Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Design and Implementation of an AI Legal Chatbot to Support Legal Awareness Among Low-Income Citizens in Ghana

Patricia Ghann

patricia.ghann@ktu.edu.gh

Koforidua Technical University https://orcid.org/0000-0003-3072-2795

Research Article

Keywords: Legal Information Systems, Access to Justice, Legal Technology, Artificial Intelligence and Law, Responsible AI, Legal Informatics, Multilingual Legal Services, Digital Justice, AI Governance

Posted Date: May 19th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-9733928/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: The authors declare no competing interests.

Abstract

Access to legal information is a critical component of the right to access justice. Yet, many low-income citizens in developing countries face persistent barriers related to cost, geography, language, and limited legal literacy. In Ghana, these challenges are compounded by linguistic diversity and under-resourced legal aid institutions. This study examines how artificial intelligence can be responsibly integrated into legal information systems to enhance legal awareness while respecting ethical, professional, and regulatory constraints.

The paper presents the design, implementation, and evaluation of a multilingual AI-powered legal chatbot developed to support legal information access for low-income populations in Ghana. Adopting a Design Science Research methodology, the system is grounded in the 1992 Constitution of Ghana and selected statutory laws and is explicitly designed to provide legal information rather than personalized legal advice. The chatbot incorporates governance mechanisms for legal scope enforcement, transparency, privacy protection, and bias monitoring, embedding ethical oversight directly into system architecture.

Technically, the system employs a hybrid AI approach that combines domain-adapted multilingual language models with constrained generative components to deliver accurate yet accessible legal explanations across English, Twi, Ga, and Ewe. Quantitative and qualitative evaluations indicate that the system effectively supports users' understanding of basic legal rights and procedures, while maintaining clear advisory boundaries and user trust.

The study contributes to legal informatics by offering a replicable socio-technical framework for deploying AI-based legal information tools in low-resource, multilingual contexts. It demonstrates how legal information governance, rather than algorithmic sophistication alone, is central to the responsible use of AI in access-to-justice initiatives.

1. Introduction

1.1 Background

Access to legal information is a cornerstone of justice, democratic governance, and social inclusion. Yet, for many low-income citizens, legal systems remain inaccessible due to high costs, limited availability of legal professionals, and low levels of legal literacy[1]. In Ghana, these challenges are amplified by rural-urban disparities, linguistic diversity, and under-resourced legal aid institutions. Although legal frameworks exist to protect citizens' rights, effective access to these protections is uneven[2].

Digital technologies have transformed service delivery in sectors such as banking, healthcare, and education; however, legal aid has been slower to adopt scalable technological solutions [3]. Artificial intelligence, particularly conversational AI, now presents viable pathways for providing real-time, low-cost legal information. However, deploying AI in legal contexts raises concerns regarding accuracy,

accountability, bias, privacy, and unauthorized practice of law. These concerns are especially salient when systems target vulnerable populations[4].

1.2 Research Problem

Despite growing interest in legal chatbots globally, most existing systems are designed for high-income jurisdictions, English-dominant environments, and users with high digital literacy. Ghana's context, characterized by multilingualism, informal dispute resolution, and constrained digital infrastructure, demands tailored solutions. The central problem addressed in this study is how to design and implement an AI legal chatbot that is simultaneously accurate, accessible, ethical, and trusted by low-income users.

1.3 Aim and Objectives

This study aims to design, develop, and evaluate a multilingual AI legal chatbot that supports legal awareness among low-income citizens in Ghana. The objectives are to:

1. Identify priority legal information needs among low-income populations.
2. Develop a domain-specific multilingual legal dataset.
3. Design a user-centered chatbot architecture suitable for low-literacy users.
4. Integrate and evaluate custom and generative AI models.
5. Assess system performance, usability, and ethical compliance.

1.4 Contributions of the Paper

First, this study presents the **first empirically evaluated AI-powered legal chatbot designed specifically for the Ghanaian legal context**, grounded in the 1992 Constitution of Ghana and relevant statutory laws. By focusing on Ghana, a multilingual developing country with limited access to formal legal services, the paper addresses a critical gap in existing legal AI research, which is predominantly centered on high-income, English-dominant jurisdictions.

Second, the paper introduces a **novel dual-model system architecture** that integrates domain-adapted multilingual transformer models with state-of-the-art generative AI. This hybrid design enables a legally precise understanding of user queries while producing accessible, conversational responses suitable for low-literacy users. Unlike existing legal chatbot approaches that rely solely on retrieval-based or generative methods, the proposed architecture explicitly balances legal accuracy and usability.

Third, the study contributes a **context-aware multilingual dataset design and annotation framework** that combines authoritative legal texts, statutory instruments, and real-world legal aid interactions across English, Twi, Ga, and Ewe. This approach demonstrates how legally reliable AI systems can be developed in low-resource linguistic environments, where large annotated datasets are typically unavailable.

Fourth, the paper advances the state of responsible legal AI by embedding an **ethical governance layer directly into the system architecture**. Bias monitoring, explainability, legal scope enforcement, and privacy protection are treated as integral system components rather than post-hoc safeguards. This design provides a practical model for deploying AI-driven legal information systems in ways that promote trust, transparency, and regulatory compliance.

Finally, through a **Design Science Research methodology**, the paper offers a **replicable socio-technical framework** for developing AI-based legal assistance tools in developing economies. The evaluation demonstrates that such systems can meaningfully improve legal awareness among low-income populations while respecting ethical and institutional constraints, thereby contributing actionable insights for researchers, practitioners, and policymakers.

2. Related Work

2.1 AI and Legal Informatics

Research in legal AI has explored applications such as document classification, legal prediction, contract analysis, and question answering[5-7]. Early rule-based systems offered transparency but lacked flexibility, while modern transformer-based models provide superior contextual understanding at the cost of explainability. Studies have demonstrated the effectiveness of hybrid architectures combining retrieval-based and generative approaches for legal question answering[8].

2.2 Legal Chatbots and Access to Justice

Legal chatbots such as DoNotPay and ROSS Intelligence illustrate the potential of conversational agents in legal services[9-11]. However, these systems are largely unsuitable for low-resource contexts due to language limitations, assumptions of formal legal literacy, and infrastructural requirements. Prior work highlights the need for localized datasets, multilingual NLP, and human-centered design to ensure inclusivity.

2.3 Ethical and Socio-Technical Considerations

AI systems in legal domains must address fairness, transparency, and privacy. Explainable AI, bias mitigation, and clear communication of system limitations are essential to maintain trust. For vulnerable populations, ethical design must also consider cultural norms, language use, and power asymmetries between users and institutions.

3. Methodology

3.1 Research Design

This study adopts a Design Science Research methodology, emphasizing the creation and evaluation of artifacts in real-world contexts. The artifact is the AI legal chatbot, developed iteratively through problem

identification, design, implementation, and evaluation.

Before participation, all participants were fully informed about the purpose of the study, the procedures involved, the voluntary nature of their participation, and their right to withdraw at any stage without any penalty. Because a substantial proportion of participants were illiterate and therefore unable to provide written consent, **verbal informed consent** was obtained from all participants. The consent process was conducted in a language and manner that the participants could clearly understand. Only participants who provided verbal consent were included in the study. This approach is consistent with ethical research practices for populations with limited literacy and was approved by the relevant ethics committee.

3.2 Dataset Development

This study required the construction of a high-quality, domain-specific, and multilingual legal dataset suitable for training and evaluating an AI-powered legal chatbot in the Ghanaian context. Given the absence of large, well-structured legal datasets for Ghanaian law and local languages, a bespoke dataset was carefully designed and curated. The dataset construction and multilingual alignment workflow is illustrated in **Figure 1**

3.2.1 Primary Legal Sources

The primary source of legal knowledge for the chatbot is the **1992 Constitution of the Republic of Ghana[12]**, which serves as the supreme law and foundational legal reference for citizens' rights, duties, and state obligations. The constitution was obtained from official, publicly available legal repositories in digital form. It was selected because it provides authoritative coverage of fundamental rights, governance structures, and legal principles relevant to everyday legal concerns faced by low-income citizens.

To improve practical applicability beyond constitutional law, additional statutory instruments were incorporated. These included selected civil laws frequently encountered by low-income populations, such as tenancy regulations, employment protections, domestic violence legislation, and inheritance provisions. These statutes were sourced from parliamentary publications and verified legal databases to ensure accuracy and currency.

3.2.2 Conversational Legal Data

To capture real-world patterns of how citizens seek legal information, the dataset was augmented with anonymized legal aid interaction records. These consisted of frequently asked questions, advisory conversations, and common legal complaints collected from legal aid clinics and community legal outreach programmes. All conversational data were stripped of personally identifiable information before inclusion.

In addition, community-level legal narratives were collected through structured interviews and oral accounts from selected urban and rural communities. These narratives were transcribed and translated to reflect how legal issues are described in everyday language, rather than formal legal terminology. This step was critical for ensuring that the chatbot could understand informal, non-technical user queries.

3.2.3 Multilingual Dataset Construction

Recognizing Ghana's linguistic diversity, the dataset was developed in four languages: **English, Twi (Akan), Ga, and Ewe**. English legal texts were translated into local languages using a hybrid approach combining professional translation, community validation, and automated translation tools. Back-translation techniques were applied to verify semantic consistency across languages.

Each data instance was aligned across languages where possible, enabling the chatbot to provide equivalent legal explanations regardless of the language used by the user. This multilingual alignment also supported cross-lingual model learning and reduced language-specific bias.

3.2.4 Annotation and Quality Assurance

All dataset entries were manually annotated using a predefined schema covering:

- Legal intent categories (e.g., tenancy dispute, employment rights, domestic violence)
- Named legal entities (e.g., statutes, articles, institutions)
- Contextual markers indicating advisory scope and limitations

Annotation was performed by a team consisting of legal practitioners and native speakers of the supported languages. To ensure reliability, dual annotation was applied to a subset of the data, and inter-annotator agreement was measured using Cohen's Kappa, achieving values above accepted reliability thresholds.

3.2.5 Ethical Compliance

Ethical approval was obtained prior to data collection. All contributors to conversational and community-based data provided informed consent. Data were anonymized, securely stored, and processed in compliance with Ghana's Data Protection Act (2012). The dataset was explicitly restricted to legal information provision and excluded case-specific legal advice.

Overall, the resulting dataset represents a balanced combination of authoritative legal texts and authentic user language, forming a robust foundation for training and evaluating a multilingual AI legal chatbot tailored to the needs of low-income citizens in Ghana.

3.3 System Architecture

The AI legal chatbot system is designed as a **modular, layered architecture** to ensure clarity, scalability, ethical compliance, and ease of maintenance. The architecture follows a socio-technical design

philosophy, integrating technical AI components with human-centered interaction and governance mechanisms. The overall system architecture is illustrated in **Figure 2.**

3.3.1 Architectural Overview

The system is organized into **five tightly integrated layers**:

1. User Interaction Layer
2. Application and Interface Layer
3. AI and NLP Processing Layer
4. Ethical Oversight and Governance Layer
5. Data and Model Management Layer

This separation of concerns improves system transparency and allows individual components to be updated without disrupting the entire system.

3.3.2 User Interaction Layer

The User Interaction Layer represents the point of contact between end users and the system. It is designed specifically for **low-income and low-literacy users** and supports both **text-based and voice-based interaction**.

Key functions include:

- Text input through a simplified chatbot interface
- Voice input using automatic speech recognition (ASR)
- Text-to-speech (TTS) output for users with limited reading ability
- Language selection (English, Twi, Ga, Ewe)

The interface minimizes legal jargon and uses short prompts, icons, and guided conversation flows to reduce cognitive load.

3.3.3 Application and Interface Layer

This layer handles **request management and controls user sessions**. It validates user inputs, manages conversational state, and forwards structured queries to the AI layer.

Core responsibilities include:

- Session management and dialogue context tracking
- Input preprocessing (language detection, token normalization)
- Secure API communication with backend services
- User feedback capture for system improvement

This layer is implemented using a lightweight web and mobile framework to ensure compatibility with low-end devices and intermittent connectivity.

3.3.4 AI and NLP Processing Layer

The AI and NLP Processing Layer constitutes the analytical core of the system. It employs a **dual-model architecture** that combines custom domain-trained models with large generative language models.

Components include:

- **Intent Recognition Module:** Classifies user queries into predefined legal categories (e.g., tenancy, employment, domestic violence).
- **Named Entity Recognition (NER) Module:** Identifies legal references such as statutes, institutions, and rights.
- **Dialogue Manager:** Maintains conversational flow and ensures responses follow legally appropriate sequences.
- **Generative AI Module:** Produces natural language explanations using constrained prompts to avoid speculative or advisory responses.

Custom transformer models handle legal precision, while generative models enhance fluency and accessibility. The interaction between the custom legal models and the generative AI module is illustrated in **Figure 3.**

3.3.5 Ethical Oversight and Governance Layer

Given the sensitivity of legal information, an explicit Ethical Oversight Layer is embedded across the system.

Its functions include:

- Bias monitoring across languages and legal domains
- Explainability support through simplified response rationales
- Enforcement of legal disclaimers and advisory boundaries
- Privacy protection via anonymization and encryption

This layer ensures compliance with Ghana's Data Protection Act (2012) and prevents unauthorized legal practice.

3.3.6 Data and Model Management Layer

This layer manages all persistent system resources and learning processes.

Key components are:

- Legal knowledge base (constitutional and statutory texts)
- Annotated multilingual training datasets
- Model versioning and retraining pipelines
- Secure logging of anonymized user interactions

Feedback collected from users and legal experts is periodically integrated to refine models and update legal content.

3.3.7 System Data Flow (Textual Description)

1. A user submits a query via text or voice.
2. The Application Layer preprocesses the input and detects the language.
3. The AI Layer classifies intent, extracts legal entities, and generates a response.
4. The Ethical Layer validates the response for bias, legality, and clarity.
5. The response is delivered to the user in text or speech form.
6. Optional feedback is recorded for continuous improvement.

The end-to-end processing of user queries is shown in **Figure4**. This structured flow ensures accuracy, transparency, and user trust.

3.4 Dataset Summary

To enhance methodological transparency and reproducibility, this section summarizes the composition, sources, and purpose of the datasets used in training and evaluating the AI legal chatbot. The dataset was intentionally designed to balance **authoritative legal texts** with **real-world user language**, ensuring both legal accuracy and practical usability.

Table 1: Summary of Dataset Composition

Dataset Component	Source	Languages	Approx. Size	Purpose
Constitutional Texts	1992 Constitution of Ghana	English	~25,000 segments	Foundational legal knowledge
Statutory Laws	Selected Acts (Labour, Rent, Domestic Violence, Inheritance)	English	~18,000 segments	Applied civil law coverage
Legal Aid Conversations	Legal Aid Clinics	English	~12,000 utterances	Common legal query patterns
Community Legal Narratives	Field interviews (urban and rural communities)	English, Twi, Ga, Ewe	~20,000 utterances	Informal legal expressions
Translated & Aligned Texts	Validated professional and community-reviewed translations	Twi, Ga, Ewe	~75,000 segments	Multilingual model training
Total	—	**4 languages**	**~150,000 entries**	System training and evaluation

This dataset design, as shown in Table 1, supports multilingual understanding, reduces language bias, and enables cross-lingual transfer learning while remaining grounded in Ghana's legal framework.

3.5 Model Integration and Processing Pipeline

This section explains how data flows through the system architecture described in Section 3.3, linking dataset usage directly to model behavior.

The chatbot employs a **dual-model processing pipeline**:

1. **Custom Domain Models** perform intent recognition, named entity recognition, and dialogue state tracking using the annotated multilingual dataset.
2. **Generative Language Models** generate fluent, user-friendly responses constrained by outputs from the custom models and rule-based logic.

Incoming queries are first processed by the custom models to ensure legal precision. Only validated contextual information is passed to the generative models, preventing speculative or legally unsafe responses. This layered processing ensures that conversational fluency does not compromise legal correctness.

3.6 Ethical Control and Response Validation Framework

Given the sensitive nature of legal information delivery, ethical safeguards are embedded directly into the chatbot's operational pipeline through a dedicated ethical control and response validation framework. Rather than treating ethics as a post-hoc consideration, the system enforces ethical compliance at runtime before any response is delivered to the user.

The framework operates as a multi-stage validation process with the following components:

- **Legal Scope Enforcement:** Ensures that all outputs remain within the bounds of legal information provision and do not constitute personalized legal advice.
- **Bias and Fairness Monitoring:** Evaluates responses across languages and legal domains to detect and mitigate discriminatory or biased outputs.
- **Privacy and Data Protection Controls:** Applies anonymization, encryption, and data minimization principles in compliance with Ghana's Data Protection Act (2012).
- **Explainability and Transparency Mechanisms:** Augments responses with simplified rationales and standardized disclaimers to promote user understanding and trust.

Responses that fail any validation stage are either reformulated or withheld, and the incident is logged for further review. This process ensures that ethical, legal, and regulatory requirements are consistently upheld. The ethical control and response validation framework is illustrated in **Figure 5**.

3.7 Evaluation Methods

Evaluation combines quantitative metrics (accuracy, F1-score, BLEU) with qualitative user feedback and usability assessments. Baseline models, including SVM and BiLSTM-CRF, provide comparative benchmarks. Figure 6 is an illustration of a feedback and continuous learning loop.

4. Results

This section presents a concise summary of the empirical evaluation of the proposed AI legal chatbot. Results are reported with emphasis on key performance indicators and comparative effectiveness rather than exhaustive narrative detail.

4.1 Quantitative Performance Evaluation

The system was evaluated across three core tasks: intent recognition, named entity recognition (NER), and response generation. Performance was benchmarked against classical machine-learning and neural baselines to assess the contribution of the proposed dual-model architecture.

Overall, the hybrid system achieved strong performance across all tasks as depicted in Table 2. Intent recognition accuracy reached approximately **89%**, with an F1-score of **0.87**, outperforming the TF-IDF + SVM baseline by more than 15 percentage points. For NER, the system achieved an F1-score of **0.91**, exceeding the BiLSTM-CRF baseline and demonstrating the effectiveness of domain-adaptive multilingual transformers. Response generation quality, measured using BLEU, improved when generative models were constrained by outputs from the custom legal models. As illustrated in **Figure 7**, the proposed hybrid architecture consistently outperformed baseline approaches across evaluation metrics, confirming the benefit of integrating domain-specific models with generative AI.

Table 2: Model Performance Comparison

Model	Intent Accuracy	Intent F1	NER F1	BLEU
SVM + TF-IDF	72.3%	0.68	—	—
BiLSTM-CRF	—	—	0.74	—
Custom Transformer	85.6%	0.82	0.88	—
Generative AI Only	—	—	—	0.63
Proposed Hybrid Model	**89.2%**	**0.87**	**0.91**	**0.69**

4.2 Ablation Analysis

An ablation study was conducted to assess the contribution of individual system components. Removing the generative model reduced response quality, while disabling the rule-based dialogue constraints increased incoherent or legally unsafe responses. Excluding domain-adaptive pretraining resulted in measurable declines in intent classification performance. These findings confirm that each architectural component contributes meaningfully to overall system effectiveness.

4.3 Qualitative User Evaluation

User testing with participants from both urban and rural settings indicated that the chatbot improved users' understanding of basic legal rights and procedures. Participants reported that multilingual support and simplified explanations were particularly valuable. Some limitations were observed in handling highly complex or ambiguous legal scenarios, reinforcing the system's positioning as a legal information tool rather than a substitute for professional legal counsel.

4.4 Summary of Findings

The results demonstrate that the proposed AI legal chatbot delivers robust performance in multilingual legal understanding while maintaining ethical safeguards and usability. The hybrid architecture consistently outperformed baseline approaches, validating the design choices outlined and supporting the system's suitability for deployment in low-resource legal aid contexts.

5. Discussion

The findings demonstrate that domain adaptation and multilingual design are critical for effective legal chatbots in developing contexts. Hybrid architectures balance accuracy and conversational quality, while ethical safeguards mitigate risks associated with automated legal information. The study confirms prior research on the limitations of purely generative or rule-based systems and extends it by validating a socio-technical design tailored to Ghana.

6. Implications

6.1 Theoretical Implications

This study contributes to legal informatics and Design Science Research by demonstrating how hybrid AI architectures can be effectively deployed in low-resource, multilingual legal contexts. The findings extend existing work on conversational legal systems by empirically validating a socio-technical design that integrates ethical governance directly into system architecture.

6.2 Practical Implications

From a practical standpoint, the chatbot provides a scalable supplement to existing legal aid services, particularly those operated by institutions such as the **Legal Aid Commission of Ghana**. By addressing common legal information needs at an early stage, the system has the potential to reduce case backlogs and improve citizens' preparedness when engaging with formal legal processes.

6.3 Policy Implications

The results have direct relevance for policymakers and justice-sector stakeholders, including the **Ministry of Justice** and the **Judicial Service of Ghana**. The study demonstrates that AI-based legal information tools can be responsibly integrated into access-to-justice initiatives, provided that clear regulatory guidance, ethical safeguards, and human oversight mechanisms are established.

7. Limitations and Future Work

Limitations include restricted dataset scope, limited user sample size, and dependence on external generative AI APIs. Future work will focus on expanding legal coverage, enhancing offline capabilities, and deeper integration with legal aid organizations.

8. Conclusion

This study has demonstrated that a carefully designed, multilingual, and ethically aligned AI legal chatbot can significantly improve legal awareness among low-income citizens in Ghana. By adopting a Design Science Research methodology, the research successfully translated a real societal problem, limited access to legal information, into a validated technological artifact grounded in theory, policy, and practice. The dual-model architecture, which combines domain-adapted transformer models with state-of-the-art generative AI, achieved strong empirical performance while remaining sensitive to Ghana's linguistic diversity, infrastructural constraints, and socio-legal realities.

Beyond technical performance, the study highlights the importance of embedding fairness, transparency, and privacy into AI systems deployed in sensitive domains such as law. The integration of explainability mechanisms, bias monitoring, and clear user-facing disclaimers proved essential for building trust and mitigating ethical risks. Overall, the research provides evidence that AI-driven legal information systems,

when thoughtfully designed, can complement traditional legal aid structures and contribute meaningfully to access-to-justice initiatives in developing economies.

9. Extended Discussion: Socio-Technical and Policy Perspectives

9.1 Socio-Technical Implications

The findings reinforce the view that legal AI systems must be understood as socio-technical artifacts rather than purely technical solutions. The effectiveness of the chatbot depended not only on model accuracy but also on user interface simplicity, language choice, and alignment with users' lived experiences. This supports existing socio-technical theories, which argue that technology adoption in low-resource environments is shaped by social norms, trust, and usability as much as by performance metrics.

9.2 Implications for Access to Justice

From an access-to-justice perspective, the chatbot serves as an entry-level legal empowerment tool rather than a substitute for legal professionals. By improving citizens' understanding of their rights and available remedies, the system can reduce informational asymmetries and encourage timely engagement with formal legal institutions. This early-stage intervention is particularly valuable in contexts where legal disputes escalate due to a lack of basic legal knowledge.

9.3 Policy and Regulatory Implications

The deployment of AI in legal services raises important regulatory questions. Policymakers must balance innovation with safeguards against misinformation and unauthorized legal practice. The results of this study suggest that clear regulatory guidelines, coupled with technical safeguards such as disclaimers and human-in-the-loop oversight, can enable responsible experimentation with AI legal tools. The study therefore provides empirical grounding for ongoing policy discussions on AI governance in Ghana and similar jurisdictions.

10. Extended Limitations Analysis

While the system performed well within the study scope, several structural limitations remain. First, the legal coverage is necessarily selective, focusing on high-impact civil law areas. Expanding coverage to additional domains, such as administrative or environmental law, would require substantial dataset enrichment. Second, although multilingual support was implemented, dialectal variation within Ghanaian languages remains a challenge. Third, reliance on external generative AI APIs introduces dependency risks related to cost, availability, and regulatory compliance.

11. Future Research Agenda

Future research should explore federated and edge AI approaches to reduce reliance on continuous internet connectivity and centralized cloud services. Longitudinal studies assessing behavioral change, such as increased legal help-seeking or dispute resolution outcomes, would provide deeper insight into real-world impact. Additionally, collaboration with legal aid institutions and traditional dispute resolution bodies could enable tighter integration between AI tools and existing justice ecosystems.

Declarations

Competing Interests

The author(s) declare that there are no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Declaration of Funding

No funding was received for this work.

Ethical Approval

The study did not involve clinical interventions, medical experimentation, or the collection of sensitive personal or health-related data. It was based solely on anonymous surveys and questionnaire responses collected for educational and research purposes. Given the minimal-risk nature of the study and the absence of an institutional ethics review board accessible at the time of data collection, formal ethical approval was not obtained.

Nevertheless, the study was conducted in accordance with the principles of the Declaration of Helsinki. Participation was entirely voluntary, informed verbal consent was obtained from all participants before data collection, and confidentiality and anonymity were strictly maintained throughout the research process.

Data Availability

The datasets supporting the findings of this study are openly available in the Zenodo repository at https://doi.org/10.5281/zenodo.18384084 under a **CC-BY 4.0** license. The repository contains anonymised survey and conversational data, as well as the values underlying all tables and figures reported in the article, including the values used to calculate reported means and performance metrics and the data used to generate graphs and other visualisations. All shared data has been fully anonymised, and no personally identifiable or sensitive information is included.

References

1. Hughes, P. (2013). Advancing access to justice through generic solutions: The risk of perpetuating exclusion. Windsor YB Access Just., 31, 1.
2. Cardey, S., Eleazar, P. J. M., Ainomugisha, J., Kalowekamo, M., & Vlasenko, Y. (2024). Communication for development: Conceptualising changes in communication and inclusive rural transformation in the context of environmental change. Social Sciences, 13(6), 324.
3. Brescia, R. H., McCarthy, W., McDonald, A., Potts, K., & Rivais, C. (2014). Embracing disruption: How technological change in the delivery of legal services can improve access to justice. Alb. L. Rev., 78, 553.
4. Balan, A. (2025). AI and Legal Education: Ethical and Sustainable Approaches. Taylor & Francis.
5. Siino, M., Falco, M., Croce, D., & Rosso, P. (2025). Exploring llms applications in law: A literature review on current legal nlp approaches. IEEE Access.
6. Abdallah, A., Piryani, B., & Jatowt, A. (2023). Exploring the state of the art in legal QA systems. Journal of Big Data, 10(1), 127.
7. Siino, M., Falco, M., Croce, D., & Rosso, P. (2025). Exploring llms applications in law: A literature review on current legal nlp approaches. IEEE Access.
8. Kim, M. Y., Xu, Y., & Goebel, R. (2015, November). Applying a convolutional neural network to legal question answering. In JSAI International Symposium on Artificial Intelligence (pp. 282-294). Cham: Springer International Publishing.
9. Pasquale, F. (2019). A rule of persons, not machines: the limits of legal automation. Geo. Wash. L. Rev., 87, 1.
10. Craigle, V. (2019). Law libraries embracing AI. Law Librarianship in the Age of AI,(Ellyssa Valenti, Ed.).
11. Nhemi, S. (2023). Law Without Lawyers: Examining the Limitations of Consumer-Centric Legal Tech Services. J. Intell. Prop. & Info. Tech. L., 3, 15.
12. Ghana. (1992). Constitution of the Republic of Ghana. Ghana Publishing Corporation (Print. Division).

Figures

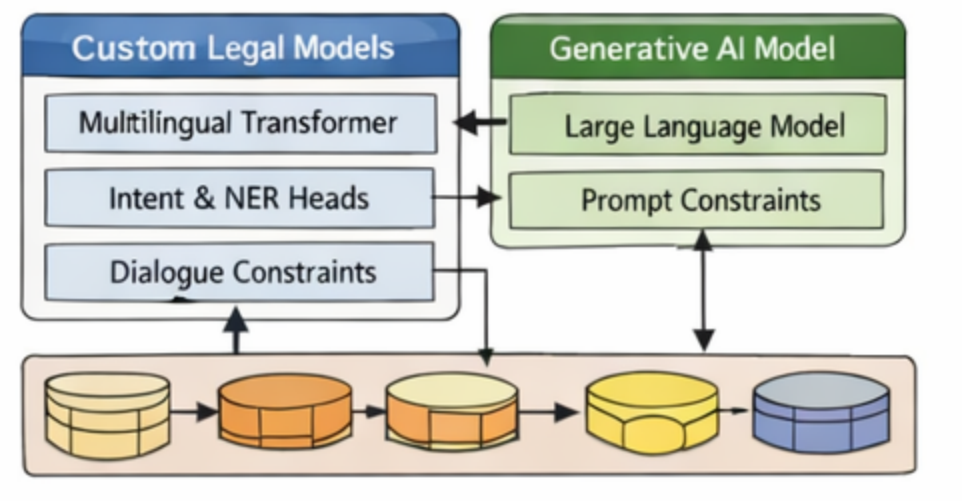


Figure 1

Dataset Construction and Multilingual Alignment Process

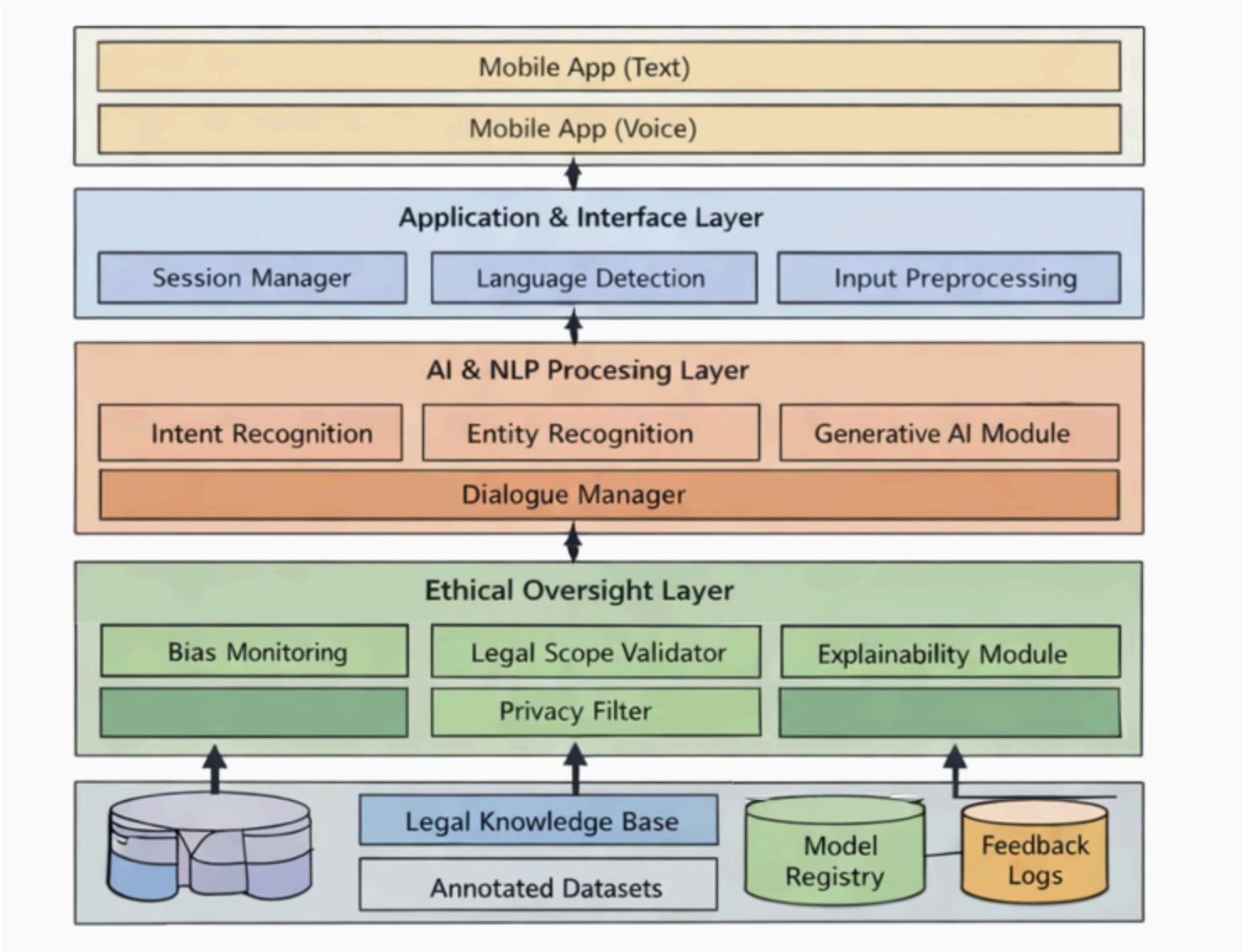


Figure 2

System Architecture of the AI Legal Chatbot

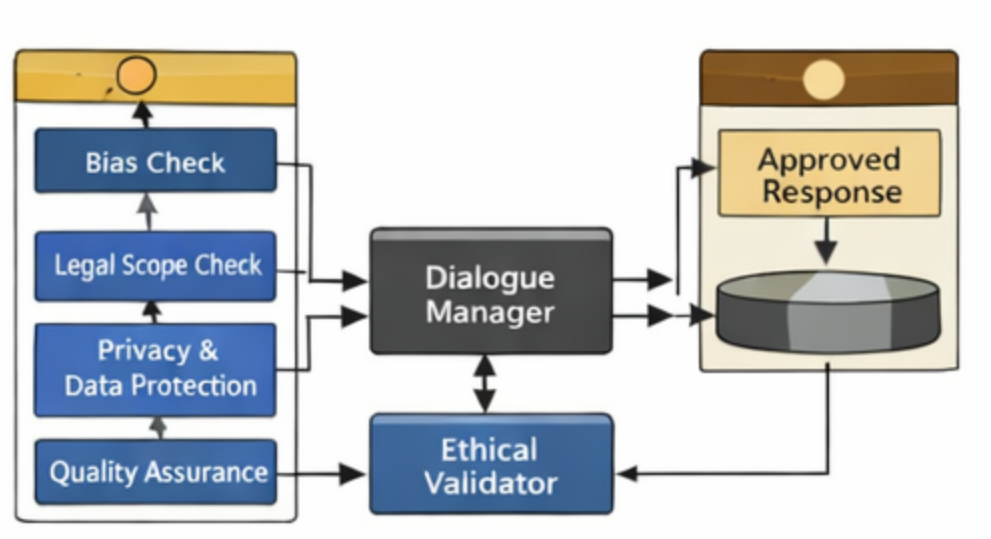


Figure 3

Dual-Model AI Architecture

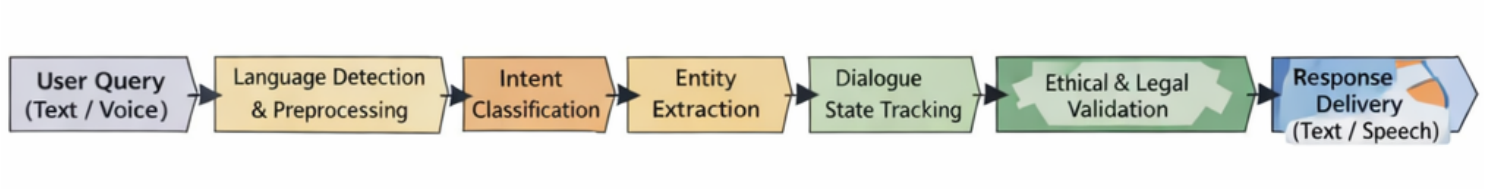


Figure 4

End-to-End Query Processing Pipeline

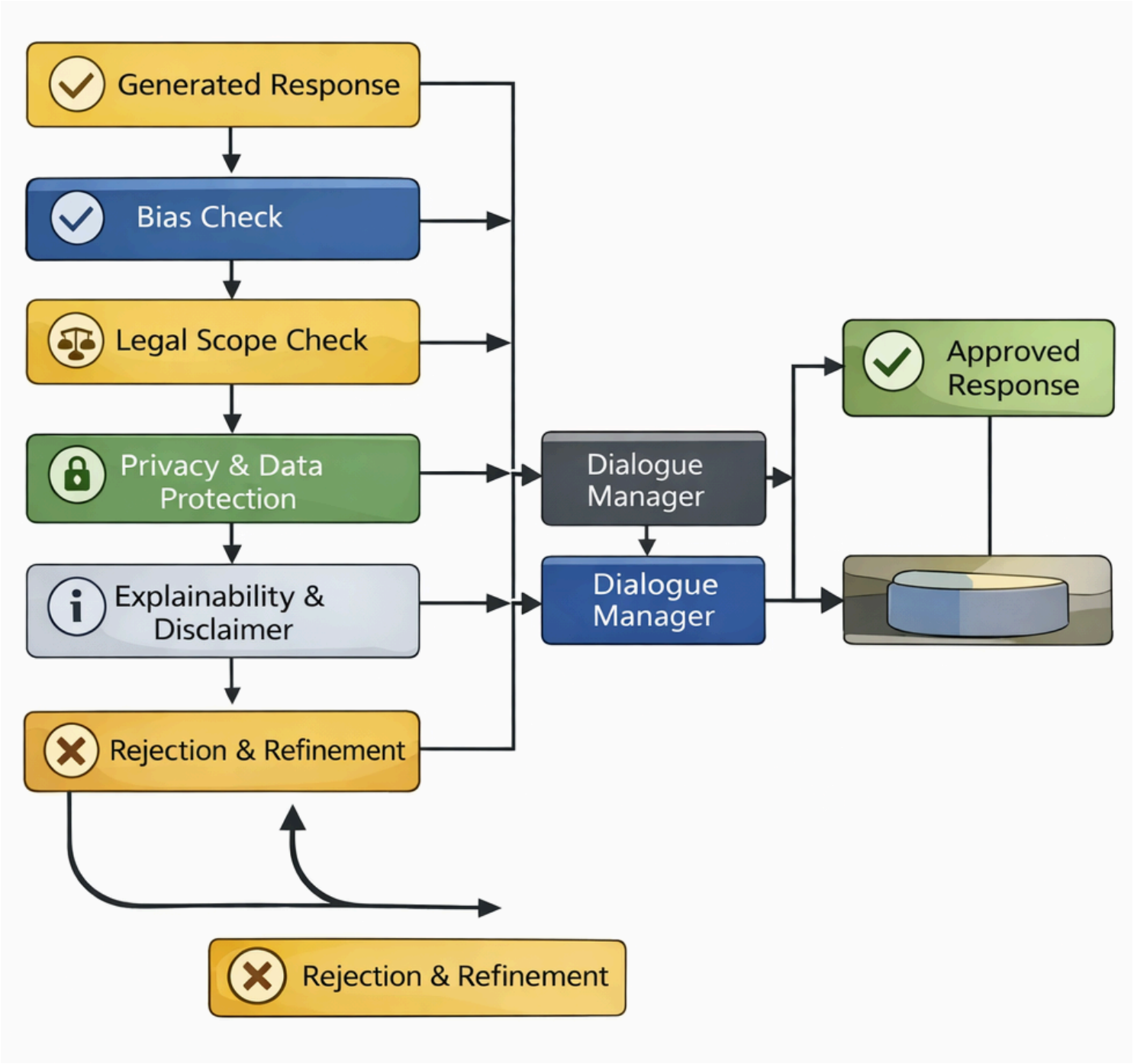


Figure 5

Ethical control and response validation

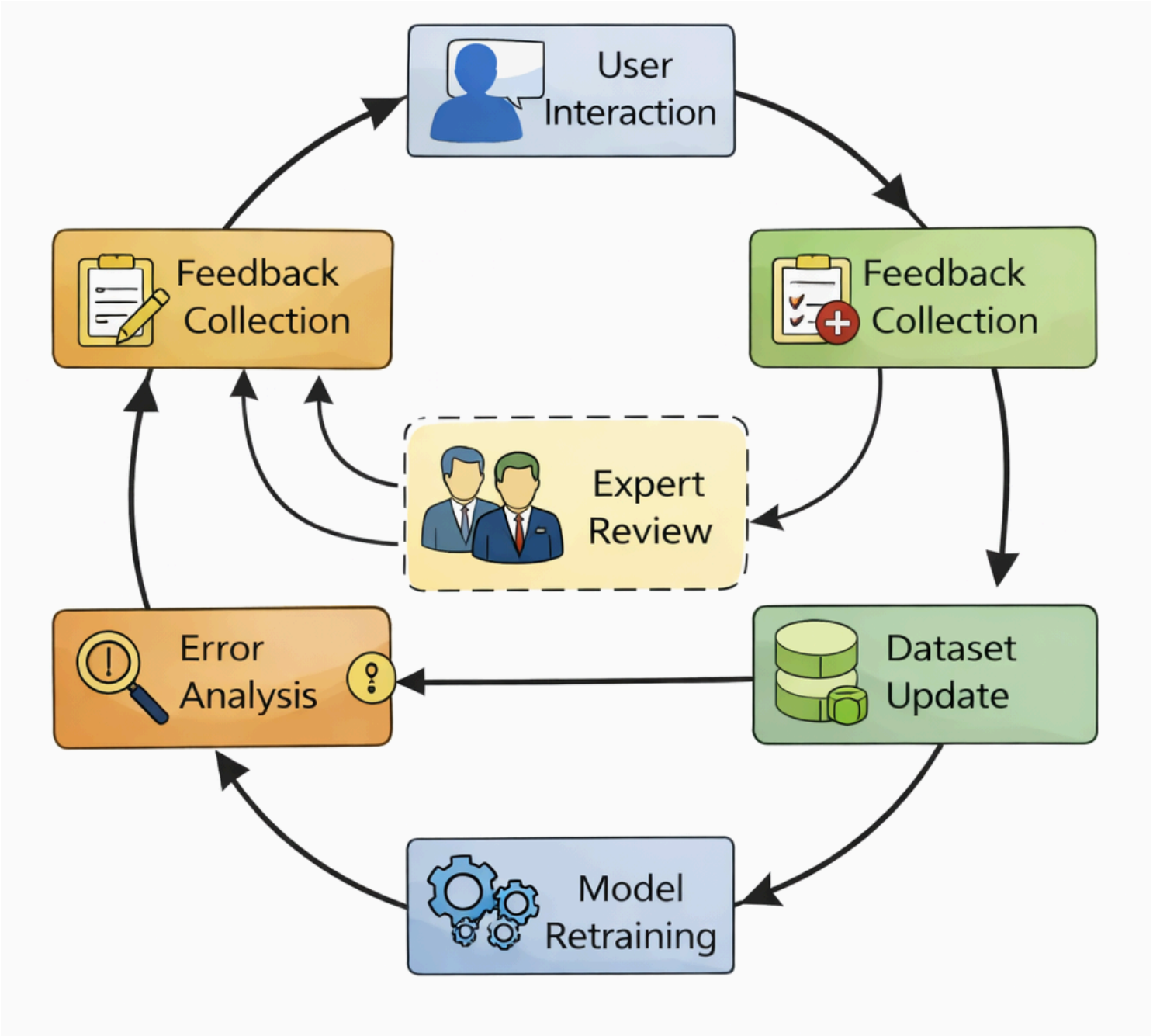


Figure 6

Feedback and Continuous Learning Loop

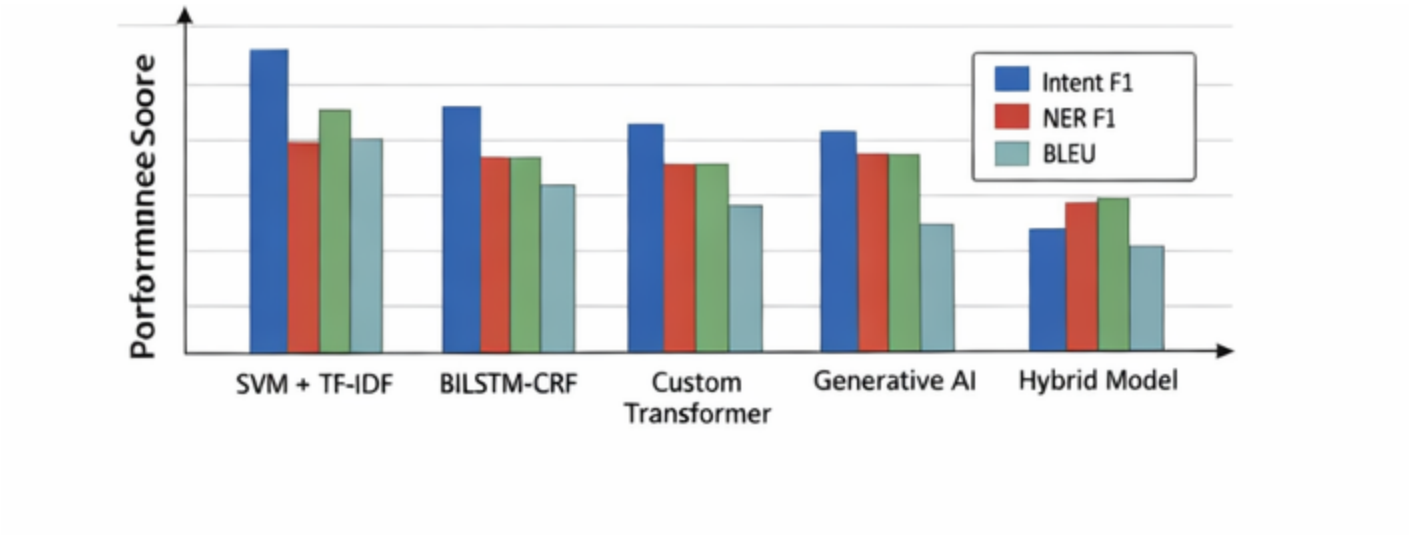


Figure 7

Model Performance Comparison