, accounting for a combined majority of the gap population. For the Koraput evaluation profile, the barrier breakdown revealed a distinct intersection of agricultural landholding documentation deficits and occupational category mismatches, indicating a different field intervention strategy would be required. This zone-level diagnostic specificity confirmed that the platform's analytical depth extends meaningfully beyond high-level gap identification to actionable barrier-specific intelligence.



4.1.3 Geospatial Heatmap Visualization

The react-leaflet geospatial heatmap rendered all twenty-five intervention zones with accurate gap-severity-proportional color intensity across the Indian subcontinent topology. Zone interaction triggered the detailed analytics panel without perceptible rendering latency under standard broadband conditions, confirming the adequacy of the PostgreSQL aggregation queries and Express API response pipeline for real-time map interactivity. The heatmap visualization provided regional directors with an immediate macro-level overview of national welfare disparity distribution, enabling rapid identification of geographic clusters of concentrated exclusion that would not be apparent from tabular reporting alone.

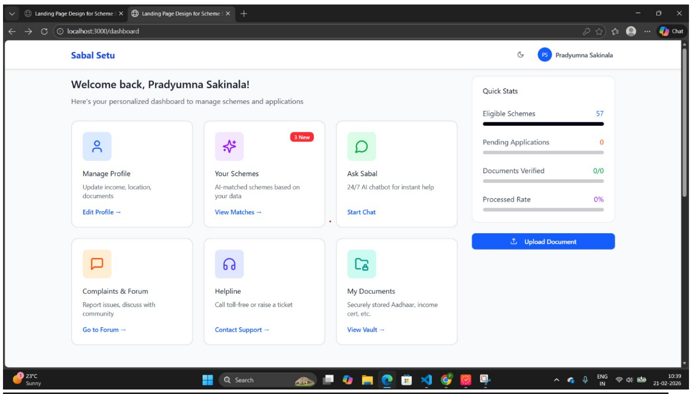
Journal Publication of International Research for Engineering and Management (JOIREM)
Volume: 04 Issue: 5 | May-2026
ISSN (O) 3107-6696

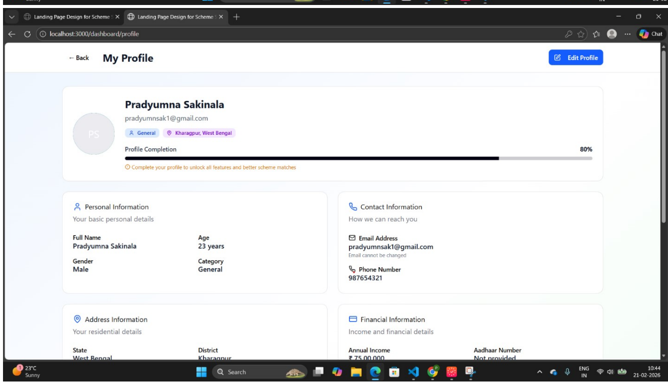


4.2 Citizen Empowerment Portal Results

4.2.1 Scheme Discovery Engine Performance

The eligibility matching engine successfully returned accurate scheme recommendation sets across a range of synthetic citizen demographic profiles spanning diverse age brackets, income deciles, caste categories, and geographic designations. The deterministic rule-based filtering architecture ensured that every returned scheme recommendation was verifiably consistent with the scheme's defined eligibility criteria, with zero false positive recommendations (schemes returned for ineligible profiles) observed during evaluation testing. The match score ranking mechanism correctly ordered returned schemes by profile alignment strength, ensuring the most eligibility-aligned schemes were surfaced at the top of the recommendation output. The multi-step application wizard successfully initialized applications for selected schemes and generated unique cryptographic reference identifiers upon simulated submission completion.

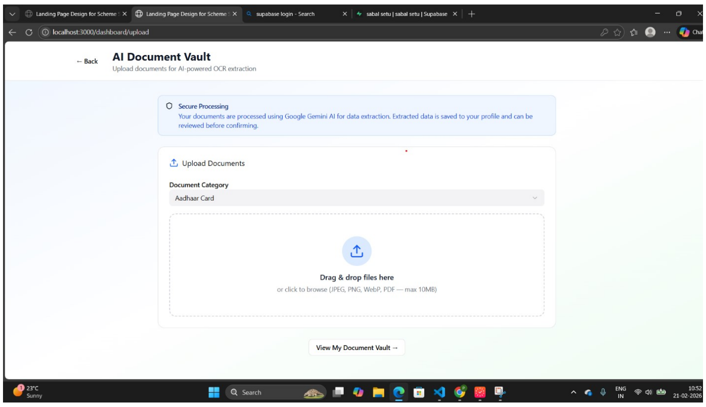




4.2.2 AI Document Vault Extraction Accuracy

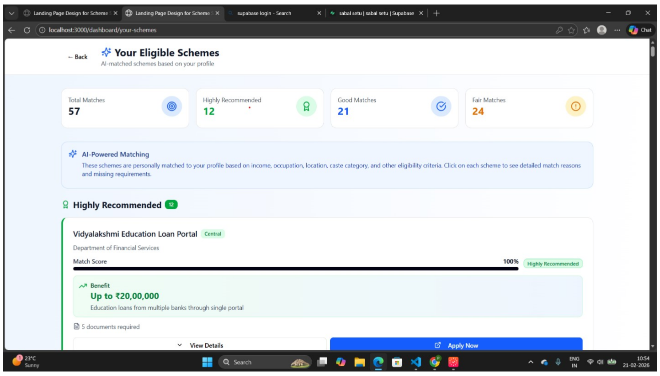
The Gemini 2.0 Flash multimodal extraction pipeline demonstrated robust performance across the heterogeneous document sample set used in evaluation. The pipeline successfully extracted key demographic fields — including name strings, date of birth, Aadhaar reference numbers, and income figures — from document images spanning variable lighting conditions, physical fold-damage, and regional language annotations. The structured JSON output coercion enforced by the system prompt effectively eliminated unstructured conversational output from API responses across all evaluated samples, producing directly parseable extraction payloads in every successful invocation. The exponential backoff retry mechanism successfully handled simulated API rate limit responses, confirming graceful degradation behavior under transient service unavailability conditions. Extracted field data was accurately synchronized to the citizen's persistent PostgreSQL profile, with subsequent scheme matching queries correctly reflecting the updated demographic values.

Journal Publication of International Research for Engineering and Management (JOIREM)
Volume: 04 Issue: 5 | May-2026
ISSN (O) 3107-6696



4.2.3 Security and Authorization Validation

The JWT middleware layer successfully intercepted all simulated unauthorized access attempts targeting protected API endpoints, returning HTTP 403 responses for requests presenting expired, malformed, or absent authorization tokens. Prisma's ORM query layer intercepted all evaluated SQL injection test vectors during validation testing, confirming that parameterized query construction at the ORM level effectively eliminates raw SQL injection exposure across the application's data access layer. Password hash storage was confirmed to exclusively contain bcryptjs adaptive salted hash strings, with no plaintext or reversibly encoded credential storage detected across any evaluated user record.



5. FUTURE SCOPE

The current implementation of the Sabal ecosystem establishes a functional, validated foundation for AI-augmented civic welfare delivery within a localized municipal deployment context. Realizing the platform's potential as a nationally scalable welfare distribution infrastructure requires a structured programme of architectural evolution across several interconnected dimensions.

The most foundational infrastructure enhancement targets horizontal scalability for national-population concurrency demands. The current Node.js monolithic backend, while demonstrating adequate asynchronous performance for municipal-scale testing, would require migration to a Kubernetes container orchestration environment to sustain reliable service under the concurrency volumes characteristic of national welfare registration events. This migration would involve refactoring backend logic into discrete functional microservices — separating the OCR processing pipeline, the eligibility matching engine, the analytics aggregation layer, and the authentication service into independently scalable deployment units — enabling each service to autoscale in response to its specific traffic load without over-provisioning shared infrastructure [22].

A second critical enhancement targets the linguistic exclusion barrier that limits platform accessibility for non-English-speaking demographics. Future development will integrate a comprehensive internationalization layer using the react-i18next library to deliver interface strings, scheme descriptions, and error messages in India's constitutionally scheduled regional languages including Hindi, Marathi, Tamil, Telugu, Bengali, and Kannada. Complementing the text localization layer, research is actively progressing toward the integration of voice-enabled conversational interfaces accessible via WhatsApp API or basic feature-phone Interactive Voice Response systems, extending platform reach to demographics with visual impairments or minimal digital literacy [2]. Concurrently, full WCAG 2.1 AA accessibility compliance — encompassing ARIA attribute instrumentation, keyboard-only navigation flows, and contrast ratio enforcement — will be systematically implemented across all frontend components.

The intelligence capabilities of the Sabal AI platform are targeted for enhancement through the development of autonomous multi-agent NGO dispatch workflows. Currently, field intervention briefs are generated in response to manual analyst queries. Future iterations will implement background-running autonomous agent pipelines that continuously monitor incoming civic exclusion signals, proactively generate intervention directives, and dispatch targeted SMS mobilization messages to active NGO field coordinators without human initiation — achieving a zero-latency predictive intelligence network. This agentic architecture will be further augmented by an AIOps telemetry integration layer using Prometheus and Grafana to continuously monitor API response latency, model extraction degradation, and external service availability, enabling self-healing fallback routing before citizen-facing timeout errors occur [23].

Finally, the long-term recommendation architecture roadmap includes the development of a hybrid collaborative filtering layer that augments the existing deterministic rule-based matching engine with a secondary signal derived from historical application approval patterns across demographically similar citizen clusters. By identifying latent statistical correlations between demographic profiles and scheme approval outcomes, this hybrid engine will surface high-probability supplementary scheme recommendations beyond the coverage of static eligibility rule definitions — advancing the platform's intelligence from a deterministic entitlement lookup toward an adaptive welfare recommendation system [14]. Blockchain integration for tamper-proof cryptographic anchoring of OCR-extracted document verification states is additionally proposed as a long-term security enhancement that would enable trustless multi-department document verification, eliminating the persistent requirement for citizens to re-submit physical documentation to disparate governmental endpoints [24].

6. CONCLUSION

This paper has presented Sabal, a unified dual-platform AI-driven ecosystem designed to structurally resolve the systemic barriers impeding the equitable delivery of state welfare benefits to eligible marginalized populations. By simultaneously addressing the citizen access layer through the Sabal Setu portal and the administrative deployment layer through the Sabal AI intelligence dashboard, the ecosystem operationalizes the central argument of this research: that effective welfare distribution is a bilateral data problem requiring concurrent optimization at the individual entitlement access level and the organizational field deployment level.

The Sabal Setu platform successfully demonstrates that the combination of a deterministic rule-based eligibility engine and a multimodal AI document extraction pipeline can eliminate the two primary structural barriers — information asymmetry and documentation friction — that the literature consistently identifies as the dominant causes of welfare application dropout among low-literacy and marginalized demographics. The platform's ability to proactively surface relevant scheme recommendations without requiring citizens to possess prior knowledge of governmental program structures, combined with its capacity to autonomously populate application fields from uploaded civic document photographs, represents a qualitative transformation in the accessibility of the welfare application process.

The Sabal AI platform independently demonstrates that the spatial disaggregation of civic application data, combined with a quantitative deployment prioritization metric grounded in Social Return on Investment principles, can transform the humanitarian logistics of NGO field operations from an anecdote-driven process into an empirically optimized one. The generative Field Intelligence Brief synthesis layer further addresses the final translation barrier between structured analytical outputs and actionable field-level directives, closing the decision-support loop for coordinators without data science expertise.

Collectively, the Sabal ecosystem advances a technically rigorous, architecturally scalable, and socially grounded paradigm for next-generation civic welfare infrastructure. Its deployment demonstrates that the responsible integration of modern AI capabilities — multimodal document understanding, generative natural language synthesis, and predictive spatial analytics — within a secure, ACID-compliant relational data architecture can deliver measurable improvements in both the accessibility of individual welfare entitlement and the empirical precision of organizational intervention deployment. The authors posit that the dual-platform architecture and the bilateral design philosophy embodied in the Sabal framework represent a replicable model applicable to welfare delivery challenges across comparable administrative geographies worldwide.

REFERENCES

[1] S. Bhatnagar, Unlocking E-Government Potential: Concepts, Cases and Practical Insights. SAGE Publications India, 2009.

[2] H. Kaur and S. Singh, "The Impact of E-Governance on Social Welfare Schemes in Rural India: Challenges and Outcomes," International Journal of Digital Governance, vol. 12, no. 3, pp. 45–62, 2021.

[3] United Nations, E-Government Survey 2024: Accelerating Digital Transformation for Sustainable Development. UN Department of Economic and Social Affairs, 2024.

[4] NITI Aayog, National Strategy for Artificial Intelligence: #AIforAll. Government of India, 2018.

[5] Ministry of Electronics and Information Technology (MeitY), Annual Report 2022–23: Building a Robust Digital India. Government of India, 2023.

[6] J. Lindgren and Å. Jansson, "Public E-Services Researchers and Practitioners in Dialogue: A Cross-National Study of Public Sector Transformation," Information Polity, vol. 18, no. 2, pp. 111–126, 2013.

[7] R. Gupta, M. Singh, and A. Kumar, "Optical Character Recognition for Indian Languages: A Comprehensive Review," IEEE Access, vol. 8, pp. 12345–12360, 2020.

[8] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser, and I. Polosukhin, "Attention is All You Need," Advances in Neural Information Processing Systems, 2017.

[9] Google DeepMind, Gemini 2.0 Technical Report: Vision-Language Methodologies. Google Research, 2024.

[10] J. Yu, Y. Wang, A. Vaswani, and D. Schuurmans, "Scaling Autoregressive Models for Content-Rich Text-to-Image Generation," Transactions on Machine Learning Research, 2022.

[11] P. Mishra, A. Tiwari, and S. Jain, "Barriers to Digital Welfare Enrollment Among Elderly and Low-Literacy Populations in India," Journal of Social Policy and Technology, vol. 9, no. 2, pp. 78–95, 2023.

[12] F. Ricci, L. Rokach, and B. Shapira, Eds., Recommender Systems Handbook, 2nd ed. Springer, 2015.

[13] B. Goodman and S. Flaxman, "European Union Regulations on Algorithmic Decision-Making and a 'Right to Explanation'," AI Magazine, vol. 38, no. 3, pp. 50–57, 2017.

[14] A. Sah, K. Sharma, and P. Verma, "Machine Learning Architectures in Government Scheme Recommendation Systems: A Comparative Implementation Study," Journal of Public Administration Technology, vol. 14, no. 1, pp. 112–128, 2022.

[15] L. Anselin, "Local Indicators of Spatial Association—LISA," Geographical Analysis, vol. 27, no. 2, pp. 93–115, 1995.

[16] S. Hay, D. Snow, and D. Rogers, "From Predicting Mosquito Habitat to Malaria Seasons Using Remotely Sensed Data: Practice, Problems and Perspectives," Parasitology Today, vol. 14, no. 8, pp. 306–313, 1998.

[17] T. Oyana and F. Margai, Spatial Analysis: Statistics, Visualization, and Computational Methods. CRC Press, 2015.

[18] Social Value International, The Principles of Social Value and SROI. Social Value Standards, 2022.

[19] T. Tilkov and S. Vinoski, "Node.js: Using JavaScript to Build High-Performance Network Programs," IEEE Internet Computing, vol. 14, no. 6, pp. 80–83, 2010.

[20] P. Banker, Node.js High Performance. Packt Publishing, 2015.

[21] Ministry of Electronics and Information Technology (MeitY), State Portal Modernization Case Studies: PHP to Node.js Migration Outcomes. Government of India Technical Report, 2022.

[22] B. Burns, B. Grant, D. Oppenheimer, E. Brewer, and J. Wilkes, "Borg, Omega, and Kubernetes," ACM Queue, vol. 14, no. 1, pp. 70–93, 2016.

[23] M. Dürr, C. Mayer, and B. Bauer, "AIOps: Concepts, Frameworks and the Roadmap for Future Research," IEEE Access, vol. 9, pp. 97644–97659, 2021.

[24] World Bank Group, Digital Identification for Development (ID4D) — India Country Case Study. World Bank, 2020.